AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Report AE10/77

*Amsterdam University Institute of actuarial
science and econometrics*

On the Unbiasedness of Iterated GLS Estimators

by F.J.H. Don

J.R. Magnus

August 1977

Instituut voor Actuariaat & Econometrie

Classification: AMS(MOS) subject classification scheme (1970):
15A51, 62J05

Keywords: Iterated GLS, unbiasedness, maximum likelihood

Abstract: In 1967 Kakwani showed that any estimator which has a sampling error of the form $H(\epsilon)\epsilon$ is unbiased, provided that the estimator has a finite expectation, H is even, and ϵ symmetrically distributed about zero. Fuller and Battese (1973) used this result to establish unbiasedness of the GLS estimator in a variance components model, and Mehta and Swamy (1976) did the same for a variant of Zellner's two-step estimator under normality.

In this paper we prove that Kakwani's result is applicable to the iterated GLS estimators at each step of the iteration, and conditions for the finiteness of the expectations are derived, implying unbiasedness of the iterated GLS estimators under certain stated conditions.

Assuming normal disturbances, unbiasedness of the maximum likelihood estimator is proved under assumptions earlier made by Oberhofer and Kmenta (1974) to guarantee convergence of the iteration procedure to the solution of the ML-equations.

GIANNINI FOUNDATION OF
AGRICULTURAL ECONOMICS

AUG 21 1978

Free copies are obtainable from:

Bibliotheek - Instituut voor Actuariat en Econometrie

Jodenbreestraat 23 - Amsterdam 'phone: 020-5254287/17

Universiteit van Amsterdam

On the Unbiasedness of Iterated GLS Estimators

By F.J. Henk Don and Jan R. Magnus*

August 1977

1. Introduction

In 1967 Kakwani showed that any estimator which has a sampling error of the form $H(\epsilon) \cdot \epsilon$ is unbiased, provided that the estimator has a finite expectation, H is even¹⁾, and ϵ is symmetrically distributed about zero. This result was subsequently used to establish that Zellner's two-step estimator for seemingly unrelated regressions, and Theil's mixed regression estimator are unbiased, if the expectations exist (Kakwani (1967, 1968)). Swamy and Mehta (1969) proved the existence of the expectation of the mixed regression estimator under normality, thus establishing its unbiasedness. In 1973 Fuller and Battese provided an elegant derivation of conditions under which the expectation of the GLS estimator in a variance components model exists, and they proved unbiasedness under normality. Finally, Mehta and Swamy (1976) studied a variant of Zellner's two-step estimator and they established its unbiasedness under normality.

In this paper we shall consider the iterated GLS estimators

$$b_j = (X' \Omega_j^{-1} X)^{-1} X' \Omega_j^{-1} y \quad (j=0, 1, \dots),$$

where Ω_j is estimated as some function of the residuals from the preceding step. Proposition 1 states that Kakwani's result holds at each step of the iteration. In the second proposition we formulate sufficient conditions for the existence of $E b_j$. The two propositions combined say that $E b_j = \beta$, if ϵ is symmetrical about zero, Ω_j is an even function of the residuals from the preceding step, and $E(\text{tr} \Omega_j)(\text{tr} \Omega_j^{-1})$ exists.

Next, we assume normal disturbances. It is then natural to ask whether the maximum likelihood estimator $\hat{\beta}$ is unbiased. We shall prove that this is the case, provided that Oberhofer and Kmenta's conditions, under which the iteration converges to a solution of the maximum likelihood equations, are satisfied.

*) Instituut voor Actuarialaats en Econometrie, Universiteit van Amsterdam, 23 Jodenbreestraat, Amsterdam, the Netherlands. We gratefully acknowledge the helpful suggestions of R.D.H. Heijmans.

1) H is an even function of ϵ if H is invariant under a change of sign of all elements in ϵ , that is $H(\epsilon) = H(-\epsilon)$.

2. Two useful propositions

We shall consider the linear regression model

$$y = X\beta + \epsilon,$$

where y is an $(n,1)$ vector of observations on the dependent variable, X is an (n,k) matrix of the values of the regressors, β is a $(k,1)$ vector of regression coefficients, and ϵ is an $(n,1)$ disturbance vector. We make the following assumptions:

ASSUMPTION 1: $E\epsilon = 0$, $E\epsilon\epsilon' = \Omega$, where Ω is a positive definite matrix whose elements are differentiable functions of a finite and constant number of parameters $\theta_1 \dots \theta_m$.

ASSUMPTION 2: X is a fixed matrix of full rank and $n > k$.

ASSUMPTION 3: The parameters in β are independent from those in θ , where $\theta = (\theta_1 \dots \theta_m)'$.

In order to estimate the parameter vectors β and θ , we choose some starting value θ_0 for θ , and apply an iteration scheme whose j th step ($j=0,1,\dots$) is defined by

$$\begin{aligned}\Omega_j^{-1} &= \Omega^{-1}(\theta_j), \\ b_j &= (X'\Omega_j^{-1}X)^{-1}X'\Omega_j^{-1}y, \\ e_j &= y - Xb_j, \\ \theta_{j+1} &= \theta(e_j).\end{aligned}$$

Thus b_j is estimated by GLS based upon the estimated θ_j , which is estimated as a function - for the moment unspecified - of the residual vector e_{j-1} .

proposition 1

If ϵ is symmetrical about zero and θ_j is estimated as an even function of the residuals e_{j-1} , then b_j is symmetrical about β at each step of the iteration.

proof 2)

At the j -th step we have

$$e_j = y - Xb_j = y - X(X'\Omega_j^{-1}X)^{-1}X'\Omega_j^{-1}y = [I - X(X'\Omega_j^{-1}X)^{-1}X'\Omega_j^{-1}]\epsilon \quad (j=0,1,\dots).$$

Following Kakwani's (1967) approach we see that, if ϵ changes sign, e_0 will change sign, but θ_1 - being an even function of e_0 - will not be affected.

2) A similar proposition in the context of maximum likelihood is proved in Magnus (lemmas 3 and 4).

Ω_1 will not be affected either. Thus Ω_1 is an even function of ϵ . This implies that b_1 is symmetrical about β . It also implies that e_1 changes sign if ϵ changes sign. Since θ_2 is an even function of e_1 , it is an even function of ϵ . Therefore b_2 is symmetrical about β . Clearly, the argument can be repeated for $j=3,4,\dots$. ▲

In other words, b_j is unbiased *if its mean exists*. Sufficient conditions for the existence of Eb_j are the subject of proposition 2. First, we will state the following lemma which will prove useful in the sequel.

Lemma 1

Let Ω be some positive semi-definite matrix of random variables, then the following statements are equivalent:

- (a) $E\Omega$ exists
- (b) $E\text{tr}\Omega$ exists
- (c) $E\mu(\Omega)$ exists,

where $\mu(\Omega)$ denotes the maximal characteristic root of Ω .

proof

We shall establish that $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (a)$. Now, $(a) \Rightarrow (b)$ is trivial. Further, $\mu(\Omega) \leq \text{tr}\Omega$ implies that $(b) \Rightarrow (c)$. Finally, since $\mu(\Omega)$ is the maximal characteristic root of Ω , we have $x'\Omega x \leq \mu(\Omega)x'x$ for any real n -vector x (n being the order of Ω). Let i_h be the h -th column of the identity matrix I_n . Then $i_h'\Omega i_h \leq \mu(\Omega)i_h'i_h$, i.e. $\omega_{hh} \leq \mu(\Omega)$. Also for $h \neq k$, $(i_h + i_k)'\Omega(i_h + i_k) \leq \mu(\Omega)(i_h + i_k)'(i_h + i_k)$, i.e. $\omega_{hh} + \omega_{kk} + 2\omega_{hk} \leq 2\mu(\Omega)$. Obviously $\omega_{hh} \geq 0$, $\omega_{kk} \geq 0$. Therefore $\omega_{hk} \leq \mu(\Omega)$ and $(c) \Rightarrow (a)$. ▲

The following proposition can now be proved.

proposition 2

A sufficient condition for the existence of Eb_j is the existence of

$$(1) \quad E\text{tr}(\Omega_j)\text{tr}(\Omega_j^{-1}) .$$

If ϵ has fourth moments, then any one of the following two conditions is sufficient too:

$$(2) \quad E\text{tr}(\Omega_j) \text{ and } E(\text{tr}\Omega_j^{-2}) \text{ exist,}$$

$$(3) \quad E\text{tr}(\Omega_j^2) \text{ and } E(\text{tr}\Omega_j^{-1}) \text{ exist.}$$

proof

The sampling error of b_j has the form

$$b_j - \beta = (X' \Omega_j^{-1} X)^{-1} X' \Omega_j^{-1} \epsilon.$$

Since X has full rank, it is necessary and sufficient to prove the existence of $E H_j \epsilon$, where

$$H_j = X(X' \Omega_j^{-1} X)^{-1} X' \Omega_j^{-1}$$

is an idempotent matrix, but *not* symmetric. This is important since it implies that the elements of H_j may not be bounded by some function of X .³⁾

We therefore introduce the idempotent and symmetric matrix P_j defined as

$$P_j = \Omega_j^{-\frac{1}{2}} X(X' \Omega_j^{-1} X)^{-1} X' \Omega_j^{-\frac{1}{2}}.$$

Let a be an arbitrary nonstochastic real n -vector. Since $H_j = \Omega_j^{\frac{1}{2}} P_j \Omega_j^{-\frac{1}{2}}$, a necessary and sufficient condition that $E b_j$ exists is that

$$|a' \Omega_j^{\frac{1}{2}} P_j \Omega_j^{-\frac{1}{2}} \epsilon|$$

has a finite expectation. Proceeding now as in Fuller and Battese (1973, p. 628-9), we can show that the following inequalities hold:

$$\begin{aligned} |a' \Omega_j^{\frac{1}{2}} P_j \Omega_j^{-\frac{1}{2}} \epsilon| &\leq (a' \Omega_j a)^{\frac{1}{2}} (\epsilon' \Omega_j^{-\frac{1}{2}} P_j \Omega_j^{-\frac{1}{2}} \epsilon)^{\frac{1}{2}} \\ &\leq (a' \Omega_j a)^{\frac{1}{2}} (\epsilon' \Omega_j^{-1} \epsilon)^{\frac{1}{2}} \leq \left[\mu(\Omega_j) a' a \right]^{\frac{1}{2}} \left[\mu(\Omega_j^{-1}) \epsilon' \epsilon \right]^{\frac{1}{2}} \\ &= (a' a)^{\frac{1}{2}} \left[\mu(\Omega_j) \mu(\Omega_j^{-1}) \epsilon' \epsilon \right]^{\frac{1}{2}}. \end{aligned}$$

The first inequality in the above expression is Cauchy-Schwarz', the second flows from the fact that P_j is idempotent and symmetric, and the third holds since the Rayleigh quotient $x' \Omega x / x' x$ attains its maximum at $\mu(\Omega)$.

Thus, if $E[\mu(\Omega_j) \mu(\Omega_j^{-1}) \epsilon' \epsilon]^{\frac{1}{2}}$ exists, then $E b_j$ exists.

Now, since for any two random variables $x, y : (Exy)^2 \leq (Ex^2)(Ey^2)$, we have

- (1) $E[\mu(\Omega_j) \mu(\Omega_j^{-1}) \epsilon' \epsilon]^{\frac{1}{2}} \leq [E\mu(\Omega_j) \mu(\Omega_j^{-1})]^{\frac{1}{2}} [\text{tr}(\Omega)]^{\frac{1}{2}},$
- (2) $E[\mu(\Omega_j) \mu(\Omega_j^{-1}) \epsilon' \epsilon]^{\frac{1}{2}} \leq [E\mu(\Omega_j)]^{\frac{1}{2}} [E\mu^2(\Omega_j^{-1})]^{\frac{1}{4}} [E(\epsilon' \epsilon)^2]^{\frac{1}{4}},$
- (3) $E[\mu(\Omega_j) \mu(\Omega_j^{-1}) \epsilon' \epsilon]^{\frac{1}{2}} \leq [E\mu(\Omega_j^{-1})]^{\frac{1}{2}} [E\mu^2(\Omega_j)]^{\frac{1}{4}} [E(\epsilon' \epsilon)^2]^{\frac{1}{4}}.$

³⁾ If $H_j = (\eta_{st})$ were bounded by, say, $|\eta_{st}| \leq \phi(X)$, then $\text{tr}(H_j H_j)$ would be bounded by $n \phi^2(X)$. Then $\epsilon' H_j H_j \epsilon \leq (\text{tr} H_j H_j) \epsilon' \epsilon \leq n \phi^2(X) \epsilon' \epsilon$, which has a finite expectation. Then $H_j \epsilon$ has a finite expectation. Of course, when $M = (m_{ij})$ is idempotent and symmetric, then $|m_{ij}| \leq 1$.

The proposition now follows from lemma 1 and the fact that $\mu^2(\Omega) = \mu(\Omega^2)$.

Summarizing this section, we have established that b_j is an unbiased estimator for β at each step of the iteration, provided only that ϵ is symmetrical about zero, θ_j is an even function of e_{j-1} , and $E(\text{tr}\Omega_j)(\text{tr}\Omega_j^{-1})$ exists.

3. The unbiasedness of the ML estimator for β

Let us now turn to the maximum likelihood (ML) estimation of β and θ , and investigate the condition under which the ML estimator $\hat{\beta}$ is unbiased. We need one more assumption.

ASSUMPTION 4: ϵ is normally distributed.

The probability density of y is

$$(2\pi)^{-\frac{n}{2}} |\Omega|^{-\frac{1}{2}} \exp^{-\frac{1}{2}\epsilon' \Omega^{-1} \epsilon}.$$

Thus, the loglikelihood takes the form

$$\Lambda(\beta, \theta) = \gamma + \frac{1}{2} \log |\Omega^{-1}| - \frac{1}{2} \epsilon' \Omega^{-1} \epsilon,$$

where γ is a constant.

As shown in Magnus (1977, theorem 1), the ML equations are

$$\begin{aligned} \text{(i)} \quad \hat{\beta} &= (X' \hat{\Omega}^{-1} X)^{-1} X' \hat{\Omega}^{-1} y \\ \text{(ii)} \quad \text{tr} \left(\frac{\partial \Omega^{-1}}{\partial \theta^{(h)}} \Omega \right)_{\theta=\hat{\theta}} &= e' \left(\frac{\partial \Omega^{-1}}{\partial \theta^{(h)}} \right)_{\theta=\hat{\theta}} e \quad (h=1 \dots m), \end{aligned}$$

where $\theta^{(h)}$ is the h -th component of θ , $e = y - X\hat{\beta}$, and $\hat{\Omega} = \Omega(\hat{\theta})$.

The ML estimates are obtained by using the iterative scheme of section two.

This means that θ_{j+1} is estimated as the solution of

$$\text{tr} \left(\frac{\partial \Omega^{-1}}{\partial \theta^{(h)}} \Omega \right) = e_j' \frac{\partial \Omega^{-1}}{\partial \theta^{(h)}} e_j \quad (h=1 \dots m).$$

Now, even though it is often impossible to express θ_{j+1} explicitly in terms of e_j , it is clear that θ_{j+1} is an even function of e_j . Further ϵ is symmetrical about zero. Thus, b_j is unbiased if $E \text{tr}(\Omega_j) \text{tr}(\Omega_j^{-1})$ exists.

When considering the ML estimator $\hat{\beta}$, we must make sure that the iterative procedure in fact converges to a solution of the ML equations. This problem has been studied by Oberhofer and Kmenta (1974). We shall need one last assumption.

ASSUMPTION 5: There exist numbers $\lambda > 0$, $M > 0$, and s , such that for any $\beta^* \in \mathbb{R}^k$ and $\theta^* \in U$ for which $\Lambda(\beta^*, \theta^*) \geq s$, we have $\hat{\theta}(\beta^*) \in U$. Here $\hat{\theta}(\beta)$ denotes the value of θ that maximizes $\Lambda(\beta, \theta)$ when β is kept fixed, and U is defined as

$$U = \{\theta \mid 0 < \lambda \leq |\Omega|, \theta' \theta \leq M\} \cap \Theta,$$

where Θ , the class of admissible values of θ , is a closed set.

The assumptions (1) - (5) are sufficient for the convergence of the iterative procedure to a solution of the ML equations (Oberhofer and Kmenta (1974, theorem 1)), but not only that. If assumption 5 is satisfied, then there exist numbers ℓ_1 and ℓ_2 such that

$$0 < \ell_1 \leq \lambda(\Omega) \leq \ell_2,$$

where $\lambda(\Omega)$ stands for any eigenvalue of Ω ⁴⁾. This implies that

$$\mu(\Omega) \mu(\Omega^{-1}) \leq \ell_2 / \ell_1.$$

Thus $\mu(\Omega) \mu(\Omega^{-1})$ is bounded. Then $\text{tr}(\Omega) \text{tr}(\Omega^{-1})$ is also bounded. Hence its expectation exists. Summarizing, we have proved the following

proposition 3

Under the assumptions (1)-(5), b_j is an unbiased estimator of β at each step of the iteration. The iterative procedure converges to a solution of the ML equations and the resulting ML estimator $\hat{\beta}$ is unbiased too.

It may be noted that assumption 5 is too strong when we are not particularly interested in the ML estimator, but rather in a two-step (or, more generally, a k-step) procedure. The following example will clarify this issue.

---

⁴⁾ Oberhofer and Kmenta (1974, p. 583).

Example : Autocorrelated errors

Let $\varepsilon_t = \rho \varepsilon_{t-1} + \xi_t$, $E\xi = 0$, $E\xi\xi' = \sigma^2 I_n$, $|\rho| < 1$, $\varepsilon_0 = 0$.

We may write

$$\xi = A\varepsilon, \text{ where } A = \begin{bmatrix} 1 & & 0 \\ -\rho & \ddots & \\ 0 & & -\rho 1 \end{bmatrix},$$

so that $\Omega = E\varepsilon\varepsilon' = \sigma^2(A'A)^{-1}$.

The ML equations for ρ and σ^2 are 5)

$$\hat{\rho} = \frac{\sum_{i=1}^{n-1} e_i e_{i+1}}{\sum_{i=1}^{n-1} e_i^2},$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (e_i - \hat{\rho} e_{i-1})^2, \quad e_0 = 0.$$

Further, it is easy to see that

$$(\text{tr}\Omega)(\text{tr}\Omega^{-1}) = (\text{tr}A'A)(\text{tr}(A'A)^{-1}) = \frac{n+(n-1)\rho^2}{1-\rho^2} \left[n - \rho^2 \frac{1-\rho^{2n}}{1-\rho^2} \right].$$

Thus a sufficient condition for the unbiasedness of b_j is the existence of a positive number α such that $\hat{\rho}_j^2 < 1 - \alpha$ with probability 1. In that case $Eb_j = \beta$ at each step of the iteration, but there is no guarantee of convergence.

5) These ML equations define the iterative Cochrane-Orcutt procedure. See Oberhofer and Kmenta (formulae 31, 32) or Magnus (formula 34).

References

- Fuller, W.A. and G.E. Battese, "Transformations for Estimation of Linear Models with Nested-Error Structure", *Journal of the American Statistical Association*, 68 (September 1973), 626-32.
- Kakwani, N.C., "The Unbiasedness of Zellner's Seemingly Unrelated Regression Equations Estimators", *Journal of the American Statistical Association*, 62 (March 1967), 141-2.
- Kakwani, N.C., "Note on the Unbiasedness of a Mixed Regression Estimator", *Econometrica*, 36 (July-October 1968), 610-1.
- Magnus, J.R., "Maximum Likelihood Estimation of the GLS Model with Unknown Parameters in the Disturbance Covariance Matrix", *Journal of Econometrics*, forthcoming.
- Mehta, J.S. and P.A.V.B. Swamy, "Further Evidence on the Relative Efficiencies of Zellner's Seemingly Unrelated Regressions Estimator", *Journal of the American Statistical Association*, 71 (September 1976), 634-9.
- Oberhofer, W. and J. Kmenta, "A General Procedure for obtaining Maximum Likelihood Estimates in Generalized Regression Models", *Econometrica*, 42 (May 1974), 579-90.
- Swamy, P.A.V.B. and J.S. Mehta, "On Theil's Mixed Regression Estimator", *Journal of the American Statistical Association*, 64 (March 1969), 273-6.