AMSTER

AE no. 11/88

INSTITUTE OF ACTUARIAL SCIENCE & ECONOMETRICS

REPORT AE 11/88

BIAS REDUCTION IN A DYNAMIC REGRESSION MODEL:
A Comparison of Jackknifed and Bias Corrected Least Squares Estimators

Jan F. Kiviet & Garry D.A. Phillips

University of Amsterdam,
Faculty of actuarial science & econometrics

Title: BIAS REDUCTION IN A DYNAMIC REGRESSION MODEL:

A Comparison of Jackknifed and Bias Corrected Least Squares Estimators

Author: Jan F. Kiviet and G.D.A. Phillips

Address: Jan F. Kiviet
Actuarial Science & Econometrics
University of Amsterdam
Jodenbreestraat 23
1011 NH Amsterdam NL

Garry D.A. Phillips
Econometrics and Social Statistics
University of Manchester
Manchester M13 9PL UK

Date: October 1988

Series and number: Report AE 11/88

Pages: 41

Price: No Charge

JEL Subject Classification: 211

Keywords: small sample bias, small sigma asymptotics, lagged dependent variables, least-squares, jackknife

Abstract:

Employing small−sigma asymptotics we approximate the small−sample bias of the ordinary least−squares (OLS) estimator of the full coefficient vector in a linear regression model which includes a one period lagged dependent variable and an arbitrary number of fixed regressors. This bias term is used to construct a corrected ordinary least−squares (COLS) estimator which is unbiased to $O(\sigma^2)$. We also consider another technique for bias reduction, viz. jackknifing, and we present a simple expression for the JOLS(m) estimator: the m−delete jackknifed OLS estimator. Then we compare the accuracy of the $O(\sigma^2)$ approximation to the bias and the efficiency of OLS, COLS and JOLS(m) in a Monte Carlo study of artificial but realistic models. It is found that the bias is extremely sensitive to the value of σ and that COLS can reduce it considerably without undue loss of efficiency if the standard deviation of the OLS lagged dependent variable coefficient estimate has a moderate value.

Bias Reduction in a Dynamic Regression Model:

A Comparison of Jackknifed and Bias Corrected Least Squares Estimators

Jan F. Kiviet

Actuarial Science & Econometrics
University of Amsterdam
23 Jodenbreestraat
1011 NH Amsterdam NL

Garry D. A. Phillips

Econometrics and Social Statistics
University of Manchester
Manchester M13 9PL UK

October 1988

Financial support from the
Netherlands Organization for the Advancement of Pure Research (NWO)
is gratefully acknowledged.

1. Introduction

If Ordinary Least Squares (OLS) is applied to estimate linear regression equations containing non−exogenous regressors, the resulting coefficient estimators will, in general, be biased. If the regressors are predetermined then the bias will usually decrease as the sample size is increased and, asymptotically, OLS will yield consistent estimators. This is the case in the model where lagged values of the dependent variable appear among the regressors. However, in practice, sample sizes are finite and may be quite small. Perhaps the fact that OLS has desirable asymptotic properties (e.g. it is the maximum likelihood estimator under normality and regularity assumptions) has served to mask the small sample problems. Practitioners appear to accept the reassurance stated in all the editions of Malinvaud (1980, p.541): "..... *in practice no serious error is committed if we apply the usual methods conceived for regression models to the treatment of autoregressive models*". However, this fallacy is based on a very limited simulation study. In fact the small sample bias problem may be very serious in autoregressive models, see for instance Kiviet (1985), Hoque and Peters (1986) and Inder (1987). Given that models which include lagged dependent variables as regressors are in frequent use, it is important that techniques for reducing the OLS bias should be examined.

In the statistical literature there has been a good deal of attention given to the development of bias reduction techniques. The most direct approach proceeds by first evaluating the bias or finding a suitable approximation to it. The bias or its approximation will usually involve unknown parameters and an estimate of the bias is obtained on replacing these unknown parameters with estimates. Once an estimate for the bias has been obtained, it is used to correct the original estimator. Typically the bias corrected estimator has a lower order bias. An alternative procedure is to use a technique such as jackknifing which does not require that the bias be estimated. The jackknife technique has widespread applicability in reducing the small sample bias of biased but consistent estimators and it is particularly helpful when the small sample distribution of the estimator is intractable. In fact, the first application of the jackknife approach for bias reduction was in the context of a simple autoregressive model, see Quenouille (1956), and the results of Orcutt and Winokur (1969) indicated that Quenouille's approach was highly successful in reducing bias.

In the econometrics literature there is considerable interest in the small sample properties of OLS estimators in models where the regressors are not all exogenous. When their distributions or their moments cannot be derived exactly, asymptotic approximations may be employed. Many articles have considered the properties of OLS estimators in autoregressive models. Originally simple AR models were investigated. Later results were obtained for the case where the process has constant but non−zero mean, and also for the model which includes −apart from one or more lagged dependent variables− some specific exogenous regressors (like a time trend) or, more generally, a number of arbitrary regressors. Various methods have been employed to derive approximations to or exact expressions for characteristics of the distribution of coefficient estimates in the multiplicity of particular types of models. We mention some of the main and most recent contributions. Sawa (1978) −see also Nankervis and Savin (1988)− obtained exact expressions for the first and second moment of the OLS estimator in the first−order model with −apart from a constant− no exogenous regressors. Tse (1982) and Maekawa (1983) consider the first−order model with exogenous regressors and employ Edgeworth expansions which yield extremely complex expressions for the coefficient bias. Pantula and Fuller (1985) consider a higher−order autoregressive process with a mean which is a function of time. Apart from approximating the coefficient bias they also use their bias expression to correct the OLS estimator for bias. Hoque and Peters (1986) present formulas for the exact bias and mean squared error of the coefficient estimates in the first−order ARMAX model with arbitrary exogenous variables and normally distributed disturbances. These formulas can be evaluated (when all parameter values are known) by numerical integration and the authors do so for a few simple models. Their results illustrate the nature of the bias for varying sample size, different values of the coefficients and of the disturbance variance, and also for varying trends in one exogenous regressor.

They also illustrate that the bias can be very serious indeed; in their calculations the bias is usually quite substantial in comparison with the corresponding actual coefficient value, and at the same time this bias is often about half as large as the standard deviation of the coefficient estimate (see their Table 1).

For practitioners it would be of great interest if feasible techniques were available that could indicate for a particular empirical model (where the actual parameter values are unknown and where the disturbances may be non−normal) whether or not the bias in the OLS estimates is serious, and which also would enable the bias to be reduced to an acceptable degree in cases where it is substantial. Approximations to the bias of OLS estimators in general dynamic regression models with predetermined regressor variables can be obtained rather straightforwardly by using a small sigma (σ) asymptotic expansion approach, see Kadane (1971). Here, see also Kiviet and Phillips (1986), we obtain an approximation to the bias of the complete OLS coefficient vector estimator in a regression model containing a first order lagged dependent variable and an arbitrary number of exogenous regressors. The approximate bias can be estimated to produce a bias corrected estimator of the type mentioned above which is unbiased to a certain order of approximation. In a Monte Carlo study we investigate the bias and mean square error of OLS and compare these with the performance of the bias corrected OLS estimator and with various implementations of the jackknifed OLS estimator.

This paper is organized as follows. In Section 2 we derive the small sigma asymptotic approximation to the OLS bias in the regression model with fixed regressors and a one period lagged dependent variable. In Section 3 we consider the sign of the bias of the lagged dependent variable coefficient estimator. In Section 4 we discuss the bias correction technique (COLS) implied by the small disturbance asymptotic approximation. In Section 5 we review the jackknife technique and we obtain a simple expression for the delete m jackknifed OLS estimator, labelled JOLS(m). We then present a small sigma approximation to the bias of the JOLS(1) estimator and it is found that the resulting bias expression is somewhat more complex than that of OLS. This prevents a direct comparison of the biases of the two techniques. However, in Section 6 a comparison is achieved through some direct numerical calculations of the respective approximate bias expressions in some artificial models with assigned parameter values. Also a Monte Carlo study is performed to investigate the accuracy of the bias approximations, to assess the actual ability of the bias reduction techniques to reduce the OLS bias, and to compare the mean square error of the OLS, COLS and JOLS(m) estimators. Our conclusions are summarized in Section 7.

2. A Small Sigma (σ) Approximation for the OLS Bias

Consider the regression model

$$y = Z\alpha + \sigma\epsilon \qquad (2.1)$$

where Z is a $T \times k$ matrix and where the disturbance term ϵ has zero mean vector and covariance matrix I_T. Let Z contain as its first column a one period lagged dependent variable and let the remaining columns be formed by a $T \times (k-1)$ non-stochastic matrix X, then (2.1) can be written as

$$y = [y_{-1} : X] \begin{bmatrix} \gamma \\ \beta \end{bmatrix} + \sigma\epsilon \qquad (2.2)$$

where

$$[y_{-1} : X] = Z \quad \text{and} \quad \begin{bmatrix} \gamma \\ \beta \end{bmatrix} = \alpha \; .$$

It is well known that least squares estimation of (2.1) may result in serious small-sample bias and for some recent evidence on this, see Kiviet (1985). We shall assume that the model is stationary, so that $|\gamma|<1$, and that the disturbances are serially independent. In an analysis of the above model based on large T asymptotics the stationarity assumption is required in order to get rid of terms which would otherwise be of higher order than $O(T^{-1})$, see Grubb and Symons (1987, p.373). The small σ analysis presented below does not require $|\gamma|<1$ as such; nevertheless we assume stationarity in order to exclude cases where the first moment of the OLS coefficient estimator does not exist. In our derivation of the approximation to the bias the assumption of normality of the disturbances is not required.

The bias in the OLS estimator of α is given by

$$E(\hat{\alpha} - \alpha) = \sigma \, E\{(Z'Z)^{-1}Z'\epsilon\} \; . \qquad (2.3)$$

By repeated substitution we find that

$$y_{-1} = y_0 c + AX\beta + \sigma A\epsilon \tag{2.4}$$

where c and A are, respectively, a $T \times 1$ vector and a $T \times T$ matrix given by

$$c = \begin{bmatrix} 1 \\ \gamma \\ \gamma^2 \\ \vdots \\ \gamma^{T-1} \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 0 & 0 & \cdots & & & 0 \\ 1 & 0 & \cdots & & & 0 \\ \gamma & 1 & & & & \vdots \\ \vdots & \gamma & \ddots & & & \vdots \\ \vdots & & \ddots & 1 & 0 & 0 \\ \gamma^{T-2} & \gamma^{T-3} & \cdots & \gamma & 1 & 0 \end{bmatrix} . \tag{2.5}$$

If we put

$$\bar{Z} = [y_0 c + AX\beta : X] \quad \text{and} \quad V = [A\epsilon : 0] \tag{2.6}$$

then

$$Z = \bar{Z} + \sigma V \tag{2.7}$$

yields a separation of Z into non–stochastic and stochastic terms, and we can prove the following:

<u>THEOREM 1</u>: *In model (2.2) with $|\gamma|<1$, y_0 and X fixed, $E\epsilon=0$ and $E\epsilon\epsilon'=I$, the approximation to order σ^2 of the bias in the OLS estimator $\hat{\alpha}$ of the complete coefficient vector α is given by*

$$E(\hat{\alpha} - \alpha) = -\sigma^2\{(\bar{Z}'\bar{Z})^{-1}\bar{Z}'A\bar{Z} + tr[(\bar{Z}'\bar{Z})^{-1}\bar{Z}'A\bar{Z}].\ I_k\}\ (\bar{Z}'\bar{Z})^{-1}e_1 + O(\sigma^3) ,$$

where $e_1 = (1 , 0 , \ldots , 0)'$ is a k x 1 unit vector.

PROOF: The term $Z'\epsilon$ in (2.3) may be written as

$$Z'\epsilon = \bar{Z}'\epsilon + \sigma V'\epsilon \ . \tag{2.8}$$

Since

$$\begin{aligned} Z'Z &= \bar{Z}'\bar{Z} + \sigma[\bar{Z}'V + V'\bar{Z}] + \sigma^2 V'V \\ &= \{I + \sigma[\bar{Z}'V + V'\bar{Z}](\bar{Z}'\bar{Z})^{-1} + \sigma^2 V'V(\bar{Z}'\bar{Z})^{-1}\}\bar{Z}'\bar{Z} \end{aligned}$$

we obtain

$$\begin{aligned} (Z'Z)^{-1} &= (\bar{Z}'\bar{Z})^{-1}\{I + \sigma[\bar{Z}'V + V'\bar{Z}](\bar{Z}'\bar{Z})^{-1} + \sigma^2 V'V(\bar{Z}'\bar{Z})^{-1}\}^{-1} \\ &= (\bar{Z}'\bar{Z})^{-1}\{I + \sigma R + \sigma^2 S\}^{-1} \end{aligned} \tag{2.9}$$

where

$$R = [\bar{Z}'V + V'\bar{Z}](\bar{Z}'\bar{Z})^{-1} \quad \text{and} \quad S = V'V(\bar{Z}'\bar{Z})^{-1} \ . \tag{2.10}$$

It follows from (2.9) that

$$(Z'Z)^{-1} = (\bar{Z}'\bar{Z})^{-1}[I - \sigma R] + O_p(\sigma^2) \tag{2.11}$$

and using (2.8), (2.9) and (2.10), the bias (2.3) may be written as

$$\begin{aligned} E(\hat{\alpha} - \alpha) &= \sigma E\{(Z'Z)^{-1}Z'\epsilon\} = \\ &= \sigma(\bar{Z}'\bar{Z})^{-1}E\{\bar{Z}'\epsilon + \sigma V'\epsilon - \sigma\bar{Z}'V(\bar{Z}'\bar{Z})^{-1}\bar{Z}'\epsilon - \sigma V'\bar{Z}(\bar{Z}'\bar{Z})^{-1}\bar{Z}'\epsilon\} + O(\sigma^3). \end{aligned} \tag{2.12}$$

To proceed further we shall evaluate the above term by term. The first term is clearly zero since $\bar{Z}$ is non-stochastic and

(i) $\qquad E(\bar{Z}'\epsilon) = 0 \ .$

Next, noting from (2.6) that $V = A\epsilon e_1'$, we obtain

(ii) $$E(V'\epsilon) = E(e_1.\ \epsilon'A'\epsilon) = e_1.E(\epsilon'A'\epsilon) = e_1.\ \text{trace } A' = 0\ ,$$

since trace A' = 0. Further, we have

(iii) $$E\{\bar{Z}'V(\bar{Z}'\bar{Z})^{-1}\bar{Z}'\epsilon\} = E\{\bar{Z}'A\epsilon.e_1'(\bar{Z}'\bar{Z})^{-1}\bar{Z}'\epsilon\} = E\{\bar{Z}'A\epsilon\epsilon'\bar{Z}(\bar{Z}'\bar{Z})^{-1}e_1\}$$
$$= \bar{Z}'A\bar{Z}(\bar{Z}'\bar{Z})^{-1}e_1\ ,$$

and

(iv) $$E\{V'\bar{Z}(\bar{Z}'\bar{Z})^{-1}\bar{Z}'\epsilon\} = E\{e_1\,\epsilon'A'\bar{Z}(\bar{Z}'\bar{Z})^{-1}\bar{Z}'\epsilon\}$$
$$= \text{trace } [A'\bar{Z}(\bar{Z}'\bar{Z})^{-1}\bar{Z}'].\ e_1\ .$$

Gathering terms from (i)–(iv), and rewriting (2.12) yields Theorem 1. □

3. The Direction of the Bias in $\hat{\gamma}$

In the pure AR(1) model (with no exogenous explanatory variables) and in the AR(1) model with an intercept the figures in Sawa (1978) and in Nankervis and Savin (1988) show that the absolute value of the bias in the least−squares estimator of the coefficient of the lagged−dependent variable γ is monotonically non−increasing with T, and with respect to the sign of the bias it is found that the OLS estimator of γ in these models has a negative bias if $\gamma > 0$. In Hoque and Peters (1986) results are presented for the model with no intercept, a one period lagged dependent variable with coefficient $\gamma > 0$ and one exogenous regressor. This regressor is either randomized and non−trended (uniformly distributed) or a simple time trend. Again it is found that the OLS estimator $\hat{\gamma}$ has negative bias. We shall now investigate whether we can obtain results on the sign of the bias in $\hat{\gamma}$ in our more general model.

We find from Theorem 1 that

$$E(\hat{\gamma} - \gamma) = -\sigma^2\{e_1'(\bar{Z}'\bar{Z})^{-1}\bar{Z}'A\bar{Z}(\bar{Z}'\bar{Z})^{-1}e_1 + tr[(\bar{Z}'\bar{Z})^{-1}\bar{Z}'A\bar{Z}].e_1'(\bar{Z}'\bar{Z})^{-1}e_1\} + O(\sigma^3). \quad (3.1)$$

On noting that $\gamma A + \gamma A' + I = \Omega$, where Ω is a T×T positive definite matrix of Toeplitz form (the familiar correlation matrix of a pure AR(1) process with autoregressive coefficient γ), it is easily shown that

$$\gamma E(\hat{\gamma} - \gamma) = -\tfrac{1}{2}\sigma^2\{e_1'[(\bar{Z}'\bar{Z})^{-1}\bar{Z}'\Omega\bar{Z}(\bar{Z}'\bar{Z})^{-1} - (\bar{Z}'\bar{Z})^{-1}]e_1 + (tr[(\bar{Z}'\bar{Z})^{-1}\bar{Z}'\Omega\bar{Z}] - k).e_1'(\bar{Z}'\bar{Z})^{-1}e_1\} + O(\sigma^3). \quad (3.2)$$

Here, some results are applicable which have been obtained −see Theil (1971, Section 6.3)− concerning the bias in the standard expressions for the variance of the OLS coefficient estimator in the regression model with fixed regressor matrix $\bar{Z}$ and disturbances with first−order serial correlation coefficient γ, and concerning the bias in the standard

OLS estimator of σ^2 in this model. It is known that for positive values of γ and smooth regressors $\bar{Z}$ the diagonal components of

$$(\bar{Z}'\bar{Z})^{-1}\bar{Z}'\Omega\bar{Z}(\bar{Z}'\bar{Z})^{-1} - (\bar{Z}'\bar{Z})^{-1}$$

are positive (the OLS expression underestimates the true variance of the OLS estimator), and also $tr[(\bar{Z}'\bar{Z})^{-1}\bar{Z}'\Omega\bar{Z}] > k$, since generally

$$tr[M_{\bar{Z}}\Omega M_{\bar{Z}}] < T - k \ , \qquad \text{with} \quad M_{\bar{Z}} = I - \bar{Z}(\bar{Z}'\bar{Z})^{-1}\bar{Z}'$$

(the standard expression involving the OLS residual sum of squares generally produces and underestimate of the disturbance variance).

The above arguments lead to the conclusion that a positive value of γ and smooth regressors in $\bar{Z}$ will generally give rise to negative values of $E(\hat{\gamma} - \gamma)$. In the simulations performed by Inder (1987), where four different empirical $\bar{Z}$ matrices are employed, negative values for this bias were found in all experiments.

4. Bias Corrected Ordinary Least Squares Estimation

From the approximation for the bias obtained in the foregoing Section a bias corrected estimator is developed along the following lines. Let $\hat{\sigma}^2$ be the OLS estimator of σ^2 given by

$$\hat{\sigma}^2 = (y - Z\hat{\alpha})'(y - Z\hat{\alpha}) \;/\; (T-k) \tag{4.1}$$

where $\hat{\alpha} = (Z'Z)^{-1}Z'y = (\hat{\gamma}\ ,\ \hat{\beta}')'$, and let $\hat{Z}$ denote the matrix

$$\hat{Z} = [y_0\hat{c} + \hat{A}X\hat{\beta} : X] \ , \tag{4.2}$$

where $\hat{c}$ and $\hat{A}$ correspond to c and A of (2.4) after substitution of $\hat{\gamma}$ for γ. Now the estimator

$$\hat{\alpha}^* = \hat{\alpha} + \hat{\sigma}^2\{(\hat{Z}'\hat{Z})^{-1}\hat{Z}'\hat{A}\hat{Z} + \text{tr}\,[(\hat{Z}'\hat{Z})^{-1}\hat{Z}'\hat{A}\hat{Z}].\ I_k\}\ (\hat{Z}'\hat{Z})^{-1}e_1 \tag{4.3}$$

will typically have smaller bias than the OLS estimator.

<u>THEOREM 2</u>: *The estimator $\hat{\alpha}^*$ given in (4.3) is unbiased to $O(\sigma^2)$ for the model of Theorem 1.*

<u>PROOF</u>: From Theorem 1 we have $\hat{\alpha} = \alpha + O_p(\sigma^2)$ and so we find

$$\hat{Z} = \bar{Z} + O_p(\sigma^2) \quad \text{and} \quad \hat{A} = A + O_p(\sigma^2) \ .$$

It follows that

$$\hat{Z}'\hat{Z} = \bar{Z}'\bar{Z} + O_p(\sigma^2) \quad ; \quad \hat{Z}'\hat{A}\hat{Z} = \bar{Z}'A\bar{Z} + O_p(\sigma^2)$$

and

$$(\hat{Z}'\hat{Z})^{-1} = (\bar{Z}'\bar{Z})^{-1} + O_p(\sigma^2) \ .$$

Furthermore we have

$$\begin{aligned} y - Z\hat{\alpha} &= \sigma\epsilon - Z(\hat{\alpha} - \alpha) \\ &= \sigma\epsilon - \sigma[\bar{Z} + \sigma V](\bar{Z}'\bar{Z})^{-1}\bar{Z}'\epsilon + O_p(\sigma^2) \\ &= \sigma M_{\bar{Z}}\epsilon + O_p(\sigma^2) \ , \end{aligned} \tag{4.4}$$

where use has been made of (2.7), (2.8) and (2.11). Upon using (4.4) we find for $\hat{\sigma}^2$ of (4.1) that

$$\hat{\sigma}^2 = \sigma^2(T-k)^{-1} \ . \ \epsilon' M_{\bar{Z}}\epsilon + O_p(\sigma^3) \ . \tag{4.5}$$

Hence, for (4.3) we obtain

$$\begin{aligned} \hat{\alpha}^* = \hat{\alpha} + \sigma^2(T-k)^{-1} \ . \ \epsilon' M_{\bar{Z}}\epsilon\{(\bar{Z}'\bar{Z})^{-1}\bar{Z}'A\bar{Z} + \\ \mathrm{tr}[(\bar{Z}'\bar{Z})^{-1}\bar{Z}'A\bar{Z}] \ . \ I_k\}(\bar{Z}'\bar{Z})^{-1}e_1 + O_p(\sigma^3) \ . \end{aligned}$$

Taking expectations and noting that $E\epsilon' M_{\bar{Z}}\epsilon = (T-k)$ we find on using Theorem 1 that

$$E\hat{\alpha}^* = \alpha + O(\sigma^3) \ . \qquad \square$$

Despite the reduction in its bias the estimator $\hat{\alpha}^*$ may suffer from several drawbacks. Firstly, it may be difficult to develop an estimator for its variance in small samples.

Secondly, its mean squared error (MSE) may be larger than for least squares, and thirdly, the incidence of estimates $\hat{\gamma}^*$ outside the stationarity region may increase, especially when γ is close to the extremes of $-1 < \gamma < 1$.

This latter problem is easily eliminated by modifying the estimator whenever $|\hat{\gamma}^*| > 0.99$. Consider the estimator $\breve{\alpha}$ of α defined by the following rule:

$$\breve{\alpha} = \hat{\alpha}^* \quad \text{if} \quad -0.99 \leq \hat{\gamma}^* \leq 0.99 \tag{4.6}$$

and otherwise

$$\breve{\gamma} = -0.99 \quad \text{if} \quad \hat{\gamma}^* < -0.99 \; ; \quad \breve{\gamma} = +0.99 \quad \text{if} \quad \hat{\gamma}^* > 0.99$$

with $\breve{\beta} = (X'X)^{-1}X(y - \breve{\gamma}y_{-1})$.

We shall call $\breve{\alpha}$ the COLS estimator of α and

$$\breve{\sigma}^2 = (y - Z\breve{\alpha})'(y - Z\breve{\alpha}) \; / \; (T - k) \tag{4.7}$$

is the COLS estimator of σ^2 .

Notice that since the distribution of $\breve{\gamma}$ is trimmed it is particularly difficult to find an estimate of the variance of $\breve{\alpha}$. From the Monte Carlo results in Section 6 we shall not only investigate the difference in bias of the OLS and COLS estimators, but also any differences in variance and in MSE.

5. The Jackknife Ordinary Least Squares Estimator

Quenouille (1956) developed a procedure which was later termed the jackknife by Tukey (1958), for reducing the bias of a consistent estimator. In addition to its bias reducing property the technique can also be used to provide approximate confidence intervals but in this study our concern is solely with bias reduction.

There have been a considerable number of published papers which examine some aspect of the jackknife procedure and the book by Gray and Schucany (1972) summarises much of the earlier work. More recently the jackknife has been studied in the context of the linear regression model and, in particular, Miller (1974) showed that the jackknifed least squares estimator of a smooth function of the regression parameters is asymptotically equivalent to the estimator itself. Hinkley (1977) also examined jackknifing in the linear regression model and he compared the standard jackknife to an alternative weighted jackknife procedure. When used to estimate a non–linear function of regression parameters, both procedures were shown to have a small bias property. Phillips and McCabe (1984) showed that the results of Miller and Hinkley could be extended and they proved that, in simultaneous equation models, the jackknifed two stage least squares estimator also has a small bias property; see also Phillips (1980).

None of the above papers were concerned with dynamic regression models and –to our knowledge– there are no published results on jackknifing in dynamic models which contain exogenous regressors, but some results have been obtained by Kew (1976). Before we investigate the jackknife in the regression model with a one period lagged dependent regressor variable, we give a brief introduction to the jackknife technique.

Let $\hat{\theta}$ be a biased but consistent estimator for an unknown parameter θ defined on a random sample $X_1, X_2, \ldots, X_T$, where

$$E(\hat{\theta}-\theta) = \frac{a_1}{T} + \frac{a_2}{T^2} + \frac{a_3}{T^3} + \ldots + \frac{a_r}{T^r} + \ldots \tag{5.1}$$

and the a_j, $j = 1, \ldots, r, \ldots$ may be functions of θ but not of T. Suppose the sample is partitioned in n subsets of size m ($n \times m = T$) and let a new random sample be formed by arbitrarily deleting a subset of size m from the original sample, then the estimator obtained when the i^{th} subset has been removed will be written as $\hat{\theta}_{-i}$, $i = 1, 2, \ldots, n$. The estimators

$$J_i(\hat{\theta}) = n\hat{\theta} - (n-1)\hat{\theta}_{-i} \qquad i = 1, 2, \ldots, n \tag{5.2}$$

are called pseudovalues of the jackknife, while the average of the pseudovalues given by

$$J(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n} J_i(\hat{\theta}) = n\hat{\theta} - \frac{n-1}{n}\sum_{i=1}^{n}\hat{\theta}_{-i} \tag{5.3}$$

is called the jackknife. Both estimators were introduced by Quenouille (1956) but the terms "jackknife" and "pseudovalues" were coined by Tukey as noted by Brillinger (1964).

Using (5.1), and noting that $nm = T$, the bias of $J_i(\hat{\theta})$ can be written as

$$E\left[J_i(\hat{\theta}) - \theta\right] = -\frac{a_2}{m^2 n(n-1)} - \frac{a_3(2T-1)}{m^3 n^2 (n-1)^2} \cdots \qquad i=1,2,\ldots,n \tag{5.4}$$

and since all the pseudovalues have the same bias, the jackknife being the average of them will also have the bias given in (5.4) and so, whereas the bias of $\hat{\theta}$ is of order $1/T$, when $a_1 \neq 0$, the bias of the jackknife is of order $1/T^2$.

The bias of the OLS estimator in the dynamic regression model (2.1) cannot, in general, be written as in (5.1) although in the simple case of the autoregressive model, bias approximations to order $1/T$ have been suggested which are of this form; see, for example, White (1961, p.89) who examines the bias in the simplest case of the conditional model examined here. Although there is no direct justification for jackknifing in the general case we will compare the bias of the jackknife ordinary least squares (JOLS) estimator with the COLS technique of Section 2.

First we present a simplified expression for the JOLS estimator in the linear regression model. As for the COLS estimator this expression will be shown to involve an adjustment to the OLS estimator. We shall write $Z_{(i)}$ and $y_{(i)}$ respectively for the m rows of Z and the m elements of y that represent the i^{th} subset. Also we shall write $Z_{(-i)}$ for the Z matrix when the i^{th} subset has been deleted; hence, $Z_{(-i)}$ is a (T−m)xk matrix. Similarly, $y_{(-i)}$ is the y vector with the components (i−1)xm+1, ... , ixm deleted. To derive the simplified expression we use the following:

LEMMA: *If A is a pxp non-singular matrix and C and D are pxq matrices such that the inverse of A+CD' exists, then this inverse is given by*

$$[A + CD']^{-1} = A^{-1} - A^{-1}C[I + D'A^{-1}C]^{-1}D'A^{-1} .$$

The JOLS estimator and its simplified (and computationally convenient) expression are given in the following:

THEOREM 3: *Let* $\hat{\alpha}_{(-i)} = [Z_{(-i)}'Z_{(-i)}]^{-1}Z_{(-i)}'y_{(-i)}$, $i = 1, \dots , n$ *(with nm=T), then according to (4.3) the JOLS(m) estimator* $\hat{\alpha}^{(m)}$ *in model (2.1) is given by*

$$\hat{\alpha}^{(m)} = n\hat{\alpha} - \frac{n-1}{n} \sum_{i=1}^{n} \hat{\alpha}_{(-i)}$$

and it can also be expressed as

$$\hat{\alpha}^{(m)} = \hat{\alpha} + \frac{n-1}{n} (Z'Z)^{-1} \sum_{i=1}^{n} Z_{(i)}'[I_m - Z_{(i)}(Z'Z)^{-1}Z_{(i)}']^{-1} \hat{\epsilon}_{(i)},$$

where $\hat{\epsilon}_{(i)}$ *contains the m elements corresponding to the* i^{th} *subset of the OLS residual vector* $\hat{\epsilon}=M_Zy$*. For m=1 this boils down to*

$$\hat{\alpha}^{(1)} = \hat{\alpha} + T^{-1}(T-1).(Z'Z)^{-1}Z'(\text{diag } M_Z)^{-1}M_Zy \; .$$

PROOF: For i = 1, 2, ... , n, we have

$$\hat{\alpha}_{(-i)} = [Z'Z - Z_{(i)}'Z_{(i)}]^{-1}[Z'y - Z_{(i)}'y_{(i)}] \; .$$

Upon using the Lemma we obtain

$$\begin{aligned}\hat{\alpha}_{(-i)} &= \\ &\{(Z'Z)^{-1} + (Z'Z)^{-1}Z_{(i)}'[I_m - Z_{(i)}(Z'Z)^{-1}Z_{(i)}']^{-1}Z_{(i)}(Z'Z)^{-1}\}[Z'y - Z_{(i)}'y_{(i)}] \\ &= \hat{\alpha} - (Z'Z)^{-1}Z_{(i)}'[I_m - Z_{(i)}(Z'Z)^{-1}Z_{(i)}']^{-1}(y_{(i)} - Z_{(i)}\hat{\alpha}) \; .\end{aligned}$$

Upon writing $y_{(i)} - Z_{(i)}\hat{\alpha} = \hat{\epsilon}_{(i)}$, the result for $\hat{\alpha}^{(m)}$ easily follows, and for m=1 and n=T the simplified expression for $\hat{\alpha}^{(1)}$ is found. □

From Theorem 3 it is seen that the jackknife involves a simple additive correction to the OLS estimator, and that this correction is a linear transformation of the OLS residuals. In the fixed regressor case, where both the OLS and JOLS estimators are unbiased, the residuals $\hat{\epsilon}$ are uncorrelated with the estimator $\hat{\alpha}$ so that $\hat{\alpha}^{(m)}$ has a covariance matrix which exceeds that of $\hat{\alpha}$ by a positive semidefinite matrix. In the context of large sample asymptotics the correction term is $O_p(1/T)$ so that as $T \to \infty$, $T^{\frac{1}{2}}$ times this term vanishes and hence, $\hat{\alpha}$ and $\hat{\alpha}^{(m)}$ are asymptotically equivalent. This large-sample result holds for the dynamic regression model in (2.1) also.

With respect to the small-sigma approximation to the bias of the simple JOLS estmator $\hat{\alpha}^{(1)}$ we have derived the following result (a proof is available from the authors upon request).

THEOREM 4: *The approximation to order σ^2 of the bias of the JOLS estimator $\hat{\alpha}^{(1)}$ of the complete coefficient vector in model (2.2) is given by*

$$
\begin{aligned}
E(\hat{\alpha}^{(1)}-\alpha) = & - \sigma^2 \{ (\bar{Z}'\bar{Z})^{-1}\bar{Z}'A\bar{Z} + tr[(\bar{Z}'\bar{Z})^{-1}\bar{Z}A'\bar{Z}].I_k \} (\bar{Z}'\bar{Z})^{-1}e_1 \\
& + \sigma^2 T^{-1}(T-1) \{ (\bar{Z}'\bar{Z})^{-1}\bar{Z}'[2 \, diag\, (M_{\bar{Z}}AM_{\bar{Z}})(diag\, M_{\bar{Z}})^{-2} \\
& \qquad -(diag\, M_{\bar{Z}})^{-1}M_{\bar{Z}}A - AM_{\bar{Z}}(diag\, M_{\bar{Z}})^{-1} \\
& \qquad + tr[(\bar{Z}'\bar{Z})^{-1}\bar{Z}'A\bar{Z}].(diag\, M_{\bar{Z}})^{-1}]\bar{Z} \\
& \qquad + tr[(M_{\bar{Z}}AM_{\bar{Z}}) (diag\, M_{\bar{Z}})^{-1}].\, I_k \} (\bar{Z}'\bar{Z})^{-1}e_1 \\
& + O(\sigma^3) .
\end{aligned}
$$

Whether or not this JOLS(1) estimator will have a smaller bias than the OLS and COLS estimators is not obvious from this result. Therefore, we shall use Monte Carlo simulation in order to assess the effectiveness of the various techniques.

6. Some Numerical Results

For some artificial models we shall examine the performance of the bias reduction techniques in practice. By Monte Carlo simulation we shall investigate the accuracy of the approximations to the actual OLS bias and we shall inspect the effectiveness of the COLS and JOLS(m) estimators with respect to bias reduction; also the efficiency in terms of the (root) mean square error of the various estimators will be assessed. All the features just mentioned before depend on y_0, X, γ, β and σ^2 in equation (2.2). For particular relevant values of these variables and these parameters we shall produce some numerical results.

We examine the first-order autoregressive model with an intercept and only one exogenous regressor

$$y_t = \gamma y_{t-1} + \beta x_t + \beta_0 + \sigma\epsilon_t , \qquad t = 1, \ldots , T \tag{6.1}$$

where $\epsilon_t \sim$ NID(0,1) . Note that β from now on denotes simply the coefficient of x_t and hence is no longer the vector containing all the coefficients of the exogenous variables as in (2.2). We shall investigate model (6.1) for cases where x_t is the realization of a random walk with drift generated according to

$$x_t = \lambda + x_{t-1} + \xi_t , \tag{6.2}$$

with x_0 fixed and $\xi_t \sim$ NID$(0,\sigma_\xi^2)$. For the starting value y_0 in (6.1) we take the "expected" non-stochastic value

$$y_0 = (\beta x_0 + \beta_0)/(1-\gamma) . \tag{6.3}$$

In Appendix A we show that in this model the major results –except those regarding the intercept– are invariant with respect to the value of the intercept β_0 itself and also with respect to the starting value x_0. We chose $\beta_0 = 0$ and $x_0 = 1$. Given these values and (6.3) we find that model (6.1) is fully characterized by the parameters $(\gamma,\beta,\sigma,\lambda,\sigma_\xi,T)$ and by the actual values $(\xi_1, \ldots ,\xi_T)$.

In all the calculations we fixed

$$\sigma_\xi = 0.02 \quad \text{and we have either} \quad \lambda = 0.00 \quad \text{or} \quad \lambda = 0.02 \; . \tag{6.4}$$

Hence, we investigate model (6.1) for the case where the x_t series is smooth (so that the results of Section 3 apply) and constitutes either a simple random walk or a trended series (random walk with drift). In our derivations the exogenous regressors are supposed to be non−stochastic. Thus, in the simulations too we should keep the x_t series fixed over the Monte Carlo replications. We shall present results for only one fixed arbitrary realization of the series $\{\xi_t\}$. For the two different values of λ two different x_t series result; these series are presented in Appendix B. Note that corresponding elements of these series simply differ 0.02t. Next to the numerical results based on these fixed x_t series we shall present particular results which indicate that our findings are not merely typical for the one and only realization of the $(\xi_1,\ldots,\xi_T)$ series, but that our general conclusions are relevant for quite a wide range of models which are met in econometric applications.

With respect to the size of the sample and the size of the omitted sub−samples when employing the m−delete Jackknife we chose

$$T = 24 \quad \text{or} \quad T = 48, \quad \text{and} \quad m = 1,\ 2,\ 3,\ 4,\ 6,\ 8,\ 12. \tag{6.5}$$

Hence, we consider a small and a moderate sample size. Moreover, the values in (6.5) are such that an integer number $n = T/m$ of sub−samples and pseudovalues is easily obtained.

We investigate various sets of values for the coefficients (γ,β). Most sets imply a long−run total multiplier of x with respect to y of unity; some calculations have been made for other values. Some sets give rise to different dynamic adjustment processes. In each experiment the value of σ is chosen such that R^2 values and t−ratios result which are rather typical for the empirical analysis of econometric time−series.

In Table 1 the main characteristics are given of the various experiments that have been carried out employing the fixed x_t series given in Appendix B. By TMP we indicate the total multiplier which in model (6.1) equals $\beta/(1-\gamma)$; MNL denotes the mean lag which amounts to $\gamma/(1-\gamma)$, and MDL gives the median lag which equals zero if $\gamma < 0.5$ and which in this model is $-\ln(2\gamma)/\ln(\gamma)$ otherwise. The chosen values for the standard deviation of the disturbances, σ, leads to values of the estimated standard deviation of the OLS coefficient estimates as given in the columns indicated by $s(\hat{\gamma})$ and by $s(\hat{\beta})$. These are averages over 1000 Monte Carlo replications of the square root of the diagonal elements of $\hat{\sigma}^2(Z'Z)^{-1}$, see (3.1). Also the averages of the R^2 values are given. We find that the chosen combination of parameter values mostly leads to regressions with reasonably accurate estimates of the coefficients and to a rather high overall fit, especially in the trended data cases ($\lambda = 0.02$). The figures presented in the last column of this Table will be discussed in due course.

Table 2 gives results on the performance of OLS and the bias correction techniques for the model with unit total multiplier and a high value ($\gamma = 0.8$) of the coefficient of the lagged dependent variable. We see that the bias of the OLS estimate of γ is negative −which is in agreement with the results of Section 3− and that the bias is also quite substantial, especially when the size of the sample is small ($T = 24$). The bias in the OLS estimate of β is positive, and −in contrast with the γ estimate− the trended x_t series gives rise to a larger bias. The values within parentheses are the standard errors of the Monte Carlo estimates of the bias, and these indicate that the bias estimates are indeed quite accurate. The COLS technique appears to lead to estimates with a smaller bias (in absolute value) than OLS. However, the remaining bias is often still quite substantial and it has opposite sign, so in fact the small σ approximation to the bias appears to give rise to an overcorrection of the actual bias in the cases here investigated. Nevertheless, on a root mean square error (rmse) criterion several of the COLS estimates are superior to the OLS estimates.

With respect to the JOLS results we find from Table 2 that the standard jackknife (m=1) never outperforms OLS. Usually the bias and the rmse have either increased a bit or they remain unchanged. Increasing the value of m (omitting more than one observation) leads initially to smaller biases and also to more favourable rmse values. However, before the bias gets really small −which is here usually the case for m values in the range 6 through 12− the efficiency of JOLS deteriorates. We find that particular JOLS(m) estimates may lead in this model to virtually unbiased estimates, but at the expense of a rmse which exceeds the OLS value considerably. We also find that the value of m which produces the best JOLS results differs for γ and for β and it is also different for the small and for the moderate sample size; in addition, the optimal m differs for the trended and non−trended data.

Thus, the results in Table 2 present a mixed picture; none of the techniques investigated outperforms all the others in all respects. Apparently for $\sigma \geqslant 0.01$ the $O(\sigma^2)$ approximation to the bias lacks accuracy for these particular experiments with the non−trended x_t series. On the other hand, in the trended data case −where the R^2 values exceed 0.99− the COLS results are favourable, both regarding the remaining bias and regarding the resulting efficiency. The surprising low value for the rmse of the COLS estimate for γ in the $\lambda=0.00$ and $T=24$ case will partly be due to the gross overstatement of the actual bias by the $O(\sigma^2)$ approximation. From the last column of Table 1 we see that in 181 replications the COLS estimate was set at 0.99; this has of course had a mitigating effect on the bias and the rmse of the COLS estimator of γ.

In the Tables 3A and 3B we present more results for the model with unit long−run elasticity; now we examine the intermediate mean lag parametrization ($\gamma=0.4$) and chose two different values of σ. Again we find that the actual OLS bias may be quite substantial, and that the COLS procedure overcorrects for bias. Here this correction never leads to an efficiency gain. Only for the models with the smaller value of σ ($=0.010$) it does appear reasonable to apply the COLS procedure. For higher values of σ the loss in efficiency is quite substantial, and with respect to the bias the major effect is that it just changes sign, without a considerable reduction in magnitude. Again it is found that the simple jackknifed estimator (m=1) is equivalent to or worse than OLS. However, for m= 2, 3 or 4 the JOLS procedure may produce estimates which are more efficient than OLS and which have a bias about half as large as the OLS bias. Higher values of m may lead to almost unbiased estimates, but these appear to be inefficient. For the cases with no dynamic adjustments ($\gamma = 0.0$) and those with a total multiplier different from unity we do not tabulate all JOLS(m) results in detail (some are given in the Tables 4A and 4B). Here much the same picture results, except that for $\gamma = 0.0$ most JOLS estimates (for any m) have a rmse inferior to OLS.

The Tables 4A and 4B contain results for various parametrizations and for the non-trended and trended data respectively. For OLS and JOLS(1) the $O(\sigma^2)$ approximation to the bias is calculated. For OLS this approximation is given in Theorem 1; from Theorem 2 it follows that this approximation is zero for COLS; and for JOLS(1) this approximation is given in Theorem 4. Note that these approximations do not result from Monte Carlo experiments and have been calculated employing the correct parameter values. The figures for the actual bias and for the relative efficiency of COLS have been obtained from Monte Carlo experiments and we recall that the simulation estimates proved to be quite accurate. As in the Tables 2, 3A and 3B we see that OLS and JOLS(1) give rise to almost identical actual bias figures. The small σ approximations to the bias of OLS and JOLS(1) are very close to each other too. They both overstate the actual bias, and they overstate it considerably if the actual bias in $\hat{\gamma}$ exceeds 0.02 in absolute value. This reveals why the COLS procedure overcorrects for bias; on the whole it is found that COLS leads to serious overcorrection if the approximation to the bias in $\hat{\gamma}$ exceeds 0.05 in absolute value. The fact that the small sigma approximation to the bias is deficient in cases where the value of σ is not trifling is also seen from the following. Since the $O(\sigma^2)$ approximations are multiples of σ^2 the bias approximations obtained at $\sigma = 0.015$ are 2.25 times the corresponding approximation values for $\sigma = 0.01$; and at $\sigma = 0.02$ they are 4 times as large as the values obtained for $\sigma = 0.01$. However, it appears from the simulation results that the actual bias is not simply a multiple of σ^2. By increasing σ from 0.01 to 0.02 in our experiments ($\gamma = 0.0$; $\beta = 1.0$) the actual bias is not blown up by a factor 4, but only by about a factor 2; and also, increasing σ from 0.01 to 0.015 (at

$\gamma = 0.4$; $\beta = 0.6$) does not even double the actual bias.

From the Tables 4A and 4B we want to find out how the actual OLS bias and the performance of COLS depends on the values of the parameters σ, γ, β, λ and T, and whether simple clues can be found which are helpful for practitioners to detect severe vulnerability with respect to small sample bias of OLS estimates and to indicate the potential success of COLS estimates. In the Tables the performance of COLS is expressed by the remaining bias and by the relative efficiency of the COLS estimators for γ and β. This relative efficiency is the ratio of the rmse of the COLS and the OLS estimators respectively. In the last column of the Tables we give an Indicator for the Effectiveness of COLS [IEC] as defined in (6.6), and we also present s.d.($\hat{\gamma}$), the Monte Carlo estimate of the actual standard deviation of the OLS estimator $\hat{\gamma}$. Upon comparison with s($\hat{\gamma}$) of Table 1 we see that −despite the bias in $\hat{\gamma}$− the dispersion of $\hat{\gamma}$ is reasonably well seized by the standard formula.

Before we concentrate on the results given in the Tables 4A and 4B we want to recall the invariance properties (see Appendix A) with respect to the value of the intercept term in the regression and regarding the level of the regressor variable, and we also want to mention yet another invariance result. In Appendix C we show that by multiplying β and σ by the same factor the relative (approximation to the) bias in both $\hat{\gamma}$ and $\hat{\beta}$ remains unaltered, and so does the relative performance of COLS and JOLS(m). Note that by such multiplication y_0 is multiplied by the same factor as well, because of (6.3). It is also easy to show that multiplying β, σ and y_0 by the same factor does not alter the values of R^2, s($\hat{\gamma}$), and s($\hat{\beta}$)/β (the square root of the variance ratio of $\hat{\beta}$). This can be checked from the results presented in Table 1 for T=24 and $\gamma = 0.4$. Because of this invariance property we omitted in Table 1 the invariant cases for T=48. For the same reason we did not include results for such parametrizations in the Tables 4A and 4B.

From the Tables it is seen that changes in the parameters β and σ and also changes in γ which induce changes in R^2 and s($\hat{\gamma}$), have a great affect on the (relative) bias and the efficiency of the various estimators. In the experiment for non−trended data with

$\gamma = 0.0$ and $\sigma = 0.020$ the inaccuracy of the bias approximation is such that the absolute value of the bias is even larger for COLS than it is for OLS. From Table 1 we see that this is the parametrization with the lowest R^2 value and with the highest value for $s(\hat{\gamma})$. Nevertheless, neither the value of R^2 nor that of $s(\hat{\gamma})$ can directly be used as a simple indicator of the potential success of the COLS procedure. This is illustrated by the following. In Table 4A reasonable COLS results are presented for a model with $R^2 = 0.887$ and $T = 24$, whereas in Table 4B unsatisfactory results are given for a model where $R^2 = 0.986$ and again $T = 24$. From Table 1 we observe that the value of R^2 depends heavily on trends in the variables (λ), and from the Tables 4A and 4B we see that the value of λ has only a moderate effect on the bias and the performance of COLS. For a given value of T, changes in the parametrization appear to lead to changes in the performance of COLS which quite well match the changes in $s(\hat{\gamma})$. However, in Table 4A reasonable COLS results are presented for a model with $s(\hat{\gamma}) = 0.111$ and $T = 24$, whereas for $T = 48$ unsatisfactory results are obtained for a model with $s(\hat{\gamma})$ as small as 0.109 .

We investigated whether the value of

$$\text{IEC} = T.\hat{\text{var}}(\hat{\gamma}) = T.\hat{\sigma}^2/y_{-1}'M_X y_{-1} = T(T-k)^{-1}.\epsilon'M_X\epsilon/y_{-1}'M_X y_{-1}, \qquad (6.6)$$

where $M_X = I - X(X'X)^{-1}X'$, is helpful in indicating the effectiveness of the COLS procedure. This value [in fact we present $T.s(\hat{\gamma})^2$] is given in the last column of the Tables 4A and 4B. We find that in the models where the IEC is below 0.20 the COLS procedure gives very satisfactory results; if this indicator is above 0.25 the COLS leads to losses in efficiency and to a substantial overcorrection for bias. Given these findings our tentative conclusions are that if for a particular model an IEC value is obtained well above 0.25 then the characteristics of the model are likely to be such that OLS will lead to estimates with a substantial bias; moreover, the $O(\sigma^2)$ approximation to this bias produces a considerable overstatement of the actual bias such that COLS estimates are inefficient and severely biased as well. If on the other hand the IEC value is around 0.25 or below then OLS can still lead to estimates with a substantial bias; again the $O(\sigma^2)$

approximation to this bias will usually overstate the actual bias, but nevertheless the correction of the OLS estimates will be worthwhile. Our results on the efficiency of COLS and on the accuracy of the standard expression for the OLS variance indicate that for models where the value of the indicator is below 0.25 the standard expression for the OLS standard deviation usually is a conservative estimate of the COLS standard deviation. Finally we conclude from the Tables 4A and 4B that if the approximation to the bias in $\hat{\gamma}$ is above 0.05 it might be a good idea to deflate the correction term in order to mitigate overcorrection and efficiency losses.

A few results for random series of the exogenous variable x_t are presented in Table 5. These results support our view that the findings for the arbitrarily chosen fixed x_t series do in fact represent quite general characteristics for models of the type (6.1) with a regressor variable of the type (6.2). For 1000 replications of the various parametrizations the approximation to the OLS bias is calculated; presented in the Table are the average and the standard deviation (in parentheses) over this Monte Carlo sample. Also the average bias of the OLS and the COLS estimators is calculated. We see that the arbitrary single realization of the x_t series used in the foregoing experiments happens to be a bit more vulnerable with respect to bias than the average x_t series of this type. From the standard deviation of the approximation values to the bias we also see that the vulnerability with respect to bias varies for different realizations of the x_t series. On average the COLS procedure proves to have bias reduction qualities which are quite well represented by the results for the fixed x_t series.

Finally we want to note here that all the numerical results in this Section have been obtained for models with normally distributed disturbances, whereas neither the small σ approximations to the bias nor the COLS and JOLS procedures actually require normality. We have not examined to what degree this has influenced our findings.

7. Conclusion

Employing small disturbance asymptotics we derive an approximation to the bias of the OLS coefficient estimator in the regression model with fixed regressors except for a one period lagged dependent variable. From this approximation we conclude that the coefficient of the lagged dependent variable will usually be underestimated if it is positive, and we show how the approximation can be used to construct a corrected estimator (COLS) which has a bias of smaller order and which is quite easy to calculate. We also derive a simple expression for the m−delete jackknifed estimator JOLS(m), and give an approximation to its bias for $m=1$. A rather complex term represents the difference in the bias of the OLS and the JOLS(1) estimators. We calculate this term for various artificial models and find it to be practically zero and so it seems that the simple jackknife does not reduce the bias of the OLS estimator in this model.

For inference purposes one would like to obtain estimates which are properly centred, and which are reasonably efficient. By performing Monte Carlo experiments we find that JOLS(m) may produce virtually unbiased estimates for $m>>1$, even if the bias is quite substantial, but that these estimates are in general inefficient. There seem to be no simple clues with respect to the choice of the optimal value for m; this is a serious drawback of the JOLS(m) estimator. From the Monte Carlo experiments we also find that the $O(\sigma^2)$ approximation to the OLS bias generally overshoots the actual bias, especially when the bias is large. As the actual bias can be very substantial −even for models with moderate values of the coefficient of the lagged dependent variable and with high values of R^2 (>0.99)− there is a serious inference problem indeed in lagged dependent variable models. The sample size, the disturbance variance, and the variance(−ratios) of the individual coefficients are found to be major determining factors of the (relative) OLS bias. We observe that in the class of models investigated the effectiveness of the COLS procedure is closely connected with the value of the product of the sample size and the variance of the OLS estimate of the coefficient of the lagged dependent variable. If in our simulation models the value of this indicator is below 0.25 then OLS estimates may show substantial bias, whereas COLS works quite well. If the indicator is above this value both OLS and COLS will produce biased estimators, where COLS overcorrects for bias and leads to losses in efficiency. Since there is no doubt that serious errors may be committed if the usual methods are applied in finite samples of dynamic models, it seems to be of great importance to enhance the development of tools for the detection of vulnerability with respect to bias due to the presence of lagged dependent variables, and of techniques which produce virtually unbiased and reasonably efficient estimators.

References

Brillinger, D.R. (1964), "The Asymptotic Behaviour of Tukey's General Method of Setting Approximate Confidence Limits (The Jackknife) when Applied to Maximum Likelihood Estimates", Review of the International Statistical Institute, 32, 202−206.

Gray, H.L. and Schucany, W.R. (1972), "The Generalised Jackknife Statistic", Marcel Dekker Inc. New York.

Grubb, D. and Symons, J. (1987), "Bias in Regressions with a Lagged Dependent Variable", Econometric Theory, 3, 371−386.

Hinkley, D.V. (1977), "Jackknifing in Unbalanced Situations", Technometrics, 19, No.3, 285−292.

Hoque, A. and Peters, T.A. (1986), "Finite Sample Analysis of the ARMAX Models", Sankhyã, 84, Series B, 266−283.

Inder, B.A. (1987), "Bias in the Ordinary Least Squares Estimator in the Dynamic Linear Regression Model with Autocorrelated Disturbances", Working Paper No.10/87, Monash University, Clayton, Australia.

Kadane, J.B. (1971), "Comparison of K−class Estimators when the Disturbances are small", Econometrica, 9, 723−737.

Kew, S.N. (1976), "Jackknifing the 2SLS Estimator in a Simultaneous Equation System Containing a Lagged Endogenous Variable", unpublished M.Sc.Thesis, University of Kent at Canterbury.

Kiviet, J.F. (1985), "Model Selection Test Procedures in a Single Linear Equation of a Dynamic Simultaneous Equation System and Their Defects in Small Samples", Journal of Econometrics, 28, 327−362.

Kiviet, J.F. and Phillips, G.D.A. (1986). "Bias Correction in Lagged Dependent Variable Models", (paper presented at ESEM 1986 Budapest), Report AE 17/86, University of Amsterdam.

Maekawa, K. (1983), "An Approximation to the Distribution of the Least Squares Estimator in an Autoregressive Model with Exogenous Variables", Econometrica, 51, 229−238.

Malinvaud, E. (1980), "Statistical Methods of Econometrics", third revised edition, North−Holland, Amsterdam.

Miller, R.G. (1974), "An Unabalanced Jackknife", Annals of Statistics, 2, 880−891.

Nankervis, J.C. and Savin, N.E. (1988), "The Exact Moments of the Least−Squares Estimator for the Autoregressive Model: Corrections and Extensions", Journal of Econometrics, 37, 381−388.

Orcutt, G.H. and Winokur, H.S. (1969), "First Order Autoregression: Inference, Estimation and Prediction", Econometrica, 37, 1−14.

Owen, A.D. and Phillips, G.D.A. (1975) "Bias Reduction and Approximate Confidence Intervals for the Jackknifed 2SLS Estimator", Paper presented to the World Congress of the Econometric Society, Toronto.

Pantula, S.G. and Fuller W.A. (1985), "Mean Estimation Bias in Least Squares Estimation of Autoregressive Processes", Journal of Econometrics, 27, 99−121.

Phillips, G.D.A. and McCabe, B.P. (1984), "Jackknifing the Two−Stage Least−Squares Estimator", Communications in Statistics: Simulation and Computation, 13, No. 2, 141−152.

Quenouille, M.H. (1956), "Notes on Bias in Estimation", Biometrika, 43, 353−369.

Sawa, T. (1978), "The Exact Moments of the Least Squares Estimator for the Autoregressive Model", Journal of Econometrics, 8, 159−172.

Theil, H. (1971), "Principles of Econometrics", North−Holland, Amsterdam.

Tukey, J.W. (1958), "Bias and Confidence in not quite Large Samples" Annals of Mathematical Statistics, 29, (Abstract), p.614.

White, J.S. (1961), "Asymptotic Expansions for the Mean and Variance of the Serial Correlation Coefficient", Biometrika, 48, 85−95.

APPENDIX A

We shall show that in the simulation model (6.1) the major results are invariant with respect to the values x_0 and β_0. Consider the case where values x_0^* and β_0^* have been chosen instead of x_0 and β_0 respectively. Then we obtain

$$x_t^* = \lambda + x_{t-1}^* + \xi_t = x_t + (x_0^* - x_0) \tag{A.1}$$

and also

$$\begin{aligned} y_t^* &= \gamma y_{t-1}^* + \beta x_t^* + \beta_0^* + \sigma\epsilon_t \\ &= y_t + \gamma(y_{t-1}^* - y_{t-1}) + \beta(x_t^* - x_t) + (\beta_0^* - \beta_0) \\ &= y_t + \gamma(y_{t-1}^* - y_{t-1}) + \beta(x_0^* - x_0) + (\beta_0^* - \beta_0) \ . \end{aligned}$$

According to (6.3)

$$y_0^* = (\beta x_0^* + \beta_0^*)/(1 - \gamma) = y_0 + [\beta(x_0^* - x_0) + (\beta_0^* - \beta_0)] \ / \ (1 - \gamma) \ ,$$

and so we can write

$$y_t^* = y_t + \gamma(y_{t-1}^* - y_{t-1}) + (1 - \gamma)(y_0^* - y_0) \ .$$

From this we find

$$y_t^* = y_t + (y_0^* - y_0) \ . \tag{A.2}$$

Thus from (A.1) and (A.2) it follows that in model (6.1) all the regressor vectors (apart from the constant) and also the regressand are affected, i.e. some constant value is added to each observation.

To investigate the consequences of these changes we consider now the more general model

$$y^* = Z^*\alpha^* + \sigma\epsilon , \tag{A.3}$$

and examine some of its differences with the original model (2.1). Suppose that the last column of both Z and Z^* relates to the intercept, and that we have

$$y^* = y + \delta_0\iota \quad \text{and} \quad Z^* = ZD \tag{A.4}$$

where ι is a Tx1 vector with all elements equal to unity, and where

$$D = \begin{bmatrix} 1 & 0 & . & . & . & . & 0 \\ 0 & 1 & . & . & & & . \\ . & . & . & . & . & & . \\ . & . & . & . & . & . & . \\ . & . & & . & . & . & . \\ 0 & . & . & . & 0 & 1 & 0 \\ \delta_1 & \delta_2 & . & . & . & \delta_{k-1} & 1 \end{bmatrix}, \text{ for which } D^{-1} = \begin{bmatrix} 1 & 0 & . & . & . & . & 0 \\ 0 & 1 & . & . & . & & . \\ . & . & . & . & . & & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & . & & & 0 & 1 & 0 \\ -\delta_1 & -\delta_2 & . & . & . & -\delta_{k-1} & 1 \end{bmatrix},$$

and where $\delta_0, \delta_1, \ldots, \delta_{k-1}$ are arbitrary constants. If α^* is such that only its last element differs from α, then by taking matching values of $\delta_0, \ldots, \delta_{k-1}$ model (A.3) can represent our simulation model (6.1) for any particular choice of x_0^* and β_0^* .

For the OLS estimator $\hat{\alpha}^*$ of α^* in (A.3) we obtain

$$\begin{aligned} \hat{\alpha}^* &= (Z^{*\prime}Z^*)^{-1}Z^{*\prime}y^* \\ &= D^{-1}(Z'Z)^{-1}Z'[y + \delta_0\iota] \\ &= D^{-1}\hat{\alpha} + \delta_0 D^{-1}e_k = D^{-1}\hat{\alpha} + \delta_0 e_k . \end{aligned} \tag{A.5}$$

Hence, we see that – due to the special structure of D^{-1} – only the estimate of the intercept is affected. The same holds for the approximation of $E(\hat{\alpha}-\alpha)$ and for the JOLS(m) estimator $\hat{\alpha}^{(m)}$.

APPENDIX B

The fixed x_t-series

t	λ=0.00	λ=0.02
0	1.000	1.000
1	1.003	1.023
2	0.981	1.021
3	0.981	1.041
4	1.009	1.089
5	0.985	1.085
6	0.980	1.100
7	1.005	1.145
8	1.012	1.172
9	0.994	1.174
10	0.980	1.180
11	0.949	1.169
12	0.978	1.218
13	0.971	1.231
14	0.954	1.234
15	0.925	1.225
16	0.930	1.250
17	0.931	1.271
18	0.916	1.276
19	0.939	1.319
20	0.944	1.344
21	0.945	1.365
22	0.961	1.401
23	0.950	1.410
24	0.938	1.418
25	0.940	1.440
26	0.921	1.441
27	0.902	1.442
28	0.906	1.466
29	0.935	1.515
30	0.917	1.517
31	0.918	1.538
32	0.943	1.583
33	0.969	1.629
34	1.000	1.680
35	0.998	1.698
36	0.996	1.716
37	1.010	1.750
38	1.007	1.767
39.	0.984	1.764
40	0.989	1.789
41	0.999	1.819
42	1.000	1.840
43	1.010	1.870
44	1.032	1.912
45	1.037	1.937
46	1.066	1.986
47	1.116	2.056
48	1.079	2.039

APPENDIX C

Suppose that β, σ and y_0 of model (2.2) are multiplied by the same factor μ, giving

$$\beta^* = \mu\beta \quad , \quad \sigma^* = \mu\sigma \quad \text{and} \quad y_0^* = \mu y_0 \ . \tag{C.1}$$

Then it follows from (2.4) and (2.2) that

$$y_{-1}^* = \mu y_{-1} \quad \text{and} \quad y^* = \mu y \ . \tag{C.2}$$

Now consider the model

$$y^* = Z^*\alpha^* + \sigma^*\epsilon \tag{C.3}$$

with $Z^* = [y_{-1}^* : X] = [\mu y_{-1} : X] = MZ$ and $\alpha^* = (\gamma \ , \ \beta^{*\prime})'$, where

$$M = \begin{bmatrix} \mu & 0 & . & . & . & . & 0 \\ 0 & 1 & 0 & . & . & . & . \\ . & 0 & 1 & . \ 0 & . & . & . \\ . & & & & & & . \\ . & . & . & . & . & 1 & 0 \\ 0 & . & . & . & . & 0 & 1 \end{bmatrix} .$$

For the bias of the OLS estimator $\hat{\alpha}^*$ of α^* in (C.3) we easily find

$$E(\hat{\alpha}^* - \alpha) = \sigma^* E\{(Z^{*\prime}Z^*)^{-1}Z^{*\prime}\epsilon\} = \mu M^{-1}E(\hat{\alpha} - \alpha) \ . \tag{C.4}$$

Thus, $E(\hat{\gamma}^* - \gamma^*) = E(\hat{\gamma} - \gamma)$ and $E(\hat{\beta}^* - \beta^*) = \mu E(\hat{\beta} - \beta)$, and the relative biases remain unaltered. Along the same lines one can derive comparable results for the small disturbance approximation to the OLS bias and for the COLS and JOLS(m) estimators.

TABLE 1 Characteristics of the various experiments with the fixed x_t series and some findings over 1000 Monte Carlo replications

	γ	β	TMP	MNL	MDL	σ	$s(\hat{\gamma})$	$s(\hat{\beta})$	R^2	% $\check{\gamma}$= 0.99
λ=0.00										
T=24										
	0.0	1.0	1.0	0.00	0.00	0.020	0.163	0.197	0.697	0.0
	0.0	1.0	1.0	0.00	0.00	0.010	0.106	0.113	0.901	0.0
	0.0	0.6	0.6	0.00	0.00	0.010	0.149	0.103	0.767	0.0
	0.4	1.2	2.0	0.67	0.00	0.020	0.111	0.223	0.887	0.0
	0.4	0.6	1.0	0.67	0.00	0.015	0.145	0.155	0.777	0.5
	0.4	0.6	1.0	0.67	0.00	0.010	0.111	0.111	0.887	0.0
	0.4	0.3	0.5	0.67	0.00	0.005	0.111	0.056	0.887	0.0
	0.8	0.3	1.5	4.00	2.11	0.010	0.086	0.103	0.920	2.1
	0.8	0.2	1.0	4.00	2.11	0.010	0.116	0.097	0.833	18.1
T=48										
	0.0	1.0	1.0	0.00	0.00	0.020	0.109	0.114	0.845	0.0
	0.0	1.0	1.0	0.00	0.00	0.010	0.070	0.068	0.956	0.0
	0.0	0.6	0.6	0.00	0.00	0.010	0.100	0.061	0.887	0.0
	0.4	0.6	1.0	0.67	0.00	0.015	0.093	0.086	0.892	0.0
	0.4	0.6	1.0	0.67	0.00	0.010	0.071	0.063	0.948	0.0
	0.8	0.3	1.5	4.00	2.11	0.010	0.047	0.047	0.957	0.0
	0.8	0.2	1.0	4.00	2.11	0.010	0.064	0.044	0.911	0.5
λ=0.02										
T=24										
	0.0	1.0	1.0	0.00	0.00	0.020	0.160	0.161	0.976	0.0
	0.0	1.0	1.0	0.00	0.00	0.010	0.103	0.103	0.994	0.0
	0.0	0.6	0.6	0.00	0.00	0.010	0.146	0.088	0.983	0.0
	0.4	1.2	2.0	0.67	0.00	0.020	0.106	0.207	0.994	0.0
	0.4	0.6	1.0	0.67	0.00	0.015	0.139	0.137	0.986	0.0
	0.4	0.6	1.0	0.67	0.00	0.010	0.106	0.103	0.994	0.0
	0.4	0.3	0.5	0.67	0.00	0.005	0.106	0.052	0.994	0.0
	0.8	0.3	1.5	4.00	2.11	0.010	0.073	0.094	0.997	1.1
	0.8	0.2	1.0	4.00	2.11	0.010	0.101	0.086	0.992	5.3
T=48										
	0.0	1.0	1.0	0.00	0.00	0.020	0.109	0.107	0.996	0.0
	0.0	1.0	1.0	0.00	0.00	0.010	0.069	0.068	0.999	0.0
	0.0	0.6	0.6	0.00	0.00	0.010	0.099	0.058	0.997	0.0
	0.4	0.6	1.0	0.67	0.00	0.015	0.091	0.088	0.998	0.0
	0.4	0.6	1.0	0.67	0.00	0.010	0.069	0.067	0.999	0.0
	0.8	0.3	1.5	4.00	2.11	0.010	0.044	0.060	0.999	0.0
	0.8	0.2	1.0	4.00	2.11	0.010	0.061	0.054	0.999	0.5

TABLE 2 Bias and root mean squared error of OLS, COLS, and JOLS(m) estimates found from 1000 replications of model (5.1) with the fixed x_t series and for γ=0.8, β=0.2, and σ=0.010

			γ bias	(s.e.)	rmse	β bias	(s.e.)	rmse
λ=0.00	T=24							
		OLS	-0.101	(.004)	0.161	0.054	(.003)	0.114
		COLS	0.045	(.004)	0.136	-0.027	(.004)	0.125
		JOLS m=1	-0.106	(.004)	0.165	0.056	(.003)	0.115
		2	-0.090	(.004)	0.154	0.049	(.003)	0.113
		3	-0.075	(.004)	0.148	0.043	(.003)	0.113
		4	-0.067	(.004)	0.154	0.045	(.003)	0.115
		6	-0.012	(.006)	0.184	0.024	(.004)	0.123
		8	0.005	(.007)	0.207	0.039	(.004)	0.131
		12	0.015	(.006)	0.199	0.077	(.004)	0.158
	T=48							
		OLS	-0.040	(.002)	0.078	0.022	(.001)	0.052
		COLS	0.032	(.002)	0.077	-0.013	(.002)	0.053
		JOLS m=1	-0.040	(.002)	0.078	0.022	(.001)	0.052
		2	-0.034	(.002)	0.074	0.019	(.001)	0.051
		3	-0.028	(.002)	0.071	0.016	(.002)	0.050
		4	-0.023	(.002)	0.070	0.014	(.002)	0.051
		6	-0.014	(.002)	0.068	0.010	(.002)	0.068
		8	-0.011	(.002)	0.069	0.009	(.002)	0.069
		12	-0.002	(.002)	0.076	0.007	(.002)	0.055
λ=0.02	T=24							
		OLS	-0.072	(.003)	0.126	0.059	(.003)	0.104
		COLS	0.009	(.004)	0.112	-0.011	(.003)	0.101
		JOLS m=1	-0.074	(.003)	0.128	0.062	(.003)	0.105
		2	-0.058	(.003)	0.121	0.048	(.003)	0.100
		3	-0.046	(.004)	0.122	0.037	(.003)	0.101
		4	-0.048	(.004)	0.123	0.039	(.003)	0.123
		6	-0.025	(.004)	0.133	0.017	(.003)	0.112
		8	-0.026	(.004)	0.133	0.023	(.003)	0.109
		12	-0.003	(.005)	0.149	0.015	(.003)	0.109
	T=48							
		OLS	-0.031	(.002)	0.067	0.028	(.002)	0.060
		COLS	0.011	(.002)	0.067	-0.009	(.002)	0.060
		JOLS m=1	-0.032	(.002)	0.067	0.028	(.002)	0.060
		2	-0.026	(.002)	0.065	0.023	(.002)	0.065
		3	-0.022	(.002)	0.064	0.019	(.002)	0.064
		4	-0.019	(.002)	0.063	0.017	(.002)	0.063
		6	-0.012	(.002)	0.063	0.011	(.002)	0.056
		8	-0.010	(.002)	0.064	0.009	(.002)	0.057
		12	-0.002	(.002)	0.071	0.003	(.002)	0.062

TABLE 3A — Bias and root mean squared error of OLS, COLS, and JOLS(m) estimates found from 1000 replications of model (5.1) with the fixed x_t series and for γ=0.4, β=0.6, and σ=0.010

				γ		β	
				bias (s.e.)	rmse	bias (s.e.)	rmse
λ=0.00	T=24						
		OLS		-0.039 (.003)	0.115	0.028 (.003)	0.114
		COLS		0.015 (.004)	0.119	-0.010 (.004)	0.119
		JOLS	m=1	-0.041 (.003)	0.116	0.030 (.003)	0.114
			2	-0.027 (.003)	0.113	0.020 (.004)	0.113
			3	-0.020 (.004)	0.114	0.015 (.004)	0.114
			4	-0.015 (.004)	0.117	0.014 (.004)	0.114
			6	-0.004 (.004)	0.122	0.005 (.004)	0.120
			8	-0.005 (.004)	0.141	0.010 (.004)	0.134
			12	0.006 (.005)	0.151	0.017 (.004)	0.142
	T=48						
		OLS		-0.017 (.002)	0.071	0.015 (.002)	0.064
		COLS		0.008 (.002)	0.072	-0.004 (.002)	0.065
		JOLS	m=1	-0.016 (.002)	0.072	0.014 (.002)	0.064
			2	-0.011 (.002)	0.071	0.010 (.002)	0.064
			3	-0.007 (.002)	0.070	0.007 (.002)	0.064
			4	-0.004 (.002)	0.070	0.006 (.002)	0.065
			6	-0.001 (.002)	0.071	0.003 (.002)	0.066
			8	-0.001 (.002)	0.071	0.003 (.002)	0.067
			12	0.002 (.002)	0.073	0.002 (.002)	0.069
λ=0.02	T=24						
		OLS		-0.033 (.003)	0.110	0.032 (.003)	0.106
		COLS		0.008 (.004)	0.113	-0.008 (.003)	0.109
		JOLS	m=1	-0.035 (.003)	0.111	0.034 (.003)	0.107
			2	-0.021 (.003)	0.108	0.021 (.003)	0.105
			3	-0.015 (.003)	0.109	0.014 (.003)	0.105
			4	-0.010 (.003)	0.108	0.010 (.003)	0.105
			6	-0.002 (.004)	0.115	0.001 (.004)	0.113
			8	0.001 (.004)	0.129	0.000 (.004)	0.125
			12	0.004 (.004)	0.121	-0.001 (.004)	0.112
	T=48						
		OLS		-0.016 (.002)	0.069	0.015 (.002)	0.067
		COLS		0.005 (.002)	0.069	-0.005 (.002)	0.067
		JOLS	m=1	-0.016 (.002)	0.069	0.015 (.002)	0.067
			2	-0.010 (.002)	0.068	0.010 (.002)	0.066
			3	-0.008 (.002)	0.068	0.008 (.002)	0.066
			4	-0.005 (.002)	0.068	0.005 (.002)	0.066
			6	-0.002 (.002)	0.069	0.002 (.002)	0.067
			8	-0.001 (.002)	0.069	0.001 (.002)	0.068
			12	0.001 (.002)	0.071	0.000 (.002)	0.069

TABLE 3B Bias and root mean squared error of OLS, COLS, and JOLS(m) estimates found from 1000 replications of model (5.1) with the fixed x_t series and for γ=0.4, β=0.6, and σ=0.015

			γ			β		
			bias	(s.e.)	rmse	bias	(s.e.)	rmse
λ=0.00	T=24							
		OLS	-0.065	(.004)	0.153	0.047	(.005)	0.161
		COLS	0.058	(.005)	0.181	-0.038	(.006)	0.183
		JOLS m=1	-0.068	(.004)	0.155	0.049	(.005)	0.162
		2	-0.044	(.004)	0.148	0.032	(.005)	0.159
		3	-0.032	(.005)	0.149	0.025	(.005)	0.161
		4	-0.025	(.005)	0.153	0.024	(.005)	0.160
		6	-0.008	(.005)	0.163	0.009	(.005)	0.169
		8	-0.009	(.006)	0.182	0.018	(.006)	0.187
		12	-0.002	(.006)	0.193	0.036	(.007)	0.214
	T=48							
		OLS	-0.028	(.003)	0.095	0.024	(.003)	0.089
		COLS	0.027	(.003)	0.102	-0.018	(.003)	0.094
		JOLS m=1	-0.028	(.003)	0.095	0.024	(.003)	0.089
		2	-0.018	(.003)	0.093	0.017	(.003)	0.088
		3	-0.011	(.003)	0.092	0.012	(.003)	0.089
		4	-0.007	(.003)	0.092	0.009	(.003)	0.090
		6	-0.002	(.003)	0.094	0.004	(.003)	0.093
		8	0.000	(.003)	0.095	0.004	(.003)	0.094
		12	0.003	(.003)	0.098	0.002	(.003)	0.096
λ=0.02	T=24							
		OLS	-0.058	(.004)	0.147	0.056	(.004)	0.142
		COLS	0.031	(.005)	0.161	-0.030	(.005)	0.156
		JOLS m=1	-0.061	(.004)	0.149	0.059	(.004)	0.144
		2	-0.037	(.004)	0.143	0.036	(.004)	0.138
		3	-0.024	(.004)	0.143	0.024	(.004)	0.138
		4	-0.016	(.005)	0.143	0.016	(.004)	0.139
		6	-0.002	(.005)	0.157	0.000	(.005)	0.154
		8	0.004	(.005)	0.172	-0.002	(.005)	0.167
		12	0.006	(.005)	0.164	-0.001	(.005)	0.152
	T=48							
		OLS	-0.027	(.003)	0.091	0.027	(.003)	0.089
		COLS	0.019	(.003)	0.096	-0.018	(.003)	0.093
		JOLS m=1	-0.027	(.003)	0.091	0.027	(.003)	0.089
		2	-0.018	(.003)	0.090	0.017	(.003)	0.087
		3	-0.012	(.003)	0.089	0.012	(.003)	0.087
		4	-0.008	(.003)	0.090	0.008	(.003)	0.087
		6	-0.003	(.003)	0.091	0.003	(.003)	0.089
		8	-0.001	(.003)	0.093	0.001	(.003)	0.090
		12	0.002	(.003)	0.097	-0.001	(.003)	0.093

TABLE 4A The Monte Carlo estimate of (and the $O(\sigma^2)$ approximation to) the JOLS(1), OLS and COLS bias, the relative efficiency of COLS, and an indicator of the effectivity of COLS in model (5.1) for the fixed x_t series with λ=0.00 (non-trended)

		JOLS(1) bias		OLS bias		COLS		s.d.($\hat{\gamma}$)
		actual	approx	actual	approx	actual bias	rel. eff.	[IEC]
T=24								
σ=0.020	γ=0.0	-0.042	-0.100	-0.042	-0.098	0.056	1.22	0.155
	β=1.0	0.035	0.079	0.035	0.079	-0.041	1.12	[0.64]
σ=0.010	γ=0.0	-0.019	-0.025	-0.019	-0.025	0.005	1.03	0.104
	β=1.0	0.015	0.020	0.016	0.020	-0.004	1.02	[0.27]
σ=0.010	γ=0.0	-0.036	-0.069	-0.035	-0.068	0.032	1.12	0.143
	β=0.6	0.017	0.033	0.018	0.033	-0.014	1.08	[0.53]
σ=0.015	γ=0.4	-0.068	-0.136	-0.065	-0.125	0.058	1.18	0.139
	β=0.6	0.049	0.098	0.047	0.091	-0.038	1.13	[0.50]
σ=0.010	γ=0.4	-0.041	-0.060	-0.039	-0.055	0.015	1.03	0.108
	β=0.6	0.030	0.044	0.028	0.040	-0.010	1.04	[0.30]
σ=0.010	γ=0.8	-0.059	-0.074	-0.056	-0.069	0.013	0.89	0.091
	β=0.3	0.049	0.064	0.047	0.061	-0.010	1.01	[0.18]
σ=0.010	γ=0.8	-0.106	-0.166	-0.101	-0.154	0.045	0.84	0.126
	β=0.2	0.056	0.096	0.054	0.091	-0.027	1.10	[0.32]
T=48								
σ=0.020	γ=0.0	-0.019	-0.041	-0.019	-0.046	0.026	1.08	0.106
	β=1.0	0.021	0.039	0.021	0.041	-0.020	1.05	[0.57]
σ=0.010	γ=0.0	-0.008	-0.010	-0.009	-0.012	0.003	1.01	0.070
	β=1.0	0.009	0.010	0.009	0.010	-0.001	1.01	[0.24]
σ=0.010	γ=0.0	-0.016	-0.029	-0.016	-0.032	0.015	1.05	0.097
	β=0.6	0.010	0.016	0.010	0.017	-0.006	1.03	[0.48]
σ=0.015	γ=0.4	-0.028	-0.055	-0.028	-0.057	0.027	1.07	0.090
	β=0.6	0.024	0.043	0.024	0.044	-0.018	1.05	[0.42]
σ=0.010	γ=0.4	-0.016	-0.025	-0.017	-0.025	0.008	1.01	0.069
	β=0.6	0.014	0.019	0.015	0.019	-0.004	1.01	[0.24]
σ=0.010	γ=0.8	-0.023	-0.032	-0.022	-0.031	0.009	0.93	0.048
	β=0.3	0.019	0.024	0.019	0.024	-0.004	0.97	[0.11]
σ=0.010	γ=0.8	-0.040	-0.071	-0.040	-0.070	0.032	1.00	0.067
	β=0.2	0.022	0.035	0.022	0.035	-0.013	1.02	[0.20]

TABLE 4B The Monte Carlo estimate of (and the $O(\sigma^2)$ approximation to) the JOLS(1), OLS and COLS bias, the relative efficiency of COLS, and an indicator of the effectivity of COLS in model (5.1) for the fixed x_t series with λ=0.02 (trended)

		JOLS(1) bias		OLS bias		COLS		s.d.($\hat{\gamma}$)
		actual	approx	actual	approx	actual bias	rel. eff.	[IEC]
T=24								
σ=0.020	γ=0.0	-0.042	-0.086	-0.041	-0.085	0.038	1.14	0.154
	β=1.0	0.041	0.085	0.040	0.084	-0.038	1.14	[0.61]
σ=0.010	γ=0.0	-0.017	-0.022	-0.017	-0.021	0.003	1.02	0.102
	β=1.0	0.017	0.021	0.017	0.021	-0.003	1.02	[0.25]
σ=0.010	γ=0.0	-0.034	-0.060	-0.034	-0.059	0.022	1.09	0.142
	β=0.6	0.020	0.035	0.020	0.035	-0.013	1.09	[0.51]
σ=0.015	γ=0.4	-0.061	-0.103	-0.058	-0.098	0.031	1.09	0.135
	β=0.6	0.059	0.100	0.056	0.095	-0.030	1.09	[0.46]
σ=0.010	γ=0.4	-0.035	-0.046	-0.033	-0.044	0.008	1.03	0.105
	β=0.6	0.034	0.045	0.032	0.042	-0.008	1.03	[0.27]
σ=0.010	γ=0.8	-0.039	-0.045	-0.037	-0.042	0.002	0.95	0.075
	β=0.3	0.049	0.056	0.047	0.053	-0.003	0.97	[0.13]
σ=0.010	γ=0.8	-0.074	-0.101	-0.072	-0.095	0.009	0.89	0.103
	β=0.2	0.062	0.084	0.059	0.079	-0.011	0.97	[0.25]
T=48								
σ=0.020	γ=0.0	-0.019	-0.040	-0.020	-0.043	0.022	1.07	0.104
	β=1.0	0.019	0.039	0.020	0.042	-0.022	1.07	[0.57]
σ=0.010	γ=0.0	-0.008	-0.010	-0.009	-0.011	0.002	1.01	0.068
	β=1.0	0.008	0.010	0.009	0.011	-0.002	1.01	[0.23]
σ=0.010	γ=0.0	-0.016	-0.027	-0.016	-0.030	0.013	1.04	0.095
	β=0.6	0.010	0.016	0.010	0.018	-0.007	1.04	[0.47]
σ=0.015	γ=0.4	-0.027	-0.048	-0.027	-0.048	0.019	1.05	0.087
	β=0.6	0.027	0.046	0.027	0.047	-0.018	1.05	[0.40]
σ=0.010	γ=0.4	-0.016	-0.021	-0.016	-0.021	0.005	1.01	0.067
	β=0.6	0.015	0.020	0.015	0.021	-0.005	1.01	[0.23]
σ=0.010	γ=0.8	-0.016	-0.020	-0.016	-0.020	0.003	0.99	0.046
	β=0.3	0.022	0.027	0.022	0.026	-0.004	0.99	[0.09]
σ=0.010	γ=0.8	-0.032	-0.045	-0.031	-0.045	0.011	1.00	0.059
	β=0.2	0.028	0.040	0.028	0.040	-0.009	1.00	[0.18]

TABLE 5 The average over 1000 Monte Carlo replications of the actual bias of OLS and COLS, and of the $O(\sigma^2)$ approximation to the OLS bias (and its standard deviation) in model (5.1) with random x_t series; T=24

			mean OLS bias	approx.(s.d.)	mean COLS bias
λ=0.00					
	σ=0.020	γ=0.0	-0.041	-0.095 (0.038)	0.053
		β=1.0	0.031	0.077 (0.035)	-0.041
	σ=0.010	γ=0.0	-0.017	-0.024 (0.009)	0.006
		β=1.0	0.013	0.019 (0.009)	-0.006
	σ=0.015	γ=0.4	-0.058	-0.112 (0.064)	0.055
		β=0.6	0.041	0.078 (0.029)	-0.033
	σ=0.010	γ=0.4	-0.034	-0.050 (0.021)	0.017
		β=0.6	0.023	0.035 (0.013)	-0.011
	σ=0.010	γ=0.8	-0.082	-0.163 (0.151)	0.043
		β=0.2	0.034	0.060 (0.032)	-0.022
λ=0.02					
	σ=0.020	γ=0.0	-0.039	-0.086 (0.039)	0.040
		β=1.0	0.036	0.085 (0.038)	-0.041
	σ=0.010	γ=0.0	-0.015	-0.021 (0.009)	0.006
		β=1.0	0.014	0.021 (0.009)	-0.007
	σ=0.015	γ=0.4	-0.053	-0.089 (0.031)	0.030
		β=0.6	0.049	0.087 (0.030)	-0.031
	σ=0.010	γ=0.4	-0.028	-0.040 (0.014)	0.010
		β=0.6	0.026	0.039 (0.013)	-0.011
	σ=0.010	γ=0.8	-0.054	-0.070 (0.037)	0.009
		β=0.2	0.044	0.058 (0.028)	-0.010