The Appropriate Autocorrelation Transformation when the Autocorrelation Process has a Finite Past

Authors	Daniel L. Thornton
Working Paper Number	1982-002B
Revision Date	February 1982
Citable Link	https://doi.org/10.20955/wp.1982.002
Suggested Citation	Thornton, D.L., 1982; The Appropriate Autocorrelation Transformation when the Autocorrelation Process has a Finite Past, Federal Reserve Bank of St. Louis Working Paper 1982-002. URL https://doi.org/10.20955/wp.1982.002

Revised May 1983

THE APPROPRIATE AUTOCORRELATION
TRANSFORMATION WHEN THE AUTOCORRELATION
PROCESS HAS A FINITE PAST

Daniel L. Thornton
Federal Reserve Bank of St. Louis
Revised 82-002

ID 3391R
DLT:cla
Disk 107R
5/2/83

When a researcher is confronted with estimating a regression model in which the error term is assumed to follow a first-order autoregressive process, he has a choice of two traditional autocorrelation transformations. In the Cochrane-Orcutt (CO) transformation, the initial observation is simply omitted and the remaining observations are differenced using ρ (the coefficient of autocorrelation), i.e., $Y_t^* = Y_t - \rho Y_{t-1}$, $t = 2, 3, \ldots, n$; the other, the Prais-Winsten transformation, is identical to the first except the initial observation is weighted by $\sqrt{1 - \rho^2}$. The second transformation yields more efficient estimates of the parameters of the model, thus it is preferred. Indeed, studies by Kadiyala (1968), Maeshiro (1976 and 1979), Chipman (1979) Beach and MacKinnon (1978), Spitzer (1979) and Doran (1979) indicate the loss in efficiency may be large, especially if the explanatory variables are time trended or if there is strong multi-collinearity. However, the Prais-Winsten transformation assumes the autocorrelation process has been in operation for an indefinite period into the past. If this assumption is too unrealistic, say because of a war or some other break in the series, then it is sometimes suggested that it may be more appropriate to simply omit the first observation rather than weighting it (e.g., Judge, et. al. (1980) p. 182, and Theil (1971) p. 253). The purpose of this paper is to derive the appropriate autocorrelation transformation when the autocorrelation process is assumed to have a finite past. This transformation is extended to the case where there is a gap in the time series.1/

This transformation is then applied to Chow's (1966) long-run money demand equation using Chow's data. The Cochrane-Orcutt transformation and another recently suggested by Savin and White (1978) are also applied. The Chow data is particularly interesting because of the missing observations during the war years.

THE MODEL

We begin with the case where the autocorrelation process is assumed to start with the initial observation. Assume a regression model of the following form:

$$(1) \qquad \begin{aligned} Y_t &= \underline{X}_t \, \underline{B} + U_t, && t = 1, 2, \ldots, n \\ U_t &= \rho U_{t-1} + \varepsilon_t, \ |\rho| < 1, && t > 1 \\ U_t &= \varepsilon_t, && t = 1. \end{aligned}$$

The elements Y_t and U_t are the t^{th} observation on the dependent variable and the error term respectively, and $\underline{X}_t$ and $\underline{B}$ are 1xk and kx1 vectors of the regressors and the parameters of the model, respectively. Assume that $\varepsilon_t \sim \text{NID}\ (0, \sigma_\varepsilon^2)$. It can be shown that

$$(2) \qquad U_t = \sum_{s=0}^{t-1} \rho^s \varepsilon_{t-s}.$$

Therefore,

$$(3) \qquad E(U_t^2) = \sigma_{ut}^2 = \sigma_\varepsilon^2 \sum_{k=1}^{t} \rho^{2(k-1)}, \qquad t=1, 2, \ldots, n$$

which becomes

$$(4) \qquad E(U_t^2) = \sigma_\varepsilon^2 \frac{(1 - \rho^{2t})}{1 - \rho^2} \qquad \text{for all } t.$$

Furthermore,

(5) $$E\,(U_t\, U_{t-s}) = \rho^s\, \sigma^2_{u_{t-s}} \text{ for } t - s \geq 1.$$

Thus, if

$$\underline{U} = (U_1\, U_2\, U_3\, \ldots\, U_n)',$$

then E ($\underline{U}\,\underline{U}'$) = V, where V is given by (6).

(6) $$\sigma^2_\varepsilon \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \cdots & \rho^{n-2} & \rho^{n-1} \\ \rho & \frac{1-\rho^4}{1-\rho^2} & \rho(\frac{1-\rho^4}{1-\rho^2}) & \rho^2(\frac{1-\rho^4}{1-\rho^2}) & \cdots & \rho^{n-3}(\frac{1-\rho^4}{1-\rho^2}) & \rho^{n-2}(\frac{1-\rho^4}{1-\rho^2}) \\ \rho^2 & \rho(\frac{1-\rho^4}{1-\rho^2}) & (\frac{1-\rho^6}{1-\rho^2}) & \rho(\frac{1-\rho^6}{1-\rho^2}) & \cdots & \rho^{n-4}(\frac{1-\rho^6}{1-\rho^2}) & \rho^{n-3}(\frac{1-\rho^6}{1-\rho^2}) \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \rho^{n-1} & \rho^{n-2}(\frac{1-\rho^4}{1-\rho^2}) & \rho^{n-3}(\frac{1-\rho^6}{1-\rho^2}) & \rho^{n-4}(\frac{1-\rho^8}{1-\rho^2}) & \cdots & \rho(\frac{1-\rho^{2n-2}}{1-\rho^2}) & (\frac{1-\rho^{2n}}{1-\rho^2}) \end{bmatrix}$$

Employing two theorems for finding the inverse of certain pattern matrices from Graybill (1969, pp. 179-80), the inverse of V is

(7) $$\frac{1}{\sigma^2_\varepsilon} \begin{bmatrix} 1+\rho^2 & -\rho & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} = \frac{1}{\sigma^2_\varepsilon}\ \Omega^{-1}$$

Also, there exists an n by n transformation matrix T such that $T'T = \Omega^{-1}$, where T is given by

$$(8)\ T = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}$$

Thus, if the autocorrelation process begins with the initial observation, the initial observation should neither be dropped nor weighted by $\sqrt{1-p^2}$, but rather simply included unweighted.

Furthermore, $|\Omega^{-1}| = 1$. Thus, if the model is transformed by T, and OLS is applied, the estimates of ρ and $\underline{B}$ which minimized the residual sum of squares are maximum likelihood estimates.[2/]

A Generalization

Unfortunately, the above transformation assumes that one begins the estimation precisely when the autocorrelation process started. The question naturally arises: what would be the appropriate transformation if the autocorrelation process were known to have started q periods ago? Fortunately, the transformation can be determined readily. Moreover, it has a form which produces the Prais-Winsten transformation and the above transformation as special cases.

To illustrate this, assume the error structure in equation (1) is known to be generated by a first-order autocorrelation process which began q periods ago. If $\underline{\mu}$ denotes the n by 1 error vector from this equation, then $E(\underline{\mu}\,\underline{\mu}')$ is given by an n by n matrix whose (1,1) element corresponds to the (q+1, q+1) element of V given by (6). That is,

$$(9)\quad E(\underline{\mu}\,\underline{\mu}')=\sigma_\varepsilon^2\begin{bmatrix} \frac{1-\rho^{2q^*}}{1-\rho^2} & \frac{\rho(1-\rho^{2q^*})}{1-\rho^2} & \frac{\rho^2(1-\rho^{2q^*})}{1-\rho^2} & \cdots & \frac{\rho^{n-1}(1-\rho^{2q^*})}{1-\rho^2} \\ \frac{\rho(1-\rho^{2q^*})}{1-\rho^2} & \frac{(1-\rho^{2(q^*+1)})}{1-\rho^2} & \frac{\rho(1-\rho^{2(q^*+1)})}{1-\rho^2} & \cdots & \frac{\rho^{n-2}(1-\rho^{2(q^*+1)})}{1-\rho^2} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{\rho^{n-1}(1-\rho^{2q^*})}{1-\rho^2} & \frac{\rho^{n-2}(1-\rho^{2(q^*+1)})}{1-\rho^2} & \frac{\rho^{n-3}(1-\rho^{2(q^*+2)})}{1-\rho^2} & \cdots & \frac{(1-\rho^{2(q^*+n-2)})}{1-\rho^2} \end{bmatrix}$$

$= \sigma_\varepsilon^2\ \phi$, where $q^* = q+1$.

Applying the same theorems from Graybill as before, it can be shown that

$$(10)\quad \phi^{-1}=\begin{bmatrix} \frac{1-\rho^{2(q+2)}}{1-\rho^{2(q+1)}} & -\rho & 0 & 0 & \cdots & 0 \\ -\rho & (1+\rho^2) & -\rho & 0 & \cdots & 0 \\ 0 & -\rho & (1+\rho^2) & -\rho & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

Consequently, the appropriate transformation matrix is the same as the T matrix above, except the (1,1) element is

$$\left[\frac{1-\rho^{2(q+2)}}{1-\rho^{2(q+1)}}-\ \rho^2\right]^{1/2} = \left[\frac{1-\rho^2}{1-\rho^{2(q+1)}}\right]^{1/2}$$

This term approaches $\sqrt{1-\rho^2}$ as $q \to \infty$, and is equal to 1 when $q = 0$. Thus, the Prais-Winsten transformation and the previous transformation given by (8) are special cases of a finite first-order autoregressive process.[3/]

The Case of Missing Observations

The above method can be extended easily to the case of missing observations, as recently considered by Savin and White (1978) for the continuous case. For example, assume, that an autoregressive process beginning in the indefinite past was interrupted, say by war, for g years, then restarted. This model corresponds to that assumption:

$$
\begin{aligned}
(11)\qquad & Y_t = \underline{X}_t \, \underline{B} + U_t, && t = 1, 2, \ldots n_1, n_1 + g + 1, \ldots, n \\
& U_t = \rho U_{t-1} + \varepsilon_t && t \leq n_1, \text{ or } t \geq n_1 + g + 2 \\
& U_t = \varepsilon_t, && t = n_1 + g + 1,
\end{aligned}
$$

where we assume that $E(U_t \, U_{t'}) = 0$ if $t \leq n_1$ and $t' \geq n_1 + g + 1$.

If we let $\underline{U} = (U_1 \, U_2 \, \ldots \, U_{n1} \, U_{n1+g+1} \, \ldots \, U_n)'$, then $E(\underline{U} \, \underline{U}')$ is given by (12).

$$
(12)\quad \sigma_\varepsilon^2 \left[\begin{array}{ccccc|ccccc}
\frac{1}{1-\rho^2} & \frac{\rho}{1-\rho^2} & \frac{\rho^2}{1-\rho^2} & \cdots & \frac{\rho^{n-1}}{1-\rho^2} & & & & & \\
\frac{\rho}{1-\rho^2} & \frac{1}{1-\rho^2} & \frac{\rho}{1-\rho^2} & \cdots & \frac{\rho^{n-2}}{1-\rho^2} & & & & & \\
\vdots & \vdots & \vdots & & \vdots & & & 0 & & \\
\frac{\rho^{n-1}}{1-\rho^2} & \frac{\rho^{n-2}}{1-\rho^2} & \frac{\rho^{n-3}}{1-\rho^2} & \cdots & \frac{1}{1-\rho^2} & & & & & \\
\hline
 & & & & & 1 & \rho & \rho^2 & \cdots & \rho^k \\
 & & & & & \rho & \frac{(1-\rho^4)}{1-\rho^2} & \frac{\rho(1-\rho^4)}{1-\rho^2} & \cdots & \frac{\rho^{k-1}(1-\rho^4)}{1-\rho^2} \\
 & & 0 & & & \vdots & \vdots & \vdots & & \vdots \\
 & & & & & \rho^k & \frac{\rho^{k-1}(1-\rho^4)}{1-\rho^2} & \cdot & \cdots & \frac{1-\rho^{2k}}{1-\rho^2}
\end{array}\right]
$$

where $k = n - (n_1 + g)$. If this expression is rewritten as

$$
(13)\quad \sigma_\varepsilon^2 \ \Psi = \sigma_\varepsilon^2 \left[\begin{array}{c|c} \Psi_1 & 0 \\ \hline 0 & \Psi_2 \end{array}\right] ,
$$

then

$$
\Psi^{-1} = \left[\begin{array}{c|c} \Psi_1^{-1} & 0 \\ \hline 0 & \Psi_2^{-1} \end{array}\right] .
$$

Ψ_1^{-1} is identical in form to V^{-1} , except the (1, 1) term of Ψ_1^{-1} is 1. Ψ_2^{-1} is, of course, identical to V^{-1}.

Thus, $\Psi^{-1} = H'H$, where

$$(14)\ H = \left[\begin{array}{cccccc|cccccc}
\sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 & & & & & & \\
-\rho & 1 & 0 & \cdots & 0 & & & & 0 & & \\
0 & -\rho & 1 & \cdots & 0 & & & & & & \\
\vdots & \vdots & \vdots & & \vdots & & & & & & \\
0 & 0 & 0 & \cdots & 1 & & & & & & \\
\hline
 & & & & & & 1 & 0 & 0 & \cdots & 0 \\
 & & & & & & -\rho & 1 & 0 & \cdots & 0 \\
 & & & & & & 0 & -\rho & 1 & \cdots & 0 \\
 & & 0 & & & & \vdots & \vdots & \vdots & & \vdots \\
 & & & & & & 0 & 0 & 0 & \cdots & 1
\end{array}\right]$$

Therefore, if Savin and White's assumption that the autocorrelation process continues over the break in the data seems too unrealistic, there exists an appropriate transformation which does not require one to delete the "initial" observations.[4/] Hence, one would not necessarily want to delete these observations as is often suggested. Estimates based on the above transformations will be more efficient in small samples, and, if Monte Carlo experiments on the Prais-Winsten and Cochrane-Orcutt transformations are any guide, they will probably be more efficient in fairly large samples as well. Unfortunately, the exact transformation for the finite, first-order autoregressive process requires that the researcher know the exact starting point in the process. Thus, one might be tempted to delete the observations rather than perhaps weight them inappropriately. The empirical results below, however, suggest that deleting the observations may be as bad or worse than weighting them.

An Illustrative Example Using Chow's Money Demand

In order to see what the effect of the treatment of initial observations might have in applied work, Chow's (1966) money demand equations were estimated using Chow's data. Chow undertook his study, in part, to test for the asset and transactions motives for holding money. His OLS estimates gave strong support to the asset over the transactions motive; current real income was statistically insignificant when included along with either permanent income or wealth. Also, the results suggested a substantial interest elasticity of the demand for money. Recently, Lieberman (1980) has argued that Chow's results are largely attributable to the fact that Chow did not adjust for serial correlation, and that Chow ignored a structural shift in the demand for money that occurred in 1933. Lieberman reestimated the Chow equations for the period 1934-1958, adjusting for autocorrelation using the iterative CO technique, and found Chow's conclusions reversed. Chow's data, however, does not contain estimates of permanent income or wealth for the war years. Thus, there are two "initial" observations which the Cochrane-Orcutt transformation deletes during the 1934-58 period.

Thornton (1982) has shown that Chow's conclusions concerning the relative importance of current and permanent income are maintained if we use the autocorrelation transformation suggested by Savin and White. In order to investigate the sensitivity of these results to the treatment of the initial observations, Chow's long-run money demand equation was estimated using different treatments of the initial observations. The following equation was estimated for the 1934-1958 period:

$$\ln (M/P)_t = \alpha_0 + \alpha_1 \ln (YP/P)_t$$
$$+ \alpha_2 \ln (Y/P)_t + \alpha_3 \ln (\text{Bond})_t + U_t \, .$$

Here, M, YP, Y and Bond denote money, permanent income, current income and the interest rate, respectively, and P denotes the level of prices.[5/]

The equation was estimated with OLS and an autocorrelation correction with seven different treatments of the initial observations: Cochrane-Orcutt, the transformation proposed above with five different restarting points (1946-1942) and the Savin-White transformation which assumes that the autocorrelation process is continuous over the break. The results are presented in table 1.[6/] The estimates appear somewhat sensitive to the treatment of the initial observations, with one of the coefficients changing sign based on the treatment of these obserations. These results suggest that one be careful with the way that initial observations are weighted when correcting for first-order autocorrelation.

It is interesting to note, however, that the autocorrelation-adjusted estimates remain consistent with each other and with the OLS estimates except when the CO transformation is used. Permanent income and the long-term bond rate are insignificant and current income is significant only when the CO transformation is used. Also, the estimates with the finite-time-horizon autocorrelation adjustments are similar to the OLS estimates and to the estimates using the Savin and White transformation. The latter result is to be expected because the finite-time-horizon transformation is very similar to the Savin-White transformation for these data.[7/]

These results suggest that one should be careful when using the CO transformation, especially if the number of initial (dropped) observations is large relative to the sample size--in our case, 10 percent.[8/] At a bare minimum, one should investigate the sensitivity of the results to the treatment of the initial observations.

FOOTNOTES

* I would like to thank Thomas B. Fomby, Dallas S. Batten, Stanley R. Johnson and Basile Goungetas for their suggestions on earlier drafts of this paper. I would also like to thank John Schulte for his assistance with computer programming.

1/ Detailed proofs of the following results are available from the author upon written request.

2/ The log of the likelihood function is given by

$$\ln(L) = n/2 \ln(2\pi) - n/2 \ln(\sigma_\varepsilon^2)$$

$$- 1/2 \ln \Omega - \frac{1}{2\sigma_\varepsilon^2} (Y - X\underline{B})' \Omega^{-1} (Y - X\underline{B}).$$

However, since the Jacobian of the transformation (Ω) is equal to one, maximizing the likelihood function is the same as minimizing the residual sum of squares from an appropriately transformed equation. This is not the case when the autocorrelation process has an indefinite past. See Beach and MacKinnon (1978) for details.

3/ Since $\left|\phi^{-1}\right| = \frac{1-\rho^2}{1-\rho^{2(q+1)}}$, maximum likelihood (ML) estimates of this model could be obtained using a grid search technique by minimizing $\frac{RSS(\rho)}{\left[\frac{1-\rho^2}{1-\rho^{2(q+1)}}\right]^{1/n}}$, where RSS($\rho$) denotes the residual sum of squares from the appropriately transformed equation, conditional on ρ. ML estimates could also be obtained by applying an appropriately modified nonlinear technique [e.g., Beach and MacKinnon (1978)].

4/ It is interesting to note that if the model conforms to (11), it would be possible to estimate a separate value of ρ over each subset of data, and to test the null hypothesis that ρ remained unchanged when the autocorrelation process restarted.

5/ For a detailed discussion of Chow's variables, see Chow (1966, pp. 128-29).

6/ All the autocorrelation adjustments were carried out using a ML grid search to an accuracy of .01 over positive values of rho.

7/ The Savin and White (1978) transformation adjusts the t^{th} observation by $g\ X_t - \rho^{(q+1)}\ g\ X_{t-q-1}$, where $g = \sqrt{(1-\rho^2)/(1-\rho^{2(q+1)})}$ and t=1946. Note that $\rho^{(q+1)}g \to 0$ as $q \to \infty$. In our case, $\rho^{(q+1)}g \approx 0$ for fairly small values of q because $\rho \approx .50$ (e.g., $(\rho^{(q+1)}g) = .041$ for q=4 and $\rho = .57$).

8/ Results from Thornton (1982) suggest that this result may hold for proportions smaller than 10 percent.

REFERENCES

Beach, C. M. and J. G. MacKinnon: "A Maximum Likelihood Procedure for Regression with Autocorrelated Errors," Econometrica, 46 (1978), pp. 51-58.

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

Chow, G. C., "On the Long-Run and Short-Run Demand for Money," Journal of Political Economy, 74 (1966), pp. 111-131.

Doran, H. E.: "Omission of Weighted First Observation In an Autocorrelated Regression Model: A Discussion of Loss of Efficiency," Working Papers in Econometrics and Applied Statistics No. 6, University of New England, Armidala, Australia, 1979.

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

Judge, G. G., et. al.: 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.

Lieberman, C., "The Long-Run and Short-Run Demand for Money, Revisited," Journal of Money, Credit and Banking, 12 (1980), pp. 43-57.

Maeshiro, A.: "On the Retention of the First Observations in Serial Correlation Adjustment of Regression Models," International Economic Review, 20 (1979), pp. 259-265.

__________: "Autoregressive Transformation, Trended Independent Variables and Autocorrelated Disturbance Terms," The Review of Economics and Statistics, 58 (1976), pp. 497-500.

Savin, N.E. and K. J. White: "Testing for Autocorrelation with Missing Observations," Econometrica, 46 (1978), pp. 59-68.

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.

Theil, H.: Principles of Econometrics, New York: John Wiley and Sons, 1971.

Thornton, D.L.: "The Long-Run and Short-Run Demand for Money: Additional Evidence," Journal of Macroeconomics 4 (1982), pp. 325-38.

Table 1: Chow's Long-Run Demand for Money, 1934-1958

Weights		Coefficients				
1934	1946	const.	YP/P	Y/P	Bond	ρ
*	*	-1.308 (2.231)	1.444 (5.240)	-.287 (1.259)	-.601 (7.140)	*
0	0	2.109 (1.462)	.097 (0.324)	.680 (3.025)	-.226 (1.728)	.76
$\sqrt{1-\rho^2}$	1	-.993 (1.372)	1.263 (4.273)	-.140 (0.575)	-.600 (6.299)	.43
$\sqrt{1-\rho^2}$	q=1	-.852 (1.100)	1.164 (3.820)	-.059 (0.236)	-.584 (5.914)	.50
$\sqrt{1-\rho^2}$	q=2	-.783 (0.971)	1.115 (3.615)	-.019 (0.077)	-.575 (5.713)	.54
$\sqrt{1-\rho^2}$	q=3	-.734 (0.881)	1.081 (3.476)	.009 (0.033)	-.568 (5.564)	.57
$\sqrt{1-\rho^2}$	q=4	-.702 (0.853)	1.058 (3.385)	.027 (0.105)	-.564 (5.463)	.59
$\sqrt{1-\rho^2}$	S-W	-.742 (0.884)	1.075 (3.463)	.015 (0.058)	-.564 (5.460)	.58

The absolute value of the t-ratio is in parentheses below each coefficient.

* indicates the equation was estimated with OLS.

S-W indicates the gap was treated with the Savin-White transformation. See Savin and White (1978).

TECHNICAL APPENDIX FOR AUTOREGRESSION TRANSFORMATION PAPER

The Model:

$$Y_t = \underline{X}_t \underline{\beta} + U_t, \qquad t = 1, 2, \ldots, n$$

$$U_t = \rho U_{t-1} + \varepsilon_t, \qquad t > 1$$

$$U_t = \varepsilon_t, \qquad t = 1$$

If $\underline{U} = (U_1 \;\; U_2 \;\; U_3 \;\; \ldots \;\; U_n)'$, then

$$V = E(\underline{U}\,\underline{U}') = \sigma_\varepsilon^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \cdots & \rho^{n-2} & \rho^{n-1} \\ \rho & \frac{1-\rho^4}{1-\rho^2} & \rho\left(\frac{1-\rho^4}{1-\rho^2}\right) & \rho^2\left(\frac{1-\rho^4}{1-\rho^2}\right) & \cdots & \rho^{n-3}\left(\frac{1-\rho^4}{1-\rho^2}\right) & \rho^{n-2}\left(\frac{1-\rho^4}{1-\rho^2}\right) \\ \rho^2 & \rho\left(\frac{1-\rho^4}{1-\rho^2}\right) & \left(\frac{1-\rho^6}{1-\rho^2}\right) & \rho\left(\frac{1-\rho^6}{1-\rho^2}\right) & \cdots & \rho^{n-4}\left(\frac{1-\rho^6}{1-\rho^2}\right) & \rho^{n-3}\left(\frac{1-\rho^6}{1-\rho^2}\right) \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho^{n-1} & \rho^{n-2}\left(\frac{1-\rho^4}{1-\rho^2}\right) & \rho^{n-3}\left(\frac{1-\rho^6}{1-\rho^2}\right) & \rho^{n-4}\left(\frac{1-\rho^8}{1-\rho^2}\right) & \cdots & \rho\left(\frac{1-\rho^{2n-2}}{1-\rho^2}\right) & \left(\frac{1-\rho^{2n}}{1-\rho^2}\right) \end{bmatrix}$$

or

$$V = \sigma_\varepsilon^2\, \Omega$$

Now we can apply two theorems from Graybill (1969, pp. 179-80). Denote the ith, jth element of Ω as b_{ij}. From this definition it is easy to verify that

$$b_{ij} = \frac{(1-\rho^{2(i)})\rho^{j-i}}{(1-\rho^2)} = b_{ji} \quad \text{for } j \geq i.$$

Likewise, it is easily verified that the conditions of Graybill's theorem 8.3.6 hold, e.g.,

$$\theta_t = \frac{b_{tj}}{b_{1j}} = \frac{(1-\rho^{2t})}{(1-\rho^2)\rho^{t-1}} \quad j \le t \le n,\ t \ge 2.$$

Now let C_{ij} denote the ith, jth element of Ω^{-1}. By theorem 8.3.7,

$$C_{11} = -\theta_2 (b_{12} - \theta_2 b_{11})^{-1}$$

$$= -\frac{(1-\rho^4)}{(1-\rho^2)\rho}\left[\rho - \frac{(1-\rho^4)}{(1-\rho^2)\ \rho}\right]^{-1}$$

$$= -\frac{(1-\rho^4)}{(1-\rho^2)\ \rho}\left[\frac{\rho^2 - \rho^4 - 1 + \rho^4}{(1-\rho^2)\ \rho}\right]^{-1}$$

$$= \frac{(1-\rho^4)}{(1-\rho^2)} = \frac{(1-\rho^2)\ (1+\rho^2)}{(1-\rho^2)}$$

$$= 1+\rho^2.$$

Likewise,

$$Cnn = -\frac{b_{1,\ n-1}}{b_{1,n}(b_{n-1,n} - \theta_n b_{1,n-1})}$$

$$= -\frac{\rho^{n-2}}{\rho^{n-1}\left[\frac{(1-\rho^{2\ (n-1)})\rho}{(1-\rho^2)} - \frac{(1-\rho^{2n})\ \rho^{n-2}}{(1-\rho^2)\ \rho^{n-1}}\right]}$$

$$= -\frac{1}{\rho\left(\frac{\rho^2 - \rho^{2n} - 1 + \rho^{2n}}{\rho(1-\rho^2)}\right)}$$

$$= -\frac{1}{\rho/-\rho} = 1.$$

Also,

$$Ctt = -\frac{b_{t-1,\ t+1} - \theta_{t+1}\, b_{1,\ t-1}}{(b_{t-1,\ t} - \theta_t\, b_{1,\ t-1})\ (b_{t,\ t+1} - \theta_{t+1}\, b_{1,\ t}}),\ t=2,\ 3,\ \ldots,\ n-1.$$

The numerator of this expression is

$$- \quad (\frac{1-\rho^{2(t-1)}}{(1-\rho^2)})\rho^2 - (\frac{1-\rho^{2(t+1)}}{(1-\rho^2)\ \rho^t})\ \rho^{t-2}$$

$$= - \quad \frac{\rho^4 - \rho^{2(t+1)} - 1 + \rho^{2(t+1)}}{(1 - \rho^2)\ \rho^2}$$

$$= \frac{(1-\rho^4)}{\rho^2(1-\rho^2)} = \frac{(1+\rho^2)\ (1-\rho^2)}{\rho^2(1-\rho^2)} = \frac{(1+\rho^2)}{\rho^2}.$$

The denominator is

$$[(\frac{1-\rho^{2(t-1)}}{(1-\rho^2)})\rho - (\frac{1-\rho^{2t})\ \rho^{t-2}}{(1-\rho^2)\rho^{t-1}}].$$

$$[(\frac{1-\rho^{2t}}{(1-\rho^2)})\ \rho - (\frac{1-\rho^{2(t+1)})\ \rho^{t-1}}{(1-\rho^2)\ \rho^t}]$$

$$= [\frac{\rho^2 - \rho^{2t} - 1 + \rho^{2t}}{\rho(1 - \rho^2)}]\ [\frac{\rho^2 - \rho^{2(t+1)} - 1 + \rho^{2(t+1)}}{\rho(1 - \rho^2)}]$$

$$= [\frac{\rho^2-1}{\rho(1-\rho^2)}][\frac{\rho^2-1}{\rho(1-\rho^2)}] = (\frac{-1}{\rho})\ (\frac{-1}{\rho}) = \frac{1}{\rho^2}.$$

Therefore,

$$C_{tt} = \frac{(1+\rho^2)}{\rho^2} \cdot \frac{\rho^2}{1} = 1+\rho^2.$$

Also,

$$C_{t,\,t-1} = (b_{t-1,t} - \theta_t\, b_{1,t-1})^{-1} = C_{t-1,\,t},\ t=2,\ 3,\ \ldots,\ n$$

$$= \left[\frac{(1-\rho^{2(t-1)})\rho}{(1-\rho^2)} - \frac{(1-\rho^{2t})\ \rho^{t-2}}{(1-\rho^2)\ \rho^{t-1}}\right]^{-1}$$

$$= \left[\frac{\rho^2 - \rho^{2t} - 1 + \rho^{2t}}{(1-\rho^2)\ \rho}\right]^{-1} = -\rho.$$

Therefore,

$$V^{-1} = \frac{1}{\sigma_\varepsilon^2}\begin{bmatrix} 1+\rho^2 & -\rho & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}$$

From this it is easy to produce the appropriate transformation matrix, T:

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

Moreover, because $\Omega^{-1} = T'T$ and $|\Omega^{-1}| = |T'||T|$, it is clear that $|\Omega^{-1}| = 1$. Therefore, the values of ρ and $\underline{\beta}$ which minimize the sum of squares of residuals from the transformed equation are maximum likelihood estimates.

Extension of the general finite duration autocorrelation process

Now the question quite naturally arises: What would be appropriate transformation if the autocorrelation process were known to have begun (q-1) periods ago? The variance-covariance matrix would be

$$E(\underline{\mu}\,\underline{\mu}')=\sigma_\varepsilon^2\begin{bmatrix} \frac{1-\rho^{2q}}{1-\rho^2} & \frac{\rho(1-\rho^{2q})}{1-\rho^2} & \frac{\rho^2(1-\rho^{2q})}{1-\rho^2} & \cdots & \frac{\rho^{n-1}(1-\rho^{2q})}{1-\rho^2} \\ \frac{\rho(1-\rho^{2q})}{1-\rho^2} & \frac{(1-\rho^{2(q+1)})}{1-\rho^2} & \frac{\rho(1-\rho^{2(q+1)})}{1-\rho^2} & \cdots & \frac{\rho^{n-2}(1-\rho^{2(q+1)})}{1-\rho^2} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{\rho^{n-1}(1-\rho^{2q})}{1-\rho^2} & \frac{\rho^{n-2}(1-\rho^{2(q+1)})}{1-\rho^2} & \frac{\rho^{n-3}(1-\rho^{2(q+2)})}{1-\rho^2} & \cdots & \frac{(1-\rho^{2(q+n-1)})}{1-\rho^2} \end{bmatrix}$$

$$= \sigma_\varepsilon^2\,\Gamma\,.$$

It is clear that the ith, jth element of Γ is

$$W_{ij} = \frac{(1-\rho^{2(q+i-1)})}{1-\rho^2}\rho^{j-i}\,.$$

Thus,

$$\theta_t = \frac{W_{tj}}{W_{1j}} = \frac{(1-\rho^{2(q+t-1)})}{(1-\rho^{2q})\ \rho^{t-1}} \qquad j \leq t \leq n,\ t \geq 2.$$

By theorem 8.3.7,

$$C_{11} = -\ \theta_2\ (W_{12} - \theta_2\ W_{11})^{-1}$$

$$= -\ \frac{(1-\rho^{2(q+1)})}{(1-\rho^{2q})\ \rho}\left[\ \frac{\rho(1-\rho^{2q})}{1-\rho^2} - \frac{(1-\rho^{2(q+1)})}{(1-\rho^{2q})\ \rho}\cdot\ \frac{(1-\rho^{2q})}{1-\rho^2}\ \right]^{-1}$$

$$= - \frac{(1-\rho^{2(q+1)})}{(1-\rho^{2q})\ \rho} \left[\frac{\rho^{2} - \rho^{2(q+1)} - 1 + \rho^{2(q+1)}}{(1-\rho^{2})\ \rho} \right]^{-1}$$

$$= \frac{(1-\rho^{2(q+1)})}{(1-\rho^{2q})}$$

Also,

$$Cnn = - \frac{W_{1,\ n-1}}{W_{1,n}\ (W_{n-1,\ n} - \theta_n\ W_{1,\ n-1})}$$

$$= - \frac{\dfrac{\rho^{n-2}(1-\rho^{2q})}{(1-\rho^{2})}}{\dfrac{\rho^{n-1}(1-\rho^{2q})}{(1-\rho^{2})} \left[\dfrac{\rho(1-\rho^{2(n+q-2)})}{(1-\rho^{2})} - \dfrac{(1-\rho^{2(n+q-1)})}{(1-\rho^{2q})\rho^{n-1}} \dfrac{(1-\rho^{2q})\rho^{n-2}}{(1-\rho^{2})} \right]}$$

$$= \frac{-1}{\rho\ \dfrac{[\rho^{2} - \rho^{2(n+q-1)} - 1 + \rho^{2(n+q-1)}]}{\rho\ (1-\rho^{2})}}$$

$$= \frac{-1}{-1} = 1\ ;$$

likewise

$$C_{t,\ t-1} = C_{t-1,t} = (W_{t-1,t} - \theta_t\ W_{1,t-1})^{-1},\ t=2,\ 3,\ \ldots,\ n-1$$

$$= \left[\frac{(1-\rho^{2(q+t-2)})\rho}{(1-\rho^{2})} - \frac{(1-\rho^{2(q+t-1)})}{(1-\rho^{2q}\ )\rho^{t-1}} \cdot \frac{(1-\rho^{2q})\rho^{t-2}}{(1-\rho^{2})} \right]^{-1}$$

$$= \left[\frac{\rho^{2} - \rho^{2(q+t-1)} - 1 + \rho^{2(q+t-1)}}{(1-\rho^{2})\ \rho} \right]^{-1}$$

$$= -\left(\frac{1}{\rho} \right)^{-1} = - \rho\ .$$

Furthermore,

$$C_{tt} = - \frac{W_{t-1,\ t+1} - \theta_{t+1}\ W_{1,\ t-1}}{(W_{t-1,\ t} - \theta_t\ W_{1\ t-1})(W_{t,\ t+1} - \theta_{t+1}\ W_{1t})},\ t=2,\ 3,\ \ldots,\ n-1$$

The numerator of this expression is

$$- \left[\frac{\rho^2 (1-\rho^{2(q+t-2)})}{(1-\rho^2)} - \frac{(1-\rho^{2(q+t)})}{(1-\rho^{2q})\ \rho^t} \cdot \frac{(1-\rho^{2q})\ \rho^{t-2}}{(1-\rho^2)} \right]$$

$$= - \left[\frac{\rho^4 - \rho^{2(q+t)} - 1 + \rho^{2(q+t)}}{(1-\rho^2)\ \rho^2} \right]$$

$$= \frac{(1-\rho^4)}{(1-\rho^2)\ \rho^2} = \frac{(1+\rho^2)}{\rho^2}.$$

The denominator of this expression is

$$\left[\frac{(1-\rho^{2(q+t-2)})}{(1-\rho^2)} \rho - \frac{(1-\rho^{2(q+t-1)})}{(1-\rho^{2q})\ \rho^{t-1}} \cdot \frac{(1-\rho^{2q})\ \rho^{t-2}}{(1-\rho^2)} \right].$$

$$\left[\frac{(1-\rho^{2(q+t-1)})}{(1-\rho^2)} \rho - \frac{(1-\rho^{2(q+t)})}{(1-\rho^{2q})\rho^t} \cdot \frac{(1-\rho^{2q})\ \rho^{t-1}}{(1-\rho^2)} \right]$$

$$= \left[\frac{\rho^2 - \rho^{2(q+t-1)} - 1 + \rho^{2(q+t-1)}}{(1 - \rho^2)\ \rho} \right].$$

$$\left[\frac{\rho^2 - \rho^{2(q+t)} - 1 + \rho^{2(q+t)}}{(1-\rho^2)\ \rho} \right]$$

$$= (- \frac{1}{\rho})\ (- \frac{1}{\rho}) = \frac{1}{\rho^2}.$$

Therefore,

$$C_{tt} = \frac{(1+\rho^2)}{\rho^2} \cdot \frac{\rho^2}{1} = (1+\rho^2).$$

Thus,

$$\Gamma^{-1} = \begin{bmatrix} \frac{1-\rho^{2(q+1)}}{1-\rho^{2q}} & -\rho & 0 & 0 & \cdots & 0 \\ -\rho & (1+\rho^2) & -\rho & 0 & \cdots & 0 \\ 0 & -\rho & (1+\rho^2) & -\rho & \cdots & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

Note now that the (1,1) element of r^{-1} is $(1+\rho^2)$ if q=1, i.e.,

$$r^{-1}_{11} = \frac{(1-\rho^4)}{(1-\rho^2)} = \frac{(1+\rho^2)(1-\rho^2)}{(1-\rho^2)} = (1+\rho^2) \ .$$

Also, note that

$$\lim_{q\to\infty} r^{-1}_{11} = \lim_{q\to\infty} \frac{(1-\rho^{2(q+1)})}{(1-\rho^{2q})} = 1.$$

Thus, we get the Prais-Winsten transformation and the transformation noted earlier as special cases of the above.