October 24, 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. THE RANDOM COEFFICIENT REGRESSION 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 where k, the number of parameters in the mean model, is assumed to be fixed 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 exogenous 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

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

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

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

where

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

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

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

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

and

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

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 $\boldsymbol{\Sigma}_{\beta\beta}$ in the case $\hat{\boldsymbol{\Sigma}}_{\beta\beta}$ is not nonnegative definite.)

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

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

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 <u>fixed n</u> and <u>t tending to infinity</u>:

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

(b) for a <u>fixed t</u> and <u>n tending to infinity</u>:

T^2_{EGLS} 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_{EGLS} 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

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

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_{EGLS}. 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,

$$\hat{\boldsymbol{\beta}}_{OLS} = (\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i)^{-1}(\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{y}_i)$$

$$= (\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i)^{-1}(\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i\mathbf{b}_i) \qquad (3.1)$$

because $\hat{\beta}_{EGLS}$ depends on the estimated variance-covariance matrix which may introduce additional variability. Note that if $\boldsymbol{\Sigma}_{\beta\beta} = \mathbf{0}$, that is, if the regression coefficients are fixed, then the BLUE $\hat{\boldsymbol{\beta}}_{GLS}$ is the same as $\hat{\boldsymbol{\beta}}_{OLS}$. 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 $\boldsymbol{\Sigma}_{\beta\beta}$.

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

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

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

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

Note that

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

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

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

Consider the statistic

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

for testing H_0: $\mathbf{L}\boldsymbol{\beta} = \boldsymbol{\lambda}_o$, where $\mathbf{L}$ is a qxk matrix of linearly independent rows. Notice that $\hat{T}^{*2}_{BAR}$ 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 $\hat{T}^{*2}_{BAR}$ has similar asymptotic properties as that of the statistic T^2_{EGLS}, we will make a few remarks.

<u>Remark 3.1</u>: 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

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

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 $\hat{T}^{*2}_{BAR}$ 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 $\hat{T}^{*2}_{BAR}$ defined in (3.4) based on the estimator $\bar{\mathbf{b}}$. Then,

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

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

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

and

(c) for the case where nt is large and q=1: $\hat{T}^{*2}_{BAR}$ 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. ∎

As either n or t tends to infinity, the asymptotic distributions of T^2_{EGLS} and $\hat{T}^{*2}_{BAR}$ are identical. The advantage of $\hat{T}^{*2}_{BAR}$ over T^2_{EGLS} is that it is simple to compute and is simple to explain. When nt is large, the Satterthwaite's approximation for the degrees of freedom for the distribution of $\hat{T}^{*2}_{BAR}$ is different from that of T^2_{EGLS}. Note however that, as in the case of T^2_{EGLS}, 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 $t \rightarrow \infty$), the estimators $\hat{\boldsymbol{\beta}}_{GLS}$, $\hat{\boldsymbol{\beta}}_{EGLS}$ and $\bar{\mathbf{b}}$ are equivalent and are efficient. Also, asymptotically (as $n \rightarrow \infty$), 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 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

$$y_{ij} = \beta_{0i}X_{0i} + \beta_{1i}X_{1ij} + e_{ij}\ , \quad i=1,2,\ldots,n;\ j=1,\ldots,t,$$

where $\boldsymbol{\beta}_i = (\beta_{0i}, \beta_{1i})'$ are NID$(\mathbf{0},\boldsymbol{\Sigma}_{\beta\beta})$; $\boldsymbol{\Sigma}_{\beta\beta} = \begin{bmatrix} 4 & 4 \\ 4 & 8 \end{bmatrix}$; $\mathbf{X}_{ij} = (X_{0i}, X_{1ij}) = \delta_i^{-1/2}(1,u_{ij})$; u_{ij}'s are NID(0,1); e_{ij}'s are NID$(0,\sigma^2)$ and $\{\boldsymbol{\beta}_i\}$, $\{u_{ij}\}$ and $\{e_{ij}\}$ are independent. The values for σ^2 are taken to be 4, 16 and 64, representing small, moderate and large ratios of within and between individual variances. The model matrices $\mathbf{X}_i$ are computed by setting $\delta_i = 1$ for i even and $\delta_i = \delta$, which takes the values 1, 3 and 9, for i odd. The different values of δ allow the model matrices to vary over different experimental units. When $\delta = 1$, the model matrices $\mathbf{X}_i$ are expected to be the same (except for random variation) for different experimental units. When $\delta = 4$ (9) the model matrices $\mathbf{X}_i$ vary moderately (greatly) over i. The values of n and t are taken to be 5, 10 and 50 to represent small, moderate and large samples. A set of 50x50 u_{ij} values is generated once for all and the same values of u_{ij}, $i=1,\ldots,n$; $j=1,\ldots,t$, are used in all of the Monte Carlo replications. For each set of values of n, t, δ and σ^2, 1000 Monte Carlo replications are used. In each replication, $\boldsymbol{\beta}_i$'s and e_{ij}'s are generated independently and the test statistics T^2_{EGLS}, $\hat{T}^{*2}_{BAR}$ and T^2_{OLS} (based on $\hat{\boldsymbol{\beta}}_{OLS}$) for testing the hypotheses (1) H_0: $\beta_0 = \beta_1 = 0$ and (2) H_0: $\beta_0 = \beta_1 = 1$ are computed. For the sake of brevity, we present only the results for $\delta = 1$, $\sigma^2 = 4$; $\delta = 1$, $\sigma^2 = 64$ and $\delta = 9$, $\sigma^2 = 4$, which provide examples of the best and worst performances of $\hat{T}^{*2}_{BAR}$. (The results of a larger Monte Carlo study, which includes the test statistics for testing H_0: $\beta_1 = 0$ and H_0: $\beta_1 = 1$, for all parametric configurations may be obtained from the authors.)

Whenever the estimate of $\boldsymbol{\Sigma}_{\beta\beta}$ is not nonnegative definite, the modified estimator suggested by Carter and Yang (1986) is used. The number of times the estimate of $\Sigma_{\beta\beta}$ is modified is presented in Table 1. The number of times different test statistics reject the two hypotheses are summarized in Tables 2 and 3.

First, note that the unbiased estimate of $\boldsymbol{\Sigma}_{\beta\beta}$ is observed to require the modification in a large number of cases, especially when σ^2 is large (compared to the elements of $\boldsymbol{\Sigma}_{\beta\beta}$) and/or nt is small to moderate. The modified estimator of $\boldsymbol{\Sigma}_{\beta\beta}$ suggested by Carter and Yang (1986) is consistent but biased. The frequent modification of the estimate of $\boldsymbol{\Sigma}_{\beta\beta}$ may be one of the reasons the number of rejections (the empirical levels) of the true hypothesis H_0: $\beta_0 = \beta_1 = 0$ is much lower than the nominal level of 5%, for the tests based on $\hat{\boldsymbol{\beta}}_{EGLS}$ and $\hat{\boldsymbol{\beta}}_{OLS}$, when δ=1 and n is small to moderate. When δ=9, the empirical level of the test criterion based on T^2_{EGLS} is observed to be larger than the nominal level, for the cases where nt is large. On the other hand, the empirical level of the F-test based on $\bar{\mathbf{b}}$ (and $\mathbf{S}_{bb}$) is observed to be much closer to the nominal level. (Recall that the test statistic $\hat{T}^{*2}_{BAR}$ based on $\bar{\mathbf{b}}$ depends only on $\mathbf{S}_{bb}$ and is not affected by the modifications of $\hat{\boldsymbol{\Sigma}}_{\beta\beta}$.)

To compare the power (the number of rejections of H_0: $\beta_0 = \beta_1 = 1$) of the test statistics, it is useful to recall the expressions of the three estimators as weighted averages of the individual regression estimators $\mathbf{b}_i$. The weight matrix $\mathbf{W}_i = [\boldsymbol{\Sigma}_{\beta\beta} + \sigma^2(\mathbf{X}_i'\mathbf{X}_i)^{-1}]^{-1}$ for the generalized least squares estimator is nearly constant if either $\sigma^2(\mathbf{X}_i'\mathbf{X}_i)^{-1}$ is small compared to $\boldsymbol{\Sigma}_{\beta\beta}$ or if the model matrices $\mathbf{X}_i$ are nearly the same for all experimental units. In either case, the simple average $\bar{\mathbf{b}}$ is expected to be as efficient as the BLUE $\hat{\boldsymbol{\beta}}_{GLS}$. On the other hand, if $\sigma^2(\mathbf{X}_i'\mathbf{X}_i)^{-1}$ is large compared to $\boldsymbol{\Sigma}_{\beta\beta}$, $\mathbf{W}_i$ is approximately $\sigma^{-2}\ \mathbf{X}_i'\mathbf{X}_i$ (which is the weight matrix for $\hat{\boldsymbol{\beta}}_{OLS}$) and hence the estimator $\hat{\boldsymbol{\beta}}_{OLS}$ is expected to be efficient. In Table 4, we summarize the theoretical minimum relative efficiency, defined as $MRE(\hat{\boldsymbol{\beta}}) = \min_{\boldsymbol{\ell}} \{Var(\boldsymbol{\ell}'\hat{\boldsymbol{\beta}}_{GLS})/Var(\boldsymbol{\ell}'\hat{\boldsymbol{\beta}})\}$, where

$\hat{\boldsymbol\beta}$ is either $\bar{\mathbf{b}}$ or $\hat{\boldsymbol\beta}_{OLS}$. Note that the MRE($\hat{\boldsymbol\beta}$) is the minimum value of the relative efficiency of $\boldsymbol\ell'\hat{\boldsymbol\beta}$ where the minimum is taken over all possible linear combinations $\boldsymbol\ell'\boldsymbol\beta$. Our choice of the simulation parameters is such that the MRE of $\bar{\mathbf{b}}$ and $\hat{\boldsymbol\beta}_{OLS}$ vary from 40% to 100% .

In the case $\delta = 1$ and $\sigma^2 = 4$, the MRE($\bar{\mathbf{b}}$) is above 90% and the test statistic $\hat{T}^{*2}_{BAR}$ based on $\bar{\mathbf{b}}$ has performed very well. For small n, the test criterion based on $\hat{T}^{*2}_{BAR}$ is observed to be more powerful than the criterion based on T^2_{EGLS}. (Recall that, in this case, the empirical level of the criterion based on T^2_{EGLS} is far below the nominal level.). For large n, the powers of the criteria based on T^2_{EGLS} and $\hat{T}^{*2}_{BAR}$ are similar and are consistently higher than that of the criterion based on T^2_{OLS}.

When $\sigma^2 = 64$, the MRE($\hat{\boldsymbol\beta}_{OLS}$) is above 90% whereas the MRE($\bar{\mathbf{b}}$) ranged from 55% to 99%. Based on the MRE($\hat{\boldsymbol\beta}_{OLS}$), the criterion based on T^2_{OLS} is expected to perform well. However, in almost every case, the criterion based on $\hat{T}^{*2}_{BAR}$ rejected the false hypothesis H_0: $\beta_0 = \beta_1 = 1$, more often than the criterion based on T^2_{OLS}. An examination of the estimated variance of $\hat{\boldsymbol\beta}_{OLS}$ reveals an upward bias for samples as large as n = 10 and t = 10. Thus, even though $\hat{\boldsymbol\beta}_{OLS}$ is an efficient estimator, the test criterion based on T^2_{OLS} which depends on $\widehat{\text{Var}}(\hat{\boldsymbol\beta}_{OLS})$, an estimate of $\text{Var}(\hat{\boldsymbol\beta}_{OLS})$, is not observed to be powerful. This is partly explained by the high number of times the modification to $\hat{\boldsymbol\Sigma}_{\beta\beta}$ is required when $\sigma^2 = 64$.

In the case $\delta = 9$ and $\sigma^2 = 4$, the model matrices $\mathbf{X}_i$ vary considerably and the MRE's of $\bar{\mathbf{b}}$ and $\hat{\boldsymbol\beta}_{OLS}$ are not very high. Except in the case n = 5 and t = 5 (small samples), the criterion based on T^2_{EGLS} is observed to be more powerful than the other two criteria. When t = 50, the MRE of $\bar{\mathbf{b}}$ is above 90% and the empirical powers of $\hat{T}^{*2}_{BAR}$ and T^2_{EGLS} are comparable. When nt is moderate, the powers of all three criteria are small.

Some general conclusions that may be drawn from our study are as follows:

1. When σ^2 is large compared to the elements of $\boldsymbol{\Sigma}_{\beta\beta}$, there is a high chance that the unbiased estimator of $\boldsymbol{\Sigma}_{\beta\beta}$ is not nonnegative definite. The modifications suggested by Carter and Yang (1986) cause an upward bias in $\hat{\boldsymbol{\Sigma}}_{\beta\beta}$ which affects the power of the test criteria based on T^2_{EGLS} and T^2_{OLS}.
2. The minimum relative efficiency of $\bar{\mathbf{b}}$ provides some guidance as to the power of the test criterion based on $\hat{T}^{*2}_{BAR}$. The higher the MRE($\bar{\mathbf{b}}$), the better the performance of $\hat{T}^{*2}_{BAR}$ is.
3. The test criterion based on $\hat{T}^{*2}_{BAR}$ maintained the proper level for <u>all</u> parametric configurations considered in our study.
4. Also, the test criterion based on $\hat{T}^{*2}_{BAR}$ is observed to have more power than the other two criteria when either (a) the sample size is small or (b) the variation in the model matrices $\mathbf{X}_i$ over i is not large or (c) the variance σ^2 is large.
5. On the other hand, when n is large the criterion based on T^2_{EGLS} is more powerful than the other two criteria. Also, the performance of T^2_{EGLS} is better than that of $\hat{T}^{*2}_{BAR}$ when the variation in $\mathbf{X}_i$ over i is large.
6. The sample sizes considered in our study are not large enough for the chi-square approximation to be reasonable.

Table 1. Proportion of Times $\hat{\Sigma}_{\beta\beta}$ is Observed Not to be a Nonnegative Definite Matrix.

n	t	$\delta = 1,\ \sigma^2 = 4$	$\delta = 1,\ \sigma^2 = 64$	$\delta = 9,\ \sigma^2 = 4$
5	5	0.425	0.810	0.740
	10	0.258	0.786	0.620
	50	0.024	0.511	0.186
10	5	0.180	0.806	0.723
	10	0.037	0.636	0.438
	50	0.000	0.205	0.030
50	5	0.003	0.640	0.384
	10	0.000	0.282	0.053
	50	0.000	0.000	0.000

Table 2. Comparisons of the Levels of Test Criteria: The Number of Times a 5% Level Test Criterion Rejects H_0: $\beta_0 = \beta_1 = 0$. (1000 Monte Carlo Replications)

n	t	Test	$\delta = 1, \sigma^2 = 4$ $F_{2,n-2}$	$\delta = 1, \sigma^2 = 4$ χ^2_2	$\delta = 1, \sigma^2 = 64$ $F_{2,n-2}$	$\delta = 1, \sigma^2 = 64$ χ^2_2	$\delta = 9, \sigma^2 = 4$ $F_{2,n-2}$	$\delta = 9, \sigma^2 = 4$ χ^2_2
5	5	EGLS	19	211	1	72	29	242
		BBAR	52	254	48	266	48	246
		OLS	10	164	1	52	10	144
	10	EGLS	24	210	2	91	48	298
		BBAR	53	239	57	247	43	237
		OLS	12	170	2	82	25	212
	50	EGLS	47	249	9	177	80	321
		BBAR	48	251	52	249	50	270
		OLS	28	205	8	163	36	257
10	5	EGLS	49	134	16	66	143	252
		BBAR	45	123	53	130	31	110
		OLS	35	115	12	51	81	168
	10	EGLS	63	138	19	69	115	228
		BBAR	64	133	43	117	36	123
		OLS	41	125	17	64	80	176
	50	EGLS	51	125	27	105	84	162
		BBAR	52	126	40	117	65	136
		OLS	49	115	21	104	53	127
50	5	EGLS	49	65	50	62	129	153
		BBAR	37	48	59	72	49	62
		OLS	32	49	51	66	96	109
	10	EGLS	51	63	52	62	87	107
		BBAR	53	60	55	69	50	62
		OLS	36	49	49	63	89	109
	50	EGLS	49	61	61	73	54	68
		BBAR	48	61	59	70	52	65
		OLS	43	61	52	68	45	60

Table 3. Comparisons of the Powers of Test Criteria: The Number of Times a 5% Level Test Criterion Rejects H_0: $\beta_0=\beta_1=1$. (1000 Monte Carlo Replications)

n	t	Test	$\delta = 1, \sigma^2 = 4$		$\delta = 1, \sigma^2 = 64$		$\delta = 9, \sigma^2 = 4$	
			$F_{2,n-2}$	χ^2_2	$F_{2,n-2}$	χ^2_2	$F_{2,n-2}$	χ^2_2
5	5	EGLS	35	303	1	110	46	317
		BBAR	80	354	57	299	76	342
		OLS	21	227	1	71	16	198
	10	EGLS	57	345	4	150	87	410
		BBAR	98	367	77	300	73	341
		OLS	32	268	4	131	41	303
	50	EGLS	92	371	36	299	118	438
		BBAR	92	371	96	355	80	385
		OLS	52	325	30	288	62	362
10	5	EGLS	166	321	52	120	252	401
		BBAR	155	313	88	206	120	272
		OLS	128	269	46	108	145	256
	10	EGLS	179	353	61	184	264	412
		BBAR	174	357	116	259	164	321
		OLS	128	285	58	172	171	318
	50	EGLS	193	375	129	282	201	358
		BBAR	191	372	147	303	167	343
		OLS	179	351	127	277	146	282
50	5	EGLS	825	850	389	428	725	759
		BBAR	812	842	305	346	607	634
		OLS	740	774	361	408	530	562
	10	EGLS	863	882	533	572	772	797
		BBAR	860	882	513	550	739	774
		OLS	778	811	499	540	571	616
	50	EGLS	876	907	778	820	863	884
		BBAR	876	906	775	817	859	884
		OLS	856	887	772	803	637	667

Table 4. Minimum Relative Efficiencies of $\bar{b}$ and $\hat{\beta}_{OLS}$.

n	t	$\delta = 1, \sigma^2 = 4$		$\delta = 1, \sigma^2 = 64$		$\delta = 9, \sigma^2 = 4$	
		$\bar{b}$	$\hat{\beta}_{OLS}$	$\bar{b}$	$\hat{\beta}_{OLS}$	$\bar{b}$	$\hat{\beta}_{OLS}$
5	5	0.96	0.75	0.77	0.97	0.46	0.53
	10	0.98	0.73	0.83	0.95	0.56	0.69
	50	1.00	0.84	0.98	0.92	0.94	0.67
10	5	0.92	0.84	0.55	0.98	0.42	0.59
	10	0.99	0.80	0.88	0.97	0.65	0.65
	50	1.00	0.90	0.99	0.95	0.96	0.61
50	5	0.92	0.72	0.59	0.94	0.43	0.63
	10	0.99	0.78	0.88	0.92	0.70	0.59
	50	1.00	0.92	0.99	0.96	0.97	0.61

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 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. Under the assumptions of our model, the exact distribution of $\bar{\mathbf{b}}$ is normal for all sample sizes. It is interesting to note that the empirical level of the test criterion based on $\hat{T}^{*2}_{BAR}$ is close to the nominal level (5%), for all parametric configurations considered in our study. The estimator $\hat{\boldsymbol{\beta}}_{EGLS}$, on the other hand, 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. The empirical levels of the test criterion based on T^2_{EGLS} are not close to the nominal level when the sample size (n) is small. It requires a moderately large number of experimental units for $\hat{\beta}_{EGLS}$ to be close to the BLUE $\hat{\beta}_{GLS}$ and the variance of $\hat{\beta}_{EGLS}$ to be close to that of $\hat{\beta}_{GLS}$. 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 ith 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 $\mathbf{X}_i$, $\boldsymbol{\Sigma}_{\beta\beta}$ and $\boldsymbol{\Sigma}_{ee}$ one can obtain the asymptotic results for the test statistic based on $\bar{\mathbf{b}}$ and $\mathbf{S}_{bb}$.

ACKNOWLEDGEMENTS

The work of S. G. Pantula was partially supported by the National Science Foundation. We are grateful to Cavell Brownie, John Rawlings and the referees for their helpful comments.

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

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

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

Also, recall, $\mathbf{b}_i = \boldsymbol{\beta}_i + \mathbf{Z}_i$ and hence

$$\mathbf{S}_{bb} = \mathbf{S}_{\beta\beta} + \mathbf{S}_{ZZ} + \mathbf{S}_{\beta Z} + \mathbf{S}'_{\beta Z} , \qquad (A.2)$$

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,

$$S_{\beta Z} = 0_p(n^{-1/2}t^{-1/2}) \ . \tag{A.3}$$

Now,

$$S_{ZZ} = (n-1)^{-1}\left[\sum_{i=1}^{n} Z_i Z_i' - n\,\bar{Z}\bar{Z}'\right] = 0_p(t^{-1})$$

since from Remark 1 we know that $Z_i = 0_p(t^{-1/2})$ and $\bar{Z} = 0_p(n^{-1/2}\ t^{-1/2})$.

Therefore, from (A.2) and (A.3), we have

$$S_{bb} = S_{\beta\beta} + 0_p(t^{-1/2}) \tag{A.4}$$

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

$$\hat{T}^{*2}_{BAR} - T^2_m = 0_p(t^{-1/2})$$

where

$$T^2_m = n(m_\beta - \beta)'\ L'[L\ S_{\beta\beta}L']^{-1}L(m_\beta - \beta)\ .$$

Now, the result (a) follows because T^2_m 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}_{BAR} is chi-square with q degrees of freedom. The difference between T^{*2}_{BAR} and $\hat{T}^{*2}_{BAR}$ is that the matrix $[n\ \mathrm{var}(\bar{b})]$ is replaced by its 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,

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

Since $\mathbf{Z}_i = O_p(t^{-1/2})$, uniformly in i, and $\bar{\mathbf{Z}} = O_p(n^{-1/2}t^{-1/2})$, we have

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

and

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

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

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

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

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

and

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

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

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

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

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

and the result (b) follows.

(c) nt large and q=1:

Consider the t-statistic

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

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

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

has a standard normal distribution. To show that

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

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

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

$$S_{bb} = S_{\beta\beta} + \sigma^2 n^{-1} \sum_{i=1}^{n} (X_i' X_i)^{-1} + 0_p(n^{-1/2} t^{-1/2})$$

$$= S_{\beta\beta} + \hat{\sigma}^2 n^{-1} \sum_{i=1}^{n} (X_i' X_i)^{-1} + 0_p(n^{-1/2} t^{-1/2})$$

where $\hat{\sigma}^2$ is defined in Section 2. Note that $(n-1)(\ell' S_{\beta\beta} \ell)(\ell' \Sigma_{\beta\beta} \ell)^{-1}$ 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, 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 $\nu^* [\ell' n \, var(\bar{b}) \ell]^{-1} \, \ell' S_{bb} \ell$ is approximately distributed as chi-square with ν^* degrees of freedom.

Now, to show the (asymptotic) independence of T^*_{BAR} and $\mathbf{S}_{bb}$, note that $\bar{\mathbf{b}} = \mathbf{m}_\beta + \bar{\mathbf{Z}}$ 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.