AMSTER

13/89

FACULTY

~~INSTITUTE~~ OF ACTUARIAL SCIENCE & ECONOMETRICS

Report AE 13/89

Bias of s^2 in the Linear Regression Model with correlated errors

Jan F. Kiviet and Walter Krämer

University of Amsterdam

Title: **Bias of s^2 in the Linear Regression Model with correlated errors**

Author: **Jan F. Kiviet*** and **Walter Krämer****

Address: University of Amsterdam
Institute of Actuarial Science and Econometrics
Jodenbreestraat 23
1011 NH Amsterdam
The Netherlands

Date: November 1989

Series and Number: Report AE 13/89

Pages: 13

Price: No charge

JEL Subject Classification: 211

Keywords: autocorrelation, linear regression, disturbance variance, bias

Abstract: We derive bounds for the relative bias of the OLS-based estimate s^2 of the disturbance variance in the linear regression model when disturbances are stationary AR(1) and show that this bias vanishes as sample size increases (i.e. s^2 is asymptotically unbiased irrespective of the particular form of the regressor sequence).

*Professor of Econometrics, Faculty of Economics and Econometrics, University of Amsterdam.

**Professor of Business and Economic Statistics, Dept of Statistics, University of Dortmund.
We are grateful to Deutsche Forschungsgemeinschaft (DFG) for financial support.

BIAS OF s^2 IN THE LINEAR REGRESSION MODEL WITH CORRELATED ERRORS

BY

JAN F. KIVIET AND WALTER KRAMER*

Abstract

We derive bounds for the relative bias of the OLS-based estimate s^2 of the disturbance variance in the linear regression model when disturbances are stationary AR(1) and show that this bias vanishes as sample size increases (i.e. s^2 is asymptotically unbiased irrespective of the particular form of the regressor sequence)

* Professor of Econometrics, Faculty of Economics and Econometrics, University of Amsterdam, and Professor of Business and Economic Statistics, Dep. of Statistics, University of Dortmund. We are grateful to Deutsche Forschungsgemeinschaft (DFG) for financial support.

1. INTRODUCTION

We consider the standard Linear Regression Model

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

where y is Tx1, X is TxK, nonstochastic and of full column rank, β is unknown Kx1, and u is an unknown Tx1 disturbance vector with expectation Eu = 0, whose components have common variance $E(u_t^2) = \sigma^2$. Our topic is the relative bias of the OLS-based estimate

$$s^2 = \frac{1}{T-K} (y-X\hat{\beta})'(y-X\hat{\beta}) \tag{1}$$

for σ^2 , where $\hat{\beta} = (X'X)^{-1}X'y$, when the disturbances have a non-scalar covariance matrix $\sigma^2 V$.

This problem is important because of the resulting distortions of t-statistics and stochastic inferences in general, and has concerned applied economists for quite some time. Cochrane and Orcutt (1949), in an early sampling experiment, find that serial correlation leads to an underestimation of σ^2, whereas Wold (1950) points out that the bias can go the other way as well. Watson (1955) and Sathe and Vinod (1974) derive the (attainable) bounds

$$\text{mean of T-K least roots of V} \le E\left[\frac{s^2}{\sigma^2}\right] \le \text{mean of T-K greatest roots of V} \tag{2}$$

which show that the bias can be both positive and negative, depending on the regressor matrix X, whatever V may be. Finally, Dufour (1985) points out that the inequalities (2) amount to

$$0 \le E\left[\frac{s^2}{\sigma^2}\right] \le \frac{T}{T-K} \tag{3}$$

when no restrictions are placed on X and V. Again these bounds are sharp and demonstrate that the underestimation of σ^2 is much more of a threat in practice than the overestimation (Incidentally, they also hold for nonlinear or stochastic regressor models ; see Dufour, 1988).

The problem with Dufour's bounds is that they are unnecessarily wide when extra information on V is available. Below we follow Sathe and Vinod (1974) and Neudecker (1977, 1978) in assuming that the components of u follow a stationary AR(1) process $u_t = \rho u_{t-1} + \varepsilon_t$, i.e. that V takes the well known form

$$V(\rho) = \begin{bmatrix} 1 & \rho & \cdots & \rho^{T-1} \\ \rho & 1 & \cdots & \rho^{T-2} \\ \rho^2 & \rho & \cdots & \cdot \\ \vdots & & & \\ \rho^{T-1} & \rho^{T-2} & \cdots & 1 \end{bmatrix}, \tag{4}$$

where $-1 < \rho < 1$. We show that Neudecker's bounds, which he derived for positive ρ, hold for negative ρ as well, and that $E(s^2/\sigma^2)$ tends to zero as $\rho \to 1$ for all regressions with an intercept. We also provide sharper bounds when certain columns of X are given (such as seasonal dummies or a linear trend) and demonstrate, contrary to the suggestion in Sathe and Vinod (1974), that the bias tends to zero as $T \to \infty$ for any given ρ and irrespective of the particular regressor sequence $\{x_t\}$.

2.RELATIVE BIAS IN FINITE SAMPLES

From $y - X\hat{\beta} = Mu$, where $M = I - X(X'X)^{-1}X'$, we have

$$E\left(\frac{s^2}{\sigma^2}\right) = E\left(\frac{1}{\sigma^2(T-K)} u'Mu\right) = \frac{1}{T-K} \operatorname{tr} MV . \tag{5}$$

We show first that

$$\lim_{\rho\to 1} E\left[\frac{s^2}{\sigma^2}\right] = 0 \tag{6}$$

whenever there is an intercept in the regression (or more generally, when $Mi = 0$, where $i = (1,1,\ldots,1)'$). To this extent, note that

$$\overline{V} \equiv \lim_{\rho\to 1} V(\rho) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & & & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} = ii' \ , \tag{7}$$

and that

$$\lim_{\rho\to 1} \operatorname{tr}(VM) = \operatorname{tr}(\overline{V}M) = \operatorname{tr}(Mii') \ . \tag{8}$$

So $Mi = 0$ implies (6). In particular, we have $Mi = 0$ whenever the regression contains an intercept, i.e. whenever i appears among the columns of X, or when X contains a full set of seasonal dummy variables.

A similar result holds for $\rho \to -1$. We have

$$\lim_{\rho\to -1} V(\rho) = ee' \tag{9}$$

where $e = (1,-1,\ldots,(-1)^{T-1})'$, so $\lim_{\rho\to -1} E(s^2/\sigma^2) = 0$ whenever e is contained in the column space of X (or more generally, when $Me = 0$). This shows that no nontrivial lower bound to $E(s^2/\sigma^2)$ exists which holds irrespective of X and V, even if V is restricted to be of the form (4).

For intermediate values of ρ, the inequalities (2) translate into

$$\frac{1}{T-K}\sum_{i=1}^{T-K}\lambda_{i+K} \leq E\left[\frac{s^2}{\sigma^2}\right] \leq \frac{1}{T-K}\sum_{i=1}^{T-K}\lambda_i , \qquad (10)$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_T$ are the eigenvalues of $V(\rho)$. These bounds hold for arbitrary but given ρ, where X can be any TxK matrix with full column rank. They can therefore be tabulated as functions of T, K and ρ, (see Neudecker, 1977, 1978) where some space can be saved by observing that the eigenvalues of $V(\rho)$ and $V(-\rho)$ are identical (since $V(-\rho) = \text{diag}(e)V(\rho)\text{diag}(e)$ and $\text{diag}(e)\text{diag}(e) = I$).

Below we present additional tables for the case where $X = [X_1 : X_2]$ and X_1 is a fixed TxK_1 matrix, such as a set of seasonal dummies or a linear trend (the important special case $X_1 = i$ having been treated by Neudecker (1978)).

Let $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$ and let $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_{T-K_1+1} = \ldots = \mu_T = 0$ be the eigenvalues of M_1VM_1 . Along the lines of King (1981, p. 1573), and generalizing Neudecker (1978), it is easily seen that

$$\frac{1}{T-K}\sum_{i=1}^{T-K}\mu_{i+K-K_1} \leq E\left[\frac{s^2}{\sigma^2}\right] \leq \frac{1}{T-K}\sum_{i=1}^{T-K}\mu_i . \qquad (11)$$

Table 1 gives the numerical values for these bounds, for various T, K and ρ, for the case where X_1 is a Tx4 matrix of seasonal dummies. Table 2 gives the analoguous bounds when X_1 is Tx2 comprising a constant and a linear trend. The bounds in both tables are narrower than the ones in Neudecker (1978), where $X_1 = i$, and also narrower than the Watson/Sathe/Vinod bounds, which do not place any restriction on X.

Contrary to the latter case, any bounds which incorporate specific features of the X-matrix are no longer symmetric in ρ. In the tables we focus on the more relevant case of positive autocorrelation. There are different bounds where correlation is negative.

For illustation, consider the regression by Rea (1983, p. 185, eq. 7) of unemployment on money supply, government expenditures and exports (which actually is part of a simultaneous equation system, but we neglect this complication here). This particular example was chosen because the data were available from an earlier paper (Krämer et. al., 1985; see also Krämer and Sonnberger, 1986, p. 144), and because the regression contains a linear trend. The original regressor matrix X has the dimension T = 90 and K = 5, but we confine ourselves to the T = 72 initial observations, which is the maximum allowed by our computer program.

Figure 1 gives the true value of $E(s^2/\sigma^2)$ plus the unrestricted Watson/Sathe/Vinod bounds and the bounds which hold for all 72x5 regressor matrices with an intercept and a linear trend. The figure shows that the unrestricted bounds are symmetric in ρ, whereas the true $E(s^2/\sigma^2)$ is not. Likewise, any bounds which incorporate additional information on X are not symmetric in ρ.

3. ASYMPTOTIC UNBIASEDNESS

Tables 1 and 2 already show that, for given $|\rho| < 1$, the bias becomes less severe as sample size increases. We demonstrate next that the bias vanishes completely as $T \to \infty$, i.e. that in the expression (10),

$$\lim_{T\to\infty} \frac{1}{T-K} \sum_{i=1}^{T-K} \lambda_{i+K} = \lim_{T\to\infty} \frac{1}{T-K} \sum_{i=1}^{T-K} \lambda_i = 1 . \qquad (12)$$

The convergence to one of the upper bound is an easy consequence of Dufour's simple formula (3). It only remains to show that the lower bound tends to one as well. To this extent, we observe (see Neudecker 1977, p. 1258) that the λ_i's can be expressed in the form

$$\lambda_i = \frac{1-\rho^2}{1-2\rho\cos(\Theta_{i,T})+\rho^2} \tag{13}$$

where $\Theta_{i,T}$ is some number which depends on i and on the sample size, but which need not concern us here. The point is that

$$\begin{aligned} 1 - 2\rho\cos(\Theta_{i,T}) + \rho^2 &\leq (1-\rho)^2 \qquad (\rho \geq 0) \\ 1 - 2\rho\cos(\Theta_{i,T}) + \rho^2 &\leq (1+\rho)^2 \qquad (\rho \leq 0) \end{aligned} \tag{14}$$

so

$$\lambda_i \leq \frac{1 + |\rho|}{1 - |\rho|} \tag{15}$$

irrespective of sample size. We note next that

$$\sum_{i=1}^{T} \lambda_i = tr(V) = T \ , \tag{16}$$

so

$$\frac{1}{T-K} \sum_{i=1}^{T-K} \lambda_{i+K} = \frac{T}{T-K} - \frac{1}{T-K} \sum_{i=1}^{K} \lambda_i \geq \frac{T}{T-K} - \frac{K}{T-K} \cdot \frac{1+|\rho|}{1-|\rho|} \tag{17}$$

where the first term tends to one and the second term tends to zero as $T \to \infty$. Therefore, the lower bound to the bias of $E(s^2/\sigma^2)$ tends to one as well, so s^2 is asymptotically unbiased, irrespective of the particular evolution of X.

The seeming paradox here is that (at least for regressions with an intercept), $E(s^2/\sigma^2)$ tends to zero as $\rho \to 1$ for any given X, but tends to one as $T \to \infty$ for any given ρ . The technical explanation for this is that the convergence to zero of $E(s^2/\sigma^2)$ as $\rho \to 1$ is not uniform in T.

4. Conclusion

We have shown that the relative bias of the least squares estimate of the disturbance variance can be quite narrowly bounded when additional information on the regressors and on the disturbance correlation structure is available. Not surprisingly, the bounds get tighter as this prior information increases. In addition, the estimate is asymtotically unbiased for AR(1) disturbances irrespective of ρ and the particular sequence of the regressors.

REFERENCES

COCHRANE, D., and ORCUTT, G.H. (1949): "*Application of least squares regression to relationships containing autocorrelated error terms*", Journal of the American Statistical Association 44, 32-61.

DUFOUR, J.M. (1985): "*Bias of s^2 in Linear Regression with Dependent Errors*", The American Statistician 40, 284-285.

DUFOUR, J.M. (1988): "*Estimators of the Disturbance Variance in Econometric Models*", Journal of Econometrics 37, 277-292.

KING, M.L. (1981): "*The Durbin-Watson Test for Serial Correlation: Bounds for Regressions with Trend and/or Seasonal Dummy Variables*", Econometrica 49, 1571-1581.

KRÄMER, W., SONNBERGER, H., MAURER, J., and HAVLIK, P. (1985): "*Diagnostic Checking in Practice*", The Review of Economics and Statistics 67, 118-123.

KRÄMER, W., and SONNBERGER, H. (1986): *The Linear Regression Model under Test*, Heidelberg (Physica Verlag).

NEUDECKER, H. (1977): "*Bounds for the Bias of the Least Squares Estimator of σ^2 in Case of a First-Order Autoregressive Process* (Positive Autocorrelation)", Econometrica 45, 1257-1262.

NEUDECKER, H. (1978): "*Bounds for the Bias of the LS Estimator in the Case of a First-Order (positive) Autoregressive Process where the Regression contains a Constant Term*", Econometrica 46, 1223-1226.

REA, J.D. (1983): "*The Explanatory Power of Alternative Theories of Inflation and Unemployment, 1895 - 1979*", The Review of Economics and Statistics 65, 183-195.

SATHE, S.T., and VINOD, H.D. (1974): "*Bounds on the Variance of Regression Coefficients due to Heteroscedastic or Autoregressive Errors*", Econometrica 42, 333-340.

WATSON, G.S. (1955): "*Serial Correlation in Regression Analysis I*", Biometrika 42, 327-341.

WOLD, H. (1950): "*On Least Squares Regression with Autocorrelated Variables and Residuals*", Bulletin de l'Institut International de Statistique 32, 277-289.

TABLE 1: Lowerbounds (lwb) and upperbounds (upb) of $E(s^2/\sigma^2)$ for various T, K and ρ, where $K_1=4$ with $X_1=[S_1:S_2:S_3:S_4]$.

	$\rho=0.30$		$\rho=0.60$		$\rho=0.80$		$\rho=0.95$	
	lwb	upb	lwb	upb	lwb	upb	lwb	upb
T=12								
K=5	0.932	1.086	0.719	1.030	0.428	0.771	0.114	0.274
K=6	0.841	1.169	0.552	1.154	0.287	0.878	0.071	0.314
K=7	0.762	1.260	0.450	1.308	0.220	1.018	0.053	0.369
K=8	0.683	1.359	0.365	1.504	0.171	1.208	0.040	0.445
T=16								
K=5	0.946	1.053	0.768	1.005	0.488	0.793	0.138	0.308
K=6	0.880	1.101	0.624	1.079	0.343	0.860	0.087	0.337
K=7	0.820	1.157	0.525	1.167	0.268	0.942	0.065	0.371
K=8	0.762	1.219	0.447	1.270	0.217	1.040	0.052	0.412
T=20								
K=5	0.958	1.039	0.806	0.999	0.541	0.820	0.162	0.346
K=6	0.906	1.073	0.679	1.052	0.393	0.870	0.103	0.368
K=7	0.857	1.112	0.584	1.112	0.311	0.927	0.077	0.394
K=8	0.809	1.155	0.508	1.181	0.256	0.994	0.062	0.425
T=24								
K=5	0.966	1.031	0.835	0.998	0.586	0.843	0.186	0.381
K=6	0.923	1.058	0.722	1.039	0.438	0.883	0.119	0.401
K=7	0.882	1.087	0.633	1.084	0.352	0.928	0.089	0.423
K=8	0.842	1.119	0.559	1.135	0.291	0.978	0.072	0.447
T=36								
K=5	0.978	1.019	0.886	0.996	0.684	0.886	0.252	0.471
K=6	0.950	1.035	0.804	1.021	0.546	0.912	0.164	0.486
K=7	0.923	1.052	0.731	1.047	0.452	0.940	0.124	0.502
K=8	0.896	1.069	0.666	1.075	0.382	0.969	0.099	0.519
T=48								
K=5	0.983	1.014	0.914	0.996	0.747	0.912	0.311	0.542
K=6	0.964	1.025	0.850	1.014	0.624	0.931	0.207	0.555
K=7	0.943	1.037	0.791	1.032	0.531	0.951	0.157	0.567
K=8	0.923	1.049	0.735	1.052	0.458	0.972	0.125	0.581
T=60								
K=5	0.987	1.011	0.931	0.996	0.790	0.928	0.364	0.599
K=6	0.971	1.019	0.879	1.010	0.681	0.943	0.247	0.609
K=7	0.955	1.028	0.830	1.025	0.594	0.959	0.188	0.620
K=8	0.939	1.038	0.783	1.039	0.521	0.975	0.151	0.632
T=72								
K=5	0.989	1.009	0.942	0.997	0.821	0.939	0.411	0.644
K=6	0.976	1.016	0.899	1.008	0.725	0.952	0.285	0.653
K=7	0.963	1.023	0.857	1.020	0.643	0.965	0.219	0.663
K=8	0.950	1.031	0.816	1.032	0.573	0.978	0.176	0.673

TABLE 2: Lowerbounds (lwb) and upperbounds (upb) of $E(s^2/\sigma^2)$ for various T, K and ρ, where $K_1=2$ with $X=[i:t]$.

		$\rho=0.30$		$\rho=0.60$		$\rho=0.80$		$\rho=0.95$	
		lwb	upb	lwb	upb	lwb	upb	lwb	upb
T=12									
	K=3	0.801	0.905	0.510	0.673	0.262	0.391	0.064	0.102
	K=4	0.737	0.948	0.421	0.724	0.202	0.425	0.048	0.112
	K=5	0.688	0.998	0.368	0.786	0.171	0.467	0.040	0.123
	K=6	0.646	1.056	0.329	0.863	0.150	0.520	0.035	0.138
T=16									
	K=3	0.845	0.927	0.583	0.734	0.319	0.460	0.080	0.128
	K=4	0.791	0.958	0.491	0.773	0.246	0.489	0.059	0.136
	K=5	0.746	0.993	0.430	0.819	0.207	0.522	0.049	0.146
	K=6	0.705	1.033	0.384	0.872	0.180	0.562	0.042	0.158
T=20									
	K=3	0.873	0.941	0.639	0.777	0.371	0.518	0.097	0.153
	K=4	0.829	0.965	0.549	0.809	0.289	0.543	0.071	0.161
	K=5	0.788	0.992	0.485	0.846	0.242	0.572	0.058	0.170
	K=6	0.751	1.022	0.434	0.887	0.209	0.604	0.050	0.180
T=24									
	K=3	0.893	0.950	0.684	0.808	0.417	0.566	0.113	0.178
	K=4	0.855	0.970	0.598	0.836	0.328	0.589	0.083	0.185
	K=5	0.820	0.992	0.533	0.866	0.276	0.614	0.067	0.194
	K=6	0.785	1.016	0.479	0.900	0.238	0.642	0.057	0.203
T=36									
	K=3	0.928	0.966	0.772	0.866	0.528	0.671	0.160	0.247
	K=4	0.902	0.979	0.701	0.885	0.430	0.688	0.117	0.254
	K=5	0.877	0.993	0.641	0.905	0.367	0.707	0.094	0.261
	K=6	0.852	1.008	0.588	0.927	0.318	0.726	0.079	0.269
T=48									
	K=3	0.946	0.974	0.823	0.897	0.607	0.736	0.206	0.309
	K=4	0.927	0.984	0.765	0.912	0.512	0.751	0.150	0.316
	K=5	0.907	0.995	0.713	0.927	0.443	0.765	0.121	0.322
	K=6	0.888	1.005	0.665	0.943	0.388	0.781	0.101	0.330
T=60									
	K=3	0.957	0.979	0.856	0.917	0.666	0.781	0.248	0.364
	K=4	0.941	0.987	0.807	0.929	0.576	0.793	0.183	0.371
	K=5	0.926	0.995	0.762	0.941	0.507	0.805	0.147	0.377
	K=6	0.910	1.004	0.720	0.954	0.449	0.818	0.123	0.383
T=72									
	K=3	0.964	0.983	0.878	0.930	0.710	0.813	0.288	0.413
	K=4	0.951	0.989	0.837	0.940	0.627	0.823	0.214	0.419
	K=5	0.938	0.996	0.798	0.950	0.560	0.834	0.173	0.425
	K=6	0.926	1.003	0.760	0.961	0.502	0.845	0.144	0.431

Figure 1: relative bias of s^2