RANK-BASED PREDICTIVE INFERENCE FOR THE FINITE POPULATION MEAN OF A SMALL AREA

M. Mushfiqur Rashid and Balgobin Nandram

M. M. Rashid, Department of Mathematical Sciences, WPI, 100 Institute Road, Worcester, MA 01609

Key words: *R*-estimates, Dispersion function, Gradient vector, Exchangeability

Abstract

Given data from a number of small areas, we use a rank-based method to estimate the finite population mean of one of these areas. As the areas are similar, we use a model which permits "borrowing of strength" from other areas. In particular, we use the nested error regression model; otherwise known as an error component model used in two-stage cluster sampling. One important feature of our method is that there is no need to assess assumptions of normality which is necessary to implement most methods for small areas. We obtain *R*-estimates of the model parameters by minimizing a dispersion function and we show that the *R*-estimators have desirable asymptotic properties. We predict the finite population mean of a small area based on the *R*-estimate and the model. Finally, we illustrate the methodology by a numerical example.

1 Introduction

Small area estimation is extremely useful to government agencies that address the issues of distribution, equity and disparity. It is particularly appropriate in a setting that involves several areas with only a small sample available from each area. One parameter of interest is the finite population mean of one area and a procedure which borrows strength from other neighboring areas is usually used. Further, it requires normality of the responses and transformation might be needed to obtain normality. The objective of this article is to develop a rank-based method to predict the mean of a small area.

In recent years model based estimation procedures for small area estimation are widely addressed. Ghosh and Rao (1994) gave an excellent review of the recent literature on small area estimation. The methods usually involve either an empirical Bayes approach or a full classical approach in which variance components are estimated. Both these procedures assume certain mixed models for prediction purpose. First, certain best linear unbiased predictors (BLUP) or EB predictors are obtained for the unknown parameters of interest assuming the variance components are known. Then the unknown variance components are estimated typically by Henderson's method of fitting constants or REML method, and the resulting estimated BLUP's are used for final prediction. Although, the above approach is usually quite satisfactory for point prediction, it is very difficult to estimate the standard errors. Ghosh and Lahiri (1989) proposed a hierarchical Bayes (HB) procedure as an alternative to the estimated BLUP or the EB procedure. The standard errors of the estimates, though complicated, can be obtained via numerical integration without any further approximation.

In this article we neither require any distributional assumptions nor estimation of variance components in predicting the mean of the small areas. However, estimateof variance components are required for the prediction errors. In Section 2 we present the model and assumptions. In Section 3 we obtain *R*-estimators together with their asymptotic properties. In Section 4 we describe residuals and model fitting. A point predictor of the finite population mean of a small area is given in Section 5. A numerical example is given in section 6, and Section 7 has conclusions.

2 Model and Assumptions

Assume that data are available from c small areas, and the i^{th} area has M_i individuals, $i = 1, 2, \ldots, c$. Let Y_{ij} be the value of the j^{th} unit in the i^{th} area. Let $\underline{Y}_i = [Y_{i1}, Y_{i2}, \ldots, Y_{iM_i}]'$ denote the vector of all values of all individuals from the i^{th} area.

We assume that from each of the c small areas a random sample of m_i individuals are drawn at random from the ith area ($\sum_i^c m_i = m_T$). Let $\underline{y}_i = (y_{i1}, \ldots, y_{im_i})'$ be a vector of m_i observations from the ith cluster on a response variable y, and let $(x_{i1k}, \ldots, x_{im_ik}, k = 1, \ldots, p-1)$ be the associated values of $p-1$ covariates which have influence on $\underline{y}_i$. Following Battese, Harter and Fuller (1988) we consider the nested error regression model

$$Y_{ij} = \alpha_0 + \underline{x}'_{ij}\underline{\beta} + v_i + u_{ij}, j = 1, \ldots, M_i, i = 1, \ldots, c \quad (2.1)$$

where $\underline{x}_{ij} = (x_{ij1}, \ldots, x_{ij,p-1})'$, $\underline{\beta} = (\beta_1, \ldots, \beta_{p-1})'$ is the vector of regression parameters, $v_i \overset{i.i.d.}{\sim} N(0, \sigma_v^2)$, $u_{ij} \overset{i.i.d.}{\sim} N(0, \sigma_u^2)$, and v_i's are independent of u_{ij}'s. In econometrics this model is known as an error component model and it is widely used for combining cross-section and time series data. For a sample of size m_i from the ith area the model (2.1) can be written as

$$Y_{ij} = \alpha_0 + \underline{x}'_{ij}\underline{\beta} + \epsilon_{ij}, j = 1, \ldots, m_i, i = 1, 2, \ldots, c. \quad (2.2)$$

For the ith area the model may be written in matrix notation as $\underline{y}_i = \alpha \underline{1}_{m_i} + X_i\underline{\beta} + \underline{\epsilon}_i$, where $\underline{y}_i = (Y_{i1}, Y_{i2}, \ldots, Y_{im_i})'$, $\underline{\epsilon}_i = [\epsilon_{i1}, \epsilon_{i2} \ldots, \epsilon_{im_i}]'$ with $\epsilon_{ij} = v_i + u_{ij}$, and X_i denoting the $m_i \times (p-1)$ matrix with rows $\underline{x}'_{ij}, j = 1, \ldots, m_i$. In parametric inference we assume that $\underline{\epsilon}_i \sim N(0, \sigma^2 V_i)$, where $\sigma^2 = \sigma_v^2 + \sigma_u^2$ and $V_i = [(1-\rho)I_{m_i} + \rho J_{m_i}]$, $\rho = \sigma_v^2/\sigma^2$ is the common intra-area correlation, I_{m_i} is the $m_i \times m_i$ identity matrix and J_{m_i} is the $m_i \times m_i$ unit matrix.

In some situations the assumption of multi-normality of $\underline{\epsilon}_i$ may not be valid. In such cases experimenters may prefer an alternative non-parametric procedure. In this article we assume that:

(1) The $\underline{\epsilon}_i$ $(i = 1, 2, \ldots, c)$ are independent and continuous random vectors, and the elements of $\underline{\epsilon}_i$ $(i = 1, 2, \ldots, c)$ are exchangeable random variables. Further, all the $\underline{\epsilon}_i$'s $(i = 1, 2, \ldots, c)$ have the same distributional form $F(.,.,\ldots,.)$;
(2) $\int_{-\infty}^{\infty} f(\epsilon, \epsilon)\, d\epsilon < \infty$, where $f(.,.)$ is the bivariate p.d.f. of any two components of $\underline{\epsilon}_i$;
(3) $x_{ijk}^2 < B$, where B is finite and x_{ijk}'s do not depend on c; and
(4) $\text{Sup}_{i=1,\ldots,c} m_i = m < \infty$.

Our objective is to use an even, location-free measure of dispersion that produces R-estimate of $(k = 1, 2, \ldots, p-1)$.

Let $\underline{\beta} = (\beta_1, \beta_2, \ldots, \beta_{p-1})'$ be the partial regression coefficients corresponding to the model (2.2). Let $n_{ij} = 1$ if $j \leq m_i$ and 0 otherwise. For the ith area we define a dispersion function as follows:

$$D_i(\underline{\beta}) = \sum_{j=1}^{m} n_{ij}\left[\frac{R(y_{ij} - \sum_{k=1}^{p-1}\beta_k x_{ijk})}{m_i + 1} - \frac{1}{2}\right]\left[y_{ij} - \sum_{k=1}^{p-1}\beta_k\, x_{ijk}\right],$$

where $R(y_{ij} - \sum_{k=1}^{p-1}\beta_k x_{ijk})$ is the intra-area rank of the residual, $y_{ij} - \sum_{k=1}^{p-1}\beta_k x_{ijk}$, corresponding to the j^{th} individual $(j = 1, 2, \ldots, m_i)$ in the ith area. The combined dispersion function for all small areas is

$$D(\underline{\beta}) = \sum_{i=1}^{c} D_i(\underline{\beta}).$$

Similar dispersion function like $D(\underline{\beta})$ was used for repeated measures incomplete block designs (see Rashid 1995) and one-way repeated measures with a changing covariate (Rashid and Nandram 1995).

This even, location free measure is a linear, rather than quadratic, function of the residuals with coefficients determined by the ranks, or sizes, of the residuals. Hence it is expected that the R-estimates generated by $D(\cdot)$ will be more robust than least squares estimators since the the influence of the outliers enters in a linear rather than a quadratic fashion. Further $D(\underline{\beta})$ is a nonnegetive, continuous and convex function of $\underline{\beta}$.

3 R-Estimators

Let $\hat{\underline{\beta}}$ minimize the dispersion function $D(\underline{\beta})$. Then $\hat{\underline{\beta}}$ is called the R-estimator of $\underline{\beta}$. Let $\Gamma = (\underline{\beta} : D(\cdot) = \text{minimum})$. Then Γ is bounded. Note that $\hat{\underline{\beta}}$ may not be unique. So in practice there may be arbitrariness in its definition. However, the diameter of the set of all possible values of $\sqrt{c}\hat{\underline{\beta}}$ tends to zero as $c \to \infty$. See Jaeckel (1972) for further discussions.

In order to develop the asymptotic distribution of $\hat{\underline{\beta}}$ we need the gradient of the dispersion function as well as its linear approximation.

The domain $\underline{\beta}$ space of $D(\cdot)$ consists of finite number of convex polygonal subsets, on each of which $D(\cdot)$ is a linear function of $\underline{\beta}$. Therefore the partial derivatives exist almost everywhere and are given by

$$\frac{\partial D(\underline{\beta})}{\partial \beta_k} = -\sum_{i=1}^{c}\sum_{j=1}^{m} n_{ij}\left[\frac{R(y_{ij} - \sum_{k=1}^{p-1}\beta_k x_{ijk})}{m_i + 1} - \frac{1}{2}\right]x_{ijk},$$

$k = 1, 2, \ldots, p-1$. Let $\underline{S}(\underline{\beta}) = [S_1(\underline{\beta}), S_2(\beta), \ldots, S_{p-1}(\beta)]'$, where

$$S_k(\underline{\beta}) = \sum_{i=1}^{c}\sum_{j=1}^{m} n_{ij}\left[\frac{R(y_{ij} - \sum_{k=1}^{p-1}\beta_l x_{ijk})}{m_i + 1} - \frac{1}{2}\right]x_{ijk}.$$

Asymptotic distributions of the R-estimators will be developed under the error component model. Therefore our results can be considered as an extension of the linear model with i.i.d. errors to nested error component model. In this article our assumptions are weaker than the i.i.d. error model in parametric inference. Even though the ranking of the residuals are done within each area separately, there is a borrowing strength from other areas since the R-estimates are obtained by minimizing the

combined dispersion function. For balanced designs, estimates and tests, achieve Friedman efficiency and as a result perform better for heavy-tailed distributions.

We will use the R-estimators to construct a predictor of the finite population mean of a small area. We are not able to estimate α because inclusion or exclusion of α in the dispersion function does not alter the ranks of the residuals. Therefore, we will only obtain the R-estimates of $(\beta_1, \beta_2, \beta_2, \ldots, \beta_{p-1})'$ by minimizing the dispersion function. However, as mentioned earlier, the estimate of α will be obtained from the residuals. Let $\underline{\beta}^0$ be the true value of $\underline{\beta}$. Without loss of generality, throughout the paper we assume that $\underline{\beta}^0 = \underline{0}$.

Under assumption (1), if $n_{ij} \neq 0$,

$$E_0[R_{ij}] = (m_i + 1)/2, \quad \text{Var}_0[R_{ij}] = (m_i^2 - 1)/12;$$

and if $i = i'$,

$$\text{Cov}_0[R_{ij}, R_{i'j'}] = -(m_i + 1)/12 \text{ and } \text{Cov}_0[R_{ij}, R_{i'j}] = 0.$$

Thus,

$$E_0[S_k(\underline{0})] = 0, \ k = 1, 2, \ldots, p-1,$$
$$\text{Cov}_0[\frac{1}{\sqrt{c}} S_k(0), \frac{1}{\sqrt{c}} S_{k'}(0)] = \sigma_{kk'},$$

where

$$\sigma_{kk'} = -\sum_{i=1}^{c} \sum_{j \neq j'=1}^{m_i} \frac{n_{ij} n_{ij'} x_{ijk} x_{ij'k'}}{c(m_i + 1)}$$

$(k \neq k' = 1, \ldots, p-1)$.

Also

$$\text{Var}_0[\frac{1}{\sqrt{c}} S_k(0)] = \sigma_{kk}$$

where

$$\sum_{i=1}^{c} \sum_{j=1}^{m} \frac{m_i n_{ij} (x_{ijk} - \bar{x}_{i.k})^2}{c(m_i + 1)} = \sigma_{kk}, \ k = 1, 2, \ldots, p-1.$$

Let Σ be the covariance matrix of $\frac{1}{\sqrt{c}} \underline{S}(\underline{0})$ under the true value $\underline{\beta}^0 = \underline{0}$. Then $\Sigma = (\sigma_{kk'})$. Note that the rank of Σ would be $p-1$. Assuming that the $\lim_{c \to \infty} \Sigma$, let

$$A = \lim_{c \to \infty} \Sigma, \tag{3.3}$$

and assume that this limit exists. To establish the properties of the R-estimators we first present Lemma 1. Let the true value $\underline{\beta}_0 = \underline{0}$.

The following lemma gives the asymptotic distribution of $\underline{S}(\underline{0})$ under the true value.

Lemma 1: Under assumption (1),

$$\sqrt{c}[\frac{1}{c} \underline{S}(\underline{0})] \xrightarrow{\mathcal{D}} MVN(\underline{0}, A) \text{ as } c \to \infty,$$

where A is defined in (3.3).

Lemma 1 follows from a multivariate central limit theorem (See Rao 1973, p. 147).

Next we present our main results about the properties of the R-estimators.

Theorem 1: Suppose assumptions (1) - (4) hold for the model (2). Let $\underline{\beta}^0$ be the true value of $\underline{\beta}$. Then,

$$\sqrt{c}(\underline{\beta} - \underline{\beta}^0)/\tau \xrightarrow{\mathcal{D}} \text{MVN}[\underline{0}, A] \text{ as } c \to \infty.$$

Proof: First we approximate the gradient vector of the dispersion function by a linear function. Let P_0 be the probability computed under the true value, $\underline{\Delta} \in R^{p-1}$ and the true value $\underline{\Delta}^0 = \underline{0}$. Using a Taylor's expansion in the neighborhood of $\underline{0}$, it can be shown that

$$E_0 \sqrt{c} \left[\frac{1}{c} S\left(\frac{1}{\sqrt{c}} \underline{\beta}\right) - \frac{1}{c} S(\underline{0})\right] \cong -\frac{1}{\tau} A \underline{\beta}$$

where E_0 is the expectation under the true value $\underline{0}$, $\tau = 1/[\sqrt{12} \int_{-\infty}^{\infty} f(\epsilon, \epsilon)\, d\epsilon]$.

Let $\theta = \frac{1}{\sqrt{c}} \underline{x}_{ij} \underline{\beta}$. Now

$$\text{Var}_0[\sqrt{c}\{\frac{1}{c}[s_k(\frac{1}{\sqrt{c}} \underline{\Delta}) - s_k(\underline{0})]\}$$

$$\leq 12mB \sum_{j=1}^{m} n_{ij} E_0[R\left(Y_{ij} - \theta_{ij}\right) - R(Y_{ij})]^2 \to 0 \text{ as } c \to \infty.$$

By appealing to Markov's inequality, we have

$$\lim_{c \to \infty} P_0 \left[\left\| \sqrt{c} \left\{ \frac{1}{c} \underline{S}\left(\frac{1}{\sqrt{c}} \underline{\Delta}\right) - \frac{1}{c} \underline{S}(\underline{0}) \right\} + \frac{1}{\tau} A \right\| \geq \varepsilon \right] = 0.$$

Since $\underline{S}(\cdot)$ is monotone, by using similar arguments given in Rashid and Nandram (1995) we have:

$$\lim_{c \to \infty} P_0[\text{Sup} \| \sqrt{c}\{\underline{\hat{\beta}} - \frac{1}{c\tau} A \underline{S}(\underline{0})\}\| \geq \epsilon, \underline{\hat{\beta}} \in \Gamma] = 0 \tag{3.4}$$

Hence, using Lemma 1 and (3.4), we prove Theorem 1.

Note that, Theorem 1 shows that our R-estimators are asymptotically unbiased and consistent. Further their asymptotic distributions are distribution free even if we replace τ by a consistent estimate $\hat{\tau}$. Let $\hat{\tau}$ be a consistent estimator of τ. Then by Slutsky's Theorem

$$\sqrt{c}[\underline{\hat{\beta}} - \underline{\beta}^0]/\hat{\tau} \xrightarrow{\mathcal{D}} MVN[\underline{0}, A] \text{ as } c \to \infty. \tag{3.4}$$

Moreover, the above asymptotic distribution is independent of the parent distribution (i.e. distribution free).

4 Model Checking

The rank-based fitted value for model (2.1) of y_{ij} is $\hat{y}_{ij} = \sum_{k=1}^{p-1} \hat{\beta}_k x_{ijk} + \hat{\nu}_i$. where

$$\hat{\nu}_i = \text{med}(y_{ij} - \sum_{k=1}^{p-1} \hat{\beta}_k x_{ijk} : j = 1, 2, \ldots, m_i).$$

Then, we obtain residuals of the model (2.1) as

$$\hat{u}_{ij} = y_{ij} - \hat{y}_{ij}.$$

This practice of computing the residuals $\hat{u}_{ij}$ is recommended in Minitab reference manual (1991, p. 10-12) for inferences concerning block designs based intra-subject ranks. In order to obtain studentized residuals of the model (2.1), we need an estimate of σ_u (The standard deviation of u_i in model (2.2)). A robust estimate is given by the mean absolute deviation (MAD) of $\hat{u}_{ij}$:

$$\hat{\sigma}_u = \sum_{i=1}^{c} \sum_{j=1}^{m_i} |\hat{u}_{ij} - \text{Med}_{ij}(\hat{u}_{ij})|\}/m_T,$$

where $m_T = \sum_{i=1}^{c} m_i$. We can plot the $\hat{u}_{ij}/\hat{\sigma}_u$ versus $\hat{y}_{ij}$ to check whether the fitted model (2.1) is adequate.

Although we are not able to estimate α using the dispersion function, one can take

$$\hat{\alpha} = \text{med}(\hat{\nu}_i, i = 1, 2, \ldots, c).$$

Let $\hat{y}(ij)$ denote the predicted value of y_{ij} when $(ij)^{th}$ is deleted from the model. Then the change in the robust fit due to the i^{th} case is

$$\text{RDFFIT}_{ij} = y_{ij} - \hat{y}(ij),$$

where RDFFIT_{ij} is our measure of the influence of case ij. Analogous to least-squares analysis, studentized RDFFIT_{ij} will detect an influential observations. Further, one can compute RDCOOK_{ij}. Notice that we have put R in front of the above measures to denote rank analog of least-squares diagnostics. See McKean et al (1990) for further discussions.

5 Point Predictor

In this section we show how to predict the finite population mean of each small area and how to estimate the prediction error. Let $\bar{Y}_i = \sum_{j=1}^{M_i} Y_{ij}$ be the finite population mean of the i^{th} small area $i = 1, 2, \ldots, c$. Letting $f_i = m_i/M_i$ be the finite population correction for the i^{th} small area, observe that

$$\bar{Y}_i = f_i \bar{y}_i^{(s)} + (1 - f_i)\bar{Y}_i^{(ns)} \tag{5.1}$$

where $\bar{y}_i^{(s)} = \sum_{j=1}^{m_i} Y_{ij}/m_i$ and $\bar{Y}_i^{(ns)} = \sum_{j=m_i+1}^{M_i} Y_{ij}/(M_i - m_i)$. Note that $\bar{y}_i^{(s)}$ is computed from the observed data and $\bar{Y}_i^{(ns)}$ is to be predicted.

First, we obtain the point predictor of $\bar{Y}_i$, the finite population mean of the i^{th} area. Let $\bar{x}_{i\cdot k}^{(s)} = \sum_{j=1}^{m_i} X_{ijk}/m_i$ and $\bar{X}_{i\cdot k}^{(ns)} = \sum_{j=m_i+1}^{M_i} X_{ijk}/(M_i - m_i)$ $i = 1, 2, \ldots, c$ and $k = 1, 2, \ldots, p-1$. Also let $\underline{x}_i^{(s)} = (\bar{x}_{i\cdot 1}^{(s)}), \ldots, \bar{x}_{i\cdot p-1}^{(s)})'$ and $\underline{X}_i^{(ns)} = (\bar{X}_{i\cdot 1}^{(ns)}, \ldots, (\bar{X}_{i\cdot p-1}^{(ns)})$. Assuming the model specification in (2.2), with appropriate expectations,

$$E(\bar{Y}_i^{(ns)}) = \alpha + \underline{\beta}' \underline{\bar{X}}_i^{(ns)}. \tag{5.2}$$

To obtain a predictor for $\bar{Y}_i^{(ns)}$ we substitute our rank-based estimators into (5.2). It follows that the predictor for the finite population mean of the i^{th} area is

$$\hat{\bar{y}}_i^{(R)} = \bar{y}_i^{(s)} + (1 - f_i)\hat{\underline{\beta}}'(\underline{\bar{X}}_i^{(ns)} - \underline{\bar{x}}_i^{(s)}). \tag{5.3}$$

The predictor in (5.3) is a rank-based regression survey predictor. It is expected to be robust. One version of this estimator was considered by Sarndal (1984). Notice that the generalized least squares estimator and corresponding predictor under the model (2.1) contain the σ_e^2 and σ_ν^2. However these variance components are unknown and have to be estimated, and in our case we do not need to estimate them. The standard error of rank-based survey regression predictor can be obtained as follows. The prediction error is $\hat{\bar{y}}_i^{(R)}$ - $\bar{Y}_i$. In order to compute the standard error of $\hat{\bar{y}}_i^{(R)}$, we need to assume that the covariance matrix of $\underline{\epsilon}_j$ exists. As our model assumes exchangeability,

$$\text{cov}(\underline{\epsilon}_i) = \sigma^2[(1 - \rho)I_{m_i \times m_i} + \rho J_{m_i \times m_i}].$$

We also need the covariances between the non-sampled observations and $\hat{\underline{\beta}}$, which can be obtained by using the one-step R-estimator defined in (3.4). However, these covariances depend on the scale parameter τ. As a the mean squared error of the predictor will contain σ^2, ρ and τ robust estimates of σ^2 and ρ can be obtained as follows:

$$\hat{\sigma}_\nu = 1.483\text{Med}_i\{|\hat{\nu}_i - \text{Med}_i\hat{\nu}_i|\},$$

$$\text{and } \hat{\sigma}^2 = \hat{\sigma}_\nu^2 + \hat{\sigma}_u^2$$

The $\hat{\sigma}_u^2$ is defined in section 4. An estimate of ρ is $\hat{\rho} = \hat{\sigma}_v^2/\hat{\sigma}_u^2$. It can be shown that parameter $\sqrt{12}/\tau$ is the density of $\epsilon_{ij} - \epsilon_{ij'} (j \neq j')$ at 0 (median) under the model (2.1). Therefore, an estimate of τ can be obtained from the residuals using Bloch and Gastwirth (1968) density estimate. See also Rashid and Nandram (1995) for further details. The standard error of the predictor of this article is nonparametric analogue of Prasad and Rao (1990).

6 Example

We consider an example described by Battese, Harter and Fuller (1988) which concerns the estimation of areas under corn and soybeans for each 12 counties in North-Central Iowa using farm-interview data in conjunction with LANDSAT satellite data. Each county was divided into area segments, and the areas under corn and soybeans were ascertained for a sample of segments by interviewing farm operators. The number of segments in the sampled counties ranged from 1 to 6. The total number of segements in the different counties ranged from 394 and 965. Thus, this is, indeed an small area problem. Auxiliary data in the form of number of pixels (a term used for "picture elements" of about 0.45 hectares) classified as corn and soybeans were also obtained for all the area segments, including the sample segments, in each county using the LANDSAT satellite readings. p Battese, Harter and Fuller (1988) employed a "nested error regression" model involving random small area effects and the segment-level data and then obtain the estimates of county areas under corn and soybeans using the classical components of variance approach. They obtained the following fitted regression equations:

$$\text{Corn} : y_{ij} = 51 + 0.329 x_{1j1} - 0.134\ x_{ij2},$$

$$\text{Soybeans} : y_{ij} = -16 + 0.028 x_{1j1} + 0.494\ x_{ij2}.$$

Datta and Ghosh (1991) applied the HB approach to these data and show that the two approaches give similar results.

We apply the method developed in this article to the same data. We obtain the R-estimates of the partial regression coefficients by minimizing the dispersion function. We have used the Nelder-Mead algorithom to minimize the dispersion function. The rank-based estimated regression equations are:

$$\text{Corn} : y_{ij} = 47.64 + 0.3154 x_{1j1} - 0.1019\ x_{ij2},$$

$$\text{Soybeans} : y_{ij} = -41.13 + 0.1029 x_{1j1} + 0.5036\ x_{ij2}.$$

Using the results of Section 4, we have plotted studentized residuals of the model (2.1) with respect to the corresponding fitted values. (The plots are omitted.) The points in the plots do not show any patterns. It appears that for the corn data the second farm in Hardin county is an outlier. Also for the soybean data the second farm in Pocahontas county is an outlier. (The studentised residuals in the first case is -6.72, and for the second case is -5.32). Otherwise, the fitted models are consistent with the data.

We have fitted the nested error regression model 2 without the observation 2 in Hardin county. The fitted models are:

$$\text{Corn} : y_{ij} = 65.89 + 0.2894 x_{1j1} - 0.1519\ x_{ij2},$$

$$\text{Soybeans} : y_{ij} = -36.82 + 0.0973 x_{1j1} + 0.4904\ x_{ij2}.$$

In Table 1 we present the rank-based survey regression predictors (with and without the second observation in the Hardin county)and the best predictor for both corn and soybeans. Notice that for both corn and soybeans these predictors are very similar to trsectionhose obtained by Battese, Harter and Fuller (1988). However, without second observation in the Hardin county the survey regression predictors are close to the best predictors which is expected.

7 Concluding Remarks

We obtain a rank-based method for estimating the finite population mean of a small area. Our R-estimates are easy to obtain using the Nelder-Mead simplex algorithm. We showed that the rank-based estimates of the regression parameters have reasonable properties. Our method is very simple and easy to compute. Although we do not assume normality, the estimate of the finite population mean of each area is close to the BLUP of Battese, Harter and Fuller (1988). We expect that this work would stimulate research in nonparametric methods for small area estimation.

REFERENCES

Battese, G. E., Harter, R. M. and Fuller, W.A. (1988), An Error Components Model for Prediction of County Crop Areas Using Survey and satellite Data. *Jour. Amer. Stat. Assoc., 83,* 28-36.

Bloch, D. A. and Gastwirth, J. S. (1968). On a Simple Estimate of the Reciprocal of the Density Function. *Annals of Mathematical Statistics, 39, No. 3,* 1083-1085

Ghosh, M. and Rao, J. N. K.(1994): Small Area Estimation: An Appraisal. *Statistical Sciences, 9,* 55 - 93.

Ghosh, M. and Lahiri, P. (1987): Robust empirical Bayes estimation of means of from stratified samples. *Jour. Amer. Stat. Assoc., 82*, 1153 - 1162.

Jaeckel, L.(1972): Estimating regression coefficients by minimising the dispersion of the residuals *The Annals of Mathematical Statistics, 43*, 1449 - 1458.

Prasad, N. G. N. and Rao, J. N. K(1990): On the Estimation of Mean squared Error of Small Predictors *Jour. Amer. Stat. Assoc., 85*. 163 -171.

Minitab Reference Manual(1991), Release 8, 10 - 12.

Rashid, M. M. (1995): Analysis of Repeated Measures Incomplete Block Designs Using *R*-Estimators. *Statistics and Probability Letters*, 22(5).

Rashid, M. M. and Nandram, B. (1995): A Rank Based Analysis of One-Way Repeated Measures Designs With a Changing Covariate *Journal of Statistical Planning and Inference*,Vol 45.

Rao, C. R. (1973): Linear Statistical Inference and Its Applications *John Wiley and Sons*, 147.

Sarndal, C. E. (1984): Design-Consistent Versus Model-Department Estimation for Small Domains *Journal of the American Statistical Association 79*, 624 - 631.

Table 1. A Comparison of the Rank-Based Survey Regression Predictor and Normal theory Predictor

	Segments		Corn		Soybean	
County	Sample	Popl	Best	R	Best	R
Cerro Gordo	1	545	122.2	127.2	77.8	67.8
				122.5		66.8
Hamilton	1	566	126.3	127.3	94.8	104.7
				126.0		104.9
Worth	1	394	106.2	92.2	86.9	84.8
				93.5		85.2
Humboldt	2	424	108.0	108.1	79.7	65.9
				106.8		65.1
Franklin	3	564	145.0	150.1	65.2	59.7
				149.7		59.6
Pocahontas	3	570	112.6	114.7	113.8	116.3
				114.4		116.4
Winnebago	3	402	112.4	110.4	98.5	101.6
				109.1		101.4
Wright	3	567	122.1	124.5	112.8	111.4
				123.9		111.1
Webster	4	687	115.8	118.1	109.6	108.9
				118.5		108.9
Hancock	5	569	124.3	122.3	101.0	104.1
				123.1		104.4
Kossuth	5	965	106.3	104.4	119.9	121.8
				104.2		121.7
Hardin	6	556	143.6	131.0	74.9	76.9
				144.6		79.4

Note: For the rank-based estimator the top entry uses all the data and in the bottom entry the outlier is removed.

Popl: Population

R:Rank