

AN ADJUSTED MAXIMUM LIKELIHOOD ESTIMATOR
OF AUTOCORRELATION IN DISTURBANCES

by

Clifford Hildreth and Warren T. Dent

Discussion Paper No. 11, October 1971

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minnesota 55455

An Adjusted Maximum Likelihood Estimator
of Autocorrelation in Disturbances

by

Clifford Hildreth and Warren T. Dent*

1. Maximum Likelihood Estimates

Although it has been shown [9] that the maximum likelihood (ML) estimator of the autocorrelation coefficient in linear models with autoregressive disturbances is asymptotically unbiased, several Monte Carlo studies [8, 6, 15] suggest that finite sample bias is usually large enough to be of some concern. In the next section an approximation to the bias is developed and used to obtain an adjusted estimator with substantially smaller bias. Section 3 presents the results of applying the adjusted ML estimator, the unadjusted ML, and two other estimators to Monte Carlo data. Some interpretations and conjectures comprise Section 4 and computing procedures are discussed in Section 5. The remainder of this section contains a brief sketch of maximum likelihood estimation.

The model is a normal linear regression model in which the disturbance (or error term) is assumed to be generated by a stationary first order autoregressive process, i.e.

$$(1) \quad y_t = \sum_{k=1}^k x_{tk} \beta_k + u_t \quad t = 1, 2, \dots, T$$

* The University of Minnesota and The University of Iowa.

This research was supported by National Science Foundation grants GS2268 and GS3317 to the University of Minnesota. The authors are indebted to The RAND Corporation for permission to use the results and the data tapes from the Hildreth-Lu study reported in Reference [8]. The development in Section 2 of an approximate mean of the maximum likelihood estimator is a condensation of the development presented in Dent's thesis [4]. The thesis also contains preliminary studies of other properties of the distribution of $\hat{\rho}$. Applications of these to statistical inference, especially testing, will be reported later.

$$(2a) \quad u_t = \rho u_{t-1} + v_t \quad t = 2, \dots, T$$

$$(2b) \quad u_1 = (1-\rho^2)^{-\frac{1}{2}} v_1$$

where y_t is an observed value of a dependent variable, x_{tk} is an observed value of the k^{th} independent variable (treated as nonrandom), β_k is an unknown constant to be estimated, and u_t is a value taken by an unobserved random disturbance. ρ is an unknown coefficient with $|\rho| < 1$. The u_t are normal with mean zero and common variance denoted by $\frac{v}{1-\rho^2}$. The v_t are normal, independent, identical with variance v .

Alternatively, one may write

$$(1^*) \quad y = X\beta + u$$

where y is a vector of order T and u is a drawing from a multivariate normal population with mean 0 and variance vA where

$$(3) \quad a_{st} = (1-\rho^2)^{-1} \rho^{|t-s|} \quad s, t = 1, 2, \dots, T.$$

The log-likelihood function is then

$$(4) \quad \begin{aligned} L(\beta, \rho, v) &= \text{const} - \frac{T}{2} \log v - \frac{1}{2} \log |A| - \frac{1}{2v} (y - X\beta)' A^{-1} (y - X\beta) \\ &= \text{const} - \frac{T}{2} \log v + \frac{1}{2} \log (1-\rho^2) - \frac{1}{2v} (y - X\beta)' B (y - X\beta) \end{aligned}$$

where $B = A^{-1}$. It may be verified that $B = I + \rho^2 I^* - 2\rho H$ where I is an identity of order T , I^* equals I except that the first element of the first row and last element of the last row of I^* are zero. H has elements each equal to $\frac{1}{2}$ in the $2(T-1)$ positions immediately adjacent to the main diagonal and zeros elsewhere. $|B| = 1 - \rho^2$.

Partial derivatives of L with respect to the unknown parameters are

$$(5) \quad \frac{\partial L}{\partial \beta} = \frac{1}{v} X' B (y - X\beta)$$

$$(6) \quad \frac{\partial L}{\partial \rho} = \frac{-2\rho}{1-\rho^2} - \frac{1}{v} (y - X\beta)' (\rho I^* - H) (y - X\beta)$$

$$(7) \quad \frac{\partial L}{\partial v} = \frac{-T}{2v} + \frac{1}{2v^3} (y - X\beta)' B (y - X\beta)$$

The normal equations obtained by equating these derivatives to zero are highly nonlinear so it has generally been found expedient to obtain maximum likelihood estimates by iterative procedures.

The normal equations may be written

$$(8) \quad \hat{\beta} = (X' \hat{B} X)^{-1} X' \hat{B} y$$

$$(9) \quad -\hat{\beta} \hat{v} - (1 - \hat{\rho}^2) (y - X\hat{\beta})' (\hat{\rho} I^* - H) (y - X\hat{\beta}) = 0$$

$$(10) \quad \hat{v} = \frac{1}{T} (y - X\hat{\beta})' \hat{B} (y - X\hat{\beta})$$

where $\hat{B} = I + \hat{\rho}^2 I^* - 2\hat{\rho}H$.

If $\hat{\rho}$ were somehow obtained, calculation of $\hat{\beta}$, \hat{v} from (8), (10) would correspond to well-known generalized least squares procedures.

Equations analogous to (8), (10) say

$$(8^*) \quad \beta_{\rho} = (X' B_{\rho} X)^{-1} X' B_{\rho} y$$

$$(10^*) \quad v_{\rho} = \frac{1}{T} (y - X\beta_{\rho})' B_{\rho} (y - X\beta_{\rho})$$

could be used to compute the maximum of L corresponding to any assumed value of ρ . The procedure used in this study (more fully explained in Section 5) is to substitute β_{ρ} , v_{ρ} for β , v in (4) obtaining a concentrated likelihood function

$$(11) \quad L^*(\rho) = \text{const} - \frac{T}{2} \log v_\rho + \frac{1}{2} \log (1 - \rho^2)$$

which is then maximized with respect to ρ . This maximizing value¹ is the ML estimator $\hat{\rho}$ of ρ and may then be used in (8), (10) or (8*), (10*) to obtain $\hat{\beta} = \beta_{\hat{\rho}}$, $\hat{v} = v_{\hat{\rho}}$. It was found convenient to note that L^* is a monotonic decreasing function of

$$(12) \quad S(\rho) = (1 - \rho^2)^{-\frac{1}{T}} (y - X\beta_\rho)' B_\rho (y - X\beta_\rho)$$

so $\hat{\rho}$ may be found by minimizing $S(\rho)$.²

Under mild assumptions about the behavior of X as T increases ([9], p. 584) it has been shown that the ML estimators $\hat{\beta}$, $\hat{\rho}$, \hat{v} are asymptotically independent and their joint asymptotic distribution is multivariate normal with means equal to the true parameter values and respective variances $v(X'BX)^{-1}$, $\frac{1 - \rho^2}{T}$, $\frac{2v^2}{T}$.

An extensive Monte Carlo study [8] was undertaken to compare ML estimators with others and to obtain hints about which properties of the asymptotic distribution might be approximately realized in moderate sized samples ($T = 30$ and 100 were considered). The results for $\hat{\rho}$ along with some subsequent trials are reported in Section 3. Although ML estimators compared favorably with the others considered, $\hat{\rho}$ did seem to be generally biased (see Table 3, pp. 20-22) and this suggested the possibility

1. The likelihood function may have multiple maxima though these have rarely occurred in practice. The scanning procedure described in Section 5 is designed to give substantial protection against the possibility of being led to a local maximum that is not global.

2. The likelihood function and the resulting estimators differ slightly according to the assumption made about the distribution of the initial observation y_1 . See Zellner and Tiao [17] for a discussion of this point. In their 1960 study of demand relations [10], Hildreth and Lu treated y_1 as a fixed number. In [8], the method described in this section was applied though the expression for $S(\rho)$ was incorrectly stated. The factor $(1 - \rho^2)^{-\frac{1}{T}}$ was omitted [8, p. 35, Equation (13)]. However, the correct form was used in making calculations.

that if an approximation to the bias could be found it might be used to construct a more accurate estimator. In the absence of a specific utility function, mean square error has been used as the indicator of accuracy.

2. The Adjusted Estimator

Finding a useful approximation to the bias means finding an approximate mean of $\hat{\beta}$ that is substantially closer than the asymptotic mean for sample sizes commonly encountered. The approximation to $E \hat{\beta}$ developed below involves a succession of simplifying alterations whose effects are not checked individually but whose combined effects are checked by using the approximation to adjust the ML estimator and then checking the effect of the adjustment on calculations from Monte Carlo data. It is seen from Table 2, pp. 16, 17, that about 2/3 of the bias in $\hat{\beta}$ was removed by this adjustment.

To examine the adjustment, consider Equation (9), p. 3 rewritten as

$$(9^*) \quad \frac{-\hat{\beta}}{(1-\hat{\beta}^2)^T} (\hat{u}'\hat{u} + \hat{\beta}^2 \hat{u}'I^*\hat{u} - 2\hat{\beta} \hat{u}'H\hat{u}) - \hat{\beta} \hat{u}'I^*\hat{u} + \hat{u}'H\hat{u} = 0$$

where $y - X\hat{\beta}$ has been written \hat{u} and the value of $\hat{\beta}$ given in Equation (10) has been substituted. Some simplification is achieved by approximating $\hat{u}'I^*\hat{u} = \sum_{t=2}^{T-1} \hat{u}_t^2$ by $\hat{u}'\hat{u} = \sum_{t=1}^T \hat{u}_t^2$, yielding

$$(13) \quad \frac{-\hat{\beta}}{T(1-\hat{\beta}^2)} [(1 + \hat{\beta}^2) \hat{u}'\hat{u} - 2\hat{\beta} \hat{u}'H\hat{u}] - \hat{\beta} \hat{u}'\hat{u} + \hat{u}'H\hat{u} \approx 0$$

where " \approx " is read "may be approximately equal to".

As T grows large, the term with the factor $\frac{1}{T}$ grows less important and $\hat{\beta}$ is determined primarily by the two latter terms. This can be made more apparent by rewriting the approximation

$$(14) \quad \hat{\rho} \approx \frac{\hat{u}'H\hat{u}}{\hat{u}'\hat{u}} + \frac{2\hat{\rho}^2}{T(1-\hat{\rho}^2)} \frac{\hat{u}'H\hat{u}}{\hat{u}'\hat{u}} - \frac{\hat{\rho}(1+\hat{\rho}^2)}{T(1-\hat{\rho}^2)} \quad \text{or}$$

$$(15) \quad \hat{\rho} \left(1 + \frac{1+\hat{\rho}^2}{T(1-\hat{\rho}^2)} \right) \approx \frac{\hat{u}'H\hat{u}}{\hat{u}'\hat{u}} \left(1 + \frac{2\hat{\rho}^2}{T(1-\hat{\rho}^2)} \right)$$

which may be restated

$$(16) \quad \hat{\rho} \approx \frac{\hat{u}'H\hat{u}}{\hat{u}'\hat{u}} \left(\frac{T(1-\hat{\rho}^2) + 2\hat{\rho}^2}{T(1-\hat{\rho}^2) + 1 + \hat{\rho}^2} \right)$$

For $-1 < \hat{\rho} < 1$, the factor in brackets lies between $\frac{T}{T+1}$ and 1. The minimum of the factor is attained at $\hat{\rho} = 0$ and, for moderate or large T , the value of this factor remains close to $\frac{T}{T+1}$ over much of the interval $(-1, 1)$. For example, if $\hat{\rho} = \pm .6$ the factor is $\frac{T+1.12}{T+1.12+1}$. In general it is $\frac{T+\alpha}{T+\alpha+1}$ where $\alpha = \frac{2\hat{\rho}^2}{1-\hat{\rho}^2}$.

It thus seems reasonable to make a further approximation by setting this factor equal to $\frac{T}{T+1}$. Let $G = \frac{T}{T+1} H$. Then

$$(17) \quad \hat{\rho} \approx \frac{\hat{u}'G\hat{u}}{\hat{u}'\hat{u}}$$

Approximation (17) represents considerable simplification, but it appears that direct evaluation of the mean of $\frac{\hat{u}'G\hat{u}}{\hat{u}'\hat{u}}$ would be difficult because little is known about the distribution of \hat{u} . A further approximation is suggested by considering the coefficient estimator and residual vector for an autoregressive model with known ρ .

Let

$$(18) \quad \tilde{\beta} = (X'BX)^{-1}X'By$$

be the ML estimator (also generalized least squares and best unbiased) for the case of known ρ . Then the residual vector

$$(19) \quad \tilde{u} = y - X \tilde{\beta} = M y = M u$$

is distributed according to the (singular) multivariate normal law with mean zero and variance

$$(20) \quad E\tilde{u}\tilde{u}' = \nu M A M' = \nu M A = \nu (A - X(X' B X)^{-1} X')$$

where

$$(21) \quad M = I - X(X' B X)^{-1} X' B$$

In connection with the proof that $\hat{\beta}$ has the same asymptotic distributions as $\tilde{\beta}$, it was conjectured [9, Section 5] that the distributions of these vectors tend to be approximately equal for moderate sample sizes and this tended to be confirmed by subsequent Monte Carlo trials [8, p. 26]. Thus one might hope that $\hat{u} = y - X \hat{\beta}$ is distributed not too differently from $\tilde{u} = y - X \tilde{\beta}$, or at least that

$$(22) \quad E \frac{\hat{u}' G \hat{u}}{\hat{u}' \hat{u}} \approx E \frac{\tilde{u}' G \tilde{u}}{\tilde{u}' \tilde{u}} .$$

To investigate the expectation on the right, let R be an orthogonal matrix that diagonalizes MA . For convenience, suppose the rows of R are chosen so that

$$(23) \quad R M A R' = \begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}$$

where P is a matrix of order $T-K$ with positive diagonal elements, P_1, P_2, \dots, P_{T-K} ($R M A R'$ is known to be of rank $T-K$ since A is nonsingular and M can readily be shown to be of rank $T-K$).

Define

$$(24) \quad z = R \tilde{u} = \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} \tilde{u} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

where R_1 comprises the first $T-K$ rows of R . It follows that z_1 is multivariate normal with mean zero, variance νP ; $z_2 = 0$; $\tilde{u}'\tilde{u} = z'RR'z = z'z = z_1'z_1$; $\tilde{u}'G\tilde{u} = z'RGR'z = z_1'R_1GR_1'z_1$. Therefore

$$(25) \quad \frac{\tilde{u}'G\tilde{u}}{\tilde{u}'\tilde{u}} = \frac{z_1'R_1GR_1'z_1}{z_1'z_1} .$$

Now consider

$$(26) \quad w = \nu^{-\frac{1}{2}} P^{-\frac{1}{2}} z_1 .$$

w is standard multivariate normal and

$$(27) \quad \frac{\tilde{u}'G\tilde{u}}{\tilde{u}'\tilde{u}} = \frac{w' P^{\frac{1}{2}} R_1 G R_1' P^{\frac{1}{2}} w}{w' P w} = \frac{w' C w}{w' P w}$$

where C is defined by the second equality. Let $Y_{st} = \frac{w_s w_t}{\sum_{r=1}^{T-K} P_r w_r^2}$ for $s, t = 1, 2, \dots, T-K$. Then

$$(28) \quad E \frac{w' C w}{w' P w} = E \sum_{t=1}^{T-K} \sum_{s=1}^{T-K} c_{st} Y_{st} = \sum_{t=1}^{T-K} c_{tt} E Y_{tt}$$

where the second equality follows from the fact that the density of Y_{st} for $s \neq t$ is symmetric about zero and therefore has mean zero.

To evaluate $E Y_{tt}$, write

$$(29) \quad Y_{tt} = \frac{w_t^2}{\sum_{r=1}^{T-K} P_r w_r^2} = \frac{k_1}{k_2}$$

Let

$$(30) \quad \psi(h_1, h_2) = E e^{h_1 k_1 + h_2 k_2} = \prod_{r=1}^{T-K} (1 - h_1 \delta_{tr} - h_2 P_r)^{-\frac{1}{2}}$$

be the joint moment generating function of k_1 and k_2 where δ_{tr} is the Kronecker delta.

By a theorem of Anderson and Anderson [1, p. 77]

$$(31) \quad E \frac{k_1}{k_2} = \int_{-\infty}^0 \left[\frac{\partial \psi}{\partial h_1} \right]_{h_1=0} dh_2 = \frac{1}{2} \int_{-\infty}^0 (1 - p_t h)^{-1} \prod_{r=1}^{T-K} (1 - p_r h)^{-\frac{1}{2}} dh$$

This is a difficult integral and obtaining the p_r can also involve substantial computation. Both tasks can be greatly simplified at the cost of still another approximation. Suppose the average of the p_r , call it p , is substituted for each individual element.³ Then $p = \frac{1}{T-K}$

tr MA and

$$(32) \quad E Y_{tt} = E \frac{k_1}{k_2} \approx \frac{1}{2} \int_0^{\infty} (1 - p h)^{-\frac{T-K+2}{2}} dh = \frac{1}{p (T-K)}$$

Substitution in (27), (26) yields

$$(33) \quad E \frac{\tilde{u}' G \tilde{u}}{\tilde{u}' \tilde{u}} = E \frac{w' C w}{w' P w} \approx \frac{\sum_{t=1}^{T-K} c_{tt}}{p (T-K)} = \frac{\text{tr } C}{\text{tr } MA}$$

which can be readily calculated for any value of p . Let $\frac{\text{tr } C}{\text{tr } MA} = \varphi(p)$.

If the various approximations sketched are not too gross, then it will be true that

$$(34) \quad E (\hat{\rho}) \approx \frac{\text{tr } C}{\text{tr } MA} = \varphi(\rho) .$$

It is seen in Section 4 that $\varphi(\rho)$ is approximately linear. Thus, if $E \hat{\rho}$ is sufficiently well approximated by $\varphi(\rho)$, then $\hat{\rho}^* = \varphi^{-1}(\hat{\rho})$ will have smaller bias than $\hat{\rho}$. For the structures considered in the next section $\hat{\rho}^*$ tends to have about 1/3 the bias of $\hat{\rho}$ though this varies

3. Dent [4, pp. 34-36] has also considered the possibility of making the p_r equal in pairs. Finding the mean of the ratio then involves finding the eigen values of MA. This should provide a somewhat more accurate approximation to the mean of $\hat{\rho}$, but at the cost of substantially heavier computations.

substantially from one structure to another. Unfortunately, the variance of $\hat{\rho}^*$ is typically larger than the variance of $\hat{\rho}$ so $\hat{\rho}^*$ does not always have lower mean square error. These comparisons and possible improvements are discussed after the results have been presented. In Section 5 it is explained how calculation of $\hat{\rho}^*$ may be conveniently combined with calculation of $\hat{\rho}$.

3. The Monte Carlo Trials

Because of formidable mathematical difficulties in determining properties of the distributions of various estimators that have been used with the autoregressive disturbance model, it was decided to compare the behavior of the adjusted maximum likelihood (AML) estimator with the ML estimator and others when applied to artificially generated samples with known parameters.

To make the hints furnished by such comparisons as useful as possible, a number of structures were chosen, 22 in all, representing a variety of circumstances that might be encountered in applications. To enhance the statistical reliability of the results a relatively large number, 300, of samples were generated for each structure.

Four estimates of ρ were computed for each sample. These included ML, AML, an estimator developed by Theil and Nagar [16] denoted in this report by TN, and an estimator suggested by Durbin [5] denoted D.⁴

The TN estimator is

4. [8] also considered the approximate Bayes estimator suggested by Zellner and Tiao [17, p. 776]. It turned out to be numerically very close to ML and to have, on the average, slightly larger mean square error so it was omitted from the present study.

$$(35) \quad \rho_{TN} = \frac{T^2 (1 - \frac{1}{2}d) + K^2}{T^2 - K^2}$$

where $d = \frac{\sum_{t=2}^T (\bar{u}_t - \bar{u}_{t-1})^2}{\sum_{t=1}^T \bar{u}_t^2}$ is the Durbin-Watson statistic calculated for

least squares residuals \bar{u}_t of the regression of y on X . The Durbin estimator is the coefficient of the lagged endogenous variable y_{t-1} in the least squares regression of y_t on y_{t-1} ; x_{tk} $k = 1, 2, \dots, K$; and $x_{(t-1)k}$ $k = 2, 3, \dots, K$ where it is assumed that $x_{t1} = 1$ $t = 1, 2, \dots, T$.

A structure is specified by choosing a design matrix X of observed values of independent variables and specific values for the parameters ρ , β , ν . Table 1 shows some properties of the structures used in this study. $x_{t1} = 1$ $t = 1, 2, \dots, T$ for every structure; thus β_1 is the constant term in each case. Structures 1 to 8 were used by Hildreth and Lu [8] and 9-12 were obtained by modifying some parts of these structures. In all of these $\beta_1 = 0$, $\beta_k = 1$ for $k \neq 1$, $\nu = 1$. Other characteristics are shown in Table 1A. Each independent variable for Structures 1-3 was generated by adding a random component to a harmonic term of low frequency. For Structures 4-6, sums of several harmonic terms including terms of high frequency were employed. Further details are given in [8, pp. 3-6, pp. 33-4]. This arrangement should make Structures 1-3 relatively favorable for least squares estimation of β and for tests based on the Durbin-Watson upper limits (see Chipman [2]).

Observations of independent variables for Structure 9 were obtained by taking the last 60 observations of corresponding variables

TABLE 1A

Some Characteristics of Structures 1 - 12*Properties of Independent Variables

| Structure Number | ρ | T | <u>Means</u> | | | <u>Variances</u> | | | <u>Correlation Coefficients</u> | | |
|---------------------|--------|-----|--------------|----------|----------|------------------|----------|----------|---------------------------------|-----------------|-----------------|
| | | | x_{t2} | x_{t3} | x_{t4} | x_{t2} | x_{t3} | x_{t4} | $x_{t2} x_{t3}$ | $x_{t2} x_{t4}$ | $x_{t3} x_{t4}$ |
| 1 | .3 | 30 | .064 | .015 | .015 | .466 | .711 | .460 | .113 | .026 | .048 |
| 2 | 0 | 30 | -.061 | -.024 | -.024 | .539 | .616 | .477 | .131 | -.153 | .089 |
| 3 | -.7 | 100 | .006 | -.007 | -.007 | .553 | .578 | .606 | .028 | -.040 | .089 |
| 4 | .7 | 30 | .000 | .025 | .025 | .750 | .727 | .736 | .010 | -.349 | .218 |
| 5 | .3 | 100 | .000 | .020 | .020 | .750 | .750 | .750 | .015 | .012 | -.025 |
| 6 | 0 | 100 | .000 | .020 | .020 | .750 | .750 | .750 | .015 | .012 | -.025 |
| 7 | .5 | 30 | .000 | .000 | .000 | .750 | .750 | .750 | .937 | .698 | .704 |
| 8 | .9 | 100 | .000 | .001 | .000 | .750 | .748 | .750 | .798 | .293 | .108 |
| 9 | -.9 | 60 | .100 | .000 | -.020 | .577 | .590 | .597 | .063 | -.143 | .157 |
| 10 | -.5 | 60 | .245 | -.012 | -.155 | .690 | .616 | .555 | .048 | .028 | -.058 |
| 11 | .4 | 60 | -.124 | -.130 | .045 | .932 | .563 | .797 | -.125 | -.220 | -.168 |
| 12 | .6 | 60 | .157 | 0.026 | -.504 | .526 | .572 | .623 | -.045 | -.014 | -.021 |

* For all of these structures $\beta_1 = 0$, $\beta_2 = \beta_3 = \beta_4 = 1$, $v = 1$.

for Structure 3. Observations for Structure 12 are the first 60 observations for Structure 3. The design matrix for Structure 10 was obtained by adding small random terms to the first 60 rows of the matrix for Structure 5 (except for the first column which has unit elements in all structures) and multiplying the sum by a constant designed to approximately preserve sample variances. The design matrix for Structure 11 was obtained similarly except that the last 60 rows of the matrix for Structure 5 were used. The independent variables for Structures 7 and 8 are series for the United States wholesale price index, numbers of immigrants to the United States, and exports of foodstuffs from the United States for the years 1928-57 (7) and 1858-1957 (8). The series were coded to produce zero means and sample variances less than unity.

The purpose of using observed series for Structures 7 and 8 was to insure that properties which might be encountered in applications would be reflected in the structures used for Monte Carlo trials. This notion was carried further in choosing Structures 13-22. Structures 13 and 14 use observations of independent variables from Prest's [14] study of demand for textiles. The data are tabulated in Theil and Nagar [16, p. 805]. For Structure 13, β , ρ , ν are one-digit approximations to the ML estimates. For Structure 14, a different value of ρ was used.

Similarly, Structures 15 and 16 come from Hoos and Shear [11], and Structures 17 and 18 from Linstrom and King [13]. The Hoos and Shear data are tabulated in Henshaw [7] and the Linstrom-King data in Hildreth and Lu [10, p. 70].⁵

5. There are errors in the Hoos-Shear data presented in [10]. These have been corrected in [7].

Structures 19 and 20 are based on Klein's [12, p. 135] simplified consumption function and Structures 21 and 22 use the data for a wage-rate equation in the Federal Reserve-MIT Econometric Model.⁶ In all cases observed values of independent variables were used with approximate ML estimates of β , ν . Two structures were specified from each study by using the approximate ML estimate of ρ for one structure and a value chosen to help insure wide selection for the other. Some characteristics of Structures 13-22 are shown in Table 1B.

For each structure described above, 300 samples were generated by drawing values of v_t from a table of random numbers and then calculating u_t , y_t from equations (1), (2a), (2b) pages 1, 2. For each sample, estimates of ρ were computed by each of the four procedures -- TN, D, ML, AML.

The main results are summarized in Table 2. Each cell corresponding to a particular structure and estimator contains three entries. The first is the calculated bias of the estimator when applied to the indicated structure. It is obtained by subtracting the true value of ρ from the mean of the 300 values estimated by the indicated procedure. The second entry is the sample variance of the estimator for the 300 trials and the third is the calculated mean square error ($\text{bias}^2 + \text{variance}$). Structures are arranged according to increasing values of the autocorrelation coefficient since this simplifies examination of properties sensitive to the coefficient.

6. The equation is No. 96 in the Appendix to reports by de Leeuw and Gramlich. See [3], p. 266.

TABLE 2

Monte Carlo Bias, Variance, and Mean
Square Error, for Various Estimators of ρ

| <u>Structural Characteristics</u> | | | | <u>Estimators</u> | | | |
|-----------------------------------|-------------------------------|----------|----------|-------------------|----------|-----------|------------|
| <u>Structure Number</u> | <u>True ρ</u> | <u>K</u> | <u>T</u> | <u>TN</u> | <u>D</u> | <u>ML</u> | <u>AML</u> |
| 9 | -.9 | 4 | 60 | .0340 | .0239 | .0244 | -.0018 |
| | | | | .0051 | .0047 | .0039 | .0045 |
| | | | | .0062 | .0053 | .0045 | .0045 |
| 17 | -.8 | 3 | 17 | .2348 | .0266 | .0101 | -.0638 |
| | | | | .0248 | .0302 | .0151 | .0280 |
| | | | | .0799 | .0309 | .0152 | .0320 |
| 3 | -.7 | 4 | 100 | .0285 | .0058 | .0042 | .0022 |
| | | | | .0059 | .0058 | .0055 | .0062 |
| | | | | .0067 | .0058 | .0055 | .0062 |
| 13 | -.5 | 3 | 17 | .0820 | -.0630 | -.0470 | -.0246 |
| | | | | .0481 | .0491 | .0411 | .0789 |
| | | | | .0548 | .0531 | .0433 | .0795 |
| 10 | -.5 | 4 | 60 | .0279 | .0084 | -.0062 | .0055 |
| | | | | .0124 | .0127 | .0117 | .0135 |
| | | | | .0132 | .0127 | .0117 | .0135 |
| 18 | -.4 | 3 | 17 | .0978 | -.0432 | -.0625 | -.0290 |
| | | | | .0384 | .0485 | .0432 | .0694 |
| | | | | .0480 | .0503 | .0471 | .0702 |
| 14 | -.1 | 3 | 17 | -.0280 | -.1775 | -.1419 | -.0046 |
| | | | | .0467 | .0608 | .0526 | .0964 |
| | | | | .0475 | .0923 | .0728 | .0964 |
| 15 | 0 | 5 | 16 | -.0150 | -.1325 | -.2112 | -.1061 |
| | | | | .0707 | .0816 | .0875 | .1487 |
| | | | | .0707 | .0991 | .1324 | .1600 |
| 2 | 0 | 4 | 30 | -.0618 | -.1360 | -.1199 | -.0252 |
| | | | | .0341 | .0369 | .0388 | .0479 |
| | | | | .0379 | .0554 | .0532 | .0486 |
| 6 | 0 | 4 | 100 | -.0264 | -.0341 | -.0389 | -.0106 |
| | | | | .0088 | .0095 | .0098 | .0103 |
| | | | | .0095 | .0106 | .0113 | .0104 |

Table 2 (Continued)

| Structural Characteristics | | | | Estimators | | | |
|----------------------------|-------------|---|-----|------------|--------|--------|--------|
| Structure Number | True ρ | K | T | TN | D | ML | AML |
| 19 | .3 | 3 | 22 | -.1364 | -.2142 | -.1802 | -.0287 |
| | | | | .0407 | .0558 | .0491 | .0606 |
| | | | | .0593 | .1017 | .0816 | .0614 |
| 1 | .3 | 4 | 30 | -.1063 | -.1920 | -.1472 | -.0260 |
| | | | | .0348 | .0419 | .0406 | .0443 |
| | | | | .0461 | .0787 | .0622 | .0450 |
| 5 | .3 | 4 | 100 | -.0369 | -.0359 | -.0298 | .0017 |
| | | | | .0104 | .0106 | .0119 | .0119 |
| | | | | .0118 | .0119 | .0128 | .0119 |
| 16 | .4 | 5 | 16 | -.1687 | -.3320 | -.3349 | -.1534 |
| | | | | .0627 | .0768 | .0912 | .1415 |
| | | | | .0912 | .1870 | .2033 | .1650 |
| 11 | .4 | 4 | 60 | -.0778 | -.0621 | -.0391 | -.0039 |
| | | | | .0147 | .0157 | .0172 | .0174 |
| | | | | .0207 | .0196 | .0187 | .0174 |
| 7 | .5 | 4 | 30 | -.1909 | -.2337 | -.2024 | -.0622 |
| | | | | .0320 | .0484 | .0423 | .0480 |
| | | | | .0684 | .1030 | .0833 | .0518 |
| 20 | .6 | 3 | 22 | -.2455 | -.3194 | -.2652 | -.1022 |
| | | | | .0485 | .0655 | .0616 | .0680 |
| | | | | .1088 | .1676 | .1319 | .0784 |
| 12 | .6 | 4 | 60 | -.1168 | -.1146 | -.0859 | -.0199 |
| | | | | .0150 | .0143 | .0156 | .0147 |
| | | | | .0286 | .0275 | .0230 | .0151 |
| 4 | .7 | 4 | 30 | -.1881 | -.1454 | -.0874 | -.0107 |
| | | | | .0264 | .0363 | .0322 | .0403 |
| | | | | .0617 | .0574 | .0398 | .0404 |
| 21 | .7 | 4 | 56 | -.1008 | -.1177 | -.0813 | -.0117 |
| | | | | .0143 | .0187 | .0162 | .0157 |
| | | | | .0244 | .0325 | .0228 | .0159 |
| 22 | .9 | 4 | 56 | -.1229 | -.1182 | -.0826 | -.0152 |
| | | | | .0108 | .0138 | .0103 | .0102 |
| | | | | .0259 | .0277 | .0171 | .0104 |
| 8 | .9 | 4 | 100 | -.1033 | -.0774 | -.0304 | .0081 |
| | | | | .0053 | .0050 | .0048 | .0047 |
| | | | | .0160 | .0110 | .0057 | .0048 |

4. Examination of Results

Looking first at the results for bias, it is seen that the adjustment was surprisingly successful considering the crudeness of some of the approximations on which it is based. AML bias is lowest (in absolute value) among the four estimators in 20 of the 22 structures (15 and 17 are the exceptions). It is lower than ML bias in all structures except Structure 17. For the 12 structures with positive ρ , AML bias is less than 1/3 of ML bias in 10.

Among the three unadjusted estimators, TN tends to have the lowest bias if ρ is zero or a small positive fraction; ML tends to have the smallest bias if $|\rho|$ is large. ML seems to improve relative to TN as sample size increases. D is frequently in between.

Results of approximate tests for bias are given in Table 3. Consider testing the null hypothesis that a particular estimator, call it $\tilde{\rho}$, for a chosen structure is unbiased, i.e. $H_0: E\tilde{\rho} = \rho$. Let $\tilde{\rho}_\alpha$ be the estimate obtained by the method in question from the α th sample. Let

$$(36) \quad \begin{aligned} \tilde{\rho}_M &= \frac{1}{300} \sum_{\alpha=1}^{300} \tilde{\rho}_\alpha \\ \tilde{\rho}_V &= \frac{1}{300} \sum_{\alpha=1}^{300} (\tilde{\rho}_\alpha - \tilde{\rho}_M)^2 \end{aligned}$$

be the sample mean and variance. From the Central Limit Theorem, the distribution of the statistic

$$(37) \quad \bar{\xi} = (\tilde{\rho}_M - \rho) \tilde{\rho}_V^{-\frac{1}{2}} (300)^{\frac{1}{2}}$$

approaches the standard normal. The value of this statistic for each

structure-estimator combination is the upper entry in the appropriate cell of Table 3. The lower entry is the marginal significance level, the approximate probability under the null hypothesis that a random value of the statistic will differ from zero by more than the calculated value.

The test emphatically (significance level .005) rejects the hypothesis of unbiasedness of TN for all structures and of D for all structures except 3 and 10. For the ML estimator, marginal significance levels above .005 are obtained for structures 3, 10, 13, 17. These are all cases in which the true value of ρ is in the vicinity of $-.6$ and it will be seen from Table 4 and Figure 1 (pp. 24, 25) that, for all of the structures considered here, $\varphi(\rho)$ tends to be close to ρ in this region.

The marginal significance level for AML is above .005 for all structures except 20 and above .05 for 14 of the 22 structures. This tends to confirm the impression obtained from Table 2 that the adjustment applied here removes much, but not all, of the bias in the ML estimator.

Variance is another matter. TN variance is lowest in 14 of the 22 cases including 9 of 12 in structures with positive ρ . This suggests that it may be worthwhile to study the distribution of the TN estimator more carefully. An adjustment that would reduce TN bias without materially increasing variance might produce typically smaller mean square error than other available estimators. ML tends to have lower variances than TN and D when $|\rho|$ is large and again shows relative improvement for large T.

TABLE 3

Test Statistics and Marginal Significance Levels for Tests of Bias

| Structure | True ρ | T | TN | D | ML | AML |
|-----------|-------------|-----|-------|--------|--------|-------|
| 9 | -.9 | 60 | 8.24 | 6.03 | 6.75 | 0.40 |
| | | | * | * | * | .69 |
| 17 | -.8 | 17 | 25.77 | 2.65 | 1.42 | -6.60 |
| | | | * | .01 | .16 | * |
| 3 | -.7 | 100 | 6.41 | 1.33 | 0.98 | 0.48 |
| | | | * | .18 | .33 | .63 |
| 13 | -.5 | 17 | 6.17 | -3.91 | -2.64 | 1.64 |
| | | | * | * | .01 | .10 |
| 10 | -.5 | 60 | 4.34 | 1.29 | -0.99 | 0.87 |
| | | | * | .20 | .32 | .38 |
| 18 | -.4 | 17 | 8.63 | -3.39 | -5.20 | -1.90 |
| | | | * | * | * | .06 |
| 14 | -.1 | 17 | -6.82 | -10.73 | -11.01 | -0.30 |
| | | | * | * | * | .76 |
| 15 | .0 | 16 | -0.97 | -8.02 | -12.38 | -4.76 |
| | | | .33 | * | * | * |
| 2 | .0 | 30 | -5.79 | -12.24 | -10.52 | -1.99 |
| | | | * | * | * | .05 |

* indicates a marginal significance level below .005.

TABLE 3 (Continued)

| Structure | True ρ | T | TN | D | ML | AML |
|-----------|-------------|-----|--------|--------|--------|-------|
| 6 | .0 | 100 | -4.87 | -6.06 | -6.79 | -1.81 |
| | | | * | * | * | .07 |
| 19 | .3 | 22 | -11.69 | -15.68 | -14.07 | -2.02 |
| | | | * | * | * | .04 |
| 1 | .3 | 30 | -9.84 | -16.23 | -12.64 | -2.14 |
| | | | * | * | * | .03 |
| 5 | .3 | 100 | -6.25 | -6.02 | -4.73 | 0.27 |
| | | | * | * | * | .78 |
| 16 | .4 | 16 | -11.65 | -20.72 | -19.18 | -7.06 |
| | | | * | * | * | * |
| 11 | .4 | 60 | -11.10 | -8.58 | -5.15 | -0.51 |
| | | | * | * | * | .61 |
| 7 | .5 | 30 | -18.46 | -18.36 | -17.01 | -4.91 |
| | | | * | * | * | * |
| 20 | .6 | 22 | -19.26 | -21.57 | -18.48 | -6.78 |
| | | | * | * | * | * |
| 12 | .6 | 60 | -16.52 | -16.56 | -11.88 | -2.84 |
| | | | * | * | * | .01 |
| 4 | .7 | 30 | -20.04 | -13.20 | -8.43 | -0.98 |
| | | | * | * | * | .33 |

TABLE 3 (Continued)

| <u>Structure</u> | <u>True ρ</u> | <u>T</u> | <u>TN</u> | <u>D</u> | <u>ML</u> | <u>AML</u> |
|------------------|-------------------------------|----------|-----------|----------|-----------|------------|
| 21 | .7 | 56 | -14.59 | -14.90 | -11.06 | -1.62 |
| | | | * | * | * | .11 |
| 22 | .9 | 56 | -20.48 | -17.43 | -14.11 | -2.61 |
| | | | * | * | * | .01 |
| 8 | .9 | 100 | -24.48 | -18.91 | -7.58 | 1.75 |
| | | | * | * | * | .08 |

For reasons that will become clearer after Table 4 and Figure 1 are inspected, variance of the AML estimator is usually larger than that of ML (18 of 22 structures).

Turning to relations among mean square errors of the various estimators, the true value of ρ makes a striking difference. ML has lowest mse for the six structures with $\rho < -.1$. TN is lowest for 7 of the 9 structures for which $-.1 \leq \rho \leq .4$. AML is lowest for six of the seven structures for which $\rho > .4$ and eight of twelve where ρ is positive.

Comparing mse for AML and ML shows that the adjustment fails to improve the mse in every case in which $\rho < 0$ and does improve the estimator for eleven of the twelve structures for which $\rho > 0$. Various comparisons that have been made between AML and ML can usefully be related to the typical behavior of the approximate mean $\varphi(\rho)$.

Some properties of $\varphi(\rho)$ for each structure are given in Table 4 and graphs for two illustrative cases are presented in Figure 1. Certain common features are apparent.

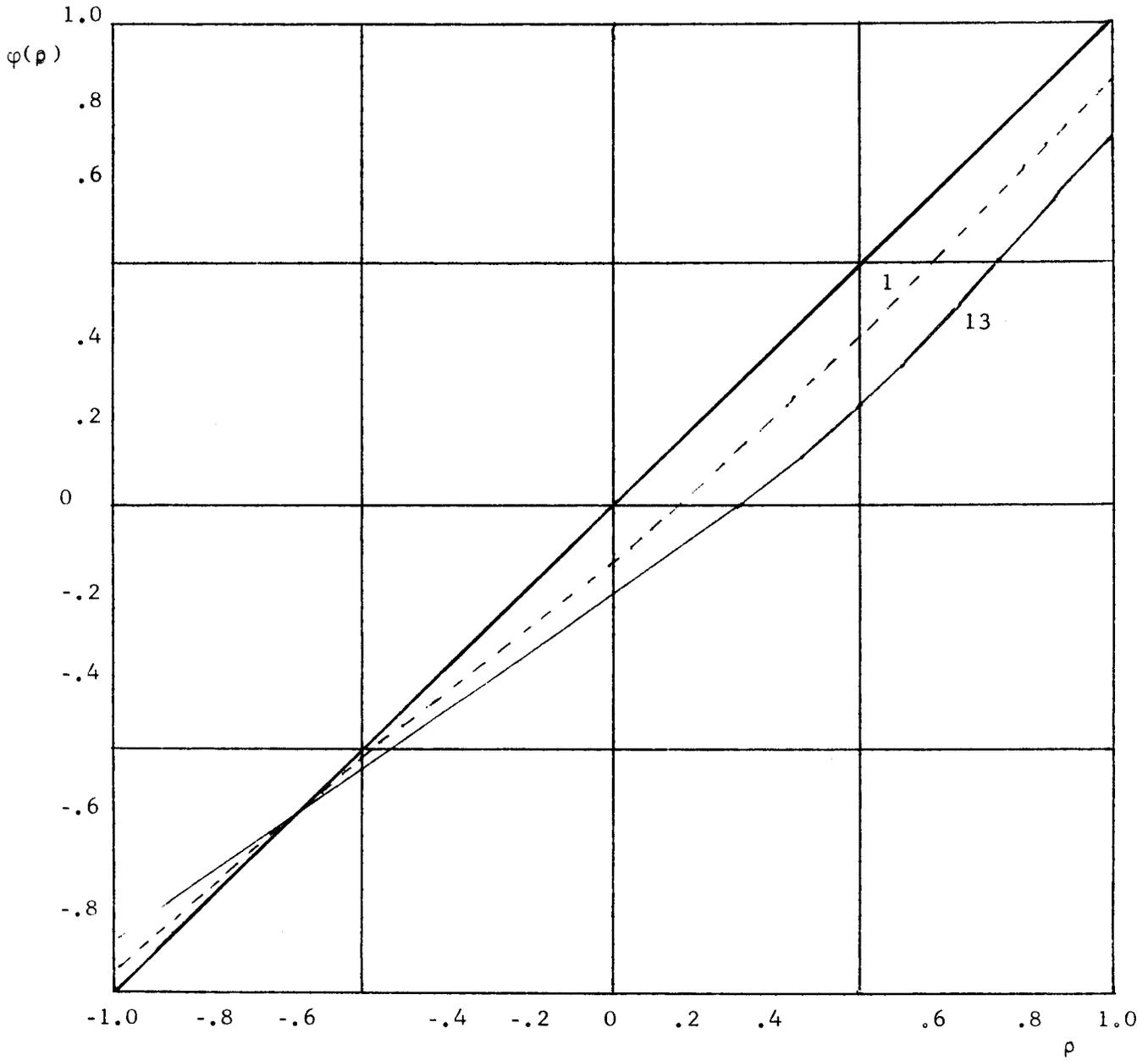
For each structure, $\varphi(\rho)$ has a fixed point ($\varphi(\rho) = \rho$) between $-.710$ and $-.560$. $\varphi(\rho) > \rho$ (indicating probably positive bias) to the left of the fixed point and $\varphi(\rho) < \rho$ to the right of the fixed point. $\varphi'(\rho)$ is predominately, but not universally, less than one and tends to increase with ρ , but not monotonically.

The most striking difference among functions determined by different structures is an approximate scale factor that may roughly be associated with the extreme difference between $\varphi(\rho)$ and ρ . Define the extreme

TABLE 4
Some Properties of $\Psi(p)$ for Alternative Structures

| Structure Number | T | Fixed Point | Extreme Difference | Abcissa/Difference | Ordinates and Slopes for Selected Values of p | | | | | | |
|------------------|-----|-------------|--------------------|--------------------|---|--------------|--------------|---------------|--------------|--------------|--------------|
| | | | | | .9 | .6 | .3 | .0 | .3 | .6 | .9 |
| 1 | 30 | -.662 | .400 | -.130 | -.855 .80 | -.612 .81 | -.364 .84 | -.106 .89 | .171 .97 | .475 1.04 | .783 .94 |
| 2 | 30 | -.650 | .300 | -.121 | -.855 .80 | -.612 .81 | -.363 .84 | -.103 .90 | .179 .98 | .487 1.04 | .786 .91 |
| 3 | 100 | -.660 | .350 | -.043 | -.886 .94 | -.604 .94 | -.320 .95 | -.034 .96 | .258 .98 | .561 1.02 | .865 .97 |
| 4 | 30 | -.630 | .990 | -.122 | -.856 .60 | -.604 .87 | -.330 .95 | -.035 1.00 | .262 .96 | .539 .82 | .795 .82 |
| 5,6 | 100 | -.657 | .990 | -.042 | -.886 .73 | -.604 .94 | -.318 .95 | -.029 .98 | .268 1.00 | .571 1.00 | .866 .96 |
| 7 | 30 | -.662 | .720 | -.155 | -.855 .80 | -.612 .81 | -.364 .84 | -.105 .88 | .166 .92 | .448 .96 | .755 1.08 |
| 8 | 100 | -.660 | .460 | -.047 | -.886 .73 | -.604 .94 | -.321 .94 | -.035 .95 | .255 .98 | .555 1.01 | .863 1.00 |
| 9 | 60 | -.496 | .390 | -.073 | -.872 .93 | -.600 .92 | -.301 .91 | -.058 .85 | .230 .97 | .524 1.03 | .838 .97 |
| 10 | 60 | -.615 | .990 | -.065 | -.875 .90 | -.604 .92 | -.327 .94 | -.036 .97 | .264 1.00 | .560 .99 | .843 .91 |
| 11 | 60 | -.615 | .990 | -.065 | -.878 .90 | -.603 .92 | -.328 .94 | -.034 .98 | .268 1.00 | .561 .98 | .844 .91 |
| 12 | 60 | -.650 | .350 | -.072 | -.876 .89 | -.608 .90 | -.340 .91 | -.058 .96 | .221 .97 | .522 1.04 | .838 .95 |
| 13,14 | 17 | -.605 | .586 | -.277 | -.816 .72 | -.600 .72 | -.387 .72 | -.170 .73 | .058 .77 | .323 .96 | .628 1.09 |
| 15,16 | 16 | -.559 | .680 | -.261 | -.815 .67 | -.584 .80 | -.359 .74 | -.137 .74 | .093 .77 | .342 .88 | .655 1.23 |
| 17,18 | 17 | -.583 | .990 | -.220 | -.815 .72 | -.595 .74 | -.369 .78 | -.121 .84 | .145 .90 | .423 .94 | .604 .82 |
| 19,20 | 22 | -.611 | .480 | -.175 | -.833 .73 | -.597 .78 | -.366 .79 | -.122 .82 | .136 .88 | .426 1.02 | .743 .93 |
| 21,22 | 56 | -.673 | .390 | -.074 | -.873 .90 | -.603 .90 | -.336 .91 | -.058 .93 | .230 .97 | .527 1.01 | .833 .90 |

FIGURE 1



$\varphi(\rho)$ for Structures 1 and 13

difference as $\varphi(\rho_0) - \rho_0$ where $-.99 \leq \rho_0 \leq .99$ and ρ_0 maximizes $|\varphi(\rho) - \rho|$ in the interval $[-.99, .99]$. Extreme differences are tabulated in the fourth column of Table 3 along with the values at which the extrema occur.

Extreme difference is clearly associated with sample size. Let Δ represent the extreme difference. Then the following hold for these trials:

$$-.281 \leq \Delta \leq -.200 \quad \text{for} \quad 16 \leq T \leq 22$$

$$-.160 \leq \Delta \leq -.121 \quad \text{for} \quad T = 30$$

$$-.090 \leq \Delta \leq -.053 \quad \text{for} \quad T = 56 \quad \text{or} \quad 60$$

$$-.052 \leq \Delta \leq -.040 \quad \text{for} \quad T = 100 \quad .$$

Consider a linear approximation, $a + b\rho$, to $\varphi(\rho)$ in the vicinity of a ML estimate $\hat{\rho}$. Then the corresponding AML estimate, $\hat{\rho}^* = \varphi^{-1}(\hat{\rho})$, is approximately $-\frac{a}{b} + \frac{1}{b}\hat{\rho}$ and $\text{Var } \hat{\rho}^*$ is approximately $\frac{1}{b^2} \text{Var } \hat{\rho}$. Since b is approximately $\varphi'(\rho)$ and the latter is typically less than unity, it is to be expected that $\text{Var } \hat{\rho}^*$ is typically greater than $\text{Var } \hat{\rho}$.

From Table 4 and Figure 1 it is seen that the bias of $\hat{\rho}$ is relatively small when $\rho < 0$ and the discrepancy between $\text{Var } \hat{\rho}^*$ and $\text{Var } \hat{\rho}$ is relatively large since $\varphi'(\rho)$ is typically smaller than when $\rho > 0$. These relations both tend to make $\text{mse } \hat{\rho}$ small relative to $\text{mse } \hat{\rho}^*$ when $\rho < 0$ and the opposite tendencies usually prevail when $\rho > 0$. This suggests a possible rule of using ML when prior knowledge and/or tests indicate negative autocorrelation and AML when positive autocorrelation is indicated. Of course, it is hoped that further study will, in time, provide superior alternatives.

For developing better estimators, two principal avenues seem to be suggested by the trials reported here. One is to study the distribution of the TN estimator with the prospect that an adjusted estimator might be found which might have substantially smaller mean square error. The second avenue is continued study of the ML estimator. Since the adjustment obtained by concentrating on bias typically increased the variance of the estimator, a more refined adjustment involving lower mean square error seems possible.

5. Computations for Maximum Likelihood and Adjusted Maximum Likelihood Estimation

From equation (12) page 4 the maximum likelihood estimator of ρ is found as that value of ρ which minimizes the function

$$(12) \quad S(\rho) = (1 - \rho^2)^{-\frac{1}{2}} (y - X\beta_\rho)' B_\rho (y - X\beta_\rho) .$$

The differences between the stationary and nonstationary versions of the linear model have been noted earlier (footnote 2, page 4). The scanning technique for the nonstationary model [10, Appendix C, page 55] has been incorporated in some software computer packages, e.g. The Econometric Software Package, available from the University of Chicago, and the Econometric Pack for the CDC 6600 at the University of Minnesota (and elsewhere). For small T , this method may yield a substantially different numerical value for the maximum likelihood estimator than the scanning procedure set forth in [8, Appendix C, page 35]. As an example, using Hildreth, and Lu's data on the Demand for California Pears [10, equation (13), page 65], Henshaw [7, page 652], the earlier scanning

procedure yields an estimate of .30 for ρ , while the later technique yields a value of .39. A brief outline of the computational procedure for the stationary model is given below. From (8*)

$$\beta_{\rho} = (X' B_{\rho} X)^{-1} X' B_{\rho} y$$

so that substituting in (12) yields

(38)

$$\begin{aligned} S(\rho) &= (1-\rho^2)^{-\frac{1}{T}} y' [I - X(X' B_{\rho} X)^{-1} X' B_{\rho}]' B_{\rho} [I - X(X' B_{\rho} X)^{-1} X' B_{\rho}] y \\ &= (1-\rho^2)^{-\frac{1}{T}} y' [B_{\rho} - B_{\rho} X(X' B X)^{-1} X' B_{\rho}] y \end{aligned}$$

In computer evaluations, it is convenient to store the $K \times 1$ vector $(X' B_{\rho} X)^{-1} X' B_{\rho} y$, since this is used eventually for least squares and maximum likelihood estimates of β . Further, $S(\rho) \cdot (1-\rho^2)^{\frac{1}{T}} / T$ yields the appropriate estimate of v for $\rho = 0$ and $\rho = \hat{\rho}$ (the ML estimate) and the second factor in the last relation above should be evaluated independently. To check for multiple minima, it is suggested that $S(\rho)$ be evaluated at values of $\rho : -.9 (.1) .9$. If the sign of successive first differences changes only once, one has reasonable assurance that the minimum is unique.

When scanning the range of ρ in this manner, it is convenient to compute and store the approximation to $E(\hat{\rho})$ given in (33). This may be written

$$(39) \quad E(\hat{\rho}) = \frac{T(T-1)\rho - (T+1)(1-\rho^2) \operatorname{tr} X' H X (X' B_{\rho} X)^{-1}}{T(T+1) - (T+1)(1-\rho^2) \operatorname{tr} X' X (X' B_{\rho} X)^{-1}}$$

For this computation one needs $X' H X$ and $X' X$, the latter being most

conveniently found as $X'B_0X$. In this initial scan suppose the minimum value of $S(\rho)$ occurs at $\rho = \rho_0^1$, and no evidence of multiple minima is found. The numerical search for the minimum value of $S(\rho)$ is then carried out as follows:

- (a) Evaluate $S(\rho)$ at $\rho_1 = \rho_0^1 - d_1$, (i corresponds to an iteration index; $i = 1, \dots, 10$ and $d_1 = .05$), and $\rho_2 = \rho_0^1 + d_1$.
- (b) Choose ρ_0^{i+1} as that value of ρ yielding the minimum of $S(\rho_0)$, $S(\rho_1)$, $S(\rho_2)$.
- (c) Set $d_{i+1} = d_i/2$ and repeat the process.

After 10 iterations, the final value of ρ emerging from step (b) is taken as the maximum likelihood estimator, and is denoted as ρ_3 .

To find the adjusted maximum likelihood estimator, two values of ρ , ρ_4 and $\rho_5 = \rho_4 + .1$ are determined to satisfy $E(\hat{\rho}_{\rho_4}) \leq \rho_3 < E(\hat{\rho}_{\rho_5})$. The technique used here is to approximate the function $\varphi(\rho)$ by a straight line segment about the point ρ^* for which $\varphi(\rho^*) = \hat{\rho}$. While the interval of approximation could be made as small as possible it has been found convenient to use intervals of length .1 with values of $E(\hat{\rho}_\rho)$ calculated using Equation (39) in the initial scan. The points ρ_4 and ρ_5 are the endpoints of such an interval. The adjusted maximum likelihood estimator ρ_6 is then computed, using linear interpolation, as

$$(40) \quad \rho_6 = \rho_4 + \frac{\rho_3 - E(\hat{\rho}_{\rho_4})}{10[E(\hat{\rho}_{\rho_5}) - E(\hat{\rho}_{\rho_4})]} .$$

A FORTRAN subroutine for the calculation of the maximum likelihood and adjusted maximum likelihood estimates is available, upon request, from the second author.

REFERENCES

- [1] Anderson, R. L. and Anderson, T. W., "The distribution of the circular serial correlation coefficient for residuals from a fitted Fourier series", Annals of Mathematical Statistics, 21, (1950), 50-81.

- [2] Chipman, John S., "The problem of testing for serial correlation in regression analysis", Technical Report 4, Department of Economics, University of Minnesota, 1965.

- [3] de Leeuw, Frank and Gramlich, Edward, "The channels of monetary policy: a further report on the Federal Reserve-MIT model", The Journal of Finance, 24, (1969), 265-90.

- [4] Dent, Warren T., "The distribution of the maximum likelihood estimator of the autocorrelation coefficient in a linear model with autoregressive disturbances", Unpublished Ph. D. Thesis, University of Minnesota, 1971.

- [5] Durbin, J., "Estimation of parameters in time-series regression models", Journal of the Royal Statistical Society, Series B, 22, (1960), 139-153.

- [6] Griliches, Zvi, and Rao, P., "Small-sample properties of several two-stage regression methods in the context of autocorrelated errors", Journal of the American Statistical Association, 64, (1969), 253-72.

- [7] Henshaw, R. C., Jr., "Testing single-equation least-squares regression models for autocorrelated disturbances", Econometrica, 34, (1966), 646-60.

- [8] Hildreth, Clifford and Lu, John Y., "A Monte Carlo study of the regression model with autoregressive disturbances", Rand Memorandum RM5728PR, Santa Monica, 1969.

- [9] Hildreth, Clifford, "Asymptotic distribution of maximum likelihood estimators in a linear model with autoregressive disturbances", Annals of Mathematical Statistics, 40, (1969), 583-94.

- [10] Hildreth, Clifford and Lu, John Y., "Demand relations with autocorrelated disturbances", Technical Bulletin 276 of the Michigan State University Agricultural Experiment Station, East Lansing, 1960.
- [11] Hoos, S., and Shear, S. W., "Relation between auction prices and supplies of California fresh Bartlett pears", Hilgardia, 14, California Agricultural Experiment Station, 1942.
- [12] Klein, Lawrence R., Economic Fluctuations in the United States, 1921-41, John Wiley and Sons, New York, 1950.
- [13] Linstrom, I., and King, Richard A., "The demand for North Carolina slicing cucumbers and green peppers", North Carolina State College, A. E. Information Series, No. 49, 1956.
- [14] Prest, A. R., "Some experiments in demand analysis", The Review of Economics and Statistics, 31, (1949), 33-49.
- [15] Reilly, D. P., "Evaluation of the small sample properties of five alternative estimation methods when the errors are correlated", Discussion Paper No. 87, Department of Economics, University of Pennsylvania, 1968.
- [16] Theil, Henri, and Nagar, A. L., "Testing the independence of regression disturbances", Journal of the American Statistical Association, 56, (1961), 793-806.
- [17] Zellner, Arnold and Tiao, G. C., "Bayesian analysis of the regression model with autocorrelated errors", Journal of the American Statistical Association, 59, (1964), 763-78.