March 8, 1988

A Simple Approach to Inference in Random Coefficient Models

Marcia Gumpertz and Sastry G. Pantula
Department of Statistics
North Carolina State University
Raleigh, NC 27695-8203

Key Words and Phrases: Repeated measures regression; Asymptotic inference; Estimated generalized least squares; Growth curve.

ABSTRACT

Random coefficient regression models have been used to analyze cross-sectional and longitudinal data in economics and growth curve data from biological and agricultural experiments. In the literature several estimators, including the ordinary least squares (OLS) and the estimated generalized least squares (EGLS), have been considered for estimating the parameters of the mean model. Based on the asymptotic properties of the EGLS estimators, test statistics have been proposed for testing linear hypotheses involving the parameters of the mean model. An alternative estimator, the simple mean of the individual regression coefficients, provides estimation and hypothesis testing procedures that are simple to compute and simple to teach. The large sample properties of this simple estimator are shown to be similar to that of the EGLS estimator. The performance of the proposed estimator is compared with that of the existing estimators by Monte Carlo simulation.

1. INTRODUCTION

Frequently in biological, medical, agricultural and clinical studies several measurements are taken on the same experimental unit over time with the objective of fitting a response curve to the data. Such studies are called growth curve, repeated measure or longitudinal studies. In many bio-medical and agricultural experiments the number of experimental units is large and the number of repeated measurements on each unit is small. On the other hand, some economic investigations and meteorological experiments involve a small number of units observed over a long period of time. Several models for analyzing such data exist in the literature; the models usually differ in their covariance structures. See Harville (1977) and Jennrich and Schluchter (1986) for a review of the models and of approaches for estimating parameters. Recently, there seems to be a renewed interest in analyzing repeated measures data using Random Coefficient Regression (RCR) models.

In this article we present a brief description of the RCR models and some of the existing results for these models. We present a simple approach that is not difficult to discuss in a course on linear models. We compare our estimation procedure with two of the existing methods. Section 2 contains the assumptions of the model and a brief review of the literature. In section 3 we present the properties of the simple estimator. A Monte Carlo comparison of the estimators is presented in Section 4. Finally, we conclude with a summary which includes some possible extensions.

2. RCR MODEL

Suppose that the t observations on the ith of n experimental units are described by the model

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{e}_i \quad , \quad i=1,2,\ldots, n \tag{2.1}$$

where $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{it})'$ is a tx1 vector of observations on the response variable, $\mathbf{X}_i$ is a txk matrix of observations on k explanatory variables, $\boldsymbol{\beta}_i$ is a kx1 vector of coefficients unique to the ith experimental unit and $\mathbf{e}_i$ is a tx1 vector of errors. Each experimental unit and its response curve is considered to be selected from a larger population of response curves, thus the regression coefficient vectors $\boldsymbol{\beta}_i$, i=1,2,...,n may be viewed as random drawings from some k-variate population and hence (2.1) is called a RCR model. In this paper we discuss the estimation and testing of such models under the following assumptions: (i) the $\mathbf{e}_i$ vectors are independent multivariate normal variables with mean zero and covariance matrix $\sigma^2\mathbf{I}_t$; (ii) the $\boldsymbol{\beta}_i$ vectors are independent multivariate normal variables with mean $\boldsymbol{\beta}$ and nonsingular covariance matrix $\boldsymbol{\Sigma}_{\beta\beta}$; (iii) the vectors $\boldsymbol{\beta}_i$ and $\mathbf{e}_j$ are independent for all i and j; (iv) the $\mathbf{X}_i$ matrices are fixed and of full row rank for each i; (v) min(n,t) > k and (vi) there exists an $M < \infty$ such that the elements of $t(\mathbf{X}_i'\mathbf{X}_i)^{-1}$ are less than M in absolute value for all i and t. The assumption (vi) is not very restrictive. It is satisfied for the models that include polynomials in time and stationary exogeneous variables.

Several authors, including Rao (1965), Swamy (1971), Hsiao (1975), Harville (1977), Laird and Ware (1982), Jennrich and Schluchter (1986) and Carter and Yang (1986) have considered the estimation and testing for the RCR models. We summarize the results of Carter and Yang (1986) since they consider the large sample distribution of the estimated generalized least squares (EGLS) estimator as n and/or t tend to infinity. For the sake of simplicity, we have assumed that equal number of repeated measurements are taken on all experimental units and that the variance of the error vector $\mathbf{e}_i$ does not depend on i. However, similar results exist for more general cases and will be discussed in the summary.

Consider the least squares estimators

$$\mathbf{b}_i = (\mathbf{X}_i'\mathbf{X}_i)^{-1}\mathbf{X}_i'\mathbf{y}_i \quad , \quad i=1,2,\ldots, n \tag{2.2}$$

of $\boldsymbol{\beta}$ computed for each individual experimental unit. Note that the $\mathbf{b}_i$'s are independent and normally distributed with mean $\boldsymbol{\beta}$ and variance $\mathbf{W}_i^{-1} = \boldsymbol{\Sigma}_{\beta\beta} + \sigma^2(\mathbf{X}_i'\mathbf{X}_i)^{-1}$. Therefore, the best (linear) unbiased estimator of $\boldsymbol{\beta}$ is the generalized least squares (GLS) estimator. Swamy (1971) showed that

$$\hat{\boldsymbol{\beta}}_{GLS} = (\sum_{i=1}^{n} \mathbf{W}_i)^{-1}(\sum_{i=1}^{n} \mathbf{W}_i\mathbf{b}_i) \quad , \tag{2.3}$$

that is, the GLS estimator is the "weighted" least squares (average) estimator of $\mathbf{b}_i$ where the weights are the inverse variance-covariance matrices of $\mathbf{b}_i$. Under the normality assumption, $\hat{\boldsymbol{\beta}}_{GLS}$ is also the maximum likelihood estimator of $\boldsymbol{\beta}$ (provided $\boldsymbol{\Sigma}_{\beta\beta}$ and σ^2 are known). The elements of $\boldsymbol{\Sigma}_{\beta\beta}$ and σ^2 are seldom known and hence we consider the estimated GLS (EGLS) estimator

$$\hat{\boldsymbol{\beta}}_{EGLS} = (\sum_{i=1}^{n} \hat{\mathbf{W}}_i)^{-1}(\sum_{i=1}^{n} \hat{\mathbf{W}}_i\mathbf{b}_i) \quad , \tag{2.4}$$

where

$$\hat{\mathbf{W}}_i^{-1} = \hat{\mathbf{\Sigma}}_{\beta\beta} + \hat{\sigma}^2(\mathbf{X}_i'\mathbf{X}_i)^{-1} \quad ,$$

$$\hat{\mathbf{\Sigma}}_{\beta\beta} = \mathbf{S}_{bb} - n^{-1}\hat{\sigma}^2 \sum_{i=1}^{n} (\mathbf{X}_i'\mathbf{X}_i)^{-1} \quad ,$$

$$\mathbf{S}_{bb} = (n-1)^{-1} \sum_{i=1}^{n} (\mathbf{b}_i - \bar{\mathbf{b}})(\mathbf{b}_i - \bar{\mathbf{b}})' \quad ,$$

$$\hat{\sigma}^2 = [n(t-k)]^{-1} \sum_{i=1}^{n} [\mathbf{y}_i'\mathbf{y}_i - \mathbf{b}_i'\mathbf{X}_i'\mathbf{y}_i] \quad ,$$

and

$$\bar{\mathbf{b}} = n^{-1} \sum_{i=1}^{n} \mathbf{b}_i \quad .$$

Carter and Yang (1986) suggested inference procedures based on the large sample distribution of the estimator $\hat{\boldsymbol{\beta}}_{EGLS}$. Their results are summarized below. (They suggested a slightly different estimator of $\mathbf{\Sigma}_{\beta\beta}$ in the case $\hat{\mathbf{\Sigma}}_{\beta\beta}$ is not nonnegative definite.)

Result 2.1: Consider the model given in (2.1) with the assumptions (i) through (vi). Consider the statistic

$$T^2 = [\mathbf{L}\,\hat{\boldsymbol{\beta}}_{EGLS} - \boldsymbol{\lambda}_0]'[\mathbf{L}(\sum_{i=1}^{n} \hat{\mathbf{W}}_i)^{-1}\mathbf{L}']^{-1}[\mathbf{L}\,\hat{\boldsymbol{\beta}}_{EGLS} - \boldsymbol{\lambda}_0] \quad , \qquad (2.5)$$

for testing H_0: $\mathbf{L}\boldsymbol{\beta} = \boldsymbol{\lambda}_0$, where $\mathbf{L}$ is a qxk matrix of q linearly independent rows. Then ,

(a) for a fixed n and t tending to infinity:

$(n-q)q^{-1}(n-1)^{-1}T^2$ is (asymptotically) distributed as $F(q,n-q)$,

(b) for a fixed t and n tending to infinity:

T^2 is (asymptotically) distributed as chi-square with q degrees of freedom,

and

(c) for the case where nt is large and q=1: T^2 is approximately distributed as $F(1,\nu)$ where

$$\nu = f^{-1}\{\boldsymbol{\ell}'\boldsymbol{\Sigma}_{\beta\beta}\boldsymbol{\ell} + t^{-1}\sigma^2\boldsymbol{\ell}'\mathbf{C}_2\boldsymbol{\ell}\}^2 ,$$

$$f = \{(n-1)^{-1}(\boldsymbol{\ell}'\boldsymbol{\Sigma}_{\beta\beta}\boldsymbol{\ell})^2 + [t^2(nt-nk)]^{-1}\sigma^4(\boldsymbol{\ell}'\mathbf{C}_2\boldsymbol{\ell})^2\}$$

$$\boldsymbol{\ell} = \mathbf{L}' ,$$

and

$$\sigma^2\mathbf{C}_2 = n\ t\left[\sum_{i=1}^{n} \mathbf{W}_i\right]^{-1} - t\ \boldsymbol{\Sigma}_{\beta\beta} .$$

Proof: See Carter and Yang (1986). []

Carter and Yang (1986) proved part (b) of the above result by observing that the distribution of $\hat{\boldsymbol{\beta}}_{EGLS}$ is asymptotically (as $n \to \infty$) equivalent to that of $\hat{\boldsymbol{\beta}}_{GLS}$. To prove part (a), they observed that the distribution of $\hat{\boldsymbol{\beta}}_{EGLS}$ is asymptotically (as $t \to \infty$) equivalent to that of

$$\mathbf{m}_\beta = n^{-1} \sum_{i=1}^{n} \boldsymbol{\beta}_i \sim N(\boldsymbol{\beta},\ n^{-1}\ \boldsymbol{\Sigma}_{\beta\beta}) , \tag{2.6}$$

which is also asymptotically equivalent to $\hat{\beta}_{GLS}$ as $t \to \infty$. Finally, when nt is large, Satterthwaite's approximation was used to approximate the distribution of T^2. In the next section we present inference procedures based on the large sample distribution of the simple estimator $\bar{\mathbf{b}}$.

3. A SIMPLE APPROACH

It is well known that the GLS estimator $\hat{\boldsymbol{\beta}}_{GLS}$ is the best (linear) unbiased estimator of $\boldsymbol{\beta}$ and that (under some regularity conditions) the EGLS estimator $\hat{\boldsymbol{\beta}}_{EGLS}$ is asymptotically (as $n \to \infty$) equivalent to the GLS estimator. However, in small samples, the distribution of $\hat{\boldsymbol{\beta}}_{EGLS}$ may be far from being normal. It is also argued that the estimator $\hat{\boldsymbol{\beta}}_{EGLS}$ may even be worse than the ordinary least squares (OLS) estimator,

$$\hat{\boldsymbol{\beta}}_{OLS} = \left(\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i\right)^{-1}\left(\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i\mathbf{b}_i\right) \tag{3.1}$$

because $\hat{\beta}_{EGLS}$ depends on the estimated variance-covariance matrix which may introduce additional variability. It is easy to see that the OLS estimator, $\hat{\boldsymbol{\beta}}_{OLS}$, is normally distributed with mean $\boldsymbol{\beta}$ and variance $(\Sigma_{i=1}^{n}\mathbf{X}_i'\mathbf{X}_i)^{-1}\left[\Sigma_{i=1}^{n}\mathbf{X}_i'\mathbf{X}_i\ \boldsymbol{\Sigma}_{\beta\beta}\mathbf{X}_i'\mathbf{X}_i + \sigma^2\ \Sigma_{i=1}^{n}\mathbf{X}_i'\mathbf{X}_i\right](\Sigma_{i=1}^{n}\mathbf{X}_i'\mathbf{X}_i)^{-1}$.
Thus, to compute either the EGLS estimate $\hat{\boldsymbol{\beta}}_{EGLS}$ or to compute the variance covariance matrices of $\hat{\boldsymbol{\beta}}_{EGLS}$ and $\hat{\boldsymbol{\beta}}_{OLS}$, it is necessary to estimate the elements of $\boldsymbol{\Sigma}_{\beta\beta}$ and σ^2. We now present the properties of the simple estimator $\bar{\mathbf{b}}$, which does not require the estimation of $\boldsymbol{\Sigma}_{\beta\beta}$ and σ^2.

Note that the GLS, EGLS and OLS estimators are weighted averages of the individual least squares estimators $\mathbf{b}_i$. The estimator

$$\bar{\mathbf{b}} = n^{-1}\sum_{i=1}^{n}\mathbf{b}_i \tag{3.2}$$

is the simple average of the individual least squares estimators. In the special case where the model matrix $\mathbf{X}_i$ is the same (=A say) for all individuals, then the GLS, EGLS and OLS estimates coincide with the estimator $\bar{\mathbf{b}}$. The estimator $\bar{\mathbf{b}}$ is normally distributed with mean $\boldsymbol{\beta}$ and variance

$$\mathrm{Var}(\bar{\mathbf{b}}) = n^{-1}\,\boldsymbol{\Sigma}_{\beta\beta} + \sigma^2 n^{-2}\sum_{i=1}^{n}(\mathbf{X}_i'\mathbf{X}_i)^{-1}\ . \tag{3.3}$$

Note that

$$
\begin{aligned}
E[\mathbf{S}_{bb}] &= E[(n-1)^{-1} \sum_{i=1}^{n} (\mathbf{b}_i - \bar{\mathbf{b}})(\mathbf{b}_i - \bar{\mathbf{b}})'] \\
&= (n-1)^{-1} E\left[\sum_{i=1}^{n} (\mathbf{b}_i - \boldsymbol{\beta})(\mathbf{b}_i - \boldsymbol{\beta})' - n(\bar{\mathbf{b}} - \boldsymbol{\beta})(\bar{\mathbf{b}} - \boldsymbol{\beta})' \right] \\
&= (n-1)^{-1} \left[\sum_{i=1}^{n} \mathrm{var}(\mathbf{b}_i) - n\ \mathrm{var}(\bar{\mathbf{b}}) \right] . \\
&= (n-1)^{-1} \left[n^2\ \mathrm{var}(\bar{\mathbf{b}}) - n\ \mathrm{var}(\bar{\mathbf{b}}) \right] \\
&= n\ \mathrm{var}(\bar{\mathbf{b}}) \ .
\end{aligned}
$$

Therefore, a simple unbiased estimator for $\mathrm{var}(\bar{\mathbf{b}})$ is $n^{-1}\mathbf{S}_{bb}$.

That is, the sample variance (covariance matrix) divided by n is an unbiased estimator for the variance of the sample mean even though the variances (of $\mathbf{b}_i$) are not homogeneous.

Consider the statistic

$$\hat{T}^{*2} = [\mathbf{L}\bar{\mathbf{b}} - \boldsymbol{\lambda}_o]'[\mathbf{L}\ n^{-1}\mathbf{S}_{bb}\mathbf{L}']^{-1}[\mathbf{L}\bar{\mathbf{b}} - \boldsymbol{\lambda}_o] \ , \tag{3.4}$$

for testing H_0: $\mathbf{L}\boldsymbol{\beta} = \boldsymbol{\lambda}_o$, where $\mathbf{L}$ is a qxk matrix of linearly independent rows. Notice that $\hat{T}^{*2}$ is the Hotelling's T^2 statistic one would compute if the variances of the $\mathbf{b}_i$'s were equal (i.e., if the X_i's were the same for all individuals).

Before we establish that the statistic $\hat{T}^{*2}$ has similar asymptotic properties as that of the statistic T^2, we will make a few remarks.

Remark 3.1: Recall that the estimators $\mathbf{b}_i$ are independently and normally distributed with mean $\boldsymbol{\beta}$ and variance $\mathbf{W}_i^{-1} = \boldsymbol{\Sigma}_{\beta\beta} + \sigma^2(\mathbf{X}_i'\mathbf{X}_i)^{-1}$. Under the assumption (vi), the elements of the matrices $t(\mathbf{X}_i'\mathbf{X}_i)^{-1}$ are uniformly (over i) bounded. Therefore, the matrices $(\mathbf{X}_i'\mathbf{X}_i)^{-1}$, i=1,...,n, converge uniformly (over i) to zero as t tends to infinity. Also, note that

$$\mathbf{b}_i = \boldsymbol{\beta}_i + \mathbf{Z}_i \tag{3.5}$$

where

$$\mathbf{Z}_i = (\mathbf{X}_i'\mathbf{X}_i)^{-1}\mathbf{X}_i'\mathbf{e}_i \sim \text{NID}(\mathbf{0},\sigma^2(\mathbf{X}_i'\mathbf{X}_i)^{-1}) , \qquad i=1,\ldots,n .$$

Since $\text{var}(\mathbf{Z}_i) = \sigma^2(\mathbf{X}_i'\mathbf{X}_i)^{-1}$ converges to zero (uniformly in i) as t tends to infinity, the difference between $\mathbf{b}_i$ and $\boldsymbol{\beta}_i$ tends to zero in probability. Therefore, for <u>n fixed and t tending to infinity</u>, $\bar{\mathbf{b}} = n^{-1}\Sigma_{i=1}^n\mathbf{b}_i$ and $\mathbf{m}_\beta = n^{-1}\Sigma_{i=1}^n\boldsymbol{\beta}_i$ are asymptotically equivalent. In fact, since

$$\text{var}(\bar{\mathbf{Z}}) = \sigma^2 n^{-2} \sum_{i=1}^{n} (\mathbf{X}_i'\mathbf{X}_i)^{-1} \le \sigma^2 n^{-1} t^{-1} M\, \mathbf{J} = O(n^{-1}t^{-1}),$$

and

$$\bar{\mathbf{Z}} = O_p(n^{-1/2}t^{-1/2}) ,$$

where $\mathbf{J}$ is a matrix with all elements equal to 1, we have

$$n^{1/2}(\bar{\mathbf{b}} - \boldsymbol{\beta}) = n^{1/2}(\mathbf{m}_\beta - \boldsymbol{\beta}) + n^{1/2}\,\bar{\mathbf{Z}}$$

$$= n^{1/2}\,(\mathbf{m}_\beta - \boldsymbol{\beta}) + O_p(t^{-1/2}) . \tag{3.6}$$

Hence $\bar{\mathbf{b}}$ is also asymptotically (as $t \to \infty$) equivalent to $\hat{\boldsymbol{\beta}}_{GLS}$ and $\hat{\boldsymbol{\beta}}_{EGLS}$. (See also Hsiao (1975) for similar comments.)

It is important to note here that the OLS estimator $\hat{\boldsymbol{\beta}}_{OLS}$, however, is <u>not</u> necessarily asymptotically equivalent to $\bar{\mathbf{b}}$. For example, suppose $\mathbf{X}_i'\mathbf{X}_i = it\mathbf{B}$ where $\mathbf{B}$ is a fixed kxk positive definite matrix. Then the assumption (vi) is satisfied. In this example, the OLS estimator $\hat{\boldsymbol{\beta}}_{OLS}$ is $(\Sigma_{i=1}^n i)^{-1}\ \Sigma_{i=1}^n i\ \mathbf{b}_i$ and hence $\hat{\boldsymbol{\beta}}_{OLS}$ is <u>not</u> asymptotically equivalent to $\bar{\mathbf{b}} = n^{-1}\ \Sigma_{i=1}^n\mathbf{b}_i$.

Remark 3.2: For a fixed t and n tending to infinity, the estimator $\bar{\mathbf{b}}$ may not be asymptotically equivalent to $\hat{\boldsymbol{\beta}}_{GLS}$ and hence may not be an efficient estimator. However, we know that the exact distribution of $\bar{\mathbf{b}}$ is normal and hence the (exact) distribution of

$$T^{*2} = (\bar{\mathbf{b}} - \boldsymbol{\beta})'\ \mathbf{L}'[\mathbf{L}\ \mathbf{var}(\bar{\mathbf{b}})\mathbf{L}']^{-1}\mathbf{L}(\bar{\mathbf{b}} - \boldsymbol{\beta})$$

is chi-square with q degrees of freedom, where $\mathbf{L}$ is a qxk matrix of rank q.

We now present the asymptotic distribution of the $\hat{T}^{*2}$ statistic as n and/or t tends to infinity.

Result 3.1: Consider the model given in (2.1) with the assumptions (i) through (vi). Consider the test statistic $\hat{T}^{*2}$ defined in (3.4) based on the estimator $\bar{\mathbf{b}}$. Then,

(a) for a fixed n and t tending to infinity:

$$(n-q)q^{-1}(n-1)^{-1}\ \hat{T}^{*2}$$ is (asymptotically) distributed as $F(q,n-q)$,

(b) for a fixed t and n tending to infinity: $\hat{T}^{*2}$ is (asymptotically) distributed as chi-square with q degrees of freedom,

and

(c) for the case where nt is large and q=1: $\hat{T}^{*2}$ is approximately distributed as $F(1,\nu^{*})$ where

$$\nu^{*} = g^{-1}[\boldsymbol{\ell}'\boldsymbol{\Sigma}_{\beta\beta}\boldsymbol{\ell} + n^{-1}\sigma^{2}\sum_{i=1}^{n} \boldsymbol{\ell}'(\mathbf{X}_i'\mathbf{X}_i)^{-1}\boldsymbol{\ell}]^{2}\ ,$$

$$g = (n-1)^{-1}[\boldsymbol{\ell}'\boldsymbol{\Sigma}_{\beta\beta}\boldsymbol{\ell}]^{2} + n^{-3}\sigma^{4}(t-k)^{-1}\sum_{i=1}^{n} [\boldsymbol{\ell}'(\mathbf{X}_i'\mathbf{X}_i)^{-1}\boldsymbol{\ell}]^{2}$$

and

$$\mathbf{L} = \boldsymbol{\ell}'\ .$$

Proof: See Appendix. []

With the exception of ν^*, the Satterthwaite's approximation for the degrees of freedom, the asymptotic distributions of T^2 and $\hat{T}^{*2}$ are identical. The advantage of $\hat{T}^{*2}$ over T^2 is that it is simple to compute and is simple to explain. Note that, as in the case of T^2, the degrees of freedom ν^* (a) tends to (n-1) as t tends to infinity and (b) tends to infinity as n tends to infinity. Also, the degrees of freedom ν^* is always greater than or equal to (n-1) and hence the approximation in (c) serves as a compromise between the F and chi-square approximations.

To summarize, we have seen that asymptotically (as $t \to \infty$), the estimators $\hat{\boldsymbol{\beta}}_{GLS}$, $\hat{\boldsymbol{\beta}}_{EGLS}$ and $\bar{\mathbf{b}}$ are equivalent and are efficient. Also, asymptotically (as $n \to \infty$), the estimators $\hat{\boldsymbol{\beta}}_{EGLS}$ and $\hat{\boldsymbol{\beta}}_{GLS}$ are equivalent and are efficient. However, for a fixed t and n large $\bar{\mathbf{b}}$ may not be as efficient as $\hat{\boldsymbol{\beta}}_{GLS}$ and hence the tests based on $\bar{\mathbf{b}}$ may not be as powerful as the tests based on $\hat{\boldsymbol{\beta}}_{GLS}$. The distribution of $\bar{\mathbf{b}}$ is exactly normal for all n and t, whereas the exact distribution of $\hat{\boldsymbol{\beta}}_{EGLS}$ is unknown. A small Monte Carlo study was conducted to compare the performance of the test statistics based on $\bar{\mathbf{b}}$ and $\hat{\boldsymbol{\beta}}_{EGLS}$. (In the study, the test statistics based on $\hat{\boldsymbol{\beta}}_{OLS}$ were also included.) The results of the study are summarized in the next section.

4. MONTE CARLO SIMULATION

Consider the model

$$y_{ij} = \beta_{0i} + \beta_{1i}x_{ij} + e_{ij}\ , \quad \begin{matrix} i=1,\ldots,n \\ j=1,\ldots,t \end{matrix}$$

where $\boldsymbol{\beta}_i = (\beta_{0i}, \beta_{1i})'$ are NID($\mathbf{0},\boldsymbol{\Sigma}_{\beta\beta}$); x_{ij}'s are independent N(0,9) random variables if i is even and N(0,4) if i is odd; e_{ij}'s are NID(0,4); $\{\boldsymbol{\beta}_i\}$, $\{x_{ij}\}$ and $\{e_{ij}\}$ are independent and

$$\Sigma_{\beta\beta} = \begin{bmatrix} 4 & 4 \\ 4 & 8 \end{bmatrix} .$$

The values for n and t are taken to be 5, 10 and 50 to represent small, moderate and large samples. A set of 250 x_{ij} values were generated once for all and the same values of x_{ij}, i=1,...,n; j=1,...,t, were used in all of the replications. For each pair of values of n and t, 100 Monte Carlo replications were used. In each replication, independent $\boldsymbol{\beta}_i$'s and e_{ij}'s were generated. Test statistics (T^2, $\hat{T}^{*2}$ and T^2_{OLS}) based on $\hat{\boldsymbol{\beta}}_{EGLS}$, $\bar{\mathbf{b}}$ and $\hat{\boldsymbol{\beta}}_{OLS}$ for testing the hypotheses (i) H_0: $\beta_1 = 0$, (ii) H_0: $\beta_0 = \beta_1 = 0$; (iii) H_0: $\beta_1 = 1$ and (iv) H_0: $\beta_0 = \beta_1 = 1$ were computed. The number of times the test statistics rejected the hypotheses are summarized in Tables 1 and 2.

From the asymptotic results in section 3 we would expect that the EGLS estimator $\hat{\boldsymbol{\beta}}_{EGLS}$ and the simple estimator $\bar{\mathbf{b}}$ perform equally well when t is large. However, we do not expect the ordinary least squares estimator to do as well as $\bar{\mathbf{b}}$ when t is large. For t=50 this expectation was borne out. At all values of n the probability of rejecting a true hypothesis (using the F-approximation) was 9% or less for all three statistics, but the power for rejecting either of the false hypotheses was always greater for $\hat{\boldsymbol{\beta}}_{EGLS}$ and $\bar{\mathbf{b}}$ than for $\hat{\boldsymbol{\beta}}_{OLS}$. Furthermore the rejection rates for $\hat{\boldsymbol{\beta}}_{EGLS}$ and $\bar{\mathbf{b}}$ were identical. A look at the true variances of $\hat{\boldsymbol{\beta}}_{GLS}$, $\bar{\mathbf{b}}$ and $\hat{\boldsymbol{\beta}}_{OLS}$ revealed that the relative efficiency of $\bar{\mathbf{b}}$ was almost 100% for both the intercept and the slope parameters, whereas for $\hat{\boldsymbol{\beta}}_{OLS}$ it was only 67% for the intercept parameter and 89% for the slope parameter in the case when n=5 and t=50. For smaller t, the efficiency of $\hat{\boldsymbol{\beta}}_{OLS}$ was even worse. However, the efficiency of $\bar{\mathbf{b}}$ was always close to 100% . Similar values for the relative efficiencies of the estimators were observed when n=5 and n=50.

Table 1. Comparison of the Levels of Test Criteria: The Number of Times a 0.05 Level Test Criterion Rejects the Hypothesis (out of 100 replications).

			(i) H_0: $\beta_1 = 0$			(ii) H_0: $\beta_0 = \beta_1 = 0$	
n	t	Estimator	$F_{1,n-1}$	χ^2_1	$F_{1,\nu*}$	$F_{2,n-2}$	χ^2_2
5	5	EGLS	8	15	11	4	26
		BBAR	9	16	10	10	32
		OLS	6	14		0	19
	10	EGLS	9	18	9	2	24
		BBAR	9	18	9	11	30
		OLS	8	17		1	14
	50	EGLS	5	12	5	8	29
		BBAR	5	12	5	9	29
		OLS	5	12		5	18
10	5	EGLS	3	5	3	3	10
		BBAR	3	5	3	2	11
		OLS	2	5		2	6
	10	EGLS	7	9	7	9	16
		BBAR	7	10	7	10	16
		OLS	6	8		3	6
	50	EGLS	4	5	4	5	13
		BBAR	4	5	4	5	13
		OLS	3	3		2	5
50	5	EGLS	8	9	8	6	10
		BBAR	7	8	7	8	9
		OLS	7	7		8	9
	10	EGLS	1	2	1	0	2
		BBAR	1	2	1	1	2
		OLS	2	3		1	1
	50	EGLS	4	4	4	2	2
		BBAR	4	4	4	2	2
		OLS	4	4		1	4

Table 2. Comparison of the Powers of Test Criteria: The Number of Times a 0.05 Level Test Criterion Rejects the Hypothesis (out of 100 replications).

			(i) $H_0: \beta_1 = 1$			(ii) $H_0: \beta_0 = \beta_1 = 1$	
n	t	Estimator	$F_{1,n-1}$	χ^2_1	$F_{1,\nu*}$	$F_{2,n-2}$	χ^2_2
5	5	EGLS	13	24	13	5	34
		BBAR	13	25	15	10	43
		OLS	13	23		3	22
	10	EGLS	12	18	12	8	34
		BBAR	12	18	13	16	41
		OLS	12	23		4	27
	50	EGLS	13	27	13	11	48
		BBAR	13	27	13	11	49
		OLS	11	26		7	31
10	5	EGLS	13	17	14	15	26
		BBAR	12	16	13	17	28
		OLS	10	15		5	16
	10	EGLS	18	27	18	19	33
		BBAR	19	27	19	19	34
		OLS	16	23		6	21
	50	EGLS	15	20	15	18	35
		BBAR	15	20	15	18	35
		OLS	14	20		9	24
50	5	EGLS	69	70	70	79	82
		BBAR	67	69	67	80	84
		OLS	51	54		52	57
	10	EGLS	70	70	70	86	87
		BBAR	70	71	70	85	88
		OLS	56	60		51	53
	50	EGLS	67	69	67	83	88
		BBAR	67	69	67	83	88
		OLS	56	58		72	76

As n approaches infinity for fixed t we would expect $\hat{\boldsymbol{\beta}}_{EGLS}$ to be more powerful than $\bar{\mathbf{b}}$. As it turned out, for n=50 the rejection rates for $\hat{\boldsymbol{\beta}}_{EGLS}$ and $\bar{\mathbf{b}}$ (using the χ^2 approximation) were nearly indistinguishable. The rejection rate for $\hat{\boldsymbol{\beta}}_{OLS}$ ranged from 14 to 39 percent lower than that of the other two estimators.

For small sample sizes none of the estimators was very powerful. However, contrary to our expectation, the performance of $\hat{\boldsymbol{\beta}}_{EGLS}$ was reasonable. The estimator $\bar{\mathbf{b}}$ may have been more powerful than $\hat{\boldsymbol{\beta}}_{EGLS}$ in rejecting H_0: $\beta_0 = \beta_1 = 1$, but by the same token, $\bar{\mathbf{b}}$ rejected the true hypotheses, H_0: $\beta_0 = \beta_1 = 0$, more often than $\hat{\boldsymbol{\beta}}_{EGLS}$. One problem that other authors (e.g., Jennrich & Schluchter (1986), Carter and Yang (1986)) have noted is that, with small sample sizes, $\hat{\boldsymbol{\Sigma}}_{\beta\beta}$ is often not a positive definite matrix. In our simulation this occurred 34% of the time for n=t=5, but for moderate sample sizes, (n=t=10) this was no longer a problem. (If $\hat{\boldsymbol{\Sigma}}_{\beta\beta}$ is not positive definite, the modified estimator suggested by Carter and Yang (1986) was used.)

In our simulation even though the $\mathbf{X}_i$ matrices were different for different individuals, the weight matrices $\mathbf{W}_i$ turned out to be close to $n^{-1}\mathbf{I}$. This may be one of the reasons why the tests based on $\bar{\mathbf{b}}$ and $\hat{\boldsymbol{\beta}}_{EGLS}$ had very similar power for all sample sizes.

5. SUMMARY

In random coefficient regression models several estimators for $\boldsymbol{\beta}$ exist in the literature. Carter and Yang (1986) derived the asymptotic distribution of the estimated generalized least squares estimator as either n, the number of experimental units, tends to infinity and/or as t, the number of repeated measurements on each unit, tends to infinity. They proposed test statistics based on the EGLS estimator. The simple average $\bar{\mathbf{b}} = n^{-1} \Sigma_{i=1}^{n} \mathbf{b}_i$ of the regression estimates from each unit has not received much attention in the literature. The main contribution of the paper is to show that inferences can be made, without much difficulty, using the simple estimator $\bar{\mathbf{b}}$. Asymptotic results for the estimator $\bar{\mathbf{b}}$, similar to those derived by Carter and Yang (1986) for $\hat{\boldsymbol{\beta}}_{EGLS}$, are derived. Also, the results of a small Monte Carlo study indicate that it is reasonable to use $\bar{\mathbf{b}}$ for inferences on $\boldsymbol{\beta}$.

It is important to emphasize the simplicity of the estimator $\bar{\mathbf{b}}$, the test statistics based on $\bar{\mathbf{b}}$ and their asymptotic properties. The estimator $\hat{\boldsymbol{\beta}}_{EGLS}$ is not as simple to compute. Also, the estimator $\hat{\boldsymbol{\Sigma}}_{\beta\beta}$ that enters the computation of $\hat{\boldsymbol{\beta}}_{EGLS}$ may need to be adjusted so that $\hat{\boldsymbol{\Sigma}}_{\beta\beta}$ is positive definite. We are, however, not suggesting that $\hat{\boldsymbol{\beta}}_{EGLS}$ be ignored. The estimator $\hat{\boldsymbol{\beta}}_{EGLS}$ may perform very well for several model matrices (especially when n is large).

Our results extend to the case where unequal number (r_i, say) of measurements are made on different individuals. In this case, part (a) of Result 3.1 should be modified to say "for a fixed n and minimum (r_i) tending to infinity." Also, when minimum (r_i) is large, the Result 3.1 (a) holds even if σ_i^2 = variance (e_{ij}) is not the same for different experimental units (provided one uses s_i^2, regression mean square error for the regression of i<u>th</u> individual, to estimate σ_i^2). When n is large, Result 3.1 (b) holds even if $\sigma_i^2 \neq \sigma^2$ for all i, provided we assume that for all i, $\sigma_i^2 \leq \sigma^2$ for some finite σ^2. Our results can also be extended to the case where the errors e_{ij} are correlated over time. For example, suppose for each i, $\{e_{ij}: j=1,\ldots,t\}$ is a stationary time series with variance covariance matrix of $\mathbf{e}_i$ given by $\boldsymbol{\Sigma}_{ee}$. It is easy to see that $n^{-1}\mathbf{S}_{bb}$ is still an unbiased estimator of var($\bar{\mathbf{b}}$). Under some regularity conditions (similar to those given in Section 9.1 of Fuller (1976)) on $\mathbf{X}_i$, $\mathbf{\Sigma}_{\beta\beta}$ and $\mathbf{\Sigma}_{ee}$ one can obtain the asymptotic results for the test statistic based on $\bar{\mathbf{b}}$ and $\mathbf{S}_{bb}$. The proofs, however, are not included for the sake of brevity.

ACKNOWLEDGEMENTS

The work of S. G. Pantula was partially supported by the National Science Foundation.

REFERENCES

Carter, R. L. and Yang, M. C. K. (1986). "Large Sample Inference in Random Coefficient Regression Models," *Communications in Statistics - Theory and Methods*, **15**(8), 2507-2525.

Fuller, W. A. (1976). *Introduction to Statistical Time Series*, New York: John Wiley and Sons.

Harville, D. A. (1977). "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems," *Journal of the American Statistical Association*, **72**, 320-340.

Hsiao, C. (1975). "Some Estimation Methods for a Random Coefficient Model," *Econometrica*, **43**, 305-325.

Jennrich, R. I. and Schluchter, M. D. (1986). "Unbalanced Repeated-Measures Models with Structured Covariance Matrices," *Biometrics*, **42**, 805-820.

Laird, N. M. and Ware, J. H. (1982). "Random-Effects Models for Longitudinal Data," *Biometrics*, **38**, 963-974.

Rao, C. R. (1965). "The Theory of Least Squares When the Parameters are Stochastic and its Applications to the Analysis of Growth Curves," *Biometrika*, **52**, 447-458.

Swamy, P. A. V. B. (1971). *Statistical Inference in Random Coefficient Regression Models*. Berlin: Springer-Verlag.

APPENDIX

In the appendix, we outline the proof of Result 3.1.

(a) n fixed and t tends to infinity:

From Remark 3.1, we know that

$$n^{1/2}(\bar{\mathbf{b}} - \boldsymbol{\beta}) = n^{1/2}(\mathbf{m}_\beta - \boldsymbol{\beta}) + O_p(t^{-1/2}) , \tag{A.1}$$

and hence the statistic $\hat{T}^{*2} = n(\mathbf{m}_\beta - \boldsymbol{\beta})\mathbf{L}'[\mathbf{L}\mathbf{S}_{bb}\mathbf{L}']^{-1}\mathbf{L}(\mathbf{m}_\beta - \boldsymbol{\beta}) + O_p(t^{-1/2})$.

Also, recall,

$$\begin{aligned} \mathbf{S}_{bb} &= (n-1)^{-1} \sum_{i=1}^{n} (\mathbf{b}_i - \bar{\mathbf{b}})(\mathbf{b}_i - \bar{\mathbf{b}})' \\ &= (n-1)^{-1} \sum_{i=1}^{n} [\boldsymbol{\beta}_i + \mathbf{Z}_i - \mathbf{m}_\beta - \bar{\mathbf{Z}}][\boldsymbol{\beta}_i + \mathbf{Z}_i - \mathbf{m}_\beta - \bar{\mathbf{Z}}]' \\ &= \mathbf{S}_{\beta\beta} + \mathbf{S}_{ZZ} + \mathbf{S}_{\beta Z} + \mathbf{S}'_{\beta Z} , \end{aligned} \tag{A.2}$$

where

$$\mathbf{S}_{cd} = (n-1)^{-1} \sum_{i=1}^{n} (\mathbf{c}_i - \bar{\mathbf{c}})(\mathbf{d}_i - \bar{\mathbf{d}})',$$

and $\mathbf{Z}_i$ is defined in Remark 3.1. Since $\boldsymbol{\beta}_i$ and $\mathbf{Z}_i$ are independent normal random variables with means $\boldsymbol{\beta}$ and $\mathbf{0}$ respectively,

$$E[\mathbf{S}_{\beta Z}] = \mathbf{0}$$

and the variance of the (ℓ,m)th element of $\mathbf{S}_{\beta Z}$ is

$$\begin{aligned} \mathrm{var}[(n-1)^{-1} \sum_{i=1}^{n} (\beta_{i,\ell} - m_{\beta,\ell})Z_{i,m}] &= (n-1)^{-2} \sum_{i=1}^{n} E(\beta_{i,\ell} - m_{\beta,\ell})^2 E(Z^2_{i,m}) \\ &= [\boldsymbol{\Sigma}_{\beta,\beta}]_{\ell,\ell}(n^2 - n)^{-1}\sigma^2[\sum_{i=1}^{n} (\mathbf{X}_i'\mathbf{X}_i)^{-1}]_{m,m} \\ &= O(n^{-1}t^{-1}) . \end{aligned}$$

Therefore,

$$S_{\beta Z} = O_p(n^{-1/2}t^{-1/2}) \quad . \tag{A.3}$$

Now,

$$S_{ZZ} = (n-1)^{-1}\left[\sum_{i=1}^{n} Z_i Z_i' - n\,\bar{Z}\bar{Z}'\right] = O_p(t^{-1})$$

since from Remark 1 we know that $Z_i = O_p(t^{-1/2})$ and $\bar{Z} = O_p(n^{-1/2}\ t^{-1/2})$. Therefore, from (A.2) and (A.3), we have

$$S_{bb} = S_{\beta\beta} + O_p(t^{-1}) + O_p(n^{-1/2}t^{-1/2})$$

$$= S_{\beta\beta} + O_p(t^{-1/2}) \tag{A.4}$$

Combining (A.1) and (A.4), we get under H_0: $L\beta = \lambda_o$,

$$\hat{T}^{*2} - T_m^2 = O_p(t^{-1/2})$$

where

$$T_m^2 = n(m_\beta - \beta)'\ L'[L\ S_{\beta\beta}L']^{-1}L(m_\beta - \beta) \quad .$$

Now, the result (a) follows because T_m^2 has the Hotelling's T^2 distribution with (n-1) degrees of freedom.

(b) <u>t fixed and n tends to infinity</u>:

From Remark 3.2, we know that the exact distribution of T^{*2} is chi-square with q degrees of freedom. The difference between T^{*2} and $\hat{T}^{*2}$ is that the matrix n var($\bar{b}$) is replaced by it's unbiased estimator S_{bb}. If we can show that S_{bb} is consistent (as $n \to \infty$), then the result (b) will follow from Slutsky's Theorem.

From (A.2) and (A.3) we have,

$$\mathbf{S}_{bb} = \mathbf{S}_{\beta\beta} + \mathbf{S}_{ZZ} + O_p(n^{-1/2}t^{-1/2}).$$

Now

$$\mathbf{S}_{ZZ} = (n-1)^{-1}[\sum_{i=1}^{n} \mathbf{Z}_i\mathbf{Z}_i' - n\mathbf{Z}\mathbf{Z}']$$

$$= n^{-1}\sum_{i=1}^{n} \mathbf{Z}_i\mathbf{Z}_i' + n^{-1}(n-1)^{-1}\sum_{i=1}^{n} \mathbf{Z}_i\mathbf{Z}_i' - (n-1)^{-1}n\,\bar{\mathbf{Z}}\bar{\mathbf{Z}}'$$

$$= n^{-1}\sum_{i=1}^{n} \mathbf{Z}_i\mathbf{Z}_i' + O_p(n^{-1}t^{-1}) \ , \tag{A.5}$$

$$\mathbf{S}_{bb} = \mathbf{S}_{\beta\beta} + n^{-1}\sum_{i=1}^{n} \mathbf{Z}_i\mathbf{Z}_i' + O_p(n^{-1/2}t^{-1/2}) \ . \tag{A.6}$$

Now, since $\boldsymbol{\beta}_i$'s are iid $N(\mathbf{0}, \boldsymbol{\Sigma}_{\beta\beta})$ variables, we have

$$\mathbf{S}_{\beta\beta} = \boldsymbol{\Sigma}_{\beta\beta} + O_p(n^{-1/2}) \ . \tag{A.7}$$

Also, since $\mathbf{Z}_i$'s are independent $N(\mathbf{0}, \sigma^2(\mathbf{X}_i'\mathbf{X}_i)^{-1})$ variables, we have

$$E[n^{-1}\sum_{i=1}^{n} \mathbf{Z}_i\mathbf{Z}_i'] = n^{-1}\sigma^2\sum_{i=1}^{n} (\mathbf{X}_i'\mathbf{X}_i)^{-1} \tag{A.8}$$

and

$$\mathrm{Var}[n^{-1}\boldsymbol{\lambda}'\sum_{i=1}^{n} \mathbf{Z}_i\mathbf{Z}_i'\boldsymbol{\lambda}] = 2n^{-2}\sigma^4\sum_{i=1}^{n} [\boldsymbol{\lambda}'(\mathbf{X}_i'\mathbf{X}_i)^{-1}\boldsymbol{\lambda}]^2$$

$$= O(n^{-1}t^{-2}) \ , \tag{A.9}$$

for any arbitrary vector $\boldsymbol{\lambda}$. Therefore,

$$\mathbf{S}_{bb} = \boldsymbol{\Sigma}_{\beta\beta} + n^{-1}\sigma^2\sum_{i=1}^{n} (\mathbf{X}_i'\mathbf{X}_i)^{-1} + O_p(n^{-1/2})$$

$$= n\ \mathrm{var}(\bar{\mathbf{b}}) + O_p(n^{-1/2}) \ ,$$

and the result (b) follows.

(c) <u>nt large and q=1</u>:

Consider the t-statistic

$$\hat{T}^* = (\boldsymbol{\ell}'\mathbf{S}_{bb}\boldsymbol{\ell})^{-1/2}\, n^{1/2}(\boldsymbol{\ell}'\bar{\mathbf{b}} - \lambda_o)$$

for testing the hypothesis H_0: $\boldsymbol{\ell}'\boldsymbol{\beta} = \lambda_o$. We know that the variable

$$T^* = [\boldsymbol{\ell}'n\ \mathrm{var}(\bar{\mathbf{b}})\boldsymbol{\ell}]^{-1/2}\, n^{1/2}(\boldsymbol{\ell}'\bar{\mathbf{b}} - \lambda_o)$$

has a standard normal distribution. To show that

$\hat{T}^* = T^*\, [\boldsymbol{\ell}'n\ \mathrm{var}(\bar{\mathbf{b}})\boldsymbol{\ell}]^{1/2}(\boldsymbol{\ell}'\mathbf{S}_{bb}\boldsymbol{\ell})^{-1/2}$

is (approximately) distributed as Student's t-distribution with ν^* degrees of freedom, we need to show that $\nu^*[\boldsymbol{\ell}'\, n\ \mathrm{var}(\bar{\mathbf{b}})\ \boldsymbol{\ell}]^{-1}\ \boldsymbol{\ell}'\mathbf{S}_{bb}\boldsymbol{\ell}$ is (approximately) a chi-square random variable with ν^* degrees of freedom and is (asymptotically, when nt is large) independent of $\boldsymbol{\ell}'\bar{\mathbf{b}}$.

From (A.6), (A.8) and (A.9) we have

$$\mathbf{S}_{bb} = \mathbf{S}_{\beta\beta} + \sigma^2\, n^{-1}\sum_{i=1}^{n}(\mathbf{X}_i'\mathbf{X}_i)^{-1} + O_p(n^{-1/2}t^{-1/2})$$

$$= \mathbf{S}_{\beta\beta} + \hat{\sigma}^2 n^{-1}\sum_{i=1}^{n}(\mathbf{X}_i'\mathbf{X}_i)^{-1} + O_p(n^{-1/2}t^{-1/2})$$

where $\hat{\sigma}^2$ is defined in Section 2. Note that $(n-1)(\boldsymbol{\ell}'\mathbf{S}_{\beta\beta}\boldsymbol{\ell})(\boldsymbol{\ell}'\boldsymbol{\Sigma}_{\beta\beta}\boldsymbol{\ell})^{-1}$ is a $\chi^2(n-1)$ random variable and $(nt-nk)\hat{\sigma}^2/\sigma^2$ is a $\chi^2(nt-nk)$ random variable. Therefore, $\mathbf{S}_{bb}$ is the sum of independent scalar multiples of chi-square random variables. Ignoring the terms of order $(nt)^{-1/2}$ and using Satterthwaite's approximation, we have that $\nu^*[\boldsymbol{\ell}'n\ \mathrm{var}(\bar{\mathbf{b}})\boldsymbol{\ell}]^{-1}\ \boldsymbol{\ell}'\mathbf{S}_{bb}\boldsymbol{\ell}$ is approximately distributed as chi-square with ν^* degrees of freedom.

Now, to show the (asymptotic) independence of T^* and $\mathbf{S}_{bb}$, note that $\tilde{\mathbf{b}} = \mathbf{m}_\beta + \bar{\mathbf{Z}}$ is independent of $\mathbf{S}_{\beta\beta}$ since $\boldsymbol{\beta}_i$'s are NID($\boldsymbol{\beta},\boldsymbol{\Sigma}_{\beta\beta}$) and are independent of $\{\mathbf{Z}_i\}$. Also, for each i, the least squares estimator $\mathbf{b}_i$ is independent of the residual sums of squares $\mathbf{y}_i'\mathbf{y}_i - \mathbf{b}_i'\mathbf{X}_i'\mathbf{y}_i$ and hence $\bar{\mathbf{b}}$ and $\hat{\sigma}^2$ are independent. Therefore, for nt large, the distribution of $\hat{T}^*$ can be approximated by Student's t-distribution with ν^* degrees of freedom.