APPLICATIONS OF VARIANCE COMPONENTS
IN ECONOMICS

by

T. D. Mount

April 1974 No. 74-7

APPLICATIONS OF VARIANCE COMPONENTS IN ECONOMICS*

T. D. Mount**

I. Introduction

The use of variance components in economics has been almost exclusively confined to analyzing pooled cross-section and time-series data. Such data are obtained by observing a set of variables for the same sample of cross-section units (e.g. firms) at different points in time (e.g. years). Consequently, the most frequently used model is a two-way crossed classification with one observation per cell. In this model, the rows correspond to time-series effects and the columns to cross-section effects. Covariates are usually present, and in fact, estimation of the corresponding slope parameters is nearly always the main objective of the analysis. With most sources of economic data, however, this objective is confounded by the high correlations that exist between the covariates. In this situation, the ordinary least squares (OLS) estimator of a particular slope parameter may lack precision even though the model fits the data well. This is true if the cross-section and time-series effects are omitted and it is even more of a problem if the effects are specified as fixed parameters. If the effects are random, however, the corresponding generalized least squares (GLS) estimator is more efficient than either of the OLS alternatives. The lure of this increased efficiency provided the primary motivation for considering variance components

---

* This paper was presented at the Spring meeting of ENAR of the Biometric Society at Tallahassee, March 1974.

** T. D. Mount is an Assistant Professor in the Department of Agricultural Economics, Cornell University, Ithaca, New York.

models in economics, although the magnitudes of the variance components have never in themselves been considered very interesting.[1]

The discussion that follows is divided into three sections. First, the simplest model used for analyzing pooled cross-section and time-series data is presented. Even though the characteristics of this type of mixed model are familiar to biometricians, it provides a convenient framework for considering two modifications of the model that are presented in the last two sections. The first of these is a dynamic model which has been widely used in economics to represent the gradual adjustment of a dependent variable over time in response to changes in the levels of the regressors. The statistical properties of this model are not nearly as tractable as those of the simple model. Finally, a model in which the slope coefficients are specified with random components is considered. This model represents a relatively straightforward generalization of the simple model, but it is interesting because covariances between the random effects are specified. In addition, this represents one example of a model in which the variance components have a useful economic interpretation, namely as measures of the risk parameters associated with a production process.

II. The Basic Model for Pooled Data

If data are obtained for a sample of N cross-section units in T time periods, the basic variance components model may be written:

(1)
$$Y = X\beta + Z_{\tau}\tau + Z_{\mu}\mu + \epsilon$$

$$E[Y] = X\beta$$

$$\mathrm{Var}[Y] = Z_{\tau}Z'_{\tau}\sigma^2_{\tau} + Z_{\mu}Z'_{\mu}\sigma^2_{\mu} + I_{NT}\sigma^2_{\epsilon}$$

1/ This may be one of the reasons why there is relatively little econometric literature on estimating variance components. See Mount and Searle [9] for a more detailed discussion of this topic.

where Y is an NTx1 vector of the dependent variable

X is an NTxK matrix of the covariates which usually include a constant term

$Z_\tau = I_N \otimes 1_T$ is a design matrix for the time effects

$Z_\mu = 1_T \otimes I_N$ is a design matrix for the cross-section effects

I_{NT} and I_N are identity matrices of order NT and N, respectively

1_T is a Tx1 vector of ones

β is a Kx1 vector of unknown slope parameters

τ is a Tx1 vector of unknown random time-series effects

μ is an Nx1 vector of unknown random cross-section effects

ϵ is an NTx1 vector of unknown random residuals

σ^2_μ, σ^2_τ and σ^2_ϵ are the unknown variances of the elements of μ, τ and ϵ, respectively

$\otimes$ represents a Kronecker product.

If the three variances are known, the best linear unbiased estimator of β in (1) is the following GLS estimator:

$$(2) \qquad \beta^* = (X'V^{-1}X)^{-1}X'V^{-1}Y$$

where $V = [Z_\tau Z'_\tau \sigma^2_\tau/\sigma^2_\epsilon + Z_\mu Z'_\mu \sigma^2_\mu/\sigma^2_\epsilon + I_{NT}]$

This type of model was first discussed in the econometric literature in 1966 by Balestra and Nerlove [2]. As the form of V is very simple, the characteristics of β^* in (2) can be investigated analytically. For example, Wallace and Hussain [16, p. 58] derived an expression for V^{-1},1/ and Nerlove [12]

1/ C. R. Henderson, Jr. [5] showed that the same expression for V^{-1} can be derived using an alternative procedure for computing β^* in (2). This procedure, which is well known to biometricians, was developed by C. R. Henderson, Sr. twenty years earlier.

derived the characteristic roots of V^{-1} in terms of the variance components. These analytical properties can be used to simplify the computation of (2) in applications.

When the variance components are unknown, standard procedures for mixed models can be used to estimate V^{-1} in (2). For example, estimators of the three variance components are derived by Mount and Searle [9] using Henderson's Method III. An estimate of β can then be obtained by replacing V^{-1} in (2) with its estimate. However, it is difficult to choose which of the alternative estimators of the variance components is best as in practice there is little difference in the efficiency of the corresponding estimators of β (e.g., see Maddala and Mount [8]).[1/] In addition, the statistical properties of β^* in (2) are difficult to determine when an estimate of V^{-1} is used. This latter statement does not apply to the maximum likelihood (ML) estimator of β, which appears, therefore, to be a sensible procedure to consider. Fortunately, the ML estimator is relatively easy to compute when Y in (1) is specified to be multivariate $N(X\beta, V\sigma_\epsilon^2)$ because of the simple form of V.

At this stage, it is convenient to simplify (1) by omitting one set of random effects. This is done to make the notation less cumbersome as all the important aspects of the material that follows can be derived from the simplified model. If $\sigma_\tau^2 = 0$ in (1), which is equivalent to omitting the time-series effects (rows), the likelihood function L may be transformed to:

$$(3) \qquad -2 \ln L = \text{const.} + NT \ln\sigma_\epsilon^2 + \ln|V| + [Y - X\beta]'V^{-1}[Y - X\beta]\sigma_\epsilon^{-2}$$

where $V^{-1} = [I_{NT} + Z_\mu Z'_\mu \sigma_\mu^2/\sigma_\epsilon^2]^{-1} = [I_{NT} - Z_\mu Z'_\mu \Phi]$, and $\Phi = \sigma_\mu^2/(\sigma_\epsilon^2 + T\sigma_\mu^2)$.

1/ If there are no covariates in the model, most of these alternatives give identical estimates as the data are balanced. This is no longer true if covariates are present, and consequently, some choice is necessary.

Following Maddala [7, p. 345], the ML estimates of β and σ_ϵ^2 are solutions to:

$$(4)\qquad \hat{\beta} = (X'[I_{NT} - Z_\mu Z_\mu'\hat{\Phi}]X)^{-1}X'[I_{NT} - Z_\mu Z_\mu'\hat{\Phi}]Y$$

$$= [W_{xx} + (1-T\hat{\Phi})B_{xx}]^{-1}[W_{xy} + (1-T\hat{\Phi})B_{xy}]$$

$$(5)\qquad \hat{\sigma}_\epsilon^2 = [Y - X\hat{\beta}]'[I_{NT} - Z_\mu Z_\mu'\hat{\Phi}][Y - X\hat{\beta}]/NT,$$

where $W_{xx} = [X'X - X'Z_\mu Z_\mu'X/T]$ and $W_{xy} = [X'Y - X'Z_\mu Z_\mu'Y/T]$ represent the sums of squares and cross products within the cross-section units

$B_{xx} = [X'X - W_{xx}]$ and $B_{xy} = [X'Y - W_{xy}]$ represent the sums of squares and cross products between the cross-section units,

and $\hat{\Phi}$ is chosen to minimize:1/

$$(6)\qquad NT\ \ln\hat{\sigma}_\epsilon^2 - \ln|\hat{V}| = NT\ \ln\hat{\sigma}_\epsilon^2 - N\ \ln(1-T\hat{\Phi}).$$

It follows from the definition of Φ in (3) that the ML estimator of σ_μ^2 is $\hat{\sigma}_\mu^2 = \hat{\sigma}_\epsilon^2\hat{\Phi}/(1-T\hat{\Phi})$, and if $0 \leq T\hat{\Phi} \leq 1$, then $\hat{\sigma}_\mu^2$ is nonnegative. Consequently, the ML estimates are generally computed by specifying values for $T\hat{\Phi}$ over the range 0 to 1, and selecting the value that minimizes (6). The form of (6) implies that a boundary solution may exist at 0 but not at 1.

In summary, the main analytical objective when applying (1) to economic data is typically to obtain estimates of the slope coefficients of the covariates. With pooled cross-section and time-series data, the ML estimates

---

1/ If $T\hat{\Phi} = 0$, then $\hat{\beta}$ in (4) is equivalent to the OLS estimator with the random effects in (1) omitted, and if $T\hat{\Phi} = 1$, to the OLS estimator with the effects specified as fixed parameters.

of these parameters and the variance components are relatively easy to compute.[1] The statistical properties of ML estimators are well known, and consequently, this method of estimation is generally more suitable for economic applications than the alternative two-step procedures.

III. A Dynamic Model

In certain situations, a dependent variable is not expected to respond immediately to changes in the levels of explanatory variables. The following model of a partial adjustment process, originally proposed by Nerlove [10], has been widely used in economics, and in the context of using pooled data, it represents the behavior of the dependent variable over time for any one of the cross-section units

$$y_t - y_{t-1} = (1-\lambda)(y_t^o - y_{t-1}) \qquad 0 \leq \lambda < 1 \tag{7}$$

$$(1-\lambda)y_t^o = \sum_{k=1}^{K} \beta_k x_{kt} + \nu_t \tag{8}$$

where y_t is the observed level of a dependent variable at time t, y_t^o is the unobserved target level of that variable which is specified as a linear function of K observable variables, x_{kt}; $k = 1,2,\ldots,K$ and a residual term, ν_t. $\beta_1, \beta_2, \ldots, \beta_K$ and λ are unknown parameters.

The implication of (7) is that the observed change in the dependent variable from one time period to the next $(y_t - y_{t-1})$ is a fixed proportion $(1-\lambda)$ of the difference between the target level (y_t^o) and the previous observed level (y_{t-1}). By substituting (8) into (7), the unobserved variable y_t^o can be eliminated to give:

1/ If both sets of random effects are included, the computations involve searching for the values of two parameters that minimize an expression corresponding to (6) (see Maddala [7, p. 354]).

$$(9) \qquad y_t = \sum_{k=1}^{K} \beta_k x_{kt} + \lambda y_{t-1} + \nu_t$$

This may be rewritten in the distributed lag form by substituting for y_{t-1}, y_{t-2}, ... on the RHS of (9) to give:

$$(10) \qquad y_t = \sum_{k=1}^{K} \beta_k \sum_{i=0}^{\infty} \lambda^i x_{k,t-i} + \sum_{i=0}^{\infty} \lambda^i \nu_{t-i}$$

The advantage of (9) is that it is a linear function of the unknown parameters. However, the presence of the lagged dependent variable y_{t-1} as a regressor creates statistical problems. For example, the OLS estimator is consistent only if the residuals are uncorrelated with each other (e.g., see Griliches [4, p. 36]). Even if this condition holds, the usual small-sample properties of the OLS estimator no longer apply. Nevertheless, the OLS estimator has been widely used in applications, and in many cases, the statistical problems have been ignored.

Although there is no lagged dependent variable in (10), the function is highly nonlinear in λ, and the computations required to estimate the parameters have been considered too cumbersome for most applications.1/ In spite of this, the approach is currently receiving more attention from economists due in part to the extensive work of Dhrymes [3] on applying ML procedures to (10).

If the same adjustment process in (7) and (8) is specified for each of the N different cross-section units, and if in addition, the residual ν_t in (8) includes a random cross-section (column) effect, then the autoregressive form (9) can be written as a modified version of the basic model for pooled data (1).

1/ Assuming that the first observation is taken at $t = 0$, the term $\sum_{i=0}^{\infty} \lambda^i x_{k,t-i}$ can be written $[\sum_{i=0}^{t} \lambda^i x_{k,t-i} + \lambda^t \eta_k]$ where $\eta_k = \sum_{i=1}^{\infty} \lambda^i x_{k,-i}$ for $k = 1, 2, \ldots, K$, and as η_k is unobserved, it can be considered as another unknown parameter. With pooled data, η_k will not be the same for all N cross-section units, and consequently, there will be an additional NK parameters in the model.

(11) $$Y = X\beta + L\lambda + Z_\mu \mu + \epsilon$$

where Y, X, Z_μ, β, μ, and ϵ are defined in (1)

λ is an unknown parameter

L is an NTx1 vector of the lagged dependent variable (the value of the dependent variable in the previous time period).1/

Treating L as though it represented another typical covariate, the GLS estimator of β and λ is

(12) $$\begin{bmatrix} \beta^* \\ \\ \lambda^* \end{bmatrix} = [W'V^{-1}W]^{-1}W'V^{-1}Y$$

$$= \begin{bmatrix} \beta \\ \\ \lambda \end{bmatrix} + [W'V^{-1}W]^{-1}W'V^{-1}[Z_\mu \mu + \epsilon]$$

where $W = [X\ L]$, $V^{-1} = [I_{NT} - Z_\mu Z'_\mu \Phi]$ and $\Phi = \sigma_\mu^2/(\sigma_\epsilon^2 + T\sigma_\mu^2)$ is defined in (3).

This estimator is consistent if the second term on the RHS of (12) has a probability limit equal to zero.2/ It is generally reasonable to assume that $\text{Plim}(NT)^{-1}W'V^{-1}W = \Psi$ is a matrix of fixed elements and that it is non-singular. Consequently, consistency depends on whether or not $\text{Plim}(NT)^{-1}W'V^{-1}[Z_\mu \mu + \epsilon] = 0$. The elements of this vector associated with the covariates in X are assumed to be zero in the limit, and it is only the single element associated with the lagged dependent variable L that is important. This element may be written:

1/ It is assumed that observations at $t = 0$ are available so that L has no missing elements.

2/ In most economic applications, this condition is actually more appropriate than the standard properties of GLS estimators as the covariates are not fixed by the analyst.

(13) $$\text{Plim}(NT)^{-1}[L'Z_\mu \mu + L'\epsilon - \Phi L'Z_\mu Z'_\mu Z_\mu \mu - \Phi L'Z_\mu Z'_\mu \epsilon]$$

$$= \text{Plim}(NT)^{-1}[(1-T\Phi)L'Z_\mu \mu - \Phi L'Z_\mu Z'_\mu \epsilon]$$

as $Z'_\mu Z_\mu = TI_N$ and $\text{Plim}(NT)^{-1}L'\epsilon = 0$. This may be further simplified to:[1/]

(14) $$\sigma_\mu^2(1-T\Phi)/(1-\lambda) - \sigma_\epsilon^2 T^{-1}\Phi[(T-1) + (T-2)\lambda + (T-3)\lambda^2 + \ldots + 2\lambda^{T-3} + \lambda^{T-2}]$$

It is clear that if $\Phi = 0$ in (12), implying that it is the OLS estimator, then (14) reduces to $\sigma_\mu^2/(1-\lambda)$, which demonstrates the inconsistency of this estimator. In contrast, if $\Phi = T^{-1}$, which is equivalent to treating the cross-section effects as fixed, the estimator is biased but consistent.[2/]

Using the definition of Φ in (12), (14) reduces to:

(15) $$\sigma_\epsilon^2\Phi\{(1-\lambda)^{-1} - T^{-1}[(T-1) + (T-2)\lambda + (T-3)\lambda^2 + \ldots + 2\lambda^{T-3} + \lambda^{T-2}]\}$$

$$= \sigma_\epsilon^2\Phi(1 + \lambda + \lambda^2 + \ldots + \lambda^{T-2} + \lambda^{T-1})/T(1-\lambda)$$

$$\doteq \sigma_\epsilon^2\Phi/T(1-\lambda)^2 \quad \text{if } T \text{ is large and } \lambda \text{ close to } 0.$$

It follows from (15) that the GLS estimator is consistent. This estimator was found to be considerably more efficient than the fixed effects model ($\Phi = T^{-1}$) in a Monte Carlo study conducted by Nerlove [11].

1/ The nth element of LZ_μ corresponding to the nth cross-section unit is:
$$\sum_{t=1}^{T} y_{t-1,n} = \sum_{k=1}^{K} \beta_k \sum_{t=1}^{T} \sum_{i=0}^{\infty} \lambda^i x_{k,t-i-1,n} + T\mu_n/(1-\lambda) + \sum_{t=1}^{T} \sum_{i=0}^{\infty} \lambda^i \epsilon_{t-i-1,n}.$$

2/ The condition $\text{Plim}(NT)^{-1}W'V^{-1}W = \Psi$ may be generalized to make Ψ dependent on the specified value of Φ.

When the variance components are unknown, standard estimation procedures cannot be directly applied. For example, procedures based on equating a quadratic form to its expected value, such as Henderson's Method III, are not suitable when the lagged dependent variable L is present. In addition, if the ML estimator discussed in Section II is derived by assuming that L is another covariate, the likelihood function is correct only if the initial values for each cross-section unit, y_{0n} $(n = 1, 2, \ldots, N)$ are independent of all subsequent observations (e.g., see Balestra and Nerlove [2, p. 598]).[1/] In fact, Nerlove [11] found that the estimator corresponding to (4) performed very poorly in a Monte Carlo study when he used data that did not conform to this specification.

At the present time, no single estimator of the variance components in (11) has been shown to be superior to others. It is not surprising that little is known about the properties of (12) when estimates of the variance components in V are used.

IV. A Stochastic Coefficient Model

A large part of economic theory is concerned with deriving the optimum solution for a given criterion, such as profits, under specified conditions. In most situations, a quantitative evaluation of this solution must be based on the estimated values of certain functional relationships. If these relationships are estimated using linear regression procedures, the results can usually be interpreted as an estimate of the optimum solution for the expected

---

1/ A crucial assumption is that $P\{\mu_n + \epsilon_{tn}\} = P\{y_{tn} | y_{t-1,n}, y_{t-2,n}, \ldots, y_{0n}\}$ where y_{tn}, μ_n and ϵ_{tn} are elements of Y, μ and ϵ, respectively, and in this situation, $P\{\mu_n + \epsilon_{1n}, \mu_n + \epsilon_{2n}, \ldots, \mu_n + \epsilon_{Tn}\} = P\{y_{1n}, y_{2n}, \ldots, y_{Tn} | y_{0n}\}$. The importance of the distribution of y_{0n} has been demonstrated by G. S. Maddala in an unpublished memorandum.

level of the criterion. In fact, the stochastic nature of a regression model has seldom contributed directly to improving the theoretical explanation of economic behavior. One exception is the theory of risk aversion.

The predominant application of risk aversion theory has been to selecting investment portfolios. Instead of maximizing the expected cash returns for a given total expenditure, investors may prefer a portfolio with a lower mean if the corresponding variance of net returns is smaller. This trade-off between the variance and the mean is, under certain conditions, consistent with risk averting behavior. The tendency of investors to diversify their stock holdings can be explained in this context if the returns to some stocks are negatively correlated.

If returns are observed for a collection of stocks in different time periods, it is easy to calculate the sample means, variances and covariances of returns. In more general situations, the means might be related to some observable variables in a regression model. The residuals of these regressions can still be used to estimate the variances and covariances using procedures which were developed by Zellner [17] for his "seemingly unrelated regressions." However, relatively little effort has been directed to determining how the structural form of the mean may be related to the theory of risk aversion.

If a production relationship is represented by a regression model with output as the dependent variable and the input levels as regressors,[1/] a standard economic objective is to determine the conditions for maximizing the expected level of profits. In a linear model, these conditions generally depend on the slope parameters, and under most specifications the optimum level of each input is a deterministic function of certain prices. However, there is generally no relationship between the mean level of profits and the

1/ Any of these variables may, in fact, be transformations of the original observations.

variance of profits due to the additive nature of the stochastic residual. Consequently, the possibility of averting risk by using inputs at levels lower than the optimum cannot be demonstrated. A standard regression model does not provide an adequate representation of the risk associated with a production process. In contrast, if the slope coefficients are specified as stochastic, the variance of output depends on the input levels,[1/] and as a result, risk behavior can be investigated in this type of model.

Regression models with stochastic coefficients have been proposed by Rubin [13], Hildreth and Houck [6], Theil [15, Section 12.4] and Swamy [14]. The usual reason for using these models in economics is to allow for some heterogeneity in the functional relationships that exist between cross-section units. However, in a production process, particularly if it is influenced by weather, it is the unexplained variability of output over time that contributes to risk. Consequently, the following model, based on pooled cross-section and time-series data, is suggested:

$$\begin{aligned} Y_1 &= X_1\beta + X_1\Delta_1 + \epsilon_1 \\ Y_2 &= X_2\beta + X_2\Delta_2 + \epsilon_2 \\ &\vdots \\ Y_T &= X_T\beta + X_T\Delta_T + \epsilon_T \end{aligned} \qquad (16)$$

Y_t is an Nx1 vector of output at time t

X_t is an NxK matrix of inputs at time t

β is a Kx1 vector of unobserved means of the slope coefficients

---

1/ The variance is identical in form to the variance of the predicted value of output for given input levels in a standard regression model, but this latter variance is related to the use of sample information and does not measure the risk faced by producers.

Δ_t is a Kxl vector of unobserved random deviations of the slope coefficients at time t from the mean vector β

ϵ_t is an Nxl vector of unobserved random residuals at time t.

If the following specifications hold,

(17)
$$\begin{aligned} E[\epsilon_t] &= 0 \text{ for all } t \\ E[\epsilon_t \epsilon_s'] &= I_N \sigma_t^2 \text{ for } t = s \\ &= 0 \text{ otherwise} \\ E[\epsilon_t \Delta_s'] &= 0 \text{ for all } s \text{ and } t \\ E[\Delta_t] &= 0 \text{ for all } t \\ E[\Delta_t \Delta_s'] &= \Omega \text{ for } t = s \\ &= 0 \text{ otherwise,} \end{aligned}$$

it follows that

(18)
$$\begin{aligned} E[Y_t] &= X_t\beta \text{ for all } t \\ Var[Y_t] &= X_t \Omega X_t' + I_N \sigma_t^2 \text{ for all } t \\ Cov[Y_t Y_s'] &= 0 \text{ for } t \neq s. \end{aligned}$$

If the cross-section effects μ in (1) are omitted, (16) can be considered as a generalization of this basic model. In fact, (16) reduces to (1) if $\sigma_1^2 = \sigma_2^2 = \ldots. = \sigma_T^2 = \sigma_\epsilon^2$ and if each Δ_t contains only one nonzero element corresponding to the constant term in X_t. This latter condition implies that Ω also has only one nonzero element, namely σ_τ^2.

In more general situations, all the elements of Ω are unknown. As a result, tne random deviations of the slope coefficients are assumed to be correlated. Although this correlation between random effects is not usually present in variance components models, it has important implications for risk behavior. Negative correlations between some of the coefficients may make certain combinations of inputs less risky than others in the same way that a diversified investment portfolio may reduce risk.

If $N > K$, the GLS estimator of β in (16) is shown by Swamy [14] to be:

$$\beta^* = [\Sigma_{t=1}^{T} \theta_t^{-1}]^{-1} \Sigma_{t=1}^{T} \theta_t^{-1} b_t \tag{19}$$

where $b_t = (X_t'X_t)^{-1}X_t'Y_t$ is the OLS estimator of $\beta + \Delta_t$ assuming that Δ_t is fixed

$\theta_t = [\Omega + \sigma_t^2(X_t'X_t)^{-1}]$.

If Ω and $\sigma_1^2, \sigma_2^2, \ldots, \sigma_T^2$ are unknown, Swamy [14] shows that unbiased estimators can be derived from the expectations of the following expressions. The estimates are computed by replacing the LHS by their observed values and solving for $\Omega, \sigma_1^2, \sigma_2^2, \ldots \sigma_T^2$.

$$E[(Y_t - X_t b_t)'(Y_t - X_t b_t)] = (N-K)\sigma_t^2 \qquad t = 1,2,\ldots,T$$

$$E[BB' - T\bar{b}\bar{b}'] = (T-1)\Omega + (T-1)T^{-1}\Sigma_{t=1}^{T} \sigma_t^2(X_t'X_t)^{-1} \tag{20}$$

where $B = [b_1 \; b_2 \ldots b_T]$ is a KxT matrix of the estimated coefficients defined in (19)

$\bar{b} = T^{-1}B1_T$ is a kx1 vector of the means of the estimated coefficients.

Hence, estimates of the variance components, which in this example represent risk parameters, are relatively easy to compute. The estimates of β in (16) can then be computed by replacing $\Omega, \sigma_1^2, \sigma_2^2, \ldots,$ and σ_T^2 in (19) by their estimated values.

V. Conclusion

A major purpose behind the use of variance components models in economics has been to obtain more efficient estimators of the slope parameters in a regression model. This can be achieved when a model includes random effects by using the GLS estimator instead of the OLS estimator. When the variance components are unknown, the usual procedure is to get initial estimates of these parameters which are then used to compute approximate GLS estimates of the slope parameters. However, with balanced cross-section and time-series data, it is relatively simple to compute the ML estimates of both the slope parameters and the variance components, and expressions for these estimates are presented in Section 2 for a model with one set of random effects.

In a model which includes a lagged dependent variable as a regressor, the OLS estimator of the slope parameters is inconsistent if random cross-section effects are present but are ignored. The consistency of the GLS estimator provides, therefore, an even more substantial reason for considering variance components procedures. In addition, although the OLS estimator is consistent if the effects are treated as fixed, this estimator proves to be relatively inefficient in practice. One disadvantage of the GLS approach is that it is difficult to obtain satisfactory estimates of the variance components. Standard estimation methods based on the expectations of different quadratic forms are not necessarily appropriate when a lagged dependent variable is present.

One of the most promising applications of variance components in economics is related to the analysis of risk behavior, as in this situation the variance components play a direct role in the theory. This contrasts with the preceding examples in which the variance components are considered for purely statistical reasons. An example of a model for a production process is presented in Section IV. The main characteristic of this model is that the slope coefficients contain random components, and this specification implies that risk is associated with the form of the production relationship as well as with the actual quantity produced for a given level of the mean.

References

1. Amemiya, T., "The Estimation of the Variances in a Variance Components Model," International Economic Review, 12, pp. 1-13, 1971.

2. Balestra, P., and M. Nerlove, "Pooling Cross-Section and Time-Series Data in the Estimation of a Dynamic Model," Econometrica, 34, pp. 585-612, 1966.

3. Dhrymes, P. J., Distributed Lags: Problems of Estimation and Formulation, Holden-Day, San Francisco, 1971.

4. Griliches, Z., "Distributed Lags: A Survey," Econometrica, 35, pp. 16-49, 1967.

5. Henderson, C. R., Jr., "Comment on 'The Use of Error Components Models in Combining Cross Section and Time Series Data'," Econometrica, 39, pp. 397-402, 1971.

6. Hildreth, C., and J. P. Houck, "Some Estimators for a Linear Model with Random Coefficients," Journal of the American Statistical Association, 63, pp. 584-595, 1968.

7. Maddala, G. S., "The Use of Variance Components Models in Pooling Cross Section and Time Series Data," Econometrica, 39, pp. 341-358, 1971.

8. Maddala, G. S., and T. D. Mount, "A Comparative Study of Alternative Estimators for Variance Components Models Used in Econometric Applications," Journal of the American Statistical Association, 68, pp. 324-328, 1973.

9. Mount, T. D., and S. R. Searle, Estimating Variance Components in Covariance Models, BU-403-M, Biometrics Unit Mimeograph Series, Cornell University, 1972.

10. Nerlove, M., "Distributed Lags and the Estimation of Long-Run Supply and Demand Elasticities: Theoretical Considerations," Journal of Farm Economics, 40, pp. 301-311, 1958.

11. __________, "Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Sections," Econometrica, 39, pp. 359-382, 1971.

12. __________, "A Note on Error Components Models," Econometrica, 39, pp. 383-396, 1971.

13. Rubin, H., "A Note on Random Coefficients," Chapter 19 in Statistical Inference in Dynamic Economic Models, T. C. Koopmans, ed., Wiley, New York, 1950.

14. Swamy, P. A. V. B., "Efficient Inference in a Random Coefficient Regression Model," Econometrica, 38, pp. 311-323, 1970.

15. Theil, H., Principles of Econometrics, Wiley, New York, 1971.

16. Wallace, T. D., and A. Hussain, "The Use of Error Components Models in Combining Cross Section with Time Series Data," Econometrica, 37, pp. 55-72, 1969.

17. Zellner, A., "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias," Journal of the American Statistical Association, 57, pp. 348-368, 1962.