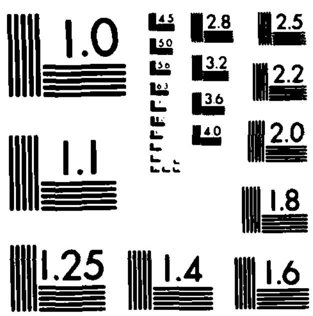
AD-A136 409 POOLING TIME-SERIES AND CROSS-SECTION DATA: AN OVERVIEW 1/1
(U) NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER SAN
DIEGO CA D C BOGER NOV 83 NPRDC-SR-84-8

UNCLASSIFIED F/G 5/2 NL

END
DATE
FILMED
-84
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

A136409

NPRDC-SR 84-8

NOVEMBER 1983

POOLING TIME-SERIES AND CROSS-SECTION DATA: AN OVERVIEW

NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER
SAN DIEGO, CALIFORNIA

NAVY PERSONNEL RESEARCH
AND
DEVELOPMENT CENTER
San Diego, California 92152

NPRDC Special Report 84-8 November 1983

POOLING TIME-SERIES AND CROSS-SECTION DATA: AN OVERVIEW

Dan C. Boger

Department of Administrative Sciences
Naval Postgraduate School

Reviewed by
Barry Siegel

Released by
J. W. Renard
Captain, U.S. Navy
Commanding Officer

Navy Personnel Research and Development Center
San Diego, California 92152

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPRDC SR 84-8	2. GOVT ACCESSION NO. AD-A136409	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) POOLING TIME-SERIES AND CROSS-SECTION DATA: AN OVERVIEW		5. TYPE OF REPORT & PERIOD COVERED Final Report 1 Oct 1980-30 Sep 1981
		6. PERFORMING ORG. REPORT NUMBER 11-82-2
7. AUTHOR(s) Dan C. Boger		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Navy Personnel Research and Development Center San Diego, California 92152		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 63707N Z1187-PN.02
11. CONTROLLING OFFICE NAME AND ADDRESS Navy Personnel Research and Development Center San Diego, California 92152		12. REPORT DATE November 1983
		13. NUMBER OF PAGES 26
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)

Pooled time-series/cross-section
Econometrics
Covariance model
Error components
Variable slope coefficients
Fixed effects
Random effects

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The literature relevant to the estimation and testing of three models that use pooled time-series and cross-section data was reviewed. These models place certain restrictions on the structure of the coefficients and are hierarchical in nature. They enable the user, via hypothesis testing of fitted models or prior information, to relax restrictions.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

FOREWORD

This research and development was conducted in response to Navy decision coordinating paper Z1187-PN (Computer-based Manpower Planning and Programming), subproject PN.02 (Officer Personnel Management Models) and was sponsored by the Deputy Chief of Naval Operations (Manpower, Personnel, and Training) (OP-01). The objective of this subproject is to develop a set of user-oriented, computer-based models and data bases to assist in the development of a Navy officer force that meets the requirements for officer manpower.

During this effort, various econometric techniques were reviewed to estimate models using pooled time-series and cross-section data. The techniques discussed herein will be evaluated in terms of their theoretical and practical utility for forecasting Navy officer personnel loss rates.

J. W. RENARD
Captain, U.S. Navy
Commanding Officer

JAMES W. TWEEDDALE
Technical Director

SUMMARY

Problem and Background

In many practical regression problems, sufficient observations are not available to estimate separate time-series or cross-section equations. One approach to this problem is to combine the time-series and cross-section data into one model. Such models have the practical advantage of allowing the estimated coefficients of the combined sample to incorporate both time-series and cross-section characteristics. For example, the officer retention forecasting model (ORFM) within the structured accession planning system for officers (STRAP-O) uses time-series (1970-1983) and cross-section (community by pay grade by length of service) data to forecast officer loss rates based on a variety of economic scenarios. Time-series observations allow the measurement of the effects of external influences (e.g., civilian unemployment) on officer loss behavior; and cross-section observations, the effects of internal influences (e.g., promotion opportunities). The search for an appropriate model for combining these data provides the motivation for this report.

Objective

The objective of this effort was to review various econometric techniques used to estimate and test models that use pooled time-series and cross-section data.

Approach

The literature relevant to the estimation and testing of three such models was reviewed.

Results

In Model I, the intercept and slope coefficients are assumed to be constant over all cross-sections and time periods. In Model II, the intercepts are allowed to vary over both cross-sections and time periods. Model III allows all coefficients to vary over cross-sections and time. Moreover, in Models II and III, the coefficients may be viewed as either fixed or random effects. In a fixed-effects model, statistical inferences outside the set of sample observations cannot be drawn. In a random-effects model, statistical inferences about the entire population can be drawn.

Conclusions

Models I, II, and III are hierarchical in nature in the sense that fewer restrictions are imposed on the model's structure as the user moves from Model I to Model III. This enables the user to move systematically down the list of models as he or she discovers, either via hypothesis testing of fitted models or prior information, that the imposed restrictions may be successively relaxed. However, it should be recognized that the satisfaction of the assumptions of the model does not validate or invalidate the appropriateness of the model for the application at hand.

CONTENTS

INTRODUCTION

Problem

Whenever observations are available for several cross-sections or individual[1] units over a period of time, the cross-sectional observations may be pooled with the time-series observations under one of several models. The primary advantage of these models is that they allow the estimated coefficients of the combined sample to incorporate both time-series and cross-section characteristics within a specified structure. Another potential advantage of this approach is that, depending upon the model, pooling may result in an increase in the degrees of freedom of the estimated coefficients that, all other things being equal, will yield more efficient estimates.

Cross-section estimates generally differ from time-series estimates made on observations drawn from the same population. In addition to the empirical observation that cross-section variation is generally greater than time-series variation, cross-sections typically will reflect long-run adjustments, whereas time-series tend to reflect short (er)-run reactions. This is because disequilibrium among individuals is generally synchronized in response to overall market forces so that many of the disequilibrium effects tend to disappear in cross-section estimates. A time series of observations on a particular individual will exhibit a less completely adjusted response to the same overall factors. Note, however, that cross-section observations, in general, will also contain some short-run disturbances so that estimates based upon cross-section data will only approximate fully adjusted, long-run coefficients. For further discussion of these points, see Klein (1974), Bass and Wittink (1975) and, especially, Kuh (1959).

The problem involved in using pooled time-series and cross-section data is to specify a model that can adequately allow for differences among cross-section units for a given time period, as well as allow for differences among time-series units for a given cross-section. After the model has been specified and estimated, hypotheses tests must be conducted on the estimated parameters to help in evaluating the model's appropriateness. If the separate time-series and cross-section observation sets are sufficiently large, it may also be advantageous to estimate separate time-series and cross-section models to provide a basis for comparison with the pooled model.

The general linear model may be written as

$$Y = X\beta + u, \tag{1}$$

where
$$Y = [y_{nt}],$$
$$X = [x_{knt}],$$
$$\beta = [\beta_{knt}], \text{ and}$$
$$u = [u_{nt}].$$

The subscript $n=1,\ldots,N$ refers to a particular cross-section unit or individual; $t=1,\ldots,T$ refers to a particular time period; and $k=1,\ldots,K$ refers to a particular nonstochastic independent variable. Therefore, y_{nt} is the observation on the dependent variable for individual n at time t. The observation of independent variable k for individual n at

[1]The terms "individual" and "cross-section" will be used interchangeably in this paper.

time t is x_{knt}. The stochastic disturbance term, u, will be assumed to have the standard Gauss-Markov properties; that is, $E(u_{nt})=0$ and $E(u_{nt}^2)=Var(u_{nt})= \sigma_u^2$. As in standard linear models, β_{knt} is an unknown parameter to be estimated.

Objective

The objective of this effort was to review various econometric techniques used to estimate and test models that use pooled time-series and cross-section data.

Background

In many applied manpower models, sufficient observations are not available to estimate separate time-series or cross-section equations. Quite often, researchers have resorted to combining the time-series and cross-section data into one model. In many instances, these models have been estimated with inappropriate econometric techniques. An equally serious problem is the failure to conduct hypotheses tests of the underlying assumptions of the model prior to estimation. This research provides a systematic review of the available techniques to estimate and test pooled time-series cross-section models.

One example of the use of pooled data in manpower modelling is the officer retention forecasting module (ORFM) within the structured accession planning model for officers (STRAP-O) (Siegel, 1983). ORFM forecasts Navy officer loss behavior based on a variety of economic scenarios. In addition to increasing the number of available observations, the use of pooled data by ORFM provides the opportunity to incorporate both external and internal variables into the model. Variations in loss behavior over time, for example, capture the effects of variations in external variables, such as unemployment rates. Conversely, variations in loss behavior within a given year across communities, pay grades, and length of service (LOS) capture the effect of internal variables, such as promotion opportunities.

The ORFM approach is actually a two-stage process. In stage I, a cost-of-leaving (COL) is calculated via dynamic programming techniques for each unrestricted line (URL) community, pay grade, and LOS cell for years 1970 through 1983. COL is defined as the difference between the present value of future earnings from remaining in the Navy 1 additional year, and then making the "optimal" stay or leave decision, and the present value of future earnings from leaving the Navy and entering the civilian labor market immediately. Among the parameters required to derive COL estimates are military basic pay, regular military compensation, civilian age-earnings profiles, military promotion, and involuntary separation probabilities.

In stage II, voluntary loss rates are related via logistic regression analysis to the COL estimates and other variables that are hypothesized to influence officer retention behavior. The choice of model specification and estimation technique is critical at this stage, since ORFM is using pooled time-series (1970-1983) and cross-section (community by pay grade by LOS) estimates of the cost of leaving the military.

APPROACH

The literature relevant to the estimation and testing of three models that use pooled time-series and cross-section data was reviewed. Each of these models places certain restrictions upon the structure of the β_{knt}. Model I assumes that the β_{knt}, the intercept and slope coefficients, are constant over all individuals and all time periods. Model II also assumes that the slope coefficients are constants but allows the intercept coefficients to vary over individuals and time. Model III assumes that all coefficients (β_{knt}) can vary over both individuals and time. Note that some further restrictions on the structure of the β_{knt} are necessary in Model III since, without any structure, there are KxNxT parameters to be estimated from NxT observations.

Additionally, in Models II and III, the β_{knt} may be viewed as either fixed or random. In the fixed-effects case, it is assumed that the sample of NxT observations is equivalent to the population under consideration; that is, there is no interest in making inferences about any set of observations other than that under consideration. In the random-effects case, there is no interest in making inferences about the population from which the observations are merely one sample. Hence, coefficients are forced to be random variables with means and variance structures that are estimated from the observations. In Model II, the fixed-effects approach is known as the covariance model, while the random-effects approach is referred to as the error components or variance components model.

RESULTS

Model I--All Coefficients Constant

This model assumes that, in equation (1), $\beta_{knt} = \beta_k$, so that there is no variation over either individuals (cross-section units) or time. This is equivalent to running one large pooled regression with NT-K degrees of freedom resulting. This model is the standard approach to pooling.

To test the implicit hypothesis $\beta_{knt} = \beta_k$ *for all k, n, and t*, it is necessary to allow the hypothesis to be violated and perform a statistical test. The way in which the hypothesis is to be violated is primarily a question of data availability. To violate the implied hypothesis, separate cross-section regressions are run for each time-period or separate time-series regressions for each individual. Either or both, if available, of these sets of regression equations can then be used to perform a Chow (Maddala, 1977) test for structural change of the regression coefficients between the individualized equations and the pooled equation. Note that the primary determinant of how the individualized equations are estimated usually depends on whether one has a small cross-section of a long-time series or a short-time series of a large cross-section. In the luxurious case of a long-time series of a large cross-section, one should perform two tests: one to test for stability across individuals; and the other, for stability across time. In the event that the null hypothesis is rejected, this testing will help determine what further model is appropriate.

As an example of testing for stability of coefficients across time, let $\hat{u}$ be the (NTx1) vector of ordinary least-squares (OLS) residuals resulting from estimation of equation (1) under the pooled hypothesis. Let $\hat{u}_t$ be the (Nx1) vector of OLS residuals resulting from the estimation of a different equation (1) for each time period, t . Then the statistic P may be formed as

$$P = \frac{(Q^* - Q) \,/\, (T-1)K}{Q/\,(N-K)T} ,$$

where $Q^* = \hat{u}' \hat{u}$ and $Q = \sum_{t=1}^{T} \hat{u}_t' \hat{u}_t$. P has the Fisher F Distribution with (T-1)K and (N-K)T degrees of freedom. It should be noted that this test assumes that all Gauss-Markov assumptions hold for each of the individualized regression equations, whether they consist of a cross-section of time-series equations or a time series of cross-section equations. Maddala (1977) discussed this particular test, as well as several conditional tests that are also available.

If it cannot be assumed that each of the individualized regression equations meets all of the Gauss-Markov assumptions, then generalized least-squares may be applied to the entire pooled sample. Kmenta (1971) gives results for a time series of cross-section equations in which the coefficients are the same for all individuals and all time periods, the disturbance vector for each individual follows a first-order autoregressive process, and the disturbances for different individuals are heteroskedastic and mutually correlated. Note that Kmenta's model can be modified to allow for an arbitrary time-series process as long as it can be identified and estimated.

Model II--Constant Slope, Variable Intercept Coefficients

The next level of complexity of modelling the pooled data is to allow the varying intercept coefficients to capture differences in behavior over individuals and over time, while holding the slope coefficients constant over individuals and time. The coefficients of this model can be written as

$$\beta_{knt} = \begin{cases} \beta_k & , k \neq 1 \\ \alpha + \gamma_n + \delta_t & , k=1 . \end{cases}$$

Therefore, the entire model can be written as

$$Y_{nt} = \alpha + \gamma_n + \delta_t + \sum_{k=2}^{K} x_{knt} \beta_k + u_{nt} ; \tag{2}$$

$$n=1,\ldots,N \; ; \; t=1,\ldots,T \; ;$$

where $\alpha + \gamma_n + \delta_t$ is the intercept for the nth individual at time t, α is the "mean" intercept for all observations, γ_n is the difference from α for the nth individual, and δ_t is the difference from α for the t^{th} time period. Note that γ_n is common across all time periods and δ_t is common across all individuals.

This discussion will closely follow that of Mundlak (1978). The appropriate estimation procedure depends upon whether γ_n and δ_t are assumed to be fixed or random. If they are assumed to be fixed, equation (2) may be estimated as a covariance model. If they are assumed to be random, equation (2) may be estimated as a variance components or error components model. Before discussing the selection of fixed versus random effects, the estimation procedures for each submodel will be discussed fully so that similarities will become apparent.

Note that equation (2) is the most general formulation of Model II. A restricted version of this model exists where either the time or the individual effect is assumed to be absent. Further discussion of this restricted version of the model may be found in Balestra and Nerlove (1966), Maddala (1971), Nerlove (1971a), and Swamy (1971). Further discussion of the model in equation (2) may be found in Mundlak (1978), Nerlove (1971b), Swamy (1971), Swamy and Arora (1972), and Wallace and Hussain (1969).

Covariance Model

When γ_n and δ_t are treated as fixed effects, one of the γ_n's and one of the δ_t's are redundant. Otherwise, the model represented by equation (2) is not of full rank and, hence, is not estimable via ordinary matrix inversion algorithms. The restrictions $\sum_n \gamma_n = 0$ and $\sum_t \delta_t = 0$ need to be imposed to maintain a model of full rank. (A discussion of estimability, restrictions, and alternative parameterizations may be found in Scheffé (1959).) The model may be reparameterized so that $\gamma_n^* = \gamma_n - \gamma_1$ for $n=2,\ldots,N$ and $\delta_t^* = \delta_t - \delta_1$ for $t=2,\ldots,T$. Then, for the nth individual, the model may be written as

$$y_n = \alpha \iota_T + \gamma_n^* \iota_T + \begin{pmatrix} 0' \\ I_{T-1} \end{pmatrix} \delta^* + X_{Rn} \beta_R + u_n ,$$

where y_n is a $(T\times1)$ vector of observations on the dependent variable for the nth individual; ι_T is a $(T\times1)$ vector of ones; δ^* is a $[(T-1)\times1]$ vector of δ_t^*; X_{Rn} is the $[T\times(K-1)]$ matrix without a constant term that is the reduced version of X_n, the $[T\times K]$ matrix implicit in equation (1); $\beta_R' = (\beta_2,\ldots,\beta_K)$, as in equation (2); $u_n' = (u_{n1},\ldots,u_{nT})$; and the other elements are conformable. Noting that $\gamma_1^* \equiv 0$, the entire set of NT observations may be written as

$$y = [\iota_{NT}\; Z_1\; Z_2\; X_R] \begin{bmatrix} \alpha \\ \gamma^* \\ \delta^* \\ \beta_R \end{bmatrix} + u, \tag{3}$$

where

$$Z_1 = \begin{pmatrix} 0' \\ I_{N-1} \end{pmatrix} \otimes \iota_T ,$$

$$Z_2 = \iota_N \otimes \begin{pmatrix} I_{T-1} \\ \theta' \end{pmatrix},$$

$$y' = (y_1,\ldots y_n),$$

$$X_R' \; (X_{R1}',\ldots,X_{RN}') ,$$

$$u' = (u_1',\ldots, u_N') , \text{ and}$$

γ^* is a $[(N-1)\times1]$ vector of γ_n^*.

This is the fixed-effects version of Model II written as equation (1). The assumptions on the disturbance term are that $E(u) = 0$ and $E(uu') = \sigma_u^2 I_{NT}$. Therefore, via the Gauss-Markov Theorem, the OLS estimator of the $[(N+T+K-2) \times 1]$ parameter vector in equation (3) is minimum-variance for all linear estimators and is unbiased.

Several hypotheses concerning the coefficients may now be tested using the usual least-squares (OLS) procedures. Of particular interest is the hypothesis of whether anything may be gained from moving from Model I to Model II. This is equivalent to the hypothesis that $\gamma_2^* = \ldots = \gamma_N^* = \delta_2^* \ldots \delta_T^* = 0$. This may be tested using the conventional F test that compares the restricted, via the hypothesis, sum of squares with the unrestricted sum of squares. The restricted sum of squares is available from a fit of Model I to the data. The test statistic is the statistic P given in the section on Model I above.

Another hypothesis of interest is that concerning whether the slope coefficients are constant over individuals and time. To test this hypothesis, the data must be arrayed in a fashion so that different slope coefficients, as well as different intercept coefficients, may be estimated. This requires estimating either a time series of cross-sections or a cross-section of time series with the degrees of freedom constraints $N+K<T$ and $T+K<N$ respectively. Then, using the F test comparing the restricted with the unrestricted sums of squares, the hypotheses of constant versus varying slope coefficients over both individuals and time, as appropriate, may be tested.

Two important problems are associated with the use of this covariance model for pooling. The first is that the dummy variables included in equation (3) for shifting the intercept of the regression equation over both time and individuals do not directly identify the variables that are causing the shifts. This is a standard problem with using dummy variables since the dummies are functioning as proxies for variables that are missing from the model. Therefore, dummy variable coefficients are difficult to interpret. The reader should recall that the dummies represent either cross-section or time-series <u>differences</u> from the overall mean, α.

The second problem involved with the covariance model is that it uses up a substantial number of degrees of freedom. In Model I, K coefficients are estimated using NT observations but, in this covariance model of Model II, $N+T+K-2$ coefficients are estimated using NT observations. An implicit assumption of the above development, now made explicit, is that $N+T+K-2<NT$. Additionally, the statistical quality of the covariance model decreases as the number of coefficients approaches the number of observations.

Another problem that is machine- and implementation-dependent is that $N+T+K-2$ may be too large for inversion. This problem may be overcome by considering the partitioned inverse that yields the OLS estimator for β_R in equation (3). Application of formulas for partitioned inverses and simplification yields:

$$\hat{\beta}_{RW} = (X_R' Q X_R)^{-1} (X_R' Q y).$$

The (NTxNT) matrix Q is idempotent and is defined by

$$Q = I_{NT} - \left(I_N \otimes \frac{\iota_T \iota_T'}{T}\right) - \left(\frac{\iota_N \iota_N'}{N} \otimes I_T\right) + \frac{\iota_{NT} \iota_{NT}'}{NT}. \tag{4}$$

This matrix is a generalization of the usual deviation-from-means matrix $A = I - \iota_N \iota_N' / N$ (see Theil (1971). This matrix arises from averaging equation (2) over n, t, and both n and t and subtracting the two single averages from the sum of equation (2) and the double average. Using the obvious notation, this yields:

$$y_{nt} - \bar{y}_{n\cdot} - \bar{y}_{\cdot t} + \bar{y}_{\cdot\cdot} = \sum_{k=2}^{K} (x_{knt} - \bar{x}_{kn\cdot} - \bar{X}_{k\cdot t} - \bar{X}_{k\cdot\cdot})\,\beta_k + (u_{nt} - \bar{u}_{n\cdot} - \bar{u}_{\cdot t} - \bar{u}_{\cdot\cdot})\,.$$

For this notation, as well as a similar approach to covariance analysis, see Sheffé (1959). In matrix formulation, this may be written as

$$Qy = QX_R\beta_{RW} + Qu\ . \tag{5}$$

Consequently, $\hat{\beta}_{RW}$ may be viewed as the OLS estimator obtained from equation (5) or the GLS estimator from equation (5) where Q is the idempotent covariance matrix of the disturbance term, Qu. Note that Q being idempotent is equivalent to it being the generalized inverse of itself. The covariance matrix for $\hat{\beta}_{RW}$ is given by $\sigma_u^2(X_R' Q X_R)^{-1}$. Because $\hat{\beta}_{RW}$ utilizes the variation of the independent and dependent variables within each individual, each time period, and both individual and time period, it is known as the "within" estimator. As shown by Mundlak (1978), this will play an important role in the error components model.

If $\hat{\beta}_{RW}$ is estimated using the partitioned model, the remaining parameters in the model may be estimated from

$$\hat{\alpha} = \bar{y}_{\cdot\cdot} - \sum_{k=2}^{K} \bar{x}_{k\cdot\cdot}\,\hat{\beta}_{kW}\ ,$$

$$\hat{\gamma}_n = (\bar{y}_{n\cdot} - \bar{y}_{\cdot\cdot}) - \sum_{k=2}^{K} (\bar{x}_{kn\cdot} - \bar{x}_{k\cdot\cdot})\,\hat{\beta}_{kW}\ , \text{ and}$$

$$\hat{\delta}_t = (\bar{y}_{\cdot t} - \bar{y}_{\cdot\cdot}) - \sum_{k=2}^{K} (\bar{x}_{k\cdot t} - \bar{x}_{k\cdot\cdot})\,\hat{\beta}_{RW}\ .$$

These are standard results from covariance analysis, as shown in Scheffé (1959).

The covariance model framework may also be used to introduce problems of heteroskedasticity and autocorrelation among residuals. A recent extension of the covariance model for time series of cross-sections, to include arbitrary intertemporal covariances, has been made by Kiefer (1980).

Error Components or Variance Components Model

The error components or variance component models treat γ_n and δ_t in equation (2) as random variables with

$$E(\gamma_n) = E(\delta_t) = 0 ;$$

$$E(\gamma_n \gamma_m) = \begin{cases} \sigma_\gamma^2 & n=m \\ 0 & n \neq m ; \end{cases}$$

$$E(\delta_t \delta_s) = \begin{cases} \sigma_\delta^2 & t=s \\ 0 & t \neq s ; \text{ and} \end{cases}$$

$$E(\gamma_n \delta_t) = E(\gamma_n u_{nt}) = E(\delta_t u_{nt}) = 0, \text{ for all n and t.}$$

Therefore, the model can be written for the nth individual as

$$y_n = \gamma_n \iota_T + I_T \delta + x_n \beta + u_i ,$$

where $\delta' = (\delta_1, \ldots, \delta_T)$ and x_n and β include a constant term and its coefficient, α. The full model with NT observations may be written as

$$y = (\gamma \otimes \iota_T) + (\iota_N \otimes I_T)\delta + X\beta + u,$$

where $\gamma' = (\gamma_1, \ldots, \gamma_N)$. The covariance matrix for the individualized equation may be written as

$$\Sigma_{nm} = E\left[(\gamma_n \iota_T + I_T\delta + u_n)(\gamma_m \iota_T + I_T\delta + u_m)'\right]$$

$$= \begin{cases} \sigma_\gamma^2 \iota_T \iota_T' + \sigma_\delta^2 I_T + \sigma_u^2 I_T , & n=m \\ \sigma_\delta^2 I_T & , n \neq m \end{cases}$$

so that the complete covariance matrix for all NT observations becomes

$$\Sigma = \sigma_\gamma^2 (I_N \otimes \iota_T \iota_T') + \sigma_\delta^2 (\iota_N \iota_N' \otimes I_T) + \sigma_u^2 I_{NT} .$$

In general, Σ does not have any simplifiable structure, except as noted at the beginning of this section.

If the three error variances, σ^2_γ, σ^2_δ, and σ^2_u, are known, then the GLS estimator for β,

$$\hat{\beta} = (X'\Sigma^{-1}X)^{-1}(X'\Sigma y) \tag{6}$$

is the minimum-variance, unbiased linear estimator. As is shown by Mundlak (1978) and others, if $\hat{\beta}$ is partitioned as $\begin{pmatrix} \hat{\alpha} \\ \hat{\beta}_{RE} \end{pmatrix}$, it is possible to show that $\hat{\beta}_R$ is a matrix weighted average of three other estimators. This partitioning results from

$$\Sigma^{-1} = \frac{Q}{\sigma^2_u} + \frac{Q_1}{\sigma^2_2} + \frac{Q_3}{\sigma^2_3}, \tag{7}$$

where Q is defined in equation (4) above,

$$Q_1 = (I_N \otimes \frac{\iota_T \iota_T'}{T}) - \frac{\iota_{NT}\iota_{NT}'}{NT},$$

$$Q_2 = (\frac{\iota_N \iota_N'}{N} \otimes I_T) - \frac{\iota_{NT}\iota_{NT}'}{NT},$$

$$Q_3 = \frac{\iota_{NT}\iota_{NT}'}{NT} = I_{NT} - Q - Q_1 - Q_2,$$

$$\sigma^2_1 = T\,\sigma^2_\gamma + \sigma^2_u,$$

$$\sigma^2_2 = N\,\sigma^2_\delta + \sigma^2_u, \text{ and}$$

$$\sigma^2_3 = T\,\sigma^2_\gamma + N\sigma^2_\delta + \sigma^2_u = \sigma^2_1 + \sigma^2_2 - \sigma^2_u.$$

Partitioning equation (6) in the manner noted above and using equation (7) yields

$$\hat{\beta}_{RE} = \left[\frac{X_R'Q_1W_R}{\sigma^2_1} + \frac{X_R'Q_2X_R}{\sigma^2_2} + \frac{X_R'QX_R}{\sigma^2_u}\right]^{-1}$$

$$\cdot \left(\frac{X_R'Q_1X_R}{\sigma^2_1}\right)\hat{\beta}^1_R + \left(\frac{X_R'Q_2X_R}{\sigma^2_2}\right)\hat{\beta}^2_R + \left(\frac{X_R'QX_R}{\sigma^2_u}\right)\hat{\beta}_{RW},$$

$$\hat{\alpha} = \bar{y}_{\cdot\cdot} - \sum_{k=2}^{K} \bar{x}_{k\cdot\cdot}\,\hat{\beta}_k$$

where $\hat{\beta}_{RW}$ is the "within" estimator from the covariance model above. The other two estimators come from OLS applied to averaging equation (2) over individuals and time respectively. To see this for $\hat{\beta}_R^1$, note that

$$\bar{y}_{n\cdot} = \alpha + \gamma_n + \sum_{k=2}^{K} \bar{x}_{kn\cdot}\ \beta_k + \bar{u}_{n\cdot}$$

is equivalent to

$$Q_1 y = Q_1 X_R \beta'_R + Q_1 u$$

so that $\quad \hat{\beta}_R^1 = (X'_R Q_1 X_R)^{-1} (X'_R Q_1 y)$.

Similarly, $\hat{\beta}_R^2 = (X'_R Q_2 X_R)^{-1} (X'_R Q_2 y)$. Consequently, $\hat{\beta}_R$ is an efficient, matrix-weighted average of the three estimators: (1) $\hat{\beta}_R^1$, which is based upon the variation over individuals, (2) $\hat{\beta}_R^2$, which is based upon the variation over time, and (3) $\hat{\beta}_{RW}$, which is based upon variation not explained by differences over individuals or over time. For further details of calculations and alternative transformations, see Fuller and Battese (1974), Mundlak (1978), Nerlove (1971b), and Swamy and Arora (1972).

In most cases, the error components, σ_u^2, σ_γ^2 and σ_δ^2, are not known. The following estimators, suggested by Swamy and Arora (1972), are unbiased:

$$\hat{\sigma}_1^2 = \hat{u}'_1\ \hat{u}_1/N\text{-}K\ ,$$

$$\hat{\sigma}_2^2 = \hat{u}'_2\,\hat{u}_2/T\text{-}K\ , \text{ and}$$

$$\hat{\sigma}_u^2 = \hat{u}'\hat{u}/\ [(N-1)(T-1)(K-1)]\ ,$$

where $\hat{u}_1 = Q_1 y - Q_1 X_R \hat{\beta}_R^1$, $\hat{u}_2 = Q_2 y - Q_2 X_R \hat{\beta}_R^2$, and $\hat{u} = Qy - QX_R\hat{\beta}_{RW}$ are the residuals from each of the above estimators. Alternative estimators have been suggested by Amemiya (1971), Fuller and Battese (1973, 1974), Maddala (1971), Nerlove (1971a), Rao (1972), Swamy (1971), and Wallace and Hussain (1969).

When X contains lagged values of the dependent variable, several problems are introduced. The most significant are that the parameters may not be identified and that the error components estimators are no longer unbiased. These problems, as well as some suggested corrections, are described in Berzec (1979), Maddala (1971), Nerlove (1967, 1971a), and Swamy (1974).

Several hypotheses are available for testing after the model has been estimated. Perhaps the most important is the hypothesis of whether anything has been gained by moving from Model I to Model II. This is equivalent to testing the vectors $\gamma=0$ and $\delta=0$, which is equivalent to $\sigma_\gamma^2 = \sigma_\delta^2 = 0$. Under the null hypothesis, the individual components do not exist, so that the OLS estimator of Model I is minimum-variance unbiased. To carry out the test, the estimator from the covariance model may be compared with the OLS estimator from Model I via the F test, which compares the restricted and unrestricted sums of squares. Details are given in the section on the covariance model above. An asymptotic test of $\sigma_\gamma^2 = \sigma_\delta^2 = 0$ based upon OLS residuals from the regression of y on X is available from Breush and Pagan (1980).

Another test of interest is whether the slope coefficients are equal over the cross-section and time-series subsamples. As pointed out by Maddala (1977), these tests are asymptotic when performed within the framework of an error components model. The test is again the standard F-test using restricted and unrestricted sums of squares, where the restriction is that the slope coefficients must be constant over individuals and time. Further details are given in the section on the covariance model above.

Attempts to generalize the error components model in the directions of a Bayesian analysis and an errors in variables approach have been suggested by Swamy and Mehta (1973) and Chamberlain and Grilichas (1975) respectively.

Conclusions for Model II

As noted above, the fixed effects assumption of the covariance model implies that it is desirable to make inferences concerning only the sample at hand. The random effects assumption of the error components model, however, assumes that the γ_n and δ_t are random variables and that the individuals and time periods can be regarded as some random samples from larger populations. The desire in the error components model is to make inferences about the larger populations. Additionally, the error components model assumes that there is no correlation between γ_n and X_n or between δ_t and X_t.

Mundlak (1978) states that, in both the fixed-effects and random-effects cases, γ_n and δ_t may be considered as random but, in the fixed-effects case, inference about the sample is conditional upon the values of γ_n and δ_t observed in the sample. On the other hand, in the random-effects case, specific distributional assumptions are made about γ_n and δ_t, so that unconditional inference is appropriate. Because no specific distributional assumptions are made on γ_n and δ_t in the covariance model, it can be used for a (conceivably) wider class of problems. If, however, the distributional assumptions of the error components model are correct, a more efficient estimator is obtained than that available from the covariance model. Mundlak (1978) points out, however, that it may be preferable to use an estimator that possesses bias but that has a lower mean square error than an available unbiased estimator.

It should be noted here that the primary interest in both the covariance model and the error components model is in obtaining good estimates of the slope coefficients. Using the Mundlak (1978) interpretation, in the covariance model the intercepts are random but are conditional on the sample values of the pure individual and pure time effects. They become the coefficients of dummy variables simply because the true way that individual and time effects enter the model is not known. In the error components model, the effects are treated as unconditionally random with assumed distributions but, again, they function as proxies for unknown cross-section and time series factors. Perhaps a skeptical interpretation of Model II is that, although it is desirable to account for the time-series and cross-section effects, it is difficult or impossible to model the process adequately.

An additional point, made by Maddala (1977), is that systematic, as opposed to random, variation in the intercepts implies that the error components model is not appropriate. He also states that it is important to check whether there is a systematic pattern in the residuals. This should reveal whether the residuals are heteroskedastic or autocorrelated and, hence, what model is appropriate.

A final point is that, since the error components model assumes no correlation between the effects and the X_{knt}, this should be checked and tested. The null hypothesis is that there is no asymptotic correlation and an asymptotic test can be carried out via a comparison of the estimator from the covariance model, $\hat{\beta}_{RW}$, and the estimator from the error components model, $\hat{\beta}_{RE}$. The test statistic, which is proposed by Hausman (1978) for use in the errors in variables problem, is

$$m = (\hat{\beta}_R - \hat{\beta}_{RE})' D^{-1} (\hat{\beta}_R - \hat{\beta}_{RE}),$$

where D is the difference between the estimated covariance matrices for $\hat{\beta}_R$ and $\hat{\beta}_{RE}$. Note that m is asymptotically distributed as a χ^2 (K-1) variable.

Model III--Variable Slope Coefficients

The final logical level of complexity of modelling the pooling process is to allow all the coefficients to reflect differences in behavior over individuals and over time. It is no longer necessary to treat intercepts differently from slopes, so that the general model can be written as

$$y = X\beta + u,$$

as in the introduction, or as

$$y_{nt} = \sum_{k=1}^{K} x_{knt} \beta_{knt} + u_{nt}, \tag{8}$$

in the two previous sections. In the most general case shown in equation (8), there is an immediate difficulty. Since there are KNT parameters to be estimated from NT observations, some further structure must be placed on the coefficients. The usual structure is to set

$$\beta_{knt} = \alpha_k + \gamma_{kn} + \delta_{kt}, \tag{9}$$

where β_k represents some mean effect over all individuals and time periods, γ_{kn} represents the effects due to specific individuals, and δ_{kt} represents the effects due to specific time periods. Note that now the individual and time effects, as well as the mean effect, are specific to individual columns of X.

As in Model II, γ_{kn} and δ_{kt} may be assumed to be either fixed or random and, again, this is how the various approaches to estimation of Model III will be dichotomized. However, in this case, there are ways of estimating the fixed effects, all of which are extensions of the covariance model presented above. The following discussion of both types of Model III will closely follow Hsiao (1974, 1975).

The version of Model III represented by equations (8) and (9) is the most general formulation of the varying slopes hypothesis. As in Model II, restricted versions of the formulation are available in which either the time or individual effects are absent. For the fixed-effects case, these will be treated below since the assumptions are crucial to the estimation procedures. In the random-effects case, further discussion of the restricted version may be found in Rosenberg (1973a) and Swamy (1970, 1971, 1973, 1974). Further discussion of the general formulation is found in Hsaio (1974, 1975).

Fixed Effects

When γ_{kn} and δ_{kt} are treated as fixed parmeters, an immediate useful simplification is to permit either γ_{kn} or δ_{kt} to disappear so that a cross-section of time-series equations or a time series of cross-section equations respectively is produced. Additionally, the coefficients can be defined as β_{kt} or β_{kn} respectively, so as to incorporate the time effects or individual effects into the mean effects of the coefficients of the explanatory variables. Once in this form, Zellner's (1962) seemingly unrelated regressions model (or joint GLS) can be applied to allow different time-series equations in each cross-section or different cross-section equations in each time series. The joint estimation can be carried out for separate β_{kn} or β_{kt}, as well as common β_k, and the hypothesis of equal β_k may be tested. Details are contained in Zellner (1962). One should check that the assumptions of the Zellner model hold for a particular application prior to estimation and testing. Estimation techniques for differing assumptions on error structures are contained in Judge, Griffiths, Hill, and Lee (1980).

The full model of equations (8) and (9) can also be viewed as a covariance model in the case where γ_{kn} and δ_{kt} are assumed to be fixed effects. In the matrix version of the model, the matrix of explanatory variables has the dimension NT x (T + N + 1)K while its rank is only (T + N - 1)K, so that 2K independent linear restrictions must be imposed on the γ_{kn} and δ_{kt}. Hsaio (1975) suggests the natural restrictions $\sum_{n=1}^{N} \gamma_n = \sum_{t=1}^{T} \delta_t = 0$ as the 2K restrictions necessary. Again, standard F tests of comparing restricted with unrestricted sums of squares may be carried out here, just as in the case of Model II. Note that the degrees of freedom constraint for this full version of the covariance model is NT>(T+N-1)K and this is likely to be violated, except in the case of large N and large T.

Random Effects

Given the model of equations (8) and (9), which can be written

$$y_{nt} = \sum_{k=1}^{K} (\alpha_k + \gamma_{kn} + \delta_{kt})X_{knt} + u_{nt} ,$$

the model for the nth individual can be written as

$$y_n = X_n\alpha + X_n\gamma_n + \bar{Z}_n\delta + u_n , \tag{10}$$

where y_n is (Tx1), X_n is (TxK), $\gamma_n' = (\gamma_{1n},\ldots,\gamma_{kn})$,

$$\delta' = (\delta_1',\ldots,\delta_T'),\ \delta_t' = (\delta_{1t},\ldots,\delta_{kt}),\ u_n' = (u_{n1},\ldots,u_{nt}) ,$$

$$\underset{(TxTK)}{\bar{Z}_n} = \begin{bmatrix} X_{n1}' & & \\ & \ddots & \\ & & X_{nt}' \end{bmatrix},$$

and $X_{nt}' = (X_{1nt},\ldots,X_{knt})$. Hsiao's (1974, 1975) assumptions are that

$$E\ [u_n] = 0 = E\,[\gamma_n]\ ,\quad E\ \delta_t = 0 ,$$

$$E\ [u_n u_m'] = \begin{cases} \sigma_u^2 I & \text{for } n\text{-}m \\ 0 & \text{for } n \neq m , \end{cases}$$

$$E\ [\gamma_n \gamma_m'] = \begin{cases} A & \text{for } n=m \\ 0 & \text{for } n \neq m \text{ , and} \end{cases}$$

$$E\ [\delta_s \delta_t'] = \begin{cases} B & \text{for } s=t \\ 0 & \text{for } s \neq t . \end{cases}$$

Additionally, he assumes γ_n, δ_t, and u_n are all uncorrelated and that A and B are diagonal with elements a_k and b_k respectively.

Rewriting (10) to include all NT observations gives

$$y = X\alpha + Z\gamma + \bar{Z}\delta + u ,$$

where Z is block diagonal with X_n as the nth diagonal block, ($\bar{Z}' = \bar{Z}_1',\ldots,\bar{Z}_N'$), and the other vectors and matrices are stacked in the natural way.

Given the assumptions above, as well as the assumption that γ and δ are random, the covariance matrix for the complete error term is

$$\Omega = E\,[(Z\gamma + \bar{Z}\delta + u)(Z\gamma + \bar{Z}\delta + u)'] = Z(I_N \otimes A)Z' + \bar{Z}\,(I_T \otimes B)\bar{Z}' + u^2 I_{NT}\,.$$

Then the GLS estimator $\hat{\alpha} = (X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}y)$ is the minimum-variance unbiased linear estimator for α and has the covariance matrix $(X'\Omega^{-1}X)^{-1}$.

Unless A, B and σ_u^2 are known, they will have to be estimated from the sample. Hsiao gives two estimation techniques for the unknown parameters of Ω. One is a direct maximum likelihood approach, while the other is an indirect minimum norm quadratic unbiased estimator (MINQUE) approach. Since the estimation procedures are both rather complicated, the reader is referred to Hsiao (1974, 1975) for details. Hypotheses concerning constancy of coefficients over time may be tested via procedures given in Hsiao (1974).

Conclusions for Model III

As in Model II, there is a choice of whether to assume the effects are fixed or random. The conclusions given previously concerning Model II carry over completely in the case of Model III. Again, the most important consideration is correlation of the random coefficients with the explanatory variables. If this correlation is present, then the GLS estimator for the random effects case is biased, so it is probably better to use the fixed-effect model. If, however, such correlation is not present and the distributional assumptions of the random-effects model are reasonable, this model will provide more efficient estimates than the fixed-effects model. Also, systematic variation of the individual time effects, as opposed to random variation, is an indication that the fixed-effects model is preferable.

Extensions of Model III in various directions have been suggested by Singh and Ullah (1974), Swamy (1974), and Swamy and Mehta (1975, 1977).

CONCLUSIONS

The random-effects models above, when combined with the coefficients of the explanatory variables, imply the mixed model of the classical analysis of variance literature. Additionally, the random-effects models can be viewed as an intermediate step between the OLS model where all the coefficients are the same for all individuals and time periods and the covariance models where the coefficients are different for all individuals and time periods.

A point that should be heavily stressed when making choices over pooling models is the importance of the sample sizes for both the cross-section and the time-series samples. These sample sizes determine both how the data can or cannot be grouped and place constraints on the choice of particular estimation methods. Specifically, estimators of the random-effects models maintain their properties only when a reasonably large sample is available.

One model, not seen in the pooling literature, is one in which the slope coefficients vary over individuals and time periods but the intercept coefficients are constant. This is a variation of Model II that can be handled by a minor extension of Model III. It is not clear that this model would have any applications in economics.

The organization of the three models discussed in this report is hierarchical in nature. This enables the user to move down the list of models as he or she discovers, via either hypothesis testing of fitted models or prior information, that imposed restrictions must be successively relaxed. It should be noted that Models II and III imply that $X\beta$ does not account for all the modelling variation in y. In the random-effects case, this remaining variation is modeled by classifying its distribution; in the fixed-effects case, it is modeled via dummy variables. A third alternative that exists in this modelling scheme is to allow random variation of the individual and time components but place it within a structural framework. This leads into the varying parameter literature. Note that this is the next logical step from Model III above. Some applications of varying parameter models to pooling problems are considered by Johnson and Rausser (1975), Rosenberg (1973b), Saxonhouse (1977), and Greenwood, Ladman, and Siegel (1981). The literature on varying parameter models is voluminous and still growing. Good overviews are available in Judge, Griffiths, Hill, and Lee (1980) and Maddala (1977).

It is important to remember that, after pooling problems have been dealt with, the application of standard regression tools to the pooled problem at hand should not be forgotton. An excellent example of this is provided by Izan (1980).

REFERENCES

Amemiya, T. The estimation of variances in a variance-components model. International Economic Review, 1971, 12, 1-13.

Balestra, P., & Nerlove, M. Pooling cross-section and time series data in the estimation of a dynamic model: The demand for natural gas. Econometrica, 1966, 34, 585-612.

Bass, F. M., & Wittink, D. R. Pooling issues and methods in regression analysis with examples in marketing research. Journal of Marketing Research, 1975, 12, 414-425

Berzec, K. The error components model: Conditions for the existence of maximum likelihood estimates. Journal of Econometrics, 1979, 10, 99-102.

Breush, T. S. & Pagan, A. R. The LaGrange multiplier test and its applications to model specification in economics. Review of Economic Studies, 1980, 47, 239-253.

Chamberlain, G., & Griliches, Z. Unobservables with a variance-components structure: Ability, schooling, and the economic success of brothers. International Economic Review, 1975, 16, 422-450.

Fuller, W. A., & Battese, G. E. Transformations for estimation of linear models with nested error structure. Journal of the American Statistical Association, 1973, 68, 626-632.

Fuller, W. A., & Battese, G. E. Estimation of linear models with crossed-error structure. Journal of Econometrics, 1974, 2, 67-78.

Greenwood, M. J., Ladman, J. R., & Siegel, B. S. Long-term trends in migratory behavior in a developing country: The case of Mexico. Demography, 1981, 18, 369-388.

Hausman, J. A. Specification tests in econometrics. Econometrica, 1978, 46, 1251-1272.

Hsiao, C. Statistical inference for a model with both random cross sectional and time effects. International Economic Review, 1974, 15, 12-30.

Hsiao, C. Some estimation methods for a random coefficient model. Econometrica, 1975, 43, 305-325.

Izan, H. Y. To pool or not to pool? A reexamination of Tobin's food demand problem. Journal of Econometrics, 1980, 13, 391-402.

Johnson, S. R., & Rausser, G. C. An estimating method for models with stochastic, time varying parameters. American Statistical Association, Proceedings of the Business and Economic Statistics Section, 1975, 356-361.

Judge, G. G., Griffiths, W. E., Hill, R. C., & Lee, T-C. The theory and practice of econometrics, New York: John Wiley, 1980.

Kiefer, N. M. Estimation of fixed effect models for time series of cross-sections with arbitrary intertenporal covariance. Journal of Econometrics, 1980, 14, 195-202.

Klein, L. R. A textbook of econometrics, (2nd Ed.). Englewood Cliffs, NJ: Prentice-Hall, 1974.

Kmenta, J. Elements of Econometrics, New York: Macmillan, 1971.

Kuh, E. The validity of cross-sectionally estimated behavior equations in time series applications. Econometrica, 1959, 27, 197-214.

Maddala, G. S. The use of variance components in pooling cross section and time series data. Econometrica, 1971, 39, 341-358.

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

Mundlak, Y. On the pooling of time series and cross section data. Econometrica, 1978, 46, 69-85.

Nerlove, M. Experimental evidence on the estimation of dynamic economic relations from a time series of cross sections. Economic Studies Quarterly, 1967, 18, 42-74.

Nerlove, M. Further evidence on the estimation of dynamic economic relations from a time series of cross sections. Econometrica, 1971, 39, 359-382. (a)

Nerlove, M. A note on error components models. Econometrica, 1971, 39, 383-396. (b)

Rao, C. R. Estimation of variance and covariance components in linear models. Journal of the American Statistical Association, 1972, 67, 112-115.

Rosenberg, B. Linear regression with randomly dispersed parameters. Biometrika, 1973, 60, 65-72. (a)

Rosenberg, B. The analysis of a cross section of time series by stochastically convergent parameter regression. Annals of Economic and Social Measurement, 1973, 2, 399-428. (b)

Saxonhouse, G. R. Regressions from samples having different characteistics. Review of Economics and Statistics, 1977, 59, 234-237.

Scheffé, H. The analysis of variance, New York: John Wiley, 1959.

Siegel, B. S. Methods for forecasting officer loss rates (NPRDC Tech. Rep. 83-30). San Diego: Navy Personnel Research and Development Center, August 1983.

Singh, B., & Ullah, A. Estimation of seemingly unrelated regressions with random coefficients. Journal of the American Statistical Association, 1974, 69, 191-195.

Swamy, P. A. V. B. Efficient inference in a random regression model. Econometrica, 1970, 38, 311-323.

Swamy, P. A. V. B. Statistical inference in random coefficient regression models, New York: Springer-Verlag, 1971.

Swamy, P. A. V. B. Criteria, constraints, and multicollinearity in random coefficient regression models. Annals of Economic and Social Measurement, 1971, 2, 429- 450.

Swamy, P. A. V. B. Linear models with random coefficients. In P. Zarembka (Ed.). Frontiers in Econometrics, New York: Academic, 1974.

Swamy, P. A. V. B., & Arora, S. S. The exact finite sample properties of the estimators of coefficients in the error components regression models. Econometrica, 1972, 40, 253-260.

Swamy, P. A. V. B., & Mehta, J. S. Bayesian analysis of error components regression models. Journal of the American Statistical Association, 1973, 68, 648-658.

Swamy, P. A. V. B., & Mehta, J. S. Bayesian and nonbayesian analysis of switching regressions and of random coefficient regression models. Journal of the American Statistical Association, 1975, 70, 593-602.

Swamy, P. A. V. B., & Mehta, J. S. Estimation of linear models with time and cross-sectionally varying coefficients. Journal of the American Statistical Association, 1977, 72, 890-898.

Theil, H. Principles of econometrics, New York: John Wiley, 1971.

Wallace, T. D., & Hussain, A. The use of error components models in combining cross section with time series data. Econometrica, 1969, 37, 55-72.

Zellner, A. An efficient method of estimating seemingly unrelated regressions and tests of aggregation bias. Journal of the American Statistical Association, 1962, 57, 348-368.

DISTRIBUTION LIST

Deputy Under Secretary of Defense for Research and Engineering (Research and Advanced Technology)

Chief of Naval Operations (OP-01), (OP-115) (2), (OP-130), (OP-130E), (OP-130E40), (OP-130E40A), (OP-130E40B), (OP-162)

Chief of Naval Research (Code 200), (Code 270), (Code 441)

Superintendent, Naval Postgraduate School

Commander, Air Force Human Resources Laboratory, Brooks Air Force Base (Scientific and Technical Information Office)

Commander, Air Force Human Resources Laboratory, Williams Air Force Base (AFHRL/OT)

Commander, Air Force Human Resources Laboratory, Wright-Patterson Air Force Base (AFHRL/LR)

Defense Technical Information Center (DDA) (12)

Commanding Officer, Air Force Personnel Center, Randolph Air Force Base

Commanding General, Army Military Personnel Center, U.S. Army

President, Center for Naval Analyses

84

C