MANCHESTER 1824

The University of Manchester

Cathie Marsh Centre for Census and Survey Research

On Robust Mean Squared Error Estimation for Linear Predictors for Domains

CCSR Working Paper 2007-10

Ray Chambers, Hukum Chandra & Nikos Tzavidis
Nikos.tzavidis@manchester.ac.uk

We discuss robust mean squared error estimation for linear predictors of finite population domain means. Our approach represents an extension of the ideas in Royall and Cumberland (1978) and appears to lead to estimators that are simpler to implement, and potentially more robust, than those suggested in the small area literature. We demonstrate the usefulness of our approach through both model-based as well as design-based simulation, with the latter based on two realistic survey data sets containing small area information.

On Robust Mean Squared Error Estimation for Linear Predictors for Domains

Ray Chambers, University of Wollongong[1]

Hukum Chandra, University of Southampton[2]

Nikos Tzavidis, University of Manchester[3]

Abstract

We discuss robust mean squared error estimation for linear predictors of finite population domain means. Our approach represents an extension of the ideas in Royall and Cumberland (1978) and appears to lead to estimators that are simpler to implement, and potentially more robust, than those suggested in the small area literature. We demonstrate the usefulness of our approach through both model-based as well as design-based simulation, with the latter based on two realistic survey data sets containing small area information.

Keywords Best linear unbiased prediction; M-quantile model; Random effects model; Small area estimation.

[1] Ray Chambers, Centre for Statistical and Survey Methodology, School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, NSW 2522, Australia. ray@uow.edu.au.

[2] Hukum Chandra, Southampton Statistical Sciences Research Institute, University of Southampton, Southampton, SO17 1BJ, UK. hchandra@soton.ac.uk.

[3] Nikos Tzavidis, Centre for Census and Survey Research, University of Manchester, Manchester, M13 9PL, UK. nikos.tzavidis@manchester.ac.uk.

1. INTRODUCTION

Linear models, and linear predictors based on these models, are widely used in survey-based inference. However, such models run the risk of misspecification, particularly with regard to second order and higher moments, and this can bias estimators of the mean squared error (MSE) of these predictors. Robust methods for estimating the MSE of predictors of finite population quantities, i.e. methods that minimise the impact of assumptions about second order and higher moments, have been developed. Valliant, Dorfman and Royall (2000, chapter 5) discuss robust MSE estimation when a population is assumed to follow a linear model.

In this paper we address a subsidiary problem, which is that of robust mean squared error estimation for linear predictors of finite population domain means. An important application, and one that motivates our approach, is small area inference. Consequently from now on we use 'small area' (or just 'area') to refer to a domain of interest. Our approach represents an extension of the ideas in Royall and Cumberland (1978) and appears to lead to simpler to implement, and potentially more robust, MSE estimators than those that have been suggested in the small area literature, see Prasad and Rao (1990) and Rao (2003, section 6.2.6).

The structure of the paper is as follows. In section 2 we discuss area-specific MSE estimation under an area-specific linear model and area-averaged MSE estimation under a linear mixed model. We then show how our approach can be used for estimating the MSE of three different small area linear predictors, (a) the empirical best linear unbiased predictor or EBLUP (Henderson, 1953); (b) the model based direct estimator (MBDE) of Chambers and Chandra (2006); and (c) the M-quantile predictor (Chambers and Tzavidis, 2006). In section 3 we present results from a series of simulation studies that illustrate the model-based and the design-based properties of our approach to MSE estimation. Finally, in section 4 we summarize our main findings.

2. ROBUST MSE ESTIMATION FOR LINEAR PREDICTORS OF AREA MEANS

2.1 MSE Estimation under an Area-Specific Linear Model

When survey-based inference relates to the characteristics of a group of *D* areas that partition the surveyed population, it is usually not realistic to assume that a linear model that applies to the population as a whole also applies within each area. We therefore consider MSE estimation for linear estimators of area means when different linear models apply at area level.

To start, let y_j denote the value of *Y* for unit *j* of the population and suppose that this unit is in area *i*. We assume an area-specific linear model for y_j of the form

$$y_j = x_j^T \beta_i + e_j \tag{1}$$

where x_j is a $p \times 1$ vector of unit level auxiliary variables for unit *j*, β_i is a $p \times 1$ vector of area-specific regression coefficients and e_j is a unit level random effect with mean zero and variance σ_j^2 that is uncorrelated between different population units. We do not assume anything about σ_j^2 at this point. Suppose also that there is a known number N_i of population units in area *i*, with n_i of these sampled. The total number of units in the population is $N = \sum_{i=1}^{D} N_i$, with corresponding total sample size $n = \sum_{i=1}^{D} n_i$. In what follows, we use *s* to denote the collection of units in sample, with s_i the subset drawn from area *i*, and use expressions like $j \in i$ and $j \in s$ to refer to the units making up area *i* and sample *s* respectively. Note that throughout this paper we assume that the sampling method used is uninformative for the population values of *Y* given the corresponding values of the auxiliary variables and knowledge of the area affiliations of the population units. As a consequence, (1) can be assumed to apply at both sample and population level.

Let $\{w_{ij}; j \in s\}$ denote a set of weights such that $\hat{m}_i = \sum_{j \in s} w_{ij} y_j$ is a linear predictor of the mean value $m_i = N_i^{-1} \sum_{j \in i} y_j$ of area *i*, with $\sum_{j \in s} w_{ij} = 1$, so that $w_{ij} = O(n^{-1})$ in general. The bias of $\hat{m}_i$ under (1) is then

$$E(\hat{m}_i - m_i) = \left(\sum_{h=1}^{D} \sum_{j \in s_h} w_{ij} x_j^T \beta_h \right) - \bar{x}_i^T \beta_i . \tag{2}$$

Here s_h denotes the sample units from area *h* and $\bar{x}_i$ denotes the vector of average values of the auxiliary variables in area *i*. Similarly, the prediction variance of $\hat{m}_i$ under (1) is

$$Var(\hat{m}_i - m_i) = N_i^{-2} \left\{ \sum_{h=1}^{D} \sum_{j \in s_h} a_{ij}^2 \sigma_j^2 + \sum_{j \in r_i} \sigma_j^2 \right\} \tag{3}$$

where r_i denotes the non-sampled units in area *i* and $a_{ij} = N_i w_{ij} - I(j \in i)$. We use $I(A)$ to denote the indicator function for event *A*, so $I(j \in i)$ takes the value 1 if population unit *j* is from area *i* and is zero otherwise. Note that since a_{ij} is $O(N_i n^{-1})$ in general, the first term within the braces in (3) is the leading term of this prediction variance.

Let sample unit j be from area *h*. We consider the important special case of linear estimation of $\mu_j = E(y_j \mid x_j) = x_j^T \beta_h$ under (1). That is, we assume an estimator $\hat{\mu}_j = \sum_{k \in s} \phi_{kj} y_k$ where the 'weights' ϕ_{kj} are $O(n^{-1})$. Then

$$y_j - \hat{\mu}_j = (1 - \phi_{jj}) y_j - \sum_{k \in s(-j)} \phi_{kj} y_k$$

and so

$$Var(y_j - \hat{\mu}_j) = \sigma_j^2 \left\{ (1 - \phi_{jj})^2 + \sum_{k \in s(-j)} \phi_{kj}^2 (\sigma_k^2 / \sigma_j^2) \right\} \tag{4}$$

under (1). Here $s(-j)$ denotes the sample s with unit j excluded. If in addition $\hat{\mu}_j$ is unbiased for μ_j under (1), i.e.

$$E(y_j - \hat{\mu}_j) = 0 \tag{5}$$

then we can adopt the approach of Royall and Cumberland (1978) and estimate (3) by

$$\hat{V}(\hat{m}_i) = N_i^{-2}\left\{\sum_{h=1}^{D}\sum\nolimits_{j\in s_h} a_{ij}^2\hat{\lambda}_j^{-1}(y_j-\hat{\mu}_j)^2 + \sum\nolimits_{j\in r_i}\hat{\sigma}_j^2\right\} \tag{6}$$

where

$$\hat{\lambda}_j = (1-\phi_{jj})^2 + \sum\nolimits_{k\in s(-j)}\hat{\gamma}_{kj}\phi_{kj}^2$$

and $\hat{\gamma}_{kj} = \hat{\sigma}_k^2 / \hat{\sigma}_j^2$. Usually, the estimates $\hat{\sigma}_j^2$ of the residual variances in (6) are derived under a 'working model' refinement to (1). In the situation of most concern to us, where the sample sizes within the different areas are too small to reliably estimate area-specific variability, a pooling assumption can be made, i.e. $\sigma_j^2 = \sigma^2$, in which case we put

$$\hat{\sigma}_j^2 = \hat{\sigma}^2 = n^{-1}\sum\nolimits_{j\in s}\left\{(1-\phi_{jj})^2 + \sum\nolimits_{k\in s(-j)}\phi_{kj}^2\right\}^{-1}(y_j-\hat{\mu}_j)^2 \tag{7}$$

and so (6) becomes

$$\hat{V}(\hat{m}_i) = N_i^{-2}\sum\nolimits_{j\in s}\left\{a_{ij}^2 + (N_i-n_i)n^{-1}\right\}\hat{\lambda}_j^{-1}(y_j-\hat{\mu}_j)^2 \tag{8}$$

where now $\hat{\lambda}_j = (1-\phi_{jj})^2 + \sum\nolimits_{k\in s(-j)}\phi_{kj}^2$. Since any assumptions regarding σ_j^2 in the working model extension of (1) only affect second order terms in (3), the estimator (8) is robust to misspecification of the second order moments of this working model.

A corresponding estimator of the MSE of $\hat{m}_i$ under (1) follows directly. This is

$$\hat{M}(\hat{m}_i) = \hat{V}(\hat{m}_i) + \hat{B}^2(\hat{m}_i)\,. \tag{9}$$

where

$$\hat{B}(\hat{m}_i) = \sum_{h=1}^{D}\sum\nolimits_{j\in s_h} w_{ij}\hat{\mu}_j \;- N_i^{-1}\sum\nolimits_{j\in i}\hat{\mu}_j \tag{10}$$

is the obvious unbiased estimator of (2).

Use of the square of the unbiased estimator (10) of the bias of $\hat{m}_i$ in the mean squared error estimator (9) can be criticised because this term is not itself unbiased for the squared bias term in the mean squared error. This can be corrected by replacing (9) by

$$\hat{M}(\hat{m}_i) = \hat{V}(\hat{m}_i) + \hat{B}^2(\hat{m}_i) - \hat{V}\left\{\hat{B}(\hat{m}_i)\right\}. \tag{11}$$

where $\hat{V}\left\{\hat{B}(\hat{m}_i)\right\}$ is a consistent estimator of the variance of (10). However, small area sample sizes may lead to this estimate becoming quite unstable, and so users may still prefer (9) over (11). Obviously (9) then represents a conservative estimator of the MSE of $\hat{m}_i$ under (1).

2.2 Average MSE Estimation under a Linear Mixed Model

The approach to MSE estimation outlined so far is *area specific*, or just specific, since the objective is to estimate the MSE of $\hat{m}_i$ on the basis that different linear models hold in different areas. As we saw in the previous sub-section, this depends on our ability to unbiasedly (or at least consistently) estimate the unit level specific expected value μ_j defined by (1). Small sample sizes in the areas of interest can make this difficult. An alternative strategy is to extend (1) by also assuming that the area specific regression parameters β_i are independent and identically distributed realisations of a random variable with expected value β and covariance matrix $\mathbf{\Omega}$. Since $\mathbf{\Omega}$ does not need to be of full rank, this is an example of a linear mixed model. Rather than estimating the MSE of $\hat{m}_i$ under (1), our aim now is to estimate the expected value of this mean squared error with respect to the distribution of the β_i. We refer to this as the *area averaged*, or just average, MSE of $\hat{m}_i$ in what follows.

Let E_B and Var_B denote expectation and variance respectively with respect to the distribution of the β_i. The average bias of $\hat{m}_i$ is then

$$\bar{B}(\hat{m}_i) = E_B\left(\sum_{h=1}^{D}\sum\nolimits_{j\in s_h} w_{ij}\mu_j - \bar{\mu}_i\right) = \left(\sum_{h=1}^{D}\sum\nolimits_{j\in s_h} w_{ij}x_j - \bar{x}_i\right)^T \beta \tag{12}$$

where $\bar{\mu}_i$ is the area i average of the μ_j, with $\bar{x}_i$ the vector of corresponding averages of the covariate X in (1). This bias is zero when the weights $\{w_{ij}; j \in s\}$ defining $\hat{m}_i$ are 'locally calibrated' on these covariates. That is,

$$\sum_{j \in s} w_{ij} x_j = \bar{x}_i . \tag{13}$$

Let w_i denote the p-vector of weights $\{w_{ij}; j \in s\}$. Then (13) is satisfied when

$$w_i = N_i^{-1} \left\{ \Delta_{is} + \mathbf{T}_i \left(N_i \bar{x}_i - n_i \bar{x}_{is} \right) + \left(\mathbf{I}_n - \mathbf{T}_i \mathbf{X}_s' \right) a_i \right\}. \tag{14}$$

Here $\mathbf{X}_s$ denotes the $n \times p$ matrix of sample X-values, $\mathbf{T}_i$ is a $n \times p$ matrix such that $\mathbf{X}_s' \mathbf{T}_i$ is the identity matrix of order p, $\bar{x}_{is}$ is the sample mean of X for area i and a_i is an arbitrary vector of size n.

Special cases of (14) are the weights that define the best linear unbiased predictor (BLUP) of m_i (Royall, 1976) under the linear mixed model implied by the assumption that the β_i are random and the weights defining the predictor for m_i generated by integrating either the Chambers and Dunstan (1986) estimator of the distribution function of Y in area i or the related estimator suggested by Rao, Kovar and Mantel (1990) under simple random sampling within the areas. See Tzavidis and Chambers (2007). Note that the weights defining the model-based direct estimator (MBDE) of m_i under this implied linear mixed model (Chambers and Chandra, 2006) do not satisfy (13), and so require an estimate of the average bias (12) to be calculated. This is straightforwardly accomplished by substituting an estimate $\hat{\beta}$ of β on the right hand side of (12). Such an estimate can be obtained by fitting the linear mixed model to the sample data. The resulting estimator of the average bias of $\hat{m}_i$ is then

$$\hat{\bar{B}}(\hat{m}_i) = \left(\sum_{h=1}^{D} \sum_{j \in s_h} w_{ij} x_j - \bar{x}_i \right)^T \hat{\beta} . \tag{15}$$

Turning now to the average MSE of $\hat{m}_i$, we note that it is equal to the average prediction variance of this statistic plus the square of its average bias (12), with the average prediction variance defined by $E_B\{Var(\hat{m}_i - m_i)\} + Var_B\{E(\hat{m}_i - m_i)\}$. Since (3) does not depend on β_i it is clear that $E_B\{Var(\hat{m}_i - m_i)\} = Var(\hat{m}_i - m_i)$ and so this term can be estimated by (8). In contrast,

$$Var_B\{E(\hat{m}_i - m_i)\} = \sum_{h=1}^{D} A_{hi}^T \Omega A_{hi} \tag{16}$$

where $A_{hi} = I(h \neq i)\left(\sum_{j \in s_h} w_{ij} x_j\right) + I(h = i)\left(\sum_{j \in s_h} w_{ij} x_j - \bar{x}_i\right)$. In most cases the variation among the y_j values within the areas will be much greater than that between the β_i, and so (16) will be of smaller order of magnitude than $Var(\hat{m}_i - m_i)$. One option therefore is to simply ignore this term when estimating the unconditional prediction variance. Another is to substitute a model-dependent estimator $\hat{\Omega}$ of Ω in (16). This term can be added to (8) to define a robust estimator of the average prediction variance of $\hat{m}_i$, of the form

$$\hat{\bar{V}}(\hat{m}_i) = \hat{V}(\hat{m}_i) + \sum_{h=1}^{D} A_{hi}^T \hat{\Omega} A_{hi} \tag{17}$$

where, as before, $\hat{V}(\hat{m}_i)$ is calculated using (8), and the claimed robustness is a consequence of the fact that this estimator is both the leading term in (17) and, by construction, robust to assumptions about the second order moments of the y_j. Combining (15) and (17), our estimator of the average MSE of $\hat{m}_i$ is therefore

$$\hat{\bar{M}}(\hat{m}_i) = \hat{\bar{V}}(\hat{m}_i) + \hat{\bar{B}}^2(\hat{m}_i). \tag{18}$$

Note that a correction to the estimate of the squared bias in (18), analogous to that used in (12) is possible. However, in this case the extra term is $O(n^{-1})$ and typically negligible.

Using $\hat{V}(\hat{m}_i)$ in (17) requires access to unbiased estimators of the area specific individual expected values μ_j. As we have already noted, such estimators may be unstable when area sample sizes are small. One might therefore be tempted to finesse this requirement of specific unbiasedness by replacing it by one of average unbiasedness. That is, we could require that

$$E_B\left\{E(\hat{\mu}_j - \mu_j)\right\} = E_B\left(\sum_{g=1}^{D}\sum\nolimits_{k\in s_g}\phi_{kj}x_k^T\beta_g - x_j^T\beta_h\right) = \left(\sum\nolimits_{k\in s}\phi_{kj}x_k - x_j\right)^T\beta = 0\,. \qquad (19)$$

The last identity on the right hand side of (19) will be true for any linear estimator $\hat{\mu}_j$ of μ_j where

$$\phi_{kj} = a_j + \left\{x_j - a_j\left(\sum\nolimits_{t\in s}x_t\right)\right\}^T\left(\sum\nolimits_{t\in s}E_{tj}x_t^T\right)^{-1}E_{kj} \qquad (20)$$

with a_j a suitably chosen constant and the $\left\{E_{kj}; k\in s\right\}$ suitably chosen *p*-vectors. In particular, any weighted least squares estimator of μ_j satisfies (20). Similarly, the BLUP of a non-sampled unit *j* is defined by weights that satisfy (20).

Unfortunately, estimators of μ_j that are unbiased on average can lead to biased estimators of $Var(y_j - \mu_j)$, and, by extension, biased estimation of $Var(\hat{m}_i - m_i)$. To illustrate this, we note that $\hat{V}(\hat{m}_i)$ uses $(y_j - \hat{\mu}_j)^2$ as an estimator of $E(y_j - \mu_j)^2$. The bias in this estimator is therefore

$$\begin{aligned} E(y_j - \hat{\mu}_j)^2 - E(y_j - \mu_j)^2 &= -2E(y_j - \mu_j)(\hat{\mu}_j - \mu_j) + E(\hat{\mu}_j - \mu_j)^2 \\ &= -E\left\{(\hat{\mu}_j - \mu_j)(2y_j - \mu_j - \hat{\mu}_j)\right\} \end{aligned}$$

so we anticipate that $\hat{V}(\hat{m}_i)$ will be negatively biased if $E\left\{(\hat{\mu}_j - \mu_j)(2y_j - \mu_j - \hat{\mu}_j)\right\}$ is positive and vice versa. Now let $j\in i$ and consider the special case of a random intercept model for y_j, i.e. $y_j = x_j^T\beta + u_i + e_j$ where u_i is a random area *i* effect and e_j is a random individual effect uncorrelated with u_i. Here $\mu_j = x_j^T\beta + u_i$. Suppose that we have a large overall sample size, allowing us to replace $\hat{\beta}$ by β. The BLUP-based estimator $\hat{\mu}_j = x_j^T\hat{\beta} + \gamma_i(\bar{y}_{is} - \bar{x}_{is}^T\hat{\beta})$ of μ_j is

then unbiased on average, and can be approximated by $\tilde{\mu}_j = x_j^T\beta + \gamma_i u_i$, where γ_i is a 'shrinkage' factor. It follows that

$$(\tilde{\mu}_j - \mu_j)(2y_j - \mu_j - \tilde{\mu}_j) = 2u_i(\gamma_i - 1)e_i - u_i^2(\gamma_i - 1)^2$$

so $E(y_j - \hat{\mu}_j)^2 - E(y_j - \mu_j)^2 \approx (\gamma_i - 1)^2\sigma_u^2$. That is, we expect $\hat{V}(\hat{m}_i)$ to be positively biased if we use the 'unbiased on average' BLUP-based estimator of μ_j to define $\hat{\mu}_j$. We also note that this bias disappears (approximately) if we 'unshrink' the residual component in this estimator. In the case of the random intercepts model, this suggests that when $j \in i$ we use

$$\hat{\mu}_j = x_j^T\hat{\beta} + (\bar{y}_{is} - \bar{x}_{is}^T\hat{\beta}) = \bar{y}_{is} + (x_j - \bar{x}_{is})^T\hat{\beta}.$$

2.3 MSE Estimation - Three Special Cases

In this section we apply the ideas set out in the previous two sub-sections to MSE estimation for three different types of linear predictors of the small area mean m_i.

MSE estimation for the EBLUP

We first consider the EBLUP for m_i based on a unit level linear mixed model extension of (1) of the form

$$\mathbf{y}_i = [\mathbf{X}_{1i}\ \mathbf{X}_{2i}]\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix} + \mathbf{X}_{1i}u_{1i} + \mathbf{e}_i = \mathbf{X}_i\beta + \mathbf{G}_i u_{1i} + \mathbf{e}_i \qquad (21)$$

where $\mathbf{y}_i$ is the N_i-vector of population values of y_j in area *i*, $\mathbf{X}_i$ is the corresponding $N_i \times p$ matrix of auxiliary variable values x_j, $\mathbf{G}_i$ is the $N_i \times q$ component of $\mathbf{X}_i$ corresponding to the *q* random components of β, u_{1i} is the associated *q*-vector of area-specific random effects and $\mathbf{e}_i$ is the N_i-vector of individual random effects. It is typically assumed that the area and individual effects are mutually independent, with the area effects independently and identically distributed as $N(\mathbf{0},\mathbf{\Omega}_1)$ and the individual effects independently and identically distributed as $N(0,\sigma^2)$. See Rao (2003, chapter 6) for development of the underlying theory of this predictor. Here we just

note that the EBLUP can be written as the indirect linear predictor defined by the n-vector of sample weights

$$\mathbf{w}_s^{(i)} = (w_{ij}) = N_i^{-1}\left[\Delta_n^{(i)} + \left\{\hat{\mathbf{H}}_s^T\mathbf{X}_r^T + \left(\mathbf{I}_n - \hat{\mathbf{H}}_s^T\mathbf{X}_s^T\right)\hat{\mathbf{\Sigma}}_{ss}^{-1}\hat{\mathbf{\Sigma}}_{sr}\right\}\Delta_{N-n}^{(i)}\right] \tag{22}$$

where $\Delta_n^{(i)}$ ($\Delta_{N-n}^{(i)}$) is the vector of size n ($N-n$) that 'picks out' the sampled (non-sampled) units in area i, $\mathbf{X}_s$ and $\mathbf{X}_r$ are the matrices of order $n \times p$ and $(N-n)\times p$ respectively of the sample and non-sample values of the auxiliary variables, $\mathbf{I}_n$ is identity matrix of order n, $\mathbf{G}_s = diag(\mathbf{G}_{is})$ and $\mathbf{G}_r = diag(\mathbf{G}_{ir})$ where $\mathbf{G}_{is}$ ($\mathbf{G}_{ir}$) is the sample (non-sample) component of $\mathbf{G}_i$, $\hat{\mathbf{\Sigma}}_{ss} = \hat{\sigma}^2\mathbf{I}_n + \mathbf{G}_s diag\left\{\hat{\mathbf{\Omega}}_1\right\}\mathbf{G}_s^T$, $\hat{\mathbf{\Sigma}}_{sr} = \mathbf{G}_s diag\left\{\hat{\mathbf{\Omega}}_1\right\}\mathbf{G}_r^T$ and $\hat{\mathbf{H}}_s = \left(\mathbf{X}_s'\hat{\mathbf{\Sigma}}_{ss}^{-1}\mathbf{X}_s\right)^{-1}\mathbf{X}_s'\hat{\mathbf{\Sigma}}_{ss}^{-1}$. Here $\hat{\sigma}^2$ and $\hat{\mathbf{\Omega}}_1$ are suitable (e.g. ML or REML) estimates of the variance components of (21). Note also that, in terms of the notation used in the previous sub-section, e.g. in (17),

$$\hat{\mathbf{\Omega}} = \begin{bmatrix} \hat{\mathbf{\Omega}}_1 & \mathbf{0}_{q\times(p-q)} \\ \mathbf{0}_{(p-q)\times q} & \mathbf{0}_{(p-q)\times(p-q)} \end{bmatrix}.$$

Given this setup, estimation of the area-specific MSE of the EBLUP can be carried out using (9) with weights defined by (22). Since the EBLUP weights (22) are locally calibrated, (12) is zero and consequently the area-averaged MSE of the EBLUP is estimated using (17).

MSE estimation for the MBDE

The second predictor we consider is the MBDE suggested by Chambers and Chandra (2006). This is a direct estimator that is an alternative to the indirect EBLUP. The weights used in the MBDE are of the form

$$w_{ij} = \frac{I(j \in s_i)\tilde{w}_j}{\sum_{k\in s_i}\tilde{w}_k}. \tag{23}$$

where the $\tilde{w}_j$ are the weights that define the EBLUP for the population total of the y_j under (21), i.e.

$$\tilde{\mathbf{w}} = (\tilde{w}_j) = \mathbf{1}_n + \left\{ \hat{\mathbf{H}}_s^T \mathbf{X}_r^T + \left(\mathbf{I}_n - \hat{\mathbf{H}}_s^T \mathbf{X}_s^T \right) \hat{\mathbf{\Sigma}}_{ss}^{-1} \hat{\mathbf{\Sigma}}_{sr} \right\} \mathbf{1}_{N-n} \tag{24}$$

where $\mathbf{1}_n$ ($\mathbf{1}_{N-n}$) denotes the unit vector of size n ($N-n$). In this case estimation of the area-specific MSE of the MBDE is carried out using (9), with weights defined by (23). Similarly, its area-averaged MSE is estimated using (18) with the same weights.

Both the EBLUP and the MBDE are based on the linear mixed model (21). As a consequence both approaches implicitly use biased estimators of the area specific unit level expectations μ_j. As noted earlier, use of such estimators to define residuals relative to (1) can bias the robust estimator (8) of the area specific prediction variance. Assuming unit *j* is from area *i*, we therefore replace these 'shrunken' estimators of μ_j by approximately unbiased ('unshrunken') estimators of the form

$$\hat{\mu}_j = x_j^T \hat{\mathbf{H}}_s \mathbf{y}_s + g_j^T \left(\mathbf{G}_{is}^T \mathbf{G}_{is} \right)^{-1} \mathbf{G}_{is}^T \left(\mathbf{y}_{is} - \mathbf{X}_{is} \hat{\mathbf{H}}_s \mathbf{y}_s \right) \tag{25}$$

when calculating (8) for either the EBLUP or the MBDE. Here g_j is the vector of q random components of x_j, and $\mathbf{X}_{is}$ and $\mathbf{y}_{is}$ are the area *i* components of $\mathbf{X}_s$ and $\mathbf{y}_s$ respectively. Note that in this case $\hat{\mu}_j = \sum_{k \in s} \phi_{kj} y_k$ where

$$\phi_{kj} = \begin{cases} c_{ijsk} + d_{ijsk}; k \in i \\ c_{ijsk}; k \notin i. \end{cases}$$

with

$$\mathbf{c}_{ijs} = \left(c_{ijsk}; k \in s \right) = \hat{\mathbf{H}}_s^T \left\{ x_j - \mathbf{X}_{is}^T \mathbf{G}_{is} \left(\mathbf{G}_{is}^T \mathbf{G}_{is} \right)^{-1} g_j \right\}$$

and

$$\mathbf{d}_{ijs} = \left(d_{ijsk}; k \in s_i \right) = \mathbf{G}_{is} \left(\mathbf{G}_{is}^T \mathbf{G}_{is} \right)^{-1} g_j .$$

MSE estimation for the M-quantile predictor

The third predictor we consider is based on the M-quantile modelling approach described in Chambers and Tzavidis (2006). This approach does not assume an underlying linear mixed

model, relying instead on characterising the relationship between y_j and x_j in area *i* in terms of the linear M-quantile model that best 'fits' the sample y_j values from this area. Thus, the alternative to (21) underpinning this approach is a model of the form

$$\mathbf{y}_i = \mathbf{X}_i \beta(q_i) + \mathbf{e}_i \tag{26}$$

where $\mathbf{e}_i$ is a vector representing within area variability not accounted for by variability in the x_j and q_i denotes the quantile coefficient of this area specific linear M-quantile fit. Given an estimate $\hat{q}_i$ of q_i, an iteratively reweighted least squares (IRLS) algorithm is used to calculate an estimate $\hat{\beta}(\hat{q}_i)$ of $\beta(q_i)$, and a non-sample value of y_j in area *i* then predicted as $\hat{y}_j = x_j^T \hat{\beta}(\hat{q}_i)$. This is equivalent to using an indirect predictor of m_i, with weights

$$\mathbf{w}_s^{iMQ1} = N_i^{-1}\left[\Delta_n^{(i)} + \mathbf{W}_s(\hat{q}_i)\mathbf{X}_s\left\{\mathbf{X}_s'\mathbf{W}_s(\hat{q}_i)\mathbf{X}_s\right\}^{-1}\mathbf{X}_r\Delta_{N-n}^{(i)}\right]. \tag{27}$$

Here $\mathbf{W}_s(\hat{q}_i)$ is the diagonal matrix of final weights used in the IRLS algorithm. Tzavidis and Chambers (2007) observe that the 'naive' M-quantile predictor with weights defined by (27) can be biased for m_i, particularly when the marginal distribution of y_j in area *i* is skewed. They therefore suggest a bias-correction to this predictor, which replaces (27) by

$$\mathbf{w}_s^{iMQ2} = n_i^{-1}\Delta_n^{(i)} + (1 - N_i^{-1}n_i)\mathbf{W}_s(\hat{q}_i)\mathbf{X}_s\left\{\mathbf{X}_s'\mathbf{W}_s(\hat{q}_i)\mathbf{X}_s\right\}^{-1}(\bar{x}_{ir} - \bar{x}_{is}) \tag{28}$$

where $\bar{x}_{is}$ and $\bar{x}_{ir}$ are the vectors of sample and non-sample means of the x_j in area *i*. It is not difficult to show that the weights defined by (28) are locally calibrated. Since (26) is an area-specific model, the area-specific MSE of the bias-corrected M-quantile predictor defined by the weights (28) can be estimated using (8). Similarly, estimation of the area-averaged MSE of this predictor can be based on (17). Finally, we observe that, in our experience, the constant $\hat{\lambda}_j$ in (8) is typically very close to one under M-quantile prediction. We therefore set it equal to this value whenever we compute values of (8) that relate to small area prediction under the M-quantile modelling approach.

Comparing these three different small area weighting methods, we see that since the EBLUP weights (22) and the M-quantile weights (28) are locally calibrated, the predictors $\hat{m}_i$ based on these weights are unbiased on average, i.e. unbiased under (21). However, since the model (26) is essentially an alternative to (1) in terms of parameterising differences in the regression relationship between areas, only the M-quantile predictor defined by (28) can be considered to be also at least approximately unbiased with respect to (1). In contrast, since the MBDE weights (23) are not locally calibrated, the MBDE predictor of m_i is biased on average. This predictor is therefore also biased with respect to the area specific model (1). Table 1 summarises how these different bias properties are accounted for when estimating the MSE of these three different types of small area estimators.

3. SIMULATION STUDIES OF THE PROPOSED MSE ESTIMATOR

In this section we describe results from four simulation studies of the performance of the approach to robust MSE estimation described in the previous section. Two of these studies use model-based simulation, with population data generated from the linear mixed model (21). The remaining two use design-based simulation, with population data derived from two real life situations where linear small area estimation is of interest.

In all four studies, the performance of an estimator of the MSE of a linear predictor of a small area mean was evaluated with respect to two basic criteria – the Monte Carlo bias of the MSE estimator and the Monte Carlo coverage of nominal 95% prediction intervals for the small area mean generated using the MSE estimator. The bias was measured by *%AvRB* and *%MedRB*, where

$$\%AvRB = \underset{i}{mean}\left\{ M_i^{-1}\left(K^{-1}\sum_{k=1}^{K}\hat{M}_{ik} \right) - 1 \right\} \times 100$$

with *%MedRB* defined similarly, but with the mean over the small areas replaced by the median. Coverage performance for prediction intervals was measured by *%AvCR* and *%MedCR*, where

$$\%AvCR = \underset{i}{mean}\left\{K^{-1}\sum_{k=1}^{K} I\left(\left|\hat{m}_{ik} - m_{ik}\right| \le 2\hat{M}_{ik}^{1/2}\right)\right\} \times 100$$

and again *%MedCR* differs from *%AvCR* only by the use of median rather than mean when averaging over the small areas. Note that the subscript of k here indexes the K Monte Carlo simulations, with m_{ik} denoting the value of the small area i mean in simulation k (this is a fixed population value in the design-based simulations), and $\hat{m}_{ik}$, $\hat{M}_{ik}$ denoting the area i estimated value and corresponding estimated MSE in simulation k. The actual MSE value (i.e. the average squared error over the simulations) is denoted M_i.

3.1 Model-Based Simulations

The first model-based simulation study used a population size of N = 15,000, with D = 30 small areas. Population sizes in the small areas were uniformly distributed over the interval [443, 542] and were kept fixed over simulations. Population values for Y were generated under the random intercepts model $y_j = 500 + 1.5x_j + u_i + e_j$, with x_j drawn from a chi squared distribution with 20 degrees of freedom. The area effects u_i and individual effects e_j were independently drawn from $N(0,\sigma_u^2)$ and $N(0,\sigma_e^2)$ distributions respectively, with σ_u^2 and σ_e^2 chosen so that $\rho = \sigma_u^2 / (\sigma_u^2 + \sigma_e^2)$ took the range of values shown in Table 2. A sample of size $n = 600$ was then selected from the simulated population, with area sample sizes proportional to the fixed area populations. Sampling was via stratified random sampling, with the strata defined by the small areas. The simulation itself was model-based, with all population values independently regenerated at each simulation and an independent sample drawn from the population each time. A total of K = 1000 simulations were carried out.

Table 2 shows the bias and coverage values obtained for this study. Note that for the EBLUP we show results for two MSE estimators. The PR estimator corresponds to the usual Prasad-Rao estimator of the MSE of the EBLUP (see Rao, 2003, section 6.2.6), while the Robust estimator

corresponds to the estimator of the specific MSE of the EBLUP defined in Table 1. Similarly, the Robust estimators of the specific MSE of the MBDE and the M-quantile predictor are defined by their entries in Table 1.

Not surprisingly, given that all its underlying assumptions are met, the PR estimator of MSE does very well in this study, with virtually no bias and empirical coverage equal to the nominal value in all cases. However, the Robust estimation method also performs well, showing only a very small negative bias and a small amount of undercoverage for all three linear predictors.

Conditions for the second model-based simulation study were the same as in the first, with the exception that the area level random effects and the individual level random effects were now independently drawn from chi squared distributions. In particular, for the combinations (a,b) shown in Table 3, values for the area effects u_i were drawn from a mean corrected chi squared distribution with b degrees of freedom and values for individual effects e_j were drawn from a mean corrected chi squared distribution with a degrees of freedom. Table 3 shows the bias and coverage values that were obtained in this second study, and we again note the good performance of all methods of MSE estimation, with the PR estimator marginally better than the different versions of the Robust estimator. This good performance of the PR estimator, even though the usual Gaussian assumptions for random terms in the linear mixed model do not hold, illustrates the insensitivity of this estimator to skewed error distributions.

The results set out in Tables 2 and 3 are averaged over the small areas and hence can obscure area level differences in performance. In Figure 1 we show the area specific values of the square root of the MSE (RMSE) that were obtained in the first model-based simulation study when $\rho = 0.2$. Similarly, in Figure 2 we show these area specific RMSE values for the second model-based simulation study when $(a,b) = (2,1)$. In both Figures we also show the average of the area-specific estimated RMSE values that were generated by the PR estimator for the EBLUP and by the Robust estimator for all three prediction methods. We see that the PR and Robust estimators

behave very similarly for the EBLUP, appearing to track the average RMSE of this predictor over the areas rather than its area specific RMSE. Similar behaviour by the Robust estimator when applied to the M-quantile predictor can also be observed. In contrast, in both Figures 1 and 2 we see that the Robust estimator performs very well in terms of tracking the area specific RMSE of the MBDE.

3.2 Design-Based Simulations

We report results from two design-based simulation studies, both based on real life populations. The first involved a sample of 3591 households spread across D = 36 districts of Albania that participated in the 2002 Albanian Living Standards Measurement Study. This sample was bootstrapped to create a population of $N = 724,782$ households by re-sampling with replacement with probability proportional to a household's sample weight. A total of $K = 1000$ independent stratified random samples were then drawn from this bootstrap population, with total sample size equal to that of the original sample and with districts defining the strata. Sample sizes within districts were the same as in the original sample, and varied between 8 and 688. The Y variable of interest was household per capita consumption expenditure and X was defined by three zero-one variables (ownership of television, parabolic antenna and land). The aim was to predict the average value of Y for each district and both the EBLUP and the MBDE used a random intercepts model with district level effects.

The second design-based simulation study was based on the same population of Australian broadacre farms as that used in the simulation studies reported in Chambers and Chandra (2006) and Chambers and Tzavidis (2006). This population was defined by bootstrapping a sample of 1652 farms that participated in the Australian Agricultural and Grazing Industries Survey (AAGIS) up to a population of $N = 81,982$ farms by re-sampling from the original AAGIS sample with probability proportional to a farm's sample weight. The small areas of interest in this case were the D = 29 broadacre farming regions represented in this sample. The design-

based simulation was carried out by selecting $K = 1000$ independent stratified random samples from this bootstrap population, with strata defined by the regions and with stratum sample sizes defined by those in the original AAGIS sample. These sample sizes vary from 6 to 117. Here Y is Total Cash Costs associated with operation of the farm, and X is a vector that includes farm area, effects for six post-strata defined by three farming climatic zones and two size bands as well as the interactions of these variables. Again, both the EBLUP and the MBDE used a random intercepts structure to define region effects and the aim was to predict the regional averages of Y.

Tables 4 and 5 show the values of *%AvRB*, *%MedRB*, *%AvCR* and *%MedCR* for the different MSE estimators investigated in these two simulation studies. We see immediately that the PR estimator of the MSE of the EBLUP has a substantial upward bias in both sets of design-based simulations. This corroborates comments by other authors (e.g. Longford, 2007) about the poor design-based properties of this estimator. In contrast, for the Albanian household population all three versions of the Robust estimator are essentially unbiased, while for the AAGIS farm population the Robust estimator is unbiased for the M-quantile predictor and biased upwards for the EBLUP and MBDE predictors, though not to the same extent as the PR estimator. In fact, if one uses *%MedRB* to measure this bias, then it is only the MSE estimators for the EBLUP (both PR and Robust) that show an upward bias.

An insight into the reasons for this difference in behaviour can be obtained by examining the area specific RMSE values displayed in Figures 3 and 5. Thus, in Figure 3 we see that all three Robust estimators track the district-specific design-based RMSEs of their respective predictors exceptionally well. In contrast, the PR estimator does not seem to be able to capture between district differences in the design-based RMSE of the EBLUP. In contrast, in Figure 5 we see that the Robust estimator of the MSE of the M-quantile predictor performs extremely well in all regions, with the corresponding estimator of the MSE of the MBDE also performing well in all regions except one (region 6) where it substantially overestimates the design-based RMSE of this

predictor. This region is noteworthy because samples taken from it that are unbalanced with respect to farm size lead to negative weights under the assumed linear mixed model. The value for *%MedRB* shown in Table 5 effectively excludes this region and consequently provides a better 'overall' picture of the performance of the Robust estimator for the MSE of the MBDE in this case. The picture becomes more complex when one considers the region specific RMSE estimation performance of the EBLUP in Figure 5. Here we see that the Robust estimator of the MSE of the EBLUP clearly 'tracks' the region specific design-based RMSE of this predictor better than the PR estimator, with the noteworthy exception of region 21, where it shows significant overestimation. This region contains a number of massive outliers (all replicated from a single outlier in the original AAGIS sample) and these lead to a 'blow out' in the value of Robust when they appear in sample (this can also be seen in the results for the MBDE and the M-quantile predictors). In contrast, with the exception of region 6 (where sample balance is a problem), there seems to be little regional variation in the value of the PR estimator of MSE.

Coverage of prediction intervals based on a particular MSE estimation method is an important aspect of the performance of the method. Values of *%AvCR* and *%MedCR* set out in Table 4 indicate that for Albanian household population all methods of MSE estimation perform well on average. In contrast, for the AAGIS farms population, we see in Table 5 that there is some undercoverage, although the differences between *%AvCR* and *%MedCR* in this table indicate that there may be region specific differences in coverage underpinning these average values. This observation is confirmed when we examine Figures 4 and 6. In particular, in Figure 4 we see that the Robust estimator of MSE 'tracks' a nominal 95% coverage value in different districts much better than the PR estimator of MSE. In contrast, in Figure 6 all MSE estimators seem to perform similarly in all regions, with the PR estimator showing a degree of overcoverage and the Robust estimators showing a degree of undercoverage. The exceptions are region 3 where the PR estimator records substantial undercoverage, and the outlier contaminated region 21 where the

Robust estimators have the same problem. Overall, if one takes coverage below 75% as 'unacceptable', then the PR estimator is unacceptable in 5 out of the 29 regions, while the Robust estimators perform slightly better, being unacceptable in 3 out of the 29 regions.

4. CONCLUSIONS AND DISCUSSION

In this paper we propose a robust and easily implemented method of estimating the mean squared error of linear estimators of small area means (and totals). The empirical results described in section 3 are evidence that this method has promise. In particular, it generally worked well in terms of estimating both model-based and design-based MSE for the three rather different linear predictors that we investigated in our simulations. This was in contrast to the more complex model-dependent approach underpinning the Prasad-Rao estimator of the MSE of the EBLUP, which worked well in our model-based simulations but then essentially failed in our design-based simulations. By implication, use of these robust methods to estimate the corresponding area-averaged MSE (see section 2.2), which is based on the 'usual' assumption that area effects are essentially exchangeable, should also work well generally.

The extension of the robust MSE approach to non-linear small area estimation situations remains to be done. However, since this approach is closely linked to robust population level MSE estimation based on Taylor series linearisation, it should be possible to develop appropriate extensions for corresponding small area non-linear estimation methods. Although the relevant results are not provided here, some evidence for this is that the robust MSE estimation method described in section 2.1 has already been used to estimate the MSE of the MBDE with both categorical and non-linear Y variables and has performed well. See Chandra and Chambers (2006). More recently, the approach has also been successfully used to estimate the MSE of geographically weighted M-quantile small area estimators in situations where the small area values are spatially correlated (Salvati *et al*., 2007). It is of also of interest to examine whether this approach to MSE estimation can be used with predictors based on non-parametric small area

models (Opsomer *et al*., 2005) or with predictors based on outlier robust mixed effects models where the development of Prasad-Rao type MSE estimators is more difficult.

As is clear from the development in this paper, our preferred approach to MSE estimation assumes that the MSE of real interest is that defined by the area specific model (1). This is in contrast to the usual approach to defining MSE in small area estimation, which adopts the area-averaged MSE concept (see section 2.2) as the appropriate measure of the accuracy of a small area estimator. As pointed out by Longford (2007), the ultimate aim in small area estimation is to make inferences about small area characteristics conditional on the realised (but unknown) values of small area effects, i.e. with respect to (1). One can consider this to be a design-based objective (as in Longford, 2007), or, as we prefer, a model-based objective that does not quite fit into the usual random effects framework for small area estimation. In either case we are interested in variability that is with respect to fixed area specific expected values. This is consistent with the concept of variability that is typically applied in population level inference. As our simulations demonstrate this allows our MSE estimator to perform well from both a model-based as well as a design-based perspective.

A final comment concerns how our method of MSE estimation relates to more nonparametric, but computational intensive, methods like the jackknife and the bootstrap. Since population level versions of robust estimators of prediction variance like (8) are well known to be first order equivalent to corresponding jackknife estimators (Valliant, Dorfman and Royall, 2000, section 5.4.2), we anticipate that a similar relationship will hold for the MSE estimator (9) in the context of small area inference. The link to modern bootstrap methods for small area inference (e.g. Hall and Maiti, 2006) is less clear and requires further research.

REFERENCES

Chambers, R.L. and Dunstan, R. (1986). Estimating distribution functions from survey data, *Biometrika* **73**, 597 - 604.

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

Chambers, R.L. and Chandra H**.** (2006). Improved Direct Estimators for Small Areas. *Southampton Statistical Sciences Research Institute*, Methodology Working Papers, M06/07, University of Southampton, U.K.
(*Available from* http://eprints.soton.ac.uk/38465/)

Chandra, H. and Chambers, R. (2006). Small area estimation for skewed data. *Southampton Statistical Sciences Research Institute*, Methodology Working Papers, M06/05, University of Southampton, U.K.
(*Available from* http://eprints.soton.ac.uk/38417/)

Hall, P. and Maiti, T. (2006). Nonparametric estimation of mean squared prediction error in nested-error regression models. *Annals of Statistics* **34**, 1733-1750.

Henderson, C.R. (1953). Estimation of variance and covariance components. *Biometrics* **9**, 226-252.

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

Opsomer, J.D., Claeskens, G., Ranalli, M.G., Kauermann, G. and Breidt, F.J. (2005). Nonparametric small area estimation using penalized spline regression. Preprint Series #05-01, Department of Statistics, Iowa State University. To appear in the *Journal of the Royal Statistical Society, Series B*.

Prasad, N.G.N and Rao, J.N.K. (1990). The estimation of the mean squared error of small area estimators. *Journal of the American Statistical Association* **85**, 163-171.

Rao, J.N.K., Kovar, J.G. and Mantel, H.J. (1990). On estimating distribution functions and quantiles from survey data using auxiliary information. *Biometrika* **77**, 365-75.

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

Royall, R.M. (1976). The linear least squares prediction approach to two-stage sampling. *Journal of the American Statistical Association* **71**, 657 - 664.

Royall, R.M. & Cumberland, W.G. (1978). Variance estimation in finite population sampling. *Journal of the American Statistical Association* **73**, 351 - 358.

Salvati, N., Tzavidis, N., Chambers, R. and Pratesi, M. (2007). Small area estimation via M-quantile geographical weighted regression. *Paper submitted for publication. A working paper is available from the authors upon request.*

Tzavidis, N. and Chambers, R. (2007). Robust prediction of small area means and distributions. *Cathy Marsh Centre for Census and Survey Research*, CCSR Working Paper series, Paper 2007-08. University of Manchester, UK.
(*Available from* http://www.ccsr.ac.uk/publications/working/#2007-08)

Valliant, R., Dorfman, A.H. and Royall, R.M. (2000). Finite Population Sampling and Inference. Wiley, New York.

Table 1 Definitions of robust MSE estimators for three weighting methods.

Weighting Method	Definition of $\hat{\mu}_j, j \in i$	MSE Estimator	
		Under (1)	Under (21)
EBLUP, (23)	(25)	(9)	(17)
MBDE, (24)	(25)	(9)	(18)
M-quantile, (28)	$x_j^T \hat{\beta}(\hat{q}_i)$	(8) with $\hat{\lambda}_j = 1$	(17) with $\hat{\lambda}_j = 1$

Table 2 Model-based simulation results with Gaussian random effects

Weighting Method	MSE Estimator	$\rho = \sigma_u^2 / (\sigma_u^2 + \sigma_e^2)$			
		0.1	0.2	0.3	0.5
			%AvRB		
EBLUP, (23)	PR	-1.52	-0.85	-0.57	-0.28
	Robust	3.27	-0.32	-0.70	-0.51
MBDE, (24)	Robust	-0.98	-0.98	-0.87	-0.87
M-quantile, (28)	Robust	-1.69	-0.79	0.07	1.57
			%AvCR		
EBLUP, (23)	PR	95	95	95	95
	Robust	93	94	94	94
MBDE, (24)	Robust	95	95	95	95
M-quantile, (28)	Robust	93	93	94	94

Table 3 Model-based simulation results with Chi squared random effects

Weighting Method	MSE Estimator	(*a*,*b*)			
		(5,1)	(2,1)	(5,2)	(5,5)
			%AvRB		
EBLUP, (23)	PR	2.33	0.30	0.20	-0.49
	Robust	5.20	0.33	0.37	-0.70
MBDE, (24)	Robust	0.04	-0.11	-0.30	0.30
M-quantile, (28)	Robust	0.06	0.59	-1.54	1.63
			%AvCR		
EBLUP, (23)	PR	95	95	95	95
	Robust	94	94	93	93
MBDE, (24)	Robust	98	99	97	98
M-quantile, (28)	Robust	92	90	92	92

Table 4 Design-based simulation results for the Albanian household population

Weighting Method	MSE Estimator	*%AvRB*	*%MedRB*	*%AvCR*	*%MedCR*
EBLUP, (23)	PR	14.4	14.7	95	98
	Robust	0.7	0.2	94	95
MBDE, (24)	Robust	-0.3	-0.6	93	94
M-quantile, (28)	Robust	-3.0	-2.9	94	94

Table 5 Design-based simulation results for the AAGIS farm population

Weighting Method	MSE Estimator	*%AvRB*	*%MedRB*	*%AvCR*	*%MedCR*
EBLUP, (23)	PR	45.6	23.7	90	98
	Robust	16.4	11.5	90	93
MBDE, (24)	Robust	19.3	-0.8	90	93
M-quantile, (28)	Robust	-1.8	-1.6	86	90

Figure 1 Area level values of true RMSE (solid line) and average estimated RMSE (dashed line) obtained in the model-based simulations with Gaussian errors and $\rho = 0.2$. Areas are ordered in terms of increasing population size. Values for the PR estimator are indicated by Δ while those for the Robust estimator are indicated by o. Plots show results for the EBLUP (top), MBDE (centre) and M-quantile (bottom) predictors.

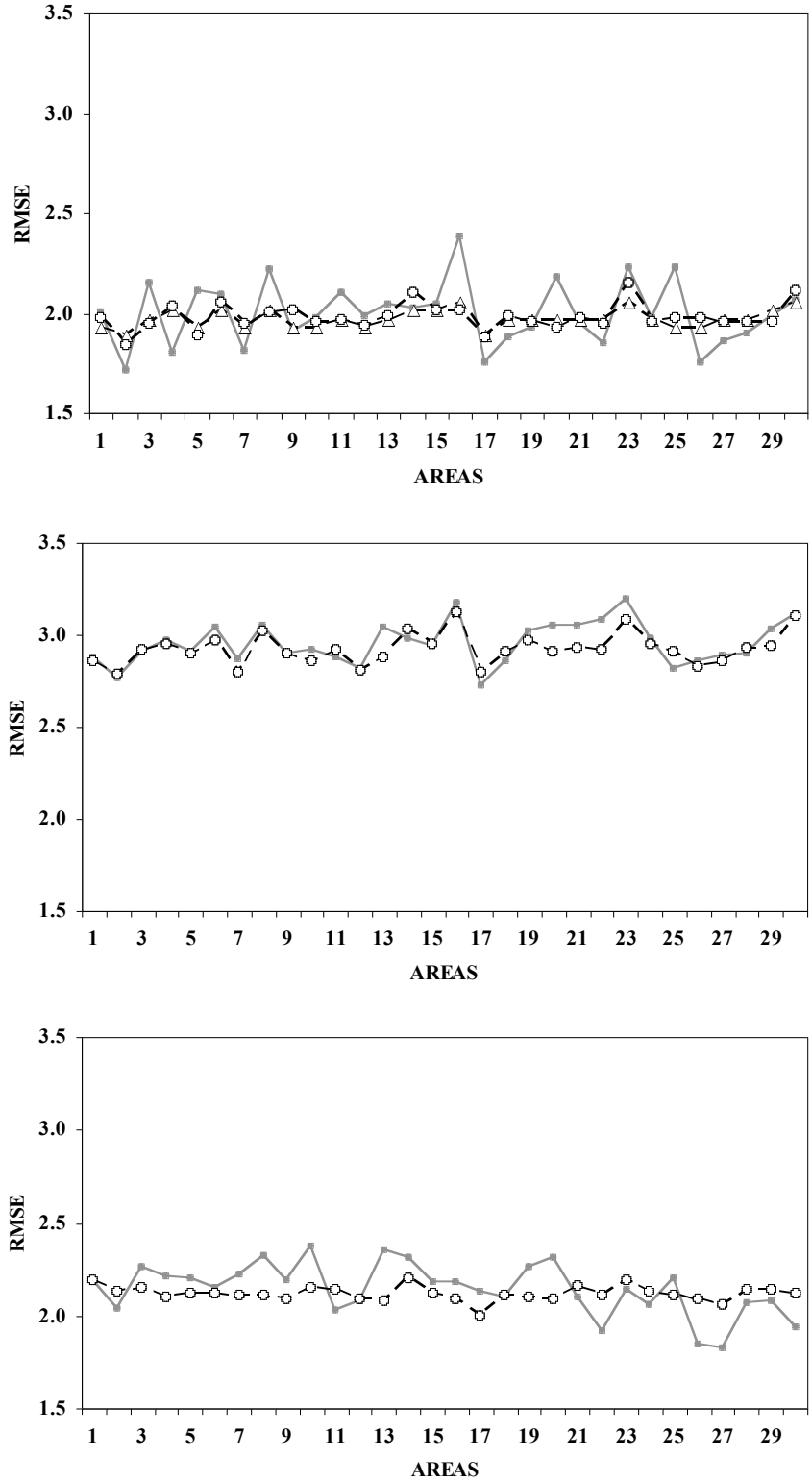

Figure 2 Area level values of true RMSE (solid line) and average estimated RMSE (dashed line) obtained in the model-based simulations with chi square errors and $(a,b)=(2,1)$. Areas are ordered in terms of increasing population size. Values for the PR estimator are indicated by Δ while those for the Robust estimator are indicated by o. Plots show results for the EBLUP (top), MBDE (centre) and M-quantile (bottom) predictors.

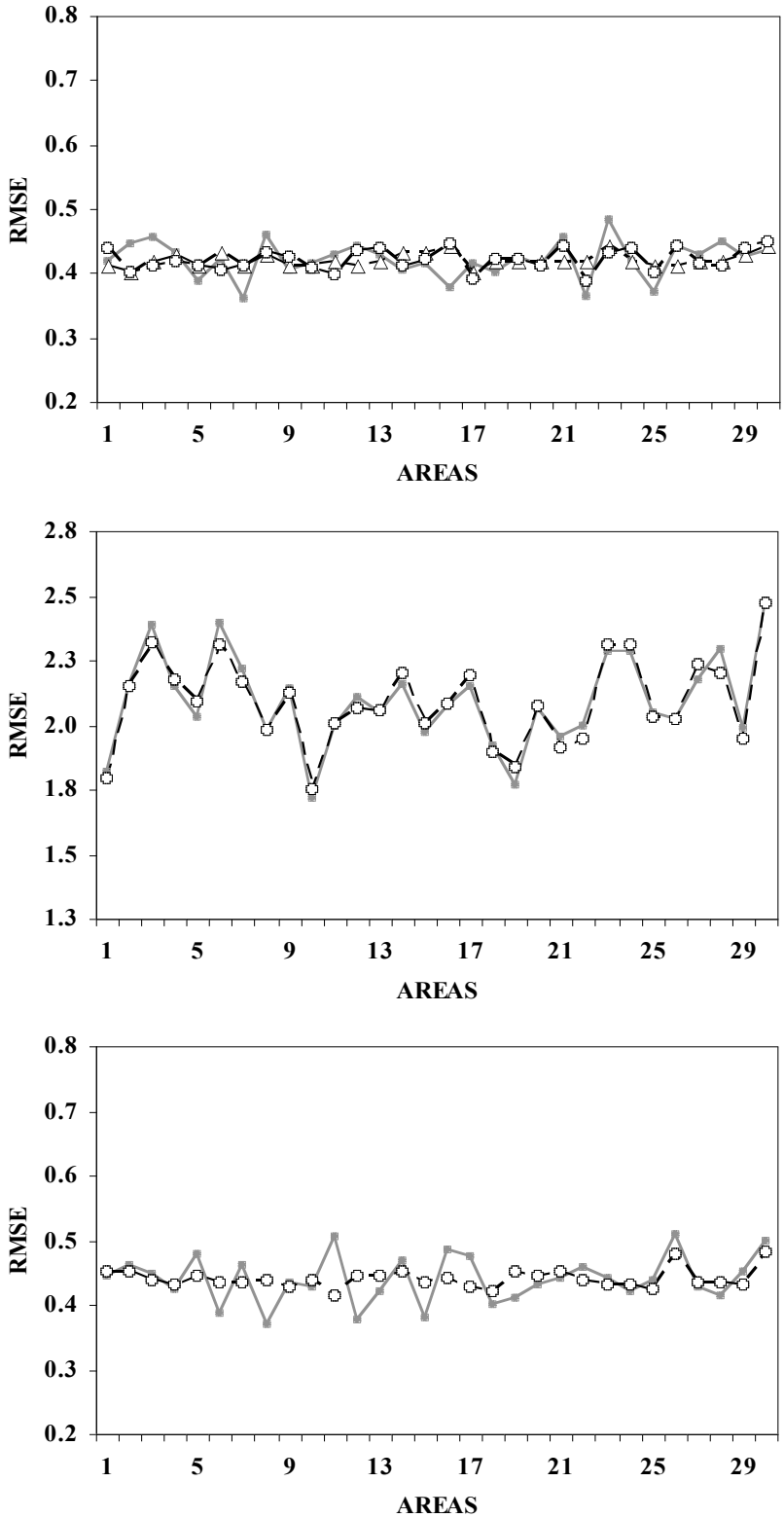

Figure 3 District level values of true design-based RMSE (solid line) and average estimated RMSE (dashed line) obtained in the design-based simulations using the Albanian household population. Districts are ordered in terms of increasing population size. Values for the PR estimator are indicated by Δ while those for the Robust estimator are indicated by o. Plots show results for the EBLUP (top), MBDE (centre) and M-quantile (bottom) predictors..

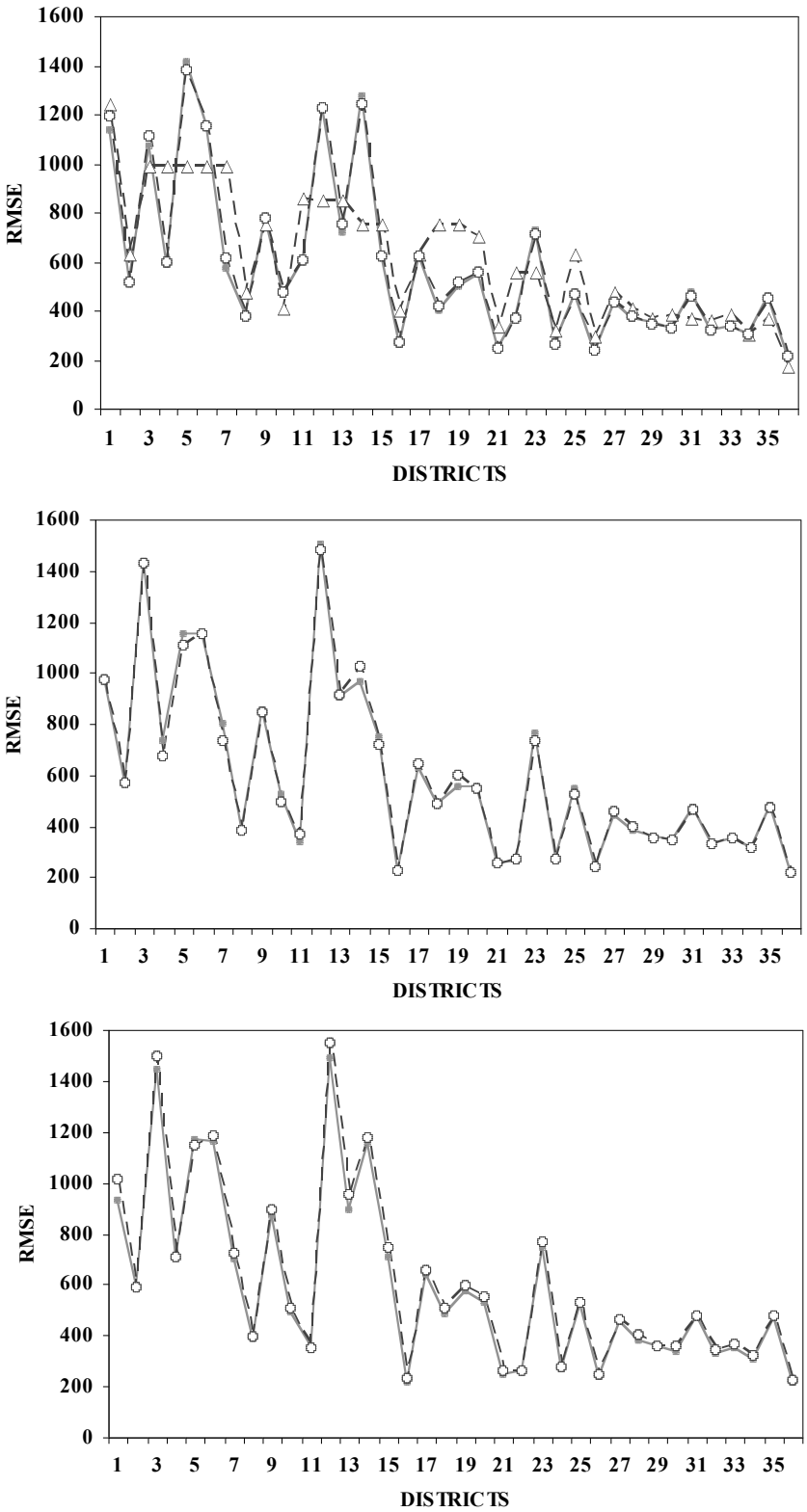

Figure 4 District level coverage rates of nominal 95% prediction intervals obtained in the design-based simulations using the Albanian household population. Districts are ordered in terms of increasing population size. MSE estimators are EBLUP/PR (solid line, Δ), EBLUP/Robust (long dashed line, o), MBDE/Robust (short dashed line, o), M-quantile/Robust (dotted line, o).

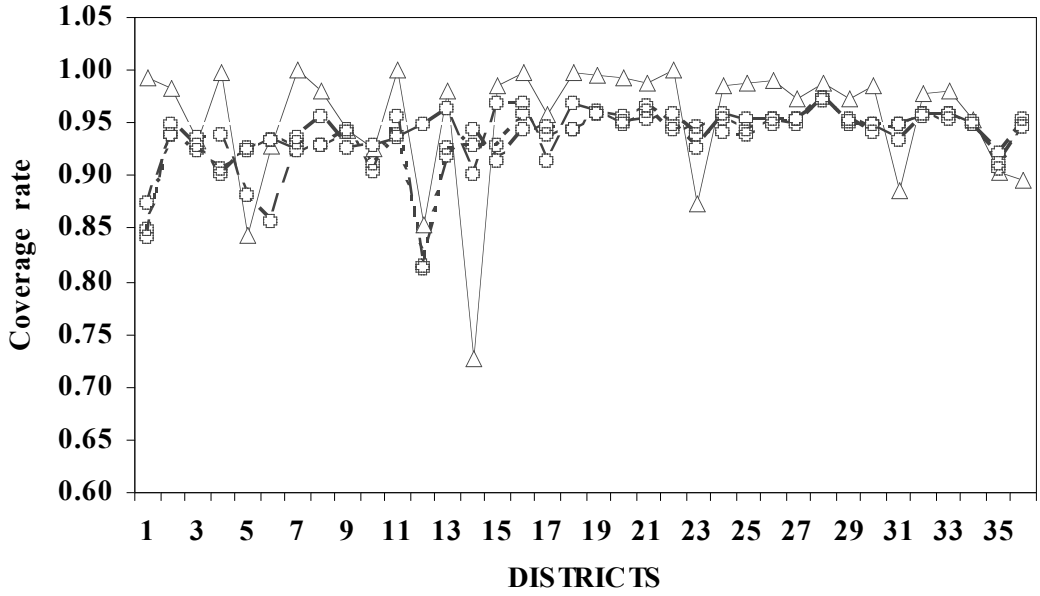

Figure 5 Regional values of true design-based RMSE (solid line) and average estimated RMSE (dashed line) obtained in the design-based simulations using the AAGIS farm population. Regions are ordered in terms of increasing population size. Values for the PR estimator are indicated by Δ while those for the Robust estimator are indicated by o. Plots show results for the EBLUP (top), MBDE (centre) and M-quantile (bottom) predictors.

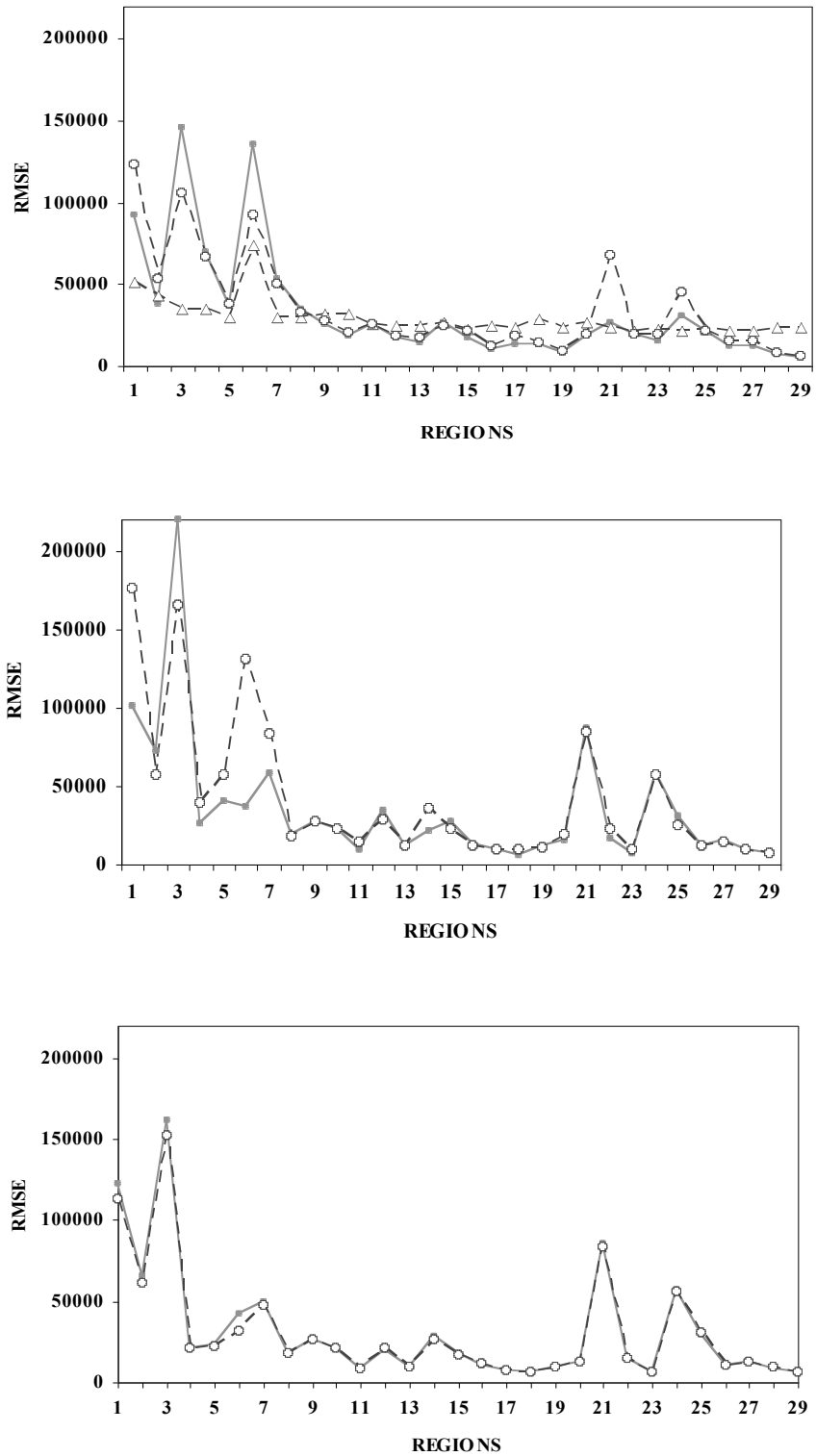

Figure 6 Regional coverage rates of nominal 95% prediction intervals obtained in the design-based simulations using the AAGIS farm population. Regions are ordered in terms of increasing population size. MSE estimators are EBLUP/PR (solid line, Δ), EBLUP/Robust (long dashed line, o), MBDE/Robust (short dashed line, o), M-quantile/Robust (dotted line, o).