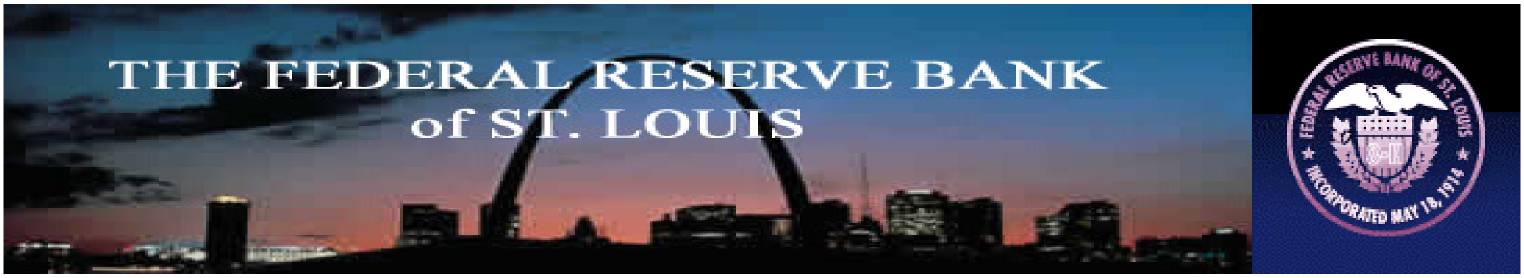

WORKING PAPER SERIES

On the Treatment of the Weighted Initial Observation in the AR(1) Regression Model

Daniel L. Thornton

Working Paper 1984-003A
http://research.stlouisfed.org/wp/1984/1984-003.pdf

1984

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
411 Locust Street
St. Louis, MO 63102

ON THE TREATMENT OF THE WEIGHTED INITIAL OBSERVATION IN THE AR(1) REGRESSION MODEL

Daniel L. Thornton
Federal Reserve Bank of St. Louis

84-003

I would like to thank Thomas Fomby, Carter Hill and Peter Schmidt for helpful comments on an earlier draft. The views expressed here are not necessarily those of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

8048R
0139R
8th draft

On the Treatment of the Weighted Initial Observation in the AR(1) Regression Model

DANIEL L. THORNTON

A number of studies have investigated the treatment of the weighted initial observation in the estimation of the AR(1) regression model, [e.g., Kadiyala (1968), Maeshiro (1976, 1979), Chipman (1979), Doran (1979), Doran and Griffiths (1983), Spitzer (1979) and Fomby and Guilkey (1983)]. It is generally conceded that the Cochrane-Orcutt (C-O) estimator, which deletes the initial observation, is inefficient relative to the Prais-Winsten (P-W) estimator, which weights it by $(1-\rho^2)^{1/2}$. Nevertheless, it is usually suggested that the C-O estimator might be preferable in cases where the past is not sufficiently long to warrant the use of the P-W estimator, [e.g., Judge, et. al, (1980) p. 182].[1/]

The purpose of this paper is to present a general AR(1) model for the case where the past is finite, to present the efficiencies of the P-W and C-O estimators for this model and to demonstrate that the P-W estimator is always more efficient than the C-O estimator when the AR(1) process has a finite past.

2. A General AR(1) Model

Consider the model

(1) $Y = X\beta + \varepsilon$,

where Y is a nx1 vector of observations on the dependent variable, X is a nxk matrix of known nonstochastic regressors, β is a kx1 vector of unknown coefficients and ε is a nx1 vector of random disturbances. Assume that

$$(2)\quad \varepsilon_t = \begin{cases} \rho\varepsilon_{t-1} + u_t & t > -q \\ u_t & t \leq -q \end{cases}$$

where u_t is nid(0, σ_u^2).2/ The parameter q is assumed known and is continuous for $q \geq 1$; q-1 denotes the number of periods since the process began.3/

It is easy to show that

$$E(\varepsilon\varepsilon') = \frac{\sigma_u^2}{1-\rho^2}\,\phi,$$

where

$$\phi^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} \mu & -\rho & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}$$

and where $\mu = (1-\rho^{2(q+1)})/(1-\rho^{2q})$.

Furthermore, there is a transformation matrix

$$C = \begin{bmatrix} \delta & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}$$

where $\delta = [(1-\rho^2)/(1-\rho^{2q})]^{1/2}$, such that $C'C = (1-\rho^2)\phi^{-1}_{.}$ Thus, equation (1) can be transformed to the classical regression model by premultiplying (1) by C.

If the customary stationarity condition is assumed (i.e., $|\rho| < 1$), then the usual covariance matrix, A, is

$$\lim_{q\to\infty} \phi = A,$$

and the usual P-W transformation matrix, M, is

$$\lim_{q\to\infty} C = M.$$ 4/

The C-O transformation matrix, Q, is C with the first row deleted.

3. The Relative Efficiencies of the C-O and P-W estimators

It is easy to show that the efficiency of the P-W estimator relative to the general Aitken's estimator is less than the efficiency of the C-O estimator relative to the Aitken's estimator. To see this, let $\tilde{\beta}$, $\hat{\beta}$ and $\hat{\hat{\beta}}$ denote the general Aitken's, P-W and C-O estimators of β, respectively. The efficiencies of $\hat{\beta}$ and $\hat{\hat{\beta}}$ with respect to $\tilde{\beta}$ are

$$E_{\hat{\beta}} = \frac{|\Sigma_{\tilde{\beta}}|}{|\Sigma_{\hat{\beta}}|} ,$$

and

$$E_{\hat{\hat{\beta}}} = \frac{|\Sigma_{\tilde{\beta}}|}{|\Sigma_{\hat{\hat{\beta}}}|} ,$$

where $\Sigma_{\tilde{\beta}}$, $\Sigma_{\hat{\beta}}$ and $\Sigma_{\hat{\hat{\beta}}}$ are their respective covariance matrices. But

$$\Sigma_{\tilde{\beta}} = \frac{\sigma_u^2}{1-\rho^2} (X'\phi^{-1}X)^{-1} ,$$

$$\Sigma_{\hat{\beta}} = \frac{\sigma_u^2}{1-\rho^2}(X'M'MX)^{-1}X'M'M\phi M'MX(X'M'MX)^{-1}$$

and

$$\Sigma_{\hat{\hat{\beta}}} = \frac{\sigma_u^2}{1-\rho^2} (X'Q'QX)^{-1}X'Q'Q\phi Q'QX(X'Q'QX)^{-1}.$$

By straightforward multiplication,

$$M'M = (1-\rho^2) A^{-1},$$

$$Q'Q = (1-\rho^2) (A^{-1}-\theta),$$

$$M\phi M' = (1-\rho^2)(I-\zeta),$$

and

$$Q\phi Q' = (1-\rho^2) I,$$

where θ and ζ are nxn matrices whose (1,1) elements are 1 and ρ^{2q}, respectively, and all other elements are zeros. Given these results, the efficiency ratios can be rewritten as

$$E_{\hat{\beta}} = \frac{|X'(A^{-1})X|^2}{|X'\phi^{-1}X|\,|X'(A^{-1}-\zeta)X|} ,$$

and

$$E\hat{\hat{\beta}} = \frac{|X'(A^{-1}-\theta)X|}{|X'\phi^{-1}X|} .$$

From a standard theorem in matrix algebra,

$|X'A^{-1}X| > |X'(A^{-1}-\zeta)X|$ or $|X'(A^{-1}-\theta)X|$, so that

$E\hat{\beta} > E\hat{\hat{\beta}}$ .5/

This intuition for this result is quite simple. Maeshiro (1980) has noted that the net gain in efficiency results from the gain in efficiency due to the reduction or elimination of autocorrelation and to the loss of efficiency due to the increased colinearity among the transformed variables.6/ In this instance the P-W estimator induces less colinearity in the transformed variables, enough less to be more efficient than the C-O estimator for any given set of fixed regressor variables.

Since both of these estimators will be most inefficient for smoothly trended data, especially for large positive values of ρ, the results for the simple time-trend model

$$y_t = \alpha + \beta t + \varepsilon_t$$

are presented. Following Chipman, let $X = \tau - \frac{n+1}{2} \ell$ where $\tau = (1\ 2\ 3\ \ldots\ n)'$; therefore,

$$X'A^{-1}X = \frac{(n-1)\ g\ (\rho,\ n)}{12\ (1-\rho^2)},$$

where $g(\rho,n) = (n-3)(n-2)\rho^2 - 2(n-3)(n+1)\rho + n(n+1)$.

By direct multiplication,

$$X'\phi^{-1}X = X'A^{-1}X + \left(\frac{n-1}{2}\right)^2 \rho^{2q}/(1-\rho^{2q}),$$

$$X' (A^{-1} - \theta) X = X'A^{-1}X - (\frac{n-1}{2})^2$$

and

$$X' (A^{-1} - \zeta) X = X'A^{-1}X - (\frac{n-1}{2})^2 \rho^{2q} ,$$

Thus, for the time-trend model

$$E_{\hat{\beta}} = \frac{[\frac{(n-1)\ q\ (\rho, n)}{12\ (1-\rho^2)}]^2}{[\frac{(n-1)q(\rho,n)}{12\ (1-\rho^2)} + (\frac{n-1}{2})^2 (\frac{\rho^{2q}}{1-\rho^{2q}})][\frac{(n-1)q(\rho,n)}{12\ (1-\rho^2)} - (\frac{n-1}{2})^2 \rho^{2q}]} .$$

Likewise,

$$E_{\hat{\hat{\beta}}} = \frac{\frac{(n-1)g(\rho,n)}{12\ (1-\rho^2)} - (\frac{n-1}{2})^2}{\frac{(n-1)g(\rho,n)}{12\ (1-\rho^2)} + (\frac{n-1}{2})^2 \frac{\rho^{2q}}{1-\rho^{2q}}} .$$

Also,

$$\lim_{q \to \infty} E_{\hat{\hat{\beta}}} = \frac{\frac{(n-1)g(\rho,n)}{12\ (1-\rho^2)} - (\frac{n-1}{2})^2}{\frac{(n-1)g(\rho,n)}{12\ (1-\rho^2)}} .$$

Values of these efficiencies are presented in Tables 1 and 2. As expected, both estimators are relatively inefficient for large positive values of ρ. The inefficiency diminishes quickly, however, for the P-W estimator as q gets large.

4. Conclusions

This paper has shown that the Prais-Winston estimator is more efficient than the Cochrane-Orcutt estimator if the usual AR(1) model is assumed to have a finite past. Thus, the usual assumption that the latter estimator is preferred in this case is shown to be wrong. Calculated efficiency ratios for the time-trend model suggest that the difference in the efficiency of the two estimators might be substantial for smoothly trended data. These results give further support to the growing belief that whenever possible the first observation should be retained in the serial correlation adjustment.

FOOTNOTES

1/Since economic time series are generally prices or quantities of commodities most (if not all) of which are the result of inventions, innovations or deregulation that have occurred in the finite past, this assumption is, strictly speaking, not valid. Nevertheless, it is reasonable to conjecture that the loss of efficiency may be small if the initial observation of the sample is "sufficiently far" from the initial observation in the time series. The results for the mean and time-trend models are consistent with this conjecture.

2/There are of course an infinite number of other finite past assumptions that could be made. For example, let ε_t be nid(0, σ_ε^2) for $t \leq -q$ and let u_t be nid(0, σ_u^2) for $t > q$, and further assume that $\sigma_\varepsilon^2 = \sigma_u^2/(1-\rho^2)$. This would eliminate the heteroskedasticity in (2), but it would also eliminate the need to distinguish between the finite and infinite past that is characteristic of nearly all discussions of the AR(1) regression model. Thus, attention is limited to the generalization of the usual model.

3/If one wished to use the feasible Aitken's estimator, q would have to be estimated simultaneously with ρ. If the wrong value of q is used, it can be shown that the resulting estimator may be more or less efficient than the C-O estimator. However, the C-O estimator is less efficient for the time-trend model.

4/Readers unfamiliar with the form of these matrices can consult Kadiyala (1969) or Judge, et. al. (1981, p. 181).

5/See Graybill (1969, p. 330).

6/See Kramer (1982) for some additional insight.

Table 1: Calculated Values of $E_{\hat{\beta}}$ for Time-Trend Model

		n				
q	ρ	10	20	50	100	250
	-0.99	0.924565	0.963070	0.985369	0.992705	0.997086
	-0.50	0.992507	0.996051	0.998369	0.999176	0.999668
	-0.20	0.999753	0.999857	0.999937	0.999968	0.999987
	0.20	0.999639	0.999745	0.999872	0.999931	0.999971
	0.50	0.979974	0.982597	0.989273	0.993688	0.997203
1	0.75	0.847964	0.848728	0.870911	0.906280	0.950776
	0.90	0.576887	0.541784	0.539708	0.562906	0.658873
	0.95	0.442701	0.358381	0.325852	0.326317	0.361145
	0.97	0.387519	0.280292	0.219723	0.211732	0.218260
	0.98	0.360330	0.242681	0.166204	0.149160	0.147969
	0.99	0.333661	0.207003	0.116815	0.087616	0.077203
	-0.99	0.99383	0.99709	0.99887	0.99944	0.99978
	-0.50	1.00000	1.00000	1.00000	1.00000	1.00000
	-0.20	1.00000	1.00000	1.00000	1.00000	1.00000
	0.20	1.00000	1.00000	1.00000	1.00000	1.00000
	0.50	1.00000	1.00000	1.00000	1.00000	1.00000
10	0.75	1.00000	1.00000	1.00000	1.00000	1.00000
	0.90	0.99644	0.99590	0.99586	0.99623	0.99748
	0.95	0.97070	0.95273	0.95273	0.95283	0.95931
	0.97	0.93599	0.90000	0.86680	0.86125	0.86581
	0.98	0.90728	0.84771	0.77591	0.75280	0.75104
	0.99	0.86806	0.77426	0.63475	0.55787	0.52364
	0.99	0.99772	0.99893	0.99958	0.99979	0.99992
	-0.50	1.00000	1.00000	1.00000	1.00000	1.00000
	-0.20	1.00000	1.00000	1.00000	1.00000	1.00000
	0.20	1.00000	1.00000	1.00000	1.00000	1.00000
	0.50	1.00000	1.00000	1.00000	1.00000	1.00000
20	0.75	1.00000	1.00000	1.00000	1.00000	1.00000
	0.90	0.99995	0.99995	0.99995	0.99995	0.99997
	0.95	0.99715	0.99596	0.99533	0.99534	0.99600
	0.97	0.98707	0.97916	0.97141	0.97007	0.97117
	0.98	0.97341	0.95419	0.92834	0.91932	0.91861
	0.99	0.94703	0.90311	0.82526	0.77421	0.74920
	-0.99	0.99964	0.99983	0.99993	0.99997	0.99999
	-0.50	1.00000	1.00000	1.00000	1.00000	1.00000
	-0.20	1.00000	1.00000	1.00000	1.00000	1.00000
	0.20	1.00000	1.00000	1.00000	1.00000	1.00000
50	0.50	1.00000	1.00000	1.00000	1.00000	1.00000
	0.75	1.00000	1.00000	1.00000	1.00000	1.00000
	0.90	1.00000	1.00000	1.00000	1.00000	1.00000
	0.95	0.99999	0.99999	0.99999	0.99999	0.99999
	0.97	0.99975	0.99959	0.99944	0.99941	0.99943
	0.98	0.99846	0.99729	0.99565	0.99506	0.99501
	0.99	0.99133	0.98350	0.96796	0.95640	0.95027

Table 2: Calculated Values of $E_{\hat{\beta}}$ for the Time-Trend Model

q	ρ	n = 10	20	50	100	250
	-0.99	0.921468	0.961555	0.984769	0.992405	0.996967
	-0.50	0.870036	0.934313	0.973503	0.986710	0.994674
	-0.20	0.812656	0.901268	0.959222	0.979391	0.991703
	0.20	0.672807	0.804607	0.912880	0.954817	0.981524
	0.50	0.497674	0.629268	0.804975	0.892147	0.954041
1	0.75	0.344461	0.357294	0.519902	0.676713	0.838702
	0.90	0.301061	0.202016	0.191553	0.268166	0.457843
	0.95	0.300654	0.177578	0.105880	0.107860	0.182183
	0.97	0.302590	0.173514	0.085836	0.064457	0.082808
	0.98	0.304002	0.172777	0.079646	0.050151	0.046660
	0.99	0.305704	0.172910	0.076171	0.041504	0.023632
	-0.99	0.990772	0.995650	0.998313	0.999164	0.999668
	-0.50	0.899254	0.949912	0.979995	0.989999	0.996000
	-0.20	0.818792	0.904841	0.960790	0.980199	0.992032
	0.20	0.681729	0.810945	0.916072	0.956546	0.982250
	0.50	0.569149	0.693548	0.846234	0.916869	0.965130
10	0.75	0.544886	0.558820	0.711597	0.826670	0.922163
	0.90	0.665712	0.539260	0.522772	0.628820	0.796097
	0.95	0.738810	0.586893	0.437934	0.443046	0.594441
	0.97	0.770072	0.618406	0.420224	0.347190	0.410700
	0.98	0.785696	0.636779	0.420755	0.307085	0.291194
	0.99	0.801153	0.656707	0.430024	0.282780	0.181318
	-0.99	0.994903	0.997602	0.999071	0.999540	0.999817
	-0.50	0.899254	0.949912	0.979995	0.989999	0.996000
	-0.20	0.818792	0.904841	0.960790	0.980199	0.992032
	0.20	0.681729	0.810945	0.916072	0.956546	0.982250
	0.50	0.569149	0.693548	0.846234	0.916869	0.965130
20	0.75	0.545671	0.559600	0.712246	0.827123	0.922390
	0.90	0.690742	0.567608	0.551291	0.655182	0.814090
	0.95	0.793502	0.658700	0.514201	0.519381	0.665683
	0.97	0.837938	0.714437	0.528067	0.450866	0.518284
	0.98	0.859430	0.745129	0.547782	0.424973	0.406561
	0.99	0.879870	0.776666	0.578333	0.418702	0.287049
	-0.99	0.997332	0.998747	0.999515	0.999760	0.999905
	-0.50	0.899254	0.949912	0.979995	0.989999	0.996000
	-0.20	0.818792	0.904841	0.960790	0.980199	0.992032
	0.20	0.681729	0.810945	0.916072	0.956546	0.982250
	0.50	0.569149	0.693548	0.846234	0.916869	0.965130
50	0.75	0.545673	0.559603	0.712248	0.827125	0.922391
	0.90	0.693908	0.571252	0.554965	0.658532	0.816330
	0.95	0.814237	0.687643	0.546969	0.552105	0.694308
	0.97	0.874880	0.771866	0.602103	0.426145	0.592671
	0.98	0.905367	0.820623	0.654637	0.536281	0.517386
	0.99	0.933454	0.869453	0.724267	0.579736	0.435373
	-0.99	0.998307	0.999205	0.999692	0.999848	0.999939
	-0.50	0.899254	0.949912	0.979995	0.989999	0.996000
	-0.20	0.818792	0.904841	0.960790	0.980199	0.992032
	0.20	0.681729	0.810945	0.916072	0.956546	0.982250
	0.50	0.569149	0.693548	0.846234	0.916869	0.965130
∞	0.75	0.545673	0.559603	0.712248	0.827125	0.922391
	0.90	0.693914	0.571259	0.554971	0.658538	0.816344
	0.95	0.815133	0.688917	0.548440	0.553573	0.695567
	0.97	0.880116	0.780332	0.613715	0.538274	0.604377
	0.98	0.916874	0.840620	0.686060	0.571422	0.552765
	0.99	0.956759	0.913084	0.805571	0.685130	0.548793

REFERENCES

Chipman, J. S., "The Efficiency of Least Squares Estimation of Linear Trend When Residuals are Autocorrelated," Econometrica 47 (1979), pp. 115-128.

Doran, H. E., "Omission of Weighted First Observation in an Autocorrelated Regression Model: A Discussion on Loss of Efficiency," Working Papers in Econometrics and Applied Statistics No. 6, University of New England, Armidale, Australia.

___________, and W. E. Griffiths, "On the Relative Efficiency of Estimators Which Include the Initial Observations in the Estimation of Seemingly Unrelated Regressions with First-Order Autoregressive Disturbances," Journal of Econometrics 23 (October 1983), pp. 165-191.

Fomby, T. B. and D. K. Guilkey, "An Examination of Two-Step Estimators for Models with Lagged Dependent Variables and Autocorrelated Errors," Journal of Econometrics 22 (August 1983), pp. 291-300.

Graybill, F. A., Introduction to Matrices with Application in Statistics, Belmont California: Wadsworth Publishing Co., Inc., 1969.

Judge, G. G., W. E. Griffiths, R. C. Hill, and T. Lee, The Theory and Practice of Econometrics, New York: John Wiley and Sons, 1980.

Kadiyala, K. R., "A Transformation Used to Circumvent the Problem of Autocorrelation," Econometrica 36 (1968), pp. 93-96.

Kramer, W., "Note on Estimating Linear Trend When Residuals are Autocorrelated," Econometrica 50 (1982), pp. 1065-67.

Maeshiro, A., "On the Retention of the First Observation in Serial Correlation Adjustment of Regression Models," International Economic Review, 20 (1979), pp. 259-65.

___________, "Autoregression Transformation, Trended Independent Variables and Autocorrelated Disturbance Terms," The Review Of Economics and Statistics 58 (1976), pp. 497-504.

___________, "Autocorrelation and Trended Explanatory Variables: A Reply." Review of Economics and Statistics, 62 (1980), pp. 487-89.

Spitzer, J. J. "Small Sample Properties of Nonlinear Least Squares and Maximum Likelihood Estimators in the Context of Autocorrelated Errors," Journal of the American Statistical Association 74 (1979), pp. 41-47.