H - A - M

11. Heteroskedasticity

Unequal

Spread or variance

One of the assumptions of the classical linear regression model (Assumption 3 on p. 116) is that the disturbances u_i appearing in the population regression function are homoscedastic. In other words, they all have the same variance. In this chapter, students will examine the validity of this assumption and figure out the consequences if this assumption is not fulfilled.

11.0.1 Nature of Heteroscedasticity

As mentioned in CHAPTER 7: Multiple Regression Analysis: The Problem of Analysis. We assume that the variance of each disturbance term u_i, conditional on the chosen values of the explanatory variables, is constant. We called this as **homoscedasticity** or equal (homo) spread (scedasticity), that is equal variance.:

$$\boxed{E(u_i^2) = \sigma^2 \quad i = 1, 2, \ldots, n} \tag{11.1}$$

→ Homoscedasticity: variance of disturbance term is constant throughout at any given X_i

Depicted by Figure 11.1, the variances of disturbance terms given any independent variable are constant and equal. Specifically, conditional on X_1, the variance is equal to σ^2 which is the same as the variance of disturbance term conditional on X_2 and on the other values of independent variable.

On the other hand, if this assumption is violated, the problem occurring is called **heteroscedasticity**. That is, conditional on X_1, the variance of disturbance term is σ_1^2; whereas, conditional on X_2, the variance of disturbance term is σ_2^2. In brief, the conditional variance of disturbance term would vary across the values of independent variable, as illustrated, mathematically, by Equation 11.2 and, graphically, by Figure 11.2.

$$\boxed{E(u_i^2) = \sigma_i^2} \tag{11.2}$$

→ Heteroscedasticity: variance of disturbance term is non-constant, depending on which level of X_i you look at!

Page 2

Thursday, May 3, 2018 8:20 AM

spread

Figure 11.1: Homoscedasticity

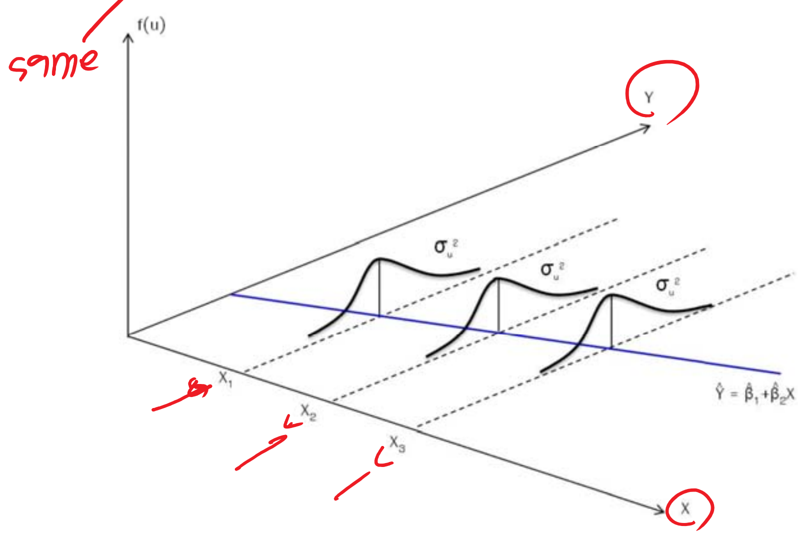
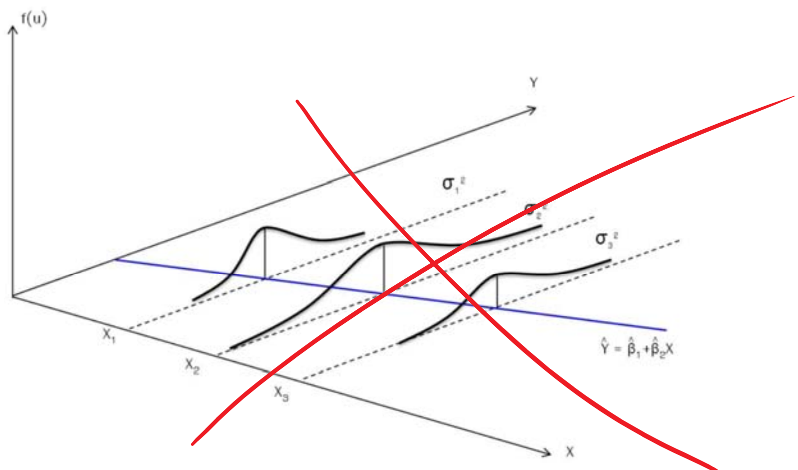
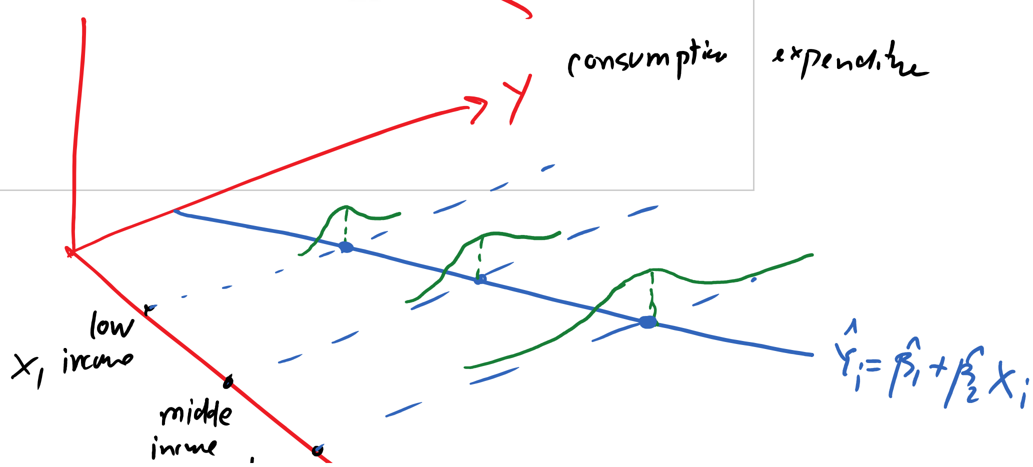
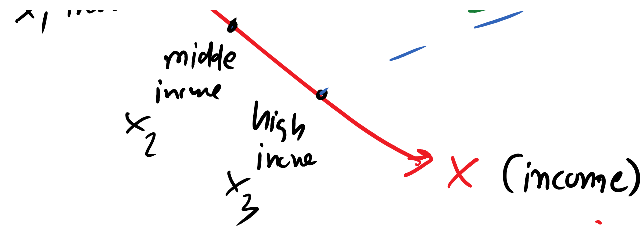


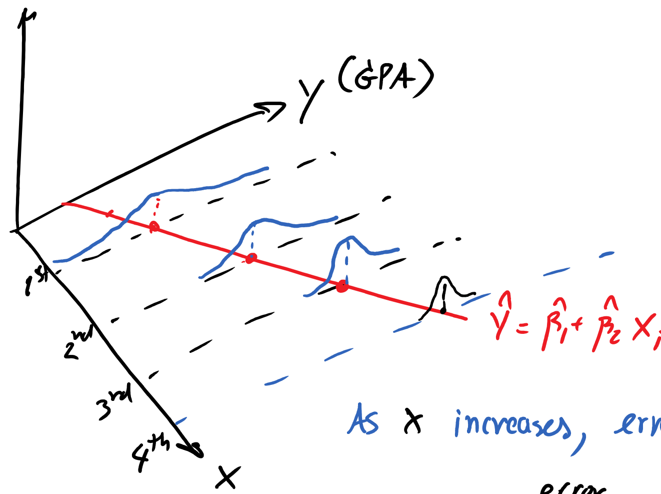
Figure 11.2: Heteroscedasticity





$Y_i = \beta_1 + \beta_2 X_i$





As X increases, error variance decreases

error

(learning model)

Generally, there is a variety of causes for the existence heteroscedasticity problem in the regression model studied by most economists. Yet, only 6 main sources are discussed here.

1. Normally, error learning is the nature of human. At the initial stage of their work, people probably commit a large number of mistakes. As they carry on working and become more specialized, the amount of errors would be reduced. In this case, it seems that the variance at the initial stage will be high but would decrease as people learn from their errors.

example: error learning model

2. Considering the relationship between independent variable X and dependent variable Y, there seems to be possible that as the value of independent variable increase, the variance of the value of dependent one will increases. The feasible reason is the characteristics of those variables such as the relationship between the profit of the company (independent variable) and dividend (dependent variable). As the profit rises, the board of director may have a variety of dividend policies. Some companies may pay a small amount of dividend in order to keep the profit for further development. Some may pay a large amount of dividend to satisfy the shareholders. On the contrary, if the profit is low, the dividend policy will not diverge across the companies since the companies seem to be at the growth stage and decide to keep their profit as retained earnings.

Dividend = f(profits)

i Div π

500 firms

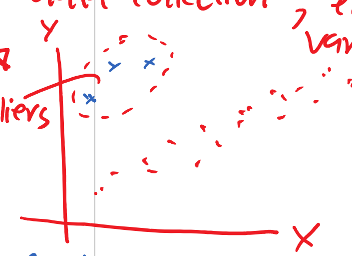
3. The collection of data is another source. As the collection technique employed by the researcher is improved, the collection error would be lower. Contrarily, with the poor collection technique, the data obtained to construct the regression model would probably incur more and more errors, causing the condition variance of disturbance term to vary.

↑ in techonology of data collection, error variance ↓

4. The existence of outliers in the independent and dependent variables may make the conditional variance of disturbance term on independent variables volatile. Mostly, if the researchers collect too small amount of data, that set of data tends to include the outliers and undermine the regression analysis. As the amount of data increases, those outliers may become normal relative to other observations.

outliers

Y



X

5. The misspecification of the model could result in the heteroscedasticity problem as well. In some cases, econometricians drop some important and necessary independent variable from the regression model. The disturbance term, then, will incorporate the characteristics of the missing variables, resulting in heteroscedasticity problem. For example, suppose the researchers want to establish the model to explain the relationship of price and quantity of good X, as suggested by the theory of demand. However, if it turns out that good Y is the substitute for good X. This mistake of failing to include price of good Y, which has the explanatory power over the demand for good X, will result in misspecification error. The disturbance term will have the characteristics of good Y, which, in turn, leads to heteroscedasticity.

IGNORE P_Y → u_i will have effect of P_Y in it.

↓

Then $\hat{u}_i$ will form a relationship with $\hat{Q}^d_X$: As $\hat{Q}^d_X$ ↑, $\hat{u}_i$ ↑

① $Q^d_X = f(\overset{-}{P_X})$ → wrong model

② $Q^d_X = f(P_X, \overset{+}{P_Y})$ → correct model

suppose X & Y are substitutes

Exclusion errors $\xrightarrow{\text{may lead to}}$ Hetero

6. Heteroscedasticity problem usually happens in the model that applies the cross-sectional data which tends to be highly diverse because the data is collected in the same time period. To illustrate, the census acquired from numerous provincial areas may cover the wide range of value and results in the stated problem. On the other hand, for the **time series data**, it is prone to be the collection of the same sample for different period of time. With the same sample, the range of the value covered seems to be narrow. Hence, the stated problem, generally, does not occur with this kind of data.

i	GPA	income	SAT	study hours	Drinking
1					
2					
3					
4					
⋮					
500					

11.0.2 OLS Estimation in the presence of Heteroscedasticity

For the estimation of regression model through OLS methods, given the assumption of heteroscedasticity by letting $E(u_i^2) = \sigma_i^2$, the model can be written as

1. NATURE
2. SOURCES
3. CONSEQUENCES
4. DETECTION
5.

t	Y	x_1	x_2	x_3
1970				
⋮				
2017				

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

The estimator β_2 can be calculated by

$$\hat{\beta}_2 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} = \frac{\sum x_i y_i}{\sum x_i^2}$$

In general, the variance of the estimator β_2 is computed by

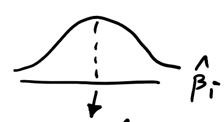
$$\boxed{Var(\hat{\beta}_2) = \frac{\sum x_i^2 \sigma_i^2}{(\sum x_i^2)^2}} \quad \text{UNDER THE PRESENCE OF HETEROSCEDASTICITY} \tag{11.3}$$

$$\boxed{var(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2}}$$ UNDER NO HETEROSCEDASTICITY (HOMOSCEDASTICITY)

2 Main points

w/ hetero..., ① OLS estimates are still UNBIASED: $E(\hat{\beta}_1) = \beta_1$

$E(\hat{\beta}_2) = \beta_2$

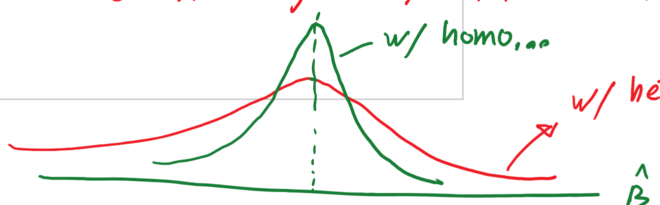


$i = 1, 2, 3, \ldots n$

② However, ... OLS estimates are no longer best

This is totally different from the usual variance formula obtained under the assumption of homoscedasticity:

or no longer w/ minimum variance :(



Page 5

Thursday, May 3, 2018 8:20 AM

$$Var(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2}$$

The $\hat{\beta}_2$ is no longer BEST and the minimum variance. It is not BLUE.

What is BLUE in the presence of Heteroscedasticity? The answer is we have to apply the method of generalized least squares (GLS) instead.

11.0.3 The Method of Generalized Least Squares (GLS)

(Why GLS?)

Via OLS, we minimize $\sum \hat{u}_i^2 = \sum (\underset{\text{Actual}}{Y_i} - \overbrace{\hat{\beta}_1 - \hat{\beta}_2 X_i}^{\hat{Y}_i \text{ predicted}})^2$

Via GLS, we minimize $\sum w_i \hat{u}_i^2 = \sum w_i (Y_i - \beta_1^* X_{0i} - \beta_2^* X_i)^2$

where $\boxed{w_i = \frac{1}{\sigma_i^2}}$

where $X_{0i} = 1$.

W/ OLS : $\hat{u}_1, \hat{u}_2, \hat{u}_3, \hat{u}_4, \ldots \hat{u}_n$

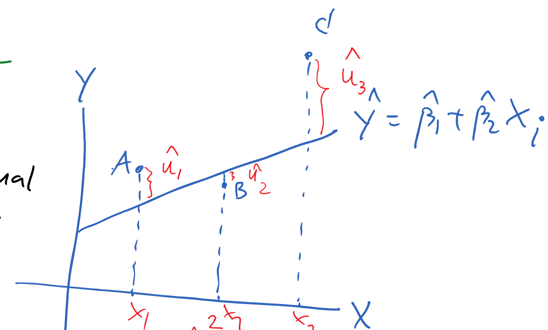
W/ GLS : $w_1\hat{u}_1 \quad w_2\hat{u}_2 \quad w_3\hat{u}_3 \quad w_4\hat{u}_4 \quad \ldots \quad w_n\hat{u}_n$

$= \frac{1}{\sigma_1^2} \quad = \frac{1}{\sigma_2^2} \quad = \frac{1}{\sigma_3^2} \quad = \frac{1}{\sigma_4^2} \quad = \frac{1}{\sigma_n^2}$

So, GLS is... A weighted sum of residual squares

when $\boxed{w_i = \frac{1}{\sigma_i^2}}$ acting as "the corresponding weights"

NOTE: Then we can say that OLS is unweighted sum of residual squares.



$\sum \hat{u}_i = 1 \cdot \hat{u}_1^2 + 1 \cdot \hat{u}_2^2 + 1 \cdot \hat{u}_3^2$

compensation to employee

(Cont.) The Method of Generalized Least Squares (GLS)

size of the firm

Consider $Y_i = \beta_1 + \beta_2 X_i + u_i$

$Y_i = \beta_1 X_{0i} + \beta_2 X_i + u_i$ where $X_{0i} = 1$ for each i.

Assume that σ_i^2 are known.

σ_i^2 = error variance

$$\left(\frac{Y_i}{\sigma_i}\right) = \beta_1 \left(\frac{X_{0i}}{\sigma_i}\right) + \beta_2 \left(\frac{X_i}{\sigma_i}\right) + \left(\frac{u_i}{\sigma_i}\right)$$

$$Y_i^* = \beta_1^* X_{0i}^* + \beta_2^* X_i^* + u_i^*$$

Then we call the above equation as "transformed OLS model"

Q: what is the benefit of doing it in this way?

In other words, why doing this helps solving hetero...?

A: Consider u_i^*, transformed disturbance term ...

$$var(u_i^*) = E(u_i^*)^2 = E\left(\frac{u_i}{\sigma_i}\right)^2$$

$$= \frac{1}{\sigma_i^2} \cdot E(u_i^2)$$

$$= \cancel{\frac{1}{\sigma_i^2}} \cdot \cancel{(\sigma_i^2)}$$ since $E(u_i^2) = \sigma_i^2$

$$= 1 \text{ !!!}$$

$var(u_i^*)$ is constant! By doing this tranformation,

$$u_i \xrightarrow{\frac{1}{\sigma_i}} u_i^*$$

hetero → no hetero (= homoscedasticity)

β_1, β_2 are not BLUE → β_1^*, β_2^* are now BLUE

So, GLS procedure is actually OLS "ON THE TRANSFORMED VARIABLES"

we call β_1^* and β_2^* as "GLS estimators"

Now, let's find recipes for β_1^* and β_2^* ...

$$Y_i - \hat{\beta}_1^* (X_{0i}) + \hat{\beta}_2^* (\quad) + (\hat{u}_i)$$

Now, let's [illegible] recipes for β_1 and β_2 ...

$$\frac{Y_i}{\sigma_i} = \hat{\beta}_1^* \left(\frac{X_{0i}}{\sigma_i}\right) + \hat{\beta}_2^* \left(\frac{X_i}{\sigma_i}\right) + \left(\frac{\hat{u}_i}{\sigma_i}\right)$$

$$Y_i^* = \hat{\beta}_1^* X_{0i}^* + \hat{\beta}_2^* X_i^* + \hat{u}_i^*$$

$$\sum \hat{u}_i^{*2} = \sum_{i=1}^{n} (\underbrace{Y_i^* - \hat{\beta}_1^* X_{0i}^* - \hat{\beta}_2^* X_i^*}_{=\hat{u}_i^*})^2$$

$$\sum \left(\frac{\hat{u}_i}{\sigma_i}\right)^2 = \sum_{i=1}^{n} \left[\left(\frac{Y_i}{\sigma_i}\right) - \hat{\beta}_1^* \left(\frac{X_{0i}}{\sigma_i}\right) - \hat{\beta}_2^* \left(\frac{X_i}{\sigma_i}\right)\right]^2$$

$$\min_{\hat{\beta}_1^*, \hat{\beta}_2^*} \sum \left(\frac{\hat{u}_i}{\sigma_i}\right)^2$$

we obtain ...

$$\boxed{\hat{\beta}_2^* = \frac{(\sum w_i)(\sum w_i X_i Y_i) - (\sum w_i X_i)(\sum w_i Y_i)}{(\sum w_i)(\sum w_i X_i^2) - (\sum w_i X_i)^2}}$$

$$\text{var}(\hat{\beta}_2^*) = \frac{\sum w_i}{(\sum w_i)(\sum w_i X_i^2) - (\sum w_i X_i)^2}$$

where $\boxed{w_i = \dfrac{1}{\sigma_i^2}}$

acting as the CORRESPONDING WEIGHTS

Page 8

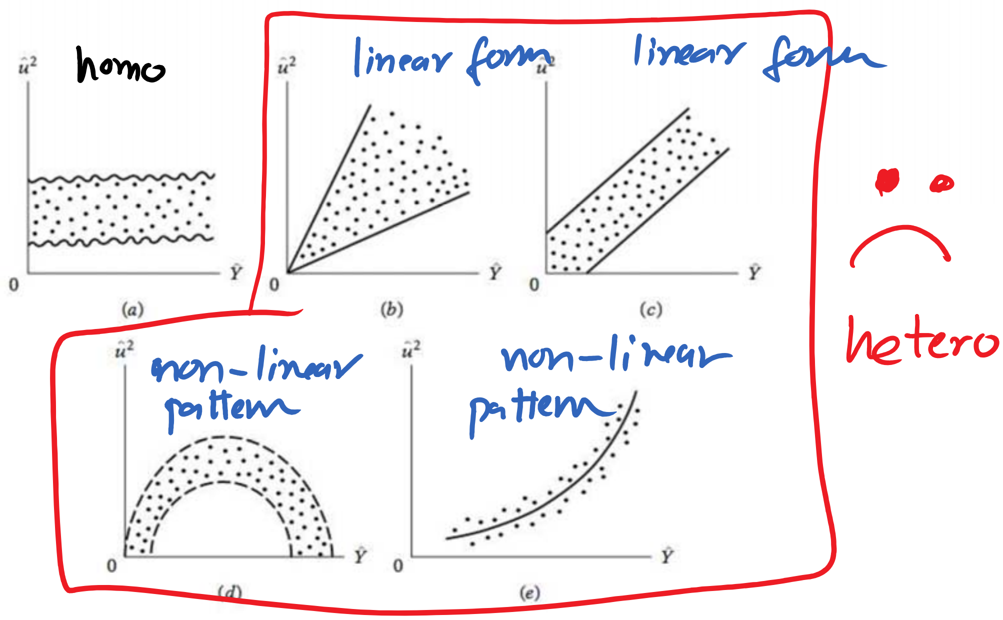
Thursday, May 3, 2018 8:20 AM

11.0.4 Detection of Heteroscedaticity

The problem of heteroscedasticity is fairly severe since it causes the estimators, which identify the relationship of regressor and regressand, to lose the minimum variance or best property despite its unbiased property. Accordingly, the statistical inference studied in the previous chapters, such as confidence interval and hypothesis test, is not applicable. It is, thus, essential to detect the problem of heteroscedasticity. Again, there are a large number of methods for detection suggested by econometricians; yet, 4 approaches are discussed here.

1. *Finding the relationship of regressor and random disturbance term by graph*: the nature of the problem is that the conditional variance of disturbance term is not constant. Hence, we can detect the problem by plotting the diagram to show the relationship between $\hat{u}_i{}^2$ and the estimated Y_i from the regression line, $\hat{Y}_i$. Also, we may plot $\hat{u}_i{}^2$ against one of the explanatory variables (X_i) to detect this problem.

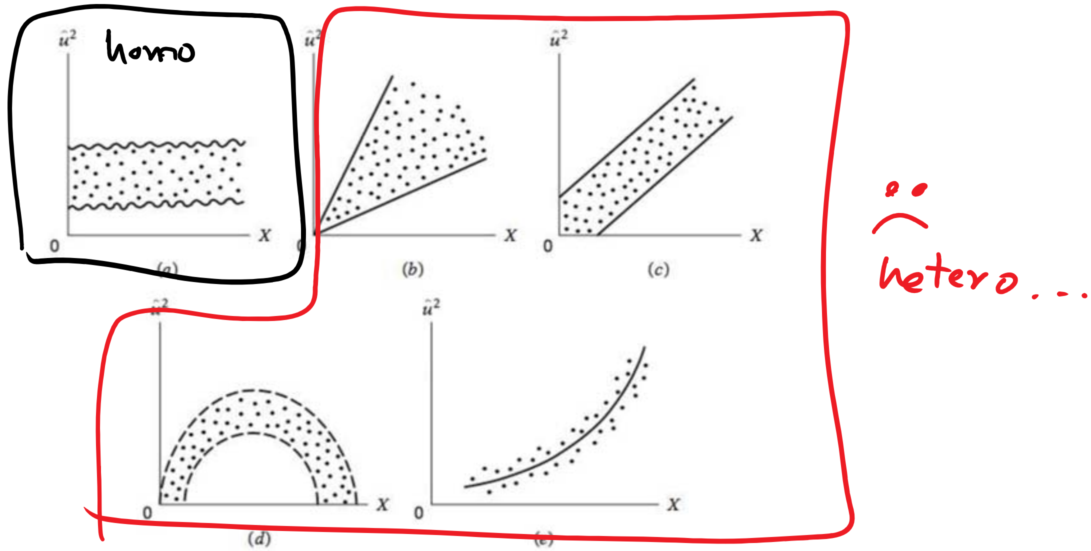
Figure 11.3: The relationship between $\hat{u}_i^2$ and $\hat{Y}_i$ of the model suffering heteroscedasticity problem



Page 9

Thursday, May 3, 2018 8:20 AM

Figure 11.4: The relationship between $\hat{u}_i^2$ and regressor X of the model suffering heteroscedasticity problem



2. **Park Test**[1]: the rationales is that, if we can construct the regression model that enables the regressor X to explains the volatility of the variance of disturbance term, that means the variance of disturbance term is not constant and depends on the regressor. The model could be constructed as Equation 11.4 and we can transform the model into the linear one as Equation 11.5

$$\sigma_i^2 = \sigma_u^2 X_i^\beta e^{v_i} \tag{11.4}$$

$$\ln \sigma_i^2 = \ln \sigma_u^2 + \beta \ln X_i + v_i \tag{11.5}$$

Practically, we would never know the true variance of the disturbance term in the model; so, we use $\hat{u}_i^2$ as the estimator and form the model similar to Equation 11.5, which is shown in Equation 11.6

$$\ln \hat{u}_i^2 = \ln \sigma_u^2 + \beta \ln X_i + v_i = \alpha + \beta \ln X_i + v_i \tag{11.6}$$

[1]Park, Rolla. E. (1966). "Estimation with Heteroscedasticity Error Terms," Econometrica, Vol.34, No.4.

Page 10

Thursday, May 3, 2018 8:20 AM

After the establishment of the model in Equation 11.6, we, then, can perform the hypothesis test to examine whether the regressor X could explain the change in the regressand $\ln \hat{u}_i^2$. In this case, we test for the statistical significance of the coefficient associated with β by finding the t-statistic. If the regressor X is statistically significantly able to describe the regressand $\ln u_i^2$, we can conclude that the model faces the problem of heteroscedasticity. In short, the procedure for Park test is as follows.

Step 1: set up the model of interest to find the relationship of regressor X and regressand Y

Step 2: calculate $\hat{u}_i^2 = (Y_i - \hat{Y}_i)^2$ from the regression model in Step 1 to be the estimator of variance in Equation 11.6

Step 3: establish the model as in Equation 11.6 and perform the hypothesis test for the relationship between the regressor and the variance of disturbance term. The hypothesis can be set as

$$H_o: \beta = 0 \quad \text{(no hetero...)}$$
$$H_a: \beta \neq 0 \quad \text{(hetero occurs...)}$$

If the null hypothesis is rejected, we can conclude that, the regressor X possess the explanatory power over the variance of disturbance term. In other word, the model suffers the heteroscedasticity prolem. Contrarily, if the null hypothesis cannot be rejected, it implies no heteroscedasticity problem in the model

3. **Breusch-Pagan Test**[2] or **LM Test**: consider the multiple regression model in Equation 11.7 and suppose that the variance of disturbance term has the linear relationship with the regressor as in Equation 11.8. To satisfy the homoscedasticity assumption such that the estimator is BLUE, all partial regression coefficients in Equation 11.8 must be zero.

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + u_i \tag{11.7}$$

$$Var(u|X_2,\ldots,X_k) = E(u^2|X_2,\ldots,X_k) = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + \cdots + \alpha_k X_{ki} \tag{11.8}$$

The procedure of Breusch-Pagan test for heteroscedasticity is as follows.

Step 1: estimate the model as in Equation 11.7 by the method of OLS and calculate the value of $\hat{u}_i^2$

Step 2: establish the regression model as in Equation 11.8 to find the coefficient of determination $R^2_{\hat{u}_i^2}$

→ AUXILIARY REGRESSION

$$\hat{u}_i^2 = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + \cdots + \alpha_k X_{ki} + u_i \tag{11.9}$$

Step 3: compute the F-statistic by Equation 11.10 and perform hypothesis test to find out whether all the regressors X's can jointly explain the variance of the disturbance term. If they have the explanatory power, we can conclude that the model in Equation 11.7 would suffer heteroscedasticity problem.

[2]Breusch, Trevor. and Pagan, Adrian. (1979). "A Simple Test for Heteroscedasticity and Random Coefficient Variation," Econometrica, Vol.47, No.5, pp.1287-94

Page 11

Thursday, May 3, 2018 8:20 AM

H_o : $\alpha_1 = \alpha_2 = \alpha_3 = ... = \alpha_k = 0$
H_a : otherwise

$$\boxed{\hat{F} = \frac{R^2_{\hat{u}_i^2}/(k-1)}{(1-R^2_{\hat{u}_i^2})/(n-k)}} \tag{11.10}$$

Furthermore, LM-statistic (**Lagrange Multiplier**) can be used to determine whether there is heteroscedasticity problem in the model and can be calculated by Equation 11.11

$$LM = nR^2_{\hat{u}_i^2} \tag{11.11}$$

The LM-statistic has the chi-square distribution with the degree of freedom $k-1$ or χ^2_{df}. We can use LM-statistic to test the following hypothesis.

H_o : Homoscedasticity
H_a : otherwise

If the LM-statistic obtained is greater than the critical value found the chi-square table, we can conclude that the model in Equation 11.7 faces the heteroscedasticity problem. Nevertheless, if the LM-statistic is less than the critical value, we can conclude that no heteroscedasticity problem exists.

4. **White Test**[3]: Consider the multiple regression model as in Equation 11.12.

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \longrightarrow \hat{u}_i^2 \tag{11.12}$$

White test has the different procedure from the third test only in the aspect of how the hypothesis is set. While Breusch-Pagan test states that the variance of disturbance term has the relationship with regressors, White test will cover the wider relationship between the variance of disturbance term and higher amount of regressors as in Equation 11.13 with the procedure as follows.

$$Var(u_i|X_2,X_3) = E(u_i^2|X_2,X_3) = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + \alpha_4 X_{2i}^2 + \alpha_5 X_{3i}^2 + \alpha_6 X_{2i}X_{3i} + v_i \tag{11.13}$$

Step 1: set up the model as in Equation 66 to obtain $\hat{u}_i^2$

Step 2: establish the regression model as Equation 11.13

$$\hat{u}_i^2 = \alpha_1 + \overbrace{\alpha_2 X_{2i} + \alpha_3 X_{3i}}^{\text{linear pattern}} + \overbrace{\alpha_4 X_{2i}^2 + \alpha_5 X_{5i}^2 + \alpha_6 X_{2i}X_{3i}}^{\text{non-linear pattern}} + v_i \tag{11.14}$$

Step 3: similar to Step-3 of Breusch-Pagan test, calculate F- or LM-statistic and set the null and alternative hypotheses. Then, compare the F- or LM-statistic with the critical value from the statistical table to test for heteroscedasticity.

STATA COMMAND FOR WHITE TEST

. estat imtest, white ↵

H_0 : homo
H_1 : hetero

[3]White, Halbert. (1980). "A Heteroscedasticity Consistent Covariance Matrix Estimator and a Direct Test of Heteroscedasticity," Econometrica, Vol.48, No.4, pp.817-38

Page 12

Thursday, May 3, 2018 8:20 AM

Through the four methods stated above, the econometricians can test whether the model that is estimated by OLS method suffers the heteroscedasticity problem. The graphical test is nothing but an initial method without any statistical test at any level of significance. The other tests involve the formulation of hypothesis and statistical test at the chosen level of significance. Hence, either Park, Breusch-Pagan or White test concerns the level of significance at 0.01, 0.05 or 0.1, depending on the situations.

11.1 Remedial Measure for Heteroscedasticity

Even though the heteroscedasticity does not make the estimators biased, the variance of the estimators obtained from the regression model will not be minimum. The loss of the best property will impair the statistical inference in which the econometricians may be interested such as the confidence interval and hypothesis test of whether there is statistically significant relationship between explanatory and explained variables.

The remedial measure for heteroscedasticity will enable the researcher to better analyze and apply the statistical tools for the study of relationship between explanatory and explained variables in the model. There are two approaches to remediation which are:

First Case where the variance of each disturbance term (σ_i^2) is known and

Second Case where the variance of each disturbance term (σ_i^2) is unknown.

First Case: When σ_i^2 is Known: The method of Weighted Least Squares

If σ_i^2 is known, the most straightforward method of correcting heteroscedasticity is to apply the generalized least squares (GLS) as we learn in section 10.3. The estimators we obtained are BLUE.

GLS ESTIMATOR

$$\hat{\beta}_2^* = \frac{(\sum w_i)(\sum w_i X_i Y_i) - (\sum w_i X_i)(\sum w_i Y_i)}{(\sum w_i)(\sum w_i X_i^2) - (\sum w_i X_i)^2} \tag{11.15}$$

$$\boxed{w_i = \frac{1}{\sigma_i^2}}$$

The variance of above estimator can be computed by :

$$Var(\hat{\beta}_2^*) = \frac{\sum w_i}{(\sum w_i)(\sum w_i X_i^2) - (\sum w_i X_i)^2} \tag{11.16}$$

where $w_i = \frac{1}{\sigma_i^2}$

Page 13

Thursday, May 3, 2018 8:20 AM

The difference between OLS and GLS is that, for GLS, the estimators obtained will minimize the weighted sum of residual squared or $\sum w_i\hat{u}_i^2$. On the other hand, for OLS, the ones obtained will minimize the sum of residual squared $\sum\hat{u}_i^2$.

For OLS

$$\sum\hat{u}_i^2 = \sum(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$$

For GLS

$$\sum w_i\hat{u}_i^2 = \sum w_i(Y_i - \hat{\beta}_1^* - \hat{\beta}_2^* X_i)^2$$

Another difference is that GLS will assign different weights to each observation of disturbance term ($\hat{u}_i^2$) according to its importance.

To be specific, if the error associated with the observation is large (that is, the variance is high), the value of the estimator in the model will greatly deviate from the value of true parameter. That observation should not be much of interest. Due to GLS method, the weight assigned will be inversely proportional to the variance of observation. Thus, that observation will be assigned a small weight of ($\frac{1}{\sigma_i^2}$). For those from a population with smaller σ_i will get proportionately larger weight in minimizing the $\sum w_i\hat{u}_i^2$.

In brief, the larger weight will be assigned to the observations that concentrate in their mean (namely, lower variance); whereas, smaller weight will be assigned to the observations that deviate from their mean (namely, higher variance).

Generally, for estimation[4] it is desirable to establish the population regression model that describes the true relationship between explanatory and explained variables. Paying more attention to the observation clustering around its (population) mean is preferable to the observation diffusing from its mean. The practice of weight assignment of GLS is a special case of least square which is also known as **weighted least squares: WLS**. Contrarily, OLS assigns the same weight for all observations.

[4] In principle, the variance of each random disturbance term is known when we have an access to the whole population data. In that case, to reflect the true relationship on average, the establishment of the population regression model should focus more on the observation close to its mean than the one far from its mean.

Page 14

Thursday, May 3, 2018 8:20 AM

11.1.1 Second Case: When σ_i^2 is not known

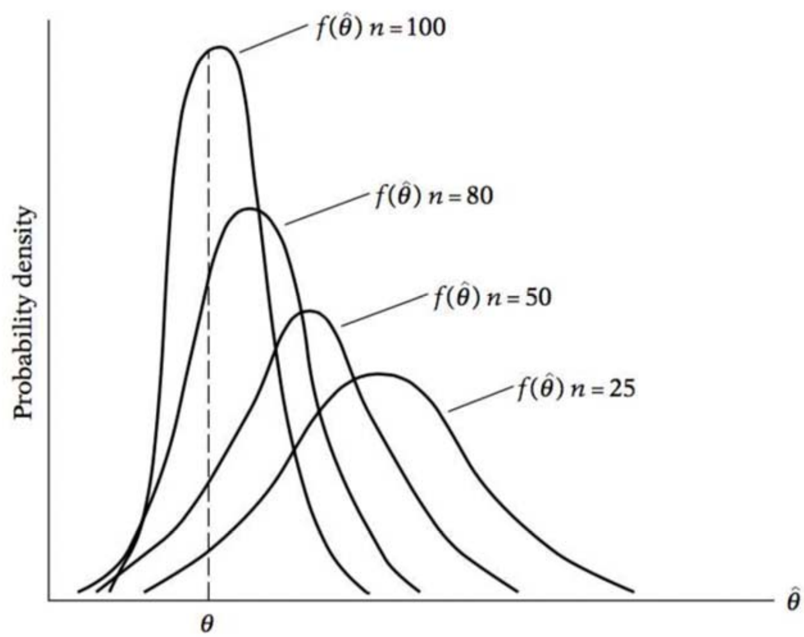
As noted earlier, if true σ_i^2 are known, we can use the WLS method to obtain BLUE estimators. However the true σ_i^2 are rarely known.

In this situation, we can apply the following methods to obtain the consistent (in the statistical sense) estimates of the variances and covariances of OLS estimators.

Consistency

θ is said to be a consistent estimator if it approaches the true value θ as the sample size gets larger and larger. Figure 11.5 illustrates this property. In this figure we have the distribution of θ based on sample sizes of 25, 50, 80, and 100. As the figure shows, θ based on n = 25 is biased since its sampling distribution is not centered on the true θ. But as n increases, the distribution of $\hat{\theta}$ not only tends to be more closely centered on θ (i.e., $\hat{\theta}$ becomes less biased) but its variance also becomes smaller. If in the limit (i.e., when n increases indefinitely) the distribution of $\hat{\theta}$ collapses to the single point θ, that is, if the distribution of $\hat{\theta}$ has zero spread, or variance, we say that $\hat{\theta}$ is a consistent estimator of θ.

Figure 11.5 : The Distribution of $\hat{\theta}$ as sample size increase



1 The distribution of $\hat{\theta}$ as sample size increases.

Page 15

Thursday, May 3, 2018 8:20 AM

There are two methods to obtain consistent (in the statistical sense) estimates of the variances and covariances of OLS estimators:

1. *White's heteroscedasticity-consistent standard errors*: by the method of White to estimate unknown standard deviation, in principle, if the sample size is large enough, this estimator can be used to represent the true standard deviation. Moreover, econometricians can apply this estimator to further statistical test as if there is no heteroscedasticity problem. Nevertheless, if the sample size is not large enough, the estimators through White method will not have t-distribution and result in false statistical conclusion.

After all, it should be aware that if the model dose not suffer heteroscedasticity problem but econometricians still use White's estimator, the conclusion from the statistical analysis will be erroneous. Accordingly, the formal test for heteroscedasticity in the model should be conducted such that the existence of heteroscedasticity is verified.

it is also called robust standard error method

reg Y X, robust

2. *Some assumptions for the distribution of random disturbance term*: consider simple regression model

$$\boxed{Y_i = \beta_1 + \beta_2 X_i + u_i}$$

$\sigma_i^2 \longrightarrow$ X's & σ_i^2

$\hat{Y}$ & σ_i^2

We can set some assumptions for the distribution of random disturbance term as follows.

Page 16

Thursday, May 3, 2018 8:20 AM

Assumption 1: let the variance of random disturbance term be proportional to X_i^2

$$E(u_i^2) = \sigma^2 X_i^2$$

With this assumption, it can be found that the variance of disturbance term of the regression model will be constant and can be shown by

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

dependent variable

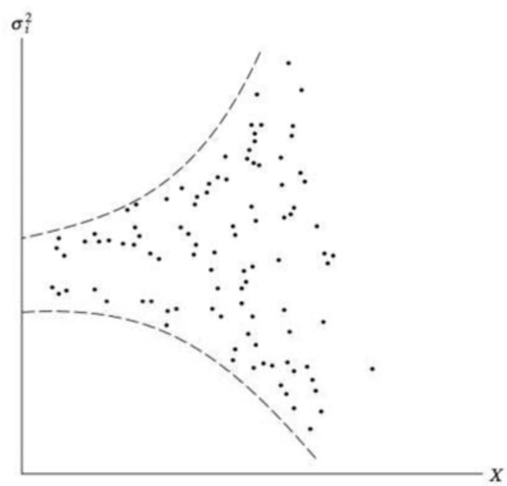
$$= \beta_1 \cdot \frac{1}{X_i} + \beta_2 + v_i$$

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

explanatory variable.

b/c $E(u_i^2) = \sigma^2 X_i^2$

Figure 11.6 : Error variance proportional to X^2



Error variance proportional to X^2.

Page 17

Thursday, May 3, 2018 8:20 AM

Assumption 2: let the variance of random disturbance term be proportional to X_i

$$\boxed{E(u_i^2) = \sigma^2 X_i}$$

With this assumption, considering the variance of disturbance term of the regression model, it can be found that the variance will be constant.

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

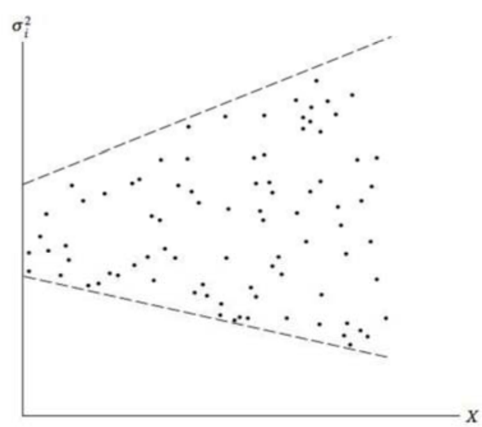
$$\frac{Y_i}{\sqrt{X_i}} = \beta_1 \frac{1}{\sqrt{X_i}} + \beta_2 \sqrt{X_i} + v_i \quad \text{where } v_i = \frac{u_i}{\sqrt{X_i}}$$

NOTICE THAT THE TRANSFORMED REGRESSION HAS NO INTERCEPT.

We can get back to the original model: $Y_i = \beta_1 + \beta_2 X_i$ by multiplying the transformed model with $\sqrt{X_i}$.

Figure 11.7: Error variance proportional to X

Check that $E(v_i^2) = \sigma^2$. D-I-Y



Error variance proportional to X.

Assumption 3: let the variance of random disturbance term be proportional to the mean of explanatory variable squared or $[E(Y_i)]^2$.

$$E(u_i^2) = \sigma^2[E(Y_i)]^2$$

In this case, $E(Y_i) = \beta_1 + \beta_2 X_i$.

Transformation will be as follows:

$$\frac{Y_i}{E(Y_i)} = \frac{\beta_1}{E(Y_i)} + \frac{\beta_2 X_i}{E(Y_i)} + \frac{u_i}{E(Y_i)}$$

$$= \beta_1 \frac{1}{E(Y_i)} + \beta_2 \left[\frac{X_i}{E(Y_i)}\right] + v_i$$

where $\boxed{v_i = \frac{u_i}{E(Y_i)}}$

Since we do not know $E(Y_i)$, then we use $\hat{Y}_i$ to replace it.

① Run OLS to obtain $\hat{Y}_i$,

② we substitute into

$$\frac{Y_i}{\hat{Y}_i} = \beta_1\left(\frac{1}{\hat{Y}_i}\right) + \beta_2\left(\frac{X_i}{\hat{Y}_i}\right) + \frac{u_i}{\hat{Y}_i}$$

This method of using $\hat{Y}_i$ for $E(Y_i)$ is ok when n is reasonably large.

Take this note.

Before the assumption about disturbance term is made, we need to identify whether the disturbance term has the relationship with other variables as in the assumption going to be made. To illustrate, the diagram depicting the relationship between the disturbance term and explanatory variable squared may be constructed to examine the validity of Assumption 1. After we explore that relationship, we, then, set the corresponding assumption to solve heteroscedasticity problem.