WORKING PAPER SERIES

W.P. No. 156

BIAS IN REGRESSIONS WITH LAGGED STOCHASTIC REGRESSORS

by

Robert F. Stambaugh
Graduate School of Business
University of Chicago

W.P. No. 156

BIAS IN REGRESSIONS WITH LAGGED STOCHASTIC REGRESSORS

by

Robert F. Stambaugh
Graduate School of Business
University of Chicago

Comments Welcome
Not for quotation.

First draft: July 1985
Revised: January 1986

Acknowledgments: I thank David Hsieh, John Huizinga, Al Madansky, Matthew Shapiro, and Arnold Zellner for helpful comments and discussions.

ABSTRACT

This study investigates the bias in a regression with a lagged and stochastic regressor. The regressor obeys a first-order autoregressive process whose innovations are correlated with the regression disturbances. When the latter correlation is positive (negative), the slope coefficient's estimator and t statistic are biased downward (upward). The bias in the OLS slope coefficient is proportional to the bias in the estimated autocorrelation coefficient of the regressor process. Combining this result with an approximate formula for the latter bias gives a new estimator which compares favorably to the OLS estimator in terms of bias and mean square error.

1. Introduction

Researchers in economics and finance often encounter time-series regressions in which the independent variables are predetermined (e.g., lagged) with respect to the dependent variable. This situation occurs, for example, when one tests whether the expected return on an asset is constant by testing whether $\beta = 0$ in the regression

$$y_t = \alpha + \beta x_{t-1} + u_t \quad , \; t = 1, \ldots, T, \tag{1}$$

where y_t is the asset's return and x_{t-1} is a variable from the information set available to investors prior to observing the return. A few examples of such tests include Fama and Schwert (1977), Fama (1984), Huizinga and Mishkin (1985), Campbell(1985), and Keim and Stambaugh (1985).

In many cases x_{t-1}, although predetermined with respect to y_t, is random and possibly correlated with previous regression disturbances. For example, suppose y_t is the holding period return on a long term bond for month t, and suppose x_{t-1} is the bond's yield to maturity observed at the end of month t-1.[1] Previous holding period returns on the bond (y's) will contain information about the most recent yield to maturity (x_{t-1}), e.g., previous negative returns will be associated with a high yield. In essence, the current price level, reflected inversely by the yield, is related to previous price changes. In such a regression, x_{t-1} is correlated with u_{t-s} ($s > 0$), or

$$E(\mathbf{u}|\mathbf{x}) \neq 0 \quad , \tag{2}$$

where $\mathbf{u} = [u_1, \ldots, u_T]'$ and $\mathbf{x} = [x_0, \ldots, x_{T-1}]'$.

If (1) represents the conditional forecast of y_t given x_{t-1}, it is easily

shown that

$$E(u_t x_{t-1}) = 0, \tag{3}$$

which, along with suitable regularity conditions, implies that the ordinary least squares (OLS) estimator of β is consistent. The condition in (2), however, generally leads to finite-sample bias of the OLS estimator of β (and α).[2]

This study investigates problems of finite-sample bias arising in (1) under the following additional specifications:

$$x_t = \mu + \rho x_{t-1} + v_t \quad , \quad t = 1, \ldots, T, \tag{4}$$

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} \sim N(0, \Sigma \otimes I_T) \tag{5}$$

$$\Sigma = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix} \tag{6}$$

where $\mathbf{v} = [v_1, \ldots, v_T]'$. Note that, for $\sigma_{uv} \neq 0$, condition (2) obtains. The specification for x_t in (4) provides a simple way to focus on the effect of serial correlation in the independent variable. For $\rho = 0$, x_{t-1} is correlated only with u_{t-1}. As ρ becomes larger, for a given correlation between u_t and v_t, the correlation between x_{t-1} and u_{t-s} $(s > 1)$ becomes larger. The regression disturbances are assumed to be serially uncorrelated in order to focus on the effect of correlation between $\mathbf{u}$ and $\mathbf{x}$.

Finite sample bias of the least squares estimator has been analyzed extensively for the first-order autoregressive process. [See, for example,

Kendall (1954), White (1961), Orcutt and Winokur (1969), Sawa (1978), Fuller (1976), Dickey and Fuller (1979, 1981), Evans and Savin (1984).] The above specification is a natural extension of that literature. In fact, the bias of $\hat{\beta}$, the OLS estimator of β, is proportional to the bias of $\hat{\rho}$, the OLS estimator of ρ in (4). Thus, previous analytical and Monte Carlo results for the bias of $\hat{\rho}$ can be extended directly to the above model.

This study combines analytical methods with Monte Carlo experiments to investigate the bias of $\hat{\beta}$, the behavior of the t statistic that tests the hypothesis $\beta = \beta_0$, and the properties of alternative estimators designed to reduce the bias. The investigation considers a range of parameter values, including various levels of correlation between the regression disturbance (u_t) and the innovation in the regressor (v_t). This correlation can vary substantially, depending on the particular application. In the example above, where a bond's return is regressed on the bond's own yield, the correlation is likely to be quite high. If the same bond yield is used to predict, say, a common stock return, then the correlation between u_t and v_t is probably lower. It is shown here that the bias of both $\hat{\beta}$ and the t statistic is increasing in this correlation.

The paper proceeds as follows. Section 2 analyzes the bias of $\hat{\beta}$ by extending previous analytical results for the first-order autoregressive process. Section 3 investigates the finite-sample distribution of the t statistic under the null hypothesis. Section 4 investigates the behavior of several estimation techniques designed to reduce small-sample bias. The first technique uses an analytical approximation for the bias of $\hat{\beta}$; jackknife techniques are also investigated. Section 5 concludes the paper.

2. The Bias of the OLS Estimator

Given the process for x_t assumed above, obtaining the bias of $\hat{\beta}$ is straightforward. Let $\hat{\rho}$ denote the OLS estimator of ρ in (4).

Theorem. $$E(\hat{\beta} - \beta) = \frac{\sigma_{uv}}{\sigma_v^2} E(\hat{\rho} - \rho) \quad . \tag{8}$$

Proof: Let $\mathbf{X} = [\iota\ \mathbf{x}]$, $\iota = [1, \ldots, 1]'$, $\mathbf{x}^+ = [x_1, \ldots, x_T]'$, $\delta = [\alpha\ \beta]'$, and $\phi = [\mu\ \rho]'$. The error in the OLS estimator of δ is given by $\hat{\delta} - \delta = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}$, and $\mathbf{u}$ can be written as $\mathbf{u} = (\sigma_{uv}/\sigma_v^2)\mathbf{v} + \varepsilon$, where $E(\varepsilon|\mathbf{x}) = \mathbf{0}$. Thus, we can write the error in $\hat{\delta}$ as $\hat{\delta} - \delta = (\sigma_{uv}/\sigma_v^2)(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{v} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon$. The expectation of the second term is zero. The first term equals $(\sigma_{uv}/\sigma_v^2)[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{x}^+ - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\phi)] = (\sigma_{uv}/\sigma_v^2)(\hat{\phi} - \phi)$, which gives the desired result.

The usefulness of (8) is that the bias of $\hat{\rho}$ has been analyzed in previous studies. Sawa (1978) gives a method for obtaining exact first and second moments of $\hat{\rho}$, and other studies have performed Monte Carlo experiments [e.g., Orcutt and Winokur (1969) and Fuller (1976)]. In general, $\hat{\rho}$ is downward biased in finite samples. From (8), $\hat{\beta}$ is downward (upward) biased if the correlation between u_t and v_t is positive (negative). As the (absolute) correlation between u_t and v_t increases, holding the ratio of their variances constant, the (absolute) bias increases. The proof above also gives the result that $E(\hat{\alpha} - \alpha) = (\sigma_{uv}/\sigma_v^2) \cdot E(\hat{\mu} - \mu)$. Although the focus here is primarily on $\hat{\beta}$, the latter result can be coupled with the analysis of previous studies that investigate the bias in the intercept [e.g., Orcutt and Winokur (1969)].

Table 1 displays values of $T \cdot E(\hat{\beta} - \beta)$ for the special case $\sigma_{uv} = \sigma_v^2$. The values for the smaller samples($T = 10, 15, 25$) are obtained using the analytical result of Sawa (1978), while the values for the larger samples ($T = 50, 100, 250, 500$) are obtained by Monte Carlo experiments based on 1000 replications for each case.[3] The analysis here is confined here to values of ρ less than unity, for which x_t is stationary. The initial value x_0 is generated randomly from the unconditional distribution. As table 1 confirms, the bias is fairly well approximated as being proportional to 1/T. It is known that

$$E(\hat{\rho}) - \rho = -\frac{1 + 3\rho}{T} + O(\frac{1}{T^2}) \tag{9}$$

[Marriott and Pope (1954) and Kendall (1954)], and table 1 indicates that the higher order terms are indeed small. Such behavior suggests that a simple jackknife estimator might eliminate most of the bias in $\hat{\beta}$, and section 4 investigates this possibility.

3. The Behavior of the t Statistic

In many studies, the central issue surrounding (1) is whether $\beta = \beta_0$. (β_0 is often zero.) This section investigates the behavior of the usual t statistic under the null hypothesis.[4] Table 2 displays estimates, based on Monte Carlo experiments of 1000 replications each, of the 2.5% fractile of the t statistic.[5] Results are shown for different values of the autocorrelation coefficient ρ and the correlation between u_t and v_t, denoted as γ. (It is easily shown that the t statistic's distribution is invariant to α, β, σ_u and σ_v.)

When either ρ or γ are large (close to unity), table 2 reveals that the t

Table 1

Values of $T \cdot E(\hat{\beta} - \beta)$ when $\sigma_{uv} = \sigma_v^2$

(For other cases, multiply value shown by σ_{uv}/σ_v^2.)

	ρ				
T	.99	.95	.80	.40	.00
			Analytical Values		
10	-4.42	-4.23	-3.62	-2.26	-1.11
15	-4.67	-4.44	-3.70	-2.26	-1.07
25	-4.88	-4.56	-3.70	-2.25	-1.04
			Monte Carlo Estimates		
50	-4.86	-4.37	-3.25	-2.21	-1.05
100	-4.98	-4.36	-3.42	-2.19	-1.03
250	-4.61	-3.94	-3.41	-1.60	-0.66
500	-4.88	-3.93	-4.13	-2.12	-0.73

statistic's distribution in samples of even 100 or more can differ substantially from the asymptotic (Normal) distribution. (The 2.5% fractile of the Normal distribution is -1.96.) In general, when γ is positive, the t statistic is downward biased and skewed slightly to the left. When both ρ and γ are high, the distributions in samples of 25 and 50 are similar to the random-walk case documented in Fuller (1976, table 8.5.2).

Table 2 reports results only for cases where γ, the correlation between u_t and v_t, is positive. When γ is negative, the density function reverses. That is, the negatives of the values in table 2 can be viewed as estimates of the 0.975 fractile. For the example mentioned in the introduction, where x_{t-1} is a bond's yield and y_t is the subsequent holding period return, γ is most likely negative--innovations in the yield are negatively correlated with returns. In that case $\hat{\beta}$ is biased upward, and to reject $\beta = 0$ against the (one-tailed) alternative $\beta > 0$ at the 0.025 significance level would require that the t statistic exceed a critical value greater then the usual value of 1.96.

4. Bias-Corrected Estimators

This section investigates the behavior of several alternative estimators designed to reduce the bias in the OLS estimator, $\hat{\beta}$. The first of these alternative estimators is based essentially on a combination of (8) and (9). This estimator is similar in spirit to an estimator for ρ considered by Orcutt and Winokur (1969). They replace $E(\hat{\rho})$ in (9) with $\hat{\rho}$, solve for ρ, and then use that solution as a new estimator. The approach taken here is to combine (8) and (9) to obtain

$$E(\hat{\beta} - \beta) = - \left(\frac{\sigma_{uv}}{\sigma_v^2}\right) \cdot \left(\frac{1 + 3\rho}{T}\right) + O\left(\frac{1}{T^2}\right) \qquad (10)$$

Table 2

Estimates of the 0.025 Fractile of the Sample t Statistic, $t(\hat{\beta} - \beta)$

Parameters		0.025 fractile of $t(\hat{\beta} - \beta)$ for sample of				
ρ	γ	25	50	100	250	500
0.99	0.99	-3.254	-3.164	-2.843	-2.648	-2.712
0.99	0.95	-3.252	-3.044	-2.986	-2.824	-2.678
0.99	0.80	-3.186	-2.998	-2.992	-2.622	-2.659
0.99	0.40	-2.701	-2.485	-2.544	-2.395	-2.225
0.99	0.00	-2.075	-2.09-	-2.159	-1.939	-2.039
0.95	0.99	-2.977	-2.942	-2.897	-2.392	-2.248
0.95	0.95	-3.054	-2.861	-2.699	-2.389	-2.367
0.95	0.80	-2.881	-2.830	-2.581	-2.393	-2.159
0.95	0.40	-2.637	-2.369	-2.349	-2.157	-2.121
0.95	0.00	-2.128	-2.030	-2.050	-2.048	-1.988
0.80	0.99	-2.715	-2.420	-2.423	-2.300	-2.205
0.80	0.95	-2.695	-2.495	-2.457	-2.139	-2.181
0.80	0.80	-2.582	-2.478	-2.367	-2.211	-1.939
0.80	0.40	-2.472	-2.220	-2.076	-1.957	-1.887
0.80	0.00	-2.097	-2.050	-1.979	-1.961	-1.950
0.49	0.99	-2.432	-2.220	-2.212	-2.145	-2.016
0.40	0.95	-2.284	-2.358	-2.081	-2.095	-1.883
0.40	0.80	-2.270	-2.192	-2.139	-2.181	-2.004
0.40	0.40	-2.339	-2.058	-2.116	-1.923	-2.027
0.40	0.00	-1.963	-2.076	-2.062	-2.051	-1.987
0.00	0.99	-1.996	-2.067	-2.024	-1.932	-2.047
0.00	0.95	-2.223	-2.150	-2.194	-1.869	-2.024
0.00	0.80	-2.051	-2.144	-1.978	-1.938	-1.849
0.00	0.40	-2.115	-1.890	-2.026	-1.989	-1.899
0.00	0.00	-1.951	-1.949	-1.950	-1.940	-1.977

and then to replace the quantities on the right-hand side with sample estimates. Thus, the new "bias-adjusted" estimator, $\hat{\beta}_A$, is given by

$$\hat{\beta}_A = \hat{\beta} + \left(\frac{s_{uv}}{s_v^2}\right)\cdot\left(\frac{1 + 3\hat{\rho}}{T}\right) \tag{11}$$

where s_{uv} and s_v^2 are estimated using residuals from the OLS regressions for (1) and (4). It should be noted here that, unlike the estimator considered by Orcutt and Winokur, $\hat{\beta}_A$ does not necessarily eliminate bias of O(1/T).[6] The addition of the estimates s_{uv} and s_v^2 complicates the analysis of the expected value. Nevertheless, given the favorable results of Orcutt and Winokur for their estimator based on (9), it seems reasonable to investigate a similar approach here.

Two additional estimators, both based on the grouped jackknife approach of Quenouille (1949), are also investigated. Such estimators eliminate bias of O(1/T), and recall from table 1 that the bias in $\hat{\beta}$ is well characterized as proportional to 1/T. Let M = T/2 [or (T-1)/2 if T is odd], and consider OLS estimators of β for subperiods of size M having contiguous observations. There are T-M+1 such subperiods. Let $\hat{\beta}_{(M),2}$ denote the average of the estimators for the first and last subperiods, and let $\hat{\beta}_{(M),ALL}$ denote the average of the estimators over all T-M+1 subperiods. The jackknife estimators are defined as

$$\hat{\beta}_{J1} = \left(\frac{T}{T-M}\right)\hat{\beta} - \left(\frac{M}{T-M}\right)\hat{\beta}_{(M),2} \qquad \text{and} \tag{12}$$

$$\hat{\beta}_{J2} = \left(\frac{T}{T-M}\right)\hat{\beta} - \left(\frac{M}{T-M}\right)\hat{\beta}_{(M),ALL} \quad . \tag{13}$$

The first jackknife estimator, which uses two subperiods, is also investigated as an estimator of ρ by Orcutt and Winokur (1969). The second estimator, which uses all of the subperiods, eliminates bias of O(1/T) in precisely the same manner as the first, but it may have less variance [Efron (1982, p. 7)].

While the above alternative estimators may be less biased than the OLS estimator, they may also possess more variance. Indeed, previous investigations of jackknife estimators have found those estimators to possess greater variance than OLS estimators in a number of different applications [e.g., Orcutt and Winokur (1969) and Huizinga (1983)]. Table 3 presents both the biases and the mean square errors of the OLS estimator and the three alternative estimators. Each row of the table is based on a set of 1000 Monte Carlo replications.

In virtually all cases, each of the three bias-corrected estimators contains less bias than the OLS estimator. In many cases the bias of the OLS estimator is five of six times greater than the bias of the alternative estimators. The biases among the three alternative estimators are fairly similar, although $\hat{\beta}_A$ tends to have less bias than the jackknife estimators in cases where both ρ and γ are high (0.95 or 0.80) and sample size is small (25 or less).

The comparisons of mean square error (variance plus the square of the bias) are less favorable to the jackknife estimators. In most cases the jackknife estimators possess substantially higher mean square error (MSE) than the OLS estimator. The two jackknife estimators themselves have similar MSE's, although $\hat{\beta}_{J2}$ (which uses all T-M+1 subperiods) is typically superior to $\hat{\beta}_{J1}$ (which uses only two subperiods) in smaller samples and when ρ is high.

The bias-adjusted estimator $\hat{\beta}_A$ generally compares favorably to the OLS estimator, $\hat{\beta}$. When both ρ and γ are high (0.80 or more), the MSE of $\hat{\beta}$ exceeds

Table 3

Monte Carlo Comparison of OLS and Bias-Corrected Estimators
(β = 0.0, σ_v = 0.10, σ_u = 0.10)

T	Bias (x 100)				Mean Square Error (x 100)			
	$\hat{\beta}$	$\hat{\beta}_A$	$\hat{\beta}_{J1}$	$\hat{\beta}_{J2}$	$\hat{\beta}$	$\hat{\beta}_A$	$\hat{\beta}_{J1}$	$\hat{\beta}_{J2}$
				ρ = 0.95, γ = 0.95				
10	-36.795	-8.608	-13.367	-13.280	23.794	18.474	28.768	22.089
15	-27.042	-6.860	-7.136	-8.163	12.950	8.629	13.200	10.621
25	-16.053	-2.814	-1.670	-1.845	4.702	2.733	5.227	3.920
50	-8.439	-1.483	-0.322	-0.356	1.488	0.893	1.648	1.471
100	-4.207	-0.644	0.128	0.084	0.414	0.254	0.476	0.411
250	-1.459	-0.007	0.218	0.211	0.080	0.060	0.093	0.105
500	-0.764	-0.036	0.020	0.094	0.033	0.027	0.034	0.041
				ρ = 0.95, γ = 0.80				
10	-30.009	-6.175	-8.771	-8.463	19.001	15.585	29.381	22.691
15	-22.793	-5.697	-7.319	-7.263	11.200	8.290	13.657	10.923
25	-14.160	-3.138	-3.093	-2.563	4.295	2.804	4.993	4.047
50	-7.140	-1.309	-0.182	0.001	1.236	0.816	1.449	1.288
100	-3.599	-0.598	0.164	0.055	0.358	0.243	0.441	0.408
250	-1.398	-0.177	-0.054	0.001	0.080	0.062	0.088	0.098
500	-0.577	0.037	0.140	0.139	0.026	0.023	0.030	0.036
				ρ = 0.95, γ = 0.40				
10	-14.180	-2.081	-2.414	-2.649	13.635	14.771	38.154	26.246
15	-11.375	-3.042	-3.596	-3.328	6.446	5.937	11.198	9.255
25	-7.162	-1.709	-0.925	-1.133	2.429	2.161	3.858	3.523
50	-3.570	-0.673	-0.490	-0.305	0.735	0.637	1.096	1.056
100	-1.845	-0.329	0.132	0.134	0.234	0.204	0.336	0.353
250	-0.598	0.009	0.216	0.094	0.057	0.054	0.075	0.091
500	-0.229	-0.078	0.093	0.118	0.023	0.023	0.026	0.035
				ρ = 0.80, γ = 0.95				
10	-32.916	-7.871	-12.095	-10.457	20.148	16.618	26.621	20.934
15	-22.537	-4.300	-4.624	-5.290	11.004	8.671	13.231	11.119
25	-13.202	-1.455	-1.203	-0.835	4.495	3.449	5.065	4.653
50	-6.933	-0.765	-0.596	-0.538	1.601	1.254	1.817	2.057
100	-3.361	-0.191	0.073	-0.168	0.555	0.468	0.564	0.662
250	-1.219	0.062	0.092	0.088	0.162	0.150	0.173	0.211
500	-0.722	-0.080	-0.700	-0.109	0.082	0.078	0.081	0.102
				ρ = 0.80, γ = 0.80				
10	-26.403	-4.777	-8.687	-7.725	18.281	17.159	28.113	23.150
15	-19.784	-4.571	-4.489	-5.139	9.726	7.891	12.511	10.461
25	-11.930	-2.027	-1.526	-1.581	4.015	3.109	5.121	4.569
50	-5.625	-0.421	0.236	-0.084	1.442	1.230	1.748	1.822
100	-2.522	0.137	0.085	0.081	0.520	0.477	0.577	0.663
250	-1.207	-0.124	-0.089	0.001	0.181	0.169	0.190	0.228
500	-0.562	-0.019	0.011	0.022	0.018	0.079	0.081	0.106
				ρ = 0.80, γ = 0.40				
10	-12.974	-2.447	-3.417	-2.950	12.038	12.569	35.879	26.969
15	-9.731	-1.981	-2.506	-1.478	7.042	6.847	12.392	10.251
25	-6.290	-1.230	-0.890	-1.075	3.092	2.904	4.592	4.590
50	-2.175	0.448	0.832	0.403	1.062	1.044	1.470	1.548
100	-1.237	0.093	0.114	0.333	0.454	0.444	0.542	0.665
250	-0.623	-0.086	-0.046	-0.033	0.156	0.153	0.166	0.216
500	-0.204	0.066	0.079	0.018	0.073	0.073	0.078	0.098
				ρ = 0.40, γ = 0.95				
10	-20.498	-3.985	-5.930	-5.672	15.153	18.206	22.967	20.559
15	-14.698	-2.874	-3.072	-2.988	8.905	9.681	11.926	10.915
25	-7.879	-0.193	-0.152	-0.116	4.194	4.418	5.478	5.186
50	-4.220	-0.205	-0.091	-0.098	1.964	1.986	2.284	2.571
100	-2.253	-0.216	-0.165	-0.151	0.940	0.939	1.004	1.219
250	-0.459	0.374	0.332	0.522	0.344	0.350	0.354	0.472
500	-0.189	0.230	0.241	0.186	0.152	0.154	0.157	0.208
				ρ = 0.40, γ = 0.80				
10	-17.257	-3.232	-4.811	-3.988	13.737	15.487	28.494	19.873
15	-10.594	-0.318	-0.848	-0.260	7.778	8.530	10.992	10.414
25	-7.739	-1.291	-1.432	-1.635	4.045	4.047	5.159	5.258
50	-3.119	0.284	0.484	0.694	1.854	1.904	2.161	2.633
100	-1.960	-0.239	-0.213	-0.167	0.883	0.878	0.943	1.222
250	-0.555	0.143	0.126	0.341	0.350	0.353	0.363	0.461
500	-0.458	-0.108	-0.104	-0.093	0.168	0.167	0.173	0.234
				ρ = 0.40, γ = 0.40				
10	-7.387	0.102	0.122	0.625	12.540	13.306	24.155	21.016
15	-4.972	0.078	0.914	0.152	8.120	8.551	12.201	12.180
25	-3.150	0.055	-0.220	0.111	4.097	4.206	4.894	5.444
50	-1.842	-0.145	-0.213	-0.115	1.799	1.801	2.049	2.509
100	-0.762	0.105	0.107	0.307	0.902	0.905	0.941	1.136
250	-0.386	-0.036	-0.015	-0.417	0.367	0.367	0.379	0.480
500	-0.232	-0.057	-0.062	-0.159	0.167	0.166	0.170	0.225

that of $\hat{\beta}_A$, often by as much as 25 percent. When ρ and γ are lower, then the MSE of $\hat{\beta}$ is typically lower than that of $\hat{\beta}_A$, but not by substantial amounts. These results are similar to those reported by Orcutt and Winokur in that they find, for high values of ρ, the estimator based on (9) is superior to the OLS estimator. (They also find the jackknife estimator to have the highest mean square error.) The superiority of $\hat{\beta}_A$ in this study is interesting in that, even though additional parameters are estimated [cf. (11)], the variance of the estimator is evidently not increased enough to produce a MSE higher than that of the OLS estimator.

5. Conclusions

This study has investigated problems arising from bias in a regression where a lagged but stochastic regressor obeys a first-order process whose innovations are correlated with the regression disturbances. When the latter correlation is positive, the OLS estimator of the regression slope coefficient and the coefficient's t statistic are biased downward. (When the correlation is negative, both quantities are biased upward.) Jackknife estimators prove to be effective in reducing the bias, but they possess substantially higher mean square error than the OLS estimator. The bias of the OLS slope coefficient is proportional to the bias of the OLS estimator for the autocorrelation coefficient of the regressor process. When this result is combined with an approximate formula for the bias of the estimated autocorrelation, a new estimator is obtained. This new bias-adjusted estimator compares favorably to the OLS estimator in terms of both bias and mean square error.

FOOTNOTES

[1]Examples of studies in which holding-period returns on assets are regressed on bond yields (or yield spreads) include Campbell (1985) and Keim and Stambaugh (1985).

[2]The failure of the orthogonality of **u** and **x** also leads to inconsistency of generalized least squares estimators that attempt to account for possible serial correlation in the u_t's. The latter large-sample problem has been analyzed by a number of studies, including Hansen and Hodrick (1980), Hansen (1982), Hansen and Sargent (1982), Cumby, Huizinga, and Obstfeld (1983), and Hayashi and Sims (1983).

[3]Some of the values shown in table 1 differ slightly from those in Sawa (1978, table 2a). The values shown here were computed in double precision using IMSL routines EIGRS and DCADRE on a Dec 20. Most of the values are also slightly closer than Sawa's to the Monte Carlo results reported by Orcutt and Winokur (1969).

[4]Independent work by Mankiw and Shapiro (1985) also reports results of Monte Carlo experiments that investigate the behavior of the t statistic in the same framework.

[5]An appendix containing estimated fractiles for both tails of the empirical distribution is available from the author on request.

[6]OLS-based estimators of σ_{uv}, σ_v^2, and ρ are used in (11) primarily for simplicity and ease of computation. The estimator of ρ could instead be that of Orcutt and Winokur (1969), and the estimator of σ_{uv} could be improved iteratively. Neither of these changes would necessarily eliminate bias of O(1/T), however.

REFERENCES

Campbell, J.Y., 1985, "Stock Returns and the Term Structure," working paper, Princeton University.

Cumby, R.E., J. Huizinga, and M. Obstfeld, 1983, "Two-Step Two-Stage Least Squares Estimation in Models with Rational Expectations," Journal of Econometrics 21, 333-355.

Dickey, D.A. and W.A. Fuller, 1979, "Distribution of the EStimators for Autoregressive Time Series with a Unit Root," Journal of the American Statistical Association 74, 427-431.

Dickey, D.A. and W.A. Fuller, 1981, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," Econometrica 49, 1057-1072.

Efron, B., 1982, The Jackknife, the Bootstrap and Other Resampling Plans (Society for Industrial and Applied Mathematics, Philadelphia).

Evans, G.B.A. and N.E. Savin, 1984, "Testing for Unit Roots: 2," Econometrica 52, 1241-1269.

Fuller, W.A., 1976, Introduction to Statistical Time Series (John Wiley & Sons, New York).

Hansen, L.P., 1982, "Large Sample Properties of Generalized Methods of Moments Estimators," Econometrica 50, 1029-1054.

Hansen, L.P. and R.J. Hodrick, 1980, "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Empirical Analysis," Journal of Political Economy 88, 829-853.

Hansen, L.P. and T. Sargent, 1982, "Instrumental Variables Procedures for Estimating Linear Rational Expectations Models," Journal of Monetary Economics 9, 263-296.

Hayashi, F. and C. Sims, 1983, "Nearly Efficient Estimation of Time Series Models with Predetermined, but not Exogenous, Instruments," Econometrica 51, 783-798.

Huizinga, J., 1983, "Small Sample Properties of Market Efficiency Tests in Models with Overlapping Data," working paper, University of Chicago.

Keim, D.B. and R.F. Stambaugh, 1985, "Predicting Returns in the Stock and Bond Markets," working paper no. 143, Center for Research in Security Prices, University of Chicago.

Kendall, M.G., 1954, "Note on Bias in the Estimation of Autocorrelation," Biometrika 41, 403-404.

Mankiw, N.G. and M.D. Shapiro, 1985, "Do We Reject Too Often? Small Sample Properties of Tests of Rational Expectations Models," working paper, Harvard University and Yale University.

Marriott, F.H.C. and J.A. Pope, 1954, "Bias in the Estimation of Autocorrelations," Biometrika 41, 390-402.

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

Quenouille, M., 1949, "Approximate Tests of Correlation in Time Series," Journal of the Royal Statistical Society, series B, 11, 18-84.

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

White, J.S., 1961, "Asymptotic Expansions for the Mean and Variance of the Serial Correlation Coefficient," Biometrika 48, 85-94.