No. 82 (80-20)

On the Simultaneous Estimation of Means and Variances of the Random Coefficient Model[1/]

by

Kazumi Asako

July 1980

Abstract

This paper develops a new estimation technique which obtains consistent and asymptotically efficient estimators of means and variances of the random coefficient model. In addition, a kind of "mixed estimation" technique is introduced in order to further increase efficiency. Some results from Monte Carlo experiments are also reported.

In this paper, I shall develop a new estimation technique which obtains consistent and asymptotically efficient estimators of means and variances of a random coefficient model:

$$(1) \qquad y_t = \sum_{i=1}^{K} \beta_{it} x_{it} + u_t, \qquad (t=1, 2, \ldots, T)$$

$$(2) \qquad \beta_{it} = \beta_i + v_{it},$$

where y_t = the dependent variable and x_{it} = nonstochastic explanatory variables; and u_t and v_{it} are mutually and serially independent random variables with means zero and variances σ_u^2 and $\sigma_{\beta_i}^2$, respectively.

Theil and Mennes [7] and Hildreth and Houck [3] have proposed an estimation technique which obtains consistent and asymptotically efficient estimators of (β_i), and $(\sigma_{\beta_i}^2 ; \sigma_u^2)$. With their approach, means (β_i) and variances $(\sigma_{\beta_i}^2 ; \sigma_u^2)$ are estimated in separate regression equations and, consequently, the entire estimation method comprises four sequences of regressions.[2/] The new approach I shall propose differs from the above in that means and variances are estimated simultaneously in the same regression equation. Consequently, it requires only three sequences of regressions to obtain the same properties of estimators (β_i) and $(\sigma_{\beta_i}^2 ; \sigma_u^2)$. Moreover, I shall propose a kind of "mixed estimation" technique of Theil and Goldberger [6] which increases efficiency of the estimators. Section I develops new estimation technique and Section 2 reports some results from Monte Carlo experiments. Concluding remarks will follow in the last section.

1. New Estimation Technique

Two equations (1) and (2) together can be written as

$$y_t = \sum_{i=1}^{K} \beta_i x_{it} + \varepsilon_t, \tag{3}$$

where

$$\varepsilon_t = \sum_{i=1}^{K} (\beta_{it} - \beta_i) x_{it} + u_t. \tag{4}$$

By assumption, $E(\varepsilon_t)=0$ and

$$E(\varepsilon_t \varepsilon_s) = \begin{cases} \sum_{i=1}^{T} \sigma_{\beta_i}^2 x_{it}^2 + \sigma_u^2 & \text{for } t = s \\ 0 & \text{for } t \neq s. \end{cases} \tag{5}$$

The ordinary least squares (OLS) residuals of (3) are given as

$$\hat{\varepsilon}_t = \sum_{j=1}^{T} d_{tj} \varepsilon_j,$$

where d_{tj} denotes the (t, j)-th element of the idempotent matrix

$$D = I - X(X'X)^{-1}X',$$

where X consists of K explanatory variables. Multiplying both sides of (3) by $\hat{\varepsilon}_t$, one obtains

$$y_t \hat{\varepsilon}_t = \sum_{i=1}^{K} \beta_i x_{it} \hat{\varepsilon}_t + \varepsilon_t \hat{\varepsilon}_t, \tag{6}$$

where it is easy to get

$$E(\varepsilon_t\hat{\varepsilon}_t) = \sum_{j=1}^{T} d_{tj}E(\varepsilon_t\varepsilon_j) = d_{tt}E(\varepsilon_t^2)$$

$$= d_{tt}(\sum_{i=1}^{K} \sigma_{\beta_i}^2 x_{it}^2 + \sigma_u^2).$$

Now, (6) is rewritten as

(7) $$y_t\hat{\varepsilon}_t = \sum_{i=1}^{K} (\beta_i x_{it}\hat{\varepsilon}_t + \sigma_{\beta_i}^2 d_{tt}x_{it}^2) + d_{tt}\sigma_u^2 + \eta_t,$$

where

$$\eta_t = \varepsilon_t\hat{\varepsilon}_t - d_{tt}(\sum_{i=1}^{K} \sigma_{\beta_i}^2 x_{it}^2 + \sigma_u^2).$$

Equation (7) includes both means (β_i) and variances $(\sigma_{\beta_i}^2 ; \sigma_u^2)$ as estimable coefficients in a single equation. However, the direct use of OLS on (7) appears to present two problems. One is the nature of the variance-covariance matrix of the η_t's and the other is possible correlation between the new explanatory variables and η_t. By OLS, the first problem causes inefficiency and the second problem causes bias and inconsistency. Let me examine these problems in more detail.

First, I shall examine the nature of random disturbance η_t of (7). By construction, $E(\eta_t)=0$ and

(8) $$E(\eta_t^2) = VAR(\varepsilon_t\hat{\varepsilon}_t) = \sum_{j=1}^{T} d_{tj}^2 VAR(\varepsilon_t\varepsilon_j)$$

$$= \sum_{j\neq t} d_{tj}^2 E(\varepsilon_t^2\varepsilon_j^2) + d_{tt}^2 VAR(\varepsilon_t^2),$$

as $E(\varepsilon_t \varepsilon_j)=0$ for $j \neq t$. Assuming henceforth that ε_t is normally distributed with mean zero and variance (5), I obtain

$$(9) \qquad E(\varepsilon_t^2 \varepsilon_j^2) = \begin{cases} (\sum_{i=1}^{K} \sigma_{\beta_i}^2 x_{it}^2 + \sigma_u^2)(\sum_{i=1}^{K} \sigma_{\beta_i}^2 x_{ij}^2 + \sigma_u^2) & \text{for } t \neq j \\ 3(\sum_{i=1}^{K} \sigma_{\beta_i}^2 x_{it}^2 + \sigma_u^2)^2 & \text{for } t = j, \end{cases}$$

so that (8) is rewritten as

$$(10) \qquad E(\eta_t^2) = (\sum_{i=1}^{K} \sigma_{\beta_i}^2 x_{it}^2 + \sigma_u^2)[\sum_{j \neq t} d_{tj}^2 (\sum_{i=1}^{K} \sigma_{\beta_i}^2 x_{ij}^2 + \sigma_u^2)$$
$$+ 2d_{tt}^2 (\sum_{i=1}^{K} \sigma_{\beta_i}^2 x_{it}^2 + \sigma_u^2)],$$

because, using (9) for $t=j$:

$$VAR(\varepsilon_t^2) = E(\varepsilon_t^4) - E(\varepsilon_t^2) = 2(\sum_{i=1}^{K} \sigma_{\beta_i}^2 x_{it}^2 + \sigma_u^2)^2.$$

Similarly, for $t \neq s$, I obtain

$$E(\eta_t \eta_s) = E(\varepsilon_t \varepsilon_s \hat{\varepsilon}_t \hat{\varepsilon}_s) - d_{tt} d_{ss} (\sum_{i=1}^{K} \sigma_{\beta_i}^2 x_{it}^2 + \sigma_u^2)(\sum_{i=1}^{K} \sigma_{\beta_i}^2 x_{is}^2 + \sigma_u^2).$$

But

$$E(\varepsilon_t\varepsilon_s\hat{\varepsilon}_t\hat{\varepsilon}_s) = E\,[\varepsilon_t\varepsilon_s(\sum_{j=1}^{T} d_{tj}\varepsilon_j)(\sum_{k=1}^{T} d_{tk}\varepsilon_k)]$$

$$= \sum_{j=1}^{T}\sum_{k=1}^{T} d_{tj}d_{tk}E(\varepsilon_t\varepsilon_s\varepsilon_j\varepsilon_k)$$

$$= (d_{tt}d_{ss} + d_{ts}^2)(\sum_{i=1}^{K}\sigma_{\beta_i}^2 x_{it}^2 + \sigma_u^2)(\sum_{i=1}^{K}\sigma_{\beta_i}^2 x_{is}^2 + \sigma_u^2),$$

so that

$$\text{(11)}\qquad E(\eta_t\eta_s) = d_{ts}^2(\sum_{i=1}^{K}\sigma_{\beta_i}^2 x_{it}^2 + \sigma_u^2)(\sum_{i=1}^{K}\sigma_{\beta_i}^2 x_{is}^2 + \sigma_u^2).$$

Equations (10) and (11) indicate that the variance-covariance matrix of the random disturbance of (7) is not diagonal. This means that OLS yields inefficient estimators and Aitken's generalized least squares (GLS) should be used to obtain efficient coefficient estimators of (7).

As for the second problem of correlation between the explanatory variables of (7) and η_t, I observe

$$E(x_{it}\hat{\varepsilon}_t\eta_t) = x_{it}E[\varepsilon_t\hat{\varepsilon}_t^2 - \hat{\varepsilon}_t E(\varepsilon_t\hat{\varepsilon}_t)]$$

$$= x_{it}E[\varepsilon_t(\sum_{j=1}^{T} d_{tj}\varepsilon_j)^2] = 0,$$

because ε_t's are independently normally distributed with means zero, implying that $E(\varepsilon_t\varepsilon_j^2)=0$ for all t and j including the case $t=j$.

Moreover, obviously, $d_{tt}x_{it}^2$ and η_t, and d_{tt} and η_t are uncorrelated because $d_{tt}x_{it}^2$ and d_{tt} are nonstochastic.

Therefore, all the explanatory variables and η_t are uncorrelated in (7). Although this seems to ensure the unbiasedness of the OLS estimators of (7), this is not so because they are not independent of each other. However, obviously, the consistency of the OLS estimators is ensured in such a case.

So far, I have shown that (7) will be consistently and efficiently estimated by GLS by taking into account (10) and (11). In practice, the variances are not known and one ought to use the estimated $(\hat{\sigma}^2_{\beta_i}; \hat{\sigma}^2_u)$ by the OLS of (7) in computing (10) and (11) for the application of GLS. Since $(\hat{\sigma}^2_{\beta_i}; \hat{\sigma}^2_u)$ by OLS are consistent as has been mentioned above, the GLS estimators of the coefficients of (7) are asymptotically efficient.[3/] Here, note that, in order to get asymptotically efficient estimators by GLS, any consistent estimator of the variance-covariance matrix will yield the same asymptotic properties. Then, one can use instead of (10) and (11)

$$(p)\lim_{T\to\infty} E(\eta_t^2) = 2\left(\sum_{i=1}^{K} \sigma^2_{\beta_i} x_{it}^2 + \sigma_u^2\right)^2,$$

$$(p)\lim_{T\to\infty} E(\eta_t \eta_s) = 0,$$

respectively, because $(p)\lim_{T\to\infty} D = I$ so that $(p)\lim_{T\to\infty} d_{ts} = 0$ for $t \neq s$ and $(p)\lim_{T\to\infty} d_{tt} = 1$.

This implies that, one needs only to correct for heteroscedasticity by dividing both sides of (7) by[4/]

$$\sqrt{2}d_{tt}\left(\sum_{i=1}^{K}\hat{\sigma}^2_{\beta_i}x^2_{it}+\hat{\sigma}^2_u\right), \tag{12}$$

or

$$\sqrt{2}\left(\sum_{i=1}^{K}\hat{\sigma}^2_{\beta_i}x^2_{it}+\hat{\sigma}^2_u\right). \tag{13}$$

After the correction for heteroscedasticity, the estimated (β_i) and $(\sigma^2_{\beta_i}; \sigma^2_u)$ will be consistent and asymptotically efficient.

The story does not end here, Additional information can be put to use to increase efficiency. Equation (3), corrected again for the heteroscedasticity of ε_t by dividing both sides of (3) by

$$\sqrt{\sum_{i=1}^{K}\hat{\sigma}^2_{\beta_i}x^2_{it}+\hat{\sigma}^2_u} \tag{14}$$

-- which is the square root of $(p)\lim_{T\to\infty} VAR(\varepsilon^2_t)$ with the true variances being replaced by their consistent estimators--should yield the same (or sufficiently close) estimators of β_i as those β_i estimated by (7).

Thus, one can apply the "mixed estimation" technique of Theil and Goldberger [6] by doubling the number of sample observations. The mixed regression equation can be written

(15)

$$\begin{matrix} T\left\{ \right. \\ T\left\{ \right. \end{matrix}\begin{bmatrix} y_t^* \\ \hline (y_t\hat{\varepsilon}_t)^* \end{bmatrix} = \begin{bmatrix} x_{1t}^* & ,\ldots, & x_{Kt}^* & , & 0 & ,\ldots, & 0 & , & 0 \\ \hline (x_{1t}\hat{\varepsilon}_t)^*, & \ldots, & (x_{Kt}\hat{\varepsilon}_t)^*, & & (d_{tt}x_{1t}^2)^*, & \ldots, & (d_{tt}x_{Kt}^2)^*, & & d_{tt}^* \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_K \\ \sigma_{\beta_1}^2 \\ \vdots \\ \sigma_{\beta_K}^2 \\ \sigma_u^2 \end{bmatrix} + \begin{bmatrix} \varepsilon_t^* \\ \hline \eta_t^* \end{bmatrix},$$

where an "asterisk" denotes the correction for heteroscedasticity by the factors (14) and (12) or (13), respectively, for the first and the second T sample observations. Since, by construction, both ε_t^* and η_t^* now have zero means and unit variances, the estimated standard error of the regression (SER) of (15) by OLS should be unity.[5/]

2. Some Results from Monte Carlo Experiments

In this section, I shall report some results from Monte Carlo experiments. The experiments were intended to check, firstly, the absolute performances of the new estimation technique developed in the previous section and, secondly, its relative performances as compared to the traditional technique proposed by Theil and Mennes [7] and Hildreth and Houck [3] (hereafter TMHH). Forty-four cases were examined in total by varying the values of means (β_i) and variances $(\sigma^2_{\beta_i}; \sigma^2_u)$ and also by altering the sample data x_{it}'s. The number of regressors, K, was three including the constant term in all the experiments. Four different sample numbers were tried: T=15, 30, 50, and 100.[6/]

The overall absolute performances turned out to be reasonably satisfactory. As for the relative performances, it first appears hard to credit one technique and discredit the other. For both techniques simultaneously yielded satisfactory results in some cases and unsatisfactory ones in others. However, it must be noted that these results are certainly in favor of the new estimation technique because it requires less amount of computation or less sequences of regressions than the TMHH in obtaining the same properties of estimators.

As for means $(\hat{\beta}_i)$, both approaches yielded apparently unbiased estimates in all the cases; and the standard errors decreased as the sample size increased. In almost all cases, more efficiency was definitely obtained by applying either technique as compared to employing the standard OLS. This was true irrespectively of the sample size.

As for variances $(\hat{\sigma}^2_{\beta_i}; \hat{\sigma}^2_u)$, however, the results were not always satisfactory. The estimates turned out to be negative in some cases by both techniques. (In such cases, the standard errors were generally large.) As compared to the estimates of means, those of variances did not appear to exhibit unbiasedness. Although it was an anticipated result (as the estimators are only consistent), the overall results did not necessarily show the improvement of performances as the number of sample increased.

Tables 1 to 4 (in ascending order by sample number) report one of the most satisfactory results obtained among 44 experiments. In each table, the top row (a) indicates true values of means and variances. Row (b) gives the estimates of means by applying standard OLS or equation (3) that is the first-step regression for both approaches.

The next three rows are: (c) the second-step (consistent variances); (d) the third-step (consistent and asymptotically efficient variances); and (e) the fourth-step (consistent and asymptotically efficient means) regression equations of the TMHH approach. By this approach the alternative R^2 (computed on the basis of moments around zero rather than around the mean) of the second- and third-step regressions must theoretically equal $\frac{1}{3}$,[7/] and this property can be utilized in checking the performance of the TMHH approach.[8/]

The next four rows summarize the results by the new estimation technique. Row (f) is the second-step regression or equation (7) and it should give consistent means and variances. The variance estimates are to be compared on theoretically the same ground with those of the second-step regression or row (c) of the TMHH approach. Row (g) reports

Table 1 (T=15)

		Const.	β_1	β_2	$\sigma^2_{\beta_1}$	$\sigma^2_{\beta_2}$	σ^2_u	R^2	AR^2	SER
	(a) True Values	100	2	-1	.09	.04	1.69			
	(b) OLS or 1st	97.2	2.11	-.42				.876		
		(47.6)	(8.58)	(1.50)						
TMHH Approach	(c) 2nd				.09	.01	2.17	.022	.529	
					(1.40)	(.12)	(.72)			
	(d) 3rd				.10	.02	2.15	.002	.522	.841*
					(1.29)	(.32)	(.85)			
	(e) 4th	97.2	2.10	-.41				.915		2.542
		(47.4)	(8.50)	(1.44)						
New Approach	(f) 2nd	97.4	2.13	-.43	.09	.01	2.18	.999		
		(45.6)	(7.19)	(1.59)	(1.15)	(.08)	(.55)			
	(g) 3rd	98.2	2.26	-.73	.10	-.03	3.30	.999		.888*
		(51.1)	(7.47)	(2.74)	(1.13)	(.33)	(.99)			
	(h) means	98.7	1.97	-.61				.988		1.007*
		(48.8)	(7.65)	(2.07)						
	(i) mixed regression	98.6	2.07	-.64	.12	.00	2.18	.999		.911*
		(74.8)	(11.4)	(3.43)	(1.34)	(.02)	(.74)			

a/ The numbers in parentheses are t-statistics in absolute value.

b/ An asterisk "*" indicates that the square of SER lies within 95% confidence interval of χ^2 distribution with proper degrees of freedom.

c/ The sample means and variances are:

$$\bar{x}_1 = 4.73,\ \bar{x}_2 = 4.80;\ \text{and}\ s^2_{x_1} = 5.81,\ s^2_{x_2} = 4.41$$

Table 2 (T=30)

		Const.	β_1	β_2	$\sigma^2_{\beta_1}$	$\sigma^2_{\beta_2}$	σ^2_u	R^2	AR^2	SER
	(a) True Values	100	2	-1	.09	.04	1.69			
	(b) OLS or 1st	98.6 (84.7)	2.06 (13.5)	-.80 (5.07)				.895		
TMHH Approach	(c) 2nd				.07 (1.54)	.07 (1.42)	-.14 (.07)	.110	.362	
TMHH Approach	(d) 3rd				.08 (1.68)	.05 (.94)	.71 (.76)	.033	.381	1.171*
TMHH Approach	(e) 4th	98.4 (77.9)	2.09 (13.2)	-.80 (5.03)				.972		2.240*
New Approach	(f) 2nd	96.0 (63.5)	2.49 (12.3)	-.61 (2.77)	.04 (.82)	.04 (.78)	1.82 (.86)	.999		
New Approach	(g) 3rd	98.0 (71.5)	2.23 (10.7)	-.70 (3.25)	.04 (.97)	.04 (.92)	1.42 (.87)	.999		.968*
New Approach	(h) means	99.2 (92.7)	1.98 (13.3)	-.84 (5.22)				.984		.972*
New Approach	(i) mixed regression	98.6 (117)	2.08 (17.5)	-.77 (6.05)	.04 (1.03)	.05 (1.01)	1.26 (.80)	.999		.963*

a/ The numbers in parentheses are t-statistics in absolute value.

b/ An asterisk "*" indicates that the square of SER lies within 95% confidence interval of χ^2 distribution with proper degrees of freedom.

c/ The sample means and variances are:

$$\bar{x}_1 = 4.77,\ \bar{x}_2 = 4.73;\ \text{and}\ s^2_{x_1} = 5.64,\ s^2_{x_2} = 5.26$$

Table 3 (T=50)

		Const.	β_1	β_2	$\sigma^2_{\beta_1}$	$\sigma^2_{\beta_2}$	σ^2_u	R^2	AR^2	SER
	(a) True Values	100	2	-1	.09	.04	1.69			
	(b) OLS or 1st	99.5 (98.7)	2.06 (15.0)	-1.03 (7.57)				.856		
TMHH Approach	(c) 2nd				.10 (2.17)	.16 (3.80)	-2.06 (1.00)	.291	.504	
	(d) 3rd				.10 (2.87)	.10 (1.96)	-.01 (.01)	.415	.534	1.231
	(e) 4th	99.0 (94.2)	2.13 (14.8)	-1.02 (7.38)				.988		2.845
New Approach	(f) 2nd	100.6 (63.5)	2.16 (12.4)	-1.31 (7.10)	.11 (2.45)	.14 (3.25)	-2.09 (.98)	.999		
	(g) 3rd	99.8 (117)	2.11 (12.6)	-1.03 (5.78)	.01 (2.56)	.11 (2.46)	-.47 (.73)	.999		1.104*
	(h) means	100.2 (149)	2.05 (16.1)	-1.10 (7.83)				.999		1.156*
	(i) mixed regression	100.1 (194)	2.07 (20.8)	-1.08 (9.96)	.10 (2.56)	.11 (2.49)	-.42 (.69)	.999		1.115*

a/ The numbers in parentheses are t-statistics in absolute value.

b/ An asterisk "*" indicates that the square of SER lies within 95% confidence interval of χ^2 distribution with proper degrees of freedom.

c/ The sample means and variances are:

$$\bar{x}_1 = 4.84,\ \bar{x}_2 = 5.02;\ \text{and}\ s^2_{x_1} = 6.05,\ s^2_{x_2} = 6.22$$

Table 4 (T=100)

		Const.	β_1	β_2	$\sigma^2_{\beta_1}$	$\sigma^2_{\beta_2}$	σ^2_u	R^2	AR^2	SER
	(a) True Values	100	2	-1	.09	.04	1.69			
	(b) OLS or 1st	99.5 (136)	2.02 (19.5)	-1.00 (9.82)				.821		
TMHH Approach	(c) 2nd				.08 (1.90)	.11 (2.66)	.51 (.25)	.106	.332	
TMHH Approach	(d) 3rd				.08 (1.90)	.02 (.45)	2.58 (3.23)	.073	.298	1.384
TMHH Approach	(e) 4th	99.7 (107)	2.03 (18.0)	-1.05 (9.24)				.981		3.329
New Approach	(f) 2nd	101.6 (80.1)	2.35 (15.0)	-1.48 (10.4)	.10 (2.47)	.10 (2.55)	.48 (.26)	.999		
New Approach	(g) 3rd	99.2 (169)	2.47 (15.2)	-1.19 (8.36)	.10 (2.23)	.03 (.61)	2.22 (3.07)	.999		1.264
New Approach	(h) means	99.3 (204)	2.06 (21.3)	-.99 (9.96)				.997		1.058*
New Approach	(i) mixed regres-sion	99.5 (263)	2.19 (24.9)	-1.09 (13.1)	.09 (2.31)	.03 (.66)	2.32 (3.46)	.999		1.181

a/ The numbers in parentheses are t-statistics in absolute value.

b/ An asterisk "*" indicates that the square of SER lies within 95% confidence interval of χ^2 distribution with proper degrees of freedom.

c/ The sample means and variances are:

$$\bar{x}_1 = 4.91,\ \bar{x}_2 = 5.03;\ \text{and}\ s^2_{x_1} = 6.06,\ s^2_{x_2} = 6.19$$

the third-step regression. Consistent and asymptotically efficient means and variances are estimated in a single equation and it replaces the third- and fourth-step regressions, i.e., rows (d) and (e), of the TMHH approach. Row (h) corresponds to row (d) and it is reported for reference purpose as it alone does not provide any additional information. The last row (i) is the mixed-regression or equation (15) and it utilizes all the data used for rows (g) and (h). It should increase the efficiency of the estimates as compared to the third-step regression or row (g) of the new estimation technique.

The intrinsic heteroscedasticity is corrected for following (12), rather than (13), in running the third-step regression (g) and the mixed-regression (i). Similarly, row (h) and the first half samples of the mixed-regression are corrected for their heteroscedasticities by (14).[9/]

From those Tables, one can see typical results summarized earlier. The estimates of means are almost exactly equal to the true values with high t-statistics. The estimated variances are reasonably close to the true values. Otherwise, their t-values are very small. The standard errors of regression for GLS equations are reasonably close to unity for rows (g), (h), and (i), as it should be so.[10/]

3. Concluding Remarks

Several observations follow. First, the present estimation technique is easily applied for testing the presence of heteroscedasticity along the line of the Glejser test [2]. Second, as has been in fact raised by Monte Carlo experiments, the estimated variances are not guaranteed to be positive. This may be especially so when the number of sample observations is small. Although one can impose constraints that they are nonnegative, as suggested by Hildreth and Houck for their own estimation technique, it is a computationally burdensome procedure. Third, the present estimation technique does not carry through when the random variable u_t of (1) is serially correlated. Typically, the serial correlation of u_t will decrease efficiency although, as is easily seen, consistency is maintained.[11/] Unfortunately, standard methods applied for correcting serial correlations cannot apply here because the coefficients are also stochastic.

References

[1] Asako, K.: An Essay on the Announcement Impact and the Role of Stabilization Policy. Unpublished Ph.D. dissertation submitted to Yale University, 1979.

[2] Glejser, H.: "A New Test for Heteroscedasticity," Journal of American Statistical Association, 64 (1969), 316-23.

[3] Hildreth, C., and J. Houck: "Some Estimators for a Linear Model with Random Coefficients," Journal of American Statistical Association, 63 (1968), 584-95.

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

[5] Theil, H.: Principles of Econometrics. New York: Wiley, 1971.

[6] Theil, H., and A. S. Goldberger: "On Pure and Mixed Statistical Estimation in Economics," International Economic Review, 2 (1961), 65-78.

[7] Theil, H., and L. B. M. Mennes: "Conception stochastique de coefficients multiplicateurs dans l'adjustement linéaire des séries temporelles," Publications de l'Institut de Statistique de l'Université de Paris, 8 (1959), 211-27.

Footnotes

1/. A part of the present paper originates from my Ph.D. dissertation [1] presented to the Faculty of the Graduate School of Yale University. I owe Y. Homma for his help in conducting much of Monte Carlo experiments.

2/. See also Theil [5], pp.622-28, from which I only come to know of [7]; and Maddala [4], pp.392-3, for the description of this approach.

3/. See, for example, Theil [5], theorem 8.4, p.399.

4/. Whether (12) or (13) is better can be judged by the standard error of the regression — which is the square root of the sum of squared residuals divided by the degree of freedom (T-2K-1) — which should be unity after the correction for heteroscedasticity.

5/. This characteristic may help check the specification error of original models.

6/. The random terms are standard normal generated by RANNRML command programed in PEC (Program for Econometric Computation) of Yale University.

7/. See Theil [5], p.626.

8/. One can see that this alternative R^2 is in fact near $\frac{1}{3}$ when T=30 and 100.

9/. Similar formulae are used for the TMHH approach.

10/. In this regard, the new estimation technique performed better than the TMHH approach in all the 44 cases, irrespectively of the sample size. In particular the fourth-step regression or row (e) of the TMHH approach performed very poorly as is reported, for example, in

Tables 1 to 4.

11/. This is so because $(p)\lim_{T\to\infty} E(\eta_t)=0$ even with serial correlation although $E(\eta_t)\neq 0$ for small samples.