



Test for Structural Change under Heteroscedastic Errors: The Case of Successive Regressions

Jagabandhu Saha

Department of Economics and Politics,
Visva-Bharati University (India)

Abstract: *The Chow test is not robust under heteroscedasticity. The presence of heteroscedasticity will affect level of significance as well as power of the test, especially when the sizes of the samples are small. The present paper not only resolves the problem of heteroscedasticity in the error terms, but also extends the existing method of comparing from two regression equations to many equations in order to make comparisons of the successive coefficients to be possible, thus generalizing Chow test in two directions. The procedure is then illustrated by an attempt to detect structural change, if any, through comparison of successive decadal growth rates of the population of India, using NSSO data.*

Key words: *Chow test, Heteroscedastic error, Comparisons of successive regression coefficients.*

I. INTRODUCTION

It is a common practice to test the equality between sets of coefficients in two linear regressions by Chow Test (Chow 1960)^[1]. Chow however assumed homoscedasticity of the regression errors. It is already demonstrated in the literature that the Chow test is not robust under heteroscedasticity (Toyoda 1974^[2], Schmidt and Sickles 1977^[3], Ali and Silver 1985^[4] and Tansel 1987^[5]). The presence of heteroscedasticity will affect level of significance as well as power of the test. This means that if there is heteroscedasticity in the errors, but we perform Chow test assuming homoscedasticity then the result may be different from the actual especially when the sizes of the samples are small.

Under homoscedasticity assumption in the Chow Test, if the null hypothesis of equality between the sets of coefficients is not rejected then there is no problem (as in the examples in his paper). But if rejected, then, naturally, one is probed to the questions: a). at which component/s the sets differ, and b). for each of those components, between the two coefficients of the two regressions concerned, which one is larger/smaller. Chow test does not provide answer to any of these questions. This problem can be resolved with some modifications of the model (Saha and Pal 2014)^[6]. Saha and Pal introduced the concept of “component wise complete comparison” (CCC)¹ in order to overcome this problem. The test procedure for CCC between every two successive regressions out of any number of given successive regressions was developed, thus generalizing Chow test in two directions. If heteroscedasticity is present then the problem of CCC aggravates and needs further modifications. The present paper extends the earlier paper by incorporating heteroscedasticity in the model and developing test procedure for CCC between every two successive regressions out of any number of given successive regressions, thus generalizing Chow test in three directions.

We may now straight go to the problem and discuss how we can arrive at a solution.

2. The Model: We consider the problem of finding test procedure for CCC between every two successive regressions out of m given successive regressions as follows:

$$\begin{aligned} y^{(1)} &= a_1^{(1)} + a_2^{(1)} x_2^{(1)} + a_3^{(1)} x_3^{(1)} + \dots + a_k^{(1)} x_k^{(1)} + u^{(1)}, \\ y^{(2)} &= a_1^{(2)} + a_2^{(2)} x_2^{(2)} + a_3^{(2)} x_3^{(2)} + \dots + a_k^{(2)} x_k^{(2)} + u^{(2)}, \\ &\dots\dots\dots \\ y^{(m)} &= a_1^{(m)} + a_2^{(m)} x_2^{(m)} + a_3^{(m)} x_3^{(m)} + \dots + a_k^{(m)} x_k^{(m)} + u^{(m)}, \end{aligned} \quad \dots\dots\dots (1)$$

where, the superscripts denote the individual regressions, n_1, n_2, \dots, n_m are the nos. of observations for these regressions and the errors are Heteroscedastic in the sense as formulated by the assumptions as given under:

¹ By complete comparison between any two parameters a and b we mean to decide whether $a < b$ or $a = b$ or $a > b$. By component wise complete comparison (CCC) between two vectors of parameters of the same size ($a_1 a_2 \dots a_m$) and ($b_1 b_2 \dots b_m$) we mean complete comparison between (a_1 and b_1), (a_2 and b_2), ... and (a_m and b_m). By CCC between/of/for two regressions with same no. of parameters we mean CCC between the two vectors of parameters of these regressions. In the paper by Saha and Pal, CCC is done between every two successive regressions out of any number of given successive regressions with same no. of parameters.

$$\begin{aligned}
 & \text{i). } E(u^{(i)}) = 0_{n \times 1}, \quad \forall i = 1, 2, \dots, m, \\
 & \text{ii). } E((u^{(i)})(u^{(i)})') = \sigma_i^2 I_{n \times n}, \quad \forall i = 1, 2, \dots, m, \\
 & \text{iii). } E((u^{(i)})(u^{(j)})') = 0_{n \times n}, \quad \forall i \neq j = 1, 2, \dots, m, \dots \dots \dots (2) \\
 & \text{where, } n_i = n, \quad \forall i = 1, 2, \dots, m,
 \end{aligned}$$

(i.e., the sample sizes for the different regressions are the same, say, n).

It is admitted that a particular type of heteroscedasticity of the errors in the regressions in (1) has been adopted --- this is in the light of the model adopted in the Zellner's (1962)^[7] SURE Estimation Procedure, and the solution here is, also, similar to that of Zellner's.

3. The Methodology: Firstly we combine the m regressions in (1) into a single regression equation model as follows.

$$\begin{pmatrix} y_1^{(1)} \\ \vdots \\ y_n^{(1)} \\ y_1^{(2)} \\ \vdots \\ y_n^{(2)} \\ \vdots \\ y_1^{(m)} \\ \vdots \\ y_n^{(m)} \end{pmatrix} = \begin{pmatrix} 1 & x_{21}^{(1)} & \dots & x_{k1}^{(1)} & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ 1 & x_{2,n}^{(1)} & \dots & x_{k,n}^{(1)} & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & x_{21}^{(2)} & \dots & x_{k1}^{(2)} & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 & 1 & x_{2,n}^{(2)} & \dots & x_{k,n}^{(2)} & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 1 & x_{21}^{(m)} & \dots & x_{k1}^{(m)} \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 1 & x_{2,n}^{(m)} & \dots & x_{k,n}^{(m)} \end{pmatrix} \begin{pmatrix} a_1^{(1)} \\ \vdots \\ a_k^{(1)} \\ a_1^{(2)} \\ \vdots \\ a_k^{(2)} \\ \vdots \\ a_1^{(m)} \\ \vdots \\ a_k^{(m)} \end{pmatrix} \begin{pmatrix} u_1^{(1)} \\ \vdots \\ u_n^{(1)} \\ u_1^{(2)} \\ \vdots \\ u_n^{(2)} \\ \vdots \\ u_1^{(m)} \\ \vdots \\ u_n^{(m)} \end{pmatrix} \dots \dots \dots (3)$$

The solution for the single equation model is same as that of finding solution separately for each equation in (1). The benefit of writing a single equation model is that the error terms in (1) are now constituents of the one error vector of this single equation model and hence we can now conceive the heteroscedasticity of these error terms easily, heteroscedasticity in the sense as formulated by (2). In addition to setting idea for heteroscedasticity of the error terms, we want to compare $a_j^{(i)}$ with $a_j^{(i+1)}$, for all $j = 1, 2, \dots, k$ and $i = 1, 2, \dots, (m-1)$. That is also possible if we slightly change the model further.

Notice that the above model does not have an intercept term. We may now introduce the intercept term in (3) and rewrite (3) as follows.

$$\begin{pmatrix} y_1^{(1)} \\ \vdots \\ y_n^{(1)} \\ y_1^{(2)} \\ \vdots \\ y_n^{(2)} \\ \vdots \\ y_1^{(m)} \\ \vdots \\ y_n^{(m)} \end{pmatrix} = \begin{pmatrix} 1 & x_{21}^{(1)} & \dots & x_{k1}^{(1)} & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ 1 & x_{2,n}^{(1)} & \dots & x_{k,n}^{(1)} & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 1 & x_{21}^{(2)} & \dots & x_{k1}^{(2)} & 1 & x_{21}^{(2)} & \dots & x_{k1}^{(2)} & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ 1 & x_{2,n}^{(2)} & \dots & x_{k,n}^{(2)} & 1 & x_{2,n}^{(2)} & \dots & x_{k,n}^{(2)} & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ 1 & x_{21}^{(m)} & \dots & x_{k1}^{(m)} & 1 & x_{21}^{(m)} & \dots & x_{k1}^{(m)} & \dots & 1 & x_{21}^{(m)} & \dots & x_{k1}^{(m)} \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{2,n}^{(m)} & \dots & x_{k,n}^{(m)} & 1 & x_{2,n}^{(m)} & \dots & x_{k,n}^{(m)} & \dots & 1 & x_{2,n}^{(m)} & \dots & x_{k,n}^{(m)} \end{pmatrix} \begin{pmatrix} c_{11} \\ \vdots \\ c_{1k} \\ c_{21} \\ \vdots \\ c_{2k} \\ \vdots \\ c_{m1} \\ \vdots \\ c_{mk} \end{pmatrix} \begin{pmatrix} u_1^{(1)} \\ \vdots \\ u_n^{(1)} \\ u_1^{(2)} \\ \vdots \\ u_n^{(2)} \\ \vdots \\ u_1^{(m)} \\ \vdots \\ u_n^{(m)} \end{pmatrix} \dots \dots \dots (4)$$

In the above model (4), c_{11} is the intercept term . This is same as $a_1^{(1)}$, the intercept term in the first regression equation in (1). Similarly, $c_{12}, c_{13}, \dots, c_{1k}$, are also same as $a_2^{(1)}, a_3^{(1)}, \dots, a_k^{(1)}$. The regression coefficients $c_{21}, c_{22}, \dots, c_{2k}$ are the changes in the intercept term and the coefficients of other variables in the second equation from the corresponding values of the first equation, $c_{31}, c_{32}, \dots, c_{mk}$ are the changes in the intercept term and the coefficients of other variables in the third equation from the corresponding values of the second equation, and so on. We can thus write $c_{ij} = a_j^{(i)}$, for all $j = 1, 2, 3, \dots, k$ and $c_{ij} = a_j^{(i)} - a_j^{(i-1)}$, for all $j = 1, 2, 3, \dots, k$ and $i = 2, 3, \dots, m$. So, our task is to device test in order to decide whether the differentials c_{ij} 's are < 0 or $= 0$ or > 0 , for all $j = 1, 2, \dots, k$ and $i = 2, 3, \dots, m$.

Let us, for convenience, rewrite (4) as:

$$Y = X c + U, \quad \dots (5)$$

where, $Y_{N \times 1}$ = the Y-vector in (4), $X_{N \times K}$ = the X-matrix in (4), $c_{K \times 1}$ = the coefficient-vector in (4) and $U_{N \times 1}$ = the disturbance-vector in (4), $N = nm$ and $K = km$.

We can now estimate c as well as perform test for $H_0 : c_{ij} = 0$ vs. $H_A : c_{ij} \neq 0$ or $H_A : c_{ij} < 0$ or $H_A : c_{ij} > 0$, for all $j = 1, 2, 3, \dots, k$ and for all $i = 2, 3, \dots, m$, and decide whether $c_{ij} < 0$ or $= 0$ or > 0 , for all $j = 1, 2, 3, \dots, k$ and for all $i = 2, 3, \dots, m$, i.e., perform CCC between every two successive regressions in m -regression equations model, since $c_{ij} = a_j^{(i)} - a_j^{(i-1)}$. And once this is done, the point/each of the points of structural change, if there is any at all, will be automatically detected --- with detailed information for that point/each of those points.

But, however, we cannot run the usual regression with (5) for this purpose simply because the model (5) is a Generalised Least Squares Model (GLSM) [8]. The estimation procedure will depend on the variance- covariance matrix of the regression error, which is given as:

$$(D(U))_{N \times N} = V, \text{ say, } = \begin{pmatrix} \sigma_1^2 I_{n \times n} & 0_{n \times n} & & 0_{n \times n} \\ 0_{n \times n} & \sigma_2^2 I_{n \times n} & & 0_{n \times n} \\ & & \ddots & \\ 0_{n \times n} & 0_{n \times n} & & \sigma_m^2 I_{n \times n} \end{pmatrix}$$

or, $V = \sum_{m \times m} \otimes I_{n \times n}$, (6)

where, $\sum_{m \times m}$ is :

$$\sum_{m \times m} = \begin{pmatrix} \sigma_1^2 & 0 & & 0 \\ 0 & \sigma_2^2 & & 0 \\ & & \ddots & \\ 0 & 0 & & \sigma_m^2 \end{pmatrix} .$$

So the model (5) is a GLSM with the variance-covariance matrix of the regression error as given by (6). The GLS estimator of c based on (5) is given as:

$$c^* = (X'V^{-1}X)^{-1}X'V^{-1}Y, \text{ (7)}$$

and its dispersion matrix is:

$$D(c^*) = (X'V^{-1}X)^{-1}. \text{ (8)}$$

But due to (6), V is unknown as $\sum_{m \times m}$ is so. So it is not possible to use (7) and (8) in practice, particularly for the testing purposes we are aimed at. Henceforth let us proceed following Zellner [7].

Firstly, we need to estimate V . For that we need to estimate $\sum_{m \times m}$ and that is done as follows. The steps are:

i). Apply OLS separately to each of the regressions in (1); let the residual vector for the i -th regression be denoted as e^i , for all $i = 1, 2, \dots, m$,

ii). Estimate σ_i^2 as : $s_i^2 = (e^{i'}e^i) / (n-k)$, for all $i = 1, 2, \dots, m$.

Then, estimated $\sum_{m \times m}$, say, $S_{m \times m}$, is:

$$S_{m \times m} = \begin{pmatrix} s_1^2 & 0 & & 0 \\ 0 & s_2^2 & & 0 \\ & & \ddots & \\ 0 & 0 & & s_m^2 \end{pmatrix} .$$



Then, V is estimated as :

$$\hat{V} = S_{m \times m} \otimes I_{n \times n} \quad \dots\dots\dots (9)$$

Now we replace V in (7) by \hat{V} as given by (9) and form the estimator:

$$c^{**} = (X'(\hat{V})^{-1}X)^{-1}X'(\hat{V})^{-1}Y \quad \dots\dots\dots (10)$$

Then it follows that $(n^{1/2})(c^{**} - c)$ has asymptotic normal distribution and the dispersion matrix of c^{**} is:

$$D(c^{**}) = (X'(\hat{V})^{-1}X)^{-1} + o_p(n^{-1}),$$

where $o_p(n^{-1})$ denotes terms of high order of smallness than n^{-1} .

So, for large value of n, c^{**} is normally distributed. Also, evidently, for large n, $o_p(n^{-1})$ is negligible and then,

$$D(c^{**}) \simeq (X'(\hat{V})^{-1}X)^{-1}.$$

So, for large n, we have:

$$c^{**} \sim N_K(c, (X'(\hat{V})^{-1}X)^{-1}) \quad \dots\dots\dots (11)$$

Now, the tests that we require are obvious, provided that n is sufficiently large which we assume for rest of the paper. Representing,

$$c^{**} = (c^{**}_1 \ c^{**}_2 \ c^{**}_3 \ \dots\dots \ c^{**}_K)' ,$$

$$c = (c_1 \ c_2 \ c_3 \ \dots\dots \ c_K)' , \text{ and}$$

$$(X'(\hat{V})^{-1}X)^{-1} = (a_{ij})_{K \times K},$$

we have: $(c^{**}_i - c_i) / (a_{ii})^{1/2} \sim N(0, 1)$, for all $i = 1, 2, 3, \dots\dots, K$.

Hence for the null hypothesis: $H_0 : c_i = 0$,

the test statistic is: $T = c^{**}_i / (a_{ii})^{1/2}$, and $\dots\dots\dots (12)$

$$T \sim N(0, 1), \text{ under } H_0, \text{ for all } i = 1, 2, 3, \dots\dots, K.$$

Needless to say that each of the tests here is a normal test.

4. Illustration: In the context of rate of growth of population in India, we consider three regression equations ($m = 3$) as follows. With state level population of India, we first define the following four variables:

- X_1 = size of the population in a state of India in 1981,
- X_2 = size of the population in a state of India in 1991,
- X_3 = size of the population in a state of India in 2001,
- X_4 = size of the population in a state of India in 2011.

The sources of these data are Census of India (1981, 1991, 2001, 2011) ^[9].

Let us now define variables Y_1, Y_2, Y_3 as follows:

$$\begin{aligned}
 Y_1 &= X_2 - X_1 \text{ (i.e., growth/increase of population during: 1981 to 1991)} \\
 Y_2 &= X_3 - X_2 \text{ (i.e., growth/increase of population during: 1991 to 2001)} \\
 Y_3 &= X_4 - X_3 \text{ (i.e., growth/increase of population during: 2001 to 2011).}
 \end{aligned}$$

We now consider three regressions as follows:

$$\left. \begin{aligned}
 Y_1 &= \beta_1 X_1 + U_1 \\
 Y_2 &= \beta_2 X_2 + U_2 \\
 Y_3 &= \beta_3 X_3 + U_3
 \end{aligned} \right\} \dots (13)$$

β_1 , β_2 and β_3 are nothing but the rates of growth of population over the decades: 1981 to 1991, 1991 to 2001 and 2001 to 2011 respectively (to be referred as first decade, second decade and so on).

Now, following Section 3, we apply OLS separately to each of the above three regressions in (13). (It may be noted that each of these regressions is a regression without an intercept term.) (The no. of observations for each regression here is $n = 32$ (no. of States in India); also, we have here: $k = 1$, $m = 3$.)

The Residual vectors of these three regressions are first obtained. Then we get the sum of squares for the residual vectors and hence the estimates of σ_i^2 's (s_i^2 's) using the formula as given already. We then use the following steps to get the estimate c^{**} in (10), $D(c^{**})$ in (11), and thence the value of the test statistic T in (12), the value of c^{**} representing as: first component gives the estimate of the growth rate in the first decade (β_1), and the second and the third components give respectively the estimates of changes in the growth rates over first decade to second one ($\beta_2 - \beta_1$) and over second decade to third one ($\beta_3 - \beta_2$).

1. Construct the matrix $S_{3 \times 3}$ and compute $(S_{3 \times 3})^{-1}$,
2. Compute: $(\hat{V})^{-1}_{96 \times 96} = S_{3 \times 3}^{-1} \otimes I_{32 \times 32}$,
3. Compute the matrix: $(a_{ij})_{K \times K} = (X'(\hat{V})^{-1}X)^{-1}$,
4. Compute c^{**} as: $c^{**} = (X'(\hat{V})^{-1}X)^{-1} X'(\hat{V})^{-1}Y$,
5. Compute the test statistic: $T = c^{**}_i / (a_{ii})^{1/2}$.

The estimate of growth rate in the first decade is 2.408 and the estimates of the changes concerned are respectively -2.270 and 0.043 with the corresponding T-values as 132.029, -86.159 and 1.362. The second T-value indicates that there is a decline in growth rate as one moves from the first decade to the second one while the third T-value indicates that the estimated change in growth rate when one moves from the second decade to the third one is insignificant. Hence there is structural change only once and the change is negative.

5. Conclusions: Our procedure extends the existing method of comparing, from two regression equations to many equations enabling one to compare successive coefficients, from assumption of homoscedastic errors to heteroscedastic errors, thus generalizing Chow test in two directions.

Firstly, we can compare whether any two coefficients are equal against the alternative hypotheses of inequality of any direction i.e., ' $<$ ' or ' $>$ ', instead of only ' \neq '. This can further be extended to vector of regression coefficients with similar alternative hypotheses for each component of vector.

Secondly, our procedure enables one to perform component wise complete comparison between the vectors of coefficients of every two successive regressions out of several given regressions. Now, one of the important implications of this is as follows. Suppose each one of the given regressions pertains to a time period/point and the regressions are arranged in increasing order of time and the investigator is in search of a) existence of structural



breakthrough and b) detection of the point/s (here, by a point we mean a time period or a time point) where it occurs, if there is any such at all. Not only the point/s of structural breakthrough, if there is any at all, through our procedure we get something more. For every such point we get component wise complete comparison of the vectors of coefficients of the two regressions associated with that point. Actually, it is not necessary that the regressions need to be ordered in increasing/decreasing order of time; it is sufficient for the regressions to be ordered in a well defined sense, e.g., (i) in order of space, e.g., regressions pertain to some states of India arranged from North to South, (ii) in increasing order of income, e.g., regressions pertain to some groups of peoples arranged in increasing order of income, etc. It seems that the concept of “Structural Change” can be extended, not pertaining to only “order of time” but pertaining to any well defined order in which the regressions can be meaningfully arranged.

Thirdly, consider the test provided by Gujarati (Gujarati 1970)^[10], Generalised Dummy Variable Approach, in order to find out whether a given set of regressions differ from one another. A moment’s reflection shows that the purpose of this test is also served by our test simply because if we arrange these regressions successively (with or without any definite meaning) then we can say that these regressions do not differ from one another iff the two vectors of coefficients of every two successive regressions coincide which is easily verifiable by our procedure. But, needless to say, the objective of this paper, i.e., developing test procedure for CCC between every two successive regressions out of any number of given successive regressions, is not served by the test due to Gujarati.

REFERENCES

1. Chow, Gregory C. (1960): Tests of Equality between Sets of Coefficients in Two Linear Regressions, *Econometrica*, Vol. 28, No. 3, pp. 591-605.
2. Toyoda, Toshihisa (1974): Use of the Chow Test under Heteroscedasticity, *Econometrica*, Vo. 42, No. 3, May, pp. 601-608.
3. Schmidt, P and Sickles, R (1977): Some Further Evidence on the Use of the Chow Test under Heteroscedasticity, *Econometrica*, Vol 45, pp. 1293-1298.
4. Ali, Mukhtar M and Silver J. Lew (1985): Tests for Equality between Sets of Coefficients in Two Linear Regressions under Heteroscedasticity, *J. of the Am. Stast. Assoc.*, Vol. 80, No. 391, pp. 730-734.
5. Tansel, Aysit (1987): Testing for Structural Change under Heteroscedasticity: A Note and an Application, *Commun. Fac Sci., Univ. Ank., Ser. A, V. 36, No. 2*, pp. 55-67.
6. Saha, J. & Pal, M (2014): A Modified Chow Test Approach Towards Testing Differences in The Engel Elasticities, *Asian-African Journal of Economics and Econometrics*, Vol.14, No.1, 2014; 57-67.
7. Zellner, Arnolds (1962): An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias, *Journal of the American Statistical Association*, Vol.57, 1962, pp.348-368.
8. Johnston J. (1984): *Econometric Methods*, Third Edn., McGraw-Hill, New York.
9. Census of India (1981, 1991, 2001, 2011): Government of India.
10. Damodar Gujarati (1970): Use of Dummy Variables in Testing for Equality Between Sets of Coefficients in Linear Regressions: A Generalisation, *The American Statistician*, December 1970, pp.18-22.