بسم الله الرحمن الرحيم

1426 هـ – 2005 م

Comparison of Estimators in Regression Models with AR(1) and AR(2) Disturbances: When is OLS Efficient?

بحث مقدم إلى المؤتمر العلمي الأول

الاستثمار والتمويل في فلسطين
بين آفـاق التنميـة والتحديـات المعـاصـرة

المنعقد بكلية التجارة في الجامعة الإسلامية

في الفترة من 8 – 9 مايو 2005م

مقدم من:

Dr. Samir Khaled Safi
James Madison University
Harrisonburg, VA 22807, USA

مايو 2005 م

Comparison of Estimators in Regression Models with AR(1) and AR(2) Disturbances: When is OLS Efficient?

Abstract:

It is well known that the ordinary least squares (OLS) estimates in the regression model are efficient when the disturbances have mean zero, constant variance and are uncorrelated. In problems concerning time series, it is often the case that the disturbances are, in fact, correlated. It is known that OLS may not be optimal in this context. We consider the robustness of various estimators, including estimated generalized least squares. We found that if the disturbance structure is autoregressive and the dependent variable is nonstochastic and linear or quadratic, the OLS performs nearly as well as its competitors. For other forms of the dependent variable, we have developed rules of thumb to guide practitioners in their choice of estimators.

Keywords: Autoregressive; Disturbances; Ordinary Least Squares; Generalized Least Squares; Relative Efficiency.

1 Introduction

Let the relationship between an observable random variable y and k explanatory variables $X_1, X_2, \ldots, X_k$ in a T-finite system be specified in the following linear regression model

$$y = X\beta + u \tag{1.1}$$

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.

In the classical linear regression model it is standard to assume that given any value of the explanatory variables the disturbances are uncorrelated with zero mean, and constant variance. Under these general assumptions it can be shown that ordinary least squares (OLS) estimators are optimal.

Note that in the case where the design matrix is stochastic the first assumption implies that X and u are uncorrelated. Finally, in order to develop inference such as the standard F-test, confidence intervals, etc. the disturbances are commonly assumed to follow a normal distribution. This assumption becomes less important in large samples due to the central limit theorem.

Regression analysis is possibly the most commonly used statistical method in data analysis. As a result, the OLS estimation developed for a standard regression models is probably the most commonly used method of estimation. The OLS estimator of β in the regression model (1.1) is

$$\hat{\beta} = (X'X)^{-1}X'y \tag{1.2}$$

Under the OLS assumptions these parameter estimates will be unbiased, $E(\hat{\beta}) = \beta$, with covariance matrix

$$\text{Cov}(\hat{\beta}) = \text{Cov}(\hat{\beta} \mid X) = \sigma_u^2 (X'X)^{-1} \tag{1.3}$$

In addition, according to the Gauss-Markov Theorem, the OLS estimator provides the *Best Linear Unbiased Estimator* (BLUE) of β. This means that the OLS estimator is the most efficient (least variance) estimator in the class of unbiased estimators which are linear in the data, y.

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

3

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. $Cov(u) = \sigma_u^2 \sum \neq \sigma_u^2 \mathbf{I}$, where $\sum$ 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. $E(\hat{\beta}) = \beta$. But it has a different covariance matrix;

$$Cov_{\sum}(\hat{\beta}) = \sigma_u^2 (X'X)^{-1} X' \sum X (X'X)^{-1}. \tag{1.4}$$

Furthermore, $\hat{\beta}$ is no longer BLUE for β. If we use the OLS estimator, $\hat{\beta}$, when the covariance of the disturbance vector u is not a scalar multiple of the identity matrix, then our inferences about β will be incorrect, since the variance-covariance matrix of $\hat{\beta}$ is no longer equal to presumed $\sigma_u^2 (X'X)^{-1}$. So, the standard OLS variance estimator $\hat{\sigma}_u^2 (X'X)^{-1}$ will be a biased estimator of $\sigma_u^2 (X'X)^{-1} X' \sum X (X'X)^{-1}$, consequently, using (1.3) in the context of the generalized regression model is incorrect and the usual inference procedures based on t and F tests are no longer appropriate. In other words, tests and confidence intervals based on $\hat{\beta}$ which assume that (1.3) will not have the nominal level. The most serious implication of autocorrelated disturbances is not the resulting inefficiency of OLS but the misleading inference when standard tests are used.

The estimator $\hat{\beta}$ ignores the autocorrelated nature of disturbances and is therefore not efficient. This is accounted for in the generalized least squares (GLS) estimator given by

$$\widetilde{\beta} = (X' \sum^{-1} X)^{-1} X' \sum^{-1} y \tag{1.5}$$

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

$$Cov(\widetilde{\beta}) = \sigma_u^2 (X' \sum^{-1} X)^{-1}. \tag{1.6}$$

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 BLUE of β in contrast to OLS. But the GLS estimator requires prior knowledge of the matrix correlation structure, $\sum$. The OLS

estimator $\hat{\beta}$ is simpler from a computational point of view and does not require a prior knowledge of $\sum$.

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

$$u_t = \rho u_{t-1} + \varepsilon_t, \ \varepsilon_t \sim \text{i.i.d. } N(0, \sigma_\varepsilon^2) \tag{1.7}$$

where ρ is the first order autoregressive disturbance parameter, for stationarity, $|\rho| < 1$. The second-order autoregressive process, or an AR(2) error process, may be written

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \varepsilon_t, \ \varepsilon_t \sim \text{i.i.d. } N(0, \sigma_\varepsilon^2) \tag{1.8}$$

where ϕ_1 and ϕ_2 are the second-order autoregressive disturbance parameters. More details about these processes will be developed in section 2.

There are numerous articles describing the efficiency of the OLS coefficient estimator $\hat{\beta}$, which ignores the correlation of the error, relative to the GLS estimator $\tilde{\beta}$, which takes this correlation into account. One strand of this literature is concerned with conditions on regressors and error correlation structure, which guarantee that OLS is asymptotically as efficient as GLS (Chipman et al., 1968 and 1979; Krämer, 1980; Ullah et al., 1983). The Cochrane-Orcutt (CO), 1949 estimator is a two-stage estimator of the regression coefficient β.

Chipman et al., 1968 give a lower bound for the relative efficiency of the sample mean, that is, they considered the special case in which X consists of a column of ones. Chipman, 1979 extended the analysis to the case in which a column representing trend is added, and he found the greatest lower bound for the efficiency of the least squares estimator of β relative to the GLS estimator over the interval $0 \le \rho < 1$ is .753763. This compares with a greatest lower bound of .535898 for the relative efficiency of the CO estimator of β. Therefore, the CO procedure is in the worst case only 71 percent as efficient as ordinary least squares. In addition, Krämer, 1980 proved that the efficiency (when measured by the trace of the variance-covariance matrix) of OLS is as great as that of GLS as ρ is close to one if the model includes only a constant term. Krämer and Marmol, 2002 show that OLS and GLS are asymptotically

equivalent in the linear regression model with AR(p) disturbance and a wide range of trending independent variables, and that OLS based statistical inference is still meaningful after proper adjustment of the test statistics.

This paper involves an important statistical problem concerning estimation in the presence of auto-correlated disturbances. Comparison of estimators in linear regression models with auto-correlated disturbances is inspired by problems that arise in meteorology and economics. It is well known that the most famous method of estimation, OLS, may be not optimal in this context. For this reason, over the years many specialized estimation techniques have been developed. These methods are more complicated than OLS and are less understood. For example, the small sample behavior is not well known. Neither is the robustness of these procedures against variations in the assumption of models underlying them. We are comparing these estimators and developing criterion under which OLS estimation perform nearly as well as the more specialized estimators.

We consider estimation of the coefficient vector in the linear model (1.1). In the case of correlated errors with known variance structure, the OLS estimator is not efficient, but an efficient estimator can be derived using GLS as in (1.5) and the variance formulas for the estimates are well-known and is given by (1.6), 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), random covariates, and extreme designs. We will try to find an answer to the following question: How robust are OLS estimates of the coefficient and its variance against different covariance structures?

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 and second order, AR(2), with autoregressive coefficients (ϕ_1, ϕ_2) for specific choices of the design vector.

Building on work on the economics and time series literature, we are investigating the price one must pay for using OLS under suboptimal conditions. We are exploring different designs under which relative efficiency of the OLS estimator to that of GLS estimator approaches to one or zero,

determining ranges of first-order autoregressive coefficient, ρ, in AR(1) disturbance and second order of autoregressive coefficients, (ϕ_1, ϕ_2) in AR(2) for which OLS is efficient and quantifying the effect of trend on the efficiency of the OLS estimator.

This paper is organized as follows. Section 2 gives an overview of autoregressive models and their properties. Section 3 presents the relative efficiency of GLS and OLS. OLS estimation in the presence of autocorrelation, GLS estimation and estimation with AR(1) and AR(2) errors are discussed. Performance comparisons of GLS and OLS for stochastic and non-stochastic designs with AR(1) and AR(2) errors are considered in section 4. Section 5 summarizes the results and offers suggestions for future research on the comparison of OLS and GLS. Finally, the complete set of tables and figures are given in Appendices A and B, respectively.

2 Overview of Autoregressive Model

2.1 Introduction

A time series is a set of ordered observations on a quantitative characteristic of an individual or collective phenomenon taken at different points of time. Although it is not a requirement, it is common for these points to be equidistant in time. For example, sales in month t could be denoted as u_t and in the previous month by u_{t-1}. The main goals of time series analysis are: first, identifying the nature of the phenomenon represented by the sequence of observations; secondly, forecasting future values of the time series variable. Both of these goals require that the pattern of observed time series data is identified and more or less formally described.

Time series analysis is used for many applications such as: economic forecasting, sales forecasting, budgetary analysis, stock market analysis, yield projections, process and quality control, inventory studies, workload projections, utility studies, census analysis, and many more.

2.2 The pth-Order Autoregressive Process

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

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \cdots + \phi_p u_{t-p} + \varepsilon_t, \varepsilon_t \sim \text{i.i.d. } N(0, \sigma_\varepsilon^2) \qquad (2.1)$$

Define the backward shift operator $L^i u_t = u_{t-i}$, then we can rewrite (2.1) in the equivalent form

$$\left(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p\right) u_u = \varepsilon_t$$

or

$$\Phi(L) u_t = \varepsilon_t \qquad (2.2)$$

where $\Phi(L)$ is the filter that converts AR(p) to white noise, u_t is the time series, and ε_t is the disturbance term (white noise).

Definition 1 (*Stationary*): *A process u_t is stationary if its statistical properties do not change over time. More precisely, the probability distributions of the process are time-invariant.*

An autoregressive model of order p, AR(p), is stationary if the p roots of the characteristic equation of the AR model, $\Phi(L) = 0$ in (2.2), lie outside the unit circle.

Definition 2 (*Invertible*): *The linear process $\{u_t\}$ is invertible and can be expressed as an infinite series of past u observations, plus an error term ε_t, i.e.*

$$u_t = \sum_{j=1}^{\infty} \psi_j u_{t-j} + \varepsilon_t \quad \text{or}$$

$$\Psi(L) u_t = \varepsilon_t \qquad (2.3)$$

where $\Psi(L) = 1 - \sum_{j=1}^{\infty} \psi_j L^j$ if the weights ψ_j are absolutely summable, that is, if $\sum_{j=0}^{\infty} |\psi_j| < \infty$. This implies that the series $\Psi(L)$ converges on or within the unit circle.

An autoregressive model of order p, AR(p), is always invertible for any values of the parameter ϕ.

The variance of the AR(p) process is

$$\sigma_u^2 = \frac{\sigma_\varepsilon^2}{1 - \rho_1 \phi_1 - \rho_2 \phi_2 - \cdots - \rho_p \phi_p} \qquad (2.4)$$

We now discuss the two important autoregressive processes, first and second order.

8

2.3 The First-Order Autoregressive Process

Most of the time in the economic time series, data generating processes is explainable in terms of autoregressive of order one or two. The most commonly assumed process in both theoretical and empirical studies is the first-order autoregressive process or briefly, AR(1). At one time, the AR(1) process was the only autocorrelation process considered by economists. Most economic data series were annual, for which the AR(1) process is reasonable. This may explain the wide use of this process, as estimation of other more complicated processes is not manageable without the aid of a computer.

So let us begin with the simplest AR process, AR(1). AR(1) has a structure similar to random walks:

$$u_t = \rho u_{t-1} + \varepsilon_t$$

where ρ is autoregressive disturbance parameter $\left(|\rho| < 1\right)$ and ε_t is another random error that is assumed to be zero mean, homoskedastic and serially uncorrelated.

The variance of an AR(1) process is

$$\sigma_u^2 = \frac{\sigma_\varepsilon^2}{1 - \rho^2} \tag{2.5}$$

2.4 The Second-Order Autoregressive Process

The second-order autoregressive process may be written

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \varepsilon_t \tag{2.6}$$

or, in lag operator notation

$$\left(1 - \phi_1 L - \phi_2 L^2\right) u_t = \varepsilon_t \tag{2.7}$$

2.4.1 Stationary Condition

The difference equation in (2.7) is stationary (stable) provided that the roots of

$$\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 = 0 \tag{2.8}$$

lie outside the unit circle. Hence, one can show that for stationarity, the parameters ϕ_1 and ϕ_2 must lie in the triangular region restricted by

$$\begin{cases} \phi_1 + \phi_2 < 1 \\ \phi_2 - \phi_1 < 1 \\ |\phi_2| < 1 \end{cases} \tag{2.9}$$

The variance of a covariance-stationary second-order autoregressive process is given by

$$\sigma_u^2 = \frac{\left(1-\phi_2\right)\sigma_\varepsilon^2}{\left(1+\phi_2\right)\left(\left(1-\phi_2\right)^2 - \phi_1^2\right)} \tag{2.10}$$

By using $\rho_1 = \phi_1\left(1-\phi_2\right)^{-1}$, the variance of the process in (2.10) can be written

$$\sigma_u^2 = \frac{\sigma_\varepsilon^2}{\left(1-\phi_2^2\right)\left(1-\rho_1^2\right)} \tag{2.11}$$

Which is substantially more complicated than the corresponding expression (2.5) for an AR(1) case. The stationary condition in (2.9) is needed to ensure a finite and positive variance.

Under stationarity condition, as the autoregressive coefficients increase in absolute value, the variance of an AR(2) process gets larger. So, with large autocorrelated disturbances, the OLS estimators can be inefficient.

3   Relative Efficiency of GLS and OLS

Economists are often interested in the relative efficiency of different estimators when the underlying assumptions of least squares breakdown. In particular, we are interested in the relative efficiency of GLS to OLS when the disturbance term, u, has mean zero but is autocorrelated, ($\text{Cov}(u) = \sigma_u^2 \Sigma \neq \sigma_u^2 \mathbf{I}$). In this section, we provide some examples to illustrate the performance of the OLS estimator as compared to the GLS estimator.

3.1   Bias in Variance Estimation of OLS

To illustrate what form the bias can take by using the variances of the standard OLS and in the presence of an AR(1) error, we present this case. Greene, 2000 presents the following example. Suppose our design vector also follows an AR(1) process but with autocorrelation parameter say θ

$$x_t = \theta\, x_{t-1} + \gamma_t \tag{3.1}$$

where

$$\text{Var}\left(x_t\right) = \frac{\text{Var}\left(\gamma_t\right)}{1-\theta^2}$$

If T is large enough it can be shown that, the ratio between the true and estimated variance of least squares, which in this case a scalar, is approximately

$$\frac{\left(\dfrac{X'X}{T}\right)^{-1}\left(\dfrac{X'\sum X}{T}\right)\left(\dfrac{X'X}{T}\right)^{-1}}{\left(\dfrac{X'X}{T}\right)^{-1}} \approx \frac{1+\rho\theta}{1-\rho\theta} \tag{3.2}$$

If ρ and θ have the same sign, then the ratio is greater than one, if both ρ and θ are positive, *the usual situation with economic time series*, the previous ratio indicates that OLS procedures will underestimate the variance.

Example 3 *If the* u's *and* x's *are autocorrelated at .5 ($\rho = \theta = .5$), then the ratio (3.2) between the true and estimated variance of least squares is about* $\dfrac{5}{3}$, *so the standard formula will underestimate the variance of the errors by about 60%. If the autocorrelation is high and positive ($\rho = \theta = .8$), then the true variance will be more than four times that given by the OLS variances. If the autocorrelation is negative ($\rho = \theta = -.5$), then the standard errors are biased downwards. So, the OLS variance estimates will be biased downward in the presence of autocorrelation.*

This is *not true* in general, because only one explanatory variable is used and the assumption that both the explanatory variable and the disturbance follow AR(1) processes with positive parameters. In addition, this result will be different when ρ and θ have opposite signs.

Definition 4 (*Estimator Efficiency*): *Let* $\hat{\beta}$ *and* $\tilde{\beta}$ *be unbiased estimators of* β, *that is,* $E(\hat{\beta}) = E(\tilde{\beta}) = \beta$. *If* $Var(\hat{\beta}) < Var(\tilde{\beta})$, *then we say that* $\hat{\beta}$ *is more efficient than* $\tilde{\beta}$.

$\hat{\beta}$ is more efficient than $\tilde{\beta}$, it means that the value of $\hat{\beta}$ that we can obtain from a particular sample would be generally closer to the true value of β than the value $\tilde{\beta}$.

Definition 5 (*Relative Efficiency*): *Let* $Var(\hat{\beta})$ *and* $Var(\tilde{\beta})$ *represent the variances of the OLS and the GLS estimators, then the relative efficiency of GLS to that of OLS in the presence of an AR(1) error, denoted by* $[RE(\beta)]_i$, $i = 1, 2, \dots, k$, *is defined as*

11

$$[RE(\beta)]_i = \frac{\left((X'\Sigma^{-1}X)^{-1}\right)_i}{\left((X'X)^{-1}X'\Sigma X(X'X)^{-1}\right)_i} \qquad (3.3)$$

If the relative efficiency in (3.3) is close to one, then OLS is nearly as efficient as GLS. Otherwise, OLS performs poorly. The efficiency of the OLS estimator $\hat{\beta}$ relative to the GLS estimator $\tilde{\beta}$ depends on both Σ and X, and so comparisons are usually made in terms of special assumptions about these matrices.

3.2   Relative Efficiency of OLS and GLS for an AR(1) Error

The regression model with autocorrelated disturbances is a generalized regression model, then it is expected that least squares to be inefficient. But just how inefficient it is depends highly on the design. Suppose that the model is

$$y_t = \beta x_t + u_t$$

and suppose u_t is derived from an AR(1) process as given in (1.7) and that x_t consists of one column that follows an independent AR(1) process as given in (3.1). If T is large enough it can be shown that (e.g., Jude, 1985)

$$\mathrm{Var}\left(\hat{\beta}\right) \approx \frac{\sigma_u^2}{\sum\limits_{i=1}^{T} x_i^2}\left[\frac{1+\rho\theta}{1-\rho\theta}\right] \qquad (3.4)$$

$$\mathrm{Var}\left(\tilde{\beta}\right) \approx \frac{\sigma_u^2}{\sum\limits_{i=1}^{T} x_i^2}\left[\frac{1-\rho^2}{1+\rho^2-2\rho\theta}\right] \qquad (3.5)$$

where θ is the sample analogue of the first-order autocorrelation coefficient of the explanatory variable. Then the efficiency of the GLS relative to OLS is approximately

$$RE(\beta) = \frac{\mathrm{Var}\left(\tilde{\beta}\right)}{\mathrm{Var}\left(\hat{\beta}\right)} \approx \frac{(1-\rho\theta)(1-\rho^2)}{(1+\rho\theta)(1+\rho^2-2\rho\theta)} \qquad (3.6)$$

The relative efficiency (3.6) compares the performance of the GLS estimator in the presence of AR(1) disturbance and the OLS estimator. Values of the relative efficiency of GLS as compared to OLS are given in Table (3.1).

Table (3.1): Efficiency of GLS to OLS

θ	ρ
	-0.9	-0.7	-0.5	-0.3	-0.1	0.1	0.3	0.5	0.7	0.9
-0.9	0.105	0.503	0.813	0.951	0.996	0.996	0.971	0.920	0.817	0.528
-0.7	0.078	0.342	0.657	0.887	0.989	0.990	0.923	0.799	0.603	0.273
-0.5	0.079	0.311	0.600	0.851	0.984	0.986	0.886	0.714	0.484	0.185
-0.3	0.086	0.311	0.584	0.835	0.981	0.982	0.858	0.655	0.409	0.141
-0.1	0.097	0.328	0.590	0.832	0.980	0.981	0.840	0.614	0.360	0.114
0.1	0.114	0.360	0.614	0.840	0.981	0.980	0.832	0.590	0.328	0.097
0.3	0.141	0.409	0.655	0.858	0.982	0.981	0.835	0.584	0.311	0.086
0.5	0.185	0.484	0.714	0.886	0.986	0.984	0.851	0.600	0.311	0.079
0.7	0.273	0.603	0.799	0.923	0.990	0.989	0.887	0.657	0.342	0.078
0.9	0.528	0.817	0.920	0.971	0.996	0.996	0.951	0.813	0.503	0.105

Example 6 *The relative efficiency in (3.6) can be extremely poor with large values of ρ for any given θ, for example, when $\rho = \theta = .7$ then the ratio is approximately equal to 0.3423. In this case if the standard error of $\tilde{\beta}$ is 1, the standard error of $\hat{\beta}$ would be around 1.71, $(0.3423)^{\frac{-1}{2}}$, which means that the OLS standard error is about 71% bigger that its GLS competitor. With large autocorrelated disturbances, the OLS estimators can be inefficient. It can be quite reasonable for a low degree of autocorrelation for any given θ, for example, when $\rho = .3$ and $\theta = .9$ then it's approximately equal to 0.9510.*

This, of course, is well known to be true in the case of the present model, but this comparison is *not fair* because we have correctly specified that both the explanatory variable and the disturbance term follow an AR(1) process when the unknown model may be different. In addition, we assume ρ is known but in practice we should estimate this parameter.

4 Performance Comparisons

In this section we present numerical results to compare the efficiency of GLS to that of OLS. We will focus on two issues; first, the relative efficiency of GLS estimator as compared with the OLS estimator when the structure of the design vector, X, is nonstochastic. For example, linear, quadratic, and exponential

design vectors with an intercept term included in the design vector. Secondly, the relative efficiency of the GLS estimator as compared with the OLS for a stochastic design vector. In the example considered here, we generated a standard Normal stochastic design vector of length 1000. The three finite sample sizes used are 50, 100, and 200 for selected values of the autoregressive coefficients. Both AR(1) and AR(2) error processes are considered to discuss the behavior of OLS as compared to GLS.

4.1 Performance Comparisons for AR (1) Process

We discuss the relative efficiencies of OLS to GLS when the disturbance term follows an AR(1) process, $u_t = \rho\, u_{t-1} + \varepsilon_t, t = 1, 2, \ldots, T$, assuming that the autoregressive coefficient, ρ, is known priori. The three finite sample sizes used are 50, 100, and 200 for the elected values of $|\rho| \le .9$ evaluated in steps of .2.

4.1.1 Linear and Quadratic Design Vectors for AR(1) Process

Tables (4.1) and (4.2) show the relative efficiencies of the variances of GLS to OLS for a regression coefficient on linear and quadratic designs with an intercept term included in the design. To demonstrate the efficiency of OLS, consider the linear design vector. For estimating an intercept term, the relative efficiency of the OLS estimator as compared to the GLS estimator decreases with increasing values of $|\rho|$. For small and moderate sample sizes, the efficiency of the OLS estimator appears to be nearly as efficient as the GLS estimator for $|\rho| \le .7$. In addition, for large size sample data, the OLS estimator performs nearly as efficiently as the GLS estimator for the additional values of $\rho = \pm.9$. Further, the efficiency for estimating the slope mimics the efficiency of the intercept, except for large sample size; the efficiency of the OLS estimator appears to be nearly as efficient as the GLS estimator for $\rho \ne \pm.9$. The efficiency of GLS estimator to the OLS estimator for the quadratic design agrees with the behavior for the linear design vector.

4.1.2 Exponential Design Vector for AR(1) Process

Table (4.3) represents the relative efficiencies of the variances of GLS to OLS for an exponential design vector, with an intercept term included in the design. For estimating an intercept term, the efficiency of the OLS estimator appears to be nearly as efficient as the GLS estimator for $\rho \ne -.9$ for small sample size. For moderate and large sample sizes, OLS appears to be as nearly as efficient as the GLS estimator for $|\rho| \le .9$. In addition for estimating the slope, regardless of

14

the sample size, OLS performs nearly as efficiently as the GLS estimator for $|\rho| \leq .3$. Otherwise, OLS performs poorly.

4.1.3 Standard Normal Design Vector for AR(1) Process

Table (4.4) represents the relative efficiencies of the variances of GLS to OLS for 1000 standard Normal design with an intercept term. We generated a standard normal vector of length 1000, which gives more accurate and fair comparisons between the OLS and GLS estimators than using one random sample of standard Normal design. The relative efficiency of the OLS estimator as compared to the GLS estimator decreases with increasing values of $|\rho|$. For estimating an intercept term, regardless of the sample size, OLS appears to be nearly as efficient as the GLS for $\rho \neq -.9$. In addition, for estimating the slope, OLS performs nearly as efficiently as the GLS estimator for $|\rho| \leq .1$. Otherwise, OLS performs poorly.

4.2 Performance Comparisons for AR (2) Process

We discuss the relative efficiencies of OLS to GLS for linear, quadratic, exponential, and 1000 standard Normal design vectors when the disturbance term follows an AR(2) process, $u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \varepsilon_t, t = 1, 2, \ldots, T$, assuming that the autoregressive coefficients ϕ_1 and ϕ_2 are known priori. The three finite sample sizes used are 50, 100, and 200 for the selected 45 pairs of the autoregressive coefficients. These coefficients were chosen according to stationary condition (2.9) and so that $\rho_1 = \phi_1 (1 - \phi_2)^{-1}$ is positive. This second condition was chosen since this is the case in most econometric studies.

4.2.1 Linear and Quadratic Design Vectors for AR(2) Process

Figures (4.1) and (4.2) represent the relative efficiencies of the variances of GLS to OLS for a regression coefficient on linear and quadratic design vectors with an intercept. To demonstrate the efficiency of OLS, consider the linear design vector. When the disturbance term follows an AR(2) process for the linear design with small sample size, OLS performs nearly as efficiently as GLS for estimating the slope for all AR(2) parametrizations except when ϕ's are close to the stationary boundary. As the sample size increases the difference between the performance of OLS and GLS decreases. Only when $\phi_2 = -.9$, does OLS perform badly regardless of the sample size. The efficiency of GLS to OLS for the quadratic design mimics the behavior for the linear design.

15

Exponential and Standard Normal Design Vectors for AR(2) Process
Figures (4.3) and (4.4) represent the relative efficiencies of the variances of GLS to OLS for a regression coefficient on exponential and standard Normal design vectors with an intercept. The efficiency of OLS appears to be nearly as efficient as GLS for $\phi_1 = .2$ and small values of ϕ_2 for all sample sizes. Otherwise, OLS performs poorly.

5   Summary and Future Research

5.1   Summary

This paper has investigated an important statistical problem concerning estimation of the regression coefficients in the presence of autocorrelated disturbances. In particular, we have discussed the comparison of efficiency of the ordinary least squares (OLS) estimation to alternative procedures such as generalized least squares (GLS) estimator in the presence of first and second order autoregressive disturbances. Both stochastic and non-stochastic design vectors were used with different sample sizes.

We have found that regardless of the sample size, design vector, and order of the auto-correlated disturbances the relative efficiency of the OLS estimator generally increases with decreasing values of the disturbance variances. In particular, if the disturbance structure is a first or second order autoregressive and the dependent variable is nonstochastic and linear or quadratic, OLS performs nearly as well as its competitors for small values of the disturbance variances. The gain in efficiency of the GLS estimator for different design vectors such as exponential and standard normal compared to the OLS estimator is substantial for moderate and large values of the autoregressive coefficient in the case of an AR(1) process and large values of the disturbance variance in the presence of an AR(2) process. However, for small values of the autoregressive coefficient and disturbance variance the OLS estimator appears to be nearly as efficient as the GLS estimator. We believe, this work is important and provides a good foundation upon which to build a research career.

5.2   Future Research

Using a computer simulations to compare performance of different estimators and characterization of the effect of the design on the efficiency of OLS. In addition, examine the sensitivity of estimators to model misspecification. In particular, how do estimators perform when the true errors are ARMA(p,q), heteroskedastic and not an AR(1) or an AR(2)?

Koreisha et al., 2004 investigated the impact that EIGLS correction may have on forecast performance. They developed a new procedure for generating forecasts for regression models with auto-correlated disturbances based on OLS and a finite AR process. They found that for predictive purposes there is not much gained in trying to identifying the actual order and form of the auto-correlated disturbances or using more complicated estimation methods such as GLS or maximum likelihood estimation (MLE) procedures which often require inversion of large matrices. Extend Koreisha et al., 2004 results for different design vectors of the independent variables including both stochastic and nonstochastic designs instead of using one independent variable generated by an AR(1) process as in their investigation.

Finally, the long range goal is the creation of guidelines or rules of thumb which will aid the practitioner when deciding which regression estimation procedure to use.

References
[1] Chipman J., Kadiyala K., Madansky A., and Pratt J., 1968 - *Efficiency of the Sample Mean when Residuals Follow a First-Order Stationary Markoff Process*, Journal of the American Statistical Association, 63, 1237-1246.

[2] Chipman J., 1979 - *Efficiency of Least Squares Estimation of Linear Trend When Residuals Are Autocorrelated*, Econometrica, 47, 115-128.

[3] Cochrane, D., and Orcutt G., 1949, *application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms*, Journal of the American Statistical Association, 44, 32-61.

[4] Greene W., 2000 - **Econometric Analysis**. Prentice Hall, 4th ed, New Jersey, p 537.

[5] Judge G., Griffiths W., Hill R., Lutkepohl H., and Lee T., 1985 - **The Theory and Practice of Econometrics**. John Wiley & Sons Inc., 2nd ed, New York, p 280

[6] Koreisha S., Fang Y., 2004 - *Forecasting with Serially Correlated Regression Models*, Journal of Statistical Computation and Simulation, 74, 625-649.

[7] Krämer W., 1980, - *Finite Sample Efficiency of Ordinary Least Squares in the Linear Regression Model with Autocorrelated Errors*, Journal of the American Statistical Association, 75, 1005-1009.

[8] Krämer W. and Marmol F., 2002 - *OLS-Based Asymptotic Inference in Linear Regression Models with Trending Regressors and AR(p)-Disturbances*, Communications in Statistics: Theory and Methods, 31, 261-270.

[9] Ullah A., Srivastava V., Magee L., and Srivastava A., 1983 - Estimation *of Linear Regression Model with Autocorrelated Disturbances*, Journal of Time Series Analysis, 4, 127-135.

A Tables

Table 4.1: Efficiency of GLS to OLS for Linear Design for an AR(1) Process

ρ	Intercept			Slope		
	T = 50	T =100	T = 200	T = 50	T =100	T = 200
-0.9	0.709722	0.827579	0.904673	0.673894	0.801225	0.888135
-0.7	0.916218	0.955223	0.976822	0.902408	0.947115	0.972414
-0.5	0.969378	0.984030	0.991843	0.963956	0.981026	0.990262
-0.3	0.990766	0.995221	0.997569	0.989077	0.994307	0.997094
-0.1	0.999057	0.999512	0.999752	0.998880	0.999417	0.999703
0.1	0.999067	0.999514	0.999752	0.998891	0.999420	0.999703
0.3	0.991111	0.995306	0.997590	0.989409	0.994387	0.997114
0.5	0.971662	0.984605	0.991988	0.966167	0.981571	0.990396
0.7	0.928789	0.958518	0.977663	0.914700	0.950257	0.973204
0.9	0.835916	0.869111	0.916408	0.800006	0.841802	0.899341

Table 4.2: Efficiency of GLS to OLS for Quadratic Design for an AR(1) Process

ρ	Intercept			Slope		
	T = 50	T =100	T = 200	T = 50	T =100	T = 200
-0.9	0.785949	0.879485	0.935543	0.665943	0.794957	0.884114
-0.7	0.942688	0.970104	0.984725	0.899079	0.945127	0.971323
-0.5	0.979452	0.989449	0.994654	0.962631	0.980284	0.989869
-0.3	0.993858	0.996857	0.998411	0.988661	0.994080	0.996975
-0.1	0.999375	0.999680	0.999838	0.998837	0.999394	0.999691
0.1	0.999384	0.999682	0.999838	0.998847	0.999396	0.999691
0.3	0.994145	0.996927	0.998428	0.988984	0.994158	0.996995
0.5	0.981370	0.989925	0.994772	0.964783	0.980813	0.989999
0.7	0.953456	0.972865	0.985420	0.911006	0.948173	0.972088
0.9	0.900305	0.915950	0.945546	0.784938	0.833951	0.894946

Table 4.3: Efficiency of GLS to OLS for Exponential Design for an AR(1) Process

ρ	Intercept			Slope		
	T = 50	T =100	T = 200	T = 50	T =100	T = 200
-0.9	0.850797	0.920629	0.958894	0.213342	0.213391	0.213392
-0.7	0.962295	0.980965	0.990437	0.546231	0.546224	0.546223
-0.5	0.986633	0.993324	0.996664	0.776286	0.776269	0.776265
-0.3	0.996014	0.998014	0.999009	0.921240	0.921225	0.921222
-0.1	0.999594	0.999797	0.999899	0.991346	0.991343	0.991342
0.1	0.999597	0.999798	0.999899	0.991349	0.991343	0.991342
0.3	0.996127	0.998041	0.999016	0.921337	0.921247	0.921227
0.5	0.987390	0.993507	0.996709	0.776854	0.776401	0.776297
0.7	0.966584	0.982027	0.990702	0.548680	0.546795	0.546361
0.9	0.904580	0.935158	0.962704	0.228933	0.217015	0.214260

Table 4.4: Efficiency of GLS to OLS for 1000 Standard Normal Design for an AR(1) Process

ρ	Intercept			Slope		
	T = 50	T =100	T = 200	T = 50	T =100	T = 200
-0.9	0.729969	0.829670	0.898115	0.138344	0.122470	0.113319
-0.7	0.912499	0.951831	0.974205	0.370766	0.356619	0.349046
-0.5	0.968064	0.983101	0.991245	0.620681	0.610232	0.604882
-0.3	0.991023	0.995325	0.997609	0.845323	0.839997	0.837356
-0.1	0.999195	0.999584	0.999788	0.981678	0.980926	0.980554
0.1	0.999333	0.999656	0.999825	0.981819	0.980997	0.980589
0.3	0.994736	0.997279	0.998617	0.848641	0.841676	0.838171
0.5	0.985596	0.992455	0.996139	0.632649	0.616263	0.607906
0.7	0.965943	0.981420	0.990293	0.394859	0.368533	0.355393
0.9	0.905753	0.935963	0.962884	0.180712	0.141191	0.123911

B Figures

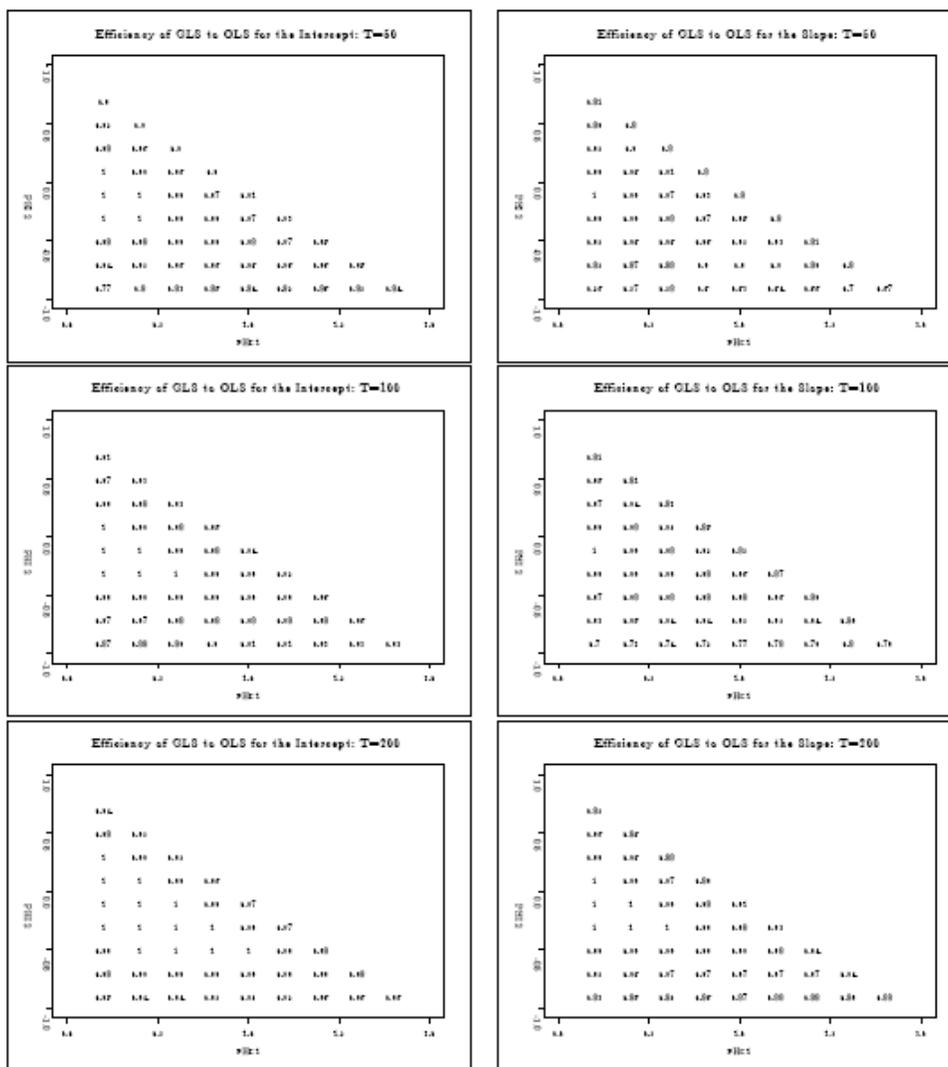Figure 4.1: Efficiency of GLS to OLS for Linear Design for AR(2) Process

Figure 4.2: Efficiency of GLS to OLS for Quadratic Design for an AR(2) Process

Efficiency of OLS to OLS for the Intercept: T=50

Efficiency of OLS to OLS for the Slope: T=50

Efficiency of OLS to OLS for the Intercept: T=100

Efficiency of OLS to OLS for the Slope: T=100

Efficiency of OLS to OLS for the Intercept: T=200

Efficiency of OLS to OLS for the Slope: T=200

22

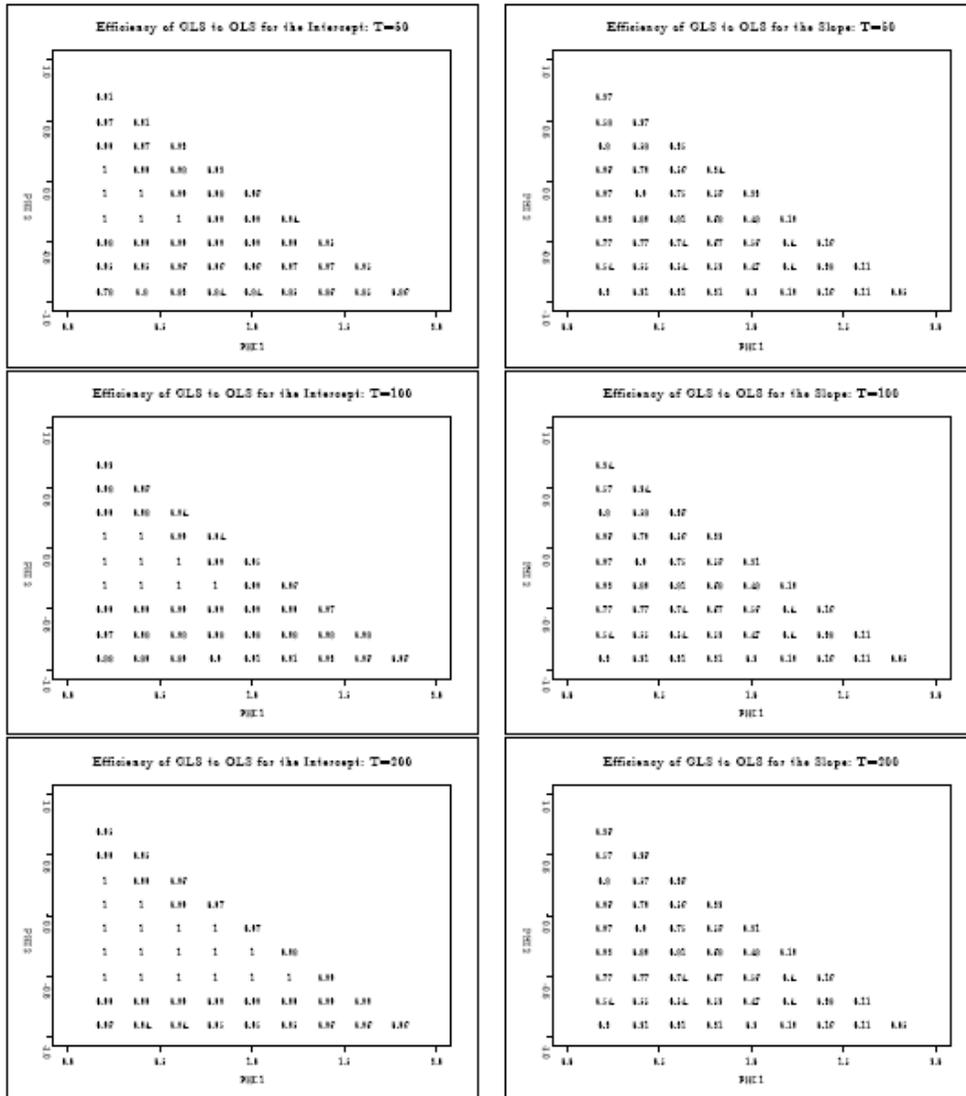Figure 4.3: Efficiency of GLS to OLS for Exponential Design for AR(2)
Process

Figure 4.4: Efficiency of GLS to OLS for 1000 Standard Normal Design for an AR(2) Process