A Note on the Estimation of a System of General Functional Forms with Random Coefficients

By G.V.S.N. Murty and Sukumari Murty*

<Contents>

I. Introduction

In the recent econometric literature, the estimation of random coefficient regression models has received wider attention. Among others, Rubin [8], Hildreth and Houch [4], Singh *et al.* [10], Swamy [12], Hsiao [5], and Singh and Amanullah [9], have presented a good rationalization of linear models with random parameter specification. The development of these models mainly stems from the illuminating discussion by Klein[6], showing that the coefficients of a regression model can be treated as random in cross-section analysis, to account for spatial and inter-individual hetrogeniety. Once the random parameter specification is accepted, the issue that might be faced is with the choice of functional form. We are however, less fortunate, to have a direct link between economic theory and the choice of functional form, as Dhrymes *et al.* [3] assert "Economic theory gives preciously few clues as to the functional forms appropriate to the specification of economic relationships, and the presence of random error terms in stochastically specified equations adds an additional element of functional ambiguity." It is then, a common practice to choose between the alternatives:

* The authors are Economists, Sardar Patel Institute of Economic and Social Research, Ahmedabad, India. They would like to thank Professors A. L. Nagar, P. K. Trivedi and R. Radhakrishna for their valuable comments and suggestions on the earlier draft of the paper, presented to the seventeenth Indian Econometric Conference.

linear, log-linear, semi-log etc. Taking into consideration, the twin problems—choice of functional form and the random parameter specification—Murty [7] has analysed the possibility of estimating the general functional form (GFF), which is non-linear in nature, with random parameter specification. This functional form—introduced by Zarembka [14](1)—is general in nature where the linear, log-linear and semi-log etc. are the special cases. The utility of this form can be enhanced in many econometric studies. The glaring examples include the studies on: (1) demand functions for food and money (Zarembka [14, 15]) (2) import demand functions (Khan and Ross) and (3) financial analysis (Chang and Lee [2]). All these attempts, however, relate to a single equation problem. Recently, Tintner and Kadekodi [13] and Spitzer [11](2) have gone a step further, and recognized the importance of the GFF formulation in a simultaneous equation framework. A natural extension, one can consider is to view GFF in the Zellner's [16] seemingly unrelated regression (SUR) framework. We consider in this paper the estimation of such a system, with random parameter specification. In the next section, we briefly outline the estimation procedure of such a system, together with a statistical test criterion for discriminating models with random coefficients in the SUR framework.

II. Model

The general functional form can be written as

$$\frac{\overset{\lambda_1}{Y}(t)-1}{\lambda_1}=\beta_1+\sum_{\delta=2}^{\Lambda}\beta_\delta\left(\frac{\overset{\lambda\delta}{X_\delta}(t)-1}{\lambda\delta}\right)+u(t) \qquad t=1, 2, \cdots, T \tag{2.1}$$

where $Y(t)$ is the t^{th} observation on the dependent variable; $X_\delta(t)$ is the t^{th} observation on the independent variable X_δ; $u(t)$ is the disturbance term, corresponding to t^{th} observation; $\beta_1, \beta_2, \ldots, \beta_\Lambda$ and $\lambda_1, \lambda_2, \ldots, \lambda_\Lambda$ are the parameters.

It can be seen that the parameters $\lambda_1, \lambda_2, \ldots\ldots, \lambda_\Lambda$ determine the way in

(1) The genesis of this functional form is based on Box-Cox [1] transformation.
(2) In a simultaneous equation framework, Box-Cox's [1] parametric transformation may lead to identification problem and this can be resolved by assuming the same transformation to all the equations in the system.

which the data enter the equation. Note that the transformation is continuous at $\lambda_1, \lambda_2, \ldots, \lambda_\Lambda = 0$ where

$$\lim_{\lambda_1 \to 0} \frac{Y^{\lambda_1}-1}{\lambda_1} = \log Y \tag{2.2}$$

$$\lim_{\lambda_i \to 0} \frac{X_i^{\lambda_i}-1}{\lambda_i} = \log X_i \qquad i=2, \cdots, \Lambda. \tag{2.3}$$

Thus (2.1) reduces to linear, log-linear, semi-log, for specific values of $\lambda_1, \lambda_2, \ldots\ldots, \lambda_\Lambda$. Let us consider a system of M seemingly unrelated General Functional Forms, of which the i^{th} $(i=1, 2, \ldots, M)$ equation is

$$\left(\frac{Y_i(t)^{\lambda_{i1}}-1}{\lambda_{i1}}\right) = \beta_{i1} + \sum_{\delta=2}^{\Lambda} \beta_{i\delta} \left(\frac{X_{i\delta}(t)^{\lambda_{i\delta}}-1}{\lambda_{i\delta}}\right) + u_i(t) \tag{2.4}$$

where $Y_i(t)$, $X_{i\delta}(t)$ are the observations on the dependent and independent variables respectively for the i^{th} equation; $\beta_{i\delta} (\delta=1, 2, \ldots, \Lambda)$ and $\lambda_{i\delta}$ $(\delta=1, 2, \ldots, \Lambda)$ are the parameters; and $u_i(t)$ is the usual disturbance term for the i^{th} equation.

Writing for convenience

$$\frac{Y_i(t)^{\lambda_{i1}}-1}{\lambda_{i1}} = Y_i^{(\lambda_{i1})}(t)$$

and

$$\frac{X_{i\delta}(t)^{\lambda_{i\delta}}-1}{\lambda_{i\delta}} = X_{i\delta}^{(\lambda_{i\delta})}(t),$$

we have (2.4) as

$$Y_i^{(\lambda_{i1})}(t) = \beta_{i1} + \sum_{\delta=2}^{\Lambda} \beta_{i\delta} X_{i\delta}^{(\lambda_{i\delta})}(t) + u_i(t). \tag{2.5}$$

If the parameters $\beta_{i1}, \beta_{i2} \ldots \beta_{i\Lambda}$ are random, they take different values for each observation and (2.5) can be written as(3)

$$Y_i^{(\lambda)}(t) = \beta_{i1}(t) + \sum_{\delta=2}^{\Lambda} \beta_{i\delta}(t) X_{i\delta}^{(\lambda)}(t) + u_i(t) \tag{2.6}$$

writing

(3) For convenience of presentation it is assumed that $\lambda_{i1} = \lambda_{i\delta}$ $(\delta=2, \ldots, \lambda)$ and it is necessary to assume further that $\lambda_{i\delta} = \lambda_{j\delta}$ $(i \neq j)$.

$$\beta_{i\delta}(t)=\bar{\beta}_{i\delta}+\epsilon_{i\delta}(t) \tag{2.7}$$

where $\bar{\beta}_{i\delta}$ is the mean regression coefficient and $\epsilon_{i\delta}(t)$ is the unobserved random disturbance term. Using (2.7) in (2.6) we have

$$\overset{(\lambda)}{Y_i}(t)=\bar{\beta}_{i1}+\sum_{\delta=2}^{\Lambda}\bar{\beta}_{i\delta}\overset{(\lambda)}{X_{i\delta}}(t)+w_i(t) \tag{2.8}$$

where

$$w_i(t)=\sum_{\delta=2}^{\Lambda}\epsilon_{i\delta}(t)\overset{(\lambda)}{X_{i\delta}}(t)+v_i(t)$$

and

$$v_i(t)=u_i(t)+\epsilon_{i1}(t).$$

Making the following assumptions

(i) $E(u_i(t))=E(v_i(t))=E(\epsilon_{i\delta}(t))=E(w_i(t))=0$ for all i's, δ's and t's

(ii) $E(u_i(t),\ \epsilon_{i\delta}(t))=0$

(iii) $E(u_i(t),\ u_j(t'))=\sigma_{ij}$ if $t=t'$
$\qquad =0$ if $t\neq t'$

(iv) $E\ (\epsilon_{i\delta}(t),\ \epsilon_{j\delta}'(t'))=\theta_{ij\delta}$ if $t=t',\ \delta=\delta'$
$\qquad =0$ if $t\neq t',\ \delta\neq\delta'$

(v) $E\ (v_i(t),\ v_j(t'))=\theta^*_{ij1}=\sigma_{ij}+\theta_{ij_1}$ if $t=t'$
$\qquad =0$ if $t\neq t'$

(vi) $E\ (w_i(t),\ w_j(t'))=w_{ij}(t)$ if $t=t'$
$\qquad =0$ if $t\neq t'$

where

$$w_{ij}(t)=\theta^*_{ij_1}+\sum_{\delta=2}^{\Lambda}\overset{(\lambda)}{X_{i\delta}}(t)\ \overset{(\lambda)}{X_{j\delta}}(t)\ \theta_{ij\delta}. \tag{2.9}$$

The system of equations in (2.8) can be written as

$$\begin{pmatrix} Y_1^* \\ Y_2^* \\ \vdots \\ Y_M^* \end{pmatrix}=\begin{pmatrix} X_1^* & 0 & \cdots & 0 \\ 0 & X_2^* & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & X_M^* \end{pmatrix}\begin{pmatrix} \bar{\beta}_1 \\ \bar{\beta}_2 \\ \vdots \\ \bar{\beta}_M \end{pmatrix}+\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \end{pmatrix} \tag{2.10}$$

where (i) Y_i^* $(i=1,2,\cdots,M)$ is a vector of order $(T\times 1)$ with the following elements

$$Y_i^*=\begin{pmatrix} \overset{(\lambda)}{Y_i}(1) \\ \vdots \\ \overset{(\lambda)}{Y^i}(T) \end{pmatrix}$$

(ii) X_i^* $(i=1,2,\ldots,M)$ is a matrix of order $(T\times \Lambda)$ with the following elements

$$X_i^*=\begin{pmatrix} 1 & \overset{(\lambda)}{X}_{i1}(1)\cdots\cdots\overset{(\lambda)}{X}_{i\Lambda}(1) \\ 1 & \overset{(\lambda)}{X}_{i1}(2)\cdots\cdots\overset{(\lambda)}{X}_{i\Lambda}(2) \\ \vdots & \vdots \qquad\qquad \vdots \\ 1 & \overset{(\lambda)}{X}_{i1}(T)\cdots \overset{(\lambda)}{X}_{i\Lambda}(T) \end{pmatrix}$$

(iii) β_i' is a vector of parameters $(\beta_{i1}, \beta_{i2}, ..., \beta_{i\Lambda})$ of order $(1\times\Lambda)$

(iv) w_i' is a vector of disturbance elements $(w_i(1), w_i(2), \ldots, w_i(T))$ of order $(1\times\Lambda)$.

More compactly (2.10) can be written as

$$Y^*=X^*\beta+w \tag{2.11}$$

where X^* is a $(MT\times M\Lambda)$ block diagonal matrix and Y^*, β and w are vectors of order $(MT\times 1)$, $(M\Lambda\times 1)$ and $(MT\times 1)$ respectively.

It can be seen from the assumption (vi) that

$$E(w_iw_j')=\begin{pmatrix} w_{ij}(1) & 0\cdots\cdots\cdots 0 \\ 0 & w_{ij}(2)\cdots\cdots 0 \\ \vdots & \qquad\qquad \vdots \\ 0 & \cdots\cdots\cdots w_{ij}(T) \end{pmatrix}=\phi_{ij}. \tag{2.12}$$

Thus we have

$$E(ww')=\begin{pmatrix} \phi_{11} & \phi_{12}\cdots\cdots\phi_{1M} \\ \phi_{21} & \phi_{22}\cdots\cdots\phi_{2M} \\ \vdots & \vdots \qquad \vdots \\ \phi_{M1} & \phi_{M2}, \cdots\phi_{MM} \end{pmatrix}=\Phi. \tag{2.13}$$

In the case where $\lambda=1$, it can be seen that (2.10) reduces to the model considered by Singh and Amanullah [9] and further if it is assumed that θ^*_{ij1}, $\theta_{ij2}, ..., \theta_{ij\Lambda}=0$ for $i, j=1, 2, \ldots, M$, then (2.9) reduces to the Zellner's [16] seemingly unrelated regression equations.

III. Estimation

If Φ is known (2.10) can be estimated for a given (λ) by using the Generalized Least-Squares (GLS) procedure in the following way.

Making the assumption of normality(4) to the disturbance term, in each

(4) We note that under the normality assumption the error term in (2.9) extends from $(-\infty, +\infty)$ and consequently the dependent variable should also extend from $-\infty$ to ∞. For some values of λ (say $\lambda=1/2$), this range may not be possible, in which case the error term can only be assumed as approximately normal.

equation we can form the likelihood function of the sample values as

$$L(\theta/\text{data})=\frac{1}{(2\pi)^{\frac{MT}{2}}}|\Phi|^{-\frac{1}{2}}e^{-\frac{1}{2}(Y^*-X^*\beta)'\Phi^{-1}(Y^*-X^*\beta)}|J| \tag{2.14}$$

where $\theta=(\beta',\lambda)$ and J is the Jacobian of transformation, such that

$$|J|=\prod_{i=1}^{M}\prod_{t=1}^{T}\left(\frac{\partial \overset{(\lambda)}{Y}_i(t)}{\partial Y_i(t)}\right).$$

For a given λ, we have from (2.3), the log-likelihood $(L^*)_\lambda$[5] value as

$$(L^*)_\lambda=-\frac{MT}{2}\log 2\pi-\frac{1}{2}\Big(\log|\Phi|\Big)_\lambda+(\lambda-1)\sum_{i=1}^{M}\sum_{t=1}^{T}\log Y_i(t) \tag{2.15}$$
$$-\frac{1}{2}\Big\{(Y^*-X^*\beta)'\Phi^{-1}(Y^*-X^*\beta)\Big\}_\lambda$$

By taking a grid of values for λ, $(L^*)_\lambda$ can be estimated from (2.15) with the information on $(\Phi)_\lambda$. From these series of λ and $(L^*)_\lambda$ values, the optimum λ (say λ^*)[6], can be chosen either by using numerical approximation procedure or by least squares procedure. Using this λ^* and $(\Phi)_{\lambda^*}$, other parameters of the system can be obtained by applying the GLS procedure to (2.11). As is evident, that this optimum λ corresponds to the maximum of the log-likelihood function of the sample values, the estimates are consistent.

In general, information on Φ is not available and thus the estimation of the parameters of (2.10) is not straightforward. Nevertheless, we can obtain a consistent estimator of Φ by using the multiequation generalization of the Hildreth-Houch [4] procedure, along the lines suggested by Singh and Amanullah [9].

It can be seen that the estimation of Φ in (2.13) essentially leads, from (2.9) and (2.12) to the estimation of $\theta^*_{ij1}, \theta_{ij2}, \ldots, \theta_{ijK}$ $(i, j=1, 2, \ldots, M)$. For a given λ, applying Ordinary Least Squares (OLS) procedure to each equation in (2.10) and writing for i^{th} equation

$$(\hat{w}_i)_\lambda=(M_i)_\lambda\,(w_i)_\lambda \tag{2.16}$$

where $(M_i)_\lambda$ is a $(T\times T)$ idempotent matrix such that

(5) The suffix λ indicates the estimate in the brackets for a specific value of λ.
(6) The optimum criterion is based on choosing λ at the maximum of the log-likelihood value L^*.

$$(M_i)_\lambda=(I-X_i^*(X_i^{*\prime}\ X_i^*)^{-1}\ X_i^{*\prime})_\lambda$$
$$=\begin{pmatrix} m_{ij}(1)\cdots\cdots m_{iT}(1) \\ \vdots \qquad\qquad \vdots \\ m_{i1}(T)\cdots\ m_{iT}(T) \end{pmatrix}_\lambda$$

writing

$(M_{ij})_\lambda$: matrix of the product of the corresponding elements of M_i and M_j for a given λ

$(X_{ij}^*)_\lambda$: matrix of the product of the corresponding elements of X_i and X_j for a given λ

$(\hat{w}_{ij})_\lambda$: vector of the product of the corresponding elements of w_i and w_j for a given λ.

$(\theta_{ij}^*)_\lambda$: vector of order $(\Lambda\times 1)$, with elements $(\theta_{ij1}^*,\ \theta_{ij2}^*,\cdots,\ \theta_{ij\Lambda}^*)$

We can have from (2.12)

$$E\ (\hat{w}_i\hat{w}_j')_\lambda=(M_i\ E(w_iw_j')M_j)_\lambda=(M_i\ \Phi_{ij}\ M_j)_\lambda \tag{2.17}$$

and

$$E\ (\hat{w}_{ij})_\lambda=(M_{ij}\ X_{ij}^*\ \theta_{ij}^*)_\lambda$$

writing

$$(\nu_{ij})_\lambda=(\hat{w}_{ij}\ -E(\hat{w}_{ij}))_\lambda$$

we have

$$(\hat{w}_{ij})_\lambda=(M_{ij}\ X_{ij}^*\ \theta_{ij}^*)_\lambda+(\nu_{ij})_\lambda=(Z_{ij}^*\ \theta_{ij}^*)_\lambda+(\nu_{ij})_\lambda \tag{2.18}$$

where

$$(Z^*_{ij})_\lambda=(M_{ij}X_{ij}^*)_\lambda. \tag{2.19}$$

Applying OLS[7] procedure to (2.18) we have the estimate of θ^*_{ij}[8], for a given λ as

$$\hat{\theta}_{ij}^*=((Z^*_{ij}{}'\ Z^*_{ij})^{-1}\ Z^*_{ij}{}'\ \hat{w}_{ij})_\lambda. \tag{2.20}$$

Using (2.20) in (2.9), (2.12) and (2.13) the GLS estimator for $\bar{\beta}$ [9] for

(7) In view of the dependence of (θ^*_{ij}) with the sample observations, it can be seen that $(\hat{\theta}^*_{ij})_\lambda$ is consistent but inefficient estimator of $(\theta^*_{ij})_\lambda$. This is because of the error term $(\nu_{ij})_\lambda$ in (2.18) being violating the homoscadastacity property of GLS. In this case it can be easily seen that $E(\nu_{ij}\nu_{ij})_\lambda=2(\Sigma_{ij})_\lambda$, (which is a matrix of squared elements of $(M_i\phi_{ij}M_j)_\lambda$ and accordingly $(\hat{\theta}^*_{ij})_\lambda$ can be obtained by applying GLS procedure to (2.18).

(8) It may sometimes be possible to arrive at the negative estimates of the diagonal elements of $(\phi_{ii},\ i=1,2,\ldots,M)$ in which case the estimation procedure can be modified suitably, along to lines suggested by Hildreth and Houch [4].

a given λ, can be obtained as

$$(\hat{\hat{\beta}})_\lambda = ((X^{*\prime}\hat{\Phi}^{-1}X^*)^{-1}\, X^{*\prime}\hat{\Phi}^{-1}Y^*)_\lambda \tag{2.21}$$

Since $(\Phi)_\lambda$ is known for a given value of λ, using (2.15), $(L^*)_\lambda$ can be formulated. Thus by considering a series of values of λ, we can have the corresponding series of the estimates of $(\Phi)_\lambda$, and a series of $(L^*)_\lambda$. The optimum value of λ can be located, from these series, either by using numerical approximation procedures or by least squares procedure.

The estimation procedure outlined above can be used for testing the validity of specific random coefficient functional form in Zellner's seemingly unrelated regression (SUR) equation framework. For example, the log-linear formulation is postulated in SUR framework with randomness in parameters. The validity of this postulation can be tested statistically by formulating the hypothesis as

$H_0 : \lambda = 0$

$H_1 : \lambda = \lambda_0$ (arrived from the GFF formulation).

Since λ discriminates the type of functional form, we can formulate the test criterion by considering the conditional and the unconditional log-likelihood values from (2.15).

In the case of different transformations for dependent and independent variables in (2.6), the estimation and testing procedures outlined above need a little modification. Since we have Λ, λ values (i.e., $\lambda_1, \lambda_2, \cdots\cdots, \lambda_\Lambda$), we obtain first, the (L^*) values by varying λ_1 (in the chosen range), and fixing $\lambda_2, \lambda_3, \cdots, \lambda_\Lambda$ and then $(L^*)_{\lambda_1}$— the maximum of the conditional likelihood function—and $\bar{\lambda}_1$, the corresponding λ_1. This procedure can be repeated by varying λ_2 and fixing $\lambda_1, \lambda_3, \cdots, \lambda_\Lambda$ and so on. Now we have a new series of $\bar{\lambda}_1, \bar{\lambda}_2, \cdots, \bar{\lambda}_\Lambda$ and $(L^*)_{\lambda_1}, (L^*)_{\lambda_2}, \ldots, (L^*)_{\lambda_\Lambda}$ values, and the final round optimum λ'^s can be located. The validity of a particular random coefficient functional form in the SUR framework can be tested accordingly.

---

(9) By making the assumption that $T \xrightarrow{U-} \infty$ $(X^{**}_{ij}\ X^*_{ij})/T$ converges to a positive definite matrix, the following properties of the estimators can be established. For details see Singh and Amanullah [9].

$$(\hat{\theta}_{ij}{}^* - \theta^*_{ij}) = O_p(T^{-\frac{1}{2}}) \qquad (\hat{\hat{\beta}} - \beta) = O_p(T^{-1})$$

where $(T^{-\frac{1}{2}})$ represents the term of which of order $T^{-\frac{1}{2}}$ in probability.

References

[1] Box, C.E.P. and D.R. Cox, "An Analysis of Transformations," *Journal of the Royal Statistical Society*, Series B, 26 (1964), 211-243.

[2] Chang Hui-Shyong, and Cheng F. Lee, "Using Pooled Time Series and Cross-section Data to Test the Firm and Time Effects in Financial Analysis," *Journal of Financial and Quantitative Analysis*, Vol. XII No. 3 (1977), 457-471.

[3] Dhrymes, P. J., E. P. Howrey, S. H. Hymans, J. Kmenta, E. E. Leamer, R. E. Quandt, J. B. Ramsey, H. T. Shapiro, and V. Zarnowitz, "Criterion for Evaluation of Econometric Models," *Ann. Econ. Soc. Measurement*, 1(1972), 261-324.

[4] Hildreth, C. and J. P. Houch, "Some Estimators for Linear Model with Random Coefficients," *Journal of the American Statistical Association*, 63(1968), 584-95.

[5] Hsiao, Cheng, "Statistical Inference for a Model with Both Random Cross-section and Time Series Effects," *International Economic Review*, XV(1974), 12-30.

[6] Klein, L. R., *A Textbook of Econometrics*, 1953, Row Peterson, Evanston, Illinois.

[7] Murty, G. V. S. N., "On the Estimation of General Functional Form with Random Coefficients," *Sankhya*, Series C, 38(1976), 37-43.

[8] Rubin, H., "Note on Random Coefficients," in T. C. Koopmans, ed., *Statistical Inference in Dynamic Economic Models, Cowles Commission Monograph 10*, 1950.

[9] Singh, B. and Amanullah, "Estimation of Seemingly Unrelated Regressions with Random Coefficients," *Journal of the American Statistical Association*, 69 (1974), 191-195.

[10] Singh, B., A. L. Nagar, N. K. Choudhari and B. Raj, "On the Estimation of Strucutral Change: A Generalization of the Random Coefficient Regression Model," *International Economic Review*, 17 (1976), 340-361.

[11] Spitzer, J. J., "A Simultaneous Equations System of Money Demand and Supply Using Generalized Functional Forms," *Journal of Econometrics*, 5 (1977), 117-128.

[12] Swamy, P. A. V. B., "Efficient Inference in a Random Coefficient Regression Model," *Econometrica*, 38 (1970), 311-323.

[13] Tintner, G. and G. Kadekodi, "Note on the Transformation of Variables in Simultaneous Equations Systems," *Journal of Indian Social and Agricultural Statistics*, 23 (1971), 163-173.

[14] Zarembka, P., "Functional Form in the Demand for Money," *Journal of the American Statistical Association*, 63 (1968), 502-511.

[15] Zarembka, P., *Toward a Theory of Economic Development*, Holden-Day, San Francisco, California, 1972.

[16] Zellner, A., "An Efficient Estimation of Seemingly Unrelated Regressions and Tests for Aggregation Bias," *Journal of the American Statistical Association*, 62 (1962), 368-384.