

# Generalized least squares with misspecified serial correlation structures

Sergio G. Koreisha and Yue Fang

*University of Oregon, Eugene, USA*

[Received April 2000. Final revision February 2001]

**Summary.** The regression literature contains hundreds of studies on serially correlated disturbances. Most of these studies assume that the structure of the error covariance matrix  $\Omega$  is known or can be estimated consistently from data. Surprisingly, few studies investigate the properties of estimated generalized least squares (GLS) procedures when the structure of  $\Omega$  is incorrectly identified and the parameters are inefficiently estimated. We compare the finite sample efficiencies of ordinary least squares (OLS), GLS and incorrect GLS (IGLS) estimators. We also prove new theorems establishing theoretical efficiency bounds for IGLS relative to GLS and OLS. Results from an exhaustive simulation study are used to evaluate the finite sample performance and to demonstrate the robustness of IGLS estimates *vis-à-vis* OLS and GLS estimates constructed for models with known and estimated (but correctly identified)  $\Omega$ . Some of our conclusions for finite samples differ from established asymptotic results.

**Keywords:** Autoregressive disturbances; Generalized least squares; Incorrect generalized least squares; Relative efficiency; Serial correlation

## 1. Introduction

The topic of serial correlation in regression has been studied extensively since Cochrane and Orcutt (1949) and Durbin and Watson (1950) developed an approximate transformation to deal with and to test for autoregressive (AR) disturbances of order 1. It is probably fair to say that during the past half-century developments in this field have tracked advances in computational power closely. In the early 1950s and 1960s most of the research focused on transformations (approximate or exact) that were developed for ordinary least squares (OLS) estimation of low order AR and moving average error structures. With the advent of bigger and faster computers since the 1970s research has been directed more towards estimation methods such as generalized least squares (GLS) (or approximations thereof) and maximum likelihood procedures which often require the inversion of large matrices. In the last two decades we have also witnessed a plethora of simulation studies dealing with more complex serial correlation structures, but for surprisingly small sample sizes, generally ranging only from 10 to 40 observations, e.g. Zinde-Walsh and Galbraith (1991). (See Hildreth (1986) and Choudhury *et al.* (1999) for more detailed chronologies of major developments in this area.)

Most of these studies assumed that the error covariance matrix  $\Omega$  from the regression model,

$$Y = X\beta + \epsilon \tag{1}$$

*Address for correspondence:* Sergio G. Koreisha, Department of Decision Sciences, Charles H. Lundquist College of Business, University of Oregon, Eugene, OR 97403-1208, USA.  
E-mail: Sergiok@Oregon.UOREGON.EDU

where  $Y$  is an  $n \times 1$  vector of observations on a dependent variable,  $X$  is an  $n \times k$  design matrix and  $\epsilon$  is a random vector with  $E(\epsilon) = 0$  and  $E(\epsilon'\epsilon) = \sigma^2\Omega$ , was either known or could be estimated consistently from data. Very few studies, however, considered the properties of estimators when the structure of  $\Omega$  was incorrectly identified or when its parameters were inefficiently estimated. Moreover, of the few studies that considered these issues (notably Amemiya (1973) and Engle (1974)) most dealt with or depended on asymptotic results which as we shall demonstrate can often be at odds with results obtained from the sample sizes that are generally available to model builders.

The challenges of identifying the structure of the covariance matrix  $\Omega$  when GLS is used to estimate the regression model with autocorrelated disturbances have been studied by Walker (1967), Kadiyala (1970), King (1983) and Koreisha and Pukkila (1987) among others. Walker (1967) and King (1983) discussed the practical difficulties of testing for AR(1) against moving average (MA(1)) disturbances. Kadiyala (1970) has shown that the null hypothesis that  $\epsilon$  follows  $N(0, \Omega_0)$  cannot be tested against the alternative hypothesis that  $\epsilon$  follows  $N(0, \Omega_1)$  under certain conditions on  $\Omega_i$  ( $i = 0, 1$ ) and the design matrix  $X$ . The test aside from a scalar factor may have the same distribution for both the null and the alternative hypotheses. Moreover, as pointed out by Thursby (1987), if the regression suffers from omitted variables, then it may not be possible to identify the form of the autocorrelation in the regression model.

For sample sizes that are generally available to model builders (50–200 observations) the ‘correct’ identification of the ARMA( $p, q$ ) process that is associated with the serially correlated disturbances can be quite elusive. On the basis of a small simulation study, for example, for the same regression models that will be described in Section 4, we found that the correct identification using automatic order selection criteria of the simulated order of the Gaussian ARMA(1, 1) structures governing the estimated residuals was at best tenuous. The identification performance of Schwarz’s Bayesian information criterion improved substantially as the sample size increased, but it was scarcely adequate, varying from nearly 50% when  $n = 50$  to 80% when  $n$  reached 200. The Akaike information criterion, as also noted in other studies, severely overestimated the order of the process and, at least for the model structures that we studied, its performance did not improve when the sample size increased. Further details of these simulation results are available from the authors.

In this paper we shall compare the finite sample efficiencies of OLS and GLS *vis-à-vis* incorrect GLS (IGLS) estimators, i.e. GLS estimators based on a wrongly identified  $\Omega$ . We shall assume that the serially correlated disturbances  $\epsilon$  in model (1) follow an ARMA process (Box and Jenkins, 1976):

$$\Phi(B)\epsilon_t = \Theta(B)a_t, \quad (2)$$

where  $\Phi(B)$  and  $\Theta(B)$  are finite polynomials of orders  $p$  and  $q$  respectively in the backshift operator  $B$ , such that  $B^j w_t = w_{t-j}$ , and  $\{a_t\}$  is a Gaussian white noise process with variance  $\sigma_a^2$ . We say that process (2) is both stationary and invertible if the roots of the characteristic equations  $\Phi(B) = 0$  and  $\Theta(B) = 0$  are outside the unit circle. Stationary and invertible ARMA( $p, q$ ) models can also be expressed as either infinite autoregressions,  $\Pi(B)\epsilon_t = a_t$ , or infinite moving averages,  $\epsilon_t = \psi(B)a_t$ . In practice, however, such representations may be approximated by processes of relatively low order because the coefficients in  $\Pi(B)$  and  $\Psi(B)$  may be effectively 0 beyond some finite lag.

The paper will be organized as follows. In Section 2 we establish conditions for IGLS to yield unbiased and consistent regression estimators. In Section 3 we prove new theorems

establishing finite efficiency bounds for IGLS relative to GLS and OLS. In Section 4 we present results from an exhaustive simulation study for sample sizes ranging from 50 to 200 observations and covering a wide spectrum of ARMA( $p, q$ ) serial correlation structures. There we contrast the quality and properties of regression estimators based on IGLS procedures which assume that the serial correlations follow different order AR processes, with OLS estimates and with estimates obtained from GLS methods with known and estimated  $\Omega$ . Finally in Section 5 we offer some concluding remarks.

**2. Some properties of the estimators**

Suppose that the linear regression model is specified by equations (1) and (2). Consider the estimator of  $\beta$

$$\hat{\beta} = (X' \Xi^{-1} X)^{-1} X' \Xi^{-1} Y. \tag{3}$$

Setting  $\Xi = I$  in equation (3) defines the OLS estimator  $\hat{\beta}_{OLS}$ ; if  $\Xi = \Omega$ , then equation (3) yields the well-known Aitken estimator  $\hat{\beta}_{GLS}$ . If the structure of  $\Omega$  is known, but its elements must be estimated, i.e.  $\Xi = \hat{\Omega}$ , then we shall refer to estimator (3) as estimated GLS and denote it as  $\hat{\beta}_{EGLS}$ . We shall refer to estimator (3) as IGLS,  $\hat{\beta}_{IGLS}$ , if  $\Xi = \Sigma \neq \Omega$ , and as the estimated IGLS,  $\hat{\beta}_{EIGLS}$ , if  $\Xi = \hat{\Sigma}$ , i.e. when  $\Sigma$  is obtained from data.

The matrix  $\Xi$  will be a Toeplitz matrix, assuming that  $\epsilon$  is stationary. Moreover, if AR correction methods (AR( $\bar{p}$ )) are used, then  $\Xi^{-1}$  will be band diagonal with  $\bar{p}$  bands above and below the main diagonal (Wise, 1955).

GLS estimates are unbiased and consistent (Theil, 1971; Schmidt, 1976; Judge *et al.*, 1985) if three conditions hold:

- (a)  $X$  is non-stochastic;
- (b) the regressors (columns of  $X$ ) are linearly independent;
- (c)  $\lim_{n \rightarrow \infty} (n^{-1} X' X)$  is finite and non-singular.

Furthermore, if  $\epsilon$  follows a Gaussian distribution with mean 0 and variance  $\Omega$ , then the GLS estimator is also the best linear unbiased estimator (BLUE) (Aitken, 1935). When  $\Xi$  is fixed, it is well known that

$$\text{cov}(\hat{\beta}) = \sigma^2 (X' \Xi^{-1} X)^{-1} X' \Xi^{-1} \Omega \Xi^{-1} X (X' \Xi^{-1} X)^{-1}. \tag{4}$$

If  $\epsilon$  is symmetrically distributed around zero, and if  $\hat{\Omega}$  is an even function of  $\hat{\epsilon}$ , then  $\hat{\beta}_{EGLS}$  is unbiased, provided that  $E(\hat{\beta}_{EGLS})$  exists (Kakwani, 1967). Sufficient conditions for the EGLS estimator to be consistent are that  $\text{plim}_{n \rightarrow \infty} (n^{-1} X' \hat{\Omega}^{-1} X)$  is finite and non-singular, and that  $\text{plim}_{n \rightarrow \infty} (n^{-1} X' \hat{\Omega}^{-1} \epsilon) = 0$ . The asymptotic distribution of  $\hat{\beta}_{EGLS}$  can also be derived, and under some regularity assumptions

$$n^{1/2} (\hat{\beta}_{EGLS} - \beta) \xrightarrow{d} n^{1/2} (\hat{\beta}_{GLS} - \beta) \xrightarrow{d} N(0, \sigma^2 V^{-1}), \tag{5}$$

where  $V = \lim_{n \rightarrow \infty} (n^{-1} X' \Omega^{-1} X)$  (Theil, 1971). It should be also noted that consistent estimates of parameters in  $\Omega$  will not be, in general, a sufficient condition for  $\hat{\beta}_{EGLS}$  and  $\hat{\beta}_{GLS}$  to have the same asymptotic distribution (Schmidt, 1976).

Now since there is a non-singular unitary matrix  $\Lambda$  such that  $\Lambda' \Lambda = \Xi^{-1}$  (Hadley, 1961; Strang, 1988), then premultiplying equation (1) by  $\Lambda$  yields

$$\Lambda Y = \Lambda X \beta + \Lambda \epsilon. \tag{6}$$

Thus, if  $\lim_{n \rightarrow \infty} \{n^{-1}(\Delta X)'(\Delta X)\} = \lim_{n \rightarrow \infty} (n^{-1}X'\Xi^{-1}X)$  is finite and non-singular, then conditions (a)–(c) are satisfied. Therefore, the conditions required for consistency of IGLS and EIGLS can be similarly derived. Below we express them as a new theorem, which is our adaptation of GLS results in Schmidt (1976).

*Theorem 1.* If  $X$  is non-stochastic and the regressors (columns of  $X$ ) are linearly independent, then

- (a)  $\hat{\beta}_{IGLS}$  is unbiased, and, if  $\lim_{n \rightarrow \infty} (n^{-1}X'\Xi^{-1}X)$  is finite and non-singular, then  $\hat{\beta}_{IGLS}$  is consistent, and,
- (b) if  $\text{plim}_{n \rightarrow \infty} (n^{-1}X'\hat{\Xi}^{-1}\hat{\Omega}^{-1}\hat{\Xi}^{-1}X)$  is finite and non-singular and  $\text{plim}_{n \rightarrow \infty} (n^{-1}X'\hat{\Xi}^{-1}\hat{\Omega}^{-1}\Lambda\epsilon) = 0$ , then  $\hat{\beta}_{EIGLS}$  is consistent.

Similar results to theorem 1 also hold even if the regressor matrix  $X$  is stochastic, but the underlying assumptions in such a case are more complicated and restrictive. See, for example, Schmidt (1976) for discussions on general issues regarding stochastic regressors.

It is possible to derive the asymptotic distribution for  $\hat{\beta}_{IGLS}$  and  $\hat{\beta}_{EIGLS}$ . When exogenous variables are stationary, under certain regularity conditions, the normality of IGLS and EIGLS estimators can be derived by using Hansen’s (1982) generalized method of moments (GMM). The critical assumption needed in using the GMM is that the regression disturbance  $\epsilon_t$  is uncorrelated with the exogenous variables. Since the IGLS and EIGLS estimators are equivalent to OLS estimators for the transformed regression model (6), it follows that IGLS and EIGLS are special cases of the GMM (see, for example, Hamilton (1994) and Mátyás (1999) for detailed discussions of the GMM procedure).

Amemiya (1973) studied the properties of IGLS and EIGLS estimators based on theoretical and estimated autocovariances of serial correlated disturbances generated by ARMA processes. Assuming that  $\Xi$  had the structure of a large order AR process he showed that, as  $n \rightarrow \infty$  and the order of the AR structure approached  $\infty$  at an appropriate rate, the asymptotic distributions of both IGLS and EIGLS were the same as the asymptotic Gaussian disturbance of the best unbiased estimator.

The OLS estimator ( $\Xi = I$ ) is less efficient than the GLS estimator ( $\Xi = \Omega$ ) since

$$\text{cov}(\hat{\beta}_{OLS}) - \text{cov}(\hat{\beta}_{GLS}) = \sigma^2(X'X)^{-1}X'\Omega X(X'X)^{-1} - \sigma^2(X'\Omega^{-1}X)^{-1}$$

is non-negative. In fact, it can be shown that the OLS and the GLS estimators are identical if and only if  $X = Q\Gamma$ , where  $Q$  contains characteristic vectors of  $\Omega$ , and  $\Gamma$  is a non-singular matrix (Zyskind, 1967). In general, however, the loss of efficiency in using OLS instead of GLS can be substantial. For example, consider the model  $y_t = \beta_0 + \beta_1x_t + \epsilon_t$  with an AR(1) disturbance:  $\epsilon_t = \phi\epsilon_{t-1} + a_t$ . For large  $n$ , Rao and Griliches (1969) have shown that

$$\frac{\text{cov}(\hat{\beta}_{GLS})}{\text{cov}(\hat{\beta}_{OLS})} \approx \frac{(1 - \phi^2)(1 - \phi\pi)}{(1 + \phi^2 - 2\phi\pi)(1 + \phi\pi)}$$

This result holds for all possible configurations of exogenous variables. IGLS estimates are also less efficient than GLS estimates since the GLS estimator is the BLUE. The relative efficiency of IGLS to OLS, however, depends on  $\Omega$ ,  $\Xi$  and  $X$ , and will be the subject of the next section.

### 3. Finite sample efficiency relationships

In this section we compare the finite sample efficiencies of OLS, GLS and IGLS estimators.

We also establish efficiency bounds for IGLS relative to GLS and OLS. We assume that  $X$  is non-stochastic.

It is well known (see, for example, Theil (1971)) that, if serial correlation is present in the disturbances of regression models, then

$$|\text{COV}(\hat{\beta}_{\text{GLS}})| \leq |\text{COV}(\hat{\beta}_{\text{OLS}})|.$$

An upper bound for the loss of the estimation efficiency of OLS relative to GLS can be established by assuming that  $\Omega$  has  $n$  ordered eigenvalues:  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ . Hannan (1970) has shown that, for any  $c$ ,

$$1 \leq \frac{\text{COV}(c' \hat{\beta}_{\text{OLS}})}{\text{COV}(c' \hat{\beta}_{\text{GLS}})} \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1 \lambda_n}.$$

Watson (1967), Bloomfield and Watson (1975) and Knott (1975) have extended this result and have shown that

$$1 \leq \frac{|\text{COV}(\hat{\beta}_{\text{OLS}})|}{|\text{COV}(\hat{\beta}_{\text{GLS}})|} \leq \prod_{s=1}^{\min(k, n-k)} \frac{(\lambda_s + \lambda_{n-s+1})^2}{4\lambda_s \lambda_{n-s+1}}, \tag{7}$$

where  $k$  is the number of columns in the design matrix  $X$ . (There are other ways to order matrices. Krämer and Baltagi (1996), for example, used traces to order matrices in analysing the limiting efficiency of OLS in the general linear regression model.) The key result used by Bloomfield and Watson (1975) and Knott (1975) to obtain inequality (7) and which we shall use later can be stated as the following lemma.

*Lemma 1.* Let  $G$  be a diagonal matrix with distinct positive elements  $\zeta_1 < \zeta_2 < \dots < \zeta_n$ , and let  $Z$  be an  $n \times k$  matrix such that  $Z'Z = I$ ; then

$$|Z'GZ||Z'G^{-1}Z| \leq \prod_{s=1}^{\min(k, n-k)} \frac{(\zeta_s + \zeta_{n-s+1})^2}{4\zeta_s \zeta_{n-s+1}}. \tag{8}$$

The upper bound is attainable for a suitable choice of  $X$ .

Armed with these preliminary results we shall now prove two new theorems establishing the bounds for the relative efficiency of IGLS *vis-à-vis* GLS and OLS for finite samples.

*Theorem 2.* Let  $\nu_1 \leq \nu_2 \leq \dots \leq \nu_n$  be eigenvalues of  $\Lambda\Omega\Lambda'$  with  $\Lambda$  satisfying  $\Lambda'\Lambda = \Xi^{-1}$ ; then

$$1 \leq \frac{|\text{COV}(\hat{\beta}_{\text{IGLS}})|}{|\text{COV}(\hat{\beta}_{\text{GLS}})|} \leq \prod_{s=1}^{\min(k, n-k)} \frac{(\nu_s + \nu_{n-s+1})^2}{4\nu_s \nu_{n-s+1}}. \tag{9}$$

Given  $\Omega$  and  $\Xi$ , the upper bound is attainable for a suitable choice of  $X$ .

*Proof.* From equation (4),

$$\frac{|\text{COV}(\hat{\beta}_{\text{IGLS}})|}{|\text{COV}(\hat{\beta}_{\text{GLS}})|} = \frac{\sigma^2 |(X'\Xi^{-1}X)^{-1}X'\Xi^{-1}\Omega\Xi^{-1}X(X'\Xi^{-1}X)^{-1}|}{\sigma^2 |(X'\Omega^{-1}X)^{-1}|}.$$

Letting  $Z = \Lambda X$  and  $\Gamma = \Lambda\Omega\Lambda'$ , then

$$\frac{|\text{cov}(\hat{\beta}_{\text{IGLS}})|}{|\text{cov}(\hat{\beta}_{\text{GLS}})|} = \frac{|Z' \Lambda \Omega \Lambda' Z| |Z' (\Lambda \Omega \Lambda')^{-1} Z|}{|Z' Z|^2} = \frac{|Z' \Gamma Z| |Z' \Gamma^{-1} Z|}{|Z' Z|^2}.$$

Since  $\Gamma$  is symmetric, there is an  $n \times n$  matrix  $H$  such that  $\Gamma = HGH'$ , where  $G$  is a diagonal matrix with the eigenvalues of  $\Gamma$ . Hence,

$$\frac{|\text{cov}(\hat{\beta}_{\text{IGLS}})|}{|\text{cov}(\hat{\beta}_{\text{GLS}})|} = \frac{|W'GW| |W'G^{-1}W|}{|W'W|^2},$$

where  $W = H'Z$ . It can be seen directly that this ratio is unchanged if we replace  $W$  by any  $n \times k$  matrix  $V$  with columns spanning the space of the columns of  $W$ . Choosing such a  $V$  with  $V'V = I$  and applying lemma 1 establishes the upper bound. The lower bound of  $|\text{cov}(\hat{\beta}_{\text{IGLS}})|/|\text{cov}(\hat{\beta}_{\text{GLS}})|$  in inequality (9) is trivial since  $\hat{\beta}_{\text{GLS}}$  is the BLUE.

*Theorem 3.* Let  $\lambda_1 < \lambda_2 < \dots < \lambda_n$  be eigenvalues of  $\Omega$  and  $\nu_1 \leq \nu_2 \leq \dots \leq \nu_n$  be the corresponding eigenvalues of  $\Lambda \Omega \Lambda'$  with  $\Lambda$  satisfying  $\Lambda' \Lambda = \Xi^{-1}$ ; then

(a)

$$\frac{|\text{cov}(\hat{\beta}_{\text{IGLS}})|}{|\text{cov}(\hat{\beta}_{\text{GLS}})|} \prod_{s=1}^{\min(k, n-k)} \frac{4\lambda_s \lambda_{n-s+1}}{(\lambda_s + \lambda_{n-s+1})^2} \leq \frac{|\text{cov}(\hat{\beta}_{\text{IGLS}})|}{|\text{cov}(\hat{\beta}_{\text{OLS}})|} \leq \frac{|\text{cov}(\hat{\beta}_{\text{GLS}})|}{|\text{cov}(\hat{\beta}_{\text{OLS}})|} \prod_{s=1}^{\min(k, n-k)} \frac{(\nu_s + \nu_{n-s+1})^2}{4\nu_s \nu_{n-s+1}}, \tag{10}$$

(b)

$$\prod_{s=1}^{\min(k, n-k)} \frac{4\lambda_s \lambda_{n-s+1}}{(\lambda_s + \lambda_{n-s+1})^2} \leq \frac{|\text{cov}(\hat{\beta}_{\text{IGLS}})|}{|\text{cov}(\hat{\beta}_{\text{OLS}})|} \leq \prod_{s=1}^{\min(k, n-k)} \frac{(\nu_s + \nu_{n-s+1})^2}{4\nu_s \nu_{n-s+1}}. \tag{11}$$

Both lower and upper bounds in inequality (10) are attainable for a suitable choice of  $X$ .

*Proof.*

$$\frac{|\text{cov}(\hat{\beta}_{\text{IGLS}})|}{|\text{cov}(\hat{\beta}_{\text{OLS}})|} = \frac{\sigma^2 |(X' \Xi^{-1} X)^{-1} X' \Xi^{-1} \Omega \Xi^{-1} X (X' \Xi^{-1} X)^{-1}|}{\sigma^2 |(X' X)^{-1} X' \Omega X (X' X)^{-1}|} = \frac{|X' \Xi^{-1} \Omega \Xi^{-1} X| |X' X|^2}{|X' \Xi^{-1} X|^2 |X' \Omega X|}.$$

Setting that  $Z = \Lambda X$  and  $\Gamma = \Lambda \Omega \Lambda'$  we have

$$\frac{|\text{cov}(\hat{\beta}_{\text{IGLS}})|}{|\text{cov}(\hat{\beta}_{\text{OLS}})|} = \frac{|Z' \Gamma Z| |Z' \Gamma^{-1} Z|}{|Z' Z|^2} \frac{|X' X|^2}{|X' \Omega X| |X' \Omega^{-1} X|}.$$

The bounds specified in inequalities (10) and (11) follow by applying the results of lemma 1 and theorem 2.

To gain some idea of the magnitude of these relative bounds consider the case for which there is only one exogenous variable and the intercept coefficient of the regression is set to 0. In this simple case,  $k = 1$ , theorems 2 and 3, part (b), lead to the following corollaries.

*Corollary 1.* Let  $\nu_1$  and  $\nu_n$  be the smallest and the largest eigenvalues of  $\Lambda \Omega \Lambda'$ , with  $\Lambda$  satisfying  $\Lambda' \Lambda = \Xi^{-1}$ ; then

$$1 \leq \frac{\text{cov}(\hat{\beta}_{\text{IGLS}})}{\text{cov}(\hat{\beta}_{\text{GLS}})} \leq \mathcal{M}_U, \tag{12}$$

where  $\mathcal{M}_U \equiv (\nu_1 + \nu_n)^2 / (4\nu_1 \nu_n)$ .

This corollary is also known as the Kantorovich inequality (see Marcus and Ming (1964)).

*Corollary 2.* Let  $\lambda_1$  and  $\lambda_n$  be the smallest and the largest eigenvalues of  $\Omega$ , and let  $\nu_1$  and  $\nu_n$  be the smallest and the largest eigenvalues of  $\Lambda\Omega\Lambda'$ , with  $\Lambda$  satisfying  $\Lambda'\Lambda = \Xi^{-1}$ ; then

$$\mathcal{M}_L \leq \frac{\text{cov}(\hat{\beta}_{\text{IGLS}})}{\text{cov}(\hat{\beta}_{\text{OLS}})} \leq \mathcal{M}_U, \tag{13}$$

where  $\mathcal{M}_U$  is given in corollary 1 and  $\mathcal{M}_L \equiv 4\lambda_1\lambda_n/(\lambda_1 + \lambda_n)^2$ .

In Table 1 we present the limits ( $\mathcal{M}_U$  and  $\mathcal{M}_L$ ) for the relative efficiency of IGLS–AR( $\bar{p}$ ),  $\bar{p} = \{1, 4, 7\}$ , *vis-à-vis* OLS and GLS when the disturbance term of the regression model follows an MA(1) process with  $|\theta_1|$  ranging from 0.1 to 0.9. Also included in Table 1 are the corresponding eigenvalues  $\lambda_1$  and  $\lambda_n$ , and  $\nu_1$  and  $\nu_n$ . Since the tabulated values are not very sensitive to the sample size due to the fact that the eigenvalues required in evaluating both  $\mathcal{M}_U$  and  $\mathcal{M}_L$  do not change much with the sample size especially when  $|\theta_1|$  is not near 1, for brevity we have only reported results associated with  $n = 50$ .

As can be seen,  $\mathcal{M}_U$  and  $\mathcal{M}_L$  do not depend on the sign of the moving average coefficient  $\theta_1$ . When  $|\theta_1|$  is small, both  $\nu_1$  and  $\nu_n$  are close to 1 and so is  $\mathcal{M}_U$ , implying that the loss of efficiency in using IGLS relative to GLS is small. For example, for  $|\theta_1| = 0.4$ ,  $\mathcal{M}_U$  is equal to 1.0638, 1.0004 and 1.0000 for  $\bar{p} = 1, 4, 7$  respectively. As  $|\theta_1|$  increases (say,  $|\theta_1| > 0.5$ ), the difference between  $\nu_1$  and  $\nu_n$  increases, thus, forcing  $\mathcal{M}_U$  to increase away from 1, indicating that the loss of efficiency in using IGLS relative to GLS can be substantial for some design matrices  $X$ . This loss of efficiency, however, decreases substantially as the order of the AR( $\bar{p}$ ) process used to represent the MA(1) simulated disturbance structure increases, as can be observed, for example, when  $|\theta_1| = 0.8$  and  $\bar{p} = 1, 4, 7$ . These are respectively 5.42, 1.49 and 1.10.

In contrast,  $\mathcal{M}_L$  is less than 1 for all values of  $\theta_1$  and orders  $\bar{p}$  studied, indicating that OLS can be considerably less efficient than IGLS. Note, for instance, when  $|\theta_1| = 0.5$ , that  $\lambda_n$  is almost 10 times larger than  $\lambda_1$ , thus yielding an  $\mathcal{M}_L$ -value of about 0.36. Although theoretically the upper bound for the relative efficiency of IGLS *vis-à-vis* OLS can be greater than 1, on the basis of the simulation study to be discussed in Section 4, the efficiency in estimating the slope of a regression model with one exogenous variable using EIGLS–AR( $\bar{p}$ )

**Table 1.** Bounds for the relative efficiency of IGLS–AR( $\bar{p}$ ) *vis-à-vis* OLS and GLS when the serially correlated disturbance follows an MA(1) process

$\bar{p}$		Results for the following values of $ \theta_1 $ :								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	$\mathcal{M}_U$	1.0025	1.0063	1.0325	1.0638	1.2795	1.4489	1.7580	5.4164	10.7587
	$\nu_1$	0.9237	0.9237	0.8361	0.8361	0.6036	0.6036	0.6036	0.2315	0.2315
	$\nu_n$	1.0101	1.0828	1.1965	1.3786	1.6628	2.1191	2.9123	4.5399	9.4921
4	$\mathcal{M}_U$	1.0000	1.0000	1.0000	1.0004	1.0037	1.0228	1.1110	1.4901	3.7109
	$\nu_1$	1.0000	0.9994	0.9953	0.9804	0.9415	0.8609	0.7233	0.5279	0.2988
	$\nu_n$	1.0000	1.0006	1.0047	1.0200	1.0624	1.1632	1.3917	1.9476	3.8138
7	$\mathcal{M}_U$	1.0000	1.0000	1.0000	1.0000	1.0001	1.0010	1.0114	1.1006	1.8710
	$\nu_1$	1.0000	1.0000	0.9999	0.9988	0.9928	0.9695	0.8995	0.7366	0.4566
	$\nu_n$	1.0000	1.0000	1.0001	1.0012	1.0072	1.0316	1.1133	1.3750	2.4179
1, 4, 7	$\mathcal{M}_L$	0.9609	0.8526	0.6981	0.5262	0.3624	0.2244	0.1204	0.0518	0.0148
	$\lambda_1$	0.8350	0.7226	0.6454	0.5923	0.5560	0.5317	0.5160	0.5067	0.5019
	$\lambda_n$	1.2463	1.6231	2.2193	3.2087	4.9624	8.3808	16.0987	38.1088	134.9336

procedures is always superior to OLS. Again, we emphasize that the relative efficiency of IGLS over OLS depends on the covariance structure of  $\epsilon$  ( $\Omega$ ), its AR approximation matrix ( $\Xi$ ) and the exogenous variables  $X$ .

There are other ways to measure the ‘relative severity of the consequences of misspecification in statistical models’. Zinde-Walsh (1990) for instance, in a study of stationary stochastic processes, used a measure based on distances (expected mean-squared error (MSE)) between two data-generating processes to determine the consequences of misrepresenting an MA(1) process by an estimated AR(1) model.

#### 4. Finite sample properties of estimated incorrect generalized least squares

In this section we shall compare the finite sample efficiencies of EIGLS relative to GLS, EGLS and OLS. We shall focus on only AR( $\tilde{p}$ ) GLS corrections for disturbances generated by mixed ARMA( $p, q$ ) processes. This type of correction is probably the most widely used by practitioners, at least on the basis of text-book coverage, and one which has been implemented in many commonly employed statistical packages such as SAS and S-PLUS. Its origin can be traced all the way back to Cochrane and Orcutt (1949) when they developed a simple transformation for the OLS estimation of linear models with an AR(1) disturbance term.

##### 4.1. The simulation set-up

To evaluate the quality of the regression estimates obtained from EIGLS–AR( $\tilde{p}$ ) procedures we conducted an exhaustive simulation study to compare their finite sample performance with the corresponding OLS estimates and with estimates obtained from GLS procedures with known and estimated  $\Omega$ . We generated for sample sizes of 50, 100 and 200 observations 1000 realizations for each of a variety of stationary and invertible Gaussian ARMA( $p, q$ ) structures with varying parameter values as the residuals for the regression model with one exogenous variable generated by an AR(1) process. The parameter values for the residual ARMA structures were chosen not only to conform with other previously published studies such as Glasbey (1982), Kallinen *et al.* (1990), Zinde-Walsh and Galbraith (1991) and Koreisha and Fang (1999) but also to provide a representative set of examples of possible autocorrelated error structures in regression models. Although we simulated tens of thousands of trials, for brevity we shall report only a subset of our simulations. A more comprehensive tabulation of the study including structures with other stationary and invertible Gaussian ARMA( $p, q$ ) disturbance terms are available from the authors.

Using the International Mathematical and Statistical Libraries’ (1999) random-number subroutine RNARM we constructed the regression model

$$y_t = 2.0 + 0.5x_t + \epsilon_t,$$

where the generating process for the exogenous variable  $x_t$  followed an AR(1) process,  $(1 - \pi B)x_t = v_t$ , with  $v_t \sim \text{IN}(0, 1)$ , and  $E(\epsilon_t, v_s) = 0, \forall t \neq s$ , and  $\pi = \{0.0, 0.5, 1.0\}$ . For a given sample size only one set of random numbers was generated for each of the AR(1) model structures of the exogenous variable. Breusch (1980) has shown that, for a fixed regressor, the distribution of  $(\hat{\beta}_{\text{EIGLS}} - \beta)/\sigma$  does not depend on  $\beta$  and  $\sigma^2$ . Moreover, the result holds even if the covariance matrix  $\Omega$  is misspecified (Breusch (1980), page 331). This implies that, in simulation studies, only one point in the parameter space for  $(\beta, \sigma^2)$  needs to be considered for EIGLS. When the regressor is stochastic, the assumption of a fixed regressor can be construed as conditioning on a given realization of the regressor, provided that the regressor is independent of  $\epsilon_t$ .



Tables 2–6 contrast the efficiency of the GLS estimates relative to that of OLS in terms of the MSE,

$$\hat{\zeta}_{\beta_i} \equiv \sum (\hat{\beta}_{i, \text{GLS}} - \beta_i)^2 / \sum (\hat{\beta}_{i, \text{OLS}} - \beta_i)^2,$$

where  $i = 0, 1$ , for five GLS estimates: the estimate based on the correct residual model structures and known ARMA coefficients (denoted GLS), the estimate based on the correct residual model structures but with estimated ARMA coefficients (denoted EGLS); the estimate based on an AR(1) correction with an estimated AR coefficient (denoted EIGLS–AR(1)) and two other EIGLS–AR( $\bar{p}$ ) estimates with lags  $\bar{p}$  equal to the closest integer part of  $n^{1/2}/2$  and  $n^{1/2}$ . Estimates for the AR parameters used in the EIGLS–AR( $\bar{p}$ ) correction were obtained by using unconditional least squares, also referred to as non-linear least squares (Spitzer, 1979). A ratio less than 1 indicates that the GLS estimates are more efficient than OLS. We have included the efficiencies associated with GLS only as a bench-mark for comparison since in practice we never really know the true structure of the covariance matrix,

**Table 2.** Relative efficiencies for AR(1) error processes

Sample size	GLS type	Relative efficiencies for the following parameter values:													
		$\phi_1 = 0.3$						$\phi_1 = 0.5$							
		$\pi = 0.0$		$\pi = 0.5$		$\pi = 1.0$		$\pi = 0.0$		$\pi = 0.5$		$\pi = 1.0$			
		$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$		
50	GLS	0.992	0.868	0.992	0.827	0.990	0.917	0.982	0.647	0.981	0.555	0.975	0.752		
	EGLS	0.992	0.873	0.994	0.870	0.994	0.961	0.983	0.648	0.985	0.580	0.983	0.829		
	EIGLS–AR(1)	0.992	0.873	0.994	0.870	0.994	0.961	0.983	0.648	0.985	0.580	0.983	0.829		
	EIGLS–AR(4)	1.006	0.924	1.006	0.935	1.008	1.009	0.999	0.678	0.999	0.630	1.000	0.885		
	EIGLS–AR(7)	1.015	1.016	1.025	1.013	1.017	1.069	1.003	0.762	1.013	0.715	1.009	0.966		
100	GLS	0.998	0.810	0.996	0.831	0.996	0.966	0.993	0.563	0.989	0.574	0.987	0.876		
	EGLS	0.999	0.821	0.998	0.850	0.996	0.976	0.994	0.570	0.992	0.591	0.988	0.898		
	EIGLS–AR(1)	1.010	0.893	1.010	0.918	1.007	1.004	1.001	0.612	1.001	0.646	0.996	0.931		
	EIGLS–AR(5)	1.010	0.893	1.010	0.918	1.007	1.004	1.001	0.612	1.001	0.646	0.996	0.931		
	EIGLS–AR(10)	1.018	0.960	1.021	0.969	1.014	1.036	1.010	0.628	1.011	0.691	1.005	0.968		
200	GLS	1.001	0.839	1.001	0.834	0.999	0.990	1.000	0.600	0.999	0.577	0.989	0.954		
	EGLS	1.001	0.847	1.000	0.847	0.998	0.995	1.000	0.602	0.998	0.584	0.988	0.962		
	EIGLS–AR(1)	0.998	0.874	0.998	0.885	0.994	0.999	0.997	0.612	0.995	0.602	0.983	0.972		
	EIGLS–AR(7)	0.998	0.874	0.998	0.885	0.994	0.999	0.997	0.612	0.995	0.602	0.983	0.972		
	EIGLS–AR(14)	1.001	0.927	1.001	0.930	0.990	1.020	1.001	0.645	1.000	0.634	0.981	0.992		
				$\phi_1 = 0.9$						$\phi_1 = -0.9$					
50	GLS	0.808	0.169	0.807	0.104	0.770	0.157	0.935	0.106	0.919	0.126	0.869	0.394		
	EGLS	0.896	0.168	0.895	0.109	0.882	0.250	0.933	0.107	0.917	0.126	0.867	0.394		
	EIGLS–AR(1)	0.915	0.174	0.904	0.120	0.900	0.275	0.940	0.121	0.927	0.143	0.871	0.404		
	EIGLS–AR(4)	0.915	0.174	0.904	0.120	0.900	0.275	0.940	0.121	0.927	0.143	0.871	0.404		
	EIGLS–AR(7)	0.965	0.201	0.941	0.158	0.948	0.360	0.963	0.155	0.946	0.171	0.896	0.426		
100	GLS	0.872	0.094	0.870	0.083	0.863	0.245	0.865	0.112	0.885	0.203	0.906	0.584		
	EGLS	0.897	0.096	0.905	0.085	0.891	0.287	0.864	0.112	0.884	0.203	0.906	0.584		
	EIGLS–AR(1)	0.901	0.097	0.904	0.089	0.898	0.313	0.871	0.123	0.893	0.217	0.913	0.589		
	EIGLS–AR(5)	0.901	0.097	0.904	0.089	0.898	0.313	0.871	0.123	0.893	0.217	0.913	0.589		
	EIGLS–AR(10)	0.916	0.108	0.920	0.101	0.915	0.341	0.888	0.139	0.907	0.236	0.918	0.619		
200	GLS	0.938	0.112	0.936	0.087	0.773	0.322	0.892	0.110	0.891	0.185	0.950	0.798		
	EGLS	0.944	0.112	0.943	0.087	0.781	0.351	0.891	0.110	0.891	0.185	0.950	0.797		
	EIGLS–AR(1)	0.942	0.114	0.940	0.089	0.790	0.370	0.892	0.115	0.893	0.191	0.954	0.798		
	EIGLS–AR(7)	0.942	0.114	0.940	0.089	0.790	0.370	0.892	0.115	0.893	0.191	0.954	0.798		
	EIGLS–AR(14)	0.944	0.118	0.949	0.094	0.805	0.404	0.898	0.132	0.899	0.212	0.951	0.816		

**Table 3.** Relative efficiencies for MA(1) error processes

Sample size	GLS type	Relative efficiencies for the following parameter values:											
		$\theta_1 = 0.3^\ddagger$						$\theta_1 = 0.5^\ddagger$					
		$\pi = 0.0$		$\pi = 0.5$		$\pi = 1.0$		$\pi = 0.0$		$\pi = 0.5$		$\pi = 1.0$	
		$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$
50	GLS	1.000	0.826	1.000	0.817	0.999	0.958	0.979	0.595	0.978	0.547	0.973	0.806
	EGLS	1.016	0.908	1.016	0.887	1.015	0.989	0.993	0.681	0.991	0.589	0.987	0.827
	EIGLS-AR(1)	0.998	0.890	0.998	0.860	0.998	0.961	0.983	0.732	0.982	0.634	0.979	0.833
	EIGLS-AR(4)	1.003	0.929	1.003	0.889	1.002	0.979	0.980	0.693	0.980	0.593	0.975	0.818
	EIGLS-AR(7)	1.027	0.997	1.024	0.933	1.021	1.007	1.000	0.737	0.996	0.621	0.990	0.830
100	GLS	0.995	0.844	0.992	0.859	0.996	0.965	0.971	0.593	0.962	0.596	0.976	0.854
	EGLS	1.000	0.878	0.998	0.890	0.996	0.978	0.976	0.628	0.967	0.626	0.975	0.862
	EIGLS-AR(1)	0.996	0.873	0.995	0.886	0.998	0.969	0.980	0.698	0.975	0.692	0.984	0.888
	EIGLS-AR(5)	1.008	0.923	1.004	0.912	1.007	0.972	0.982	0.648	0.973	0.628	0.983	0.855
	EIGLS-AR(10)	1.013	0.967	1.013	0.946	1.011	1.007	0.987	0.681	0.980	0.659	0.986	0.884
200	GLS	0.992	0.828	0.992	0.855	0.993	0.981	0.974	0.585	0.970	0.590	0.974	0.926
	EGLS	0.990	0.852	0.989	0.876	0.993	0.985	0.972	0.603	0.968	0.603	0.973	0.929
	EIGLS-AR(1)	0.992	0.854	0.991	0.872	0.993	0.983	0.979	0.683	0.975	0.673	0.978	0.939
	EIGLS-AR(7)	0.991	0.875	0.991	0.889	0.994	0.981	0.975	0.613	0.972	0.605	0.976	0.924
	EIGLS-AR(14)	0.996	0.921	0.997	0.939	0.990	1.001	0.979	0.648	0.977	0.643	0.973	0.941
		$\theta_1 = 0.9^\S$						$\theta_1 = -0.9^\S\S$					
50	GLS	0.365	0.155	0.366	0.106	0.339	0.187	0.989	0.142	0.988	0.186	0.968	0.502
	EGLS	0.401	0.233	0.408	0.147	0.363	0.199	0.989	0.206	0.988	0.267	0.969	0.575
	EIGLS-AR(1)	0.687	0.570	0.681	0.396	0.641	0.393	0.995	0.547	0.995	0.609	0.991	0.891
	EIGLS-AR(4)	0.518	0.349	0.520	0.214	0.477	0.256	1.007	0.329	1.005	0.421	0.999	0.810
	EIGLS-AR(7)	0.430	0.295	0.442	0.188	0.386	0.217	1.016	0.264	1.019	0.368	1.008	0.782
100	GLS	0.326	0.099	0.310	0.084	0.369	0.191	0.998	0.098	0.993	0.177	0.984	0.662
	EGLS	0.370	0.147	0.352	0.120	0.387	0.213	0.998	0.138	0.993	0.220	0.985	0.700
	EIGLS-AR(1)	0.669	0.514	0.624	0.426	0.697	0.522	1.000	0.514	0.997	0.654	0.997	0.960
	EIGLS-AR(5)	0.435	0.205	0.403	0.160	0.458	0.277	1.009	0.201	1.007	0.344	0.998	0.839
	EIGLS-AR(10)	0.399	0.181	0.378	0.144	0.406	0.231	1.016	0.182	1.017	0.320	1.007	0.834
200	GLS	0.448	0.098	0.393	0.076	0.441	0.237	0.998	0.111	0.996	0.185	0.979	0.888
	EGLS	0.462	0.121	0.407	0.093	0.447	0.247	0.997	0.137	0.994	0.217	0.981	0.906
	EIGLS-AR(1)	0.713	0.509	0.661	0.416	0.665	0.531	1.003	0.523	1.001	0.650	0.995	0.988
	EIGLS-AR(7)	0.487	0.153	0.428	0.119	0.471	0.297	0.998	0.183	0.995	0.280	0.982	0.937
	EIGLS-AR(14)	0.470	0.135	0.413	0.104	0.452	0.260	0.997	0.174	0.994	0.268	0.970	0.944

$\ddagger$   $\Pi$ -weights of the AR representation,  $\pi_1, \pi_2, \dots, \pi_{15}$ :  $-3.0 \times 10^{-1}, -9.0 \times 10^{-2}, -2.7 \times 10^{-2}, -8.1 \times 10^{-3}, -2.4 \times 10^{-3}, -7.3 \times 10^{-4}, -2.2 \times 10^{-4}, -6.6 \times 10^{-5}, -2.0 \times 10^{-5}, -5.9 \times 10^{-6}, -1.8 \times 10^{-6}, -5.3 \times 10^{-7}, -1.6 \times 10^{-7}, -4.9 \times 10^{-8}, -1.4 \times 10^{-8}$ .

$\ddagger$   $\Pi$ -weights of the AR representation,  $\pi_1, \pi_2, \dots, \pi_{15}$ :  $-5.0 \times 10^{-1}, -2.5 \times 10^{-1}, -1.3 \times 10^{-1}, -6.3 \times 10^{-2}, -3.1 \times 10^{-2}, -1.6 \times 10^{-2}, -7.8 \times 10^{-3}, -3.9 \times 10^{-3}, -2.0 \times 10^{-3}, -9.8 \times 10^{-4}, -4.9 \times 10^{-4}, -2.4 \times 10^{-4}, -1.2 \times 10^{-4}, -6.1 \times 10^{-5}, -3.1 \times 10^{-5}$ .

$\S$   $\Pi$ -weights of the AR representation,  $\pi_1, \pi_2, \dots, \pi_{15}$ :  $-0.900, -0.810, -0.729, -0.656, -0.590, -0.531, -0.478, -0.430, -0.387, -0.349, -0.314, -0.282, -0.254, -0.229, -0.206$ .

$\S\S$   $\Pi$ -weights of the AR representation,  $\pi_1, \pi_2, \dots, \pi_{15}$ :  $0.900, -0.810, 0.729, -0.656, 0.590, -0.531, 0.478, -0.430, 0.387, -0.349, 0.314, -0.282, 0.254, -0.229, 0.206$ .

let alone its parameters. To gain additional insights into the efficiency of the EIGLS estimates over OLS we have included in each table a finite AR representation of the ARMA( $p, q$ ) error process.

#### 4.2. The empirical results

In examining the results from Tables 2–6 six general conclusions emerge. First and foremost we see that regardless of the sample size for all model structures and parameterizations the

**Table 4.** Relative efficiencies for AR(2) and MA(2) error processes

Sample size	GLS type	Relative efficiencies for the following parameter values:											
		$(\phi_1, \phi_2)=(1.42, -0.73)$						$(\phi_1, \phi_2)=(1.60, -0.64)$					
		$\pi = 0.0$		$\pi = 0.5$		$\pi = 1.0$		$\pi = 0.0$		$\pi = 0.5$		$\pi = 1.0$	
		$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$
<i>(a) AR(2)</i>													
50	GLS	0.950	0.066	0.949	0.055	0.917	0.215	0.739	0.012	0.737	0.007	0.687	0.013
	EGLS	0.962	0.065	0.959	0.056	0.930	0.234	2.080	0.014	10.55	0.013	0.823	0.061
100	EIGLS-AR(1)	1.123	0.096	1.109	0.112	1.103	0.482	32.60	0.014	5.068	0.013	118.5	0.073
	EIGLS-AR(4)	0.967	0.070	0.966	0.064	0.941	0.265	2.009	0.014	3.211	0.013	2.039	0.073
	EIGLS-AR(7)	0.975	0.080	0.975	0.082	0.954	0.333	1.785	0.018	2.496	0.018	2.375	0.111
	GLS	0.969	0.046	0.962	0.060	0.934	0.393	0.829	0.006	0.829	0.006	0.811	0.023
200	EGLS	0.973	0.046	0.965	0.061	0.938	0.410	0.857	0.007	0.863	0.008	0.850	0.031
	EIGLS-AR(1)	1.058	0.083	1.046	0.162	1.040	1.045	3.351	0.008	13.63	0.011	56.89	0.076
	EIGLS-AR(5)	0.976	0.053	0.970	0.074	0.942	0.448	0.971	0.009	0.937	0.010	1.002	0.055
	EIGLS-AR(10)	0.986	0.055	0.982	0.080	0.952	0.474	0.953	0.010	0.966	0.011	0.912	0.055
200	GLS	0.977	0.051	0.966	0.059	0.921	0.674	0.914	0.007	0.910	0.006	0.672	0.041
	EGLS	0.978	0.051	0.967	0.059	0.920	0.679	0.920	0.007	0.919	0.006	0.684	0.043
	EIGLS-AR(1)	1.057	0.086	1.047	0.146	1.202	1.578	1.053	0.009	1.049	0.010	0.785	0.107
	EIGLS-AR(7)	0.976	0.053	0.964	0.063	0.911	0.691	0.960	0.008	0.993	0.008	0.728	0.060
	EIGLS-AR(14)	0.981	0.055	0.969	0.065	0.915	0.703	0.941	0.009	0.987	0.008	0.727	0.059
<div style="display: flex; justify-content: space-around;"> <span><math>(\theta_1, \theta_2)=(1.42, -0.73)^\dagger</math></span> <span><math>(\theta_1, \theta_2)=(1.80, -0.90)^\ddagger</math></span> </div>													
<i>(b) MA(2)</i>													
50	GLS	0.739	0.059	0.738	0.049	0.716	0.200	0.142	0.006	0.143	0.036	0.139	0.012
	EGLS	0.749	0.091	0.753	0.078	0.721	0.211	0.174	0.056	0.193	0.041	0.162	0.022
100	EIGLS-AR(1)	0.788	0.300	0.785	0.172	0.758	0.263	0.373	0.318	0.367	0.168	0.322	0.109
	EIGLS-AR(4)	0.752	0.108	0.752	0.076	0.730	0.215	0.198	0.084	0.202	0.040	0.176	0.024
	EIGLS-AR(7)	0.758	0.122	0.757	0.086	0.730	0.220	0.174	0.074	0.199	0.045	0.164	0.021
	GLS	0.688	0.047	0.691	0.060	0.755	0.312	0.128	0.004	0.132	0.004	0.164	0.022
200	EGLS	0.691	0.061	0.698	0.069	0.756	0.313	0.148	0.025	0.148	0.019	0.174	0.031
	EIGLS-AR(1)	0.749	0.233	0.743	0.180	0.810	0.411	0.347	0.245	0.310	0.165	0.394	0.211
	EIGLS-AR(5)	0.700	0.075	0.704	0.085	0.772	0.339	0.151	0.027	0.149	0.018	0.177	0.036
	EIGLS-AR(10)	0.717	0.077	0.718	0.082	0.768	0.330	0.157	0.032	0.156	0.022	0.180	0.034
200	GLS	0.788	0.050	0.773	0.059	0.801	0.522	0.216	0.004	0.195	0.003	0.236	0.067
	EGLS	0.789	0.052	0.775	0.062	0.803	0.524	0.228	0.012	0.207	0.001	0.238	0.068
	EIGLS-AR(1)	0.824	0.237	0.802	0.173	0.841	0.593	0.405	0.250	0.355	0.164	0.373	0.212
	EIGLS-AR(7)	0.794	0.062	0.780	0.070	0.812	0.540	0.226	0.014	0.206	0.011	0.246	0.079
	EIGLS-AR(14)	0.798	0.063	0.784	0.072	0.814	0.549	0.231	0.017	0.209	0.013	0.249	0.077

$\dagger$   $\Pi$ -weights of the AR representation,  $\pi_1, \pi_2, \dots, \pi_{15}$ :  $-1.420, -1.286, -0.790, -0.183, 0.317, 0.584, 0.597, 0.422, 0.163, -0.076, -0.227, -0.267, -0.214, -0.108, 0.002$ .

$\ddagger$   $\Pi$ -weights of the AR representation,  $\pi_1, \pi_2, \dots, \pi_{15}$ :  $-1.800, -2.340, -2.592, -2.560, -2.274, -1.790, -1.176, -0.505, 0.149, 0.723, 1.167, 1.450, 1.560, 1.503, 1.301$ .

efficiency in estimating the slope of the regression model ( $\beta_1$ ) is higher for the GLS procedures including those based on incorrectly identified error structures (IGLS) than for OLS when the exogenous variable is stationary. The relative efficiency gain of GLS and IGLS over OLS in estimating  $\beta_1$  seems to depend not only on the ARMA( $p, q$ ) error structure but also on the magnitude of the parameters themselves. For pure AR processes (Tables 2 and 4, part (a)), the improvement in relative efficiency ranges from nearly 0 to about 300%. For example, when  $\pi = 0$  and  $n = 100$  for the AR(1) model with  $\phi = 0.5$ ,  $\hat{\zeta}_{\beta_1}$  changes from 0.563 to 0.628 whereas the variation for the AR(2) parameterization with  $\phi_1 = 1.60$  and  $\phi_2 = -0.64$  ranges from 0.006 to 0.010. From Table 2, it can also be seen that as the magnitude of the AR(1) parameter increases so does the relative efficiency of the GLS and IGLS estimates over

**Table 5.** Relative efficiencies for ARMA(1, 1), ARMA(1, 2) and ARMA(2, 1) error processes

Sample size	GLS type	Relative efficiencies for the following parameter values:											
		$(\phi_1, \theta_1)=(0.8, 0.5)^\dagger$						$(\phi_1, \theta_1)=(-0.8, 0.7)^\ddagger$					
		$\pi = 0.0$		$\pi = 0.5$		$\pi = 1.0$		$\pi = 0.0$		$\pi = 0.5$		$\pi = 1.0$	
		$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$	$\hat{\zeta}_{\beta_0}$	$\hat{\zeta}_{\beta_1}$
<i>(a) ARMA(1, 1)</i>													
50	GLS	1.013	0.758	1.013	0.595	1.008	0.656	0.861	0.031	0.859	0.023	0.820	0.052
	EGLS	1.011	0.794	1.010	0.693	1.008	0.844	0.894	0.042	0.896	0.041	0.870	0.139
100	EIGLS-AR(1)	0.994	0.775	0.995	0.685	0.993	0.866	0.910	0.069	0.910	0.062	0.896	0.195
	EIGLS-AR(4)	1.007	0.834	1.005	0.714	1.007	0.855	0.921	0.056	0.935	0.051	0.921	0.190
	EIGLS-AR(7)	1.015	0.935	1.012	0.826	1.013	0.975	0.946	0.059	0.951	0.063	0.948	0.237
	GLS	1.007	0.664	1.005	0.588	1.000	0.798	0.923	0.019	0.921	0.023	0.903	0.103
200	EGLS	1.010	0.689	1.012	0.632	1.002	0.863	0.934	0.027	0.936	0.033	0.913	0.140
	EIGLS-AR(1)	0.998	0.707	0.997	0.670	0.995	0.903	0.931	0.056	0.930	0.073	0.904	0.295
	EIGLS-AR(5)	1.009	0.726	1.010	0.662	1.000	0.875	0.940	0.041	0.940	0.053	0.925	0.224
	EIGLS-AR(10)	1.011	0.775	1.012	0.703	1.005	0.915	0.952	0.039	0.954	0.051	0.937	0.210
200	GLS	1.013	0.706	1.011	0.593	0.982	0.898	0.965	0.025	0.962	0.025	0.767	0.228
	EGLS	1.013	0.719	1.010	0.612	0.972	0.915	0.967	0.028	0.964	0.030	0.768	0.242
	EIGLS-AR(1)	1.001	0.753	1.000	0.672	0.990	0.956	0.971	0.060	0.967	0.066	0.812	0.467
	EIGLS-AR(7)	1.006	0.743	1.004	0.632	0.971	0.926	0.968	0.045	0.965	0.048	0.791	0.356
EIGLS-AR(14)	1.013	0.772	1.012	0.658	0.977	0.958	0.969	0.040	0.966	0.043	0.800	0.343	
<div style="display: flex; justify-content: space-around;"> <span><math>(\phi_1, \theta_1, \theta_2)=(-0.8, 1.4, -0.6)^\S</math></span> <span><math>(\phi_1, \phi_2, \theta_1)=(1.4, -0.6, -0.8)^\S\S</math></span> </div>													
<i>(b) ARMA(1, 2) and ARMA(2, 1)</i>													
50	GLS	0.105	0.002	0.091	0.002	0.060	0.006	0.909	0.003	0.908	0.002	0.874	0.009
	EGLS	0.132	0.018	0.129	0.018	0.071	0.015	0.931	0.003	0.931	0.002	0.899	0.009
100	EIGLS-AR(1)	0.172	0.038	0.147	0.018	0.094	0.015	1.439	0.025	1.751	0.036	1.412	0.165
	EIGLS-AR(4)	0.135	0.018	0.132	0.014	0.070	0.009	1.121	0.013	1.142	0.018	1.074	0.083
	EIGLS-AR(7)	0.150	0.024	0.157	0.018	0.095	0.015	1.108	0.017	1.282	0.026	1.085	0.132
	GLS	0.061	0.002	0.073	0.003	0.088	0.014	0.950	0.002	0.946	0.002	0.909	0.017
200	EGLS	0.072	0.010	0.084	0.010	0.092	0.018	0.956	0.002	0.955	0.002	0.922	0.017
	EIGLS-AR(1)	0.101	0.029	0.109	0.020	0.143	0.043	1.035	0.027	1.026	0.057	0.979	0.388
	EIGLS-AR(5)	0.073	0.011	0.084	0.009	0.093	0.017	0.962	0.009	0.959	0.013	0.919	0.053
	EIGLS-AR(10)	0.076	0.014	0.088	0.012	0.101	0.020	0.969	0.009	0.965	0.013	0.929	0.058
200	GLS	0.081	0.002	0.077	0.003	0.145	0.035	0.975	0.002	0.967	0.003	0.707	0.057
	EGLS	0.084	0.005	0.082	0.006	0.145	0.035	0.974	0.002	0.967	0.003	0.708	0.057
	EIGLS-AR(1)	0.109	0.031	0.101	0.021	0.176	0.060	1.031	0.028	1.024	0.049	0.973	0.800
	EIGLS-AR(7)	0.086	0.007	0.084	0.007	0.146	0.036	0.978	0.059	0.970	0.007	0.719	0.078
EIGLS-AR(14)	0.087	0.008	0.084	0.008	0.158	0.042	0.980	0.061	0.973	0.008	0.722	0.083	

$^\dagger$   $\Pi$ -weights of the AR representation,  $\pi_1, \pi_2, \dots, \pi_{15}$ :  $3.0 \times 10^{-1}, 1.5 \times 10^{-1}, 7.5 \times 10^{-2}, 3.8 \times 10^{-2}, 1.9 \times 10^{-2}, 9.4 \times 10^{-3}, 4.7 \times 10^{-3}, 2.3 \times 10^{-3}, 1.2 \times 10^{-3}, 5.9 \times 10^{-4}, 2.9 \times 10^{-4}, 1.5 \times 10^{-4}, 7.3 \times 10^{-5}, 3.7 \times 10^{-5}, 1.8 \times 10^{-5}$ .

$^\ddagger$   $\Pi$ -weights of the AR representation,  $\pi_1, \pi_2, \dots, \pi_{15}$ :  $-1.500, -1.050, -0.735, -0.515, -0.360, -0.252, -0.176, -0.124, -0.086, -0.061, -0.042, -0.030, -0.021, -0.015, -0.010$ .

$^\S$   $\Pi$ -weights of the AR representation,  $\pi_1, \pi_2, \dots, \pi_{15}$ :  $-2.200, -2.480, -2.152, -1.525, -0.844, -0.266, 0.134, 0.347, 0.405, 0.359, 0.260, 0.148, 0.052, -0.017, -0.054$ .

$^\S\S$   $\Pi$ -weights of the AR representation,  $\pi_1, \pi_2, \dots, \pi_{15}$ :  $2.200, 0.350, -0.280, 0.224, -0.179, 0.144, -0.115, 0.092, -0.073, 0.059, -0.047, 0.038, -0.030, 0.024, -0.019$ .

OLS. Similar observations on the estimates of  $\beta_1$  can be made for other error structures and parameterizations.

When the exogenous variable is non-stationary the only times when we observed the relative efficiency of GLS to be lower than OLS in estimating the slope of the regression were when EIGLS was used to correct AR(1) and MA(1) disturbances with  $\phi = 0.3$  (Table 2) and  $\theta = 0.3$  (Table 3) respectively. However, even in these cases the maximum value observed for  $\hat{\zeta}_{\beta_1}$  was only 1.069 (Table 2, EIGLS-AR(7) and  $n = 50$ ).

**Table 6.** MSEs of GLS estimators when the serially correlated disturbance follows AR(2) processes

$\pi$	Estimator	$MSE(\hat{\beta}_0)$				$MSE(\hat{\beta}_1)$			
		$n = 50$	$n = 100$	$n = 200$	$n = 400$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
<i>(a) <math>(\phi_1, \phi_2) = (1.42, -0.73)</math></i>									
0	OLS	0.217	0.113	0.054	0.027	0.063	0.056	0.027	0.015
	GLS	0.206	0.110	0.053	0.026	0.005	0.003	0.001	0.001
	EGLS	0.209	0.110	0.053	0.026	0.005	0.003	0.001	0.001
	EIGLS-AR(1)	0.243	0.119	0.057	0.027	0.007	0.005	0.002	0.001
	EIGLS-AR( $[n^{1/2}/2]$ )	0.210	0.110	0.052	0.026	0.005	0.003	0.001	0.001
	EIGLS-AR( $[n^{1/2}]$ )	0.211	0.111	0.053	0.026	0.006	0.003	0.001	0.001
0.5	OLS	0.217	0.114	0.055	0.030	0.159	0.081	0.042	0.025
	GLS	0.206	0.110	0.053	0.029	0.009	0.005	0.002	0.001
	EGLS	0.208	0.110	0.053	0.029	0.009	0.005	0.002	0.001
	EIGLS-AR(1)	0.241	0.119	0.057	0.030	0.018	0.013	0.006	0.003
	EIGLS-AR( $[n^{1/2}/2]$ )	0.210	0.110	0.053	0.029	0.010	0.006	0.003	0.001
	EIGLS-AR( $[n^{1/2}]$ )	0.212	0.112	0.053	0.028	0.013	0.006	0.003	0.001
1	OLS	0.226	0.120	0.077	0.074	0.055	0.013	0.002	0.001
	GLS	0.207	0.112	0.071	0.058	0.012	0.005	0.001	0.001
	EGLS	0.210	0.113	0.070	0.059	0.012	0.005	0.001	0.001
	EIGLS-AR(1)	0.249	0.125	0.092	0.091	0.027	0.014	0.003	0.002
	EIGLS-AR( $[n^{1/2}/2]$ )	0.213	0.113	0.070	0.059	0.015	0.006	0.001	0.001
	EIGLS-AR( $[n^{1/2}]$ )	0.216	0.114	0.070	0.057	0.018	0.006	0.001	0.001
<i>(b) <math>(\phi_1, \phi_2) = (1.60, -0.64)</math></i>									
0	OLS	9.672	5.936	2.968	1.693	0.360	0.392	0.176	0.065
	GLS	7.144	4.918	2.712	1.590	0.003	0.002	0.001	0.001
	EGLS	20.12	5.085	2.729	1.588	0.005	0.003	0.001	0.001
	EIGLS-AR(1)	315.3	19.89	3.127	1.610	0.005	0.003	0.002	0.001
	EIGLS-AR( $[n^{1/2}/2]$ )	19.43	5.760	2.848	1.582	0.005	0.003	0.001	0.001
	EIGLS-AR( $[n^{1/2}]$ )	17.26	5.655	2.792	1.638	0.006	0.004	0.002	0.001
0.5	OLS	9.700	5.931	2.980	1.700	1.145	0.729	0.381	0.143
	GLS	7.114	4.919	2.713	1.590	0.008	0.004	0.002	0.001
	EGLS	7.144	5.120	2.739	1.589	0.015	0.005	0.002	0.001
	EIGLS-AR(1)	102.3	40.42	3.126	1.618	0.014	0.008	0.004	0.001
	EIGLS-AR( $[n^{1/2}/2]$ )	49.16	5.558	2.958	1.585	0.015	0.007	0.003	0.001
	EIGLS-AR( $[n^{1/2}]$ )	31.15	5.729	2.941	1.638	0.020	0.008	0.003	0.001
1	OLS	10.41	6.069	4.161	3.705	1.019	0.308	0.087	0.042
	GLS	7.147	4.919	2.796	1.718	0.498	0.007	0.004	0.001
	EGLS	8.571	5.161	2.847	1.731	0.516	0.009	0.004	0.002
	EIGLS-AR(1)	1234	345.3	3.267	2.013	0.516	0.023	0.009	0.003
	EIGLS-AR( $[n^{1/2}/2]$ )	21.23	6.082	3.030	1.749	0.524	0.017	0.005	0.002
	EIGLS-AR( $[n^{1/2}]$ )	24.72	5.536	3.023	1.729	0.534	0.017	0.005	0.002

Second, the relative efficiency of GLS and IGLS over OLS in estimating the intercept term  $\beta_0$  is more dependent on the error structure that is used to obtain  $\Xi$  in equation (3). With one exception, i.e. the AR(2) model with parameterization (1.60, -0.64) (Table 4, part (a)), the relative efficiency of GLS and EGLS in estimating  $\beta_0$  is either comparable with or, more commonly, much higher than OLS, up to more than 10 times, as demonstrated by the ARMA(1, 2) process with the parameterization  $(\phi_1, \theta_1, \theta_2) = (-0.8, 1.4, -0.6)$  in Table 5. For the ‘incorrectly’ identified error structures, i.e. EIGLS based on AR( $\tilde{\rho}$ ) models, the relative superiority of OLS over GLS is only observed in a few cases when the sample size is relatively small (i.e.  $n = 50$ ) and when the AR order of the error structure used in the GLS estimation is lower than the simulated AR structure (see the AR(2) model with parameterization (1.60, -0.64) in Table 4, part (a)) or lower than an ‘appropriate’ finite AR representation of the simulated structure (see the ARMA(2, 1) parameterization (1.40, -0.60, -0.80) in Table 5).

To understand why the relative efficiency of EIGLS in estimating  $\beta_0$  can be lower than that of OLS at times, consider the situation when the disturbances of the regression model with an AR(1) exogenous variable follow an AR(2) process, i.e.

$$\begin{aligned} y_t &= \beta_0 + \beta_1 x_t + \epsilon_t, \\ \epsilon_t &= \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + a_t \end{aligned} \tag{14}$$

where the  $a_t$ s are IID(0, 1).

If we use an AR(1) instead of the AR(2) GLS correction, then, we are essentially performing an OLS estimation of the regression model

$$y_t^* = \beta_0^* + \beta_1 x_t^* + \epsilon_t^* \tag{15}$$

where  $y_t^* = y_t - \rho_1 y_{t-1}$ ,  $x_t^* = x_t - \rho_1 x_{t-1}$  and  $\epsilon_t^* = \epsilon_t - \rho_1 \epsilon_{t-1}$ , and  $\rho_1$  is the first-order autocorrelation of  $\epsilon_t$  in model (14), i.e.  $\rho_1 = \phi_1 / (1 - \phi_2)$ . The constant of the new regression,  $\beta_0^*$ , is  $\beta_0(1 - \rho_1)$ . Thus, an estimate of  $\beta_0$  can be recovered from  $\hat{\beta}_0^*$  by using the equation  $\hat{\beta}_0 = \hat{\beta}_0^* / (1 - \rho_1)$ . The estimation error of  $\beta_0$ , i.e.  $\text{var}(\hat{\beta}_0)$ , is approximately  $\text{var}(\hat{\beta}_0^*) / (1 - \rho_1)^2$ . Consequently, when  $\rho_1$  is close to 1, it is possible for  $\text{var}(\hat{\beta}_0)$  to ‘blow up’, leading at times to much higher values than  $\text{var}(\hat{\beta}_0^*)$ . For example, consider the case where  $(\phi_1, \phi_2) = (1.6, -0.64)$ . Since  $\rho_1 \approx 0.98$ ,

$$\text{var}(\hat{\beta}_0) \approx \text{var}(\hat{\beta}_0^*) / 0.02^2 = 2500 \text{var}(\hat{\beta}_0^*).$$

If the estimation error  $\text{var}(\hat{\beta}_0^*)$  is of the same magnitude as that of the OLS estimator of  $\beta_0$  based on model (14), then the estimation error of the OLS estimator of  $\beta_0$  based on model (15) will inevitably be higher as shown in a few cases in Table 4, part (a). (If  $\rho_1 \approx 1$ , it can be shown that  $x_t^*$  is an AR(2) process unless the first-order autocorrelation of  $x_t$  is  $\rho_1$ , and  $\epsilon_t^*$  is close to an AR(1) process with AR coefficient  $-\phi_2$ .)

The third general conclusion from the simulation study is that the differences in the relative efficiency of EIGLS *vis-à-vis* GLS and EGLS in estimating  $\beta_1$  with a few expected exceptions is not very large. In fact, when the error structure is assumed to follow an AR( $\bar{p}$ ) model with  $\bar{p} = \lceil n^{1/2} / 2 \rceil$  the relative efficiency of EIGLS is comparable with that of EGLS regardless of the sample size particularly when the exogenous variable is stationary. Some of the most noticeable exceptions to this general conclusion occur when the order of the EIGLS correction is much smaller than a ‘reasonable’ corresponding finite AR representation of the simulated autocorrelation structure. For example, consider the MA(1) process with  $\theta = 0.9$  in Table 3. When  $n = 50$ , the  $\hat{\zeta}_{\beta_1}$ s for EGLS are 0.233, 0.147 and 0.199 for  $\pi$  equal to 0, 0.5 and 1 respectively, whereas the corresponding  $\hat{\zeta}_{\beta_1}$ s for EIGLS based on the AR(1) process are 0.570, 0.396 and 0.393. Note, however, that as the order of the error AR structure increases, say to  $\lceil n^{1/2} / 2 \rceil$ , the differences in efficiencies between EIGLS and EGLS (GLS) decrease considerably. For the MA(2) parameterizations (Table 4, part (b)) even when  $\pi = 1$  the differences in GLS efficiencies between correctly and incorrectly identified error structures are very small.

When the order of the AR error structure is set to  $\lceil n^{1/2} / 2 \rceil$ , the only times that we observed the relative efficiency of OLS to be superior to that of EIGLS in estimating the constant  $\beta_0$  was when  $n = 50$  for the AR(2) and ARMA(2, 1) model structures mentioned earlier.

These observations differ from those in many previous studies, particularly from those implied by Engle (1974), and others subsequently such as Judge *et al.* (1985) and Choudhury *et al.* (1999). Engle (1974) demonstrated that, asymptotically, under certain conditions OLS ‘should generally [provide] superior estimate[s]’ (page 145) compared with GLS when an

incorrect truncation of the order of the autocorrelation of the disturbances is used. He illustrated this by using AR(2) disturbances that were approximated by AR(1) processes. His results depended on the asymptotic distribution of eigenvalues of Toeplitz form matrices (Szegő, 1917). They do not appear to hold for finite samples even when  $n = 200$  as highlighted by the results found in Table 4, part (a), and in particular by the same AR(2) parameterization (1.6,  $-0.64$ ) used by Engle (1974) to exemplify his proofs. Admonitions such as ‘OLS may often be better than assuming another incorrect truncation of the actual process’ (Judge *et al.* (1985), page 281) and ‘sometimes it is better to ignore the problem altogether and use OLS rather than to incorrectly assume the process is AR(1)’ (Choudhury *et al.* (1999), page 347) should be viewed with caution since they are based on asymptotic not finite sample theoretical results; truncations should depend on the sample size (Berk, 1974).

Fourth, our study corroborates the findings of Zinde-Walsh and Galbraith (1991), among others, that the relative efficiency of GLS over OLS decreases as the magnitude of the AR parameter  $\pi$  associated with the exogenous variable of the regression model increases. Unlike Zinde-Walsh and Galbraith’s results, however, in many cases (often for  $n = 50$ ) we observed a small improvement in efficiency when  $\pi$  increased from 0 to 0.5. A decline in efficiency was usually observed when  $\pi$  changed from 0.5 to 1. This pattern can be easily seen by checking some of the results associated with say the ARMA(1, 1) error structures of Table 5.

Fifth, for stationary exogenous variables the relative efficiency of GLS and IGLS over OLS does not appear to be affected significantly by the sample size. For non-stationary exogenous variables the efficiency generally deteriorates as the sample size increases, as illustrated in Table 5, part (b), which contains the relative efficiencies for ARMA(2, 1) error processes. As can be seen for the parameterization (1.40,  $-0.60$ ,  $-0.80$ ), the  $\hat{\zeta}_{\beta_i}$ s for EGLS for  $n = 50, 100, 200$  are 0.009, 0.017 and 0.057 respectively. For non-stationary exogenous variables, gains in efficiency of GLS and IGLS over OLS are less than those for the corresponding stationary cases. These results are in general agreement with Krämer’s (1986) asymptotic results showing that, when  $\epsilon$  is an AR( $p$ ) process and  $x_t$  follows an integrated process, OLS and GLS are asymptotically equivalent. Krämer’s results depend crucially on the fact that sample autocorrelations of the independent variable tend to 1 in probability as  $n \rightarrow \infty$ , a property that makes OLS asymptotically as efficient as GLS (EGLS), and, consequently, as efficient as any asymptotically efficient IGLS (EIGLS).

Although OLS, GLS and IGLS yield consistent estimators (Aitken (1935) and theorem 1 of Section 2), the rate of convergence of these estimators can be quite different. In Table 6 we report MSEs of OLS, GLS and IGLS estimators for sample sizes 50, 100, 200 and 400 when the serially correlated disturbance follows two AR(2) processes. As can be seen when the roots of the disturbance term are far from the unit circle, i.e.  $(\phi_1, \phi_2) = (1.42, -0.73)$ , the MSEs for both  $\beta_0$  and  $\beta_1$  obtained from OLS, GLS and IGLS appear to decline at relatively similar rates. However, when the roots of the error structure are close to the unit circle, i.e.  $(\phi_1, \phi_2) = (1.60, -0.64)$ , the MSEs for  $\beta_1$  decline faster for GLS and IGLS than for OLS. Moreover, there seems to be very little difference in the declining rates between the GLS and IGLS procedures. The MSEs associated with  $\beta_0$  based on OLS, GLS and IGLS decline at similar rates.

## 5. Concluding remarks

We have examined the finite sample performance of GLS and IGLS methods *vis-à-vis* OLS for the estimation of regression models with autocorrelated disturbances. We have established theoretical efficiency bounds for GLS procedures based on ‘incorrectly’ identified error

structures relative to OLS and GLS methods based on ‘correctly’ identified and known or estimated error structures. From a large simulation study we found that GLS estimation based on AR representations of ARMA( $p, q$ ) disturbances yields more efficient estimates than OLS does particularly when the order of the autoregression is set near  $[n^{1/2}/2]$  and the exogenous variables are stationary. Moreover, we observed that in most cases the differences in estimation efficiency between EIGLS and EGLS are small. This suggests that there may not be much to be gained in trying to identify the ‘correct’ order of OLS residuals when performing GLS estimation especially in the light of the tenuous identification performance of commonly used procedures based on finite, especially small, samples. We are currently investigating the effect that EIGLS corrections may have on forecasting performance.

## Acknowledgements

The authors wish to thank the Joint Editor and the reviewers for their helpful and insightful comments and suggestions.

## References

- Aitken, A. C. (1935) On least squares and linear combinations of observations. *Proc. R. Soc. Lond.*, **55**, 42–48.
- Amemiya, T. (1973) Generalized least squares with an estimated autocovariance matrix. *Econometrica*, **41**, 723–732.
- Berk, K. N. (1974) Consistent autoregressive spectral estimates. *Ann. Statist.*, **2**, 489–502.
- Bloomfield, P. and Watson, G. (1975) The inefficiency of least squares. *Biometrika*, **62**, 121–128.
- Box, G. and Jenkins, G. (1976) *Time Series Analysis: Forecasting and Control*, 2nd edn. San Francisco: Holden-Day.
- Breusch, T. (1980) Useful invariance results for generalized regression models. *J. Econometr.*, **13**, 327–340.
- Choudhury, A., Hubata, R. and Louis, R. (1999) Understanding time-series regression estimators. *Am. Statistn*, **53**, 342–348.
- Cochrane, D. and Orcutt, G. (1949) Application of least squares regression to relationships containing auto-correlated error terms. *J. Am. Statist. Ass.*, **44**, 32–61.
- Durbin, J. and Watson, G. (1950) Testing for serial correction in least squares regression. *Biometrika*, **37**, 409–428.
- Engle, R. (1974) Specification of the disturbance for efficient estimation. *Econometrica*, **42**, 135–146.
- Glasbey, C. A. (1982) A generalization of partial autocorrelation useful in identifying ARMA models. *Technometrics*, **24**, 223–228.
- Hadley, G. (1961) *Linear Algebra*. Reading: Addison-Wesley.
- Hamilton, J. (1994) *Time Series Analysis*. Princeton: Princeton University Press.
- Hannan, E. (1970) *Multiple Time Series*. New York: Wiley.
- Hansen, L. (1982) Large sample properties of generalized method of moments estimators. *Econometrica*, **50**, 1029–1054.
- Hildreth, C. (1986) The Cowles Commission in Chicago, 1939–1955. *Lect. Notes Econ. Math. Syst.*, **271**.
- International Mathematical and Statistical Libraries (1999) *Visual Numerics*. Houston: International Mathematical and Statistical Libraries.
- Judge, G., Griffiths, W., Hill, R., Lütkepohl, H. and Lee, T. C. (1985) *The Theory and Practice of Econometrics*, 2nd edn. New York: Wiley.
- Kadiyala, K. (1970) Testing for the independence of regression disturbances. *Econometrica*, **38**, 97–117.
- Kakwani, N. (1967) The unbiasedness of Zellner’s seemingly unrelated regression equation estimators. *J. Am. Statist. Ass.*, **62**, 141–142.
- Kallinen, A., Koreisha, S. and Pukkila, T. (1990) The identification of ARMA models. *Biometrika*, **77**, 537–549.
- King, M. (1983) Testing for autoregressive against moving average errors in the linear regression model. *J. Econometr.*, **21**, 35–51.
- Knott, M. (1975) On the minimum efficiency of least squares. *Biometrika*, **62**, 129–132.
- Koreisha, S. and Fang, Y. (1999) The impact of measurement errors on ARMA prediction. *J. Forecast.*, **18**, 95–110.
- Koreisha, S. G. and Pukkila, T. M. (1987) Identification of nonzero elements in the polynomial matrices of mixed VARMA processes. *J. R. Statist. Soc. B*, **49**, 112–126.
- Krämer, W. (1986) Least squares regression when the independent variable follows an ARIMA process. *J. Am. Statist. Ass.*, **81**, 150–154.
- Krämer, W. and Baltagi, B. (1996) A general condition for an optimal limiting efficiency of OLS in the general linear regression model. *Econ. Lett.*, **50**, 13–17.
- Marcus, M. and Ming, H. (1964) *A Survey of Matrix Theory and Matrix Inequalities*. Boston: Allyn and Bacon.
- Mátyás, L. (1999) *Generalized Method of Moments Estimation*. Cambridge: Cambridge University Press.



- Rao, P. and Griliches, Z. (1969) Small-sample properties of several two-stage regression methods in the context of auto-correlated errors. *J. Am. Statist. Ass.*, **64**, 253–272.
- Schmidt, P. (1976) *Econometrics*. New York: Dekker.
- Spitzer, J. (1979) Small-sample properties of nonlinear least squares and maximum likelihood estimators in the context of autocorrelated errors. *J. Am. Statist. Ass.*, **74**, 41–47.
- Strang, G. (1988) *Linear Algebra and Its Applications*. New York: Academic Press.
- Szegő, G. (1917) A Toeplitz fële formákról. *Math. Természett. Értes.*, **35**, 185–222.
- Theil, H. (1971) *Principles of Econometrics*. New York: Wiley.
- Thursby, J. (1987) OLS or GLS in the presence of specification error? *J. Econometr.*, **35**, 359–374.
- Walker, A. (1967) Some tests of separate families of hypotheses in time series analysis. *Biometrika*, **54**, 39–68.
- Watson, G. (1967) Linear least squares regression. *Ann. Math. Statist.*, **38**, 1679–1699.
- Wise, J. (1955) Autocorrelation function and the spectral density function. *Biometrika*, **42**, 151–159.
- Zinde-Walsh, V. (1990) The consequences of misspecification in time series processes. *Econ. Lett.*, **32**, 237–241.
- Zinde-Walsh, V. and Galbraith, J. (1991) Estimation of a linear regression model with stationary ARMA(p,q) errors. *J. Econometr.*, **47**, 333–357.
- Zyskind, G. (1967) On canonical forms, non-negative covariance matrices and best and simple linear least squares estimators in linear models. *Ann. Math. Statist.*, **38**, 1092–1109.