HETEROSCEDASTICITY

3

Q.1. What is meant by the problem of heteroscadasticity? What are the plausibility of the assumption of heteroscadasticity? What are the consequences of the violation of the assumption of homoscadasticity?

2. How will you detect the presence of heteroscadastic disturbances. Discuss some methods of solving this problem.

Heteroscadasticity

Ans:- One important assumption about the random variable U is that its probability distribution remains the same over all observations of x, and in particular that the variance of each U is the same for all values of the explanatory variable. Symbolically we have $Var(U) = E\{(u_i - E(u_i)\}^2 = E(u_i)^2 = \sigma_u^2 = \text{constant}$

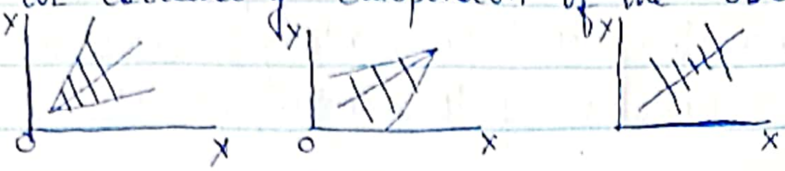
This assumption is known as the assumption of homoscadasticity as the assumption of constant variance of U's. If it is not satisfied in any particular case, we say that U's are heteroscadastic. $Var(u_i) = \sigma_{u_i}^2 \neq \text{constant}$.

Here the subscript (i) signifies the fact that the individual variances may all be different.

The meaning of the assumption of homoscadasticity is that the variation of each u_i around its zero mean does not depend on the values of x. The variance of each u_i remain the same, irrespective of small as large values of the explanatory variable, i.e. $\sigma_u^2 \neq f(x_i)$.

If $\sigma_u^2 \neq$ constant, and its values depend on the values of x, we may write $\sigma_{u_i}^2 = f(x_i)$.

The case of heteroscadasticity is shown by the increasing as decreasing dispersion of the observations from the regression line.



It should be clear that the pattern of heteroscadasticity depends on the signs and values of the co-efficients of the relationship $\sigma_{u_i}^2 = f(x_i)$. Since the U's are not observable we do not know the true pattern of heteroscadasticity. In applied research econometricians usually make the convenient assumption that the heteroscadasticity is of the form $\sigma_{u_i}^2 = k^2 x^2$, where K is a constant to be estimated from the model.

Plausibility of the assumption of Heteroscadasticity

In many econometric studies, especially those based on cross section data, the assumption of a constant variance for the disturbance term is unrealistic. This can easily be understood if we take into the factors whose influences are ~~absorbed~~ observed by the disturbance term. V expresses the influences on the dependent variables of errors in its measurements of and of omitted variables. On both accounts there are reasons for expecting the variance of U to vary over time, as to vary systematically with the explanatory variable x. Thus as Y increases ~~errors~~ errors of measurement tend to increase, because it becomes more difficult to collect data and check their consistancy and reliability. Furthermore, the errors of measurement tend to be cumulative over time, so that their size tends to increase. In this case the variance of u_i increases with the increasing values of x. On the otherhand, the sampling techniques and various other methods of collecting data are continuously improving, and thus errors of measurement may decrease. In this case $var(u_i)$ decreases over time. Most important, many of the variables which are omitted from the function tend to change in the same direction with x, thus causing an increase of the variation of the observations from the regression line.

For example, let us suppose that we have a cross section sample of family budgets from which we want to measure the saving function.

$S_i = b_0 + b_1 Y_i + u_i$, where S_i = saving of the ith household
Y_i = income of the ith household

In this case the assumption of constant variance of the u's is not ~~appropriate~~ appropriate, because high income families show a much greater variability in their savings behaviour than do low income families. Families with high income stick to a certain standard of living and when their income falls they cut down their savings rather than their consumption expenditure. On the otherhand, low income families save for certain purposes (for example to pay some instalments or to repay debts) and thus their savings pattern are more regular. This implies that at high incomes u_i's will be higher, while at low incomes u_i's will be small.

ESTIMATION.

We are considering a model where disturbance term is heteroscedastic but pairwise unautocorrelated. If in our model, $E(uu') = \sigma^2 I_n$, we apply OLS method. But if $E(uu') = \sigma^2 \Omega$, the 2nd assumption of classical linear model is violated. So OLS cannot be applied, because if we apply OLS method, the corresponding estimator will be linear and unbiased,

(***) We consider the model $Y = X\beta + u$ ---- (1), where Y is n-element column vector
β is K- " "
u is n- " "
X is a matrix with n rows and K columns.

Our assumptions are — (i) $E(u_i) = 0$ ..... (2a)
(ii) $E(uu') = \sigma^2 \Omega$, since there is heteroscedasticity --- (2b)
(iii) X is a non-stochastic matrix —— (2c)
(iv) $\rho(X) = K < n$.

but its minimum variance property will be destroyed. So, here we have to apply GLS method. In order to do so we require a transformation matrix. Here all the OLS assumption remain intact.

(***)

~~[In the case of heteroscedasticity $E(uu') = \sigma^2\Omega$, where]~~

Now Ω is a positive definite matrix.

Let $\Omega = \begin{bmatrix} 1/\lambda_1 & 0 & \cdots & 0 \\ 0 & 1/\lambda_2 & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & 1/\lambda_n \end{bmatrix}$ and $\Omega^{-1} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & & & \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$ —— (3a)

$$\therefore \quad E(uu') = \sigma^2 \begin{bmatrix} 1/\lambda_1 & 0 & \cdots & 0 \\ 0 & 1/\lambda_2 & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & 1/\lambda_n \end{bmatrix} \quad \text{------ (3b)}$$

Now any positive definite matrix is expressed in the form PP', where P is any $n \times n$ nonsingular matrix.

$\therefore PP' = \Omega$ .... (4)

and $P^{-1'}P^{-1} = \Omega^{-1}$ ... (5)

Now, $P^{-1}\Omega P^{-1'} = I$ --- (6).

From (5), we can write $P^{-1} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & \sqrt{\lambda_n} \end{bmatrix}$ .... (7)

This is our transformation matrix. We transform the original model $Y = X\beta + u$ by using transformation matrix P^{-1}.

$$P^{-1}Y = P^{-1}X\beta + P^{-1}u.$$

or, $Y_* = X_*\beta + u_*$ ...(8), where Y_* = transformed observation on Y
u_* = " on random disturbance ~~term~~ u.

Scanned with CamScanner

Our resultant estimator of β will be

$$b = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}y \quad \ldots\ldots\ (9)$$

and b is the blue with variance covariance matrix,

$$var(b) = \sigma^2 (X'\Omega^{-1}X)^{-1} \quad \ldots\ldots\ (10)$$

b is the OLS estimator of parameters in equ^n (8), but it is GLS estimator of parameters in (1). Now model (8) satisfies all the classical assumption. So 'b' is best, linear, unbiased estimator. This is implies the GLS estimators are blue. This is Aitken's generalisation of Gauss-Markov Theorem.

For the case of a single explanatory variable –

$$X'\Omega^{-1}X = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

$$= \begin{pmatrix} \Sigma\lambda_i & \Sigma\lambda_i x_i \\ \Sigma\lambda_i x_i & \Sigma\lambda_i x_i^2 \end{pmatrix}$$

Normal Equ^n –

$$\begin{pmatrix} \Sigma\lambda_i & \Sigma\lambda_i x_i \\ \Sigma\lambda_i x_i & \Sigma\lambda_i x_i^2 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} \Sigma\lambda_i y_i \\ \Sigma\lambda_i x_i y_i \end{pmatrix}$$

$$V(b) = \sigma^2 (X'\Omega^{-1}X)^{-1}$$

$$= \sigma^2 \frac{\begin{pmatrix} \Sigma\lambda_i x_i^2 & -\Sigma\lambda_i x_i \\ -\Sigma\lambda_i x_i & \Sigma\lambda_i \end{pmatrix}}{(\Sigma\lambda_i)(\Sigma\lambda_i x_i^2) - (\Sigma\lambda_i x_i)^2}$$

$$V(b_1) = \frac{\sigma^2 (\Sigma\lambda_i x_i^2)}{(\Sigma\lambda_i)(\Sigma\lambda_i x_i^2) - (\Sigma\lambda_i x_i)^2} \quad \ldots\ldots\ (11)$$

$$V(b_2) = \frac{\sigma^2 \Sigma\lambda_i}{(\Sigma\lambda_i)(\Sigma\lambda_i x_i^2) - (\Sigma\lambda_i x_i)^2} \quad \ldots\ldots\ (11')$$

$$Cov(b_1, b_2) = \frac{-\sigma^2 \Sigma \lambda_i x_i}{(\Sigma \lambda_i)(\Sigma \lambda_i x_i^2) - (\Sigma \lambda_i x_i)^2} \quad \text{------} \quad (12)$$

If we apply OLS method directly to $y = X\beta + u$, the estimator would be

$\hat{\beta} = (X'X)^{-1}X'y$ .... (13), where $\hat{\beta}$ is the OLS estimator of β.

$$\hat{\beta} = \beta + (X'X)^{-1}X'u$$

$$\begin{aligned} Var(\hat{\beta}) &= E(\hat{\beta} - \beta)(\hat{\beta} - \beta)' \\ &= E\left[(X'X)^{-1}X'uu'X(X'X)^{-1}\right] \\ &= (X'X)^{-1}X' E(uu') X (X'X)^{-1} \\ &= \sigma^2\left[(X'X)^{-1}X'\Omega X(X'X)^{-1}\right] \quad \text{....} \quad (14) \end{aligned}$$

$$Var(\hat{\beta}) = \frac{\sigma^2 \begin{bmatrix} n & -\Sigma x_i \\ -\Sigma x_i & \Sigma x_i^2 \end{bmatrix} \begin{bmatrix} \Sigma 1/\lambda_i & \Sigma x_i/\lambda_i \\ \Sigma x_i/\lambda_i & \Sigma x_i^2/\lambda_i \end{bmatrix} \begin{bmatrix} n & -\Sigma x_i \\ -\Sigma x_i & \Sigma x_i^2 \end{bmatrix}}{\left[n\Sigma x_i^2 - (\Sigma x_i)^2\right]^2}$$

The sampling variance of the regression slope is

$$Var(\hat{\beta}_2) = \frac{\sigma^2\left[(\Sigma x_i)^2 \Sigma\left(\frac{1}{\lambda_i}\right) - 2(\Sigma x_i)\left(\frac{\Sigma x_i}{\lambda_i}\right)(\Sigma x_i^2) + (\Sigma x_i^2)^2 \Sigma\left(\frac{x_i^2}{\lambda_i}\right)\right]}{\left[n\Sigma x_i^2 - (\Sigma x_i)^2\right]^2} \quad \text{---} \quad (15)$$

comparing (15) and (11) we conclude that it is preferable to estimate the b of (11) rather than the simple least squares of $\hat{\beta}$ since the former is a best linear unbiased estimate.

Solutions For Heteroscadastic Disturbance :-

When heteroscadasticity is established on the basis of any test, the appropriate solution is to transform the original model in such a way as to obtain a form in which the transformed disturbance term has constant variance. We then may apply the classical least squares to the transformed model.

In general, the transformation of the original model consists in dividing through the original relationship by the square root of the term which is responsible for the heteroscedasticity.

CASE - A

Let us assume that the variance of the disturbance term is proportional to the square of x, i.e.

$E(u_i^2) = \sigma^2 x_i^2$, $i = 1, 2, \ldots n$.

Here σ^2 denotes some unknown constant.

Our original model is $y_i = \beta_1 + \beta_2 x_i + u_i$, $i = 1, 2, \ldots n$.

~~(We obtain the transformed model by (divind) dividing the original relationship)~~

The appropriate transformation of the original model consists of the division of the original relationship by $\sqrt{x^2} = x$, which means that the appropriate transformation version is

$$\frac{y_i}{x_i} = \frac{\beta_1}{x_i} + \beta_2 + \frac{u_i}{x_i}$$

The new transformed random term u_i/x_i is homoscedastic since,

$$E\left(\frac{u_i}{x_i}\right)^2 = \frac{1}{x_i^2} E(u_i^2) = \frac{1}{x_i^2} \cdot \sigma^2 x_i^2 = \sigma^2$$

In original model disturbance term was heteroscedastic but in transformed model disturbance term is homoscedastic. Hence we may apply classical least squares to the transformed model.

CASE - B

Let us assume that the form of heteroscedasticity to be $E(u_i^2) = \sigma^2 x_i$; i.e. variance of the disturbance term is directly proportional to the explanatory variable. Here the appropriate transformation of the original model consists of the division of original relationship by $\sqrt{X_i}$.

$$\frac{Y_i}{\sqrt{X_i}} = \frac{\beta_1}{\sqrt{X_i}} + \frac{\beta_2 X_i}{\sqrt{X_i}} + \frac{u_i}{\sqrt{X_i}}$$

The transformed random term $u_i/\sqrt{X_i}$ is homoscadastic with constant variance equal to σ^2.

$$E\left(u_i/\sqrt{X_i}\right)^2 = \frac{1}{X_i} E(u_i^2) = \frac{1}{X_i} X_i \sigma^2 = \sigma^2 .$$

Hence with the above transformation we avoid heteroscadasticity and we may apply classical least squares to the transformed model.

CASE-C

Let us assume that the variance of disturbance term is directly proportional to $E(Y_i)$.

i.e. $E(u_i^2) = \sigma^2 E(Y_i)$

We have to find out $E(Y_i)$. Now $E(Y_i) = \beta_1 + \beta_2 X_i$.

It is not practically acceptable because β_1 and β_2 are unknown. Here we have to apply first the OLS method, since there is heteroscadasticity so that we can find out $\hat{\beta}_1$ and $\hat{\beta}_2$.

In this way we get $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$

$E(\hat{Y}_i) = \beta_1 + \beta_2 X_i$

So here we have to devide both sides by $\sqrt{E(\hat{Y}_i)}$ or simply $\sqrt{\hat{Y}_i}$

CASE-D

Let us assume the form of heteroscadasticity is

$E(u_i^2) = \sigma^2 (a_0 + a_1 X_i)^2$

The appropriate transformation implies the division of the original equⁿ by $(a_0+a_1x_i)$, i.e.

$$\frac{y_i}{a_0+a_1x_i} = \frac{\beta_1}{a_0+a_1x_i} + \beta_2\frac{x_i}{a_0+a_1x_i} + \frac{u_i}{a_0+a_1x_i}$$

The new random term is homoscadastic with constant variance equal to σ^2

$$E\left[\frac{u_i}{a_0+a_1x_i}\right]^2 = \frac{1}{(a_0+a_1x_i)^2} E(u_i^2) = \sigma^2$$

Hence with the above transformation we avoid heteroscedasticity.

How can we detect the Heteroscadasticity ?

Various tests have been suggested for detecting heteroscadasticity.

BREUSCH – PAGON TEST

This is a large sample test and it is the most generalised test of heteroscadasticity.

They start with the simple linear model – $Y = X\beta + u$ and $\sigma_i^2 = h(z_i'\alpha)$

This implies that the disturbance term is heteroscadastic.

Here z_i' = a set of explanatory variable that influence the homoscadasticity.

α = column vector of P element i.e. $\alpha = \begin{pmatrix}\alpha_1 \\ \alpha_i \\ \vdots \\ \alpha_p\end{pmatrix}_{p\times1}$

$$\sigma_i^2 = h(\alpha_1 z_1 + \alpha_2 z_2 + \cdots\cdots + \alpha_p z_p)$$

Our null hypothesis is H_0: $\alpha_2 = \alpha_3 = \cdots\cdots = \alpha_p = 0$.

which means that $\sigma_i^2 = h(\alpha_1)$ = function of a constant.

We have to test this null hypothesis against the alternative hypothesis $\sigma_i^2 \neq$ constant.

We have to carry out this test in the following way—
Given the general linear model $y = X\beta + u$

(i) We have to apply first the maximum likelihood method to this model to obtain e_i.

(ii) Then we have to calculate $\hat{\sigma}^2 = \frac{e'e}{n}$

(iii) Then we form a series $g_i = \frac{e_i^2}{\hat{\sigma}}$, $i = 1, 2, \dots n$

(iv) We regress g_i on z_i and obtain ESS

(v) We have to show that $\frac{ESS}{2} \sim \chi^2_{p-1}$

GOLDFIELD AND QUANDT TEST

This is the test which is applicable to large samples. They consider the linear model $y = X\beta + u$, where the departure from homoscedasticity takes the form $E(u_i^2) = \sigma^2 X_{ji}^2$, i.e. the disturbance variance increases with the square of one of the explanatory variables.

The hypothesis to be tested is the null hypothesis
H_0: u_i's are homoscedastic
and is tested against the alternative hypothesis
H_1: u_i's are heteroscedastic.

~~(They then proposed two alternative tests, one parametric and the other nonparametric)~~.
The steps involved in the Goldfield and Quandt test may be outlined as follows.

① We have to order the observations according to the magnitude of the explanatory variable X.

② We select arbitarily a certain number (c) of central observations which we omit from the analysis. The remaining $(n-c)$ observations are divided into two subsamples of equal size $(n-c)/2$, one including the small values of X, and the other including the large values of X.

③. We fit separate regressions by the ordinary least squares to each sub-samples and we obtain the sum of squared residuals from each of them.
Σe_1^2 = residuals from the subsamples of low values of X, with $[(n-c)/2]-K$ degrees of freedom, where K is the total number of parameters in the model.
Σe_2^2 = residuals from the subsample of high values of X, with the same degrees of freedom, $[(n-c)/2]-K$

④ Let RSS_1 denotes the sum of squared residuals from the smaller X_i values and RSS_2 denotes the sum of the squared residuals from the larger ~~val~~ x_i values. Then we have to find out $R = \frac{RSS_2}{RSS_1}$, and this on the assumption of homoscadasticity have the F distribution with $[(n-c-2K)/2, (n-c-2K)/2]$ degrees of freedom. Under the alternative hypothesis R will tend to be larger.

GLEJER TEST

It is based on the regression of absolute value of least squares errors on some explanatory variable (Z) which is thought to influence homoscadasticity.

$$|e_i| = a_0 + a_1 Z_i^h + v_i$$

Here v_i is assumed to have zero mean and constant variance and zero covariance.
Then firstly we perform the regression of Y on all the

explanatory variables and we compute the residuals, e's.

2ndly, we regress the absolute values of e's ($|e_i|$) on the explanatory variable. The actual form of this regression is usually not known, so that one may experiment with various formulation containing various powers of z_i.

For example – Let $h=1$, $z_i = x_i$

So $|e_i| = a_0 + a_1 x_i + v_i$

If, $h=-1$, $z_1 = x_i$. $|e_i| = a_0 + a_1/x_1 + v_i$

If, $h = 1/2$, $z_1 = x_i$ $|e_i| = a_0 + a_1\sqrt{x_i} + v_i$

In this way we can construct infinite number of equation depending on values of h and z. ~~(Which)~~ Which equation should we accept depends on the notion of investigator.

We have to choose the form of regression which gives the best fit in the light of the correlation coefficient and the standard errors of the coefficients a_0 and a_1. Heteroscedasticity is judged in the light of the statistical significance of a_0 and a_1. We perform any standard for these coefficients and if they are found significantly different from zero we accept that the u_i's are heteroscedastic.

$H_0: a_1 = 0$ is to be tested against the alternative hyp. $H_1: a_1 \neq 0$.

Now, if $a_1 = 0$ is accepted e is not affected by x. So there is no heteroscedasticity.

Now, (i) If $a_1 = 0$, $a_0 = 0$ this is the case of Pure Homoscedasticity

(ii) If $a_1 = 0$, $a_0 \neq 0$ this " " " " mixed "

(iii) If $a_1 \neq 0$, $a_0 \neq 0$ " " " " " " Heteroscedasticity

(iv) If $a_1 \neq 0$, $a_0 = 0$, " " " " " Pure "

This is the advantage of Glejser test that we can distinguish the case of mixed and pure homoscedasticity.