Small-Area Estimation by Combining Time-Series and Cross-Sectional Data

Author(s): J. N. K. Rao and Mingyu Yu

Source: *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, Dec., 1994, Vol. 22, No. 4, Special Issue on the Analysis of Health and Environmental Data / Numéro spécial sur l'analyse des données sur l'hygiéne et l'environnement (Dec., 1994), pp. 511-528

Published by: Statistical Society of Canada

Stable URL: https://www.jstor.org/stable/3315407

The Canadian Journal of Statistics
Vol. 22, No. 4, 1994, Pages 511–528
La Revue Canadienne de Statistique

Small-area estimation by combining time-series and cross-sectional data*

J.N.K. RAO and Mingyu YU

Carleton University and Statistics Canada

Key words and phrases: Autocorrelated errors; best linear unbiased prediction; Fay-Herriot model; mean-squared error estimation; sampling errors; two-stage estimator.
AMS 1985 subject classifications: Primary 62D05; secondary 62M10.

ABSTRACT

A model involving autocorrelated random effects and sampling errors is proposed for small-area estimation, using both time-series and cross-sectional data. The sampling errors are assumed to have a known block-diagonal covariance matrix. This model is an extension of a well-known model, due to Fay and Herriot (1979), for cross-sectional data. A two-stage estimator of a small-area mean for the current period is obtained under the proposed model with known autocorrelation, by first deriving the best linear unbiased prediction estimator assuming known variance components, and then replacing them with their consistent estimators. Extending the approach of Prasad and Rao (1986, 1990) for the Fay-Herriot model, an estimator of mean squared error (MSE) of the two-stage estimator, correct to a second-order approximation for a small or moderate number of time points, T, and a large number of small areas, m, is obtained. The case of unknown autocorrelation is also considered. Limited simulation results on the efficiency of two-stage estimators and the accuracy of the proposed estimator of MSE are presented.

RÉSUMÉ

Un modèle impliquant des effets aléatoires autocorrélés et des erreurs d'échantillonnages est proposé pour l'estimation des petites surfaces, utilisant à la fois des séries chronologiques et des données transversales. Les erreurs d'échantillonnages sont présumées avoir une matrice connue de variance-covariance bloc diagonale. Ce modèle est une extension d'un modèle bien connu dû à Fay et Herriot (1979) pour données transversales. Un estimateur à deux niveaux pour la moyenne d'une petite surface pour la période en cours est obtenu sous les hypothèses du modèle proposé avec autocorrélation connue, en dérivant d'abord l'estimateur de la meilleure prédiction linéaire non biaisée (MPLNB), en assumant connues les variances et en les remplaçant par leurs estimateurs consistants. Généralisant l'approche de Prasad et Rao (1986, 1990) pour le modèle de Fay-Herriot, on a obtenu un estimateur de l'erreur quadratique moyenne (EQM) de l'estimateur à deux niveaux, qui est une bonne approximation d'ordre deux lorsque le nombre de points dans le temps, T, est petit ou modérément grand, et que le nombre de petites surfaces, m, est relativement grand. Le cas où l'autocorrélation est inconnue, est aussi considéré. Des résultats limités basés sur des études de simulations et portant sur l'efficacité des estimateurs à deux niveaux et la précision de l'EQM, sont présentés.

1. INTRODUCTION

Most large-scale sample surveys are designed to provide reliable estimates for large geographical regions and large subgroups of a population. For example, the U.S. National

*This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

Health Interview Survey provides reliable estimates of health characteristics, such as mortality, disability, and utilization of health surveys, for the nation and four broad geographical regions. However, estimates are also needed for small areas, such as counties and Health Service Areas (National Center for Health Statistics 1979). Such estimates are increasingly being used in formulating policies and programs, in allocation of government funds, and in regional programs, etc.

Direct survey estimates for a small area, based on the data only from the sample units in the area, are likely to yield unacceptably large standard errors due to unduly small size of the sample in the area. Alternative estimates that borrow strength from related small areas are therefore needed to improve efficiency. Such estimates are based on either implicit or explicit models which provide a link to related small areas through supplementary data such as administrative records and recent census counts.

Most of the research on small-area estimation has focused on cross-sectional data at a given point in time. Rao (1986) and Ghosh and Rao (1994), among others, have given an account of this research. Estimators proposed in the literature include (1) synthetic estimators (Gonzalez 1973, Ericksen 1974), and structure-preserving estimators (Purcell and Kish 1980); (2) sample-size-dependent estimators (Drew *et al.* 1982, Särndal and Hidiroglou 1989); (3) empirical Bayes estimators (Fay and Herriot 1979, Ghosh and Lahiri 1987), and two-stage estimators or empirical BLUP (best linear unbiased prediction) estimators (Prasad and Rao 1986, 1990; Battese *et al.* 1988); (4) hierarchical Bayes estimators (Datta and Ghosh 1991).

Scott and Smith (1974), Jones (1980), Binder and Dick (1989), and others used time-series methods to develop efficient estimators of aggregates (e.g., overall means) from repeated surveys, by combining the direct survey estimates over time. Tiller (1992) used the Kalman filter to combine a current-period state-wide estimate from the U.S. Current Population Survey with past estimates for the same state and auxiliary data from the Unemployment Insurance System and the Current Employment Statistics Payroll Survey. However, neither Scott and Smith (1974) nor Tiller (1992) considered small-area estimation by combining time-series and cross-sectional data.

The main purpose of this article is to propose a combined cross-sectional and time-series model involving autocorrelated random effects and sampling errors with an arbitrary covariance matrix over time. Using this model, empirical BLUP estimators (or two-stage estimators) of small-area means and associated standard errors for the current period are obtained. Pfeffermann and Burck (1990) and Singh, Mantel, and Thomas (1991), among others, have also considered cross-sectional and time-series models for small-area estimation using the Kalman filter, but they assumed specific models for the sampling errors over time (e.g., independent sampling errors). As a result, their approach cannot be used with an arbitrary covariance matrix for sampling errors.

The remainder of this article is organized as follows. Section 2 reviews some work on regression synthetic estimators and empirical Bayes or two-stage estimators obtained solely from cross-sectional data for the current period. Some combined cross-sectional and time-series models are considered in Section 3, and an extension of the Fay-Herriot model (Fay and Herriot 1979) is proposed. Two-stage estimators of small-area means are given in Section 4, assuming known autocorrelation. Extending the approach of Prasad and Rao (1986, 1990) for the Fay-Herriot model, an estimator of the MSE of the two-stage estimator is obtained in Section 5, correct to a second-order approximation for a small or moderate number of time points, T and a relatively large number of small areas, m. The case of unknown autocorrelation is considered in Section 6. The results of a limited simulation study on the efficiency of two-stage estimators and relative bias of estimators of MSE are reported in Section 7. Finally, some concluding remarks are given in Section 8.

2. CROSS-SECTIONAL ESTIMATORS

We first obtain regression synthetic estimators, assuming a deterministic model on the small-area means. We then allow uncertainty into the model and obtain empirical Bayes or two-stage estimators of small-area means.

2.1. Regression Synthetic Estimators.

Let y_{iT} be the direct survey estimator of the ith small-area mean for the current period T, say θ_{iT} ($i = 1, \ldots, m$). We assume that y_{iT} is design-unbiased for θ_{iT}, i.e.,

$$y_{iT} = \theta_{iT} + e_{iT}, \tag{2.1}$$

where the e_{iT}'s are sampling errors with $\mathcal{E}(e_{iT}|\theta_{iT}) = 0$. We assume that a vector of fixed concomitant variables $\mathbf{x}_{iT} = (x_{iT1}, \ldots, x_{iTp})^{\mathsf{T}}$ related to θ_{iT} is available for all i at time T—for example, time-varying supplementary data such as administrative records and other auxiliary data such as census data.

A simple regression synthetic estimator of the current small-area mean θ_{iT}, based solely on the cross-sectional data $\{(y_{iT}, \mathbf{x}_{iT}), i = 1, \ldots, m\}$ for time T, is obtained by assuming the following deterministic model on θ_{iT}:

$$\theta_{iT} = \mathbf{x}_{iT}^{\mathsf{T}} \boldsymbol{\beta}_T, \tag{2.2}$$

where $\boldsymbol{\beta}_T = (\beta_{T1}, \ldots, \beta_{Tp})^{\mathsf{T}}$ is the vector of regression coefficients. It is given by

$$\tilde{\theta}_{iT}(\text{reg}) = \mathbf{x}_{iT}^{\mathsf{T}} \tilde{\boldsymbol{\beta}}_T, \tag{2.3}$$

where $\tilde{\boldsymbol{\beta}}_T$ is the ordinary least-squares estimator of $\boldsymbol{\beta}_T$ obtained from the combined model based on (2.1) and (2.3):

$$y_{iT} = \mathbf{x}_{iT}^{\mathsf{T}} \boldsymbol{\beta}_T + e_{iT}, \qquad i = 1, \ldots, m,$$

for time T.

Synthetic estimators like (2.3) could lead to substantial design biases, since they give a zero weight to the direct estimator y_{iT}. On the other hand, empirical Bayes or two-stage estimators give "proper" weights to the survey estimator and the synthetic estimator. As a result they can lead to smaller biases relative to synthetic estimators.

2.2. Empirical Bayes or Two-Stage Estimators.

Following Fay and Herriot (1979), we can introduce uncertainty into the model (2.2) as follows:

$$\theta_{iT} = \mathbf{x}_{iT}^{\mathsf{T}} \boldsymbol{\beta}_T + v_{iT},$$

where the error terms v_{iT} are independent random variables with mean 0 and unknown variance σ_{vT}^2. For sampling errors, we assume that the e_{iT}'s are independent normal variables with $\mathcal{E}\, e_{iT} = 0$ and $\mathcal{V}ar\ e_{iT} = \sigma_{iT}^2$, where σ_{iT}^2 is known. The combined model is given by

$$y_{iT} = \mathbf{x}_{iT}^{\mathsf{T}} \beta_T + v_{iT} + e_{iT}. \tag{2.4}$$

Under the model (2.4), the empirical Bayes estimator or the two-stage estimator of θ_{iT} is given as a weighted sum of the direct estimator y_{iT} and the regression synthetic estimator $\hat{\theta}_{iT}(\text{reg}) = \mathbf{x}_{iT}^{\mathsf{T}}\hat{\beta}_T$:

$$\hat{\theta}_{iT}(\hat{\sigma}_{vT}^2, \mathbf{y}_T) = w_{iT}y_{iT} + (1 - w_{iT})\hat{\theta}_{iT}(\text{reg}). \tag{2.5}$$

Here $\mathbf{y}_T = (y_{1T}, \ldots, y_{mT})^{\mathsf{T}}$, $w_{iT} = \hat{\sigma}_{vT}^2/(\hat{\sigma}_{vT}^2 + \sigma_{iT}^2)$, $\hat{\boldsymbol{\beta}}_T$ is the weighted least-squares estimator of $\boldsymbol{\beta}_T$ under the combined model (2.4), and $\hat{\sigma}_{vT}^2$ is an estimator of σ_{vT}^2. A simple moment estimator of σ_{vT}^2 (Prasad and Rao 1986, 1990) or a more complicated estimator, such as the maximum-likelihood estimator, may be employed. Using a moment estimator of σ_{vT}^2, Prasad and Rao (1986, 1990) obtained an estimator of the MSE of $\hat{\theta}_{iT}(\hat{\sigma}_{vT}^2, \mathbf{y}_T)$, correct to a second order of approximation for large m, by taking into account the uncertainty in estimating σ_{vT}^2.

Note that the ratio $\hat{\sigma}_{vT}^2/\sigma_{iT}^2$ measures between-small-area variation relative to sampling variance. More weight w_{iT} is given to the direct estimator y_{iT} as this ratio increases.

Fay and Herriot (1979) used estimators of the form (2.5) to estimate per capita income for small areas (with population less than 500 or 1000). They presented empirical evidence that (2.5) leads to smaller average error than either the direct survey estimate (based on a 20% sample from the 1970 U.S. Census of Population and Housing) or the county average. Datta, Fay, and Ghosh (1991) extended the Fay-Herriot model to multiple characteristics of interest, and derived empirical Bayes and hierarchical Bayes estimators of small-area means.

3. CROSS-SECTIONAL AND TIME-SERIES MODELS

The methods in Section 2 use only cross-sectional data for the current period T. As a result, they do not exploit the information in the data at other time points. Cross-sectional and time-series models are needed to take advantage of data at other time points. We consider one such model here, and using this model we extend the Fay-Herriot approach to time series of direct small-area estimators $\{y_{it}\}$, in conjunction with supplementary data $\{\mathbf{x}_{it}\}$, $i = 1, \ldots, m$ $t = 1, \ldots, T$. We assume that

$$y_{it} = \theta_{it} + e_{it}, \qquad i = 1, \ldots, I, \quad t = 1, \ldots, T, \tag{3.1}$$

where the e_{it}'s are sampling errors with $\mathcal{E}(e_{it}|\theta_{it}) = 0$.

3.1. Previously Proposed Models.

Extensive econometric literature exists on modelling and estimating relationships that combine time-series and cross-sectional data (see Judge *et al.* 1985, Chapter 13), but sampling errors are seldom taken into account. Following Anderson and Hsiao (1981), we assume that

$$\begin{aligned} \theta_{it} &= \mathbf{x}_{it}^{\mathsf{T}}\boldsymbol{\beta} + v_i + u_{it}, \\ u_{it} &= \rho u_{i,t-1} + \epsilon_{it}, \qquad |\rho| < 1, \end{aligned} \tag{3.2}$$

where $\mathbf{x}_{it} = (x_{it1}, \ldots, x_{itp})^{\mathsf{T}}$ is the vector of fixed concomitant variables for the ith area at time t, $v_i \overset{\text{iid}}{\sim} \mathsf{N}(0, \sigma_v^2)$, $\epsilon_{it} \overset{\text{iid}}{\sim} \mathsf{N}(0, \sigma^2)$, and $\{v_i\}$, $\{\epsilon_{it}\}$ are independent. Here the v_i's are random small-area effects, and the u_{it}'s follow a common AR(1) process for each i. The relevance of the model (3.2) may be better seen by rewriting it as a distributed-lag model:

$$\theta_{it} = \rho\theta_{i,t-1} + (\mathbf{x}_{it} - \rho\mathbf{x}_{i,t-1})^{\mathsf{T}}\boldsymbol{\beta} + (1 - \rho)v_i + \epsilon_{it}. \tag{3.2a}$$

The alternative form (3.2a) relates θ_{it} to the previous-period mean $\theta_{i,t-1}$, the values of the auxiliary variables for the time points t and $t-1$, and the small-area effect v_i. The special case of $\rho = 0$, which gives a nested error-variance-components model, has also been studied in the econometric literature.

More complex models than (3.2) may be formulated by assuming an ARMA process for the u_{it}'s instead of the simple AR(1) process, but the resulting efficiency gains relative to (3.2) are unlikely to be significant. Similarly, random slopes $\{\boldsymbol{\beta}_{it}\}$ obeying an autoregressive process may be used in place of constant slopes $\boldsymbol{\beta}$ (Pfeffermann and Burck 1990, Singh *et al.* 1991), but empirical evidence seems to suggest that the resulting efficiency gains relative to (3.2) are likely to be small (Singh *et al.* 1991).

The θ_{it}'s are related to direct survey estimators y_{it} through (3.1) and (3.2). Choudhry and Rao (1989) treated the composite error $w_{it} = e_{it} + u_{it}$ as an AR(1) process and assumed $\theta_{it} = \mathbf{x}_{it}^{\mathsf{T}}\boldsymbol{\beta} + v_i$, i.e., time effects on true area means are only through $\mathbf{x}_{it}$, unlike (3.2). The resulting nested error regression model may be written as

$$\begin{aligned} y_{it} &= \mathbf{x}_{it}^{\mathsf{T}}\boldsymbol{\beta} + v_i + w_{it}, \\ w_{it} &= \tilde{\rho} w_{i,t-1} + \tilde{\epsilon}_{it}, \qquad |\rho| < 1, \end{aligned} \tag{3.3}$$

where $\tilde{\epsilon}_{it} \overset{\text{iid}}{\sim} \mathsf{N}(0, \tilde{\sigma}^2)$. They obtained a two-stage estimator of the small-area mean θ_{it} under (3.3), and evaluated its efficiency relative to two synthetic estimators and the direct estimator y_{it}, using monthly survey estimates of unemployment for census divisions (small areas) from the Canadian Labour Force Survey in conjunction with monthly administrative counts from the Unemployment Insurance System and monthly survey estimates of population in the labour force as auxiliary variables.

3.2. Proposed Model.

The model (3.3) does not explicitly depend on the sampling error, but it is less realistic than the combined model based on (3.1) and (3.2), viz.,

$$\begin{aligned} y_{it} &= \mathbf{x}_{it}^{\mathsf{T}}\boldsymbol{\beta} + v_i + u_{it} + e_{it}, \\ u_{it} &= \rho u_{i,t-1} + \epsilon_{it}, \qquad |\rho| < 1. \end{aligned} \tag{3.4}$$

Following Fay and Herriot (1979), we assume that the sampling errors e_{it} are normally distributed with zero mean and block-diagonal covariance matrix $\boldsymbol{\Sigma}$ with arbitrary but known blocks $\boldsymbol{\Sigma}_i$ where $\boldsymbol{\Sigma}_i$ is a $T \times T$ matrix, i.e., we assume independence of sample errors across small areas. The covariance matrices $\boldsymbol{\Sigma}_i$ may be obtained by using the method of generalized variance functions (see Wolter 1985, Chapter 5) and prior estimates of panel correlations (see Lee 1990); a direct estimator of $\boldsymbol{\Sigma}_i$ is likely to be unstable due to small sample sizes in the areas. In practice, specification of $\boldsymbol{\Sigma}_i$ may not be easy, and further work on this problem would be useful.

The model (3.4) provides an extension of the Fay-Herriot model to cross-sectional data and time-series data. As mentioned earlier, Pfeffermann and Burck (1990) and Singh *et al.* (1991) also considered models of the form (3.4) and more general models, but they assumed specific models for the sampling errors e_{it}.

We focus on the extended Fay-Herriot model (3.4), and obtain a two-stage estimator of the current small-area means θ_{iT}. We also obtain an estimator of their MSE, correct to a second-order approximation for small or moderate T and relatively large m. This estimator of the MSE may be used as a measure of uncertainty in the two-stage estimator.

4. TWO-STAGE ESTIMATOR: ρ KNOWN

Arranging the data $\{y_{it}\}$ as $\mathbf{y} = (y_{11}, \ldots, y_{1T}; \ldots; y_{m1}, \ldots, y_{mT})^{\mathsf{T}} = (\mathbf{y}_1^{\mathsf{T}}, \ldots, \mathbf{y}_m^{\mathsf{T}})^{\mathsf{T}}$, the proposed model (3.4) may be written in matrix form as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{v} + \mathbf{u} + \mathbf{e} \tag{4.1}$$

with

$$\mathbf{X} = (\mathbf{X}_1^{\mathsf{T}}, \ldots, \mathbf{X}_m^{\mathsf{T}})^{\mathsf{T}}, \qquad \mathbf{X}_i^{\mathsf{T}} = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{iT}),$$
$$\mathbf{Z} = \mathbf{I}_m \otimes \mathbf{1}_T,$$
$$\mathbf{v} = (v_1, \ldots, v_m)^{\mathsf{T}}, \qquad \mathbf{u} = (\mathbf{u}_1^{\mathsf{T}}, \ldots, \mathbf{u}_m^{\mathsf{T}})^{\mathsf{T}},$$
$$\mathbf{e} = (\mathbf{e}_1^{\mathsf{T}}, \ldots, \mathbf{e}_m^{\mathsf{T}})^{\mathsf{T}},$$

where $\mathbf{u}_i^{\mathsf{T}} = (u_{i1}, \ldots, u_{iT})$, $\mathbf{e}_i^{\mathsf{T}} = (e_{i1}, \ldots, e_{iT})$, $\mathbf{1}_T$ is a T-vector of 1's, $\mathbf{I}_m$ is the identity matrix of order m, and $\otimes$ denotes the direct product. Further,

$$\begin{aligned} \mathcal{E}\mathbf{v} &= \mathbf{0}, \qquad \mathcal{Cov}\ \mathbf{v} = \sigma_v^2 \mathbf{I}_m, \\ \mathcal{E}\mathbf{u} &= \mathbf{0}, \qquad \mathcal{Cov}\ \mathbf{u} = \sigma^2 \mathbf{I}_m \otimes \boldsymbol{\Gamma} = \sigma^2 \mathbf{R}, \\ \mathcal{E}\mathbf{e} &= \mathbf{0}, \qquad \mathcal{Cov}\ \mathbf{e} = \boldsymbol{\Sigma} = \text{block diag}(\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_m), \end{aligned} \tag{4.2}$$

and $\mathbf{v}$, $\mathbf{u}$, and $\mathbf{e}$ are mutually independent, where $\boldsymbol{\Gamma}$ is a $T \times T$ matrix with elements $\rho^{|i-j|}/(1-\rho^2)$. It follows from (4.1) and (4.2) that

$$\begin{aligned} \mathcal{Cov}\ \mathbf{y} = \mathbf{V} &= \boldsymbol{\Sigma} + \sigma^2 \mathbf{R} + \sigma_v^2 \mathbf{Z}\mathbf{Z}^{\mathsf{T}} \\ &= \text{block diag}_i(\boldsymbol{\Sigma}_i + \sigma^2 \boldsymbol{\Gamma} + \sigma_v^2 \mathbf{J}_T) = \text{block diag}_i(\mathbf{V}_i), \qquad \text{say}, \end{aligned}$$

with $\mathbf{J}_T = \mathbf{1}_T \mathbf{1}_T^{\mathsf{T}}$.

4.1. BLUP Estimator.

The current small-area mean $\theta_{iT} = \mathbf{x}_{iT}^{\mathsf{T}}\boldsymbol{\beta} + v_i + u_{iT}$ is a special case of the linear combination $\tau = \mathbf{l}^{\mathsf{T}}\boldsymbol{\beta} + \mathbf{l}_1^{\mathsf{T}}\mathbf{v} + \mathbf{l}_2^{\mathsf{T}}\mathbf{u}$, where $\mathbf{l} = \mathbf{x}_{iT}$, $\mathbf{l}_1$ is the m-vector with 1 in the ith position and 0 elsewhere, and $\mathbf{l}_2$ is the mT-vector with 1 in the iTth position and 0 elsewhere. Noting that the model (4.1) is a special case of the general mixed linear model, the BLUP estimator of θ_{iT} can be obtained from Henderson's general results (Henderson 1975) for arbitrary linear combinations τ.

Assuming first that σ_u^2, σ_v^2, and ρ are known, the BLUP estimator of τ is given by

$$\tilde{\tau} = \mathbf{l}^{\mathsf{T}}\tilde{\boldsymbol{\beta}} + (\sigma_v^2 \mathbf{l}_1^{\mathsf{T}} \mathbf{Z}^{\mathsf{T}} + \sigma^2 \mathbf{l}_2^{\mathsf{T}} \mathbf{R}) \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}), \tag{4.3}$$

where $\tilde{\boldsymbol{\beta}} = (\mathbf{X}^{\mathsf{T}}\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^{\mathsf{T}}\mathbf{V}^{-1}\mathbf{y}$ is the generalized least-squares estimator of $\boldsymbol{\beta}$. Using the special structures of $\mathbf{l}_1$, $\mathbf{l}_2$, $\mathbf{Z}$, $\mathbf{R}$, and $\mathbf{V}$, it is easily seen that (4.3) reduces to

$$\begin{aligned} \tilde{\theta}_{iT} &= a(\sigma^2, \sigma_v^2, \rho, \mathbf{y}) \\ &= \mathbf{x}_{iT}^{\mathsf{T}}\tilde{\boldsymbol{\beta}} + (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)^{\mathsf{T}} (\boldsymbol{\Sigma}_i + \sigma^2 \boldsymbol{\Gamma} + \sigma_v^2 \mathbf{J}_T)^{-1} (\mathbf{y}_i - \mathbf{X}_i \tilde{\boldsymbol{\beta}}), \end{aligned} \tag{4.4}$$

where $\boldsymbol{\gamma}_T$ is the Tth row of $\boldsymbol{\Gamma}$. The BLUP estimator may also be written as a weighted sum of the direct estimator y_{iT}, the synthetic estimator $\mathbf{x}_{iT}^{\mathsf{T}}\tilde{\boldsymbol{\beta}}$, and the residuals $y_{ij} - \mathbf{x}_{ij}^{\mathsf{T}}\tilde{\boldsymbol{\beta}}$, $j = 1, \ldots, T-1$:

$$\tilde{\theta}_{iT} = w_{iT}^* y_{iT} + (1 - w_{iT}^*)\mathbf{x}_{iT}^{\mathsf{T}}\tilde{\boldsymbol{\beta}} + \sum_{j=1}^{T-1} w_{ij}^*(y_{ij} - \mathbf{x}_{ij}^{\mathsf{T}}\tilde{\boldsymbol{\beta}}),$$

where

$$(w_{i1}^*, \ldots, w_{iT}^*) = (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)^{\mathsf{T}} \mathbf{V}_i^{-1}.$$

4.2. Two-Stage Estimator.

In practice, the parameters, σ^2, σ_v^2, and ρ are usually unknown. We first assume that ρ is known and replace σ^2 and σ_v^2 in (4.4) by their consistent estimators $\hat{\sigma}^2(\rho)$, $\hat{\sigma}_v^2(\rho)$ to obtain a two-stage estimator $\hat{\theta}_{iT}(\rho)$. The case of unknown ρ is studied in Section 6.

Pantula and Pollock (1985) estimated σ^2, σ_v^2 in the nested error-regression model (3.3) with autocorrelated errors w_{it} by extending a simple transformation method for the special case of independent errors (Fuller and Battese 1973). We now extend their method to the more general model (3.4) with both autocorrelated errors u_{it} and sampling errors e_{it}, assuming ρ is known.

We first obtain an unbiased estimator of σ^2. For this purpose, we transform the model (3.4) to eliminate the random effect v_i. First transform $\mathbf{y}_i$ to $\mathbf{z}_i = \mathbf{P}\mathbf{y}_i$ such that the covariance matrix of $\mathbf{P}\mathbf{u}_i$ is $\sigma^2 \mathbf{I}_T$, i.e., $\boldsymbol{\Gamma} = \mathbf{P}^{-1}(\mathbf{P}^{-1})^{-1}$. The $T \times T$ matrix $\mathbf{P}$ has the following form: first diagonal element $(1-\rho^2)^{\frac{1}{2}}$; remaining diagonal elements 1; $(t+1, t)$th element $-\rho$ for $t = 1, \ldots, T-1$; and remaining elements 0 (see Judge *et al.* 1985, p. 285). The transformed model is given by

$$\mathbf{z}_i = \mathbf{P}\mathbf{X}_i\boldsymbol{\beta} + \mathbf{f}v_i + \mathbf{P}(\mathbf{u}_i + \mathbf{e}_i), \qquad i = 1, \ldots, m, \tag{4.5}$$

where $\mathbf{f} = (f_1, \ldots, f_T)^{\mathsf{T}}$ with $f_1 = (1-\rho^2)^{\frac{1}{2}}$ and $f_t = 1 - \rho$ for $2 \leq t \leq T$. Next we transform $\mathbf{z}_i$ to $\mathbf{z}_i^{(1)} = (\mathbf{I}_T - \mathbf{D})\mathbf{z}_i$, where $\mathbf{D} = (\mathbf{f}\mathbf{f}^{\mathsf{T}})/c$ with $c = \mathbf{f}^{\mathsf{T}}\mathbf{f} = (1-\rho)\{T-(T-2)\rho\}$. This leads to the following reduced model:

$$\mathbf{z}_i^{(1)} = \mathbf{H}_i^{(1)}\boldsymbol{\beta} + \mathbf{e}_i^*, \qquad i = 1, \ldots, m, \tag{4.6}$$

where $\mathbf{H}_i^{(1)} = (\mathbf{I}_T - \mathbf{D})\mathbf{P}\mathbf{X}_i$ and $\mathbf{e}_i^* = (\mathbf{I}_T - \mathbf{D})\mathbf{P}(\mathbf{u}_i + \mathbf{e}_i)$. Since

$$Cov\ \mathbf{e}_i^* = (\mathbf{I}_T - \mathbf{D})(\sigma^2\mathbf{I}_T + \mathbf{P}\boldsymbol{\Sigma}_i\mathbf{P}^{\mathsf{T}})(\mathbf{I}_T - \mathbf{D})^{\mathsf{T}} \tag{4.7}$$

does not involve σ_v^2, we can estimate σ^2 through the reduced model (4.6) using the residual sum of squares. Let $\mathbf{z}^{(1)} = [(\mathbf{z}_1^{(1)})^{\mathsf{T}}, \ldots, (\mathbf{z}_m^{(1)})^{\mathsf{T}}]^{\mathsf{T}}$, $\mathbf{H}^{(1)} = [(\mathbf{H}_1^{(1)})^{\mathsf{T}}, \ldots, (\mathbf{H}_m^{(1)})^{\mathsf{T}}]^{\mathsf{T}}$, and $\hat{\mathbf{e}}^{\mathsf{T}}\hat{\mathbf{e}}$ be the residual sum of squares obtained by regressing $\mathbf{z}^{(1)}$ on $\mathbf{H}^{(1)}$ using ordinary least squares. An unbiased estimator σ^2 is then given by

$$\begin{aligned}\tilde{\sigma}^2(\rho) = \big(\hat{\mathbf{e}}^{\mathsf{T}}\hat{\mathbf{e}} - \operatorname{tr}[\{\text{block diag}_i(\mathbf{I}_T - \mathbf{D}) - \mathbf{H}^{(1)}(\mathbf{H}^{(1)\,\mathsf{T}}\mathbf{H}^{(1)})^{-}\mathbf{H}^{(1)\,\mathsf{T}}\} \\ \times \text{block diag}_i(\mathbf{P}\boldsymbol{\Sigma}_i\mathbf{P}^{\mathsf{T}})]\big)\{m(T-1) - \mathrm{r}(H^{(1)})\}^{-1},\end{aligned} \tag{4.8}$$

where $\mathbf{A}^-$ is a generalized inverse of $\mathbf{A}$. The unbiasedness of $\tilde{\sigma}^2(\rho)$ follows by noting that $\hat{\mathbf{e}}^{\mathsf{T}}\hat{\mathbf{e}} = \mathbf{e}^{*T}(\mathbf{I}_{mT} - \mathbf{H}^{(1)}(\mathbf{H}^{(1)\,\mathsf{T}}\mathbf{H}^{(1)})^{-}\mathbf{H}^{(1)\,\mathsf{T}})\mathbf{e}^*$ and then using (4.7).

Turning to the estimation of σ_v^2, we transform (4.5) by changing $\mathbf{z}_i$ to $c^{-\frac{1}{2}}\mathbf{f}^{\mathsf{T}}\mathbf{z}_i$ such that $u_i^* = c^{-\frac{1}{2}}\mathbf{f}^{\mathsf{T}}\mathbf{P}\mathbf{u}_i$ has mean 0 and variance σ^2. The transformed model is given by

$$c^{-\frac{1}{2}}\mathbf{f}^{\mathsf{T}}\mathbf{z}_i = c^{-\frac{1}{2}}\mathbf{f}^{\mathsf{T}}\mathbf{P}\mathbf{X}_i\boldsymbol{\beta} + c^{\frac{1}{2}}v_i + u_i^* + c^{-\frac{1}{2}}\mathbf{f}^{\mathsf{T}}\mathbf{P}\mathbf{e}_i$$

with error variance $c\sigma_v^2 + \sigma^2 + c^{-1}\mathbf{f}^{\mathsf{T}}\mathbf{P}\boldsymbol{\Sigma}_i\mathbf{P}^{\mathsf{T}}\mathbf{f}$. Let $\hat{\mathbf{u}}^{\mathsf{T}}\hat{\mathbf{u}}$ be the residual sum of squares obtained by regressing $c^{-\frac{1}{2}}\mathbf{f}^{\mathsf{T}}\mathbf{z}_i$ on $c^{-\frac{1}{2}}\mathbf{f}^{\mathsf{T}}\mathbf{P}\mathbf{X}_i$ using ordinary least squares. An unbiased estimator $\tilde{\sigma}_v^2$ is then given by

$$\tilde{\sigma}_v^2(\rho) = c^{-1}\{m - \mathrm{r}(\mathbf{F})\}^{-1}(\hat{\mathbf{u}}^{\mathsf{T}}\hat{\mathbf{u}} - \mathrm{tr}[\{\mathbf{I}_m - \mathbf{F}(\mathbf{F}^{\mathsf{T}}\mathbf{F})^{-}\mathbf{F}^{\mathsf{T}}\}\,\mathrm{diag}_i(c^{-1}\mathbf{f}^{\mathsf{T}}\mathbf{P}\boldsymbol{\Sigma}_i\mathbf{P}^{\mathsf{T}}\mathbf{f})\}) - c^{-1}\tilde{\sigma}^2(\rho), \quad (4.9)$$

where $\mathbf{F} = (\mathbf{X}_1^{\mathsf{T}}\mathbf{P}^{\mathsf{T}}\mathbf{f}, \ldots, \mathbf{X}_m^{\mathsf{T}}\mathbf{P}^{\mathsf{T}}\mathbf{f})^{\mathsf{T}}$. The unbiasedness of $\tilde{\sigma}_v^2(\rho)$ follows by noting that

$$\mathcal{E}(\hat{\mathbf{u}}'\hat{\mathbf{u}}) = (c\sigma_v^2 + \sigma^2)\{m - \mathrm{r}(\mathbf{F})\} + \mathrm{tr}[\{\mathbf{I}_m - \mathbf{F}(\mathbf{F}^{\mathsf{T}}\mathbf{F})^{-}\mathbf{F}^{\mathsf{T}}\}\,\mathrm{diag}_i(c^{-1}\mathbf{f}^{\mathsf{T}}\mathbf{P}\boldsymbol{\Sigma}_i\mathbf{P}^{\mathsf{T}}\mathbf{f})].$$

If follows from (4.8) and (4.9) that only ordinary least squares is required for calculating $\tilde{\sigma}_v^2(\rho)$ and $\tilde{\sigma}^2(\rho)$.

Since $\tilde{\sigma}^2(\rho)$ and $\tilde{\sigma}_v^2(\rho)$ can take negative values, we truncate them at zero and use

$$\hat{\sigma}^2(\rho) = \max\{0, \tilde{\sigma}^2(\rho)\}, \qquad \hat{\sigma}_v^2(\rho) = \max\{0, \tilde{\sigma}_v^2(\rho)\}.$$

The truncated estimators $\hat{\sigma}^2(\rho)$ and $\hat{\sigma}_v^2(\rho)$ are no longer unbiased, but they are consistent as $m \longrightarrow \infty$.

A two-stage estimator of θ_{iT} is now obtained from (4.4) by substituting $\hat{\sigma}^2(\rho)$ and $\hat{\sigma}_v^2(\rho)$ for σ^2 and σ_v^2 respectively:

$$\hat{\theta}_{iT}(\rho) = a[\hat{\sigma}^2(\rho), \hat{\sigma}_v^2(\rho), \rho, \mathbf{y}]. \quad (4.10)$$

The two-stage estimator $\hat{\theta}_{iT}(\rho)$ remains unbiased by noting that $\hat{\sigma}^2(\rho)$ and $\hat{\sigma}_v^2(\rho)$ are even functions of $\mathbf{y}$ and translation-invariant, i.e., they remain unchanged when $\mathbf{y}$ is changed to $-\mathbf{y}$ or to $\mathbf{y} - \mathbf{X}\mathbf{a}$ for all $\mathbf{y}$ and $\mathbf{a}$; see Kackar and Harville (1984).

It is not necessary to assume normality of the errors in the model (3.4); only symmetric distributions are needed. However, we need normality to derive an estimator of the MSE of $\hat{\theta}_{iT}(\rho)$, correct to a second-order approximation as $m \longrightarrow \infty$.

5. ESTIMATOR OF THE MSE

We first obtain a second-order approximation to the MSE of the two-stage estimator $\hat{\theta}_{iT}(\rho)$, in the sense that the neglected terms are $o(m^{-1})$ for large m. Using the approximation, an estimator of $\mathrm{MSE}[\hat{\theta}_{iT}(\rho)]$, correct to the same order of approximation, is then obtained.

5.1. Second-Order Approximation to the MSE

Following Kackar and Harville (1984), we have, under normality of the errors v_i, u_{it}, and e_{it},

$$\mathrm{MSE}[\hat{\theta}_{iT}(\rho)] = \mathrm{MSE}(\tilde{\theta}_{iT}) + \mathcal{E}\{\tilde{\theta}_{iT} - \hat{\theta}_{iT}(\rho)\}^2, \quad (5.1)$$

where $\mathrm{MSE}(\tilde{\theta}_{iT}) = \mathcal{E}(\tilde{\theta}_{iT} - \theta_{iT})^2$. Further, using Henderson's general result (Henderson 1975), an exact expression for $\mathrm{MSE}(\tilde{\theta}_{iT})$ is given by

$$\mathrm{MSE}(\hat{\theta}_{iT}) = g_{1iT}(\sigma^2, \sigma_v^2, \rho) + g_{2iT}(\sigma^2, \sigma_v^2, \rho), \quad (5.2)$$

where

$$g_{1iT}(\sigma^2, \sigma_v^2, \rho) = \sigma_v^2 + \frac{\sigma^2}{1-\rho^2} - (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)^{\mathsf{T}} \mathbf{V}_i^{-1} (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)$$

and

$$g_{2iT}(\sigma^2, \sigma_v^2, \rho) = \{\mathbf{x}_{iT} - \mathbf{X}_i^{\mathsf{T}} \mathbf{V}_i^{-1} (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)\}^{\mathsf{T}} (\mathbf{X}^{\mathsf{T}} \mathbf{V}^{-1} \mathbf{X})^{-1} \times \{\mathbf{x}_{iT} - \mathbf{X}_i^{\mathsf{T}} \mathbf{V}_i^{-1} (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)\}.$$

The second term $g_{2iT}(\sigma^2, \sigma_v^2, \rho)$ in (5.2), due to estimating $\boldsymbol{\beta}$, is of order $O(m^{-1})$, while the first term $g_{1iT}(\sigma^2, \sigma_v^2, \rho)$ is $O(1)$.

It remains to evaluate the term $\mathcal{E}\{\tilde{\theta}_{iT} - \hat{\theta}_{iT}(\rho)\}^2$ in (5.1). Following Kackar and Harville (1984), we propose a Taylor approximation to this term:

$$\mathcal{E}\{\tilde{\theta}_{iT} - \hat{\theta}_{iT}(\rho)\}^2 \simeq \mathcal{E}\left(\frac{\partial \tilde{\theta}_{iT}}{\partial \sigma^2}\{\tilde{\sigma}^2(\rho) - \sigma^2\} + \frac{\partial \tilde{\theta}_{iT}}{\partial \sigma_v^2}\{\tilde{\sigma}_v^2(\rho) - \sigma_v^2\}\right)^2.$$

Following Prasad and Rao (1990), a further approximation is obtained as

$$\mathcal{E}[\tilde{\theta}_{iT} - \hat{\theta}_{iT}(\rho)]^2 \approx \operatorname{tr}(\boldsymbol{\Delta}^{\mathsf{T}} \mathbf{V} \boldsymbol{\Delta} \boldsymbol{\Sigma}^*) = g_{3iT}(\sigma^2, \sigma_v^2, \rho), \qquad \text{say,} \tag{5.3}$$

where $\boldsymbol{\Sigma}^*$ is the 2×2 covariance matrix of the unbiased estimators $\tilde{\sigma}^2(\rho)$ and $\tilde{\sigma}_v^2(\rho)$, and $\boldsymbol{\Delta} = (\partial \mathbf{b}/\partial \sigma^2, \partial \mathbf{b}/\partial \sigma_v^2)$ with $\mathbf{b}' = (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)^{\mathsf{T}} \mathbf{V}_i^{-1}$. Calculating the derivatives $\partial \mathbf{b}/\partial \sigma^2$ and $\partial \mathbf{b}/\partial \sigma_v^2$, we obtain, after simplification, $\boldsymbol{\Delta}^{\mathsf{T}} \mathbf{V} \boldsymbol{\Delta} = \mathbf{A}$, where $\mathbf{A}$ is a 2×2 symmetric matrix with diagonal elements

$$a_{11} = \{\boldsymbol{\gamma}_T - \boldsymbol{\Gamma} \mathbf{V}_i^{-1} (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)\}^{\mathsf{T}} \mathbf{V}_i^{-1} \{\boldsymbol{\gamma}_T - \boldsymbol{\Gamma} \mathbf{V}_i^{-1} (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)\}$$

$$a_{22} = \{\mathbf{1}_T - \mathbf{J}_T \mathbf{V}_i^{-1} (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)\}^{\mathsf{T}} \mathbf{V}_i^{-1} \{\mathbf{1}_T - \mathbf{J}_T \mathbf{V}_i^{-1} (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)\}$$

and off-diagonal elements

$$a_{12} = a_{21} = \{\boldsymbol{\gamma}_T - \boldsymbol{\Gamma} \mathbf{V}_i^{-1} (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)\}^{\mathsf{T}} \mathbf{V}_i^{-1} \{\mathbf{1}_T - \mathbf{J}_T \mathbf{V}_i^{-1} (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)\}.$$

The term $g_{3iT}(\sigma^2, \sigma_v^2, \rho)$ is of the same order as $g_{2iT}(\sigma^2, \sigma_v^2, \rho)$. Combining (5.2) and (5.3), we get a second-order approximation to $\mathrm{MSE}[\hat{\theta}_{iT}(\rho)]$ as

$$\mathrm{MSE}[\hat{\theta}_{iT}(\rho)] \simeq g_{1iT}(\sigma^2, \sigma_v^2, \rho) + g_{2iT}(\sigma^2, \sigma_v^2, \rho) + g_{3iT}(\sigma^2, \sigma_v^2, \rho). \tag{5.4}$$

The neglected terms in the approximation (5.4) are of lower order, $o(m^{-1})$, for large m. A rigorous proof of this assertion is quite involved, but follows along the lines of Prasad and Rao (1990), using the following regularity conditions:

(1) There exist positive constants c_0 and c_1 such that

$$c_0 \le (mT)^{-1}(\text{smallest eigenvalue of } \mathbf{X}^{\mathsf{T}}\mathbf{X})$$

and $\|x_{it}\| \le c_1$ for all i and t, where $\|\cdot\|$ denotes the Euclidean norm.

(2) There exist positive constants a_0 and a_1 such that

$$a_0 \mathbf{I}_T \le \boldsymbol{\Sigma}_i \le a_1 \mathbf{I}_T$$

for all i.

It remains to obtain the covariance matrix $\mathbf{\Sigma}^*$. For this purpose we rewrite $\tilde{\sigma}^2(\rho)$ and $\tilde{\sigma}_v^2(\rho)$ as

$$\tilde{\sigma}^2(\rho) = \{(m-1)T - \mathrm{r}(\mathbf{H}^{(1)})\}^{-1}\mathbf{a}^\mathsf{T}\mathbf{C}_1\mathbf{a} + \text{const}, \tag{5.5}$$

$$\tilde{\sigma}_v^2(\rho) = c^{-1}\{m - \mathrm{r}(\mathbf{F})\}^{-1}\mathbf{a}^\mathsf{T}\mathbf{C}_2\mathbf{a} - c^{-1}(m-1)T - \mathrm{r}(\mathbf{H}^{(1)})\}^{-1}\mathbf{a}^\mathsf{T}\mathbf{C}_1\mathbf{a} + \text{const}, \tag{5.6}$$

where $\mathbf{a} = \mathbf{Zv} + \mathbf{u} + \mathbf{e} \sim \mathrm{N}(\mathbf{0}, \mathbf{V})$,

$$\mathbf{C}_1 = \mathbf{C}^\mathsf{T}\{\mathbf{I}_n - \mathbf{CX}(\mathbf{X}^\mathsf{T}\mathbf{CC}^\mathsf{T}\mathbf{X})^{-1}\mathbf{X}^\mathsf{T}\mathbf{C}^\mathsf{T}\}\mathbf{C}$$

with $\mathbf{C} = \text{block diag}_i[(\mathbf{I}_T - \mathbf{D})\mathbf{P}]$, and

$$\mathbf{C}_2 = \mathbf{C}^{*\mathsf{T}}\{\mathbf{I}_m - \mathbf{C}^{*\mathsf{T}}\mathbf{X}(\mathbf{X}^\mathsf{T}\mathbf{C}^{*\mathsf{T}}\mathbf{C}^*\mathbf{X})^{-1}\mathbf{X}^\mathsf{T}\mathbf{C}^{*\mathsf{T}}\}\mathbf{C}^*$$

with $\mathbf{C}^* = \text{block diag}_i(c_i^{-\frac{1}{2}}\mathbf{f}^\mathsf{T}\mathbf{P})$. We can now evaluate the elements of $\mathbf{\Sigma}^*$ using (5.5) and (5.6) in the following well-known lemma on the covariance of two quadratic forms of normally distributed variables.

LEMMA . *If* $\mathbf{y} \sim \mathrm{N}(\mathbf{0}, \mathbf{\Omega})$, *then* $Cov(\mathbf{y}^\mathsf{T}\mathbf{G}_1\mathbf{y}, \mathbf{y}^\mathsf{T}\mathbf{G}_2\mathbf{y}) = 2\,\mathrm{tr}(\mathbf{G}_1\mathbf{\Omega}\,\mathbf{G}_2\mathbf{\Omega})$, *where* $\mathbf{G}_1$ *and* $\mathbf{G}_2$ *are two symmetric matrices.*

5.2. Second-Order Approximation to Estimator of MSE.

Following Prasad and Rao (1990), it can be shown that

$$\mathcal{E}g_{1iT}\big(\hat{\sigma}^2(\rho), \hat{\sigma}_v^2(\rho), \rho\big) \simeq g_{1iT}(\sigma^2, \sigma_v^2, \rho) - g_{3iT}(\sigma^2, \sigma_v^2, \rho), \tag{5.7}$$

$$\mathcal{E}g_{2iT}\big(\hat{\sigma}^2(\rho), \hat{\sigma}_v^2(\rho), \rho\big) \simeq g_{2iT}(\sigma^2, \sigma_v^2, \rho), \tag{5.8}$$

$$\mathcal{E}g_{3iT}\big(\hat{\sigma}^2(\rho), \hat{\sigma}_v^2(\rho), \rho\big) \simeq g_{3iT}(\sigma^2, \sigma_v^2, \rho). \tag{5.9}$$

The neglected terms in the approximations (5.7)–(5.9) are of lower order, $o(m^{-1})$. A rigorous proof of this assertion is quite involved, but essentially involves showing that $\mathcal{E}\,o_p(m^{-1}) = o(m^{-1})$ under regularity conditions.

It now follows from (5.7)–(5.9) and (5.4) that an estimator of $\mathrm{MSE}[\hat{\theta}_{iT}(\rho)]$ with expectation correct to $O(m^{-1})$ is given by

$$\mathrm{mse}[\hat{\theta}_{iT}(\rho)] \simeq g_{1iT}\big(\hat{\sigma}^2(\rho), \hat{\sigma}_v^2(\rho), \rho\big) + g_{2iT}\big(\hat{\sigma}^2(\rho), \hat{\sigma}_v^2(\rho), \rho\big) + 2g_{3iT}\big(\hat{\sigma}^2(\rho), \hat{\sigma}_v^2(\rho), \rho\big). \tag{5.10}$$

Although (5.10) is derived under the assumption of a small or moderate T and a relatively large m, limited simulation results in Section 7 with $T = 40$ and $m = 40$ suggest that it might remain fairly accurate even when T and m are of the same order.

A naive estimator of the MSE, ignoring the uncertainty in the estimators $\hat{\sigma}^2(\rho)$ and $\hat{\sigma}_v^2(\rho)$, is given by

$$\mathrm{mse}_N[\hat{\theta}_{iT}(\rho)] = g_{1iT}\big(\hat{\sigma}^2(\rho), \hat{\sigma}_v^2(\rho), \rho\big) + g_{2iT}\big(\hat{\sigma}^2(\rho), \hat{\sigma}_v^2(\rho), \rho\big). \tag{5.11}$$

This estimator is computationally simpler than (5.10), since it does not require $g_{3iT}\big(\hat{\sigma}^2(\rho), \hat{\sigma}_v^2(\rho), \rho\big)$, which involves the estimation of covariance matrix of the variance components

$\hat{\sigma}^2(\rho)$ and $\hat{\sigma}_v^2(\rho)$. However, it can lead to severe underestimation of the true MSE (see Section 7).

6. TWO-STAGE ESTIMATOR; ρ UNKNOWN

In Sections 4 and 5 we obtained a two-stage estimator of θ_{iT} and an estimator of its MSE, assuming known ρ. However, in practice ρ is seldom known. We propose here three possible methods for handling the case of unknown ρ. In method 1, a two-stage estimator $\hat{\theta}_{iT}(\rho)$ based on a prior guess ρ_0 is used, and its MSE is estimated by substituting ρ_0 for ρ in (5.10). Denote this estimator of the MSE as $\text{mse}[\hat{\theta}_{iT}(\rho_0)]$.

In method 2, we ignore the sampling errors e_{it} and obtain a moment estimator of ρ, along the lines of Pantula and Pollock (1985). This naive estimator of ρ is given by

$$\hat{\rho}_N = \frac{\sum_{i=1}^{m}\sum_{t=1}^{T-2}\hat{a}_{it}(\hat{a}_{i,t+1} - \hat{a}_{i,t+2})}{\sum_{i=1}^{m}\sum_{t=1}^{T-2}\hat{a}_{it}(\hat{a}_{i,t} - \hat{a}_{i,t+1})}, \qquad T > 2, \tag{6.1}$$

where $\hat{a}_{it} = y_{it} - \mathbf{x}_{it}^{\mathsf{T}}(\mathbf{X}^{\mathsf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathsf{T}}\mathbf{y}$ is the itth ordinary least-squares residual. The estimator $\hat{\rho}_N$ is inconsistent and typically underestimates ρ (see Section 7). Nevertheless, the resulting two-stage estimator $\hat{\theta}_{iT}(\hat{\rho}_N)$ remains unbiased, because $\hat{\rho}_N$ is an even function of $\mathbf{y}$ and translation-invariant. The MSE of $\hat{\theta}_{iT}(\hat{\rho}_N)$ is estimated by substituting $\hat{\rho}_N$ for ρ in (5.10). Denote this estimator of the MSE as $\text{mse}[\hat{\theta}_{iT}(\hat{\rho}_N)]$.

Turning to method 3, we obtain a consistent moment estimator of ρ by taking into account the sampling errors. It is given by

$$\hat{\rho} = \frac{\sum_{i=1}^{m}\sum_{t=1}^{T-2}\{\hat{a}_{it}(\hat{a}_{i,t+1} - \hat{a}_{i,t+2}) - (\sigma_{t,t+1}^{(i)} - \sigma_{t,t+2}^{(i)})\}}{\sum_{i=1}^{m}\sum_{t=1}^{T-2}\{\hat{a}_{it}(\hat{a}_{i,t} - \hat{a}_{i,t+1}) - (\sigma_{t,t}^{(i)} - \sigma_{t,t+1}^{(i)})\}}, \qquad T > 2, \tag{6.2}$$

where $\sigma_{t,t}^{(i)} = \mathcal{V}ar\ e_{it}$, $\sigma_{t,t+1}^{(i)} = \mathcal{C}ov(e_{it}, e_{i,t+1})$ and $\sigma_{t,t+2}^{(i)} = \mathcal{C}ov(e_{it}, e_{i,t+2})$. The resulting two-stage estimator $\hat{\theta}_{iT}(\hat{\rho})$ remains unbiased, and its MSE is estimated by substituting $\hat{\rho}$ for ρ in (5.10). Denote this estimator of the MSE as $\text{mse}[\hat{\theta}_{iT}(\hat{\rho})]$. Rao and Yu (1992) obtained a second-order approximation to the MSE which accounts for the uncertainty in the estimator $\hat{\rho}$, unlike $\text{mse}[\hat{\theta}_{iT}(\hat{\rho})]$. Our simulation results indicated difficulties with method 3, since $\hat{\rho}$ often took values outside the admissible range, $(-1, 1)$, especially for small T or small σ^2 relative to the sampling variation.

7. SIMULATION STUDY

To study the efficiency of two-stage estimators and relative biases of estimators of the MSE, we conducted a limited simulation study using the following simple model:

$$\begin{aligned} y_{it} &= v_i + u_{it} + e_{it}, \\ u_{it} &= \rho u_{i,t-1} + \epsilon_{it}, \qquad |\rho| < 1, \end{aligned} \tag{7.1}$$

with $\rho = 0.2$ and 0.4, $e_{it} \overset{\text{iid}}{\sim} \mathrm{N}(0, 1)$, $v_i \overset{\text{iid}}{\sim} \mathrm{N}(0, \sigma_v^2)$, $\epsilon_{it} \overset{\text{iid}}{\sim} \mathrm{N}(0, \sigma^2)$. [Choudhry and Rao (1989) obtained $\text{est}(\tilde{\rho}) = 0.36$ under the model (3.3) using data from the Canadian Labour Force Survey.] We assumed uncorrelated sampling errors e_{it} and sampling variances all equal to one (i.e., $\boldsymbol{\Sigma}_i = \mathbf{I}_T$) for simplicity, but extensive simulation studies with more realistic $\boldsymbol{\Sigma}_i$ are clearly needed.

We used $m = 40$ small areas and both small T $(= 5)$ and moderate T $(= 10)$, and generated 5000 independent samples $\{y_{it}\}$ for each selected pair (σ_v^2, σ^2). We also

TABLE 1: Gain in efficiency of the two-stage estimator $\hat{\theta}_{1T}(\rho)$ over the Fay-Herriot estimator: $\rho = 0.2$.

	Efficiency gain (%)			
σ_v^2	$\sigma^2 = 0.25$	0.5	1.0	2.0
	$T = 5$			
0.25	32	19	10	5
0.5	51	29	15	7
1.0	77	45	22	10
2.0	107	64	32	15
	$T = 10$			
0.25	49	26	13	6
0.5	76	40	18	8
1.0	114	60	28	11
2.0	157	86	41	27

considered large T ($= 40$) for $\rho = 0.4$ to study the effect of T. Note that σ_v^2 and σ^2 represent between-small-area variation relative to sampling variation and between-time variation relative to sampling variation, respectively, since $\mathcal{V}ar\ e_{it} = 1$.

From each simulated sample, the two-stage estimators and the Fay-Herriot estimator of $\theta_{1T} = v_1 + u_{1T}$ and the estimators of the MSE were computed. Note that it is sufficient to consider θ_{1T}, since the MSE values will be the same for all θ_{iT}, $i = 1, \ldots, 40$. Simulated values of the MSE of any estimator $\hat{\theta}_{1T}$ (say) and relative bias of an estimator of the MSE, say mse, were computed as follows:

$$\mathrm{MSE}(\hat{\theta}_{1T}) = \frac{1}{5000} \sum_{s=1}^{5000} (\hat{\theta}_{1T}^s - \theta_{1T}^s)^2,$$

$$\mathrm{RB}[\mathrm{mse}] = \frac{1}{\mathrm{MSE}(\hat{\theta}_{1T})} \left(\frac{1}{5000} \sum_{s=1}^{5000} (\mathrm{mse}_s) - \mathrm{MSE}(\hat{\theta}_{1T}) \right),$$

where $\hat{\theta}_{1T}^s$ θ_{1T}^s, and mse_s are the values of $\hat{\theta}_{1T}$ θ_{1T}, and mse, respectively, of the sth simulation. Simulated percentage values of gain in efficiency of a two-stage estimator over the Fay-Herriot (FH) estimator were computed as follows:

$$\mathrm{GE} = \left(\frac{\mathrm{MSE(FH\ estimator)}}{\mathrm{MSE(two\text{-}stage\ estimator)}} - 1 \right) \times 100.$$

7.1. ρ *Known.*

Tables 1 and 2 report the values of GE for the two-stage estimator $\hat{\theta}_{1T}(\rho)$ with $\rho = 0.2$ and 0.4 respectively. We have the following results from Tables 1 and 2:

(1) Substantial gains in efficiency (GE) are achieved when the between-time variation relative to sampling variation is small ($\sigma^2 = 0.25$, 0.5), especially when the between-small-area variation is substantial ($\sigma_v^2 = 1$, 2). For example, GE $= 105\%$ when $T = 10$, $\sigma^2 = 0.25$, $\rho = 0.4$ and $\sigma_v^2 = 1.0$.

(2) GE values increase significantly with T, especially for small σ^2, i.e., the use of data for more time points improves the efficiency of the two-stage estimator. For example, for

TABLE 2: Gain in efficiency of the two-stage estimator $\hat{\theta}_{1T}(\rho)$ over the Fay-Herriot estimator: $\rho = 0.4$.

	Efficiency gain (%)			
σ_v^2	$\sigma^2 = 0.25$	0.5	1.0	2.0
		$T = 5$		
0.25	33	21	14	8
0.5	50	30	17	10
1.0	74	44	24	12
2.0	101	62	33	16
		$T = 10$		
0.25	47	27	15	9
0.5	71	39	20	10
1.0	105	57	28	13
2.0	144	80	40	19
		$T = 40$		
0.25	65	33	17	9
0.5	99	48	23	11
1.0	144	72	34	15
2.0	195	98	47	21

$\rho = 0.4$, $\sigma^2 = 0.25$, $\sigma_v^2 = 1.0$, we have GE $= 74\%$ with $T = 5$ vs. 105% with $T = 10$ and 144% with $T = 40$.

(3) For a fixed σ_v^2, GE decreases as σ^2 increases, whereas it increases with σ_v^2 for a fixed σ^2.

Table 3 and 4 report the values of the relative bias (RB) of the naive estimator of the MSE, $\text{mse}_N[\hat{\theta}_{1T}(\rho)]$, given by (5.11), and the improved estimator of the MSE, $\text{mse}[\hat{\theta}_{1T}(\rho)]$, given by (5.10), for $\rho = 0.2$ and 0.4 respectively. We have the following results from Tables 3 and 4:

(1) The improved estimator of the MSE performs well, leading to slight overestimation for $T = 5$ or 10 (RB $\leq 3\%$ for $T = 5$, and RB $\leq 5\%$ for $T = 10$). It seems to also perform well for large T ($= 40$), although it is derived under the assumption of a small or moderate T and a relatively large m. It leads to slight underestimation for small σ^2.

(2) The naive estimator of the MSE leads to significant underestimation for small σ^2 and small T ($= 5$), but for other values it is quite satisfactory. For example, RB $= -9.1\%$ when $\sigma^2 = 0.25$, $\sigma_v^2 = 0.25$, $\rho = 0.4$, and $T = 5$.

7.2. ρ Unknown.

(a) Method 1

We computed the GE values for the two-stage estimator $\hat{\theta}_{1T}(\rho_0)$ using prior guesses $\rho_0 = 0.1$, 0.3 for true $\rho = 0.2$ and $\rho_0 = 0.2$, 0.3, 0.5, 0.6 for true $\rho = 0.4$. Our result indicate that the GE values are virtually unaffected by the choice of ρ_0, i.e., the two-stage estimator $\hat{\theta}_{1T}(\rho_0)$ retains its efficiency even when the prior guess ρ_0 deviates significantly from ρ.

We have also computed the RB values for $\text{mse}[\hat{\theta}_{1T}(\rho_0)]$, the estimator of the MSE obtained by substituting ρ_0 for ρ in (5.10). Our results in Table 5 and 6 for $T = 5$

TABLE 3: Relative bias of $\text{mse}[\hat{\theta}_{1T}(\rho)]$ and $\text{mse}_N[\hat{\theta}_{1T}(\rho)]$ (in parentheses):

	Relative bias (%)			
σ_v^2	$\sigma^2 = 0.25$	0.5	1.0	2.0
		$T = 5$		
0.25	2.4(−7.2)	1.7(−4.1)	1.8(−1.5)	1.6(−0.1)
0.5	2.5(−5.6)	1.8(−3.3)	1.7(−1.2)	1.4(−0.2)
1.0	2.7(−4.3)	1.9(−2.5)	1.7(−0.8)	1.3(−0.1)
2.0	2.7(−3.6)	1.9(−2.1)	1.7(−0.6)	1.2(−0.0)
		$T = 10$		
0.25	4.5(−0.6)	4.1(1.2)	3.5(2.0)	2.8(2.0)
0.5	4.4(−0.1)	4.0(1.5)	3.5(2.1)	2.8(2.0)
1.0	4.2(0.2)	3.9(1.6)	3.4(2.1)	2.7(2.0)
2.0	4.1(0.3)	3.8(1.7)	3.3(2.1)	2.6(2.0)

TABLE 4: Relative bias of $\text{mse}[\hat{\theta}_{1T}(\rho)]$ and $\text{mse}_N[\hat{\theta}_{1T}(\rho)]$ (in parentheses):

	Relative bias (%)			
σ_v^2	$\sigma^2 = 0.25$	0.5	1.0	2.0
		$T = 5$		
0.25	2.9(−9.1)	1.3(−5.5)	1.6(−1.8)	1.7(−0.1)
0.5	2.8(−7.7)	1.2(−4.9)	1.4(−1.8)	1.4(−0.3)
1.0	2.9(−6.5)	1.3(−4.2)	1.4(−1.5)	1.2(−0.3)
2.0	2.9(−5.7)	1.3(−3.7)	1.4(−1.2)	1.2(−0.2)
		$T = 10$		
0.25	4.3(−2.6)	3.9(0.4)	3.5(1.7)	2.9(2.1)
0.5	4.3(−1.9)	3.9(0.7)	3.4(1.9)	2.8(2.0)
1.0	4.2(−1.6)	3.8(0.9)	3.4(2.0)	2.7(2.0)
2.0	4.1(−1.5)	3.7(1.0)	3.3(2.0)	2.6(2.0)
		$T = 40$		
0.25	−1.1(−2.8)	−0.6(−1.3)	−0.1(−0.5)	0.1(−0.1)
0.5	−1.2(−2.8)	−0.6(−1.3)	−0.1(−0.5)	0.1(−0.1)
1.0	−1.2(−2.8)	−0.6(−1.3)	−0.1(−0.4)	0.0(−0.1)
2.0	−1.2(−2.8)	−0.6(−1.3)	−0.1(−0.4)	0.0(−0.1)

suggest that $\text{mse}[\hat{\theta}_{1T}(\rho_0)]$ performs well when $\rho_0 < \rho$, but it can lead to significant overestimation when ρ_0 is significantly larger than ρ and σ^2 is small. For example, RB $\simeq 10\%$ when $\rho_0 = 0.6$, $\rho = 0.4$ and $\sigma^2 = 0.25$. Our method 1, therefore, appears to be satisfactory provided the prior guess ρ_0 is not significantly larger than the true ρ.

(b) Method 2

We computed the GE values for the two-stage estimator $\hat{\theta}_{1T}(\hat{\rho}_N)$, using the naive estimator $\hat{\rho}_N$ given by (6.1). Again, the GE values are close to those under the true ρ; i.e., the two-stage estimator retains efficiency even when a naive estimator $\hat{\rho}_N$ is used. It may be noted that $\hat{\rho}_N$ leads to significant underestimation of ρ, which increases with ρ.

Tables 7 and 8 report the RB values for $\text{mse}[\hat{\theta}_{1T}(\hat{\rho}_N)]$, the estimator of the MSE obtained by substituting $\hat{\rho}_N$ for ρ in (5.10). Our results in Tables 7 and 8 suggest that $\text{mse}[\hat{\theta}_{1T}(\hat{\rho}_N)]$ performs well for small ρ, but it can lead to underestimation, which

TABLE 5: Relative bias of $\text{mse}[\hat{\theta}_{1T}(\rho_0)]$: $\rho_0 = 0.1, 0.3$, $\rho = 0.2$; $T = 5$.

σ_v^2	Relative bias (%) $\sigma^2 = 0.25$	0.5	1.0	2.0
	$\rho_0 = 0.1$			
0.25	1.0	0.5	0.8	0.8
0.5	1.1	0.5	0.6	0.6
1.0	1.1	0.5	0.5	0.5
2.0	1.1	0.4	0.4	0.3
	$\rho_0 = 0.3$			
0.25	4.6	3.4	3.0	2.5
0.5	4.7	3.5	2.9	2.1
1.0	4.8	3.7	3.0	2.0
2.0	4.9	3.8	3.1	2.1

TABLE 6: Relative bias of $\text{mse}[\hat{\theta}_{1T}(\rho_0)]$: $\rho_0 = 0.2, 0.3, 0.5, 0.6$, $\rho = 0.4$; $T = 5$.

σ_v^2	Relative bias (%) $\sigma^2 = 0.25$	0.5	1.0	2.0
	$\rho_0 = 0.2$			
0.25	−1.4	−2.0	−1.1	−0.4
0.5	−1.5	−2.2	−1.3	−0.5
1.0	−1.5	−2.3	−1.5	−0.7
2.0	−1.5	−2.4	−1.7	−1.0
	$\rho_0 = 0.3$			
0.25	0.3	−0.6	0.2	0.6
0.5	0.3	−0.7	0.0	0.5
1.0	0.4	−0.7	−0.1	0.3
2.0	0.3	−0.7	−0.2	0.1
	$\rho_0 = 0.5$			
0.25	6.4	3.6	3.5	2.9
0.5	6.1	3.4	2.9	2.4
1.0	6.2	3.6	2.9	2.0
2.0	6.2	3.7	3.0	2.0
	$\rho_0 = 0.6$			
0.25	11.2	7.0	5.8	4.0
0.5	10.2	6.8	4.6	3.5
1.0	10.2	5.9	4.1	2.8
2.0	10.3	6.1	4.3	2.6

increases with ρ. (For $T = 5$ and $\rho = 0.7$ we obtained RB $= -10\%$.) Thus, our method 2 may be satisfactory when ρ is expected to be small.

(c) Method 3

Method 3 uses the estimator $\hat{\rho}$ given by (6.2). It takes into account the sampling errors, but often leads to values outside the admissible range, $(-1, 1)$, especially for small T

TABLE 7: Relative bias of mse$[\hat{\theta}_{1T}(\hat{\rho}_N)]$: $\rho = 0.2$.

σ_v^2	Relative bias (%) $\sigma^2 = 0.25$	0.5	1.0	2.0
	$T = 5$			
0.25	−1.0	−0.7	0.3	0.7
0.5	−0.8	−0.7	0.1	0.5
1.0	−0.6	−0.6	0.0	0.4
2.0	−0.4	−0.5	0.0	0.3
	$T = 10$			
0.25	1.4	1.8	2.1	2.0
0.5	1.2	1.6	2.0	1.9
1.0	1.0	1.4	1.8	1.8
2.0	0.8	1.2	1.6	1.7

TABLE 8: Relative bias of mse$[\hat{\theta}_{1T}(\hat{\rho}_N)]$: $\rho = 0.4$.

σ_v^2	Relative bias (%) $\sigma^2 = 0.25$	0.5	1.0	2.0
	$T = 5$			
0.25	−4.0	−3.8	−1.8	−0.3
0.5	−4.0	−3.9	−2.1	−0.5
1.0	−3.9	−4.0	−2.3	−0.7
2.0	−3.7	−4.0	−2.4	−0.9
	$T = 10$			
0.25	−2.3	−1.6	−0.2	1.1
0.5	−2.6	−1.9	−0.4	0.9
1.0	−3.0	−2.3	−0.7	0.8
2.0	−3.2	−2.5	−0.9	0.6

or small σ^2 relative to sampling variation. Similar difficulties are encountered under measurement-error models (see Fuller 1987, Section 2.5). It would be useful to develop suitable modifications of $\hat{\rho}$, using methods similar to those in Fuller (1987), that lead to more efficient estimators taking values in the admissible range, $(-1, 1)$.

8. CONCLUSIONS

In this paper, we have proposed a model involving autocorrelated random effects and sampling error for small-area estimation, using both time-series and cross-sectional data. Under this model, we first obtained a two-stage estimator of a small-area mean for the current period, assuming a known autocorrelation ρ. We then proposed three methods of estimating ρ. An estimator of the MSE of the resulting two-stage estimator is also given.

Our simulation results in Section 7 have shown that the two-stage estimator under methods 1 and 2 can lead to substantial gain in efficiency over the Fay-Herriot estimator which uses only the current cross-sectional data, especially when the between-time variation relative to sampling variation is small. Further, our method 1 of estimating the MSE, based on a prior guess ρ_0 of the true ρ, may be satisfactory, provided ρ_0

is not significantly larger than ρ. Our method 2 of estimating the MSE, based on a naive estimator of ρ that ignores the sampling errors, may also be satisfactory when ρ is expected to be small. We have also noted that our method 3, based on an estimator $\hat{\rho}$ that takes sampling errors into account, runs into difficulties, since $\hat{\rho}$ often leads to values outside admissible range $(-1, 1)$, especially for a small number of time points T or when the between-time variation relative to sampling variation is small. Suitable modifications of $\hat{\rho}$ that can lead to more efficient estimators taking values in the admissible range are needed.

Rao and Yu (1992) proposed a hierarchical Bayes (HB) approach for inference on the small-area means θ_{iT}, under the model (3.4). In the HB approach, a prior distribution on all the model parameters is specified and the posterior distribution of the parameters of interest is then obtained. Inferences are based on the posterior distribution; in particular, θ_{iT} is estimated by its posterior mean, and its precision is measured by its posterior variance. The HB approach is computer-intensive, but the computations can be handled using Gibbs sampling and related Monte Carlo numerical integration methods. It appears that the difficulties associated with the frequentist approach in estimating ρ may be circumvented using the HB approach, but we need to study its robustness to choice of prior distributions on the model parameters, especially on ρ.

Ghosh and Nangia (1993) also considered a HB approach for inference on the small-area means θ_{iT}, but assumed independent sampling errors e_{it}. The results were applied to the estimation of the median income for four-person families for the fifty states in the U.S.A. and the District of Columbia, using a model of the form

$$\theta_{it} = \mathbf{x}_{it}^{\mathsf{T}}\boldsymbol{\beta} + b_t; \qquad b_t = b_{t-1} + \epsilon_t,$$

where $\epsilon_t \overset{\text{iid}}{\sim} \mathsf{N}(0, \sigma^2)$. Small area effects are absent in this model but it leads to closed-form expressions for all the conditional distributions needed in implementing Gibbs sampling, unlike model (3.4); see Rao and Yu (1992) for the relevant conditional distributions under (3.4). Ghosh and Nangia (1993) actually considered the multivariate case $\theta_{it} = (\theta_{it1}, \ldots, \theta_{itk})^{\mathsf{T}}$, a k-vector of parameters associated with unit it.

ACKNOWLEDGEMENT

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada. The authors wish to thank John Eltinge and the referees for several constructive comments.

REFERENCES

Anderson, T.W., and Hsiao, C. (1981). Formulation and estimation of dynamic models using panel data. *J. Econometrics*, 18, 67–82.

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. *J. Amer. Statist. Assoc.*, 80, 28–36.

Binder, D.A., and Dick, J.P. (1989). Modelling and estimation for repeated surveys. *Survey Methology*, 15, 29–45.

Choudhry, G.H., and Rao, J.N.K. (1989). Small area estimation using models that combine times series and cross-sectional data. *Proc. Statistics Canada Symp. on Analysis of Data in Time* (A.C. Singh and P. Whitridge, *eds.*), 67–74.

Datta, G.S., and Ghosh, M. (1991). Bayesian prediction in linear models: Application to small area estimation. *Ann. Statist.*, 19, 1748–1770.

Datta, G.S., Fay, R.E., and Ghosh, M. (1991). Hierarchical and empirical Bayes analysis in small-area estimation. *Proc. Annual Research Conference*, U.S. Bureau of the Census, 63–78.

Drew, J.D., Singh, M.P., and Choudhry, G.H. (1982). Evaluation of small area estimation techniques for the Canadian Labour Force Survey. *Survey Methodology*, 8, 17–47.

Ericksen, E.P. (1974). A regression method for estimating population of local areas. *J. Amer. Statist. Assoc.*, 69, 867–875.

Fay, R.E., and Herriot, R.A. (1979). Estimates of income for small places: An application of James-Stein procedures to census data. *J. Amer. Statist. Assoc.*, 74, 269–277.

Fuller, W.A. (1987). *Measurement Error Model*. Wiley, New York.

Fuller, W.A., and Battese, G.E. (1973). Transformations for estimation of linear models with nested error structures. *J. Amer. Statist. Assoc.*, 68, 626–632.

Ghosh, M., and Lahiri, P. (1987). Robust empirical Bayes estimation of means from stratified samples. *J. Amer. Statist. Assoc.*, 82, 1153–1162.

Ghosh, M., and Nangia, N. (1993). Estimation of median income of four-person families: A Bayesian time series approach. Technical Report 429, Department of Statistics, University of Florida.

Ghosh, M., and Rao, J.N.K. (1994). Small area estimation: An appraisal. *Statist. Sci.*, 9, 55–93.

Gonzales, M.E. (1973). Use and evaluation of synthetic estimates. *Proc. Amer. Statist. Assoc. Soc. Statist. Sec.*, 33–6.

Henderson, C.R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics*, 31, 423–447.

Jones, R.G. (1980). Best linear unbiased estimators for repeated surveys. *J. Roy. Statist. Soc. Ser. B*, 42, 221–226.

Judge, G.G., Griffiths, W.E., Hill, R.C., Lütkepohl, H., and Lee, T. (1985). *The Theory and Practice of Econometrics*. Second Edition. Wiley, New York.

Kackar, R.N., and Harville, D.A. (1984). Approximations for standard errors of estimators of fixed and random effects in mixed linear models. *J. Amer. Statist. Assoc.*, 79, 853–862.

Lee, H. (1990). Estimation of panel correlations for the Canadian Labour Force Survey. *Survey Methodology*, 16, 283–292.

National Center for Health Statistics (1979). *Small Area Estimation: An Empirical Comparison of Conventional and Synthetic Estimators for States*. Vital and Health Statistics. Ser. 2, No. 82. U.S. Government Printing Office, Washington.

Pantula, S.G., and Pollock, K.H. (1985). Nested analysis of variance with autocorrelated errors. *Biometrics*, 41, 909–920.

Pfeffermann, D., and Burck, L. (1990). Robust small area estimation combining time series and cross-sectional data. *Survey Methodology*, 16, 217–237.

Prasad, N.G.N., and Rao, J.N.K. (1986). On the estimation of mean squared error of small area predictors. Technical Report 97, Carleton University, Laboratory for Research in Statistics and Probability, Ottawa.

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

Purcell, N.J., and Kish, L. (1980). Postcensal estimates for local areas (or domains). *Internat. Statist. Rev.*, 48, 3–18.

Rao, J.N.K. (1986). Synthetic estimators, SPREE and best model-based predictors of small area means. In Technical Report 97, Carleton University, Laboratory for Research in Statistics and Probability, Ottawa.

Rao, J.N.K., and Yu, M. (1992). Small area estimation combining time series and cross-sectional data. *Proc. Survey Research Methods Section*. Amer. Statist. Assoc., 1–9.

Särndal, C.E., and Hidiroglou, M.A. (1989). Small domain estimation: A conditional analysis. *J. Amer. Statist. Assoc.*, 84, 266–275.

Scott, A.J., and Smith, T.M.P. (1974). Analysis of repeated surveys using time series methods. *J. Amer. Statist. Assoc.*, 69, 674–678.

Singh, A.C., Mantel, H., and Thomas, B.W. (1991). Time series generalizations of Fay-Herriot estimator for small areas. *Proc. Survey Research Methods Section*. Amer. Statist. Assoc., 455–460.

Tiller, R.B. (1992). Times series modelling of sample survey data from the U.S. Current Population Survey. *J. Official Statist.*, 8, 149–166.

Wolter, K. (1985). *Introduction to Variance Estimation*. Springer-Verlag, New York.

Received 4 August 1993
Revised 28 June 1994
Accepted 9 August 1994

Department of Mathematics and Statistics
Carleton University
Ottawa, Ontario
Canada K1S 5B6