RANDOM COEFFICIENT MODELS:

A SIMPLE ILLUSTRATION OF A NEW ESTIMATION TECHNIQUE*

ELMAR R. DA CRUZ

HAMILTON, OCTOBER 1974

* Term Paper presented at McMaster University-Canada, based on the author's M.A. Thesis from the University of Manchester - England.

III-2

Mc MASTER UNIVERSITY

ELMAR R. DA CRUZ

ECON - 768 - TERM PAPER

RANDOM COEFFICIENT MODELS

A SIMPLE ILLUSTRATION OF A NEW ESTIMATION TECHINIQUE

1. PRESENTATION OF THE NEW TECHNIQUE

In a recent paper, Burns (1974) introduced a simplified technique to estimate random coefficient models, and used it to investigate the results obtained by Friedman and Meiselman (1963), on monetary policy. Burns claimed that this method, tough computationally simpler, is equivalent to Hildreth and Houck's (1968) in the sense that it will generally approach the same final solution over a number of iterations.

The simplicity of Burns method stems from the fact that it uses standard regression techniques. It is my purpose to evaluate Burn's claim by comparing his results with those obtained by the adoption of Hildreth and Houck's method, since in his article Burns made no such attempt. Unfortunately there isn't at the moment any computer programme available at Mc Master that provides Hildreth and Houck's method. Therefore I will defer to a later date the comparison I referred to, and I will restrict the analysis here to a mere presentation of Burn's method, and compare it with OLS results presented in an econometrics textbook. In doing so I shall attempt to draw your attention to certain weaknesses in his method.

The rationale of Burn's method is that one can choose simultaneously the values of the estimaton $\hat{\beta}$ and $\left[\dfrac{\sigma_1^2}{\sigma_1^2}\right]$ [see (1.2.3) and (1.2.10) pages 9 and 12 - in my survey, respectively] that minimise the following sum of squares:

$$(1) \qquad S = \sum_{t=1}^{T} \left[\tilde{e}_t^2 - E\,\tilde{e}_t^2\right]^2$$

Where $\tilde{e}_t$ is an undertermined residual, since we do not know yet the value of β..

It is evident that Theil (1971) and Hildreth and Houck (1968), chose $\begin{bmatrix} \hat{\sigma}^2_\mu \\ \hat{\sigma}^2_1 \end{bmatrix}$ to minimise an expression similar to (1) but which contained the OLS residuals ℓ_t instead fo the undetermined $\tilde{e}_t$.

To obtain this estimator it is only necessary to racall that

(2) $y_t = (\beta + \delta_t)\, x_t + u_t \qquad t = 1 \ldots\ldots t$

[see (1.2.1) p.9 of the survey for explanation of the symbols]. Will give us the following expression:

3) $u^{*2}_t = \left[y_t - \beta x_t\right]^2 \qquad t = 1 \ldots\ldots t$ where $u^*_t := \delta x_t + ut$

and

4) $E\, u^{*2}_t = \sigma^2_1 x^2_t + \sigma^2_u \qquad t = 1 \ldots\ldots T$

Combining (3) and (4)

5) $\left[y_t - \beta xt\right]^2 - \sigma^2_1 x^2_t - \sigma^2_u = u^{*2}_t - E\, u^{*\,2}_t$

Expanding and rearranging terms:

6) $y^2_t = 2\beta xt\, yt - \beta^2 x^2_t + \sigma^2_1 \quad x^2_t + \sigma^2_u + u^{*2}_t - E\, u^{*2}_t$

7) $y^2_t = \sigma^2_u + 2\beta x_t\, y_t - (\beta^2 - \sigma^2_1)\, x^2_t + \tilde{u}t$

Where $\tilde{u}t = u^{*2}_t - E\, u^{*2}_t$

Burns assumes that for estimation purposes these transformed explanatory variables ($x_t\, y_t$ and x^2_t) can be taken as uncorralated with $\tilde{u}_t$ and since $E\, \tilde{u}_t = 0$, one can apply OLS to(7) and result will be unbiased estimates of the parameters involved. However when it comes to the retrieval of the estimators of the original parameters (i.e., $\hat{\beta}$ and $\hat{\sigma}_1^2$) non-linearities are involved when obtaining $\hat{\sigma}_1^2$ and one should no loger expect the property of unbiasedness of the latter parameter.

Turning now to the problem of the accuracy of the estimates one might possibly hope that Burn's method would eliminate the so called accuracy problem of RcR (see page 15 of the survey). Unfortunately this is not the case. The estimates of $\hat{\sigma}_1^2$ and $\hat{\sigma}_\mu^2$ can still take negative values with positive probability as it is the case with the RCR models developed by Theil and Mennes (1959). We shall however expect high values of R^2 in the estimation of equation (7) (see the numerical example below) but this is almost solely due to the relative magnitude of its variables, in comparison with other estimations methods (see Burns p-25 for an elaboration of this point).

The problems do not end here. As Burn's admits, the heteroscedastic term in (4) invalidates the usual tests of significance and t statistics stemming form the conventional OLS package [he doesn't speculate about the possibility of using GLS to estimate (7) but this could be done, specially on an iterative basis].

Another complication is the presence of the disturbance term. If there is a constant term then its corresponding regressor is a column of units. The implication is that in this particular model (equation 7) multicollinearity is likely to be a serious problem. This is due to the term $x_t\ y_t$. It involves the dependent variable, and hence it is highly collinear with the intercept of (7).

To overcome this problem Burns proposes an alternative estimating equation (p.25 - equation 50 in his article), using 2SLS. what I used are here data in deviation from the mean, and thus the results present only the slopes. This enables a direct comparison with Wonnacott and Wonnacott's example, in which they used OLS in a fixed coefficient context. Therefore I used OLS to estimate (7), rather than 2SLS.

2. THE DATA AND RESULTS

The data used are given in both Wonnacott's books. They refer to their yield/fertilizer example.

It has very nice classroom features due to its simplicity, and furthermore it has an appealing interpretation of the underlying randomness of the regression coefficients. The reason is that one could invoke the well estabilished fact that nature is random and that fertilizer application (x_t) is affected by natural causes. Therefore if fertilizer application changes, the dependent variable(yield = yt) will react with a random response rate βt, i.e.,

$$8) \quad y_t = \beta t \; x_t + u_t \qquad u_t = 1 \; .... \; T$$

Where we expressed y_t and x_t as deviations from their respective means.

Note that (8) corresponds to (1-1.1) in the survey and the specification of the probability distribution of β_t is

$$\beta t \sim D(\beta, \sigma_1^2)$$

Wonnacott and Wonnacott gave the following data:

Y	X	$Y = Y - \bar{Y}$	$X = X - \bar{X}$
40	100	- 20	- 300
45	200	- 15	- 200
50	300	- 10	- 100
65	400	5	0
70	500	10	100
70	600	10	200
80	700	20	300

$\bar{Y} = 60$ $\qquad \bar{X} = 400$

Fitted equation:

(9) $\hat{Y}_t = .068 \quad x_t$

$R^2 = .957 \qquad \bar{R}^2 = .973$

Using the same data, and transforming them accordingly to satisfy(7) we obtained the following fit.

(10) $\hat{Y}_t^2 = .125\, x_t\, y_t + .4107\, x_t^2$

(.559) (.363)

$R^2 = \bar{R}^2 = .999$

Noting that $.125 = 2\hat{\beta}$ and that, $.4107 = -(\hat{\beta}^2 - \hat{\sigma}_1^2)$, after straightforward algebraic manipulations we conclude that:

(11) $\hat{\beta} = .063$

(12) $\hat{\sigma}_1^2 = .405$

Two things must be stressed here:

i) As we stated earlier, although the coefficient in (10) are unbiased, only the value of $\hat{\beta}$ will be unbiased. The estimate .405 of $\hat{\sigma}_1^2$ will probably be biasea since it was obtained from a non-linear transformation. This is a major drawback in Burns method.

ii) The standard errons given below (10) are only suggestive. Strictly speaking they are inappropriate, since the term $\tilde{u}_t$ in(7) is heteroscedastic.

The values of the estimates of $\hat{\beta}$ are reasonably close to each other [see (9) and (11)], despite the fact that we are dealing with only seven observations.

The estimate of $\hat{\sigma}_1^2$ (the variance of the coefficient β_1) is .405, which doesn't reveal anything important, unless we could compute valid confidence intervals for $\hat{\sigma}_1^2$.

This is an important point overlooked by Burns. We can only attach singificance to certain parometers and draw useful inferences only after:

i) Certifying ourselves that these parameters are significantly different from zero (given reasonable significance levels).

ii) Or satisfying ourselves that the sign and magnitude of the parameters do indeed confirm our a priori expectations, given that the latter are suficiently strong (this seems to be a highly subjective matter but it is a very important point in applied econometrics).

The lesson we can draw from this simple illustration is that here we are reasonably sure that nature is indeed random, and that the positive value of $\hat{\sigma}_1^2$ merely reinforced our a priori idea.

However, when dealing with real economic data, things are not so simple. One needs in general a lot more support from the data to confirm or reject certain existing theories. In this respect Burns method has indeed to be improved, since it is not sufficiently well equiped to provide condition (i) above.

Hamilton, Ont.
December 1974
E. R. DA CRUZ

Part 2. BRIEF SURVEY OF THE LITERATURE

I	Model I - A Simple Version	0
	1.1. - Specification And Uses	7
	1.2. - Estimation Methods	8
II	Model II- A Simple Version	
	2.1. - Specification And Uses	15
	2.2. - Estimation Methods	20
III	Illustrative Examples	
	3.1. - Model I - Monetary Policy	24
	3.2. - Model II - AGGR. Consumption FN.	27
IV	REFERENCES	31

I Model I - A Simple Version

1.1. - Specification And Uses

Theil (1971 - PP. 622-627) considered the following model:

$$(1\text{-}1\text{-}1)\quad y_t = b_t\, x_t + u_t \qquad t = 1 \,\ldots\ldots\ldots t$$

Its interpretation is that if an explanatory variable x_t increases by one unit, ceteris paribus, the dependent variable y_t reacts with a random change expressed by b_t with a certain mean and a positive variance.

b_t, the random response rate, fluctuates from one observation to the next and follws the distribuition,

$b_t \sim D\ (\beta\ ,\ \sigma_1^2)$

Expressing b_t in terms of its expectation β and its random element (here defined as δt) we have:

$$(1\text{-}1\text{-}2)\quad b_t = \beta + \delta t \qquad t = 1 \,\ldots\ldots\ldots\, t$$

It is assumed that:

(1-1-3) X_t is non-stochastic

(1-1-4) $E\begin{bmatrix} u \\ \delta \end{bmatrix} = 0$ and $E\begin{bmatrix} u \\ \delta \end{bmatrix}\begin{bmatrix} u' & \delta' \end{bmatrix} = \begin{bmatrix} \sigma_{\mu}^{2} I & 0 \\ 0 & \sigma_{1}^{2} I \end{bmatrix}$

Where =

$$u = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_t \end{bmatrix} \quad \text{and} \quad \delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_t \end{bmatrix} = \begin{bmatrix} b_1 - \beta \\ b_2 - \beta \\ \vdots \\ b_t - B \end{bmatrix}$$

There are theree parameters to be estimated:

(i) β

(ii) Var $b_t = \sigma_1^2$

(iii) Var $u_t = \sigma_u^2$

If we set $b_t = \beta$ for all t then $\sigma_1^2 = 0$ i.e., equation (1.1.1.) collapses to the classical two- variable linear model with fixed coefficients. In that case only two parameters would be estimated (β and $\hat{\sigma}_u^2$).

Theil and Mennes (1959) used this model to analyse aggregate time series data on British import and export prices from 1870 to 1952.

Klein (1953 - p.p. 211-225) developed the basic ideas of RCR model I in the context of cross- section data (in individuals). Burns (1974) uses this model to evaluate Friedman and Meiselman's results on monetary policy . We shall examine this application in our illustrative example.

1.2 ESTIMATION METHODS

METHOD I - THEIL AND MENNES (1959)

Combining (1.1.) and (1.1.2) :

(1-2-1) $$Y_t = (\beta + \delta_t) X_t + u_t \qquad t = 1$$

or

(1-2-2) $Y_t = \beta X_t + u_t^*$ t= 1..........t

Where $u_t^* = \;_t X_t + u_t$

Under (1-1-4) we conclude that

$$u_t^* \sim D(0, \sigma_1 X_t^2 + \sigma_u^2)$$

This implies that we are in a heteroscedastic situation and application of OLS to (1-2-2) will result in consistent but inefficient estimates of β. This situation mggests the use of

$$(1\text{-}2\text{-}3) \quad \hat{\beta} = \left(\sum_{t=1}^{t} \frac{X_t \quad Y_t}{\sigma_1^2 X_t^2 + \sigma_u^2} \right) / \left(\sum_{t=1}^{t} \frac{X_t^2}{\sigma_1^2 X_t^2 + \sigma_u^2} \right)$$

as the G.L.S. estimator of β, where the lower case X_t and Y_t indicate deviations from the means of Y_t and X_t respectively.

The variance of this estimator is :

$$(1\text{-}2\text{-}4) \quad E\,(\hat{\beta} - \beta)^2 = \sum_{t=1}^{t} \frac{\sigma_1 X_t^2 + \sigma_u^2}{X_t^2}$$

Which hasto be distinghished from the variance of $b_t = \sigma_1^2$. The latter refers to the variance of the population random coefficient (the estimation procedure will be presented below) and the former refers to the variance of the G.L.S. estimator $\hat{\beta}$.

Needless to say, σ_1^2 and σ_u^2 are not Known. We have to operate with their estimates and therefore the results will be only approx. valid.This method is suggested by theil. Compute the O.L.S. residuals from (1-2-2)

(1-2-5) $e_t = y_t - \hat{b} X_t$ t= 1..........t

$= \;_t X_t - (\hat{b} - \beta) X_t + M_t$ t= 1..........t

Where $\;_t = (b_t - \beta)$ and $\hat{b}$ is the O.L.S estimator of β.

The variance of the t th residual is:

(1-2-6) $\text{var } e_t = \sigma_1 X_t^2 + X_t^2 \text{ Var } \bar{b} + \sigma_u^2 - 2 X_t^2 \text{ cob } (b_t - \bar{b})$

$+ 2 X_t \text{ cov } (b_t\ t_{1\ t}) - 2 X_t \text{ cov } (\bar{b}\ u_t)$

$$= \sigma_1^2 \quad X_t^2 + \frac{\sigma_u^2 \quad X_t^2}{\sum X_t^2} + \frac{\sigma_1^2 \quad X_1^2 \quad X_1^4}{(\sum X_1^2)} + \sigma_u^2 - \frac{2\sigma_1^2 \quad X_t^4}{\sum X_t^{20}} - \frac{2\sigma_u^2 \quad X_t^2}{\sum X_t^2} = \sigma_u^2 \ P_t + \sigma_1^2 \quad QT$$

Where P_t and Q_t are defined as the following Known functions of X_t=

$$P_t = 1 - \frac{X_t^2}{\sum X_t^2}$$

$$Q_t \quad X_t^2 \left[1 - 2 \frac{X_t^2}{\sum X_t^2} + \frac{X_t^4}{(\sum X_1^2)^2} \right]$$

Given that $F\, \ell_t = 0$, it follous that $\text{var } \ell_t = E_{e_1}^2$ so:

(1-2-7) $\ell_t^2 = \sigma_u^2 \quad P_i + \sigma_1^2 \quad Q_t + f_x$

Where $f_t = \ell_t^2 - E \ \ell_t^2$ and $E \quad f_t = 0$

Since P_t and Q_t are known, σ_u^2 and σ_1^2 can be estimated. Just run a regressias of ℓ_t^2 on P_t and Q_t with f_t treated as the disturbance term the regression method depends on the form of the variance - covariance matrix of f_t. If we assume that b_t and u_t (and hence ℓ_t) are normally distributed, and if we neglect the covariances between two different $f_{t}s$.

(see theil 1971; P. 624 for a justifications of this procedure) then we are left with:

$$(1\text{-}2\text{-}8)\quad \operatorname{var} f_t = E(\ell_t^2 - E\,\ell_t^2)^2 = E\,\ell_t^4 - (E\,\ell_t^2)^2 = 2\,(E\,\ell_t^2)^2$$

$$= 2\,(\sigma_u^2 P_t + \sigma_1^2 Q_t)^2$$

Which is twice the square of (1-2-6). Inspection of (1-2-8) reveals that f_t is heteroscedastic and therefore G.L.S. should be used to estimate σ_u^2 and σ_1^2.

Take (1-2-7) and run a L.S. regression of ℓ_t^2 on P_t and Q_{t-2} This step leads to preliminary variance estimates S_u and S_1^2 (of σ_u^2 and σ_1^2) :

$$(1\text{-}2\text{-}9)\quad \begin{bmatrix} \sum P_t\,\ell_t^2 \\ \sum Q_t\,\ell_t^2 \end{bmatrix} = \begin{bmatrix} \sum P_t^2 & \sum P_t Q_t \\ \sum P_t Q_t & \sum Q_t^2 \end{bmatrix} \begin{bmatrix} S_u^2 \\ S_1^2 \end{bmatrix}$$

Now apply wrighted L.S. to (1-2-7)

$$(1\text{-}2\text{-}10)\quad \begin{bmatrix} \sum W_t\, P_t\,\ell_t^2 \\ \sum W_t\, Q_t\,\ell_t^2 \end{bmatrix} = \begin{bmatrix} \sum W_t\, P_t^2 & \sum W_t\, P_t\, Q_t \\ \sum W_t\, P_t\, Q_t & \sum W_t\, Q_t^2 \end{bmatrix} \begin{bmatrix} \hat{\sigma}_u^2 \\ \sigma_1^2 \end{bmatrix}$$

$$\text{Where } W_t = \frac{1}{2} \cdot (S_u^2 P_t + S_1^2 Q_t)^{-2}$$

This step results in the final variance estimates $\hat{\sigma}_\mu^2$ and $\hat{\sigma}_1^2$. The 2 x 2 matrix in the r.h.o. of (1.2.10), when inverted, provides an approximate variance - covariance matrix of $\hat{\sigma}_u^2$ and $\hat{\sigma}_1^2$.

Finally estimate β, the mean response rate. Formula (1.2.3.) is used for this purpose with the final estimates $\hat{\sigma}_u^2$ and $\hat{\sigma}_1^2$ replacing the unknowns.

SOME METHODS COSIDERED BY HILDRETH AND HOUCK (1968)

They suggested several alternative estimator of RCR modèls.

The estimating equation being considered here is (1.2.2). Let ris start with the O.L.S. estimator:

$$(1.2.11)\quad \hat{b} = \frac{\sum x_t y_t}{\sum x_t^2}$$

$$= \beta + \frac{\sum x_t u_t^*}{\sum x_t^2}$$

$$= \beta + \frac{\sum x_t^2 (b_t - \beta) + \sum x_t u_t}{\sum x_t^2}$$

Were all summations are over t, and t=1 .... T.

Under assumptions (1.1.4) it is easily seen that this estimator is unbiased and it has the following variance:

$$\text{var } \hat{b} = \frac{\sigma_u^2}{\sum x_t^2} + \frac{\sigma_1^2 \sum x_t^4}{(\sum x_t^2)^2}$$

Recalling that $u_t^* = \sigma_t X_t + M_t$, it followsthat var $u_t^* =$ $\sigma_1 X_1^2 + \sigma_u^2$, i.e., $\hat{b}$ is inefficient in view of this heteroscedasticity.

If σ_1^2 and σ_u^2 were known we could of course obtain efficient estimates of β (see 1-2-3). Unfortumately this is not the case, and Hildreth and Hyuck devote a great deal of attention in deviloping estimaton that can replace σ_1^2 and σ_u^2.

One of the simplest is to take the L.S. residuals ℓ_t in (1-2-5) and compute their variance (see 1-2-6). This we can obtain (1-2-7), repeated here for the sake of convenience

$$\ell_t^2 = \sigma_u^2 P_t + \sigma_1^2 Q_t + f_t \qquad \text{(1-2-7) repeated}$$

$$= \begin{bmatrix} \sigma_u^2 & \sigma_1^2 \end{bmatrix} \begin{bmatrix} P_t \\ Q_t \end{bmatrix} + f_t$$

or

$$\text{(1-2-12)} \quad \ell_1^2 = G_t + f_t$$

Where $\bar{\alpha} = \begin{bmatrix} \sigma_u^2 & \sigma_1^2 \end{bmatrix}$ and $G_t = \begin{bmatrix} P_t \\ Q_t \end{bmatrix}$

This is merely a simplified notation of equation 13 (page 586 in their article). They maggest the application of O.L.S. to (1.2.12)

$$\text{(1-2-13)} \quad \alpha^* = \begin{bmatrix} G_t^1 & G_t \end{bmatrix}^{-1} G_t^1 \ \ell_t^2$$

This is precidly the preliminary variance estimates $\begin{bmatrix} S_u^2 \\ S_1^2 \end{bmatrix}$

given in (1.2.9). This the only difference with the method outlined by theil and Menns is that Hildreth and Houck don't bother about the final variance estimates $\hat{\sigma}^2_1$, $\hat{\sigma}^2_u$. In any event both cases present a common problem namely that $\hat{\sigma}^2_1$, $\hat{\sigma}^2_u$ may take negative values. We shall mention this problem in the next sub-section.

Bearning this problem in mind, Hildreth and Houch considered a more elaborate estimator, obtained by restricted least-squares, the elements of α are restricted to be non-negatice, and this is achieved by means a quadratic programming algorithm. Although intuitevely appealing, this method has not yet received widespread support, partly because of the complicated computed programme that it requires.

We have yet to consider Burns (1974) method. It will be discussed in section (3.1.) together with the illustrative example.

A NOTE ON THE ACCURACY OF THE ESTIMATES

It is important to note that (1.2.7) or (1.2.12) has a very poor fit. In fact theil (1971 - p.626) shows that its sistematic part ($\sigma^2_u P_t + \sigma^2_1 Q_t$) accounts for only ONE-THIRD of the behaviour of ℓ^2_t. This fact leads to very large sampling variability and as a consequence $\hat{\sigma}^2_u$ and $\hat{\sigma}^2_1$ (or α^*) may take negative values, despite the fact that we know a priori that they should be non-negative. [1]

Therefore a sizable sample is needed to estimate these parameters with a reasonable degree of precision (and also because more degress of freedom are needed - see pag 2). The alternative method of constrained least-squares (quadratic programming) is unfortunately seldom adopted in practice. The conclusion is that further research is necessary in this topic.

MODEL I I – ASIMPLE VERSION

2.1. – Specification and Uses

Model II was developed to cater for certain specifie situations, specially in the case where we have panel data. In his classic work on aggregation, theil (1954) showed that there is in general an aggregation bias in micro-relations estimated from aggregated data. This bias stems from the presence of non-corresponding micro-parameters as determining factors of macro coefficients.

Zellner (1966) inspired by the pioneer work of klein (1953) reconsidered the aggregation problem in terms of RCR models, rather than the fisced coefficient approach used in theil's original work. Zellner was able to show that under the following model there is no aggregation bias (we shall Keep zell ner's notation to facilitate cross reference) Let:

2.1.1.) $Y_i = X_i \beta_i + u_i$ $\qquad$ i= 1........n

Where:

Y_i is a KX_1 victor of observations on a dependent variable X_i is a TXK matrix of observations on K non- stochestu explanatory variables. β_i is a KX1 vector of random coefficients. u_i is a T X 1 victor of disturbance terms n is the number of micro units.

We are interested in estomating β, the mean response rate, given by:

$\beta = E\beta_i$ $\qquad$ i= 1......n

let:

(2-1-2) $\beta_i + \delta_i$ $\qquad$ i= 1......n

Where δ_i is a random victor (K X 1) with $E\delta_i = 0$

Combining (2-1-1) and (2-1-2)

(2-1-3) $Y_i = X_i (\beta + \delta_i) + n_i$ $\qquad$ i=1........n

Summing it over i:

$$\sum Y_i = \sum X_i \beta + \sum X_i \delta_i + E n_i$$

or

$$(2.1.4)\ Y = X\beta + \sum X_i \delta_i + n$$

The expectation of the macro- least squares estimator $\bar{\beta}$ is :

$$E\,\bar{\beta} = E\,(X'X)^{-1} X' Y$$

$$= (X'X)^{-1} X' (X\beta + \sum X_i \delta_i + n)$$

$$= \beta$$

Since $E\,(X'X)^{-1} X' \sum X_i \delta_i = 0$ and $E\,(X'X)^{-1} X'n = 0$, i.e., there is no aggregation bras in the estimation of β, the mean coefficient vector of the micro parameters β_i .

Zellner's assumption that coefficient vectors of different individuals are random deawinp from the rame multivariate distribution, is a compromise between the limiting assumptions that coefficient vectors are fixed and the same for all individuals and that coefficient vector are fixed and different from one individual to another. The former case is often found to be too restrictive given that micro-units usually differ in their bchavionr, whill the latter assumation involves the use of an excessively large number of parameters since separate regressions have to be used for each unit. In this sense Zellner's model is indeed very ingenions. The coefficient vector in (2.1.3) varies randomly across units, and once an individual is selected, a drawing on its coefficient vector is kept the same for all observations on that individual.

This the randonness of the coefficients may be attributed to the random selection of units. Recall however that theil (or klein) assumed in his model that the coefficients are random from one observation to the next (among different individuals or among different time periods).

This latter assumption is rather more general and hence model I has a wider applicability than model II.

To complete the specification of model II we rewrite equation (2.1.4) as (2.1.5) $Y = X\beta + u^*$

Where =

$$y' = \sum_{i=1}^{n} Y_i$$

$$X = \sum_{i=1}^{n} X_i$$

and

(2.1.6) $$u^* = \sum_{i=1}^{n} X_i \delta_i + u$$

and we formalise the assumptions:

(2.1.7) $$\text{rank } X_i = K \text{ and } \begin{cases} n > K \\ T > K \end{cases}$$

(2.1.8) $$E\, u_i = 0 \quad \text{and } E\, u_i u_j' = \begin{cases} 0 & \text{if } i \neq J \\ \sigma_{ii} I & \text{if } i = J \end{cases}$$

(2.1.9) β_i' are random coefficient vectors independently and indentically distributed across the micro-units with:

$$\beta_i = \beta + \delta_i$$

$$E\beta_i = \beta$$

$$E\, \delta_i \delta_i^1 = \Delta$$

$$E\, u_i \delta_j^1 = 0 \text{ for all } i \text{ and } J$$

Δ is a KXK positive definite varaiance -covariance matrix of the coefficient vector of the $i^{\underline{th}}$ individual

(2-1-10) $$\Delta = \begin{bmatrix} \sigma_{11} & \sigma_{12} \cdots\cdots\cdots & \sigma_{1K} \\ \sigma_{21} & \sigma_{22} \cdots\cdots\cdots & \sigma_{2K} \\ \vdots & & \\ \sigma_{K1} & \sigma_{K2} & \sigma_{KK} \end{bmatrix}$$

The diagonal elements indicate the variance of the $K\underline{th}$ coefficient. The off-diagonal elements denote the contemporaneous covariances between any two pairs of different coefficients.

Note that (2-1-8) allows different disturbance variances for each individual. The implication is that we have the following parameters to be estimated:

(i) K elements of β.

(ii) $\frac{1}{2}$ K (K+1) distinct elements of $\triangle$

(iii) N disturbance variances σ_{ii} i= 1.........n.

Before we proceed to the tedions estimation details of madel II , it would be very intructive at this moment to diverge from the main path and to look briefly at the main uses of R C R models since this will help us to understand more clearly the limitations of model II.

The main uses of RCR models are as fallows:

a) In cross section studies, as suggested by Klein (1953), where it is very difficult to justify the absence of parameter variation across units.

b) In panel data studies where variations both across individuals and through time render the use of fixed coefficient models a very donbtful task.

c) In analyses which are affected by variables too for ontside their scope to be successfully handled by a conventional fixed coefficient model.

A typical example mentioned in Bowden (1968) is the partial adjustament mechanism embodied in the flexible accelerator in the inventory investment models. The coefficient of adjustment is usually assumed fixed in the published literature, but it is not defficult to find situations where it should be allowed to become random. Under tight business, conditions inventory adjustaments to a desired level are hender to be attained, giving a smaller adjustament coefficient value than in other periods of the cycle.. Accounting for this variation under a conventional fixed coefficient model is no easy kask. Therefore to avoid depatong too much from the scope of the inventory model, the adjustament coefficient should be regarded as a random variable and estimated by

RCR model I.

d) When the relative stability of the response rates is being studied. This case will be discussed in our illustrative example of model I.

In general, when a study involves many economic units with different riations it is only natural to treat the parameters of the representative relationship as stochastic variables.

A quastion that naturally arises at this point is in which cases should model I be used and in which situations model II is more appropiate. As we hinted before, in some cases model II is not appropriate due to its restrictive assumptions. Nevertheles it is generally accepted that the estimation of model II is computationally simpler (since it uses standard multiviriate techniques as we shall see below) whenever panel data is involved.

When the cross-sections involve different units from year to year (and as a consequence we are NOT dealing with panel data), the parameters cannot be considered as being random drawings from the same probability distribution. If this the case then model I should be used.

When dealing with panel data of firms that possess high rates of technological change, model I is again recommended, since model II assumes that a random drawing on the same coefficient vector is kept the same for the whole observation period. However model II could be used if the time period of analysis is relatively short. An obvions advice is that each case has to be analysed carefully before deciding which RCR model to use.

Model II is likely to give satisfactory results in the following cases:

(i) In the analysis of consumption and income data for different regions. Swamy (1971-ch- VI) studied a simple Keynesian consumption function together with a modified version of Friedman's permanent income hypothesis, in a sample of 24 countries for the period 1955-1963 Hissimple Keynesian model is the subject of our illustrative example II.

(ii) In the analysis of corporate investment. Swamy (1971 - chapter V) has added a new dimension to Grunfeld's (1958) corporate investment model. He confirmed Grunfeld's theony bit showed that the assumption of same coefficient vector for all corporations is not appropiate.

(iii) In constant elasticity functions for bank deposits. Feige (1964) used annual time series for 48 continental states in the U.S. over the period 1949-1959 to estimate the demand functions for liquid assets. He used dummy variables in the estimation process and found that demand deposits and savings and association shares are n substitutes. Lee (1966) used a conventional fixed coefficient model with the same data and concluded that savings and loan association shares are close substitutes for demand deposits. Feige and Swamy (1972) using R C R model II confirmed Feige's results, reversed Lee's conclusions, and demonstrated that the assumption of identical regression coefficient vectors for all 48 states is inappropriate.

2.2. ESTIMATION METHODS OF MODEL II

Following Zellner (1966) we can apply O.L.S. to (2.1.4). The L.S. macro-estimator $\tilde{\beta} = (X'X)^{-1} X'Y$ is an unbiased estimator of β, ie., there is no aggregation bias as we saw insection 2.1.

However the sampling error of $\tilde{\beta}$ is =
$\tilde{\beta} - \beta = (X'X)^{-1} X' \sum_{i=1}^{n} X_i \delta_i + (X'X)^{-1} X'n$ which reflects two sources of random ness:

(i) That arising from δ_i, the random element of the coefficient vector;

(ii) That arising from the macro-disturbances u.

The conslusion is that the O.L.S standard errons are inappropriate since they reflect only one source of randomness, specifically (ii) mentioned above. Therefore one cannot rely on inferences stimming from L.S. standard errons, and an alternative estimation method had be divised.

Swamy (1971 - ch IV) proposed the following method of estimating β.

Let H (θ) be the variance - covariance matrix of n^* i.e., from (2.1.6) and the assumptions of the model we have:

(2-2-1)

$$E\, n^{*}\, n^{*\prime} = H(\theta) = \begin{bmatrix} X_1 \Delta X_1' + \sigma_{11} I & 0 \ldots\ldots\ldots\ldots 0 \\ 0 & X_2 \Delta X_2^1 + \sigma_{22} I \ldots\ldots 0 \\ \vdots & \vdots \\ 0 & 0 \ldots\ldots x_n \Delta x_n^1 + \sigma_{nn} I \end{bmatrix}$$

Where is a T X T unit matrix; 0 is a T X T null matrix and Δ and 6_{ii} were defined in section 2.1..

To estimate β (the mean response rate vector) in (2.1.5) Swamy suggests the application of Aitken's G.L.S. :

$$(2.2.2) \quad \hat{\beta}(\theta) = \left[X^1 H(\theta)^{-1} X \right]^{-1} X^1 H(\theta)^{-1} y$$

$$\left[\sum_{j=1}^{n} x_j' (x_j \Delta x_j' + \sigma_{jj} I)^{-1} x_j \right]^{-1} \left[\sum_{i=1}^{n} x_i' (x_i \Delta x_i' + \sigma_{ii} I)^{-1} y_i \right]$$

A matrix result given in Rao (1965 a - p.29) is:

$$(2.2.3) \quad (X_i \Delta X_i^1 + \sigma_{ii} I)^{-1} = \frac{Mi}{\sigma_{ii}} + X_i (X_i' X_i)^{-1} \left[\Delta + \sigma_{ii} (X_i^1 X_i)^{-1} \right]^{-1}$$

$(X'_i X_i)^{-1} \quad X'_i$

Where $M = I - X_i (X'_i X_i)^{-1} X'_i$. This result can be verified by premultiplying both sides by $(X'_i \Delta X'_i + \sigma_{ii} I)$.

Swamy uses this result to obtain a macro-coefficient estimator which consists in a weighted average of the individual micro-coefficients β_i, with weights taken directly from (2.2.3) :

(2.4.4.)

$$w_i(\theta) = \left[\sum_{j=1}^{n} \left\{ \Delta + \sigma_{jj} \quad (x'_j x_j) ^{-1} \right\}^{-1} \right]^{-1} \left\{ \Delta + \sigma_{ii} \quad (x'_i x_i)^{-1} \right\}^{-1}$$

This the macro estimator (2.2.2) is alternatively expressed as:

$$(2\text{-}2\text{-}5) \quad \hat{b} (\theta) = \sum_{i=1}^{n} W_i (\theta) \quad \hat{b}_i$$

Where $\hat{b}_i$ is the estimator of the maioro coefficients β_i specified in (2.1.1).

There are two ways of estimating these $\hat{b}_i$:

(i) Run separate L.S. regressions of each Y_i on the corresponding X_i, i.e., $\hat{b}_i = (X'_i \quad X_i)^{-1} \quad X'_i \quad Y_i$

(ii) Use Zellner's SURE (Seemingly Unrelated Regression Eqns.) Which may be appropriate under certain circunstances.

Both cases will be examined in the illustrative example of model II.

The variance - covariance matrix of $\hat{b}$ (0) is:

$$(2\text{-}2\text{-}6) \quad E \; (\hat{b} \; (\theta) - \beta) \; (\hat{b} \; (\underline{0}) - \beta)^1 = (X^1 \; H(\theta)^{-1} \quad X)^{-1}$$

Recall that we assumed identically and independently distributed coefficient vector b_i. This the b_i ' s (i= 1.... n) provide n different linear unbiased and uncorrelated estimator of the same parametric vector β.

Since the variances of these $\hat{b}_i$ are:

$$(2\text{-}2\text{-}7) \quad Var \; \hat{b}_i = \sigma_{ii} \; (X'_i \quad X_i) ^{-1} \qquad i= 1 \ldots\ldots\ldots\ldots\ldots n$$

There are n different variances of the estimators. Thus the best way to pool them into a single estimator is to take a wlighted average of all $\hat{b}_i$'s with weights inversely proportional to their variances (given in 2.2.7). This is precisely what Swanny proposes in (2.2.5).

As expected Δ and σ_{ii} (the variance covariance matrix of the coefficient vector and the variances of the disturbances) are unknown. An unbiased estimator of σ_{ii} is:

$$(2.2.8) \quad S_{ii} = \frac{n_i' M n_i}{T-K} = \frac{\ell_i' \ell_i}{T-K} \qquad i = 1 \ldots\ldots n$$

Where $\ell_i = M_i n_i$ and $M_i = I - X_i (X_i' X_i)^{-1} X_i'$

i.e., ℓ_i is the L.S. residuals from a fit of y_i on x_i.

To estimate Δ Swamy suggests treating the least squares estimators b_i as a random sample of size n.

Define:

$$(2\text{-}2\text{-}9) \quad S_{\hat{e}} = \sum_{i=1}^{n} \hat{b}_i \hat{b}_i^1 - \frac{1}{n} \sum_{i=1}^{n} \hat{b}_i \sum_{i=1}^{n} \hat{b}_i^1$$

Where $\frac{S\hat{b}}{m-1}$ is simple the sample varaince-covariance matrix of the $\hat{b}_i$'s (for a quicker understanding of these steps see the illustrative example in the section 3.2)

Express $\hat{b}_i$ as :

$$\hat{b}_i = b_i + (X_i' X_i)^{-1} X' n_i$$

so that it can be used in the expectation of (2.2.9)

$$E \; S\hat{b} = n (\Delta + \beta\beta') + \sum_{i=1}^{n} \sigma_{ii} (X'_i \; X_i)^{-1} - (\Delta + \beta\beta')$$

$$- (n-1) \beta\beta' - \frac{1}{n} \sum_{i=1}^{n} \sigma_{ii} (X_i' X_i) ^{-1}$$

Rearranging terms:

$$E \; S_{\hat{b}} = n - 1 \Delta + \frac{n-1}{n} \sum_{i=1}^{n} \sigma_{ii} (X_i' X_i)^{-1}$$

This, an unbiased estimator of Δ can be expressed as:

$$(2.2.10)\quad \hat{\Delta} = \frac{S\hat{b}}{n-1} - \frac{1}{n} \sum_{i=1}^{n} S_{ii} \; (X' X)^{-1}$$

Note that it depends critically on the assumption (2.1.7) to (2.1.10) .

Using the above estimatiors of 6_{ii} and Δ we are now able to compute w_i (θ) and hence $\hat{b}$ (θ) (see 2.2.5). The latter is the swamy estimator of β, the mean coefficient vector of the micro b_i' s.

I I I - ILLUSTRATIVE EXAMPLES

3.1 - MODEL I - MONETARY POLICY

Burns (1974) analyses Friedman and Meiselman's (1963) model on monetary policy. They used a fixed coefficient approach with variables expressed in levels. After careful considerations Burns concluded that the model should be expressed in terms of first differences under R C R model I. Burn's estimation method is slightly different from that of Hildreth and Houck, but it is computationally simpler (since a standard OLS computer programme can be used) and it will approach the same solution over successive iterations.

Withont going into estimation details, we shall attempt to highlight Burn's results. He draws the attention of some sizable errors of inference that have been made in the past when fixed coefficient models were used in situations where R C R models would be more appropriate. In particular he analyses Friedman and Meiselman's data (1963-p.260) from 1929 to 1958, which gave rise to their conclusion that monetary effects were a great deal more stable on the level of economic activity than expenditure factors (fiscal policy).

Their estimating equation is:

$$(3.1.1)\quad C_t = \beta_1 + \beta_2 A_t + \beta_3 M_t + M_t \qquad t= 1 \ldots\ldots\ldots t$$

where =

C_t = Indicator of the level of economic activity.

A_t = Indicator of the level of expenditure factors.

M_t = Indicator of the level of monetary factors.

β_2 and β_3 are the respective fixed response rates, and β_1 is the intercept.

Burns specifies:

$$C_t^2 = \beta_1 + 2\beta_2 C_t A_t + 2\beta_3 C_t M_t - (\beta_2^2 - \sigma_{22}) A_t^2 - (\beta_3 - \sigma_{33}) M_t^2 - 2(\beta_2\beta_3 - \sigma_{23}) A_t M_t \qquad t=1\ldots\ldots t \tag{3-1-2}$$

Which conforms to his estimations method.

It was used by Burns in five regressions, with the results given in table (1) below. These are only first round parameter estimates, but they are sufficient for the purpore in hand.

Table (2) converts the direct estimates into estimates of β_2 and β_3 for a comparison with Friedman and Meiselman's results.

Table (1) - DIRECT PARAMETER ESTIMATES

DEPENDENT VARIABLE	β_1	$2\beta_2$	$2\beta_3$	$\beta_2^2 - \sigma_{22}$	$\beta_3^2 - \sigma_{33}$	$\beta_2\beta_3 -$ 33	R^2
C	-34.70	-7.74	2.88	0.23	2.02	-2.20	1.00
C	-	-1.72	2.88	0.32	2.02	-2.21	0.99
C	45.00	1.86	2.28	-1.20	1.25	2.51	1.00
Δ C	-	0.26	3.26	0.11	1.56	0.15	0.93
Δ C	2. 32	0.25	3.23	0.11	1.55	0.15	0.93

TABLE (2) - CONVERTED PARAMETER ESTIMATES

EQUATION NUMBER	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\sigma}_{22}$	$\hat{\sigma}_{33}$	$\hat{\rho}_{23}$
1	- 0. 87	1. 44	0. 53	0. 06	------ 0.9
2	- 0. 86	1. 44	0. 42	0. 06	------ 0.9
3	.0. 93	1. 14	2. 06	0. 05	0.6
4	0. 13	1. 63	< 0	1. 10	---------
5	0. 13	1. 61	< 0	1. 05	---------
FRIEDMAN	- 0. 87	1. 52	-----	----	---------

Equation (1) : Variables measured in levels, with a constant term.

Equation (2) : Similar to (1) but with the intercept constrained to be zero.

Equation (3) : Uses two Stage Least Squares.

Equation (4) : Variables are expressed in first differences rather than levels. The intercept is constrained to be zero.

Equation (5) : Similar to (4) but with unconstrained intercept.

Inspection of (1) and (2) reveals that they are a like in terms of the response rate estimates. Furthermore these estimates do not contradict Friedman Meirelman's conclusions. However equation (3) exhibits a quite different picture. The expenditure response rate β_2 is no longer negative, and in addition its associated variance σ_{22} increases considerably. These values suggest considerable instability, which implies that further investigations are necessary before valid inferences can be made.

Equations (4) and (5) show that the variance of the expenditure response rate (σ_{22}) is negative, when we know a priori that it should be at least zero. This when first differences are used, no evidence exists to support Friedman and Meiselman's theory that increased use of expenditure factors will provoke larger random oscillations in the level of economic activity than the adoption of monetary policy. Furthermore the variance σ_{33}(of the response rate of monetary factors) is positive in both equations. These positive variances support the idea that an increased use of monetary factors will be associated with increased random fluctuations (i.e. instability) in the level of economic activity. Therefore, as Burns points ont, it is quite impossible with this set of data to make any definite inference to establish the superioritiy of monetary policy over expenditure policy If anything the apposite view seems to be more plausible. This , there is some evidence which suggests that Friedman and Meiselman's conclusions are not valid. Withont a RCR model this evidence probably would not be available.

We have seen that for policy purposes the estimation of the variances of the response rates may be quite important. If this is the case then a RCR model should be used, since fixed-coefficient models do not convey such information.

3.2 MODEL I I - AGGREGATE CONSUMPTION FUNCTION

We shall consider an example of model II developed by Swamy (1971 - chapter VI). He analyses cross country data on aggregate consumption expenditures and in his simple Keynesion consumption model he assumes that such a model is defined for each country appearing in the sample.

These he tests whether the marginal propensity to consume (MPc) is the same for all countries. If the MPc instead of being fixed is distributed randomly across countries with the same mean and variance, then the application of model II is valid, i.e., the data on all countries can be pooled since they contain information on the same probability distribution.

Having pooled the data we shall be able to estimate the mean MPc with considerable more precision this that obtained by separate regressions on each country.

Let consumption of the $i^{\underline{th}}$ country in the $t^{\underline{th}}$ year (it) be related to measured disposable.

Income $(X_i\ t)$ as:

(3.2.1.) $$C_{it} = \alpha_i + \beta_i\, x_{it} + n_{it}$$

$$(i = 1 \ldots\ldots n) \qquad (t=1 \ldots T)$$

α_i is the intercept

β_i is the m.p.c. of each country.

The source of data for Swamy's study is the U.N. Yearbook of National Statistics (1965). He collected data on consumption expenditure and disposable income relating to 24 countries for the period 1955-1963.

In the fixed coefficient model there are two ways to handle the problem. Fist we assume:

(3.2.2.) $\beta_1 = \beta_2 = \ldots\ldots\ldots \quad \beta_n = \beta$

i.e. same (and fixed) m.p.c. for all countries.

(3.2.3) $E\, n_i \quad t=0$ and $E\, n_i t\, n_{js} = \begin{cases} \zeta_n^2 & \text{if } i=j \text{ and } t=0 \\ 0 & \text{otherwise.} \end{cases}$

(3. 2-4) X_{it} is non - stochastic.

In the first case, by pooling the time series data from all the 24 countries we obtain 192 observations on each variable. The O.L.S estimator of the pooled model is:

$$(3.2.5) \quad \tilde{\beta} = \left[\sum_{j=1}^{n} X_j' \; X_j \right]^{-1} \left[\sum_{i=1}^{n} X_j^i \; X_i \; \bar{b}_i \right]$$

where:

$$(3.2.6\ ) \quad \bar{b}_i = (X_i' \; X_i)^{-1} X_i' \; Y_i$$

is obtained by using separate L.S. regressions for each country.

The use of (3.2.5) gave rise to the following L.S. estimates:

$$(3.2.7) \quad C_{it} = 10.538 + 0.893\, X_i t$$

$$(3.071) \qquad (0.004)$$

The figures in parentheses indicate the standard errors calculated by the square roots of the diagonal elements of:

$$S^2 \left[\sum_{i=1}^{n} X_i' \; X_i \right]^{-1}$$

Where $S^2 = \dfrac{e'\ell}{nt-k}$ is computed from the L.S residuals resulting from 3.2.5..

As first glance the fit seemst to be sucessful. The std-errors of the m p.c. is remarkably low, only 0.004. However assumption (3.2.3) is very restrictive. It is unlikely that all countries have the same disturbance variances.

So let us see the second case, by relaxing that assumption to:

(3.2.8) $E\ n_i t = 0 \qquad E\ n_i t\ n\ js = \begin{cases} \sigma_{ii} & \text{if } i = j \text{ and } t = S \\ 0 & \text{otherwise} \end{cases}$

This we are now to consider Zellner's SURE approach. The joint estimator of the pooled model becomes:

$$(3.2.9)\quad \beta_j = \left[\sum_{j=1}^{n} \frac{X_j' \ X_j}{S_{jj}} \right]^{-1} \left[\sum_{i=1}^{n} \frac{X'_i \ X_i}{S_{ii}} \ \hat{b}_i \right]$$

Where:

S_{ii} was defined in (2-2-8)

$\hat{b}_i$ was defined in (3.2.6)

The use of $\hat{\beta}_j$ gives the following estimates:

(3.2.10) $C_{it} = \underset{(0.278)}{7.429} + \underset{(0.001)}{0.906}\ X_{it}$

As expected the precision in the estimation of the mpc increased considerably.

Let us test Now the hypothesis of a fixed coefficient vector across countries. (i.e., same intercept and same M.P.C. for all countries)

For this purpose, in addition to the previons assumptions we considerer:

(3.2.11) u_{it}'s are normally distributed.

(3.2.12) $\quad H_o = \beta_1 = \beta_2 = \ldots\ldots\ldots \beta_n = \beta$

Under H_o, Zellner (1962) shows that:

$$(3.2.13)\ \frac{1}{(m-1)K} \cdot \sum_{i=1}^{n} \frac{(\hat{b}_i - \beta)^1\ x_i'\ xi\ (\hat{b}_i - \beta)}{S_{ii}}$$

has an F distribution with $(n-1)K$ and $(T-K)n$ degrees of freedom.

The value of (3.2.13) is 46.94, well above the 5% of the F statistic with 46 and 144 degrees of freedom. We are forced to reject the hypothesis of fixed and same m.p.c. for all countries. One way of relaxing condition (3.2.12) is to allow different but fixed mpc for different countries. This of course rules ont the possibiloty of pooling data-from different countries. On the other hand if we can assume that the mpc of different countries is not a fixed parameter, but a random variable which is independently distributed with the same means and the same variance across countries, then we can use model II and pool the data on all these countries.

Therefore our final approach is the RCR model II. Take (3.2.1) and run separate L.S. regression for each country using (3.2.6).

The standard errons of these estimates are obtained by taking the square roots of $S_{ii}(X_i' X_i)^{-1}$ $i = 1 \ldots n$ (see 2.2.7). When [S_{ii} is given by (2-2-8)]. From (2.2.10) compute $\hat{\Delta}$ which is:

$$(3.2.14) \quad \hat{\Delta} = \begin{array}{c} \\ \alpha_i \\ \beta_i \end{array} \begin{array}{c} \begin{array}{cc} \alpha_i & \beta_i \end{array} \\ \begin{bmatrix} 2864.8 & 2.0778 \\ & 0.0075 \end{bmatrix} \end{array}$$

$W_i(\hat{\theta})$ in (2.2.4) can then be obtained, and hence $\hat{b}(\theta)$ (defined in 2-2-5) can be solved, giving the estimates of the means of α_i and β_i as follows

$$(3.2.15) \quad C_{it} = \underset{(10.925)}{89.553} + \underset{(0.0168)}{0.7368}\, X_{it}$$

The figures in parenthesis are large sample standard errons, calailated by taking the square roots of the diagonal elements of Δ / n.

Swamy computes the following confidence intervals for α and β, using Δ

$$(3.2.16) \quad \Pr(-17.49 < \alpha < 196.60) \approx 0.95$$

$$\Pr(0.56 < \beta < 0.91 \approx 0.95$$

Beraning these values in mind, we test the hypothesis:

$$(3.2.17) \quad H_o = \triangle = 0 \quad \text{given that} \quad E\begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

Using a X^2 statistic (see swamy - (1971-p-124) we obtain the value of 104-15, which is weell above the 5% value of X^2 with 3 depres freedom. This H_0 is ryected. This implies that the variances of α_i and β_i are positive and that this combined with test (3.2.13) implies that the fixed coefficient approaches (first and second) are inappropriate since they assume from the start that $\triangle = 0$.

However (3.2.17) depends crucially on the condition that

$$E\begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

If this is not true then data on all these countries cannot be pooled and model II is also inappropriate. Swamy argues that specification errons due to the falsity of the assumption of identically distributed coefficient vectors is unavoidable in panel data analysis. This is indeed an issue that deserves further investigation.

REFERENCES

BOWDEN, R,J. (1968) "Some Formulations of Stichastic Regression coefficients in Economics" Discussion paper # 24 in Economics and Econometrics, University of Manchester.

BURNS, M.E(1974) " The Investigation of Aggregate Economic Ralations; Friedman and Meiselman Revisited " Manchester School (Forthcoming) (Xerox copy in INIS ROOM)

FEIGE,E. (1964) "The Demand for Liquid Assets: A Temporal Cross-section Analysis "Prentice-hall Englewood Cliffs, New Jersey.

FEIGE, E.L. and P.A.V.B. SWAMY (1972) "A Random Coefficient Model of the Demand for Liquid Assets" Workshop paper ≠ 7211 - Univ. of Wisconsin J. Money, Credit and Banking (Forthcoming)

FISK, P.R. (1967) "Models of the second Order in Regression Analysis "Journal of the Royal Statistical Soc. - B - Vol. 28 p.p. 266-281.

FRIEDMAN AND MEISELMAN (1963) "The Relative Stability of Monetary Velocity and the Investiment Multiplier in the U.S." in "Stabilization Policies" by E.C. Brown etal. - Englewood Cliffs 1963 - p.p. 165-268.

GRUNFELD, Y. (1968) "The Determinants of corporate Investiment" Unpublished Ph.D. Thesis - University of Chicago.

HILDRETH, C. and J. P. HOUCK (1968) "Some Estimatons for a Linear Model with Random Coefficients
J.A.S.A. Vol. 63, p.p. 584-595

KLEIN, L. R. (1953) "A Textbook of Econometrics" (p.p. 211-225) Row Peterson - Evanston, Illinois.

LEE, T.H. (1966) "Substitutabilily of Non-Bank Intermediary Liabilities for Money: the Empirical Evidence" J. Finance, 21, - 441 - 457.

RAO, C. R. (1965) "Linear Statiscal Inference and its Applications" - Wiley - New York.

RAO, C. R. (1965) "The Theory of Least Squares when the Parameters are Stochastic and its applications to the Analysis of Gronth Curves" Biometrika, 52.

ROSEMBERG (1970) "Varying Parameter Regression in the Analysis of a Cross-Section of Time-Series-I" Mimeo, Univ. of California, Berkeley.

SWAMY, P. A. V. B. (1971) "Statistical Inference in Random Coefficient Regression Models Springer-Verlag-New York.

SWAMY, P.A.V.B. (1974) "Linear Models with Random Coefficients" in "Frontcers in Econometrias" P - Zarembka (Ed.) Academic Press Inc., New York.

THEIL, H - (1971) "Principles of Econometrics
Wiley New York (Section 12-4)

ZELLNER,A- (1962) "An Efficient Method of
Estimating Seemingly Unrelated Regressions"
J.A.S.A. Vol. 57, PP.348.368.

ZELLNER,A (1966) " On the Aggregation Problem::
A Nen Approach to a Trouble some Problem" in "Economic Models,
Estimation and Risk Programming: Essays in Honor of Gerhard Iintner",
K.A.Fox etal (eds.)
Springer -Verlag,New York, 1969.