Technical Report

SOME IMPLICATIONS OF STRUCTURAL CHANGES WITHIN THE SAMPLE

by

P. N. Misra

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD

SOME IMPLICATIONS OF STRUCTURAL CHANGES WITHIN THE SAMPLE

by

P. N. Misra

July 1973

Indian Institute of Management
Ahmedabad

No. 5

To

Chairman (Research)
IIMA

Technical Report

Title of the report ... SOME IMPLICATIONS OF STRUCTURAL CHANGES WITHIN THE SAMPLE

Name of the Author ... Prof.P.N.Misra

Under which area do you like to be classified? ... [illegible]

ABSTRACT (within 250 words)

An implicit assumption underlying least squares estimation procedure is that the unknown coefficients remain invariant over sample observations. In actual practice, however, one tends to use larger and larger number of observations without verifying as to whether this assumption holds true for the entire set of sample observations. Present article examines the consequence of ignoring this fact under the framework of a general linear regression model. We find that in the presence of parametric shift within the sample, the least squares estimators are biased as well as inefficient and that the explanatory power of the model is reduced. Theoretical findings are supported by empirical evidence.

Please indicate restrictions if any that the author wishes to place upon this note ...

Date 3/8/73 .....

Signature of the Author

SOME IMPLICATIONS OF STRUCTURAL CHANGES WITHIN THE SAMPLE

By

P. N. MISRA*

Indian Institute of Management
Ahmedabad

1. Introduction

An implicit assumption underlying the well known Ordinary Least Square (OLS) procedure of estimation is that the unknown coefficients remain invariant over sample observations. For instance, coefficients in the general linear regression model

$$(1.1)\quad y(t) = \sum_{\lambda=1}^{\wedge} \beta_\lambda x_\lambda(t) + u(t);\ t=1, \ldots\ldots\ldots\ldots, T,$$

are assumed to be constant for all the T observations irrespective of whether they are time series or cross section. What happens if this assumption is violated by the observations at hand? In actual practice, one tends to use larger and larger number of observations for achieving higher degrees of freedom and the same may, often, increase the chances of this assumption to break down. It could be that a structural change – in the sense that β's get changed, may take place somewhere within the sample period, say, after T_1 ($<T$) but one is not aware of such a possibility and goes ahead for estimation and onward inference by using all the T observations. In that case a

*The author is thankful to Miss Anurag for programming and efficient computations.

new type of specification error is being introduced implicitly. This is because as a matter of fact, we have two different models, namely,

$$(1.2)\quad y(t) = \sum_{\lambda=1}^{\wedge} \beta_{\lambda 1} x_{\lambda}(t) + u(t);\ t = 1, \ldots\ldots, T_1$$

and

$$(1.3)\quad y(t) = \sum_{\lambda=1}^{\wedge} \beta_{\lambda 2} x_{\lambda}(t) + u(t);\ t = T_1+1, \ldots\ldots, T$$

but the model to be estimated is incorrectly specified as (1.1).

The purpose of the present article is to analyse the consequence of estimating the model (1.1) while the true specifications are (1.2) and (1.3). In this connection, we analyse bias, efficiency and explanatory power in Sections 2, 3 and 4, respectively. The final section contains some empirical results in support of our theoretical deductions.

2 Bias of OLS Estimator

We may rewrite models (1.1), (1.2) and (1.3) in matrix notation as

$$(2.1)\quad y = X\beta + u$$

$$(2.2)\quad y_1 = X_1\beta_1 + u_1 \text{ and}$$

$$(2.3)\quad y_2 = X_2\beta_2 + u_2$$

where observations are arranged so that

$$(2.4)\quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},\quad X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix},\quad u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Ordinary Least Squares estimator

$$(2.5) \quad b = (X'X)^{-1}X'y$$

of the parameter vector β in model (2.1) can be expressed as

$$(2.6) \quad b = (X'X)^{-1} (X_1'Y_1 + X_2'Y_2)$$
$$= (X'X)^{-1} X_1'X_1\beta_1 + (X'X)^{-1} X_2'X_2\beta_2 + (X'X)^{-1}X_1'u_1 + (X'X)^{-1} X_2'u_2$$

where use has been made of the true relations (2.2) and (2.3). Further, using (2.4) and making suitable adjustment in (2.6) we have

$$(2.7) \quad b = \beta_1 + (X'X)^{-1} X_2'X_2 (\beta_2 - \beta_1) + (X'X)^{-1} X_1'u_1 + (X'X)^{-1} X_2'u_2$$

Finally, taking mathematical expectation on both sides of (2.7) under the usual least squares assumptions namely that x's are non-stochastic and u's have zero mean, we obtain

$$(2.8) \quad E(b) = \beta_1 + (X'X)^{-1} X_2'X_2 (\beta_2 - \beta_1)$$

which makes it clear that the estimator b is biased unless

$$(2.9) \quad \beta_1 = \beta_2 = \beta$$

that is, there is no structural change within the sample period.

3 Efficiency

We can write the bias B of estimator b around β as

(3.1) $$B = \beta_1 - \beta + (X'X)^{-1} X_2'X_2 (\beta_2 - \beta_1)$$

Now variance-covariance matrix of b can be written as

(3.2) $$E(b-\beta)(b-\beta)' = E\left[B + (X'X)^{-1} X_1'u_1 + (X'X)^{-1} X_2'u_2\right]\left[B + (X'X)^{-1} X_1'u_1 + (X'X)^{-1} X_2'u_2\right]' = BB' + \sigma^2 (X'X)^{-1}$$

where, in addition to assumptions used while deriving (2.8), use has been made of the assumption

(3.3) $$E\left[u(t)\, u(t')\right] = \sigma^2 \text{ if } t = t'$$
$$= 0 \text{ if } t \neq t'$$

for all $t = 1, \ldots, T$ and the relation

(3.4) $$X'X = X_1'X_1 + X_2'X_2$$

Obviously, $\sigma^2(X'X)^{-1}$ is the expression one would have obtained if (2.9) were true. Therefore, for the situation when (2.9) is not true the efficiency of the estimator, b goes down because the matrix BB' is positive definite.

4 Explanatory Power

Using estimator b, defined in (2.5), we can derive estimated values of y_1 and y_2 as

(4.1) $y_1^* = X_1 b$

(4.2) $y_2^* = X_2 b$

which yield

(4.3) $y_1^{*\prime} y_1^* = y'X(X'X)^{-1} X_1'X_1 (X'X)^{-1} X'y$

$y_2^{*\prime} y_2^* = y'X(X'X)^{-1} X_2'X_2 (X'X)^{-1} X'y$

Similarly, defining

(4.4) $y^* = Xb$

and using (3.4) and (4.3) we have

(4.5) $y^{*\prime} y^* = y_1^{*\prime} y_1^* + y_2^{*\prime} y_2^*$

We may now define three ratios in terms of amount of variation explained to total variation as

(4.6) $R^2 = (y'y)^{-1} y^{*\prime}y^*$, $R_1^{*2} = (y_1'y_1)^{-1} y_1^{*\prime} y_1^*$, $R_2^{*2} = (y_2'y_2)^{-1} y_2^{*\prime} y_2^*$

It can be easily verified with the help of relation (4.5) that these ratios are related as

(4.7) $R^2 = (y'y)^{-1} y_1'y_1 R_1{*}^2 + (y'y)^{-1} y_2'y_2 R_2{*}^2$

which means that R^2 is weighted average of $R_1{*}^2$ and $R_2{*}^2$ where the sum of the weights is unity. Therefore, we have

(4.8) $R^2 <$ maximum of $(R_1{*}^2, R_2{*}^2)$

Alternatively, considering the OLS estimators

(4.9) $b_1 = (X_1'X_1)^{-1} X_1'y_1$

and

(4.10) $b_2 = (X_2'X_2)^{-1} X_2'y_2$

of the parameter vectors in the true equations (2.2) and (2.3), respectively, we get estimated y's as

(4.11) $\hat{y}_1 = X_1 b_1$

and

(4.12) $\hat{y}_2 = X_2 b_2$

Further, according to least squares principle as applied to model (2.2) we have

(4.13) $y_1'y_1 - y_1{*}'y_1{*} \leq y_1'y_1 - \hat{y}_1'\hat{y}_1$

because error sum of squares is least corresponding to estimator b_1 only.

Relation (4.13) means

$$(4.14) \quad y_1^{*\prime} y_1^* < \hat{y}_1' \hat{y}_1$$

or simply

$$(4.15) \quad R_1^{*2} < R_1^2$$

where

$$(4.16) \quad R_1^2 = (y_1' y_1)^{-1} \hat{y}_1' \hat{y}_1$$

Similarly, it can be shown that

$$(4.17) \quad R_2^{*2} < R_2^2$$

where

$$(4.18) \quad R_2^2 = (y_2' y_2)^{-1} \hat{y}_2' \hat{y}_2$$

Finally, combining (4.15) and (4.17) with (4.8) we obtain

$$(4.19) \quad R^2 < \text{maximum of } (R_1^2, R_2^2)$$

which implies that atleast one of the true models (2.2) and (2.3) has higher explanatory power than the mis-specified model (2.1).

5 Some Empirical Results

Our theoretical deductions in previous sections indicate that an incorrectly specified model like (2.1) yields not only biased and inefficient estimators but its explanatory power also gets reduced. Empirical evidence to these results is found to be substantial.

For this purpose, we considered three equations of Klein's model I (1, pp.432-33) and estimated the results as follows: Estimates of bias of the estimator b have been obtained around the true parametre vectors β_1 and β_2. To obtain this, as in (2.8), we require knowledge of β_1 and β_2 themselves which have been estimated by b_1 and b_2. Sample observations (T=21) have been divided into two groups (T_1=10) and (T_2=11) and the same have been used to estimated the models (2.2) and (2.3).

As each one of the three equations contains four unknown coefficients, we denote them, for the sake of simplicity by b_o, b_1, b_2 and b_3 in each case. Estimates of these coefficients are provided in Table 1.

Table 1

ESTIMATION OF COEFFICIENTS

Model	Sample size	b_o	b_1	b_2	b_3
Consumption	21	16.236	0.193	0.090	0.796
	10	13.646	0.148	0.280	0.794
	11	22.209	0.656	0.231	0.454
Investment	21	8.365	0.515	0.226	-0.098
	10	9.055	0.508	0.330	-0.109
	11	47.700	-0.012	0.408	-0.267
Labour	21	1.497	0.439	0.146	0.136
	10	11.920	0.609	-0.160	0.657
	11	0.321	0.479	0.143	-0.820

Results given in Table 1 clearly indicate that the two sets of observations imply two sets of coefficients in case of each equation. Strictly speaking, one could argue that such a statement should be made after testing, for a chosen significance level, the difference between the two sets of coefficients for each equation. But the issue under question relates to absolute difference, if any, and the situation when such a difference tends to vanish creates no problem.

Table 2

ESTIMATES OF BIAS

Model	Around	b_0	b_1	b_2	b_3
Consumption	β_1	16.943	1.328	1.365	-1.292
	β_2	6.807	0.396	2.088	-1.348
Investment	β_1	30.141	-0.232	-0.055	-0.127
	β_2	-8.487	0.289	-0.134	0.031
Labour	β_1	-0.449	-0.032	-0.012	-1.652
	β_2	14.809	0.237	-0.443	0.621

Magnitude of bias, as given in Table 2, indicates that the vector is neither equal to β_1 nor β_2. This is what one would expect on apriori reasoning because combination of heterogenous observations implies altogether different parametric vector.

We now turn to examine, numerically, the efficiency of the estimator b. As indicated earlier, we have two sets of biases, namely, around β_1 and β_2. Therefore, substituting these biases in (3.2) we derive two sets of sampling variances of the elements of b. These results are given in Table 3. The column titled 'specification' indicates that the variances correspond to individual models if the same were the true models but the model (2.1) was estimated.

Table 3

ESTIMATES OF VARIANCES OF OLS ESTIMATORS

Model	Specification	b_0	b_1	b_2	b_3
Consumption	(2.1)	1.698	0.008	0.008	0.016
	(2.2)	288.755	1.771	1.872	1.685
	(2.3)	48.034	0.164	4.367	1.832
Investment	(2.1)	26.235	0.009	0.008	0.005
	(2.2)	934.689	0.063	0.011	0.017
	(2.3)	98.261	0.092	0.026	0.001
Labour	(2.1)	1.613	0.001	0.001	0.001
	(2.2)	1.815	0.002	0.001	0.280
	(2.3)	220.907	0.057	0.197	0.387

Obviously, from the results given in Table 3, one notes that efficiency of the estimator b goes down substantially in case of each equation. In fact, the more the magnitude of bias, the further down goes the efficiency of the estimator.

Table 4

ESTIMATES OF R^2

Model	R^2	R_1^2	R_2^2
Consumption	0.9994	0.9996	0.9973
Investment	0.9387	0.9497	0.8370
Labour	0.9997	0.9992	0.9999

Loowking at the magnitude of R^2 given in Table 4, one notes that the empirical results are entirely in tune with what was expected according to the result (4.19).

REFERENCE

1. Theil Henry, Princples of Econometrics, North-Holland Publishing company, Amsterdam - London, 1971.