94-F-36

Asymptotic Comparison of Variances of OLS and Feasible GLS Estimators in Heteroscedastic Regression Model under Contiguous Heteroscedasticity

by

Arinobu Nakao and Akimichi Takemura
University of Tokyo

December 1994

Asymptotic Comparison of Variances of OLS and Feasible GLS Estimators in Heteroscedastic Regression Model under Contiguous Heteroscedasticity

Arinobu Nakao and Akimichi Takemura*

Abstract

This paper compares asymptotic variances of OLS and feasible GLS estimators in heteroscedastic regression model under the condition that heteroscedasticity approaches 0 asymptotically. Based on asymptotic expansion of variances we will confirm that OLS performs better than feasible GLS when heteroscedasticity is small and feasible GLS performs better when heteroscedasticity is large. Our simulation study shows that the finite sample behavior of the estimators is largely consistent with our asymptotic results.

Keywords: OLS estimator, feasible generalized least estimator, contiguous heteroscedasticity

1 Introduction

Consider the heteroscedastic regression model

$$y_{ij} = \beta_0 + x_{ij1}\beta_1 + x_{ij2}\beta_2 + \cdots + x_{ijp}\beta_p + \varepsilon_{ij} , \tag{1.1}$$

$$i = 1, \ldots, k, \quad j = 1, \ldots, n_i ,$$

$$\varepsilon_{ij} \sim N(0, \sigma_i^2) ,$$

*Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyoku, Tokyo 113, Japan.

where ε_{ij}'s are mutually independently distributed according to normal distribution $N(0, \sigma_i^2)$, $i = 1, \dots, k$. Letting

$$Y_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{in_i} \end{pmatrix}, \quad X_i = \begin{pmatrix} 1 & x_{i11} & \dots & x_{i1p} \\ 1 & x_{i21} & \dots & x_{i2p} \\ \vdots & \vdots & & \vdots \\ 1 & x_{in_i1} & \dots & x_{in_ip} \end{pmatrix},$$

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \varepsilon_i = \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{in_i} \end{pmatrix},$$

(1.1) can be written in matrix form as

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_k \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix} \beta + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_k \end{pmatrix}, \tag{1.2}$$

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_k \end{pmatrix} \sim N\left(0, \begin{pmatrix} \sigma_1^2 I_{n_1} & & & \mathbf{0} \\ & \sigma_2^2 I_{n_2} & & \\ & & \ddots & \\ \mathbf{0} & & & \sigma_k^2 I_{n_k} \end{pmatrix}\right), \tag{1.3}$$

where Y_i is an $n_i \times 1$ vector of random observations, X_i is an $n_i \times (p+1)$ matrix of fixed regressors, β is a $(p+1) \times 1$ vector of common regression coefficients, and ε_i is an $n_i \times 1$ vector of disturbances.

If σ_i^2's are known, the best linear unbiased estimator of β is

$$\widehat{\beta}_{GLS} = \left(\sum_{i=1}^{k} \frac{X_i' X_i}{\sigma_i^2}\right)^{-1} \sum_{i=1}^{k} \frac{X_i' Y_i}{\sigma_i^2}. \tag{1.4}$$

However if σ_i^2's are unknown we can not use the generalized least square estimator $\widehat{\beta}_{GLS}$. In this case we may use

$$\widehat{\beta}_{OLS} = \left(\sum_{i=1}^{k} X_i' X_i\right)^{-1} \sum_{i=1}^{k} X_i' Y_i, \tag{1.5}$$

or "feasible generalized least square estimator"

$$\widehat{\beta}_{FGLS} = \left(\sum_{i=1}^{k} \frac{X_i' X_i}{\widehat{\sigma}_i^2}\right)^{-1} \left(\sum_{i=1}^{k} \frac{X_i' Y_i}{\widehat{\sigma}_i^2}\right), \tag{1.6}$$

where

$$\widehat{\sigma}_i^2 = \frac{Y_i' M Y_i}{n_i - d_i}, \quad M = I - X_i(X_i'X_i)^- X_i',$$
$$d_i = \text{rank } X_i . \tag{1.7}$$

The obvious question here is when $\widehat{\beta}_{OLS}$ performs better than $\widehat{\beta}_{FGLS}$ and vice versa. The purpose of our paper is to derive meaningful results on the relative performances of $\widehat{\beta}_{OLS}$ and $\widehat{\beta}_{FGLS}$ based on asymptotic expansion of their variances.

Throughout this paper we use the following notation.

$$t_{Ni} = \frac{n_i}{N}, \quad N = \sum_{i=1}^{k} n_i, \quad A_{Ni} = \frac{X_i'X_i}{n_i}. \tag{1.8}$$

Furthermore we assume that t_{Ni}, A_{Ni} approach a finite limit as $N \to \infty$.

Assumption 1

$$\lim_{N\to\infty} t_{Ni} = t_i ,$$
$$0 < t_i < 1, \quad \sum_{i=1}^{k} t_i = 1, \tag{1.9}$$
$$\lim_{N\to\infty} A_{Ni} = A_i, \quad i = 1, \ldots, k, \tag{1.10}$$

where A_i, $i = 1, \ldots, k$, are positive semidefinite.

If $\sigma_1^2, \ldots, \sigma_k^2$ are fixed and not equal to each other, then it is well known (see e.g. Theil (1971) Chapter 6, Judge et al. (1982) Chapter 10, Amemiya (1985) Chapter 6, Rothenberg (1984)) that $\text{Var}(\widehat{\beta}_{FGLS}) = \text{Var}(\widehat{\beta}_{GLS}) + o(N^{-1})$ and $\widehat{\beta}_{FGLS}$ is first order efficient under some reasonable conditions, whereas $\widehat{\beta}_{OLS}$ is not first order efficient by Gauss-Markov theorem. Hence $\widehat{\beta}_{FGLS}$ performs better than $\widehat{\beta}_{OLS}$ for this type of asymptotic framework. However for small sample $\widehat{\beta}_{OLS}$ may perform better than $\widehat{\beta}_{FGLS}$ because $\widehat{\beta}_{FGLS}$ has added variability due to estimation errors in $\widehat{\sigma}_1^2, \ldots, \widehat{\sigma}_k^2$ (see Taylor (1977,1978)). At first it seems difficult to treat this problem by an asymptotic argument. However by considering the "contiguous heteroscedasticity" where $\sigma_1^2 = \sigma_{N1}^2, \ldots, \sigma_k^2 = \sigma_{Nk}^2$ approach a common value as $N \to \infty$ and comparing the terms of order N^{-2} in the variances we can make meaningful comparison of $\widehat{\beta}_{OLS}$ and $\widehat{\beta}_{FGLS}$. More specifically we assume

Assumption 2

$$\frac{1}{\sigma_i^2} = \eta_{Ni} = c\left(1 + \frac{1}{\sqrt{N}}\tau_i\right), \quad i = 1, \ldots, k, \tag{1.11}$$

where c, τ_i are constants satisfying

$$\sum_{i=1}^{k} t_i \tau_i = 0\,, \tag{1.12}$$

$$c = \sum_{i=1}^{k} t_i \eta_{Ni}\,. \tag{1.13}$$

Note that c^{-1} is the limiting common value of σ_i^2 and σ_i^2 approaches c^{-1} with order $O(N^{-\frac{1}{2}})$. Intuitively this is the correct order of convergence, because to obtain non-trivial comparison of $\widehat{\beta}_{OLS}$ and $\widehat{\beta}_{FGLS}$, $\sigma_i^2 - c^{-1}$ has to be of the same order as the estimation error in $\widehat{\sigma}_i^2$.

2 Asymptotic expansion of variances of the estimators of the coefficient vector

It is well known that

$$\mathrm{Var}(\widehat{\beta}_{GLS}) = \left(\sum_{i=1}^{k} \frac{X_i'X_i}{\sigma_i^2}\right)^{-1} = \frac{1}{N}\left(\sum_{i=1}^{k} t_{Ni}\eta_{Ni}A_{Ni}\right)^{-1}, \tag{2.1}$$

$$\begin{aligned}\mathrm{Var}(\widehat{\beta}_{OLS}) &= \left(\sum_{i=1}^{k} X_i'X_i\right)^{-1} \sum_{i=1}^{k} \sigma_i^2 X_i'X_i \left(\sum_{i=1}^{k} X_i'X_i\right)^{-1} \\ &= \frac{1}{N}\left(\sum_{i=1}^{k} t_{Ni}A_{Ni}\right)^{-1} \sum_{i=1}^{k} \frac{t_{Ni}A_{Ni}}{\eta_{Ni}} \left(\sum_{i=1}^{k} t_{Ni}A_{Ni}\right)^{-1}. \end{aligned} \tag{2.2}$$

Since $\widehat{\sigma}_1^2, \ldots, \widehat{\sigma}_k^2, X_1'Y_1, \ldots, X_k'Y_k$, are mutually independent, $\widehat{\beta}_{FGLS}$ is an unbiased estimator and its variance can be expressed as

$$\begin{aligned} &\mathrm{Var}(\widehat{\beta}_{FGLS}) \qquad (2.3) \\ &= E\left[E\left\{\left(\sum_{i=1}^{k} \frac{X_i'X_i}{\widehat{\sigma}_i^2}\right)^{-1} \left(\sum_{i=1}^{k} \frac{X_i'\varepsilon_i}{\widehat{\sigma}_i^2}\right)\left(\sum_{i=1}^{k} \frac{X_i'\varepsilon_i}{\widehat{\sigma}_i^2}\right)' \right.\right. \\ &\quad \left.\left. \times \left(\sum_{i=1}^{k} \frac{X_i'X_i}{\widehat{\sigma}_i^2}\right)^{-1} \middle| \widehat{\sigma}_1^2, \ldots, \widehat{\sigma}_k^2 \right\}\right] \\ &= \frac{1}{N} E\left\{\left(\sum_{i=1}^{k} \frac{t_{Ni}A_{Ni}}{\widehat{\sigma}_i^2}\right)^{-1} \sum_{i=1}^{k} \left(\frac{\sigma_i^2 t_{Ni}A_{Ni}}{\widehat{\sigma}_i^4}\right)\left(\sum_{i=1}^{k} \frac{t_{Ni}A_{Ni}}{\widehat{\sigma}_i^2}\right)^{-1}\right\}. \qquad (2.4) \end{aligned}$$

Note that

$$\frac{\sigma_i^2}{\widehat{\sigma}_i^2} \sim \frac{n_i - d_i}{\chi^2(n_i - d_i)}, \quad d_i = \text{rank } X_i\,, \tag{2.5}$$

where $\chi^2(n_i - d_i)$ denotes χ^2 distribution with $n_i - d_i$ degrees of freedom. The right-hand side of (2.4) has the following asymptotic expansion.

Lemma 1 *Under Assumptions 1 and 2, as $N \to \infty$*

$$\begin{aligned}\text{Var}(\widehat{\beta}_{FGLS}) &= \frac{1}{N} B_{N0}^{-1} \tilde{B}_{N0} B_{N0}^{-1} \\ &+ \frac{2}{N^2} B_{N0}^{-1} \left\{ \sum_{i=1}^{k} \eta_{Ni} A_{Ni} - \sum_{i=1}^{k} t_{Ni} \eta_{Ni}^2 A_{Ni} B_{N0}^{-1} A_{Ni} \right\} B_{N0}^{-1} \\ &+ o(N^{-2})\,,\end{aligned} \tag{2.6}$$

where

$$\begin{aligned}B_{N0} &= \sum_{i=1}^{k} t_{Ni} \eta_{Ni} s_{Ni} A_{Ni}\,, \\ \tilde{B}_{N0} &= \sum_{i=1}^{k} t_{Ni} \eta_{Ni} s_{Ni}^2 A_{Ni}\,, \\ s_{Ni} &= \frac{n_i - d_i}{n_i - d_i - 2}.\end{aligned} \tag{2.7}$$

Note that we assume nonsingularity of B_{N0}. Proof of Lemma 1 and Theorem 1 below are given in Section 5. Combining (2.1), (2.2) and Lemma 1 we have the following main theorem.

Theorem 1 *Under Assumption 1 and 2,*

$$\begin{aligned}&\lim_{N\to\infty} N^2 c \left\{ \text{Var}(\widehat{\beta}_{FGLS}) - \text{Var}(\widehat{\beta}_{OLS}) \right\} \\ &= B_0^{-1} \left\{ 2 \left(\sum_{i=1}^{k} A_i - \sum_{i=1}^{k} t_i A_i B_0^{-1} A_i \right) \right. \\ &\quad \left. - \left(\sum_{i=1}^{k} t_i \tau_i^2 A_i - \sum_{i=1}^{k} t_i \tau_i A_i B_0^{-1} \sum_{i=1}^{k} t_i \tau_i A_i \right) \right\} B_0^{-1}\,,\end{aligned} \tag{2.8}$$

where

$$B_0 = \lim_{N\to\infty} \frac{B_{N0}}{c} = \sum_{i=1}^{k} t_i A_i\,. \tag{2.9}$$

In section 5 we also prove the validity of the asymptotic expansions in Lemma 1 and Theorem 1, namely the remainder terms are finite and actually of the order $o(N^{-2})$.

Theorem 1 shows that the limiting relative performances of $\widehat{\beta}_{FGLS}$ and $\widehat{\beta}_{OLS}$ depend on the difference of two matrices $C = 2\left(\sum A_i - \sum t_i A_i B_0^{-1} A_i\right)$, $D = \sum t_i \tau_i^2 A_i - \sum t_i \tau_i A_i B_0^{-1} \sum t_i \tau_i A_i$. C and D are both nonnegative definite. If $\tau_1 = \ldots = \tau_k = 0$, then $C - D = C$ which is nonnegative definite. If $|\tau_1|, \ldots, |\tau_k|$ are large, then $C - D$ is usually nonpositive definite. Therefore asymptotically $\widehat{\beta}_{OLS}$ performs better than $\widehat{\beta}_{FGLS}$ when $|\tau_1|, \ldots, |\tau_k|$ are small and vice versa. However $C - D$ may not be of definite sign for intermediate values of $|\tau_1|, \ldots, |\tau_k|$ as shown in Example 2 of the following section.

Nonnegative-definiteness of C can be shown as follows. Let $B_0^{-1} = T'T$ where T is a nonsingular matrix. Let $TA_iT' = F_i$. Then

$$\begin{aligned} & B_0^{-1}\left(\sum_{i=1}^k A_i - \sum_{i=1}^k t_i A_i B_0^{-1} A_i\right) B_0^{-1} \\ & = T' \sum_{i=1}^k (F_i - t_i F_i^2) T \\ & = T' \sum_{i=1}^k \frac{1}{t_i}\left(t_i F_i - (t_i F_i)^2\right) T\,, \end{aligned} \tag{2.10}$$

which is non-negative definite because $\sum_{i=1}^k t_i F_i = I$ implies that all the eigenvalue of $t_i F_i$ are in [0,1] and hence $t_i F_i - t_i^2 F_i^2$ is non-negative definite.

Also D is non-negative definite, because

$$\begin{aligned} & B_0^{-1}\left(\sum_{i=1}^k t_i \tau_i^2 A_i - \sum_{i=1}^k t_i \tau_i A_i B_0^{-1} \sum_{i=1}^k t_i \tau_i A_i\right) B_0^{-1} \\ & = T'\left(\sum_{i=1}^k t_i \tau_i^2 F_i - \left(\sum_{i=1}^k t_i \tau_i F_i\right)^2\right) T\,, \end{aligned} \tag{2.11}$$

and this can be written as

$$T' \sum_{i=1}^k (\sqrt{t_i}\tau_i G_i - \sqrt{t_i} G_i \sum_{j=1}^k t_j \tau_j F_j)'(\sqrt{t_i}\tau_i G_i - \sqrt{t_i} G_i \sum_{j=1}^k t_j \tau_j F_j) T\,,$$

where

$$F_i = G_i' G_i\,.$$

3 Examples

In this section we specialize Theorem 1 to two simple cases. The first example is the problem of estimating a common mean.

Example 1

We will estimate the "common mean" μ in the situation

$$X_{ij} \sim N(\mu, \sigma_i^2)\,, \quad i = 1, \ldots, k\,, \quad j = 1, \ldots, n_i\,,$$

where each X_{ij} is independent. In this case the estimators are

$$\widehat{\mu}_{GLS} = \frac{n_1 \frac{\bar{X}_1}{\sigma_1^2} + n_2 \frac{\bar{X}_2}{\sigma_2^2} + \cdots + n_k \frac{\bar{X}_k}{\sigma_k^2}}{\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2} + \cdots + \frac{n_k}{\sigma_k^2}}\,, \tag{3.1}$$

$$\widehat{\mu}_{OLS} = \frac{1}{N} \sum_{i=1}^{k} n_i \bar{X}_i\,, \tag{3.2}$$

$$\widehat{\mu}_{FGLS} = \frac{n_1 \frac{\bar{X}_1}{\widehat{\sigma}_1^2} + n_2 \frac{\bar{X}_2}{\widehat{\sigma}_2^2} + \cdots + n_k \frac{\bar{X}_k}{\widehat{\sigma}_k^2}}{\frac{n_1}{\widehat{\sigma}_1^2} + \frac{n_2}{\widehat{\sigma}_2^2} + \cdots + \frac{n_k}{\widehat{\sigma}_k^2}}\,, \tag{3.3}$$

where

$$\bar{X}_i = \frac{\sum_{j=1}^{n_i} X_{ij}}{n_i}\,, \quad \widehat{\sigma}_i^2 = \frac{\sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2}{n_i - 1}\,.$$

The variances of estimators are

$$\mathrm{Var}(\widehat{\mu}_{GLS}) = \frac{1}{N} \left(\frac{t_{N1}}{\sigma_1^2} + \frac{t_{N2}}{\sigma_2^2} + \cdots + \frac{t_{Nk}}{\sigma_k^2} \right)^{-1}\,, \tag{3.4}$$

$$\mathrm{Var}(\widehat{\mu}_{OLS}) = \frac{1}{N} \sum_{i=1}^{k} t_{Ni} \sigma_i^2\,, \tag{3.5}$$

$$\mathrm{Var}(\widehat{\mu}_{FGLS}) = \frac{1}{N} E \left(\frac{t_{N1} \frac{\sigma_1^2}{\widehat{\sigma}_1^4} + t_{N2} \frac{\sigma_2^2}{\widehat{\sigma}_2^4} + \cdots + t_{Nk} \frac{\sigma_k^2}{\widehat{\sigma}_k^4}}{\left(\frac{t_{N1}}{\widehat{\sigma}_1^2} + \frac{t_{N2}}{\widehat{\sigma}_2^2} + \cdots + \frac{t_{Nk}}{\widehat{\sigma}_k^2} \right)^2} \right)\,. \tag{3.6}$$

From Theorem 1 it follows that

$$\lim_{N \to \infty} N^2 c(\mathrm{Var}(\widehat{\mu}_{FGLS}) - \mathrm{Var}(\widehat{\mu}_{OLS})) = 2(k-1) - \sum_{i=1}^{k} t_i \tau_i^2 \tag{3.7}$$

This result shows that when the magnitude of the heteroscedasticity is expressed in terms of the weighted sum of τ_i^2, then $\widehat{\mu}_{FGLS}$ is better than $\widehat{\mu}_{OLS}$ if this weighted sum exceeds $2(k-1)$ and vice versa.

Example 2

Here we specialize our model (1.1) to simple regression model with equal sample sizes $n_1 = \ldots = n_k$.

$$y_{ij} = \alpha + x_{ij}\beta + \varepsilon_{ij}\ , \quad i = 1, 2, \ldots, k, \quad j = 1, 2, \ldots, n. \tag{3.8}$$

Without loss of generality let $\sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij} = 0$ and assume $c = 1$.

Then from (2.8)

$$\begin{aligned} &\lim_{n\to\infty} n^2k^2 \left(\mathrm{Var}\begin{pmatrix} \widehat{\alpha}_{FGLS} \\ \widehat{\beta}_{FGLS} \end{pmatrix} - \mathrm{Var}\begin{pmatrix} \widehat{\alpha}_{OLS} \\ \widehat{\beta}_{OLS} \end{pmatrix} \right) \\ &= \begin{pmatrix} 1 & 0 \\ 0 & \dfrac{k}{\sum_i w_i} \end{pmatrix} \left\{ 2 \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} - \frac{1}{k} \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} \right\} \\ &\quad \times \begin{pmatrix} 1 & 0 \\ 0 & \dfrac{k}{\sum_i w_i} \end{pmatrix}, \end{aligned} \tag{3.9}$$

where

$$\begin{aligned} c_{11} &= k - 1 - \frac{\sum_i \bar{x}_i^2}{\sum_i w_i}, \qquad c_{12} = -\frac{\sum_i w_i \bar{x}_i}{\sum_i w_i}, \\ c_{22} &= \sum_i w_i - \frac{\sum_i \bar{x}_i^2}{k} - \frac{\sum_i w_i^2}{\sum_i w_i}, \\ d_{11} &= \sum_i {\tau_i}^2 - \frac{\left(\sum_i \tau_i \bar{x}_i\right)^2}{\sum_i w_i}, \quad d_{12} = \sum_i {\tau_i}^2 \bar{x}_i - \frac{\left(\sum_i \tau_i \bar{x}_i\right)\left(\sum_i \tau_i w_i\right)}{\sum_i w_i}, \\ d_{22} &= \sum_i {\tau_i}^2 w_i - \frac{\sum_i \tau_i \bar{x}_i^2}{k} - \frac{\left(\sum_i \tau_i w_i\right)^2}{\sum_i w_i}, \end{aligned} \tag{3.10}$$

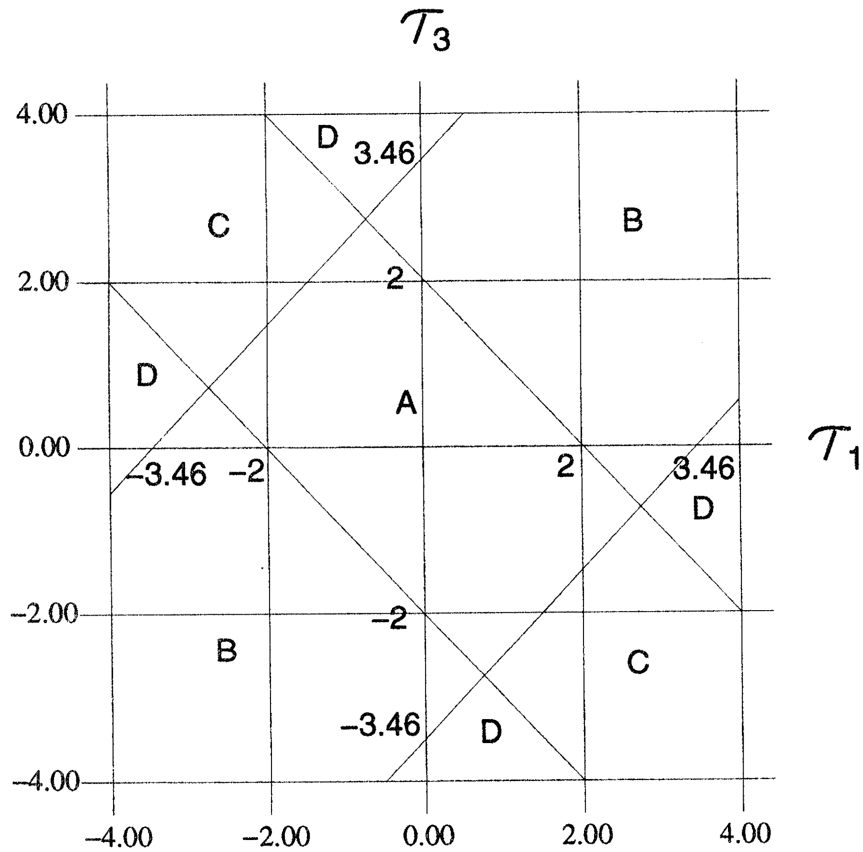
with

$$\bar{x}_i = \frac{\sum_j x_{ij}}{n}, \quad w_i = \frac{\sum_j x_{ij}^2}{n}.$$

We are interested in the sign of the right hand side of (3.9). Let us consider a very simple case $x_{ij} = x_i$, $k = 3$, $x_1 = -1, x_2 = 0, x_3 = 1$, Note that $\tau_2 = -(\tau_1 + \tau_3)$. (3.9) becomes

$$\lim_{n\to\infty} N^2 \left(\mathrm{Var}\begin{pmatrix} \widehat{\alpha}_{FGLS} \\ \widehat{\beta}_{FGLS} \end{pmatrix} - \mathrm{Var}\begin{pmatrix} \widehat{\alpha}_{OLS} \\ \widehat{\beta}_{OLS} \end{pmatrix} \right)$$
$$= \begin{pmatrix} 2 - \frac{1}{2}(\tau_1 + \tau_3)^2 & \frac{(\tau_1{}^2 - \tau_3{}^2)}{4} \\ \frac{(\tau_1{}^2 - \tau_3{}^2)}{4} & \left(\frac{3}{2} - \frac{1}{8}(\tau_1 - \tau_3)^2\right) \end{pmatrix} \tag{3.11}$$

Figure 1



(3.11) shows that the region where $\widehat{\alpha}_{OLS}$ performs better than $\widehat{\alpha}_{FGLS}$ and the region where $\widehat{\beta}_{OLS}$ performs better than $\widehat{\beta}_{FGLS}$ are different. In Figure 1, region A is the region where both $\widehat{\alpha}_{OLS}$ and $\widehat{\beta}_{OLS}$ perform better than $\widehat{\alpha}_{FGLS}$ and $\widehat{\beta}_{FGLS}$ asymptotically. Region D is the region where both $\widehat{\alpha}_{FGLS}$ and $\widehat{\beta}_{FGLS}$ performs better than $\widehat{\alpha}_{OLS}$ and $\widehat{\beta}_{OLS}$. On regions B or C relative performance of OLS and FGLS depends on the component of the estimated regression vector.

4 Numerical results

Example 1

We conducted a Monte Carlo simulation to investigate finite sample behavior of the estimators for the common mean problem with equal sample sizes $n_1 = \ldots = n_k = n$. The entries of the tables below are as follows.

τ_i	values of τ_i
Var(OLS)	$\mathrm{Var}(\hat{\mu}_{OLS})$=a
Var(FGLS)	$\mathrm{Var}(\hat{\mu}_{FGLS})$=b
simulate value	$N^2(b-a)$
theoretical value	equation (3.7)
sd(simulated value)	std. dev. of sim. val.

Variance of FGLS is evaluated by Monte Carlo simulation of replication size 100000 in each case. We take $c = 1$ without loss of generality. $N = nk$ and Var(GLS) = $\mathrm{Var}(\hat{\mu}_{GLS})$ = N^{-1}.

From these tables we see that our theoretical value of equation (3.7) largely agrees with simulated finite sample value. These values tend to differ for large k and asymmetrically distributed $\tau_1, \ldots, \tau_k$.

Case 1: $n = 10$, $k = 2$, Var(GLS) = 0.050000

τ_1	0.0000	1.0000	1.2000	1.3300
τ_2	0.0000	-1.0000	-1.2000	-1.3300
Var(OLS)	0.050000	0.052632	0.053879	0.054851
Var(FGLS)	0.054988	0.054918	0.054885	0.054860
simulated value	1.995105	0.914428	0.402355	0.003442
theoretical value	2.000000	1.000000	0.560000	0.231100
sd(simulated value)	0.000388	0.000414	0.000425	0.000432

τ_1	1.4200	1.6000	1.8000	2.0000
τ_2	-1.4200	-1.6000	-1.8000	-2.0000
Var(OLS)	0.055606	0.057339	0.059666	0.062500
Var(FGLS)	0.054840	0.054795	0.054735	0.054663
simulated value	-0.306361	-1.017642	-1.972266	-3.134989
theoretical value	-0.016400	-0.560000	-1.240000	-2.000000
sd(simulated value)	0.000438	0.000451	0.000465	0.000480

Case 2: $n = 50$, $k = 2$, Var(GLS) = 0.010000

τ_1	0.0000	1.0000	1.2000	1.3300
τ_2	0.0000	-1.0000	-1.2000	-1.3300
Var(OLS)	0.010000	0.010101	0.010146	0.010180
Var(FGLS)	0.010201	0.010199	0.010199	0.010198
simulated value	2.011040	0.984673	0.526464	0.181299
theoretical value	2.000000	1.000000	0.560000	0.231100
sd(simulated value)	0.000087	0.000087	0.000087	0.000086
τ_1	1.4200	1.6000	1.8000	2.0000
τ_2	-1.4200	-1.6000	-1.8000	-2.0000
Var(OLS)	0.010206	0.010263	0.010335	0.010417
Var(FGLS)	0.010198	0.010197	0.010196	0.010194
simulated value	-0.079951	-0.658359	-1.390941	-2.221844
theoretical value	-0.016400	-0.560000	-1.240000	-2.000000
sd(simulated value)	0.000086	0.000086	0.000086	0.000086

Case 3: $n = 100$, $k = 2$, Var(GLS) = 0.005000

τ_1	0.0000	1.0000	1.2000	1.3300
τ_2	0.0000	-1.0000	-1.2000	-1.3300
Var(OLS)	0.005000	0.005025	0.005036	0.005045
Var(FGLS)	0.005050	0.005050	0.005050	0.005050
simulated value	2.003528	0.989561	0.540158	0.202932
theoretical value	2.000000	1.000000	0.560000	0.231100
sd(simulated value)	0.000044	0.000044	0.000044	0.000044
τ_1	1.4200	1.6000	1.8000	2.0000
τ_2	-1.4200	-1.6000	-1.8000	-2.0000
Var(OLS)	0.005051	0.005065	0.005082	0.005102
Var(FGLS)	0.005050	0.005050	0.005049	0.005049
simulated value	-0.051567	-0.612765	-1.319116	-2.114324
theoretical value	-0.016400	-0.560000	-1.240000	-2.000000
sd(simulated value)	0.000044	0.000044	0.000043	0.000043

Case 4: $n = 10$, $k = 3$, Var(GLS) = 0.033333

τ_1	0.0000	1.5500	1.3300	2.8200
τ_2	0.0000	0.9000	1.3300	-1.4100
τ_3	0.0000	-2.4500	-2.6600	-1.4100
Var(OLS)	0.033333	0.038307	0.039483	0.037261
Var(FGLS)	0.038480	0.038308	0.038274	0.038328
simulated value	4.632279	0.001258	-1.088096	0.960768
theoretical value	4.000000	0.928333	0.462200	0.023800
sd(simulated value)	0.000522	0.000539	0.000540	0.000555
τ_1	2.4500	2.9600	3.0000	3.1800
τ_2	0.0000	-1.5300	-1.5000	-1.5900
τ_3	-2.4500	-1.4300	-1.5000	-1.5900
Var(OLS)	0.038892	0.037668	0.037782	0.038342
Var(FGLS)	0.038283	0.038313	0.038309	0.038288
simulated value	-0.547499	0.580732	0.474080	-0.047786
theoretical value	-0.001667	-0.382467	-0.500000	-1.056200
sd(simulated value)	0.000548	0.000558	0.000559	0.000563

Case 5: $n = 50$, $k = 3$, Var(GLS) = 0.006667

τ_1	0.0000	1.5500	1.3300	2.8200
τ_2	0.0000	0.9000	1.3300	-1.4100
τ_3	0.0000	-2.4500	-2.6600	-1.4100
Var(OLS)	0.006667	0.006821	0.006848	0.006829
Var(FGLS)	0.006851	0.006850	0.006849	0.006849
simulated value	4.156257	0.651558	0.032134	0.452625
theoretical value	4.000000	0.928333	0.462200	0.023800
sd(simulated value)	0.000090	0.000090	0.000090	0.000090
τ_1	2.4500	2.9600	3.0000	3.1800
τ_2	0.0000	-1.5300	-1.5000	-1.5900
τ_3	-2.4500	-1.4300	-1.5000	-1.5900
Var(OLS)	0.006852	0.006845	0.006850	0.006872
Var(FGLS)	0.006849	0.006849	0.006849	0.006849
simulated value	-0.064922	0.084488	-0.020589	-0.521382
theoretical value	-0.001667	-0.382467	-0.500000	-1.056200
sd(simulated value)	0.000090	0.000090	0.000090	0.000089

Case 6: $n = 100$, $k = 3$, Var(GLS) = 0.003333

τ_1	0.0000	1.5500	1.3300	2.8200
τ_2	0.0000	0.9000	1.3300	-1.4100
τ_3	0.0000	-2.4500	-2.6600	-1.4100
Var(OLS)	0.003333	0.003370	0.003376	0.003375
Var(FGLS)	0.003379	0.003378	0.003378	0.003378
simulated value	4.068143	0.725555	0.163010	0.319591
theoretical value	4.000000	0.928333	0.462200	0.023800
sd(simulated value)	0.000044	0.000043	0.000043	0.000043
τ_1	2.4500	2.9600	3.0000	3.1800
τ_2	0.0000	-1.5300	-1.5000	-1.5900
τ_3	-2.4500	-1.4300	-1.5000	-1.5900
Var(OLS)	0.003379	0.003379	0.003379	0.003386
Var(FGLS)	0.003378	0.003378	0.003378	0.003378
simulated value	-0.041655	-0.053710	-0.160644	-0.668639
theoretical value	-0.001667	-0.382467	-0.500000	-1.056200
sd(simulated value)	0.000043	0.000043	0.000043	0.000043

Case 7: $n = 10$, $k = 4$, Var(GLS) = 0.025000

τ_1	0.0000	2.3500	1.4200	2.5000
τ_2	0.0000	1.0000	1.3200	2.3700
τ_3	0.0000	0.0000	1.2200	-2.3700
τ_4	0.0000	-3.3500	-3.9600	-2.5000
Var(GLS)	0.025000	0.025000	0.025000	0.025000
Var(OLS)	0.025000	0.029492	0.032231	0.029357
Var(FGLS)	0.029719	0.029557	0.029497	0.029534
simulated value	7.550794	0.102672	-4.375059	0.283693
theoretical value	6.000000	1.563750	0.767800	0.066550
sd(simulated value)	0.000597	0.000599	0.000591	0.000607
τ_1	3.5000	4.6000	2.4900	5.3700
τ_2	0.0000	-1.3000	2.4800	-1.7900
τ_3	-0.0500	-1.5000	-2.4800	-1.7900
τ_4	-3.4500	-1.8000	-2.4900	-1.7900
Var(GLS)	0.025000	0.025000	0.025000	0.025000
Var(OLS)	0.030324	0.028415	0.029564	0.029532
Var(FGLS)	0.029522	0.029558	0.029526	0.029507
simulated value	-1.284075	1.828577	-0.060883	-0.039001
theoretical value	-0.038750	-1.085000	-0.175250	-3.612300
sd(simulated value)	0.000605	0.000622	0.000608	0.000632

Case 8: $n = 50$, $k = 4$, Var(GLS) = 0.005000

τ_1	0.0000	2.3500	1.4200	2.5000
τ_2	0.0000	1.0000	1.3200	2.3700
τ_3	0.0000	0.0000	1.2200	-2.3700
τ_4	0.0000	-3.3500	-3.9600	-2.5000
Var(GLS)	0.005000	0.005000	0.005000	0.005000
Var(OLS)	0.005000	0.005127	0.005166	0.005153
Var(FGLS)	0.005158	0.005157	0.005157	0.005157
simulated value	6.320094	1.183119	-0.376598	0.145962
theoretical value	6.000000	1.563750	0.767800	0.066550
sd(simulated value)	0.000087	0.000087	0.000087	0.000087
τ_1	3.5000	4.6000	2.4900	5.3700
τ_2	0.0000	-1.3000	2.4800	-1.7900
τ_3	-0.0500	-1.5000	-2.4800	-1.7900
τ_4	-3.4500	-1.8000	-2.4900	-1.7900
Var(GLS)	0.005000	0.005000	0.005000	0.005000
Var(OLS)	0.005160	0.005150	0.005159	0.005199
Var(FGLS)	0.005157	0.005156	0.005156	0.005156
simulated value	-0.129651	0.238117	-0.113194	-1.749000
theoretical value	-0.038750	-1.085000	-0.175250	-3.612300
sd(simulated value)	0.000087	0.000087	0.000087	0.000087

Case 9: $n = 100$, $k = 4$, Var(GLS) = 0.002500

τ_1	0.0000	2.3500	1.4200	2.5000
τ_2	0.0000	1.0000	1.3200	2.3700
τ_3	0.0000	0.0000	1.2200	-2.3700
τ_4	0.0000	-3.3500	-3.9600	-2.5000
Var(OLS)	0.002500	0.002530	0.002538	0.002538
Var(FGLS)	0.002539	0.002538	0.002538	0.002538
simulated value	6.162396	1.294817	0.016341	0.107146
theoretical value	6.000000	1.563750	0.767800	0.066550
sd(simulated value)	0.000041	0.000041	0.000041	0.000041
τ_1	3.5000	4.6000	2.4900	5.3700
τ_2	0.0000	-1.3000	2.4800	-1.7900
τ_3	-0.0500	-1.5000	-2.4800	-1.7900
τ_4	-3.4500	-1.8000	-2.4900	-1.7900
Var(OLS)	0.002539	0.002539	0.002539	0.002552
Var(FGLS)	0.002538	0.002538	0.002538	0.002538
simulated value	-0.072745	-0.126023	-0.143164	-2.210645
theoretical value	-0.038750	-1.085000	-0.175250	-3.612300
sd(simulated value)	0.000041	0.000041	0.000041	0.000041

Example 2

Here we evaluate accuracy of (3.11) by Monte Carlo simulation. The replication size of Monte Carlo simulation is again 100000 in each case. The entries of the tables below are as follows.

Var(GLS)	$\text{Var}(\widehat{\alpha}_{GLS})$	$\text{Var}(\widehat{\beta}_{GLS})$
Var(OLS)	$\text{Var}(\widehat{\alpha}_{OLS})=e$	$\text{Var}(\widehat{\beta}_{OLS})=f$
Var(FGLS)	$\text{Var}(\widehat{\alpha}_{FGLS})=g$	$\text{Var}(\widehat{\beta}_{FGLS})=h$
simulated value	$N^2(g-e)$	$N^2(h-f)$
theoretical value	(1,1) element of (3.11)	(2,2) element of (3.11)
sd(simulated value)	std. dev. of sim. val.	std. dev. of sim. val.

where k=3, $N = nk = 3n$. We again see that our asymptotic theoretical values largely agree with simulated finite sample values.

Case 1: region A, $\tau_1 = 0.0000, \tau_2 = 0.0000, \tau_3 = 0.0000$

	n=50		n=100	
Var(GLS)	0.00666667	0.01000000	0.00333333	0.00500000
Var(OLS)	0.00666667	0.01000000	0.00333333	0.00500000
Var(FGLS)	0.00675942	0.01007077	0.00335597	0.00501712
simulated value	2.08704296	1.59233009	2.03736621	1.54122873
theoretical value	2.00000000	1.50000000	2.00000000	1.50000000
sd(simulated value)	0.00000559	0.00002590	0.00001411	0.00001542

Case 2: region A, $\tau_1 = 2.7200, \tau_2 = -1.9900, \tau_3 = -0.7300$

	n=50		n=100	
Var(GLS)	0.00674922	0.00936315	0.00335431	0.00475813
Var(OLS)	0.00683480	0.00940827	0.00337565	0.00477069
Var(FGLS)	0.00683667	0.00942107	0.00337588	0.00477284
simulated value	0.04208046	0.28800368	0.02125421	0.19390237
theoretical value	0.01995000	0.01218750	0.01995000	0.01218750
sd(simulated value)	0.00000390	0.00001719	0.00001331	0.00001208

Case 3: region A, $\tau_1 = 1.7300, \tau_2 = 0.0000, \tau_3 = -1.7300$

	n=50		n=100	
Var(GLS)	0.00675654	0.01013481	0.00335565	0.00503348
Var(OLS)	0.00675715	0.01020359	0.00335573	0.00505038
Var(FGLS)	0.00685164	0.01020873	0.00337858	0.00505099
simulated value	2.12593985	0.11572405	2.05697351	0.05477480
theoretical value	2.00000000	0.00355000	2.00000000	0.00355000
sd(simulated value)	0.00000218	0.00002139	0.00001502	0.00001432

Case 4: region B, $\tau_1 = -1.0900, \tau_2 = 2.1800, \tau_3 = -1.0900$

	n=50		n=100	
Var(GLS)	0.00666667	0.01097693	0.00333333	0.00533579
Var(OLS)	0.00676508	0.01097693	0.00335836	0.00533579
Var(FGLS)	0.00676480	0.01106741	0.00335709	0.00535628
simulated value	-0.00632759	2.03590043	-0.11418712	1.84431133
theoretical value	-0.37620000	1.50000000	-0.37620000	1.50000000
sd(simulated value)	0.00000416	0.00003272	0.00001572	0.00001843

Case 5: region B, $\tau_1 = 1.0100, \tau_2 = -2.0200, \tau_3 = 1.0100$

	n=50		n=100	
Var(GLS)	0.00666667	0.00923816	0.00333333	0.00472450
Var(OLS)	0.00676698	0.00923816	0.00335758	0.00472450
Var(FGLS)	0.00675193	0.00929340	0.00335462	0.00473885
simulated value	-0.33858490	1.24291344	-0.26659246	1.29166127
theoretical value	-0.04020000	1.50000000	-0.04020000	1.50000000
sd(simulated value)	0.00000660	0.00002044	0.00001250	0.00001294

Case 6: region C, $\tau_1 = -1.8000, \tau_2 = 0.0000, \tau_3 = 1.8000$

	n=50		n=100	
Var(GLS)	0.00676407	0.01014610	0.00335751	0.00503626
Var(OLS)	0.00676479	0.01022077	0.00335760	0.00505459
Var(FGLS)	0.00685934	0.01022030	0.00338046	0.00505381
simulated value	2.12752990	-0.01056520	2.05799261	-0.07034169
theoretical value	2.00000000	-0.12000000	2.00000000	-0.12000000
sd(simulated value)	0.00000947	0.00003168	0.00001342	0.00001656

Case 7: region C, $\tau_1 = 1.8000, \tau_2 = 0.0000, \tau_3 = -1.8000$

	n=50		n=100	
Var(GLS)	0.00676407	0.01014610	0.00335751	0.00503626
Var(OLS)	0.00676479	0.01022077	0.00335760	0.00505459
Var(FGLS)	0.00685936	0.01022029	0.00338046	0.00505381
simulated value	2.12793963	-0.01076962	2.05799542	-0.07020994
theoretical value	2.00000000	-0.12000000	2.00000000	-0.12000000
sd(simulated value)	0.00000204	0.00002122	0.00001506	0.00001428

Case 8: region D, $\tau_1 = 3.1400, \tau_2 = -2.3000, \tau_3 = -0.8400$

	n=50		n=100	
Var(GLS)	0.00677569	0.00929112	0.00336107	0.00472771
Var(OLS)	0.00689064	0.00934787	0.00338959	0.00474376
Var(FGLS)	0.00686240	0.00934755	0.00338249	0.00474212
simulated value	-0.63546419	-0.00720852	-0.63821766	-0.14711869
theoretical value	-0.64500000	-0.48005000	-0.64500000	-0.48005000
sd(simulated value)	0.00000374	0.00001613	0.00001317	0.00001162

Case 9: region D, $\tau_1 = 0.0000, \tau_2 = 3.4700, \tau_3 = -3.4700$

	n=50		n=100	
Var(GLS)	0.00677222	0.01183489	0.00335830	0.00559822
Var(OLS)	0.00705457	0.01197666	0.00342625	0.00562633
Var(FGLS)	0.00687500	0.01194347	0.00338289	0.00562144
simulated value	-4.04039015	-0.74675757	-3.90314305	-0.43995882
theoretical value	-4.02045000	-0.00511250	-4.02045000	-0.00511250
sd(simulated value)	0.00000158	0.00003001	0.00001774	0.00001886

5 Proofs

Let

$$r_i = \frac{\sigma_i^2}{\hat{\sigma}_i^2} - \frac{\nu_i}{\nu_i - 2}, \quad \nu_i = n_i - d_i . \tag{5.1}$$

Note that r_i are distributed as $\nu_i/\chi^2(\nu_i) - \nu_i/(\nu_i - 2)$ and of the order $O_p(N^{-\frac{1}{2}})$. We work with the reciprocal of the χ^2 random variable rather than χ^2 random variable itself because validity of the asymptotic expansion is more easily established. The moments of r_i are easily computed as

$$\begin{aligned} E(r_i) &= 0 , \\ \mathrm{Var}(r_i) &= \frac{2\nu_i^2}{(\nu_i - 2)^2(\nu_i - 4)} = O(N^{-1}) , \\ E(r_i^h) &= O(N^{-2}), \quad h \geq 3 . \end{aligned} \tag{5.2}$$

(5.1) can be expressed as

$$\frac{1}{\hat{\sigma}_i^2} = \frac{1}{\sigma_i^2}\left(\frac{\nu_i}{\nu_i - 2} + r_i\right) = \eta_{Ni}(s_{Ni} + r_i) , \tag{5.3}$$

where $s_{Ni} = \nu_i/(\nu_i - 2)$ which converges to 1 as $N \to \infty$.

Proof of the Lemma 1

In this paper we evaluate variances to the order $O(N^{-2})$. The variance of $\widehat{\beta}_{FGLS}$ is

$$
\begin{aligned}
&\mathrm{Var}(\widehat{\beta}_{FGLS}) \\
&= \frac{1}{N} E\left(\left(\sum \frac{t_{Ni}A_{Ni}}{\widehat{\sigma}_i^2} \right)^{-1} \sum \frac{t_{Ni}\sigma_i^2 A_{Ni}}{\widehat{\sigma}_i^4} \left(\sum \frac{t_{Ni}A_{Ni}}{\widehat{\sigma}_i^2} \right)^{-1} \right) \\
&= N^{-1} E\left\{ \left(\sum t_{Ni}\eta_{Ni}(s_{Ni}+r_i)A_{Ni} \right)^{-1} \sum t_{Ni}\eta_{Ni}(s_{Ni}+r_i)^2 A_{Ni} \right. \\
&\qquad \left. \times \left(\sum t_{Ni}\eta_{Ni}(s_{Ni}+r_i)A_{Ni} \right)^{-1} \right\} . \qquad (5.4)
\end{aligned}
$$

We expand the right hand side of (5.4) in Taylor series as in Zellner (1962). The first term can be expressed as

$$
\begin{aligned}
&\left(\sum t_{Ni}\eta_{Ni}(s_{Ni}+r_i)A_{Ni} \right)^{-1} \\
&\quad = \left(\sum t_{Ni}\eta_{Ni}s_{Ni}A_{Ni} + \sum t_{Ni}\eta_{Ni}r_iA_{Ni} \right)^{-1} \\
&\quad = (B_{N0} + B_{N1})^{-1} \\
&\quad = B_{N0}^{-1} - B_{N0}^{-1}B_{N1}B_{N0}^{-1} + B_{N0}^{-1}B_{N1}B_{N0}^{-1}B_{N1}B_{N0}^{-1} + R_1 , \qquad (5.5)
\end{aligned}
$$

where

$$
B_{N0} = \sum t_{Ni}\eta_{Ni}s_{Ni}A_{Ni} , \quad B_{N1} = \sum t_{Ni}\eta_{Ni}r_iA_{Ni} . \qquad (5.6)
$$

$$
R_1 = -B_{N0}^{-1}(B_{N1}B_{N0}^{-1})^3(I + B_{N1}B_{N0}^{-1})^{-1} , \qquad (5.7)
$$

and the second term can be expressed

$$
\sum t_{Ni}\eta_{Ni}(s_{Ni}+r_i)^2A_{Ni} = \tilde{B}_{N0} + B_{N2} , \qquad (5.8)
$$

where

$$
\tilde{B}_{N0} = \sum t_{Ni}\eta_{Ni}s_{Ni}^2A_{Ni} , \quad B_{N2} = \sum t_{Ni}\eta_{Ni}(2s_{Ni}r_i + r_i^2)A_{Ni} . \qquad (5.9)
$$

Note that B_{N0}, $\tilde{B}_{N0}$ are constant matrices and B_{N1}, B_{N2} are random matrices.

The complicated part is $(I + B_{N1}B_{N0}^{-1})^{-1}$. However we can show that

$$
E(B_{N0}^{-1}(B_{N1}B_{N0}^{-1})^3(I + B_{N1}B_{N0}^{-1})^{-1}) = O(N^{-\frac{3}{2}}) . \qquad (5.10)
$$

Validity of (5.10) is proved below in a rigorous manner. Therefore for large N it suffices to calculate

$$
\begin{aligned}
&N^{-1}E\left\{ (B_{N0}^{-1} - B_{N0}^{-1}B_{N1}B_{N0}^{-1} + B_{N0}^{-1}B_{N1}B_{N0}^{-1}B_{N1}B_{N0}^{-1})(\tilde{B}_{N0} + B_{N2}) \right. \\
&\quad \left. \times (B_{N0}^{-1} - B_{N0}^{-1}B_{N1}B_{N0}^{-1} + B_{N0}^{-1}B_{N1}B_{N0}^{-1}B_{N1}B_{N0}^{-1}) \right\} . \qquad (5.11)
\end{aligned}
$$

Expanding (5.11) and taking expectation term by term we obtain

$$\begin{aligned}
&\mathrm{Var}\,(\widehat{\beta}_{FGLS}) \qquad\qquad (5.12)\\
&= N^{-1}E(B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1} - B_{N0}^{-1}B_{N1}B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1} - B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1}B_{N1}B_{N0}^{-1}\\
&\quad + B_{N0}^{-1}B_{N2}B_{N0}^{-1} + B_{N0}^{-1}B_{N1}B_{N0}^{-1}B_{N1}B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1}\\
&\quad + B_{N0}^{-1}B_{N1}B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1}B_{N1}B_{N0}^{-1} + B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1}B_{N1}B_{N0}^{-1}B_{N1}B_{N0}^{-1}\\
&\quad - B_{N0}^{-1}B_{N1}B_{N0}^{-1}B_{N2}B_{N0}^{-1} - B_{N0}^{-1}B_{N2}B_{N0}^{-1}B_{N1}B_{N0}^{-1}) + o(N^{-2})\\
&= \frac{1}{N}\left\{B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1} + 2\sum \frac{t_{Ni}\eta_{Ni}\nu_i^2}{(\nu_i-2)^2(\nu_i-4)}B_{N0}^{-1}A_{Ni}B_{N0}^{-1}\right.\\
&\quad + 6\sum \frac{t_{Ni}^2\eta_{Ni}^2\nu_i^2}{(\nu_i-2)^2(\nu_i-4)}B_{N0}^{-1}A_{Ni}B_{N0}^{-1}A_{Ni}B_{N0}^{-1}\\
&\quad \left. - 8\sum \frac{t_{Ni}^2\eta_{Ni}^2\nu_i^2}{(\nu_i-2)^2(\nu_i-4)}B_{N0}^{-1}A_{Ni}B_{N0}^{-1}A_{Ni}B_{N0}^{-1}\right\} + o(N^{-2})\\
&= \frac{1}{N}\left(B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1} + 2\sum \frac{t_{Ni}\eta_{Ni}}{\nu_i}B_{N0}^{-1}A_{Ni}B_{N0}^{-1}\right.\\
&\quad \left. - 2\sum \frac{t_{Ni}^2\eta_{Ni}^2}{\nu_i}B_{N0}^{-1}A_{Ni}B_{N0}^{-1}A_{Ni}B_{N0}^{-1}\right) + o(N^{-2})\\
&= \frac{1}{N}\left(B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1} + 2\sum \frac{\eta_{Ni}}{N}B_{N0}^{-1}A_{Ni}B_{N0}^{-1}\right.\\
&\quad \left. - 2\sum \frac{t_{Ni}\eta_{Ni}^2}{N}B_{N0}^{-1}A_{Ni}B_{N0}^{-1}A_{Ni}B_{N0}^{-1}\right) + o(N^{-2})\,. \qquad ∎ \qquad (5.13)
\end{aligned}$$

In the above derivation we used the following lemma.

Lemma 2

$$\begin{aligned}
E(B_{N1}B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1}B_{N1}) &= E(B_{N1}B_{N0}^{-1}B_{N1}B_{N0}^{-1}\tilde{B}_{N0})\\
&= E(\tilde{B}_{N0}B_{N0}^{-1}B_{N1}B_{N0}^{-1}B_{N1})\\
&= 2\sum \frac{t_{Ni}^2\eta_{Ni}^2\nu_i^2}{(\nu_i-2)^2(\nu_i-4)}A_{Ni}B_{N0}^{-1}A_{Ni}\\
&\quad + o(N^{-1})\,, \qquad (5.14)\\
E(B_{N1}B_{N0}^{-1}B_{N2}) &= E(B_{N2}B_{N0}^{-1}B_{N1})\\
&= 4\sum \frac{t_{Ni}^2\eta_{Ni}^2\nu_i^2}{(\nu_i-2)^2(\nu_i-4)}A_{Ni}B_{N0}^{-1}A_{Ni}\\
&\quad + o(N^{-1})\,. \qquad (5.15)
\end{aligned}$$

Proof of the Lemma 2

We prove only one case of (5.13) and (5.15). Other cases are similarly proved. From (5.6)

$$\begin{aligned} \tilde{B}_{N0} - B_{N0} &= \sum t_{Ni}\eta_{Ni}(s_{Ni}^2 - s_{Ni})A_{Ni} \\ &= \sum t_{Ni}\eta_{Ni}\frac{2\nu_i}{(\nu_i - 2)^2}A_{Ni} = O(N^{-1}). \end{aligned} \tag{5.16}$$

Therefore

$$\begin{aligned} &E(B_{N1}B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1}B_{N1}) \\ &\quad = E\left(\sum t_{Ni}\eta_{Ni}r_i A_{Ni}B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1}\sum t_{Ni}\eta_{Ni}r_i A_{Ni}\right) \\ &\quad = \sum t_{Ni}^2\eta_{Ni}^2 E(r_i^2 A_{Ni}B_{N0}^{-1}A_{Ni}) \\ &\qquad + \sum t_{Ni}^2\eta_{Ni}^2 E(r_i^2 A_{Ni}B_{N0}^{-1}(\tilde{B}_{N0} - B_{N0})B_{N0}^{-1}A_{Ni}) \\ &\quad = 2\sum \frac{t_{Ni}^2\eta_{Ni}^2\nu_i^2}{(\nu_i - 2)^2(\nu_i - 4)}A_{Ni}B_{N0}^{-1}A_{Ni} + o(N^{-1}), \end{aligned}$$

$$\begin{aligned} E(B_{N1}B_{N0}^{-1}B_{N2}) &= E\left(\sum t_{Ni}\eta_{Ni}r_i A_{Ni}B_{N0}^{-1}\sum t_{Ni}\eta_{Ni}(2s_{Ni}r_i + r_i^2)A_{Ni}\right) \\ &= \sum t_{Ni}^2\eta_{Ni}^2 E(2s_{Ni}r_i^2 A_{Ni}B_{N0}^{-1}A_{Ni}) + o(N^{-1}) \\ &= 4\sum \frac{t_{Ni}^2\eta_{Ni}^2\nu_i^2}{(\nu_i - 2)^2(\nu_i - 4)}A_{Ni}B_{N0}^{-1}A_{Ni} + o(N^{-1}). \end{aligned}$$

∎

Now we give a proof of Theorem 1

Proof of the Theorem 1

From (1.11) and (5.6) B_{N0}^{-1} can be expressed as

$$\begin{aligned} B_{N0}^{-1} &= \frac{1}{c}\left(\sum t_{Ni}s_{Ni}A_{Ni} + \frac{1}{\sqrt{N}}\sum t_{Ni}\tau_i s_{Ni}A_{Ni}\right)^{-1} \\ &= \frac{1}{c}\left(\sum t_{Ni}s_{Ni}A_{Ni}\right)^{-1} + o(N^{-1}) \\ &= \frac{1}{c}\left(\sum t_{Ni}A_{Ni}\right)^{-1} + o(N^{-1}). \end{aligned} \tag{5.17}$$

Denote $\bar{B}_{N0} = \sum t_{Ni}\eta_{Ni}A_{Ni}$. Then

$$\begin{aligned} &\mathrm{Var}(\widehat{\beta}_{FGLS}) - \mathrm{Var}(\widehat{\beta}_{GLS}) \\ &\quad = \frac{1}{N}\left(B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1} - \bar{B}_{N0}^{-1}\right) \\ &\quad + \frac{2}{N^2}\left\{\sum \eta_{Ni}B_{N0}^{-1}A_{Ni}B_{N0}^{-1} - \sum t_{Ni}\eta_{Ni}^2 B_{N0}^{-1}A_{Ni}B_{N0}^{-1}A_{Ni}B_{N0}^{-1}\right\} \\ &\quad + o(N^{-2}) \end{aligned}$$

$$
\begin{aligned}
&= \frac{2}{N^2c^2}\Big\{\sum \eta_{Ni}\left(\sum t_{Ni}A_{Ni}\right)^{-1} A_{Ni}\left(\sum t_{Ni}A_{Ni}\right)^{-1} \\
&\quad - \sum t_{Ni}\eta_{Ni}^2\left(\sum t_{Ni}A_{Ni}\right)^{-1} A_{Ni}\left(\sum t_{Nj}A_{Nj}\right)^{-1} A_{Ni}\left(\sum t_{Ni}A_{Ni}\right)^{-1}\Big\} \\
&\quad + o(N^{-2}) \\
&= \frac{2}{N^2c}\left(\sum t_iA_i\right)^{-1}\left(\sum A_i - \sum t_iA_i\left(\sum t_jA_j\right)^{-1} A_i\right)\left(\sum t_iA_i\right)^{-1} \\
&\quad + o(N^{-2})\,. \qquad (5.18)
\end{aligned}
$$

The second equality holds because $B_{N0}^{-1}\tilde{B}_{N0}B_{N0}^{-1} - \bar{B}_{N0}^{-1} = O(N^{-2})$, which follows directly from the following relations.

$$
\begin{aligned}
B_{N0} &= \bar{B}_{N0} + 2\sum t_{Ni}\eta_{Ni}\frac{1}{\nu_i - 2}A_{Ni}\,, \\
B_{N0}^{-1} &= \bar{B}_{N0}^{-1} - 2\bar{B}_{N0}^{-1}\sum t_{Ni}\eta_{Ni}\frac{1}{\nu_i - 2}A_{Ni}\bar{B}_{N0}^{-1} + O(N^{-2})\,, \\
\tilde{B}_{N0} &= \bar{B}_{N0} + 4\sum t_{Ni}\eta_{Ni}\frac{\nu_i - 1}{(\nu_i - 2)^2}A_{Ni}\,. \qquad (5.19)
\end{aligned}
$$

On the other hand $\bar{B}_{N0}^{-1}$ is expanded as

$$
\begin{aligned}
\bar{B}_{N0}^{-1} &= \frac{1}{c}\left(\sum t_{Ni}A_{Ni} + \frac{1}{\sqrt{N}}\sum t_{Ni}\tau_iA_{Ni}\right)^{-1} \\
&= \frac{1}{c}\left(\sum t_{Ni}A_{Ni}\right)^{-1} - \frac{1}{c\sqrt{N}}\left(\sum t_{Ni}A_{Ni}\right)^{-1}\sum t_{Ni}\tau_iA_{Ni}\left(\sum t_{Ni}A_{Ni}\right)^{-1} \\
&\quad + \frac{1}{c}\left(\sum t_{Ni}A_{Ni}\right)^{-1}\frac{1}{\sqrt{N}}\sum t_{Ni}\tau_iA_{Ni}\left(\sum t_{Ni}A_{Ni}\right)^{-1} \\
&\quad \times \frac{1}{\sqrt{N}}\sum t_{Ni}\tau_iA_{Ni}\left(\sum t_{Ni}A_{Ni}\right)^{-1} + R_2\,, \qquad (5.20)
\end{aligned}
$$

where

$$
\begin{aligned}
R_2 &= \frac{1}{c}\left(\sum t_{Ni}A_{Ni}\right)^{-1}\left(\frac{1}{\sqrt{N}}\sum t_{Ni}\tau_iA_{Ni}\left(\sum t_{Ni}A_{Ni}\right)^{-1}\right)^3 \\
&\quad \times \left(I - \frac{1}{\sqrt{N}}\sum t_{Ni}\tau_iA_{Ni}\left(\sum t_{Ni}A_{Ni}\right)^{-1}\right)^{-1} \\
&= o(N^{-1})\,. \qquad (5.21)
\end{aligned}
$$

Hence

$$
\mathrm{Var}(\widehat{\beta}_{OLS}) - \mathrm{Var}(\widehat{\beta}_{GLS})
$$

$$
\begin{aligned}
&= \frac{1}{N}\left\{\left(\sum t_{Ni}A_{Ni}\right)^{-1}\sum \frac{t_{Ni}A_{Ni}}{c\left(1+\frac{\tau_i}{\sqrt{N}}\right)}\left(\sum t_{Ni}A_{Ni}\right)^{-1} - \bar{B}_{N0}^{-1}\right\} \\
&= \frac{1}{Nc}\left\{\left(\sum t_{Ni}A_{Ni}\right)^{-1}\sum\left(1-\frac{\tau_i}{\sqrt{N}}+\frac{\tau_i^2}{N}\right)t_{Ni}A_{Ni}\left(\sum t_{Ni}A_{Ni}\right)^{-1}\right. \\
&\quad - \left(\sum t_{Ni}A_{Ni}\right)^{-1} + \frac{1}{\sqrt{N}}\left(\sum t_{Ni}A_{Ni}\right)^{-1}\sum t_{Ni}\tau_i A_{Ni}\left(\sum t_{Ni}A_{Ni}\right)^{-1} \\
&\quad - \left(\sum t_{Ni}A_{Ni}\right)^{-1}\frac{1}{\sqrt{N}}\sum t_{Ni}\tau_i A_{Ni}\left(\sum t_{Ni}A_{Ni}\right)^{-1} \\
&\quad \left.\times\frac{c}{\sqrt{N}}\sum t_{Ni}\tau_i A_{Ni}\left(c\sum t_{Ni}A_{Ni}\right)^{-1}\right\} + o(N^{-2}) \\
&= \frac{1}{N^2c}\left(\sum t_iA_i\right)^{-1}\left\{\sum t_i\tau_i^2A_i - \sum t_i\tau_iA_i\left(\sum t_iA_i\right)^{-1}\sum t_i\tau_iA_i\right\} \\
&\quad \times\left(\sum t_iA_i\right)^{-1} + o(N^{-2}) . \qquad (5.22)
\end{aligned}
$$

Substituting (5.22) from (5.18) we obtain

$$
\begin{aligned}
&\mathrm{Var}(\widehat{\beta}_{FGLS}) - \mathrm{Var}(\widehat{\beta}_{OLS}) \\
&\quad = -\frac{1}{N^2c}\left(\sum t_iA_i\right)^{-1}\left\{2\left(\sum A_i - \sum t_iA_i\left(\sum t_jA_j\right)^{-1}A_i\right)\right. \\
&\quad - \left.\left(\sum t_i\tau_i^2A_i - \sum t_i\tau_iA_i\left(\sum t_iA_i\right)^{-1}\sum t_i\tau_iA_i\right)\right\}\left(\sum t_iA_i\right)^{-1} \\
&\quad + o(N^{-2}) . \qquad (5.23)
\end{aligned}
$$

■

Finally we establish the validity of the asymptotic expansions.

Proof of the validity.

The argument above can be made rigorous, by proving the validity of the asymptotic expansion. Here we prove the validity of an essential step (5.10). The validity of other terms can be shown similarly. We need to show

$$
E\left|B_{N0}^{-1}(B_{N1}B_{N0}^{-1})^3(I+B_{N1}B_{N0}^{-1})^{-1}\right| = O(N^{-\frac{3}{2}}) , \qquad (5.24)
$$

where

$$
|A| = \sum_i\sum_j |a_{ij}| .
$$

From (5.4), (5.6), $(B_{N0}+B_{N1})^{-1} = (\sum t_{Ni}A_{Ni}/\widehat{\sigma}_i^2)^{-1}$. From Cramer formula the elements of $(B_{N0}+B_{N1})^{-1}$ can be expressed as

$$\left(\sum_l \frac{t_{Nl}A_{Nl}}{\widehat{\sigma}_l^2}\right)^{-1}_{ij} = \frac{\mathrm{Adj}\left(\sum_l \frac{t_{Nl}A_{Nl}}{\widehat{\sigma}_l^2}\right)_{ij}}{\left|\sum_l \frac{t_{Nl}A_{Nl}}{\widehat{\sigma}_l^2}\right|}, \tag{5.25}$$

where Adj denotes the cofactor of a matrix. The numerator is a polynomial of degree p in $(\widehat{\sigma}_i^2)^{-1}$'s and the denominator is a polynomial of degree $p+1$ in $(\widehat{\sigma}_i^2)^{-1}$'s. Hence there exists a constant $c_{Nij}>0$ such that

$$\frac{\mathrm{Adj}\left(\sum_l \frac{t_{Nl}A_{Nl}}{\widehat{\sigma}_l^2}\right)_{ij}}{\left|\sum_l \frac{t_{Nl}A_{Nl}}{\widehat{\sigma}_l^2}\right|} \le c_{Nij}\frac{\max_l \widehat{\sigma}_l^{2p+2}}{\min_m \widehat{\sigma}_m^{2p}} \le c_{Nij}\sum_{l,m}\frac{\widehat{\sigma}_l^{2p+2}}{\widehat{\sigma}_m^{2p}}. \tag{5.26}$$

Furthermore $\lim_{N\to\infty} c_{Nij} = c_{ij}$ where c_{ij} is a positive constant. Denoting $(B_{N0}^{-1}B_{N1})_{ij} = (b_{ij})$ it suffices to show that

$$E\left(|b_{i_1j_1}||b_{i_2j_2}||b_{i_3j_3}|\, c_{Ni_4j_4}\sum_{l,m}\frac{\widehat{\sigma}_l^{2p+2}}{\widehat{\sigma}_m^{2p}}\right) = O(N^{-\frac{3}{2}}). \tag{5.27}$$

for any $i_1,\ldots,i_4,j_1,\ldots,j_4$. But (5.27) follows easily from Cauchy-Schwarz inequality. ∎

References

[1] Amemiya, T., *Advanced Econometrics.* Blackwell (1985).

[2] Judge, G.G., Hill, R., Griffiths, W., Lütkepohl, H., and Lee, T., *Introduction to the theory and practice of econometrics.* Wiley, New York (1982).

[3] Rothenberg, T.J., Approximate Normality of Generalized Least Squares Estimates. *Econometrica* **52** (1984), 811-825.

[4] Taylor, W.E., Small Sample Properties of a Class of Two Stage Aitken Estimators. *Econometrica* **45** (1977), 497-508.

[5] Taylor, W.E., The Heteroscedastic Linear Model: Exact Finite Sample Results. *Econometrica* **46** (1978), 663-675.

[6] Theil, H., *Principles of Econometrics.* Wiley, New York (1971).

[7] Zellner, A., An Efficient Method of Estimating Seemingly Unrelated Regression and Tests for Aggregation Bias. *Journal of the American Statistical Association* **57** (1962), 348-368.