

CONSISTENCY OF REGRESSION ESTIMATES
WHEN SOME VARIABLES ARE SUBJECT TO ERROR

by

Paul P. Gallo*

University of North Carolina at Chapel Hill

Abstract

For a general univariate "errors-in-variables" model, the maximum likelihood estimate of the parameter vector (assuming normality of the errors), which has been described in the literature, can be expressed in an alternative form. In this form, the estimate is computationally simpler, and deeper investigation of its properties is facilitated. In particular, we demonstrate that, under conditions a good deal less restrictive than those which have been previously assumed, the estimate is weakly consistent.

AMS 1980 Subject Classifications: 62F10, 62F12, 62J99.

Key Words and Phrases: Errors-in-variables model, maximum likelihood estimate, weak consistency.

*This research was supported in part by a National Science Foundation Graduate Fellowship, and by the Air Force Office of Scientific Research under Grant AFOSR-80-0080.

1. Introduction.

The estimation of linear regression parameters when some variables cannot be ascertained due to measurement or observation error is a problem with a long history in the statistical literature, yet one with a considerable recent emphasis. We consider a general "errors-in-variables" model in which some subset of the variables is observed with error (much of the literature concerns the case in which all variables are subject to error, with particular emphasis on models with just one independent variable; see Moran (1971) and Kendall and Stuart (1961, Chapter 27)).

Our model is

$$Y = X_1 \beta_1 + X_2 \beta_2 + \epsilon \qquad C = X_2 + U ,$$

$n \times 1$ $n \times p_1$ $p_1 \times 1$ $n \times p_2$ $p_2 \times 1$ $n \times p_2$

where β_1 and β_2 are vectors of regression parameters to be estimated, Y and C consist of observable random variables, X_1 and X_2 consist of constants but X_1 is known and X_2 is not, and ϵ and U are composed of random variables such that the rows of $[U \ \epsilon]$ are i.i.d. with mean zero

and unknown non-singular covariance matrix $\Sigma = \begin{bmatrix} \Sigma_u & \Sigma_{\epsilon u} \\ \Sigma'_{\epsilon u} & \sigma^2 \end{bmatrix}$. (Models such

as this with the independent variables being constants have generally been referred to under the title "linear functional relationship." A related model in which the variables are stochastic has been called a "linear structural relationship"; see Madansky (1959) for discussion.) Although in our discussion n will vary, there should be no confusion if we do not subscript the matrices involved.

We consider maximum likelihood estimation under the assumption that the errors are jointly normally distributed. It is well-known that

the supremum of the likelihood is infinite unless we impose additional structure on Σ (and furthermore, that under any conditions on Σ which yield a solution to the likelihood equations, the estimate obtained is the same as that obtained by the method of weighted least squares). The assumption most frequently made in the literature, and one which we will adopt, is

$$(1.1) \quad \Sigma = \sigma^2 \Sigma_0 = \sigma^2 \begin{bmatrix} \Sigma_{\epsilon u_0} & \Sigma_{\epsilon u_0} \\ \Sigma'_{\epsilon u_0} & 1 \end{bmatrix} \text{ with } \Sigma_0 \text{ known .}$$

The most detailed results along these lines can be obtained from the work of Gleser and his students, who considered multivariate regression models. In our model, let

$$(1.2) \quad W = [C \ Y]' R [C \ Y] \text{ with } R = I - X_1 (X_1' X_1)^{-1} X_1' ,$$

$$\theta = \lambda_{p_2+1} (\Sigma_0^{-1} W) \quad (\lambda_i(A) \text{ denotes } i^{\text{th}} \text{ largest eigenvalue of } A) ,$$

and

$$g' = \begin{pmatrix} g_1' & g_2' \\ 1 \times 1 \end{pmatrix} \text{ is an eigenvector associated with } \theta .$$

Healy (1975) has shown that if $g_2 \neq 0$, then the MLE's of β_1 and β_2 exist and are given by:

$$(1.3) \quad \hat{\beta}_2 = -g_1 g_2^{-1}$$

$$\hat{\beta}_1 = (X_1' X_1)^{-1} X_1' (Y - C \hat{\beta}_2) .$$

In Section 2, we demonstrate that the MLE can be expressed in alternate forms which are easier to interpret than (1.3), as well as computationally simpler. These "simpler forms" also facilitate deeper investigation of certain properties of the estimate. In Section 3, we consider one such aspect: we demonstrate that $\hat{\beta}_1$ and $\hat{\beta}_2$ are weakly consistent estimates under conditions weaker than those which have been previously shown.

2. The MLE under Normality.

In this section, we will make use of the following obvious notation:

$$\begin{aligned} X &= [X_1 \quad X_2] \\ C^* &= [X_1 \quad C] \\ \beta' &= [\beta_1' \quad \beta_2'] \\ U^* &= [0 \quad U] \end{aligned} \quad \begin{aligned} \Sigma^* &= \begin{bmatrix} \Sigma_u^* & \Sigma_{\epsilon u}^* \\ \Sigma_{\epsilon u}^{*\prime} & \sigma^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma \end{bmatrix} \\ p &= p_1 + p_2 ; \end{aligned}$$

we define $\Sigma_o^*, \Sigma_{uo}^*, \Sigma_{\epsilon uo}^*$ analogously. Also, let $H = [C^* \quad Y]' [C^* \quad Y]$.

The main result of this section is:

Theorem 1. *In our model, if the joint distribution of the errors is absolutely continuous with respect to Lebesgue measure, then the normality-MLE of β exists almost surely and is given by*

$$\begin{aligned} \hat{\beta}_2 &= (C'RC - \theta \Sigma_{uo}^*)^{-1} (C'RY - \theta \Sigma_{\epsilon uo}^*) \\ (2.1) \quad \hat{\beta}_1 &= (X_1'X_1)^{-1} X_1'(Y - C\hat{\beta}_2) , \end{aligned}$$

with θ and R given by (1.2); we also have

$$(2.2) \quad \hat{\beta} = (C^*{}' C^* - \theta \Sigma_{u0}^*)^{-1} (C^*{}' Y - \theta \Sigma_{\epsilon u 0}^*) ,$$

with $\theta = \gamma^{-1}$, $\gamma =$ largest root of $|\Sigma_0^* - \gamma H| = 0$.

In the form (2.2), $\hat{\beta}$ can be viewed as a modification of the ordinary least squares regression estimate, which is known to be inconsistent in the errors-in-variables (E.I.V.) case. In fact, the estimate seems to operate much like the "method-of-moments" estimate described by Fuller (1980). In an E.I.V. model in which Σ_u and $\Sigma_{\epsilon u}$ can be consistently and independently estimated but are otherwise unknown, Fuller has proposed estimates such as

$$\hat{\beta} = (C^*{}' C^* - n \hat{\Sigma}_u^*)^{-1} (C^*{}' Y - n \hat{\Sigma}_{\epsilon u}^*) .$$

Under the assumption that $n^{-1} X' X$ converges to a finite matrix, Healy (1975) showed that $n^{-1} \theta$ consistently estimates σ^2 ; hence $n^{-1} \theta \Sigma_{u0} \xrightarrow{p} \Sigma_u$ in our model. Thus, while Fuller's method requires an "external" variance estimate, the maximum likelihood approach in effect produces its own "internal" estimate. Of course we do not get this for free; the price we have paid is the additional structure that we have imposed upon Σ .

In proving Theorem 1, we will make use of the following result:

Lemma 1. Under the conditions of Theorem 1, $\theta = \lambda_{p_2+1}(\Sigma_0^{-1} W)$ is not an eigenvalue of $C' RC$ with probability one.

Proof. Let $G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \\ & & 1 \times 1 \end{bmatrix}$ be the matrix of normalized eigenvectors associated with the (ordered) eigenvalues of $\Sigma_o^{-1} W$, with $F = G^{-1}$ partitioned similarly. Thus

$$(2.3) \quad \Sigma_o^{-1} W = GDF$$

with $D = \begin{bmatrix} \lambda & 0 \\ 0 & \theta \end{bmatrix}$, $\lambda = \text{diag}(\lambda_1(\Sigma_o^{-1} W), \dots, \lambda_{p_2}(\Sigma_o^{-1} W))$. Equation (2.3) implies that

$$(2.4) \quad C' RC = (\Sigma_{uo} G_{11} + \Sigma_{\epsilon uo} G_{21})(\lambda - \theta I_{p_2}) F_{11} + \theta I_{p_2}.$$

From Gleser (1981), we infer that $\Sigma_{uo} G_{11} + \Sigma_{\epsilon uo} G_{21}$ and F_{11} are non-singular a.s. if the error distribution is absolutely continuous; it follows from a result of Okamoto (1973) that the eigenvalues of $\Sigma_o^{-1} W$ are distinct with probability one (all we need is $\theta \neq \lambda_{p_2}(\Sigma_o^{-1} W)$), in which case $\lambda - \theta I_{p_2}$ is non-singular. The result follows since (2.4) implies that $C' RC - \theta I_{p_2}$ is non-singular a.s. \square

Proof of Theorem 1. From the definition of θ and G ,

$$(2.5) \quad \begin{bmatrix} C' RC - \theta \Sigma_{uo} & C' RY - \theta \Sigma_{\epsilon uo} \\ Y' RC - \theta \Sigma'_{\epsilon uo} & Y' RY - \theta \end{bmatrix} \begin{bmatrix} G_{12} \\ G_{22} \end{bmatrix} = 0.$$

Gleser (1981) has shown that $G_{22} \neq 0$ a.s., in which case the MLE exists. As mentioned above, θ has multiplicity one a.s., so the left-hand matrix has rank p_2 w.p. 1, and solutions to (2.5) will be determined by

equations corresponding to any p_2 linearly independent rows of that matrix. In light of Lemma 1, the first p_2 rows will do:

$$(2.6) \quad (C'RC - \theta\Sigma_{uo})G_{12} + (C'RY - \theta\Sigma_{\epsilon uo})G_{22} = 0$$

$$\Rightarrow -G_{12}G_{22}^{-1} = (C'RC - \theta\Sigma_{uo})^{-1} (C'RY - \theta\Sigma_{\epsilon uo}) .$$

By (1.3), this is $\hat{\beta}_2$, which demonstrates (2.1).

For the second part of the theorem, note first that

$$H^{-1}\Sigma_o^* = \begin{bmatrix} 0 & 0 \\ 0 & W^{-1}\Sigma_o \end{bmatrix} ,$$

from which it follows that θ of eq. (2.2) is the same as that of (1.2) (in this part of the theorem, we want to express $\hat{\beta}$ in a form which does not explicitly refer to our partitions of the matrices involved).

Now according to (2.2),

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} X_1' X_1 & X_1' C \\ C' X_1 & C' C - \theta\Sigma_{uo} \end{bmatrix}^{-1} \begin{bmatrix} X_1' Y \\ C' Y - \theta\Sigma_{\epsilon uo} \end{bmatrix}$$

$$= \begin{bmatrix} (X_1' X_1)^{-1} + (X_1' X_1)^{-1} X_1' C C C' X_1 (X_1' X_1)^{-1} & -(X_1' X_1)^{-1} X_1' C \theta \\ -\theta C' X_1 (X_1' X_1)^{-1} & \theta \end{bmatrix} \begin{bmatrix} X_1' Y \\ C' Y - \theta\Sigma_{\epsilon uo} \end{bmatrix}$$

(with $Q = (C'RC - \theta\Sigma_{uo})^{-1}$)

$$= \begin{bmatrix} (X_1' X_1)^{-1} X_1' Y - (X_1' X_1)^{-1} X_1' C [Q(C' Y - \theta\Sigma_{\epsilon uo}) - \theta C' X_1 (X_1' X_1)^{-1} X_1' Y] \\ Q(C' Y - \theta\Sigma_{\epsilon uo}) - \theta C' X_1 (X_1' X_1)^{-1} X_1' Y \end{bmatrix} ,$$

which implies

$$\hat{\beta}_2 = Q(C'RY - \theta \Sigma_{\epsilon u_0}) = (C'RC - \theta \Sigma_{u_0})^{-1} (C'RY - \theta \Sigma_{\epsilon u_0})$$

$$\hat{\beta}_1 = (X_1'X_1)^{-1} X_1'(Y - C\hat{\beta}_2) .$$

These agree with (1.3) and (2.6). □

3. Consistency.

Various results concerning weak and strong consistency of $\hat{\beta}$ in our model and related models have been described by Healy (1975), Bhargava (1975), and Gleser (1981). Generally, all require that

$$(3.1) \quad \lim_{n \rightarrow \infty} n^{-1} X'X \text{ exists and is positive definite .}$$

Such a condition on X is much stronger than conditions which have been shown to be sufficient for consistency of the usual linear regression estimate (the special case of our model with $p_2 = 0$). In recent years, results of increasing strength and generality on this matter have been produced: see Eicker (1963), Drygas (1976), Anderson and Taylor (1976), Lai et al. (1979). Conditions on the errors vary somewhat among these papers, but the condition on X which is crucial to all of them is

$$(3.2) \quad \lambda_p(X'X) \rightarrow \infty \text{ as } n \rightarrow \infty .$$

We would like to find conditions "intermediate" between (3.1) and (3.2) which are sufficient for weak consistency of $\hat{\beta}$.

Theorem 2. *If the following conditions on X are satisfied:*

$$(A.1) \quad n^{-\frac{1}{2}} \lambda_p(X'X) \rightarrow \infty \text{ as } n \rightarrow \infty$$

$$(A.2) \quad \lambda_1^{-1}(X'X)\lambda_p^2(X'X) \rightarrow \infty \text{ as } n \rightarrow \infty,$$

and the joint distribution of the errors possesses finite fourth moment, then $\hat{\beta} \xrightarrow{P} \beta$ as $n \rightarrow \infty$. (Note that we are obtaining consistency without using the assumption that the errors are normally distributed.)

The following simple lemma will be useful:

Lemma 2. (i) $\lambda_1(X_2'RX_2) \leq \lambda_1(X'X)$;

(ii) letting $(X'X)^{-1} = \begin{bmatrix} L_1 & L_2 \\ \text{p} \times \text{p}_1 & \text{p} \times \text{p}_2 \end{bmatrix}$, $\lambda_1(L_2L_2') \leq \lambda_1^2(X'X)^{-1}$.

Proof. Since $(X_2'RX_2)^{-1}$ is the lower right-hand submatrix of $(X'X)^{-1}$,

$$\lambda_p(X'X)^{-1} = \inf_{\|Z\|=1} Z'(X'X)^{-1}Z \leq \inf_{\|Z\|=1} Z'(X_2'RX_2)^{-1}Z = \lambda_{p_2}(X_2'RX_2)^{-1}$$

$$\Rightarrow \lambda_1(X'X) \geq \lambda_1(X_2'RX_2).$$

Noting that the non-zero eigenvalues of L_2L_2' and $L_2'L_2$ are identical, (ii) follows similarly since $L_2'L_2$ is a lower right-hand submatrix of $(X'X)^{-2}$. □

Proof of Theorem 2.

$$\begin{aligned} \hat{\beta} &= (C^*{}'C^* - \theta\Sigma_{u0}^*)^{-1} (C^*{}'Y - \theta\Sigma_{\epsilon u0}^*) \\ &= (I_p + (X'X)^{-1} (X'U^* + U^*{}'X + (U^*{}'U^* - n\Sigma_u^*) + (n\sigma^2 - \theta)\Sigma_{u0}^*))^{-1} \\ &\quad \times (X'Y + U^*{}'XB + (U^*{}'\epsilon - n\Sigma_{\epsilon u}^*) + (n\sigma^2 - \theta)\Sigma_{\epsilon u0}^*). \end{aligned}$$

Clearly, it will suffice to show that:

- (i) $(X'X)^{-1} X'U^* \xrightarrow{P} 0$
- (ii) $(X'X)^{-1} U^{*'} X \xrightarrow{P} 0$
- (iii) $(X'X)^{-1} (U^{*'}U^* - n\Sigma_u^*) \xrightarrow{P} 0$
- (iv) $|\text{no}^2 - \theta| (X'X)^{-1} \xrightarrow{P} 0$
- (v) $(X'X)^{-1} X'Y \xrightarrow{P} \beta$
- (vi) $(X'X)^{-1} (U^{*'}\epsilon - n\Sigma_{\epsilon u}^*) \xrightarrow{P} 0$.

Eicker (1963) has shown (v) when $\lambda_p(X'X) \rightarrow \infty$, which is of course true by (A.1); (i) also follows immediately from his work under the same condition.

Note that $U^{*'}U^* - n\Sigma_u^* = O_p(n^{-1/2})$ if the errors have finite fourth moments, so (iii) holds if $(X'X)^{-1} = o(n^{-1/2})$, which follows from (A.1). The same argument demonstrates (vi). $(X'X)^{-1} U^{*'}X = L_2 U'X$; the $(i,j)^{\text{th}}$ element has mean zero and variance $\sum_k X_{kj}^2 \cdot P_i' \Sigma_u P_i$, where P_i is the i^{th} column of L_2' . Thus (ii) is satisfied if $\max \text{diag}(X'X) \cdot \max \text{diag}(L_2 L_2') \rightarrow 0$; this is seen to be equivalent to (A.2) using Lemma 2(ii).

Letting k henceforth denote $\lambda_1(X'X)^{-1}$, we need only demonstrate (iv), which is equivalent to $k(\theta - \text{no}^2) \xrightarrow{P} 0$. Note that $\theta = \lambda_{p_2+1}(\Sigma_0^{-1}W)$ if and only if $k(\theta - \text{no}^2) = \lambda_{p_2+1}(k(\Sigma_0^{-1}W - \text{no}^2 I_{p_2+1}))$; we will show that this converges to zero in probability. Let

$$D = k\Sigma_0^{-1} \begin{bmatrix} X_2' R X_2 & X_2' R X_2 \beta \\ \beta_2' X_2' R X_2 & \beta_2' X_2' R X_2 \beta \end{bmatrix}. \quad \text{As the product of a positive definite}$$

matrix and a positive semi-definite matrix of rank p_2 , D has p_2 positive

eigenvalues, its other eigenvalue being zero. Now

$$\begin{aligned}
& k(\Sigma_0^{-1}W - n\sigma^2 I_{p_2+1}) - D \\
&= k\Sigma_0^{-1} \begin{bmatrix} X_2'RU + U'RX_2 & U'RX_2\beta + X_2'Re \\ \beta_2'X_2'RU + \varepsilon'RX_2 & \beta_2'X_2'Re + \varepsilon'RX_2\beta_2 \end{bmatrix} \\
&\quad + k\Sigma_0^{-1}\{[U \ \varepsilon]' [U \ \varepsilon] - n\Sigma\} + k\Sigma_0^{-1}[U \ \varepsilon]' (R - I_n)[U \ \varepsilon] \\
&= M_1 + M_2 + M_3, \quad \text{say.}
\end{aligned}$$

Using arguments essentially the same as before, $M_1 \xrightarrow{p} 0$ by (A.2) and Lemma 2(i). M_2 does likewise since $k = o(n^{-1/2})$. Finally, noting that $I_n - R$ is idempotent, we deduce that $E[M_3] = -\sigma^2 kp_1 I_{p_2+1}$. The diagonal elements of the positive definite matrix $\Sigma_0 M_3$ are positive with expectations going to zero; thus they are $o_p(1)$ themselves. Consequently, $M_3 \xrightarrow{p} 0$.

Since eigenvalues are continuous functions of a sequence of matrices, it follows from the above discussion that

$$\lambda_{p_2+1}(k(\Sigma_0^{-1}W - n\sigma^2 I_{p_2+1})) \xrightarrow{p} \lambda_{p_2+1}(D) = 0, \text{ and hence } k(\theta - n\sigma^2) \xrightarrow{p} 0. \quad \square$$

Our assumptions (A.1) and (A.2) are intermediate in the sense mentioned earlier: either one implies (3.2), while both are implied by (3.1). Condition (A.1) requires that $X'X$ "gets large" at a faster rate than does (3.1) (it can be seen by considering the demonstration of (iii), e.g. in the proof of Theorem 2, that (3.2) is too weak a condition for our model). A simple example in which (3.1) is too weak to ensure consistency, but where (A.1) suffices, is a situation where $p = p_2 = 1$,

and the independent variable varies linearly with n . Condition (A.2) will also hold much more generally than (3.1); it is satisfied, for example, if (A.1) holds and the independent variables are bounded. Finally, while our requirement of fourth moments of the errors is not particularly restrictive, we could weaken it if we were willing to strengthen (A.1) (for example, we would require only finite $(2+\delta)^{\text{th}}$ moment, $0 \leq \delta \leq 2$, if $n^{-[2(2+\delta)^{-1}]}$ $\lambda_p(X'X) \rightarrow \infty$ as $n \rightarrow \infty$).

Acknowledgment

The author gratefully acknowledges the value of many discussions on this topic with Professor R.J. Carroll, as well as helpful suggestions from Professors Stamatis Cambanis and P.K. Sen.

References

- Anderson, T.W. and Taylor, J.B. (1976). Strong consistency of least squares estimates in normal linear regression. *Ann. Statist.* 4, 788-790.
- Bhargava, A.K. (1975). Estimation of a linear transformation and associated distributional problems. Ph.D. thesis, Purdue University.
- Drygas, H. (1976). Weak and strong consistency of the least squares estimators in regression models. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 34, 119-127.
- Eicker, F. (1963). Asymptotic normality and consistency of the least squares estimators for families of linear regressions. *Ann. Math. Statist.* 34, 447-456.
- Fuller, W.A. (1980). Properties of some estimators for the errors-in-variables model. *Ann. Statist.* 8, 407-422.
- Gleser, L.J. (1981). Estimation in a multivariate "errors in variables" regression model: large sample results. *Ann. Statist.* 9, 24-44.
- Healy, J.D. (1975). Estimation and tests for unknown linear restrictions in multivariate linear models. Ph.D. thesis, Purdue University.

- Kendall, M.G. and Stuart, A. (1961). *The Advanced Theory of Statistics*. Griffin, London.
- Lai, T.L., Robbins, H. and Wei, C.Z. (1978). Strong consistency of least squares estimates in multiple regression. *Proc. Natl. Acad. Sci. USA* 75, 3034-3036.
- Madansky, A. (1959). The fitting of straight lines when both variables are subject to error. *J. Amer. Statist. Assoc.* 54, 173-205.
- Moran, P.A.P. (1971). Estimating structural and functional relationships. *J. Multivariate Anal.* 1, 232-255.
- Okamoto, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *Ann. Statist.* 1, 763-765.

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|-----------------------|---|
| 1. REPORT NUMBER | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) Consistency of Regression Estimates When Some Variables Are Subject to Error | | 5. TYPE OF REPORT & PERIOD COVERED TECHNICAL |
| | | 6. PERFORMING ORG. REPORT NUMBER Mimeo Series No. 1328 |
| 7. AUTHOR(s) Paul P. Gallo | | 8. CONTRACT OR GRANT NUMBER(s) AFOSR Grant AFOSR-80-0080 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Bolling Air Force Base Washington, DC 20332 | | 12. REPORT DATE February 1981 |
| | | 13. NUMBER OF PAGES 13 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | | 15. SECURITY CLASS. (of this report) UNCLASSIFIED |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |
| 16. DISTRIBUTION STATEMENT (of this Report) Approved for Public Release -- Distribution Unlimited | | |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) | | |
| 18. SUPPLEMENTARY NOTES | | |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) errors-in-variables model, maximum likelihood estimate, weak consistency | | |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) For a general univariate "errors-in-variables" model, the maximum likelihood estimate of the parameter vector (assuming normality of the errors), which has been described in the literature, can be expressed in an alternative form. In this form, the estimate is computationally simpler, and deeper investigation of its properties is facilitated. In particular, we demonstrate that, under conditions a good deal less restrictive than those which have been previously assumed, the estimate is weakly consistent. | | |

