Explicit Formulas to Determine the Efficiency of OLS in the Presence of First Order Autoregressive Disturbances

Samir Khaled Safi

James Madison University

Harrisonburg, VA 22807, USA

February 20, 2005

Abstract

In problems concerning time series, it is often the case that the disturbances are, in fact, correlated. It is known that the ordinary least squares (OLS) may not be optimal in this context. We have proved that the relative efficiency of the variance of the generalized least squares (GLS) to that of OLS is invariant to scaling and shifting of the design vectors. We have derived explicit formulas for the relative efficiencies of the GLS estimator to that of OLS estimator in some important special cases. We consider both linear and quadratic design vectors in the presence of AR(1) disturbances with and without an intercept term included in the design and use these formulas to show some asymptotic properties of the estimators.

Keywords: Autocorrelation; Autoregressive; Ordinary Least Squares; Generalized Least Squares; Efficiency.

1 Introduction

Let the relationship between an observable random variable y and k explanatory variables $X_1, \ldots, X_k$ in a T-finite system be specified in the following linear regression model

$$y = X\ \beta + u \tag{1}$$

where

y is a $(T \times 1)$ vector of observations on a response variable, X is a $(T \times k)$ design matrix, β is a $(k \times 1)$ vector of unknown regression parameters, and u is a $(T \times 1)$ random vector of disturbances. For convenience we assume that X is full column rank $k < T$ and its first column is 1's.

The *ordinary least squares* (OLS) estimator of β in the regression model (1) is

$$\widehat{\beta} = (X'X)^{-1}\, X'y \tag{2}$$

In problems concerning time series, it is often the case that the disturbances are, in fact, correlated. Practitioners are then faced with a decision, use OLS anyway, or try to fit a more complicated disturbance structure. The problem is difficult because the properties of the estimators depend highly on the structure of the independent variables in the model. For more complicated disturbance structures, many of the properties are not well understood. If the disturbance term has mean zero, i.e. $E\,(u) = \mathbf{0}$, but is in fact, autocorrelated, i.e. $Cov\,(u) = \sigma_u^2 \Sigma$, where Σ is a $T \times T$ positive definite matrix and the variance σ_u^2 is either known or unknown positive and finite scalar, then the OLS parameter estimates will continue to be unbiased, i.e. $E\left(\widehat{\beta}\right) = \beta$. But it has a different covariance matrix;

$$Cov\left(\widehat{\beta}\right) = \sigma_u^2\, (X'X)^{-1}\, X'\, \Sigma\ X\, (X'X)^{-1}\,. \tag{3}$$

The most serious implications of autocorrelated disturbances is not the resulting inefficiency of OLS but the misleading inference when standard tests are used. The autocorrelated nature of disturbances is accounted for in the *generalized least squares* (GLS) estimator given by

$$\widetilde{\beta} = \left(X'\Sigma^{-1}X\right)^{-1} X'\Sigma^{-1}y \tag{4}$$

which is unbiased, i.e. $E\left(\widetilde{\beta}\right) = \beta$, with covariance matrix

$$Cov\left(\widetilde{\beta}\right) = \sigma_u^2 \left(X' \ \Sigma^{-1}X\right)^{-1}. \tag{5}$$

The superiority of GLS over OLS is due to the fact that GLS has a smaller variance. According to the Generalized Gauss Markov Theorem, the GLS estimator provides the *Best Linear Unbiased Estimator* (BLUE) of β. But the GLS estimator requires prior knowledge of the matrix correlation structure, Σ. The OLS estimator $\widehat{\beta}$ is simpler from a computational point of view and does not require a prior knowledge of Σ.

A common approach for modeling univariate time series is the autoregressive model. The general finite order *autoregressive process of order* p or briefly, $AR(p)$, is

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \cdots + \phi_p u_{t-p} + \varepsilon_t,\ \varepsilon_t \sim i.i.d.\ N\left(0, \sigma_\varepsilon^2\right) \tag{6}$$

An important special case is the first-order autoregressive disturbance, or an $AR(1)$ error process, which is the simplest AR process, commonly seen in economic and environmental studies and easier to handle mathematically. $AR(1)$ is represented in the autoregressive form as

$$u_t = \rho u_{t-1} + \varepsilon_t,\ \varepsilon_t \sim i.i.d.\ N\left(0, \sigma_\varepsilon^2\right) \tag{7}$$

For a linear regression model with first order autocorrelated disturbances, several alternative estimators for the regression coefficients have been discussed in the literature, and their efficiency properties have been investigated with respect to the OLS and GLS estimators (e.g. Kadiyala, 1968; Maeshiro, 1979; Chipman, 1979; Spitzer, 1979; Ullah et al., 1983).

This article involves an important statistical problem concerning estimation in the presence of auto-correlated disturbances.

We consider estimation of the coefficient vector in the linear model (1). In the case of correlated errors with known variance structure, an efficient estimator can be derived using GLS as in (4) and the variance formula for the estimates are well-known and is given by (5), but depend in a non-trivial way upon the design matrix X as well as the covariance structure of the disturbances.

An interesting question occurs naturally in many cases: what happens to the relative efficiency of OLS to that of GLS for different design vectors? We will look for families of designs that we can use to characterize the efficiency ratio, such as deterministic polynomials (linear, quadratic). We have investigated the relative efficiency of GLS to OLS in the important cases of autoregressive disturbances of order one, $AR(1)$, with autoregressive coefficient ρ.

This article is organized as follows. In section 2, we look for families of designs that can we use to characterize the relative efficiency of GLS to that of OLS. In section 3, we use the explicit formulas derived in section 2 to show some asymptotic properties of the estimators. Finally in section 4 we offer some conclusion remarks and suggestions for future research on the comparison of OLS and GLS.

2 DESIGN VECTORS

In this section we look for families of designs that we can use to characterize the relative efficiency of GLS to OLS such as deterministic polynomials. The relative efficiency in a linear regression containing an autocorrelated disturbance term depends on the structure of the matrix of observations on the independent variables, i.e. it depends on a specific design which makes it difficult to characterize in general. We prove that the relative efficiency is invariant to scaling and shifting of the design vectors. In addition, explicit formulas for the relative efficiencies of the GLS estimator to that of OLS estimator in some important special cases are derived. We consider both linear and quadratic design vectors in the presence of $AR(1)$ disturbances with and without an intercept term included in the design.

2.1 Scaling and Shifting the Design Vectors

Suppose X is a vector uniformly spaced, i.e.

$$X' = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ \frac{1-a}{b} & \frac{2-a}{b} & \cdots & \frac{(T-1)-a}{b} & \frac{T-a}{b} \end{bmatrix} \tag{8}$$

where $a \in \mathbb{R}$, $b > 0$. As special case when $a = 0$ and $b = 1$, i.e.

$$X' = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & T \end{bmatrix} \tag{9}$$

We prove that using design (8) or (9) has no effect in the relative efficiency of the variance of GLS to that of OLS. In other words, the relative efficiency is invariant to scaling and shifting of the design vectors.

Theorem 1 *The relative efficiency of the variance of GLS to that of OLS is* *invariant to scaling and shifting of the design vectors.*

Proof. Recall the definitions for the variances of OLS and GLS estimators,

$V = Var\left(\widehat{\beta}\right) = \sigma_u^2 \left(X'\ X\right)^{-1} X'\ \Sigma\ X \left(X'\ X\right)^{-1}$

$\widetilde{V} = Var\left(\widetilde{\beta}\right) = \sigma_u^2 \left(X'\ \Sigma^{-1} X\right)^{-1}$

Suppose X is a $(T \times k)$ design matrix with full column rank $k < T$, and Σ is a $T \times T$ positive definite matrix. For scaled design, consider new linear regression design $Y = X^{*}\beta + u$, where $X^{*} = XD$ scaled design, $D = diag\left(b_1,\ b_2,\ \cdots,\ b_k\right),\ b_i \neq 0,\ i = 1, 2, ..., k.$

Let $Var\left(\widehat{\beta}^{*}\right)$ and $Var\left(\widetilde{\beta}^{*}\right)$ are the variances of OLS and GLS estimators for the scaled design, respectively.

$Var\left(\widehat{\beta}^{*}\right) = \sigma_u^2 \left(X^{*\prime} X^{*}\right)^{-1} X^{*\prime}\Sigma\ X^{*} \left(X^{*\prime} X^{*}\right)^{-1}$

$= \sigma_u^2 D^{-1} \left(X'\ X\right)^{-1} D^{-1} D X'\ \Sigma\ X D D^{-1} \left(X'\ X\right)^{-1} D^{-1}$

$= \sigma_u^2 D^{-1} \left(X'\ X\right)^{-1} X'\ \Sigma\ X \left(X'\ X\right)^{-1} D^{-1}$

$= \left[\frac{1}{b_i b_j} V_{i\ j}\right]_{i,\ j=1}^{k}$

Similarly,

$Var\left(\widetilde{\beta}^{*}\right) = \sigma_u^2 \left(X^{*\prime}\ \Sigma^{-1} X^{*}\right)^{-1}$

$= \sigma_u^2 D^{-1} \left(X'\ \Sigma^{-1} X\right)^{-1} D^{-1}$

$= \left[\frac{1}{b_i b_j} \widetilde{V}_{i\ j}\right]_{i,\ j=1}^{k}$

Hence the relative efficiencies of the variances of GLS to that of OLS is invariant to scaling of the design vectors.

Consider the linear regression design $Y = X\beta + u$, $\widehat{\beta}_j$ is determined solely by the relationship between Y and $x_j^{\perp}$ where $x_j^{\perp} = x_j - p\left(x_j | V_j^{\perp}\right),\ j \neq 0,$ and $V_j^{\perp} = \mathcal{L}\left(x_i,\ i = 0, 1, ..., k,\ i \neq j\right)$, i.e. $\widehat{\beta}_j = f\left(y, x_j^{\perp}\right)$. For shifted design, consider new linear regression design $Y = X^{*}\beta + u$ where

$X^{*} = \left[x_0^{*},\ x_1^{*},\ ...,\ x_k^{*}\right]$, $x_i^{*} = x_i + a_i\ x_0$, $x_0' = [1,\ 1,\ ...,\ 1]$, and $x_j^{*\perp} = x_j^{*} - p\left(x_j^{*} | V_j^{*\perp}\right),\ j \neq 0$, i.e. $\widehat{\beta}_j^{*} = f\left(y, x_j^{*\perp}\right)$

$V_j^{\perp} = \{c_0 x_0 + \cdots + c_{j-1} x_{j-1} + c_{j+1} x_{j+1} + \cdots + c_k x_k + \cdots, c_k \in \mathbb{R}\}$

$$V_j^{*\perp} = \left\{c_0 x_0^* + \cdots + c_{j-1} x_{j-1}^* + c_{j+1} x_{j+1}^* + \cdots + c_k x_k^* + \cdots, c_k \in \mathbb{R}\right\}$$
$$= \{c_0 (x_0 + a_0 x_0) + \cdots + c_{j-1} (x_{j-1} + a_{j-1} x_0) + \cdots + c_k (x_k + a_k x_0) + \cdots\}$$
$$= \{[c_0 (1 + a_0) + \cdots + c_k a_k] x_0 + c_1 x_1 + \cdots + c_{j-1} x_{j-1} + \cdots + c_k x_k + \cdots\}$$
$$= \{c_0 x_0 + c_1 x_1 + \cdots + c_{j-1} x_{j-1} + c_{j+1} x_{j+1} + \cdots + c_k x_k + \cdots\}$$
$$= V_j^{\perp}$$

Thus, in any linear regression analysis which includes the vector **J** of all ones as an **x**-vector, the $\widehat{\beta}' s$ corresponding to other vectors are not affected by adding the same intercept to all elements of those vectors, hence shifted design which is linearly dependent on **J** only effects an intercept, $\widehat{\beta}_0$. Therefore, the relative efficiency of the variance of GLS to that of OLS is invariant to scaling and shifting of the design vectors. ■

2.2 Linear Design Vector

Consider the linear regression model with first-order auto-correlated disturbances as given in (1) and (7), and assume that the true value of ρ is known. We will consider a special case, namely that in which the regression takes the form of a simple trend with and without an intercept term included in the design.

2.2.1 Linear Design Vector without an Intercept Term

Consider the simple linear trend

$$y_t = X\ \beta + u_t,\ \ t = 1, 2, ..., T \tag{10}$$

where X is a vector uniformly spaced, i.e. $X' = \begin{bmatrix} 1 & 2 & \cdots & T \end{bmatrix}$. The variance matrix of disturbance terms can be written as

$$E\left(u\ u'\right)\ = \sigma_u^2 \Sigma = \sigma_u^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{bmatrix} \tag{11}$$

where $\sigma_u^2 = \sigma_\varepsilon^2 \left(1-\rho^2\right)^{-1}$ is the variance of an AR(1) process. It is known that the inverse of Σ in (11) is given by

$$\Sigma^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix} \tag{12}$$

with Σ given in (11),

$X'\Sigma X = \sum_{i=1}^{T} i^2 + 2 \sum_{k=1}^{T-1} \sum_{i=1}^{T-k} \rho^k i\left(i+k\right),\ X'X = \sum_{i=1}^{T} i^2$

with $Var\left(\widehat{\beta}\right)$ given in (3) and collecting terms, we find that

$$Var\left(\widehat{\beta}\right) = \frac{\left[\sum_{i=1}^{T} i^2 + 2 \sum_{k=1}^{T-1} \sum_{i=1}^{T-k} \rho^k i\left(i+k\right)\right] \sigma_u^2}{\left[\sum_{i=1}^{T} i^2\right]^2} \tag{13}$$

with Σ^{-1} given in (12),

$X'\ \Sigma^{-1} X = \frac{1}{1-\rho^2} \left[\left(1+\rho^2\right) \sum_{i=1}^{T} i^2 - 2\rho \sum_{i=1}^{T-1} i\left(i+1\right) \rho^2 \left[1+T^2\right]\right]$

with $Var\left(\widetilde{\beta}\right)$ given in (5) and collecting terms, we find that

$$Var\left(\widetilde{\beta}\right) = \frac{\left(1-\rho^2\right)\ \sigma_u^2}{\left(1+\rho^2\right)\sum\limits_{i=1}^{T} i^2 - 2\rho\sum\limits_{i=1}^{T-1} i\left(i+1\right) - \rho^2\left(1+T^2\right)} \tag{14}$$

The variances in (13) and (14) are functions of ρ and T. The relative efficiency of the variance of GLS to that of OLS is given by

$$RE\left(\beta\right) = \frac{\left(1-\rho^2\right)\left[\sum\limits_{i=1}^{T} i^2\right]^2}{\Psi\left(\rho,T\right)\left[\sum\limits_{i=1}^{T} i^2 + 2\sum\limits_{k=1}^{T-1}\sum\limits_{i=1}^{T-k}\rho^k i\left(i+k\right)\right]} \tag{15}$$

where

$$\Psi\left(\rho,T\right) = \left[\left(1+\rho^2\right)\sum_{i=1}^{T} i^2 - 2\rho\sum_{i=1}^{T-1} i\left(i+1\right) - \rho^2\left(1+T^2\right)\right]$$

Equation (15) shows that upon taking the ratio of the variances, σ_u^2 is cancelled and the relative efficiency is therefore completely determined by two factors the values of ρ and T.

2.2.2 Linear Design Vector with an Intercept Term

Consider the simple trend given in (10), where X is a vector uniformly spaced, i.e.

$$X' = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & T \end{bmatrix}$$

then

$$X'\ X = \begin{bmatrix} T & \sum\limits_{t=1}^{T} t \\ \sum\limits_{t=1}^{T} t & \sum\limits_{t=1}^{T} t^2 \end{bmatrix},\ \left[X'X\right]^{-1} = \frac{2}{T\left(T-1\right)}\begin{bmatrix} 2T+1 & -3 \\ -3 & \frac{6}{T+1} \end{bmatrix}$$

with Σ given in (11),

$$X'\ \Sigma X = \begin{bmatrix} f\left(\rho,T\right) & g\left(\rho,T\right) \\ g\left(\rho,T\right) & h\left(\rho,T\right) \end{bmatrix}$$

where

$$f(\rho,T) = 2\sum_{i=0}^{T-1}\rho^i + \sum_{j=1}^{T-2}\sum_{i=1}^{j}\rho^i + \sum_{j=2}^{T-1}\sum_{i=0}^{T-j}\rho^i$$

$$g(\rho,T) = (T+1)\sum_{i=0}^{T-1}\rho^i + \sum_{j=1}^{T-2}\sum_{i=1}^{j}(j+1)\rho^i + \sum_{j=2}^{T-1}\sum_{i=0}^{T-j}j\rho^i$$

$$h(\rho,T) = \sum_{i=0}^{T-1}(i+1)\left(\rho^i + T\rho^{T-i-1}\right) + \sum_{j=1}^{T-2}\sum_{i=1}^{j}(j+1)\,i\rho^{j-i+1} + \sum_{j=2}^{T-1}\sum_{i=j}^{T}ji\rho^{i-j}$$

Collecting terms, we find that the variances of an intercept and the slope of the OLS estimator are

$$Var\left(\widehat{\beta}_0\right) = \frac{4\left\{(2T+1)\left[(2T+1)f(\rho,T) - 6g(\rho,T)\right] + 9h(\rho,T)\right\}\sigma_u^2}{\left[T(T-1)\right]^2} \tag{16}$$

$$Var\left(\widehat{\beta}_1\right) = \frac{36\left\{(T+1)\left[(T+1)f(\rho,T) - 4g(\rho,T)\right] + 4h(\rho,T)\right\}\sigma_u^2}{\left[T(T^2-1)\right]^2} \tag{17}$$

respectively, with Σ^{-1} given in (12), $X'\,\Sigma^{-1}X = \frac{1}{1-\rho^2}\begin{bmatrix} k(\rho,T) & l(\rho,T) \\ l(\rho,T) & m(\rho,T) \end{bmatrix}$

where

$$k(\rho,T) = (1-\rho)\left[T - \rho(T-2)\right]$$

$$l(\rho,T) = \tfrac{1}{2}(T+1)\,k(\rho,T)$$

$$m(\rho,T) = (1-2\rho) + T\left[T - (T-1)\rho\right] + \tfrac{1}{6}(1-\rho)^2(T-2)\left(2T^2+T+3\right)$$

Collecting terms, we find that the variances of an intercept and the slope of the GLS estimator are

$$Var\left(\widetilde{\beta}_0\right) = \frac{\left(1-\rho^2\right)m(\rho,T)\,\sigma_u^2}{d(\rho,T)} \tag{18}$$

$$Var\left(\widetilde{\beta}_1\right) = \frac{\left(1-\rho^2\right)k(\rho,T)\,\sigma_u^2}{d(\rho,T)} \tag{19}$$

respectively, where $d(\rho,T) = k(\rho,T)\,m(\rho,T) - \left[l(\rho,T)\right]^2$

Then the relative efficiencies of the variances of GLS to that of OLS for an intercept and the slope are given by

$$RE\left(\beta_0\right)=\frac{\left[T\left(T-1\right)\right]^2\left(1-\rho^2\right)m\left(\rho,T\right)}{4d\left(\rho,T\right)\left\{\left(2T+1\right)\left[\left(2T+1\right)f\left(\rho,T\right)-6g\left(\rho,T\right)\right]+9h\left(\rho,T\right)\right\}} \tag{20}$$

$$RE\left(\beta_1\right)=\frac{\left[T\left(T^2-1\right)\right]^2\left(1-\rho^2\right)k\left(\rho,T\right)}{36d\left(\rho,T\right)\left\{\left(T+1\right)\left[\left(T+1\right)f\left(\rho,T\right)-4g\left(\rho,T\right)\right]+4h\left(\rho,T\right)\right\}} \tag{21}$$

2.3 Quadratic Design Vector

Consider the quadratic regression design with first-order auto-correlated disturbances as given in (1) and (7), and assume that the true value of ρ is known. We will consider a special case, namely that in which the regression takes the form of a simple trend with and without an intercept term included in the design.

2.3.1 Quadratic Design Vector without an Intercept Term

Consider the simple trend given in (10), where X is a quadratic vector, i.e. $X'=\begin{bmatrix} 1 & 4 & \cdots & T^2 \end{bmatrix}$. With Σ given in (11),

$X'\,\Sigma X=\sum_{i=1}^{T} i^4+2\sum_{k=1}^{T-1}\sum_{i=1}^{T-k}\rho^k i^2\left(i+k\right)^2$, $X'X=\sum_{i=1}^{T} i^4$

with $Var\left(\widehat{\beta}\right)$ given in (3) and collecting terms, we find that the variance of the OLS estimator is

$$Var\left(\widehat{\beta}\right)=\frac{\left[\sum_{i=1}^{T} i^4+2\sum_{k=1}^{T-1}\sum_{i=1}^{T-k}\rho^k\left[i\left(i+k\right)\right]^2\right]\sigma_u^2}{\left[\sum_{i=1}^{T} i^4\right]^2} \tag{22}$$

with Σ^{-1} given in (12),

$X'\Sigma^{-1}X=\frac{1}{1-\rho^2}\left[\left(1+\rho^2\right)\sum_{i=1}^{T} i^4-2\rho\sum_{i=1}^{T-1} i^2\left(i+1\right)^2-\rho^2\left(1+T^4\right)\right]$

with $Var\left(\widetilde{\beta}\right)$ given in (5) and collecting terms, we find that the variance of the GLS estimator is

$$Var\left(\widetilde{\beta}\right)=\frac{\left(1-\rho^2\right)\sigma_u^2}{\Psi^*\left(\rho,T\right)} \tag{23}$$

where

$$\Psi^*(\rho,T) = \left[\left(1+\rho^2\right)\sum_{i=1}^{T} i^4 - 2\rho\sum_{i=1}^{T-1}\left[i\,(i+1)\right]^2 - \rho^2\left(1+T^4\right)\right]$$

The variances in (22) and (23) are functions of ρ and T. The relative efficiency of the variance of GLS to that of OLS is given by

$$RE(\beta) = \frac{\left(1-\rho^2\right)\left[\sum_{i=1}^{T} i^4\right]^2}{\Psi(\rho,T)\left[\sum_{i=1}^{T} i^4 + 2\sum_{k=1}^{T-1}\sum_{i=1}^{T-k}\rho^k\left[i\,(i+k)\right]^2\right]} \tag{24}$$

Equation (24) shows that upon taking the ratio of the variances, σ_u^2 is cancelled and the relative efficiency is therefore completely determined by two factors the values of ρ and T.

2.3.2 Quadratic Design Vector with an Intercept Term

Consider the simple trend given in (10), where X is a quadratic vector, i.e.

$$X' = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 4 & \cdots & (T-1)^2 & T^2 \end{bmatrix}$$

Then

$$X'X = \begin{bmatrix} T & a \\ a & b \end{bmatrix}$$

where

$$a = \sum_{t=1}^{T} t^2 = \frac{T\,(2T+1)\,(T+1)}{6}, \text{ and } b = \sum_{t=1}^{T} t^4 = \frac{T\,(2T+1)\left(3T^2+3T-1\right)}{30}$$

$$\left(X'X\right)^{-1} = \tfrac{1}{bT^2-a^2}\begin{bmatrix} b & -a \\ -a & T \end{bmatrix}$$

with Σ given in (11),

$$X'\Sigma X = \begin{bmatrix} f^*(\rho,T) & g^*(\rho,T) \\ g^*(\rho,T) & h^*(\rho,T) \end{bmatrix}$$

where

$$f^*(\rho,T) = 2\sum_{i=0}^{T-1}\rho^i + \sum_{j=1}^{T-2}\sum_{i=1}^{j}\rho^i + \sum_{j=2}^{T-1}\sum_{i=0}^{T-j}\rho^i = f(\rho,T)$$
$$g^*(\rho,T) = \left(T^2+1\right)\sum_{i=0}^{T-1}\rho^i + \sum_{j=1}^{T-2}\sum_{i=1}^{j}(j+1)^2\rho^i + \sum_{j=2}^{T-1}\sum_{i=0}^{T-j}j^2\rho^i$$
$$h^*(\rho,T) = \sum_{i=0}^{T-1}(i+1)^2\left(\rho^i + T^2\rho^{T-i-1}\right) + \sum_{j=1}^{T-2}\sum_{i=1}^{j}\left[i(j+1)\right]^2\rho^{j-i+1} + \sum_{j=2}^{T-1}\sum_{i=j}^{T}\left[ij\right]^2\rho^{i-j}$$

Collecting terms, we find that the variances of an intercept and the slope of the OLS estimator are

$$Var\left(\widehat{\beta}_0\right) = \frac{\left[b^2 f^*(\rho,T) - 2ab\ g^*(\rho,T) + a^2\ h^*(\rho,T)\right]\sigma_u^2}{\left[bT-a^2\right]^2} \tag{25}$$

$$Var\left(\widehat{\beta}_1\right) = \frac{\left[a^2 f^*(\rho,T) - 2aT\ g^*(\rho,T) + T^2\ h^*(\rho,T)\right]\sigma_u^2}{\left[bT-a^2\right]^2} \tag{26}$$

respectively, with Σ^{-1} given in (12),

$$X'\ \Sigma^{-1}X = \tfrac{1}{1-\rho^2}\begin{bmatrix} k^*(\rho,T) & l^*(\rho,T) \\ l^*(\rho,T) & m^*(\rho,T) \end{bmatrix}$$

where

$$k^*(\rho,T) = (1-\rho)\left[T-\rho(T-2)\right] = k(\rho,T)$$
$$l^*(\rho,T) = (1-\rho)\left[T^2+1+(1-\rho)\sum_{i=2}^{T-1}i^2\right]$$
$$m^*(\rho,T) = (1-4\rho) + T^2\left[T^2-\rho(T-1)^2\right] + (1-\rho)^2\sum_{i=2}^{T-1}i^4 - 2\rho\sum_{i=2}^{T-1}i^2$$

Collecting terms, we find that the variances of an intercept and the slope of the GLS estimator are

$$Var\left(\widetilde{\beta}_0\right) = \frac{\left(1-\rho^2\right)m^*(\rho,T)\,\sigma_u^2}{d^*(\rho,T)} \tag{27}$$

$$Var\left(\widetilde{\beta}_1\right) = \frac{\left(1-\rho^2\right)k^*(\rho,T)\,\sigma_u^2}{d^*(\rho,T)} \tag{28}$$

respectively, where $d^*(\rho,T) = k^*(\rho,T)\,m^*(\rho,T) - \left[l^*(\rho,T)\right]^2$

Then the relative efficiencies of the variances of GLS to that of OLS for an intercept and the slope are given by

$$RE\left(\beta_0\right) = \frac{\left(bT - a^2\right)^2 \left(1-\rho^2\right) m^*\left(\rho, T\right)}{d^*\left(\rho, T\right)\left[b^2 f^*\left(\rho, T\right) - 2ab\ g^*\left(\rho, T\right) + a^2\ h^*\left(\rho, T\right)\right]} \tag{29}$$

$$RE\left(\beta_1\right) = \frac{\left(bT - a^2\right)^2 \left(1-\rho^2\right) k^*\left(\rho, T\right)}{d^*\left(\rho, T\right)\left[a^2 f^*\left(\rho, T\right) - 2aT\ g^*\left(\rho, T\right) + T^2\ h^*\left(\rho, T\right)\right]} \tag{30}$$

respectively.

The previous derived formulas are designed to compute the relative efficiencies of the variances of regression coefficients of GLS to that of OLS for linear and quadratic design vectors when the disturbance term follows an $AR(1)$ process.

The features of these particular cases will, first and foremost, reduce the required time for computations by about 70% and, second, be easier to use than the regular formulas for the covariances of the regression coefficients, $\boldsymbol{\beta}$, by OLS and GLS in (3) and (5). In addition, these derived formulas enable us to derive the following properties of the relative efficiency of GLS to OLS.

3 Properties of the Relative Efficiency Function

Using the previous explicit expressions for the relative efficiencies of GLS to OLS, we prove that the following limiting properties hold for the relative efficiencies of the variances of GLS to OLS of (β_0, β_1) for linear and quadratic design vectors with an intercept term included in the design. The other cases for these designs without an intercept term are similar and satisfy these properties.

3.1 Linear Design Vector

The following limiting properties hold for the relative efficiencies, $RE\left(\boldsymbol{\beta}\right)$, for the linear design vector with an intercept term.

(*i*) $\lim_{|\rho|\to 1} RE\left(\beta_0\right) = \lim_{|\rho|\to 1} RE\left(\beta_1\right) = 0$

That is, if the first order autoregressive coefficient of the disturbance term is large enough, then OLS estimators of estimating an intercept and the slope perform poorly.

(ii) $\lim_{T\to\infty} RE(\beta_0) = \lim_{T\to\infty} RE(\beta_1) = 1$, for any fixed ρ, $|\rho| < 1$

That is, if the sample size is large enough, then OLS estimators of estimating an intercept and the slope perform nearly as efficiently as GLS estimators.

Proof. Case (1): For an intercept, the relative efficiency of the variance of GLS to OLS of β in (20) can be written as

$$RE(\beta_0) = \frac{f_0(\rho)}{g_0(\rho)\ h_0(\rho)} \tag{31}$$

where

$f_0(\rho) = \rho T^2 \left[2(\rho-1)^2 T^3 + 3\left(1-\rho^2\right)T^2 + \left(\rho^2+4\rho+1\right)T - 6\rho^2\right]$
$(T-1)(\rho+1)(\rho-1)^4$

$g_0(\rho) = -\rho(\rho-1)\left[(\rho-1)\left(\rho^2+2\rho-1\right) + 4\rho^{T+1}(4\rho+5)\right]T +$
$\left(8\rho^{T+1} - \rho^2 + 10\rho + 1\right)(\rho-1)^2 \rho T^2 + 2\rho(\rho+1)(\rho-1)^3 T^3 -$
$2\rho^2\left[\rho^3 + 4\rho^2 + 10\rho + 4 - 2\rho^T(\rho+2)(2\rho+1)\right]$

$h_0(\rho) = \left[(\rho-1)^2 T^2 - (5\rho+1)(\rho-1)T + 6\rho(\rho+1)\right]\left[(\rho-1)T - 2\rho\right]$

As $|\rho| \to 1$, $f_0(\rho) \to 0$, $g_0(\rho) \nrightarrow 0$ and $h_0(\rho) \nrightarrow 0$ in (31), then $\lim_{|\rho|\to 1} RE(\beta_0) = 0$.

The coefficient of T^6, the maximum power of $f_0(\rho)$ and $g_0(\rho)\ h_0(\rho)$, is $2\rho(\rho+1)(\rho-1)^6$. Divide numerator and denominator by T^6, as $T \to \infty$ in (31), then $\lim_{T\to\infty} RE(\beta_0) = 1$.

Case (2): For the slope, the relative efficiency of the variance of GLS to OLS of β in (21) can be written as

$$RE(\beta_1) = \frac{f_1(\rho)}{g_1(\rho)\ h_1(\rho)} \tag{32}$$

where

$f_1(\rho) = \rho(\rho+1)(\rho-1)^5 T^2 (T+1)\left(T^2-1\right)$

$g_1(\rho) = \rho\left(1-\rho^2\right)\left[(1-\rho)^2 + 12\rho^{T+1}\right] T + 6\rho^2(\rho-1)^2\left(1+\rho^T\right) T^2 + \rho(\rho+1)$
$(\rho-1)^3 T^3$

$h_1(\rho) = (\rho-1)^2 T^2 - (5\rho+1)(\rho-1) T + 6\rho(1+\rho)$

Applying L'Hôpital's rule, as $|\rho| \to 1$, in (32), then $\lim\limits_{|\rho|\to 1} RE(\beta_1) = 0$.

The coefficient of T^5, the maximum power of $f_1(\rho)$ and $g_1(\rho)\ h_1(\rho)$, is $\rho(\rho+1)(\rho-1)^5$. Divide numerator and denominator by T^5, as $T \to \infty$ in (32), then $\lim\limits_{T\to\infty} RE(\beta_1) = 1$. ■

3.2 Quadratic Design Vector

The following limiting properties hold for the relative efficiencies, $RE(\boldsymbol{\beta})$, for the quadratic design vector with an intercept term.

(*i*) $\lim\limits_{|\rho|\to 1} RE(\beta_0) = \lim\limits_{|\rho|\to 1} RE(\beta_1) = 0$

(*ii*) $\lim\limits_{T\to\infty} RE(\beta_0) = \lim\limits_{T\to\infty} RE(\beta_1) = 1$, for any fixed ρ, $|\rho| < 1$

Proof. Case (1): For an intercept, the relative efficiency of the variance of GLS to OLS of β in (29) can be written as

$$RE(\beta_0) = \frac{f_0^*(\rho)}{g_0^*(\rho)\ h_0^*(\rho)} \tag{33}$$

where

$f_0^*(\rho) = \left[\rho T^2 (1+\rho)(\rho-1)^6 (T-1)(8T+11)^2\right]$
$\left[6(\rho-1)^2 T^5 + 15\left(1-\rho^2\right) T^4 + 10(\rho+1)^2 T^3 - \left(\rho^2+8\rho+1\right) T - 30\rho^2\right]$

$g_0^*(\rho) = 16(\rho-1)^3 T^5 - 2(43\rho+23)(\rho-1)^2 T^4 + (\rho-1)\left(89\rho^2+290\rho+41\right) T^3$
$+\left(119\rho^3 - 309\rho^2 - 279\rho - 11\right) T^2 - 24\rho\left(3\rho^2+14\rho+3\right) T - 180\rho^2(\rho+1)$

$h_0^*(\rho) = \delta_0 + \delta_1 T + \delta_2 T^2 + \delta_3 T^3 + \delta_4 T^4 + \delta_5 T^5$

where

$\delta_0 = -12\rho^2\left(\rho^2+3\rho+6\right)\left(6\rho^2+3\rho+1\right)\left(\rho^T-1\right)$

$\delta_1 = \rho\left(\rho - 1\right)$

$\left[\left(11\rho^4 - 104\rho^3 - 336\rho^2 - 280\rho - 11\right)\left(\rho - 1\right) + 12\rho^{T+1}\left(21\rho^3 + 57\rho^2 + 33\rho - 11\right)\right]$

$\delta_2 = -12\rho\left(\rho - 1\right)^2\left[\left(3\rho^4 + 6\rho^3 + 41\rho^2 + 13\rho - 3\right) + \rho^{T+1}\left(18\rho^2 + 9\rho - 37\right)\right]$

$\delta_3 = -4\rho\left(\rho - 1\right)^3\left[8\rho^3 - 27\rho^2 + 51\rho + 8 + 9\left(\rho + 9\right)\rho^{T+1}\right]$

$\delta_4 = 3\rho\left(\rho - 1\right)^4\left(11\rho^2 + 26\rho - 11 + 24\rho^{T+1}\right)$

$\delta_5 = 24\rho\left(\rho + 1\right)\left(\rho - 1\right)^5$

Applying L'Hôpital's rule, as $|\rho| \to 1$, in (33), then $\lim\limits_{|\rho| \to 1} RE\left(\beta_0\right) = 0$

The coefficient of T^{10}, the maximum power of $f_0^*\left(\rho\right)$ and $g_0^*\left(\rho\right)\ h_0^*\left(\rho\right)$ is $386\rho\left(\rho + 1\right)\left(\rho - 1\right)^8$. Divide numerator and denominator by T^{10}, as $T \to \infty$ in (33), then $\lim\limits_{T \to \infty} RE\left(\beta_0\right) = 1$

Case (2): For the slope, the relative efficiency of the variance of GLS to OLS of β in (30) can be written as

$$RE\left(\beta_1\right) = \frac{f_1^*\left(\rho\right)}{g_1^*\left(\rho\right)\ h_1^*\left(\rho\right)} \tag{34}$$

where

$f_1^*\left(\rho\right) = \rho\left(\rho - 1\right)^7\left(\rho + 1\right)T^2\left(8T + 11\right)\left(8T^2 + 3T - 11\right)\left(2T + 1\right)^2\left(T + 1\right)^2$
$\left[T\left(\rho - 1\right) - 2\rho\right]$

$g_1^*\left(\rho\right) = g_0^*\left(\rho\right)$

$h_1^*\left(\rho\right) = \delta_0^* + \delta_1^* T + \delta_2^* T^2 + \delta_3^* T^3 + \delta_4^* T^4 + \delta_5^* T^5$

where

$\delta_0^* = 10\rho^2\left(5\rho^2 + 8\rho - 1\right)\left(\rho^2 - 8\rho - 5\right)\left(\rho^T - 1\right)$

$\delta_1^* = -\rho\left(\rho^2 - 1\right)\left[\left(\rho - 1\right)\left(11\rho^3 + 147\rho^2 + 213\rho - 11\right) - \rho^{T+1}\left(\rho^2 + 7\rho - 2\right)\right]$

$\delta_2^* = -10\rho\left(\rho - 1\right)^2\left[3\rho^4 - \rho^3 + 38\rho^2 + 11\rho - 3 + \rho^{T+1}\left(31\rho^2 + 34\rho - 41\right)\right]$

$\delta_3^* = -5\rho\left(\rho - 1\right)^3\left[\left(\rho^2 - 36\rho - 1\right)\left(\rho - 1\right) - 12\rho^{T+1}\left(\rho - 5\right)\right]$

$\delta_4^* = 10\rho\left(\rho - 1\right)^4\left(3\rho^2 + 10\rho - 3 + 8\rho^{T+1}\right)$

$\delta_5^* = 16\rho\left(\rho + 1\right)\left(\rho - 1\right)^5$

Applying L'Hôpital's rule, as $|\rho| \to 1$ in (34), then $\lim_{|\rho|\to 1} RE(\beta_1) = 0$.
The coefficient of T^{10}, the maximum power of $f_1^*(\rho)$ and $g_1^*(\rho)\ h_1^*(\rho)$, is $256\rho(\rho+1)(\rho-1)^8$. Divide numerator and denominator by T^{10}, as $T \to \infty$ in (34), then $\lim_{T\to\infty} RE(\beta_1) = 1$. ■

4 Conclusion and Future Research

In this article we have proved that the relative efficiency of the variance of GLS to that of OLS is invariant to scaling and shifting of the design vectors. We have derived explicit formulas for the relative efficiencies of the GLS estimator to OLS estimator for linear and quadratic design vectors in the presence of first order autoregressive disturbances in the regression models and their asymptotic properties.

We are currently using these explicit formulas to investigate the performance of the GLS estimator to that of the OLS estimator of the regression coefficient when the disturbance term follows $AR(1)$ or $AR(2)$ process for linear and quadratic design vectors.

References

[1] CHIPMAN, J. S. (1979) Efficiency of Least Squares Estimation of Linear Trend When Residuals Are Autocorrelated. *Econometrica* 47, 115-128.

[2] KADIYALA, K. R. (1968) A Transformation used to Circumvent the Problem of Autocorrelation. *Econometrica* 36, 93-96.

[3] MAESHIRO, A. (1979) On the Retention of the First Observations in Serial Correlation Adjustment of Regression Models. *International Economic Review* 20, 259-65.

[4] SPITZER, J. J. (1979) Small-Sample Properties of Nonlinear Least Squares and Maximum Likelihood Estimators in the Context of Autocorrelated Errors. *Journal of the American Statistical Association* 74, 41-47.

[5] ULLAH, A.; SRIVASTAVA, V. K.; MAGEE, L. and SRIVASTAVA, A. (1983) Estimation of Linear Regression Model with Autocorrelated Disturbances. *Journal of Time Series Analysis* 4, 127-135.