Sankhyā : The Indian Journal of Statistics
1985, Volume 47, Series B, Pt. 1, pp. 141-157.

ESTIMATION AND TESTING IN AN AUTOCORRELATED LINEAR REGRESSION MODEL WITH DECOMPOSED ERROR TERM : THE CASE OF TWO AR(1) COMPONENTS

By NITYANANDA SARKAR
Indian Statistical Institute

SUMMARY. This paper assumes a linear regression model where the disturbance term is decomposed into two independent components each following autoregressive process of order 1. Such a specification of the disturbance term of an autocorrelated linear regression model is expected to be more complete from the point of view of sources of autocorrelation (say, errors-in-observations and/or misspecification) being incorporated into the model. A large sample test is suggested to identify different situations that are characterized by different combinations of the parameters involved in the variance-covariance matrix of the error term. Methods for obtaining consistent and efficient estimators for all the cases are also discussed.

1. Introduction

In the standard econometric literature estimation of a linear regression model with autocorrelated errors is done by assuming some particular stationary stochastic process [e.g., an autoregressive (AR) or a moving average (MA) process] for the errors. In fact, until recently the common practice has been to assume an AR(1) process. This kind of an approach does not seem to be quite satisfactory [see Newbold and Davies (1978) in this connection] and may not even be easy to justify often. For example, if the error term is viewed as the sum of two independent components—one representing the effects of omission of variables and the other errors-in-observations-then the autocorrelation in the composite error term can be due to autocorrelation in either or both the components. If these independent additive error components are assumed to follow AR(1) processes, the composite error term will, indeed, not follow an AR(1) process.[1] It will, in fact, follow an ARMA (2, 1) process [*vide* Granger and Morris (1976) and Rose (1977)]. Some of the standard two-step estimation methods like those of Cochrane-Orcutt and Prais-Winsten which are based on the assumption of

[1]The same conclusion will be valid even if one of the two components is not autocorrelated. Further, the same type of conclusion will hold if the individual components are assumed to follow other stationary stochastic processes.

AMS (1980) *subject classification* : 62J02, 62P20.

Key words and phrases : Autocorrelation, Bartlett's test, Composite error term, Generalized least squares, Misspecification.

AR(1) errors might then be inefficient. It seems, therefore, necessary to examine whether an observed autocorrelation in a regression equation is due to one or both of these factors, and accordingly make suitable assumptions about the error processes for efficient estimation.

It may be argued that in a suitation like the one considered above, one can straightforwardly assume the composite error term to follow an ARMA (2, 1) process and then estimate the parameters of the model by available methods [see, for example, Pierce (1971) and Judge *et al.* (1980)]. In our opinion, however, there are still valid reasons for undertaking a study of the present type. Specifically, when one of the error components arises out of misspecification [cf. Ramsey (1969), Chaudhuri (1977, 1979) etc.] it is worthwhile to see if we can ascertain whether the observed autocorrelation is due partly or wholly to mispecification. Such an information is helpful for various reasons. Firstly, one can then try to respecify the data matrix for proper reestimation. Secondly, an appropriate error variance-covariance matrix can then be obtained depending on the nature of the observed autocorrelation. And finally, if autocorrelation in the error results from misspecification of the data matrix, it is likely that some standard assumptions of the classical linear regression model will be violated so that the usual two-step methods of estimation will no longer be efficient [*vide* Maddala (1977) and Judge *et al.* (1980)]. To devise an efficient estimation technique for such a situation a knowledge of the source(s) of observed autocorrelation would be useful.

This apart, for the type of error structure considered here, there is no *a priori* basis for using an ARMA(2, 1) error process since one would not know beforehand if the error is, in fact, sum of two independent AR(1) components. Also, the presumption of an ARMA(2, 1) process would lead to an unduly complicated structure of the error variance-covariance matrix compared to the one resulting from the assumption of AR(1) processes for the two independent additive components of the error term.

In this paper we consider a model whose error term is decomposed into two independent additive components with the possibility of both being generated by AR(1) processes, and develop a large sample test for identifying various autocorrelated situations. We also suggest methods for obtaining consistent and efficient estimators for the parameters of the model according to situations. These include estimators for situations where standard methods turn out to be inapplicable. The test procedure as well as the estimators suggested here are based on ordinary least squares (OLS) residuals and hence are computationally simple.

In what follows in Section 2 we present the model. While in Section 3 the nature of the error process is discussed, the nature of autocorrelation of the error term is characterized in Section 4. Section 5 deals with the development of a large sample test for identifying the source(s) of autocorrelation. And lastly, the estimation method is described in Section 6.

2. The model

We consider a k-variable linear regression model which is written in matrix notation as

$$y = X\beta + \epsilon^{+} \quad \ldots \quad (2.1)$$

where X is the $(n \times k)$ data matrix on k regressors, β is the $(k \times 1)$ vector of associated regression coefficients, y is the $(n \times 1)$ vector of observations on the regressand, and ϵ^{+} is the $(n \times 1)$ vector of disturbances.[2]

We make the following assumptions[3] :

(i) $\varepsilon_t^{+} = \varepsilon_t + z_t$ for all $t = 1, 2, \ldots, n$. ... (2.2)

(ii) X is stochastic but disturbuted independently of both ϵ and z[4] ... (2.3)

where $\epsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)'$ and $z = (z_1, z_2, \ldots, z_n)'$.

(iii) ϵ is independent of z. ... (2.4)

---

[2]One of the regressors in the equation in (2.1) may be taken to be unity for all t incorporating thereby an intercept term in the equation.

[3]If the regression equation in (2.1) is considered as a misspecified equation then it can be shown that ε_t^{+} is indeed the sum of ε_t, the eror term in the true regression equation, and z_t which is due to misspecification [see, for example, Ramsey (1969) and Chaudhuri (1979)]. Also, as noted by Judge *et al.* (1980) and actually shown by Chaudhuri (1979), the Durbin-Watson (DW) statistic may come out to be significantly less than 2 more often than the predetermined first kind of error. There is thus a risk in applying DW test for detecting autocorrelation in the error term in this situation.

[4]The assumption that X and z are independent may appear to be restrictive. It may, however, be noted here that the assumption is needed only for estimation and not for the test proposed. For the later it is enough to assume that $\operatorname{plim}_{n \to \infty} \left(\frac{1}{n} X'z \right) = 0$. This is a much weaker assumption, and it gets satisfied for a misspecified model where β is redefined (as compared to the true model) so as to capture as much of the additional influence of the excluded regressors on the regressand as possible [see Chaudhuri (1979) and Gupta and Maasoumi (1979) for details]. On the other hand, if one tries to estimate the model by directly assuming an error process for ϵ^{+}—an ARMA (2, 1) process in our case—then he would necessarily have to assume independence of X and ϵ^{+} which, given our decomposition of ϵ^{+}, would also imply independence of X and z. Hence standard methods of estimation available for estimating models with ARMA processes are equally vitiated by the restrictiveness of this assumption.

(iv) $$\varepsilon_t = \rho_\varepsilon \varepsilon_{t-1} + u_t \quad \ldots \quad (2.5)$$

where $|\rho_\varepsilon| < 1$ and u_t's are independently distributed with zero mean and constant variance σ_u^2.

(v) $$z_t = \rho_z z_{t-1} + v_t \quad \ldots \quad (2.6)$$

where $|\rho_z| < 1$ and v_t's are independently distributed with zero mean and constant variance σ_v^2.

(vi) $$\operatorname*{plim}_{n \to \infty} \left(\frac{1}{n} X'X \right) = \Sigma_{xx}, \text{ a positive definite matrix of rank } k < n. \quad \ldots \quad (2.7)$$

Obviously from (2.5) and (2.6),

$$\left.\begin{array}{l} V(\varepsilon_t) = \sigma_\varepsilon^2 = \dfrac{\sigma_u^2}{1-\rho_\varepsilon^2} \\ V(z_t) = \sigma_z^2 = \dfrac{\sigma_v^2}{1-\rho_z^2} \end{array}\right\} \text{ for all } t = 1, 2, \ldots, n. \quad \ldots \quad (2.8)$$

3. Nature of the error process

It has been mentioned earlier that the error term will not, in general, follow the same process as that generating the individual components. We may now illustrate this for an AR(1) process. While the general result on this is well-known, the following arguments may shed some light on the issues.

Let us assume that ε_t^+'s are given by

$$\varepsilon_t^+ = \rho \varepsilon_{t-1}^+ + w_t \quad \ldots \quad (3.1)$$

where $|\rho| < 1$ and w_t's are independently distributed with zero mean and constant variance σ_w^2.

Now, since $\varepsilon_t^+ = \varepsilon_t + z_t$, we have from (3.1),

$$\varepsilon_t + z_t = \rho(\varepsilon_{t-1} + z_{t-1}) + w_t$$

or,

$$w_t = \varepsilon_t - \rho\varepsilon_{t-1} + z_t - \rho z_{t-1}. \quad \ldots \quad (3.2)$$

Let us now consider the first-order autocovariance of w_t given as

$$\begin{aligned} \operatorname{cov}(w_t, w_{t-1}) &= \operatorname{cov}(\varepsilon_t - \rho\varepsilon_{t-1} + z_t - \rho z_{t-1},\ \varepsilon_{t-1} - \rho\varepsilon_{t-2} + z_{t-1} - \rho z_{t-2}) \\ &= \sigma_\varepsilon^2(1-\rho\rho_\varepsilon)(\rho_\varepsilon - \rho) + \sigma_z^2(1-\rho\rho_z)(\rho_z - \rho). \quad \ldots \quad (3.3) \end{aligned}$$

Clearly, this is not, in general, equal to zero, as required for assumption (3.1) to be true. The covariance will be equal to zero only when

(i) $\rho = \rho_z = \rho_\varepsilon$ or, (ii) either $\sigma_z^2 = 0$ or $\sigma_\varepsilon^2 = 0$[5].

[5] Both σ_ε^2 and σ_z^2 cannot be equal to zero because that would mean there is no disturbance term in the model. If one of them (say, σ_z^2) is equal to zero, then obivously $\rho = \rho_\varepsilon$ and hence cov $(w_t, w_{t-1}) = 0$.

We thus find that even if both ε and z separately follow AR(1) processes, ε^+ would not, in general, follow the same process. As noted earlier ε^+ would, in fact, follow an ARMA (2, 1) process.

4. Characterization of the nature of autocorrelation

We may now present a method of investigating the nature of the disturbance term[6] under fairly mild assumptions. In order to describe the method conveniently we may first enumerate the possible situations as shown in the following table :

TABLE 1 : DIFFERENT MODELS FOR THE ERROR TERM

	$\rho_\varepsilon = 0$	$\rho_\varepsilon \neq 0$
$\sigma_z^2 = 0$	case 3*	case 5*
$\sigma_z^2 > 0$	case 1 : $\rho_z \neq 0$	case 2 : $\rho_z = 0$
	case 4 : $\rho_z = 0$	case 6 : $\rho_z \neq 0$ and $\rho_z \neq \rho_\varepsilon$
		case 7 : $\rho_z \neq 0$ and $\rho_\varepsilon = \rho_z = \rho$ (say)

* For cases 3 and 5, $\sigma_z^2 = 0$ which implies $z_t = 0$ for all t. Therefore ρ_z is not defined for those two cases.

It is possible to further characterize the different possibilities in terms of the parameters $\rho_\varepsilon, \rho_z, \sigma_\varepsilon^2$ and σ_z^2. For this purpose we consider the OLS residuals $\boldsymbol{e}^+$ where

$$\boldsymbol{e}^+ = \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}} \text{ and } \hat{\boldsymbol{\beta}} \text{ is the OLS estimator of } \boldsymbol{\beta}$$

Now,

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$$

$$= \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{z}$$

and thus

$$\operatorname*{plim}_{n\to\infty} \hat{\boldsymbol{\beta}} = \boldsymbol{\beta} \text{ by assumptions (2.3) and (2.7).} \quad \ldots \quad (4.1)$$

Therefore, $\boldsymbol{e}^+$ converges in distribution to $(\boldsymbol{\varepsilon}+\boldsymbol{z})$ as $n \to \infty$.

Let us now define

$$\hat{\rho}_1 = \frac{\sum_{t=2}^{n} e_t^+ e_{t-1}^+}{\sum_{t=1}^{n} e_t^{+2}}.$$

[6] In order to ensure that the error term has at least one random component, autocorrelated or not, we assume, without any loss of generality, that $\sigma_\varepsilon^2 > 0$.

Since e^+ converges in distribution to $(\epsilon+z)$,

$$\underset{n\to\infty}{\text{plim}}\ \hat{\beta}_1 = \underset{n\to\infty}{\text{plim}}\ \frac{\sum\limits_{2}^{n} (\epsilon_t+z_t)(\epsilon_{t-1}+z_{t-1})}{\sum\limits_{1}^{n} (\epsilon_t+z_t)^2}. \qquad \dots \quad (4.2)$$

Assuming that u_t's in (2.5) and v_t's in (2.6) have finite fourth-order moments[7], we have

$$\underset{n\to\infty}{\text{plim}}\ \frac{1}{n}\sum_{1}^{n} \epsilon_t^2 = \sigma_\epsilon^2 \qquad \underset{n\to\infty}{\text{plim}}\ \frac{1}{n}\sum_{2}^{n} \epsilon_t\epsilon_{t-1} = \rho_\epsilon\sigma_\epsilon^2$$

$$\underset{n\to\infty}{\text{plim}}\ \frac{1}{n}\sum_{1}^{n} z_t^2 = \sigma_z^2 \quad \text{and} \quad \underset{n\to\infty}{\text{plim}}\ \frac{1}{n}\sum_{2}^{n} z_t z_{t-1} = \rho_z\sigma_z^2.$$

Therefore (4.2) can be shown to reduce to

$$\underset{n\to\infty}{\text{plim}}\ \hat{\rho}_1 (= \bar{\rho}_1, \text{say}) = \frac{\rho_\epsilon\sigma_\epsilon^2+\rho_z\sigma_z^2}{\sigma_\epsilon^2+\sigma_z^2} \qquad \dots \quad (4.3)$$

Let us next define

$$\hat{\rho}_2 = \frac{\sum\limits_{3}^{n} e_t^+ e_{t-2}^+}{\sum\limits_{2}^{n} e_t^+ e_{t-1}^+}.$$

Again, by similar algebraic simplifications and making the same assumptions, we get

$$\underset{n\to\infty}{\text{plim}}\ \hat{\rho}_2 = (\bar{\rho}_2, \text{say}) = \frac{\rho_\epsilon^2\sigma_\epsilon^2+\rho_z^2\sigma_z^2}{\rho_\epsilon\sigma_\epsilon^2+\rho_z\sigma_z^2}. \qquad \dots \quad (4.4)$$

The values of $\underset{n\to\infty}{\text{plim}}\ \hat{\rho}_2$ for the different cases listed earlier may now be presented in the following table.

TABLE 2 : VALUES OF $\underset{n\to\infty}{\text{plim}}\ \hat{\rho}_2 (= \bar{\rho}_2)$ FOR THE DIFFERENT CASES OF TABLE 1

	$\rho_\epsilon = 0$	$\rho_\epsilon \neq 0$
$\sigma_z^2 = 0$	case 3* : 0	case 5 : ρ_ϵ
$\sigma_z^2 > 0$	case 1: ρ_z	case 2 : ρ_z
	case 4* : 0	case 6 : $\frac{\rho_\epsilon^2\sigma_\epsilon^2+\rho_z^2\sigma_z^2}{\rho_\epsilon\sigma_\epsilon^2+\rho_z\sigma_z^2}$
		case 7 : ρ where $\rho_\epsilon = \rho_z = \rho \neq 0$

*Strictly speaking, for cases 3 and 4, $\bar{\rho}_2 = \frac{0}{0}$ and hence undefined. In these two cases, what is really meant by the entries in the table is that $\underset{n\to\infty}{\text{plim}}\ \text{cov}\ (e_t^+, e_{t-2}^+) = 0$.

[7] *Vide* Goldberger (1963), pp. 149-153.

5. A large sample test for detecting the sources of autocorrelation

Let us define the random variable

$$\phi_t = e_t^+ - \hat{\beta}_2 e_{t-1}^+. \qquad \ldots \quad (5.1)$$

As set out in Table 2, $\hat{\beta}_2$ takes different values in large samples for different cases.

Now, let C_s be defined as

$$C_s = \text{cov}(\phi_t, \phi_{t-s}), \quad \text{for all } s > 0.$$

Then

$$C_s = \text{cov}(e_t^+, e_{t-s}^+) - \hat{\beta}_2 \,\text{cov}(e_t^+, e_{t-s-1}^+) + \hat{\beta}_2^2 \,\text{cov}(e_{t-1}^+, e_{t-s-1}^+) - \hat{\beta}_2 \,\text{cov}(e_{t-1}^+, e_{t-s}^+). \qquad \ldots \quad (5.2)$$

Now in large samples

$$\text{cov}(e_t^+, e_{t-s}^+) \simeq \text{cov}(z_t, z_{t-s}) + \text{cov}(\varepsilon_t, \varepsilon_{t-s}) = \rho_z^s \sigma_z^2 + \rho_\varepsilon^s \sigma_\varepsilon^2. \qquad \ldots \quad (5.3)$$

Therefore from (5.2), we have[8]

$$C_s = \rho_z^s \sigma_z^2 + \rho_\varepsilon^s \sigma_\varepsilon^2 - \bar{\beta}_2 \rho_z^{s+1} \sigma_z^2 - \bar{\beta}_2 \rho_\varepsilon^{s+1} \sigma_\varepsilon^2 + \bar{\beta}_2^2 \rho_z^s \sigma_z^2 + \bar{\beta}_2^2 \rho_\varepsilon^s \sigma_\varepsilon^2 - \bar{\beta}_2 \rho_z^{s-1} \sigma_z^2 - \bar{\beta}_2 \rho_\varepsilon^{s-1} \sigma_\varepsilon^2. \qquad \ldots \quad (5.4)$$

In order to examine the autocovariance structures of ϕ_t's, we obtain C_s's for $s > 0$ for each of the seven cases and present the values in the following table where the seven cases are arranged under three groups. The algebra is straightforward for all the cases and hence is omitted.

TABLE 3 : AUTOCOVARIANCE STRUCTURES OF ϕ_t FOR THE DIFFERENT CASES

group	case	values of parameters	autocovariance structure
I	1	$\rho_z = 0, \rho_\varepsilon \neq 0, \sigma_z^2 > 0$	$C_1 \neq 0, C_s = 0 \ \forall s \geq 2$
	2	$\rho_z \neq 0, \rho_\varepsilon = 0, \sigma_z^2 > 0$	
II	3	$\rho_\varepsilon = 0, \sigma_z^2 = 0$	$C_s = 0 \ \forall s \geq 1$
	4	$\rho_z = 0, \rho_\varepsilon = 0, \sigma_z^2 > 0$	
	5	$\rho_\varepsilon \neq 0, \ \sigma_z^2 = 0$	
	7	$\rho_z = \rho_\varepsilon = \rho \neq 0, \sigma_z^2 > 0$	
III	6	$0 \neq \rho_z \neq \rho_\varepsilon \neq 0, \sigma_z^2 > 0$	$C_s \neq 0 \ \forall s \geq 1$

[8] Henceforth, without any loss of generality, we shall take C_s to be $\underset{n\to\infty}{\text{plim}}\ C_s$.

It is now clear from the above table that we may distinguish among three broad groups of cases on the basis of the first two autocovariances of ϕ_t's. However, the discrimination among these groups of cases on the basis of zero/non-zero values of the autocovariances of ϕ_t's can equivalently be done by using autocorrelation coefficients of ϕ_t's and hence we suggest application of Bartlett's well-known test [see Box and Jenkins (1976), pp. 34–36 and Malinvaud (1980), pp. 442–444] to examine whether the population autocorrelation coefficients of ϕ_t's ($t = 1, 2, \ldots$) are effectively zero beyond a certain lag.

It can be noted that all assumptions of Bartlett's test[9] are satisfied if we assume ϕ_t's to be normal (this will automatically mean assumption of normality for ϵ_t^+'s) and hence it can be used to discriminate among our three broad groups of cases. This, however, would not solve our problem completely. To achieve complete identification we have to further distinguish among the different cases falling under the same group. It does not seem possible to distinguish between the two cases in Group I which are observationally equivalent. But for Group II, we can divide the four cases into two subgroups each consisting of two indistinguishable cases. While in Cases (5) and (7), ϵ_t's follow an AR(1) process, in Cases (3) and (4), ϵ_t^+ is a random series. In fact, we test the null hypothesis

$$H_0 : \begin{cases} \text{either } \rho_\epsilon = 0, \quad \sigma_z^2 = 0 \\ \text{or} \quad \rho_\epsilon = 0, \quad \rho_z = 0, \quad \sigma_z^2 > 0 \end{cases}$$

against the alternative

$$H_1 : \begin{cases} \text{either } \rho_\epsilon \neq 0, \quad \sigma_z^2 = 0 \\ \text{or} \quad \rho_\epsilon = \rho_z = \rho \neq 0, \quad \sigma_z^2 > 0 \end{cases}$$

by using the conventional DW and other tests.

We may now restate the final groupings of the seven exhaustive cases according to their identifiability on the basis of the tests proposed here. The broad descriptions of the situations are also given here.

Group I : (i) $\rho_\epsilon = 0, \quad \rho_z \neq 0, \quad \sigma_z^2 > 0$ and (ii) $\rho_\epsilon \neq 0, \quad \rho_z = 0, \quad \sigma_z^2 > 0$ } AR(1)+Random

Group II : (subgroup 1) (i) $\rho_\epsilon = 0, \quad \sigma_z^2 = 0$ and (ii) $\rho_\epsilon = 0, \quad \rho_z = 0, \quad \sigma_z^2 > 0$ } Random

[9]As ϵ_t's and z_t's are stationary processes and $\epsilon_t^+ \to \epsilon_t + z_t$ in distribution as $n \to \infty$, ϵ_t^+'s (and hence ϕ_t's) would be stationary in large samples.

Group II : (subgroup 2)	(i) $\rho_e \neq 0, \; \sigma_z^2 = 0$ and (ii) $\rho_e = \rho_z = \rho \neq 0, \; \sigma_z^2 > 0$	AR(1)
Group III :	(i) $0 \neq \rho_e \neq \rho_z \neq 0, \; \sigma_z^2 > 0$	AR(1)+AR(1).

It may be noted that the failure to achieve further discrimination does not really affect the conclusion regarding the nature of autocorrelation in the error term. This may be seen by looking at the interpretations of the cases belonging to the four groups/subgroups stated above. Thus, if an observed situation is found to fall in Group I, the conclusion would be that while both the components are present, the autocorrelation is due to only one of them. The situation represented by Group III, on the other hand, implies that the autocorrelation is due to both the components being autocorrelated. Cases coming under subgroup 1 of Group II indicate that there is no autocorrelation in the error term. The first case in the subgroup 2 of Group II would mean that the error term is autocorrelated and that it is due to one component only, the other being absent.[10] Clearly, the autocorrelated linear regression model commonly considered in the literature really deals with situations described in subgroup 2 of Group II.

6. Estimation

We have seen in the last section that it is not possible to know exactly which particular case out of the seven possible cases, a given set of data represents. (The only exception is Case (6) under Group III.) The most that could be done by using our tests is to classify a given situation into one of the four broad groups/subgroups. However, although it is not possible to distinguish between the cases in the first three groups/subgroups, one can still consistently and efficiently estimate the regression coefficients for each of the seven cases if the broad group/subgroup to which a particular case belongs could be identified.

We now describe the method of estimation to be applied for the different situations. For this purpose, it would be convenient to first obtain the structure of the variance-covariance matrix of ε^+ for the most general case where $\sigma_z^2 > 0$ and $0 \neq \rho_e \neq \rho_z \neq 0$.

Since $$\varepsilon_t^+ = e_t + z_t,$$

[10]In the other case of this subgroup, both the components are present and yet the error can follow an AR(1) process. But then $\rho_e = \rho_z = \rho \neq 0$ and this is unlikely to happen in observed situations.

we have

$$\operatorname{cov}(\varepsilon_t^+, \varepsilon_{t-s}^+) = \rho_\varepsilon^s \sigma_\varepsilon^2 + \rho_z^s \sigma_z^2 \quad \text{for } s = 0, 1, \ldots, t-1 \text{ and } t = 1, 2, \ldots, n.$$

Hence

$$V(\boldsymbol{\varepsilon}^+) = \begin{bmatrix} \sigma_\varepsilon^2+\sigma_z^2 & \rho_\varepsilon\sigma_\varepsilon^2+\rho_z\sigma_z^2 & \ldots & \rho_\varepsilon^{n-1}\sigma_\varepsilon^2+\rho_z^{n-1}\sigma_z^2 \\ \rho_\varepsilon\sigma_\varepsilon^2+\rho_z\sigma_z^2 & \sigma_\varepsilon^2+\sigma_z^2 & \ldots & \rho_\varepsilon^{n-2}\sigma_\varepsilon^2+\rho_z^{n-2}\sigma_z^2 \\ \ldots\ldots\ldots & \ldots\ldots\ldots & \ldots & \ldots\ldots\ldots \\ \rho_\varepsilon^{n-1}\sigma_\varepsilon^2+\rho_z^{n-1}\sigma_z^2 & \rho_\varepsilon^{n-2}\sigma_\varepsilon^2+\rho_z^{n-2}\sigma_z^2 & \ldots & \sigma_\varepsilon^2+\sigma_z^2 \end{bmatrix}$$

$$= \sigma_\varepsilon^2 \begin{bmatrix} 1+\sigma_0^2 & \rho_\varepsilon+\rho_z\sigma_0^2 & \ldots & \rho_\varepsilon^{n-1}+\rho_z^{n-1}\sigma_0^2 \\ \rho_\varepsilon+\rho_z\sigma_0^2 & 1+\sigma_0^2 & \ldots & \rho_\varepsilon^{n-2}+\rho_z^{n-2}\sigma_0^2 \\ \ldots\ldots\ldots & \ldots\ldots\ldots & \ldots & \ldots\ldots\ldots \\ \rho_\varepsilon^{n-1}+\rho_z^{n-1}\sigma_0^2 & \rho_\varepsilon^{n-2}+\rho_z^{n-2}\sigma_0^2 & \ldots & 1+\sigma_0^2 \end{bmatrix} \quad \ldots \quad (6.1)$$

where $\sigma_0^2 = \sigma_z^2/\sigma_\varepsilon^2$.

Clearly, depending upon the particular values of the parameters involved, this matrix will assume different forms for the different cases.

Let us first consider subgroup 1 of Group II which comprises Cases (3) and (4). It may be seen that the variance-covariance matrix for this group is

$$V(\boldsymbol{\varepsilon}^+) = \sigma^{+2} I_n$$

where

$$\sigma^{+2} = \begin{cases} \sigma_\varepsilon^2 & \text{for Case (3)} \\ \sigma_\varepsilon^2+\sigma_z^2 & \text{for Case (4).} \end{cases}$$

It is obvious that in either case the best linear unbiased (and consistent) estimator of the regression coefficient $\boldsymbol{\beta}$ for such a model is given by OLS estimator and the estimate of the asymptotic variance-covariance matrix of $\hat{\boldsymbol{\beta}}$ is given by

$$n^{-1}\hat{\sigma}^{+2}\left(\frac{1}{n} X'X\right)^{-1}$$

where

$$\hat{\sigma}^{+2} = \frac{\sum_{1}^{n} e_t^{+2}}{n-k}.$$

It may be noted that $\hat{\sigma}^{+2}$ will estimate σ_ε^2 in Case (3) and $(\sigma_\varepsilon^2+\sigma_z^2)$ in Case (4) consistently. It is therefore a consistent estimator of the disturbance variance in each case. Thus, even if discrimination between Cases (3) and (4) is not possible, application of OLS will yield best estimates irrespective of the actual situation.

Now, for all the other cases, we can apply generalized least squares (GLS) taking $\sigma^2 V_0$ (where σ^2 and V_0 are different for the different cases and V_0 is assumed to be positive definite) as the form of the variance-covariance matrix of the disturbances. It is well-known that the application of GLS will yield consistent and asymptotically efficient estimate of $\boldsymbol{\beta}$. The trouble, however, is that V_0 is unknown and hence a straightforward application of GLS is not possible.

This problem can be tackled by using a result given in Theil. The result[11] is that given a consistent estimator $\hat{V}_0$ of V_0, $\boldsymbol{\beta}$, σ^2 and the asymptotic variance of $\widetilde{\boldsymbol{\beta}}$, the GLS estimator of $\boldsymbol{\beta}$ using $\hat{V}_0$, can under certain conditions be consistently and efficiently estimated by GLS with V_0 replaced by $\hat{V}_0$. The problem then reduces to finding a consistent estimator of V_0 for each of the possible cases.

If the tests suggested earlier indicate that a given situation falls under Group I, then the variance-covariance matrix of ϵ^+ is given by

$$V(\epsilon^+) = \sigma_z^2 \begin{bmatrix} \left(1+\frac{1}{\sigma_0^2}\right) & \rho_z & \rho_z^2 & \cdots & \rho_z^{n-1} \\ \rho_z & \left(1+\frac{1}{\sigma_0^2}\right) & \rho_z & \cdots & \rho_z^{n-2} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & & \cdots\cdots \\ \rho_z^{n-1} & \rho_z^{n-2} & \rho_z^{n-3} & \cdots & \left(1+\frac{1}{\sigma_0^2}\right) \end{bmatrix}$$

$$= \sigma_z^2 V_1.$$

or,

$$V(\epsilon^+) = \sigma_\epsilon^2 \begin{bmatrix} (1+\sigma_0^2) & \rho_\epsilon & \rho_\epsilon^2 & \cdots & \rho_\epsilon^{n-1} \\ \rho_\epsilon & (1+\sigma_0^2) & \rho_\epsilon & \cdots & \rho_\epsilon^{n-2} \\ \cdots\cdots & \cdots\cdots & & \cdots & \cdots\cdots \\ \rho_\epsilon^{n-1} & \rho_\epsilon^{n-2} & \rho_\epsilon^{n-3} & \cdots & (1+\sigma_0^2) \end{bmatrix}$$

$$= \sigma_\epsilon^2 V_1$$

[11]The result and the conditions can be found in Theil (1971, p. 399). It can easily be seen that the conditions are satisfied in the present case.

where $\sigma_0^2 = \sigma_z^2/\sigma_\epsilon^2$, as before, and V_1 is defined as above according to the specific case.

We now suggest the following estimator $\hat{V}_1$ for V_1:

$$\hat{V}_1 = \begin{bmatrix} \hat{\rho}_2/\hat{\rho}_1 & \hat{\rho}_2 & \hat{\rho}_2^2 & \cdots & \hat{\rho}_2^{n-1} \\ \hat{\rho}_2 & \hat{\rho}_2/\hat{\rho}_1 & \hat{\rho}_2 & \cdots & \hat{\rho}_2^{n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \hat{\rho}_2^{n-1} & \hat{\rho}_2^{n-2} & \hat{\rho}_2^{n-3} & \cdots & \hat{\rho}_2/\hat{\rho}_1 \end{bmatrix}$$

It can easily be seen from the definitions of $\hat{\rho}_1$ and $\hat{\rho}_2$ in Section 4 that $\hat{V}_1$ is a consistent estimator of V_1 irrespective of whether we have Case (1) or Case (2).

It may be noted that if a particular situation corresponds to either of Cases (5) and (7) in subgroup 2 of Group II, then the error actually follows an AR(1) process with autocorrelation coefficient ρ_ϵ or ρ as the case may be. For this the consistent estimator of V_0 is very well-known.

We may finally consider Group III which consists of only one case viz., Case (6). We have already presented the variance-covariance matrix for this case in (6.1). Now to get a consistent estimator of this variance-covariance matrix, we first define

$$S_i = \frac{1}{n-i} \sum_{t=i+1}^{n} e_t^+ e_{t-i}^+, \; i = 0, 1, 2, 3.$$

Then assuming that for both u_t and v_t fourth-order moments are finite [*vide* Goldberger (1963), pp. 149–153] and using the fact proved earlier that e_t^+ converges in distribution to $\epsilon_t + z_t$ as $n \to \infty$, it is easy to see that

$$\operatorname*{plim}_{n\to\infty} S_i = \rho_\epsilon^i \sigma_\epsilon^2 + \rho_z^i \sigma_z^2, \; i = 0, 1, 2, 3.$$

We may then obtain consistent estimators of $\rho_\epsilon, \rho_z, \sigma_0^2 (= \sigma_z^2/\sigma_\epsilon^2)$ by solving the following equations:

$$S_0 = \sigma_\epsilon^2 + \sigma_z^2$$
$$S_1 = \rho_\epsilon \sigma_\epsilon^2 + \rho_z \sigma_z^2$$
$$S_2 = \rho_\epsilon^2 \sigma_\epsilon^2 + \rho_z^2 \sigma_z^2$$
$$S_3 = \rho_\epsilon^3 \sigma_\epsilon^2 + \rho_z^3 \sigma_z^2.$$

Thus, $$\hat{\rho}_z = \frac{S_2-\hat{\rho}_\epsilon S_1}{S_1-\hat{\rho}_\epsilon S_0}$$

and $$\hat{\sigma}_0^2 = 1 \div \left\{\frac{S_0}{(S_1-\hat{\rho}_\epsilon S_0)/(\hat{\rho}_z-\hat{\rho}_\epsilon)}-1\right\}$$

where $\hat{\rho}_\epsilon$ is a solution of the quadratic equation

$$\rho_\epsilon^2(S_1^2-S_0S_2)+\rho_\epsilon(S_3S_0-S_1S_2)+(S_2^2-S_3S_1) = 0.$$

In case the quadratic equation yields two real roots each lying between -1 and $+1$, we shall choose that solution as the estimate of ρ_ϵ for which the residual sum of squares is minimum. Even though this general method of obtaining consistent estimates of the parameters ρ_ϵ, ρ_z and σ_0^2 could, in principle, be applied in all the cases, one may not use this method for the other cases because once the relevant group is identified, the methods for obtaining consistent estimates of the relevant parameters for these cases are fairly easy and efficient, at least asymptotically.[12]

Acknowledgements. The author is indebted to Dr. Dipankar Coondoo and Professors Robin Mukherjee and Nikhilesh Bhattacharya for their advice and criticisms in preparing this paper. Thanks are also due to an anonymous referee for some useful suggestions. The author is, of course, responsible for any errors.

References

Box, G. E. P. and G. M. Jenkins (1976): *Time Series Analysis : Forecasting and Control*, Revised Edition, San Francisco : Holden-Day.

Chaudhuri (Mukherjee), M. (1977): Autocorrelated disturbances in the light of specification analysis, *Journal of Econometrics*, 5, 301-313.

——— (1979): Some problems of econometric regression analysis: *Unpublished Ph.D. thesis, Indian Statistical Institute*, Calcutta.

Goldberger, A. S. (1963): *Econometric Theory*, John Wiley and Sons.

Granger, C. W. J. and M. J. Morris (1976): Time series modelling and interpretation, *J. Roy Statist Soc*, Series A, 139, 246-257.

Gupta, Y. P. and E. Maasoumi (1979): Omitted variables, variability of estimated parameters and the appearance of autocorrelated disturbances, *Journal of Econometrics*, 9, 379-385.

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

Maddala, G. S. (1977): *Econometrics*, New York : McGraw-Hill.

[12] Since we obtain consistent estimates of the parameters and since we have already assumed errors to be normal for the application of Bartlett's test, it may seem that efficient estimates could as well be obtained by maximum likelihood method of estimation. Actually this has been done in another paper in the context of a more general error process of which AR(1) is a special case.

MALINVAUD, E. (1980) : *Statistical Methods in Econometrics*, Third Revised Edition, Amsterdam : North Holland.

NEWBOLD, P. and N. DAVIES (1978) : Error mis-specification and spurious regressions, *International Economic Review*, 19, 513-519.

PIERCE, D. A. (1971) : Least squares estimation in the regression model with autoregressive-moving average errors, *Biometrika*, 58, 299-312.

RAMSEY, J. B. (1969) : Tests for specification errors in classical least squares regression analysis, *J. Roy Statist Soc*, Series B, 31, 35-71.

ROSE, D. E. (1977) : Forecasting aggregates of independent ARIMA processes, *Journal of Econometrics*, 5, 323-345.

THEIL, H. (1971) : *Principles of Econometrics*, New York : Wiley.

Paper received : November, 1982.
Revised : June, 1983.