Ann. Inst. Statist. Math.
36 (1984), Part A, 395-401

ESTIMATION OF A RANDOM COEFFICIENT MODEL UNDER LINEAR STOCHASTIC CONSTRAINTS

V. K. SRIVASTAVA, BALDEV RAJ AND KULDEEP KUMAR

(Received Jan. 6, 1983; revised Mar. 30, 1984)

Summary

A random coefficient model in which means of random coefficients are subject to a set of linear stochastic constraints is considered and estimators for the means of coefficients are proposed. Their asymptotic properties are presented and some remarks on efficiency are placed.

1. Introduction

This paper examines the problem of estimation in a random coefficient model when means of random coefficients are subject to a set of linear stochastic constraints. The proposed estimators may be useful, for instance, in estimating the heritability coefficient encountered in animal biology. The regression model commonly used for the purpose of estimating the heritability coefficient assumes that the response coefficient is fixed. Given the complex nature of biological mechanism, the assumption of a fixed response coefficient appears both restrictive and unreasonable. A more reasonable assumption would be to treat the regression coefficient as randomly varying instead of fixed constant. Also, since the heritability coefficient is expected to have a range of zero to one half of one, an appropriate procedure in estimating this coefficient would be to impose a linear stochastic restriction on the mean of randomly varying response coefficient, (see Nigam, Srivastava, Jain and Gopalan [2]). This framework may also have other applications in medical, agriculture and social sciences.

The paper is organized as follows. In the following section the random coefficient model under linear stochastic constraints is formulated. Section 3 presents its estimators and properties. The final section contains a brief summary and suggestion for further research.

2. The random coefficient model under linear stochastic constraints

Following Hildreth and Houck [1], we may write the random coefficient model as

$$y_i=\beta_{i1}+\sum_{k=2}^{K} x_{ik}\beta_{ik} \tag{2.1}$$

such that

$$\beta_{ik}=\beta_k+\omega_{ik} \qquad (i=1, 2,\cdots, N;\ k=1, 2,\cdots, K) \tag{2.2}$$

where y_i is the ith observation on the variable to be explained, x_{ik} is the ith observation on the kth explanatory variables and β_{ik} is the random regression coefficient associated with it. The usual disturbance term is assumed to be subsumed in the random intercept term β_{i1}. We shall assume that x's are linearly independent fixed numbers. Further, we shall assume that the random coefficient β_{ik}'s are independently and identically distributed with means β_k's and variance σ_k^2's such that:

$$\begin{aligned} &\mathrm{E}(\omega_{ik})=0 \qquad \text{for all } i \text{ and } k \\ &\mathrm{E}(\omega_{ik}\,\omega_{jk'})=\begin{cases} \sigma_k^2 & \text{for } i=j \text{ and } k=k' \\ 0 & \text{for } i\neq j \text{ and/or } k\neq k'. \end{cases} \end{aligned} \tag{2.3}$$

The random coefficient model under these specifications can be compactly written as:

$$y=X\beta+u \tag{2.4}$$

where y is an $N\times 1$ vector of observations on the variable to be explained, X is an $N\times K$ matrix of observations on K explanatory variables including the intercept term, β is a $K\times 1$ vector of fixed means for the randomly varying coefficients and u is an $N\times 1$ vector of disturbances.

It is easily verified that:

$$\begin{aligned} &\mathrm{E}(u)=0 \\ &\mathrm{E}(uu')=\Theta=\mathrm{diag.}(\theta_{11}, \theta_{22},\cdots, \theta_{NN}) \end{aligned} \tag{2.5}$$

where

$$\theta_{ii}=\sigma_1^2+\sum_{k=2}^{K} x_{ik}^2\sigma_k^2 .$$

The problem of estimation in the random coefficient model (2.4) when σ_k^2's are unknown is well known and is documented, among other places, in the econometrics book by Raj and Ullah [6]. However, the problem to be analysed in this paper is different since the mean regression vector β is subject to a set of linear stochastic constraints specified by

$$r=R\beta+v \tag{2.6}$$

where r is a $G\times 1$ vector of known elements, R is a $G\times K$ matrix of full row rank whose elements are known and v is a $G\times 1$ vector of disturbance terms reflecting the uncertainty about the prior restrictions. The disturbance vector v is assumed to have mean vector 0 and variance-covariance matrix Ψ. We shall further assume that disturbances in (2.4) and (2.6) are uncorrelated. Specifically, we have assumed that:

$$\mathrm{E}(v)=0\,, \qquad \mathrm{E}(vv')=\Psi\,, \qquad \mathrm{E}(vu')=0\,. \tag{2.7}$$

Once again, the problem of estimation in a fixed coefficient model under linear stochastic constraints is well known and is documented in Theil and Goldberger [10] among other places. See Srivastava and Singh [9] and Srivastava [7], respectively, for a useful extension and annotated bibliography on the mixed estimation procedure. The object of this paper is to extend Theil and Goldberger's mixed estimation method to a situation where regression coefficients are randomly varying and the variances of random coefficients are not known.

3. Alternative estimation procedures and their properties

The random coefficient model (2.4) under the linear stochastic constraint (2.6) can be written as:

$$\begin{bmatrix} y \\ r \end{bmatrix}=\begin{bmatrix} X \\ R \end{bmatrix}\beta+\begin{bmatrix} u \\ v \end{bmatrix}\,. \tag{3.1}$$

The mixed estimator (ME) of β in (3.1), which accounts for the linear stochastic constraint (2.6) but ignores the randomness of regression coefficients in (2.4) is given as:

$$\hat{\beta}=(X'X+R'\Psi^{-1}R)^{-1}(X'y+R'\Psi^{-1}r)\,. \tag{3.2}$$

The estimator $\hat{\beta}$ is unbiased and its variance-covariance matrix is given by:

$$\mathrm{E}(\hat{\beta}-\beta)(\hat{\beta}-\beta)'=(X'X+R'\Psi^{-1}R)^{-1} \times(X'\Theta X+R'\Psi R)(X'X+R'\Psi^{-1}R)^{-1}\,. \tag{3.3}$$

The random mixed estimator (RME) of β in (3.1), which accounts for randomness of coefficients and linear stochastic constraints, is given as:

$$\tilde{\beta}=(X'\Theta^{-1}X+R'\Psi^{-1}R)^{-1}(X'\Theta^{-1}y+R'\Psi^{-1}r) \tag{3.4}$$

which is unbiased and its variance-covariance matrix is given below,

$$\mathrm{E}\,(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'=(X'\Theta^{-1}X+R'\Psi^{-1}R)^{-1}\,. \tag{3.5}$$

From (3.3) and (3.5), it is easy to verify that $\tilde{\beta}$ is more efficient than $\hat{\beta}$. Furthermore, we may note that the RME estimator (3.4) is non-operational because the elements of diagonal matrix Θ involve unknown parameters, σ_k^2's $(k=1,2,\cdots,K)$. An operational estimator of β, which has the same distribution as the RME estimator $\tilde{\beta}$ may be obtained by replacing Θ in (3.4) by its consistent estimator. Below we shall derive a number of consistent estimators of Θ, and corresponding to each estimator of Θ an operational RME estimator of β can be defined. For this purpose, we observe that the ME estimator of the disturbance vector u in model (3.1) is given as:

$$\tilde{u}=My-X(X'X+R'\Psi^{-1}R)^{-1}R'\Psi^{-1}r$$

where

$$M=I_N-X(X'X+R'\Psi^{-1}R)^{-1}X'\,.$$

It is easily verified that

$$\mathrm{E}\,(\tilde{u})=0\,,\qquad \mathrm{E}\,(\tilde{u}\tilde{u}')=M\Theta M+D \tag{3.6}$$

where $D=X(X'X+R'\Psi^{-1}R)^{-1}R'\Psi^{-1}R(X'X+R'\Psi^{-1}R)^{-1}X'$. Now collecting the diagonal terms from the variance-covariance matrix of $\tilde{u}$ as a column vector, we write,

$$\mathrm{E}\,(\tilde{\tilde{u}})=\dot{M}\dot{X}\sigma+d \tag{3.7}$$

where d is a column vector formed by the diagonal elements of D, $\tilde{\tilde{u}}=\tilde{u}*\tilde{u}$, $\dot{M}=M*M$, $\dot{X}=X*X$, (* denotes the Hadamard matrix product) and $\sigma'=(\sigma_1^2,\sigma_2^2,\cdots,\sigma_K^2)$.

Thus, we can express

$$g=G\sigma+\varepsilon \tag{3.8}$$

where $g=(\tilde{\tilde{u}}-d)$, $G=\dot{M}\dot{X}$ and ε is an $N\times 1$ disturbance vector with mean vector 0 and variance-covariance matrix Ω.

The least squares (LS) estimator of σ in (3.7) is given as:

$$\hat{\sigma}=(G'G)^{-1}G'g\,. \tag{3.9}$$

On the lines of Hildreth and Houck [1] it is easily shown that the estimator (3.9) is consistent. Now we can utilize the elements of σ to construct a consistent estimator $\hat{\Theta}$ of Θ from (2.5). Thus, an operational random mixed estimator (ORME) corresponding to the consistent estimator $\hat{\Theta}$ may be defined as

$$\tilde{\beta}_1=(X'\hat{\Theta}^{-1}X+R'\Psi^{-1}R)^{-1}(X'\hat{\Theta}^{-1}y+R'\Psi^{-1}r) . \tag{3.10}$$

Another estimator of σ may be defined by setting at zero the ultradiagonal elements in the variance-covariance matrix Ω and applying the weighted least squares (WLS) method to (3.8); thus we get the following estimator:

$$\hat{\hat{\sigma}}=[G'(I_N^* \Omega)^{-1}G]^{-1}G'(I_N^* \Omega)^{-1}g . \tag{3.11}$$

Following Theil [11], it can be shown that the consistent estimator $\hat{\hat{\sigma}}$ is more efficient than the LS estimator $\hat{\sigma}$. Once again, the elements of (3.11) may be utilized to produce another consistent estimator $\hat{\hat{\Theta}}$ of Θ from (2.5). Thus another ORME estimator of β corresponding to $\hat{\hat{\Theta}}$ is given by

$$\tilde{\beta}_2=(X'\hat{\hat{\Theta}}^{-1}X+R'\Psi^{-1}R)^{-1}(X'\hat{\hat{\Theta}}^{-1}y+R'\Psi^{-1}r) . \tag{3.12}$$

Finally, we may apply the generalized least squares (GLS) procedure to (3.8) to obtain yet another consistent estimator $\tilde{\sigma}$ of σ, where

$$\tilde{\sigma}=(G'\Omega^{-1}G)^{-1}G'\Omega^{-1}g . \tag{3.13}$$

The estimator $\tilde{\sigma}$ can be shown to be more efficient than $\hat{\sigma}$ and $\hat{\hat{\sigma}}$; see Raj [3].

An ORME of β corresponding to the consistent estimator (3.13) may be defined as

$$\tilde{\beta}_3=(X'\tilde{\Theta}^{-1}X+R'\Psi^{-1}R)^{-1}(X'\tilde{\Theta}^{-1}y+R'\Psi^{-1}r) \tag{3.14}$$

where $\tilde{\Theta}$ is a consistent estimator of Θ obtained by utilizing the ORME estimates of σ_k^2's from (3.13). The asymptotic variances of the elements of $\tilde{\beta}_3$ can be obtained from

$$(X'\Theta^{-1}X+R'\Psi^{-1}R)^{-1} . \tag{3.15}$$

It may be noted that the LS estimator (3.9) is needed in the WLS and GLS estimators of σ to operationalize them since they involve the unknown Θ. We shall assume that it has been done.

Following Hildreth and Houck ([1], Sections 4 and 5) it is straightforward to show that all three estimators $\tilde{\beta}_1$, $\tilde{\beta}_2$ and $\tilde{\beta}_3$ of β are consistent. Furthermore, in the first order sense, all the three estimators are asymptotically equivalent to the best linear unbiased estimator $\tilde{\beta}$ of β, and each ORME is more efficient than the ME estimator $\hat{\beta}$ of β.

On the lines of Theil ([11], pp. 622–628) it is easily shown that the ultradiagonal elements of variance-covariance matrix Ω of the ME residual vector $\tilde{u}$ are of lower order of magnitude than its diagonal elements. Therefore, the WLS and GLS estimators $\hat{\sigma}$ and $\tilde{\sigma}$ are asymptotically equally efficient whereas the LS estimator $\hat{\sigma}$ is inefficient. It might therefore appear that it is preferable to choose either the WLS or the GLS estimators of Θ over the LS estimator $\hat{\Theta}$ in deriving the ORME of β. However, in a large sample situation, there would be no efficiency gain in utilizing either the GLS or WLS over LS estimator of Θ in selecting a suitable ORME of β.

This result should not be surprising because all three estimators of Θ are consistent. Similar results were obtained by Raj and Srivastava [5] in the context of random coefficient model without linear stochastic constraints. The small sample efficiency rankings of alternatives ORMEs could differ. In a simulation study on alternative random coefficient models, Raj [3] has shown that for the sample of size 50 the operational GLS estimator of β, which utilizes GLS estimator of Θ, is to be prefered over other operational GLS estimators of β, which utilize either the WLS or LS estimator of Θ. Thus, we should expect that, in small samples, the ORME $\tilde{\beta}_3$ would be more efficient than either $\tilde{\beta}_2$ or $\tilde{\beta}_1$ estimators in small samples.

It should also be noted that in defining each of the three ORMEs for β we have utilized unrestricted estimators of σ_k^2's, and these estimates could sometimes be negative. An appealing solution to the problem of negative estimates of σ_k^2 is either to use truncated estimates of σ_k^2's wherein negative estimates of σ_k^2's are replaced by zero or to use restricted estimates of σ_k^2's wherein estimators of σ_k^2's are obtained by optimizing an appropriate criterion function subject to the restrictions $\sigma_k^2 \geqq 0$ for $k=1, 2, \cdots, K$. However, a restricted (or truncated) operational estimator would be biased but more efficient than the corresponding unrestricted ORME, see Hildreth and Houck [1]. Furthermore, the use of a truncated or restricted estimator could complicate (if not invalidate) the usual testing procedure. See Srivastava, Mishra and Chaturvedi [8] for an alternative approach to ensuring almost nonnegativity in σ_k^2's.

4. Concluding remarks

We have proposed three operational RME (or operational GLS) estimators of β that are asymptotically equivalent to the best linear unbiased estimator for β. Based on the evidence in Raj [3], we can conclude that in small samples, estimator $\tilde{\beta}_3$ (which utilizes most efficient estimators of σ_k^2's), is most efficient estimator of β in the random coefficient model under linear stochastic constraints.

In a similar situation it has been shown by Raj, Srivastava and Upadhaya [6] that the variances of the operational GLS estimator of β obtained from the asymptotic moment matrix (3.15) would underestimate its finite sample approximate moment matrix. Thus, a finite sample approximation to the asymptotic moment matrix of limiting distribution of the ORME $\tilde{\beta}_3$ needs to be derived.

Acknowledgement

The authors wish to thank an anonymous referee for helpful suggestions for improving the exposition of this paper.

LUCKNOW UNIVERSITY, INDIA
WILFRID LAURIER UNIVERSITY, CANADA
UNIVERSITY OF KENT AT CANTERBURY, ENGLAND

REFERENCES

[1] Hildreth, C. and Houck, J. P. (1968). Some estimators for a linear model with random coefficients, *J. Amer. Stat. Ass.*, **63**, 584-595.

[2] Nigam, A. K., Srivastava, V. K., Jain, J. P. and Gopalan, R. (1979). A note on estimation of heritability by regression analysis, *Biometrical Journal*, **21**, 667-673.

[3] Raj, B. (1975). Linear regression with random coefficients, the finite sample and convergence properties, *J. Amer. Stat. Ass.*, **70**, 127-137.

[4] Raj, B., Srivastava, V. K. and Upadhyaya, S. (1980). The efficiency of estimating a random coefficient model, *Journal of Econometrics*, **12**, 285-299.

[5] Raj, B. and Srivastava, V. K. (1981). Asymptotic distributions of estimators in a random coefficient model, *Indian Journal of Physical and Natural Sciences*, **1**, 31-34.

[6] Raj, B. and Ullah, A. (1981). *Econometrics, A Varying Coefficient Approach*, (London: Croom Helm Ltd.).

[7] Srivastava, V. K. (1980). Estimation of linear single-equation and simultaneous—equation models under stochastic linear constraints, *Int. Statist. Rev.*, **48**, 79-82.

[8] Srivastava, V. K., Mishra, G. D. and Chaturvedi, A. (1981). Estimation of linear regression model with random coefficients ensuring almost non-negativity of variance estimators, *Biometrical Journal*, **23**, 3-8.

[9] Srivastava, V. K. and Singh, R. C. (1971). Utilization of prior information in the estimation of simultaneous linear stochastic equations, *Proceedings of the All-India Seminar on Demography and Statistics*, Banaras Hindu University, 95-109.

[10] Theil, H. and Gordberger, A. S. (1961). On pure and mixed statistical estimation in economics, *Int. Econ. Rev.*, **2**, 65-78.

[11] Theil, H. (1971). *Principles of Econometrics*, John Wiley and Sons, New York, 622-628.