HETEROSCHEDASTICITY

Chapter four

Continued.....

HETEROSKEDASTICITY

what happens if the error variance is non constant?

- One of the important assumptions of the classical linear regression model is that the *variance of each disturbance term* u_i, conditional on the chosen values of the explanatory variables, ***is some constant number equal to σ².*** This is the assumption of **homoscedasticity**, or ***equal* (homo) *spread* (scedasticity), that is, *equal variance*.** Symbolically,

- $E(u^2_i) = \sigma^2 \qquad i = 1, 2, \ldots, n \qquad (11.1.1)$
- Look at Figure 11.1. In contrast, consider Figure 11.2, the variances of Y_i are not the same. Hence, there is heteroscedasticity.
- Symbolically, $Eu^2_i = \sigma^2_i$ (11.1.2)

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

Homoskedasticity: $\mathrm{Var}(U_i) = \sigma^2$

Or $E(U_i^2) = \sigma^2$

Heteroskedasticity: $\mathrm{Var}(U_i) = \sigma_i^2$

Or $E(U_i^2) = \sigma_i^2$

4.2 HETEROSKEDASTICITY

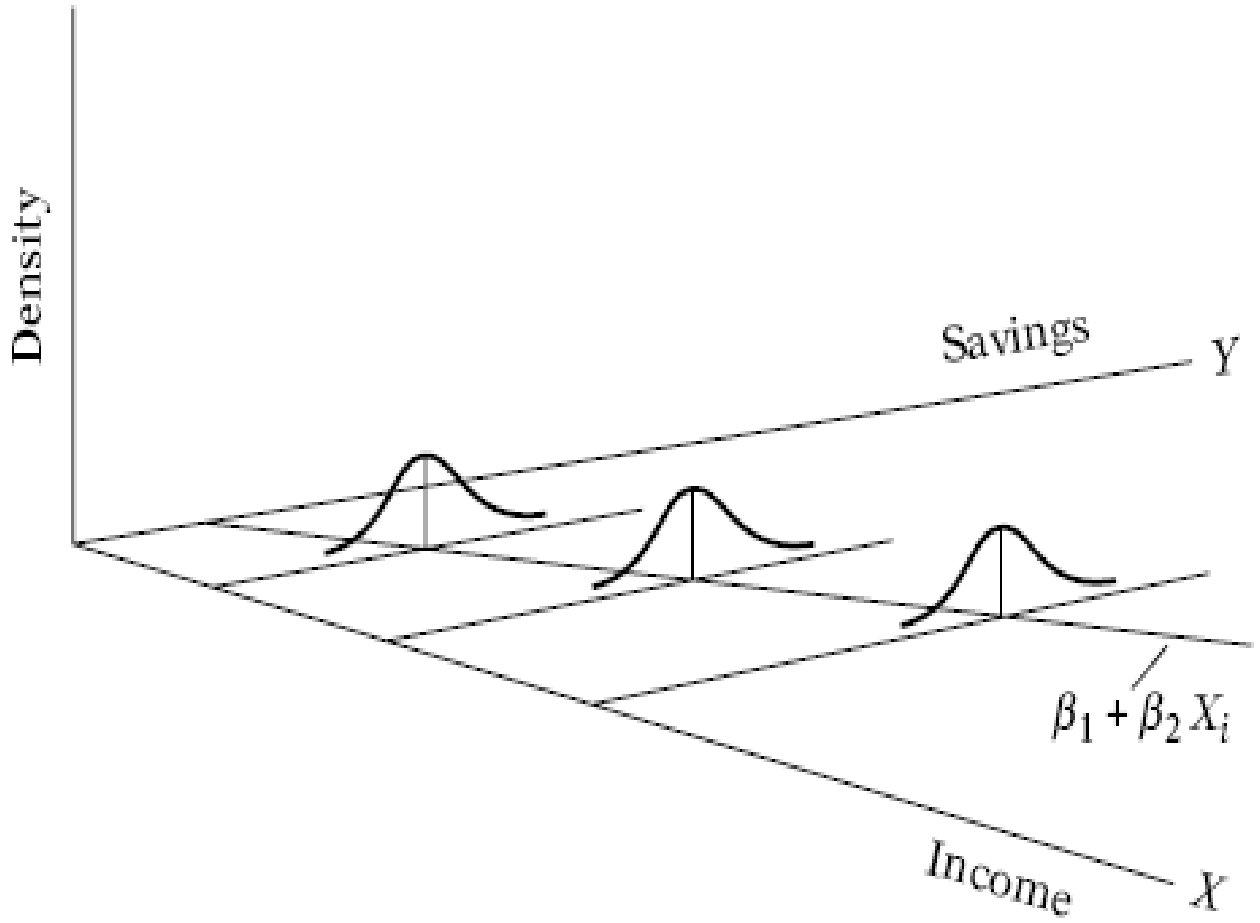


FIGURE 11.1 Homoscedastic disturbances.

4.2 HETEROSKEDASTICITY

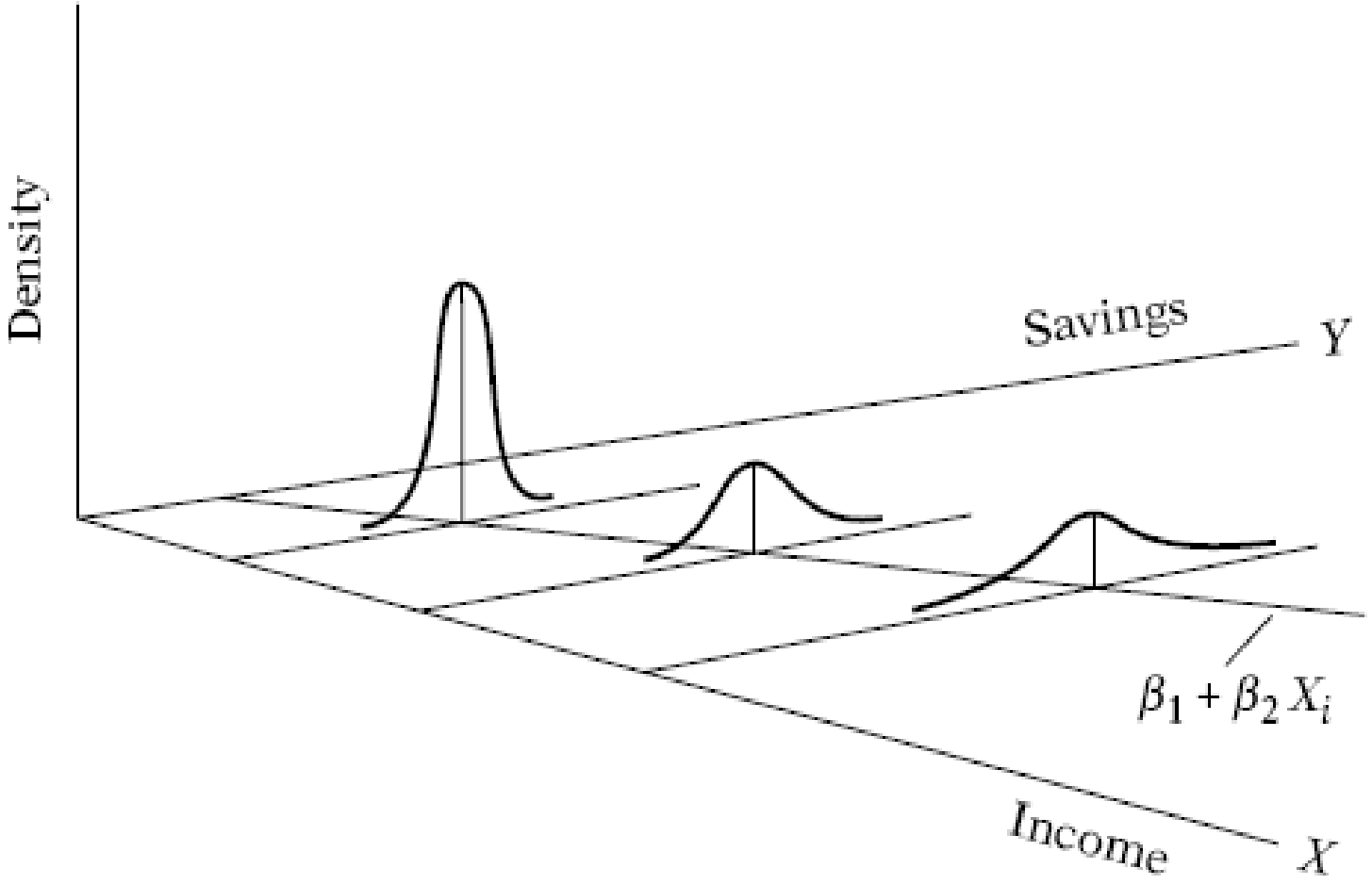


FIGURE 11.2 Heteroscedastic disturbances.

4.2 HETEROSKEDASTICITY

- Assume that in the two-variable model $Y_i = \beta_1 + \beta_2 X_i + u_i$, Y represents savings and X represents income. Figures 11.1 and 11.2 show that as income increases, savings on the average also increase. But in Figure 11.1 the variance of savings remains the same at all levels of income, whereas in Figure 11.2 it increases with income. It seems that in Figure 11.2 the *higher income families on the average save more than the lower-income families*, but there is also *more variability* in their savings.
- There are several reasons why the variances of u_i may be variable, some of which are as follows.

1. Following the *error-learning models*, as people learn, their errors of behavior become smaller over time. In this case, σ_i^2 is expected to decrease. As

4.2 HETEROSKEDASTICITY

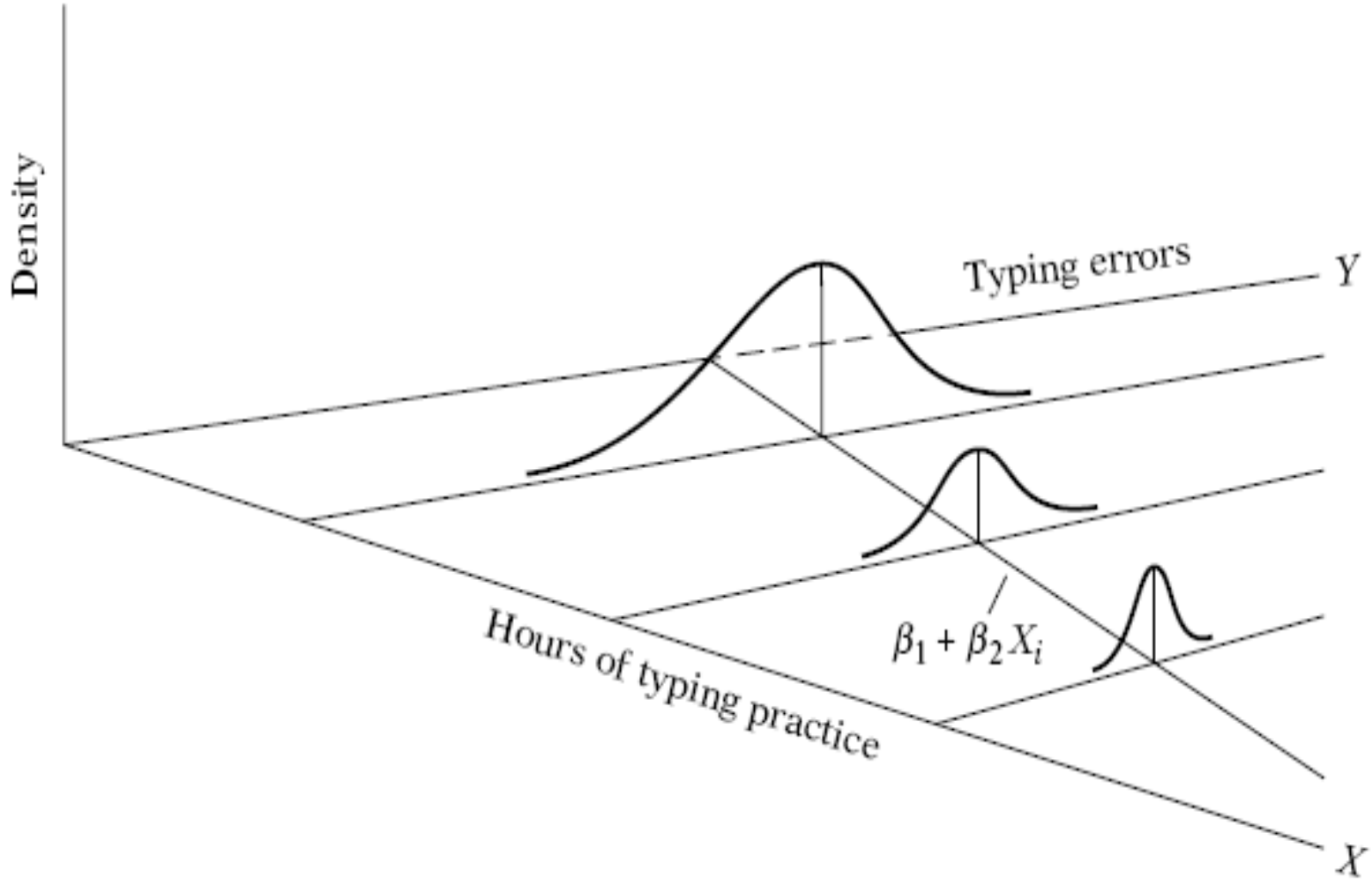


FIGURE 11.3 Illustration of heteroscedasticity.

4.2 HETEROSKEDASTICITY

2. As incomes grow, people have more *discretionary income* and hence more scope for choice about the disposition of their income. Hence, σ^2_i is likely to increase with income. Similarly, companies with larger profits are generally expected to show greater variability in their dividend policies than companies with lower profits.

3. As data collecting techniques improve, σ^2_i is likely to decrease. Thus, banks that have sophisticated data processing equipment are likely to commit *fewer errors* in the monthly or quarterly statements of their customers than

4.2 HETEROSKEDASTICITY

4. Heteroscedasticity can also arise as a result of the presence of *outliers*, (either very small or very large) in relation to the observations in the sample Figure 11.4. The inclusion or exclusion of such an observation, especially if the sample size is small, can substantially alter the results of regression analysis. Chile can be regarded as an outlier because the given Y and X values are much larger than for

4.2 HETEROSKEDASTICITY

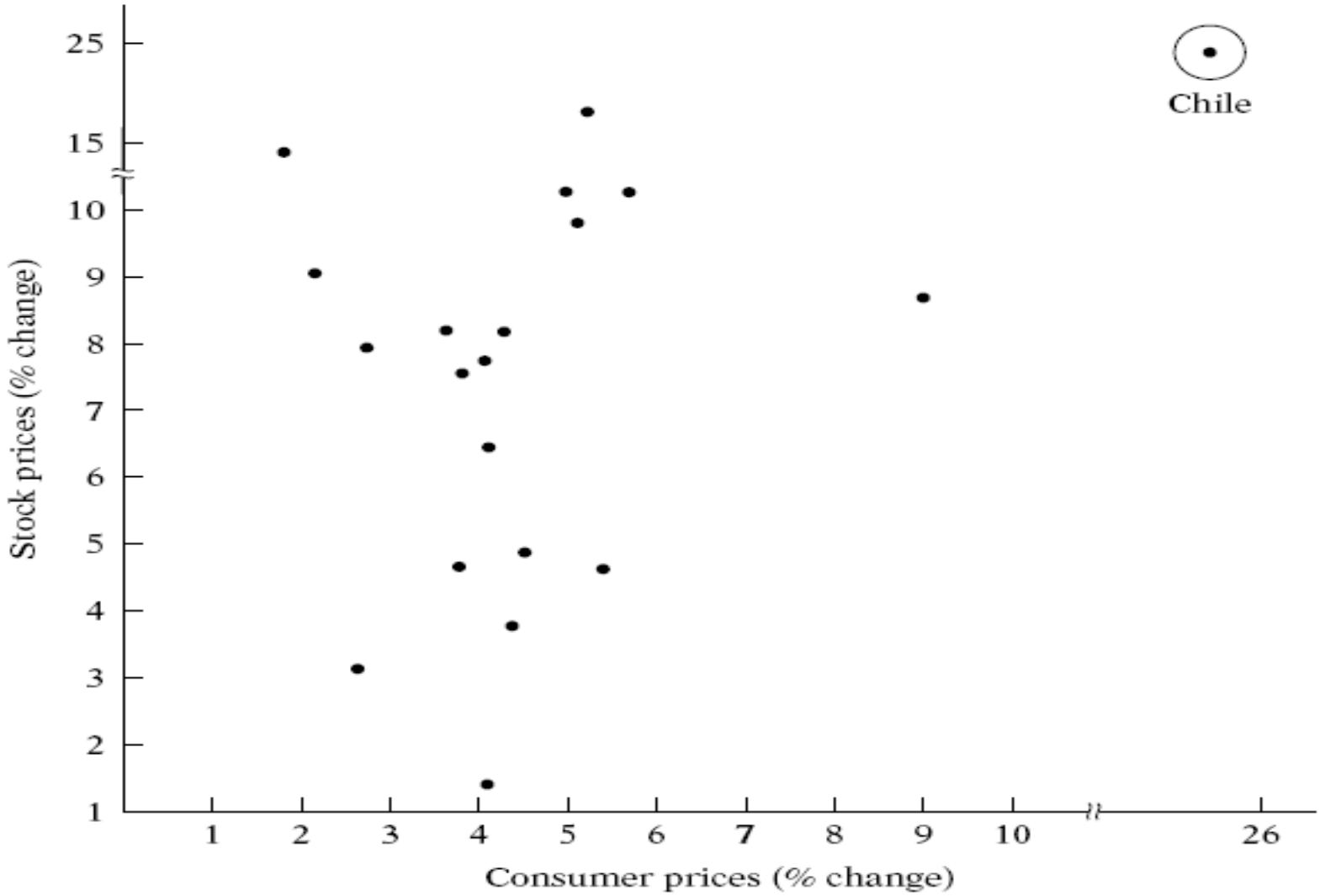


FIGURE 11.4 The relationship between stock prices and consumer prices.

4.2 HETEROSKEDASTICITY

5. Another source of heteroscedasticity arises from violating Assumption 9 of CLRM, namely, that the regression model *is correctly specified*, very often what looks like heteroscedasticity may be due to the fact that some important variables are omitted from the model. But if the omitted variables are included in the model, that impression may disappear.

4.2 HETEROSKEDASTICITY

6. **Another source of heteroscedasticity is *skewness* in the distribution of one or more regressors included in the model.** Examples are economic variables such as income, wealth, and education. It is well known that the distribution of income and wealth in most societies is *uneven*, with the bulk of the income and wealth being owned by a few at the top.
7. **Other sources of heteroscedasticity: As David Hendry notes, heteroscedasticity can also arise because of**
   (1) incorrect data transformation (e.g., ratio or first difference transformations)
   (2) incorrect functional form (e.g., linear versus log–linear models).

4.2 HETEROSKEDASTICITY

Note that the problem of heteroscedasticity is likely to be more common in cross-sectional than in time series data.

In cross-sectional data, members may be of different sizes, such as small, medium, or large firms or low, medium, or high income.

In time series data, on the other hand, the variables tend to be of similar

OLS ESTIMATION IN THE PRESENCE OF HETEROSCEDASTICITY

What happens to OLS estimators and their variances if we introduce heteroscedasticity by letting $E(u_i^2) = \sigma_i^2$ but retain all other assumptions of the classical model? To answer this question, let us revert to the two-variable model:

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

Ap
e

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2}$$

Is

$$= \frac{n\sum X_i Y_i - \sum X_i \sum Y_i}{n\sum X_i^2 - (\sum X_i)^2} \qquad \textbf{(11.2.1)}$$

4.2 HETEROSKEDASTICITY

but its variance is now given by the following

$$\operatorname{var}(\hat{\beta}_2) = \frac{\sum x_i^2 \sigma_i^2}{(\sum x_i^2)^2}$$

which is obviously different from the usual variance formula obtained under the assumption of homosce mely,

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

Of course, if $\sigma_i^2 = \sigma^2$ for each i, the two formulas will be identical.

Consequences of Heteroscedasticity

1. Ordinary least squares estimators are still **linear** and **unbiased**.
2. Ordinary least squares estimators **not efficient**.
3. Usual formulas give **incorrect** standard errors for least squares.
4. Confidence intervals and hypothesis tests based on usual standard errors are **wrong**.

THE METHOD OF GENERALIZED LEAST SQUARES (GLS)

- *If U_i is hetroschedastic $\hat{\beta}^2$ is no longer best, although it is still unbiased.*
- Consider the model: $\boldsymbol{Y_i = \beta_1 + \beta_2 X_i + U_i}$

which for ease of algebraic manipulation we write as

$$\boldsymbol{Y_i = \beta_1 X_{0i} + \beta_2 X_i + u_i} \quad \textbf{(11.3.2)}$$

wh

$$\frac{Y_i}{\sigma_i} = \beta_1 \left(\frac{X_{0i}}{\sigma_i}\right) + \beta_2 \left(\frac{X_i}{\sigma_i}\right) + \left(\frac{u_i}{\sigma_i}\right) \qquad (11.3.3)$$

Divide through by $\boldsymbol{\sigma_i}$ to obtain:

4.2 HETEROSKEDASTICITY

which for ease of exposition we write as

$$Y^*_i = \beta^*_1 X^*_{0i} + \beta^*_2 X^*_i + U^*_i$$

where the starred, or transformed, variables are the original variables divided by (the known)σi.

We use the notation β^*_1 *and* β^*_2*, the parameters of the transformed model, to distinguish them from the usual OLS parameters β1 and β2.*

4.2 HETEROSKEDASTICITY

What is the purpose of transforming the original model? To see this, notice the following feature of the transformed

$$\begin{aligned}
\operatorname{var}(u_i^*) = E(u_i^*)^2 &= E\left(\frac{u_i}{\sigma_i}\right)^2 \\
&= \frac{1}{\sigma_i^2} E(u_i^2) \qquad \text{since } \sigma_i^2 \text{ is known} \qquad \textbf{(11.3.5)} \\
&= \frac{1}{\sigma_i^2}(\sigma_i^2) \qquad \text{since } E(u_i^2) = \sigma_i^2 \\
&= 1
\end{aligned}$$

which is a constant. **That is, *the variance of the transformed disturbance***

4.2 HETEROSKEDASTICITY

The finding that it is u^* that is homoscedastic suggests that if we apply OLS to the transformed model (11.3.3) it will produce estimators that are ***BLUE.***

This procedure of transforming the original variables in such a way that the transformed variables satisfy the assumptions of the classical model and then applying OLS to them is known as ***the method of generalized least squares (GLS).***

In short, ***GLS*** *is* ***OLS on the transformed variables that satisfy the standard least-***

4.2 HETEROSKEDASTICITY

The actual ***mechanics*** of estimating β^*_1 and β^*_2 are as follows. First, we write down the ***SRF*** of (11.3.3)

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

$$\boldsymbol{Y^*_i = \hat{\beta}^*_1 X^*_{0i} + \hat{\beta}^*_2 X^*_i + \hat{u}^*_i}$$

(11.3.6)

Now, to obtain the GLS estimators, we minimize

$$\sum \hat{u}_i^2 = \sum \left[\left(\frac{Y_i}{\sigma_i}\right) - \hat{\beta}^*_1 \left(\frac{X_{0i}}{\sigma_i}\right) - \hat{\beta}^*_2 \left(\frac{X_i}{\sigma_i}\right) \right]^2 \qquad (11.3.7)$$

that

4.2 HETEROSKEDASTICITY

The actual mechanics of minimizing (11.3.7) follow the standard calculus techniques the GLS estimator of β^*_2 is

$$\hat{\beta}_2^* = \frac{(\sum w_i)(\sum w_i X_i Y_i) - (\sum w_i X_i)(\sum w_i Y_i)}{(\sum w_i)(\sum w_i X_i^2) - (\sum w_i X_i)^2} \qquad (11.3.8)$$

And its variance is given

$$\text{var}\,(\hat{\beta}_2^*) = \frac{\sum w_i}{(\sum w_i)(\sum w_i X_i^2) - (\sum w_i X_i)^2} \qquad (11.3.9)$$

where $w_i = 1/\sigma^2_i$.

4.2 HETEROSKEDASTICITY

The Difference between OLS and GLS

Recall from Chapter 3 that in OLS we minimize:

- $\Sigma \hat{u}^2_i = \Sigma(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$ **(11.3.10)**

but in GLS we minimize the expression (11.3.7), which can also be written as

$$\Sigma w_i \hat{u}^2_i = \Sigma w_i (Y_i - \hat{\beta}^*_1 X_{0i} - \hat{\beta}^*_2 X_i)^2 \quad \textbf{(11.3.11)}$$

where $w_i = 1/\sigma^2_i$

- **Thus, *in GLS we minimize a weighted sum of residual squares* with $w_i = 1/\sigma^2_i$ acting as the weights, *but in OLS we minimize an***

4.2 HETEROSKEDASTICITY

- As (11.3.7) shows, ***in GLS the weight assigned to each observation is inversely proportional to its*** σ_i , that is, observations coming from a population with **larger** σ_i will get relatively **smaller** weight and those from a population with **smaller** σ_i will get proportionately **larger weight** in minimizing the RSS (11.3.11).
- To see the difference between OLS and GLS clearly, consider the hypothetical

4.2 HETEROSKEDASTICITY

- In the (unweighted) OLS, each $\hat{u}_i^2$ associated with points *A*, *B*, and *C* will receive the same weight in minimizing the RSS. Obviously, in this case the $\hat{u}_i^2$ associated with point *C* will dominate the RSS. But in GLS the extreme observation *C* will get relatively smaller weight than the other two observations.
- Since (11.3.11) minimizes a weighted RSS, it is appropriately known as weighted least squares (WLS), and the estimators thus obtained and given in (11.3.8) and (11.3.9) are known as WLS estimators.
- But WLS is just a special case of the more general estimating technique, GLS.
- **In the context of heteroscedasticity, one can treat**

4.2 HETEROSKEDASTICITY

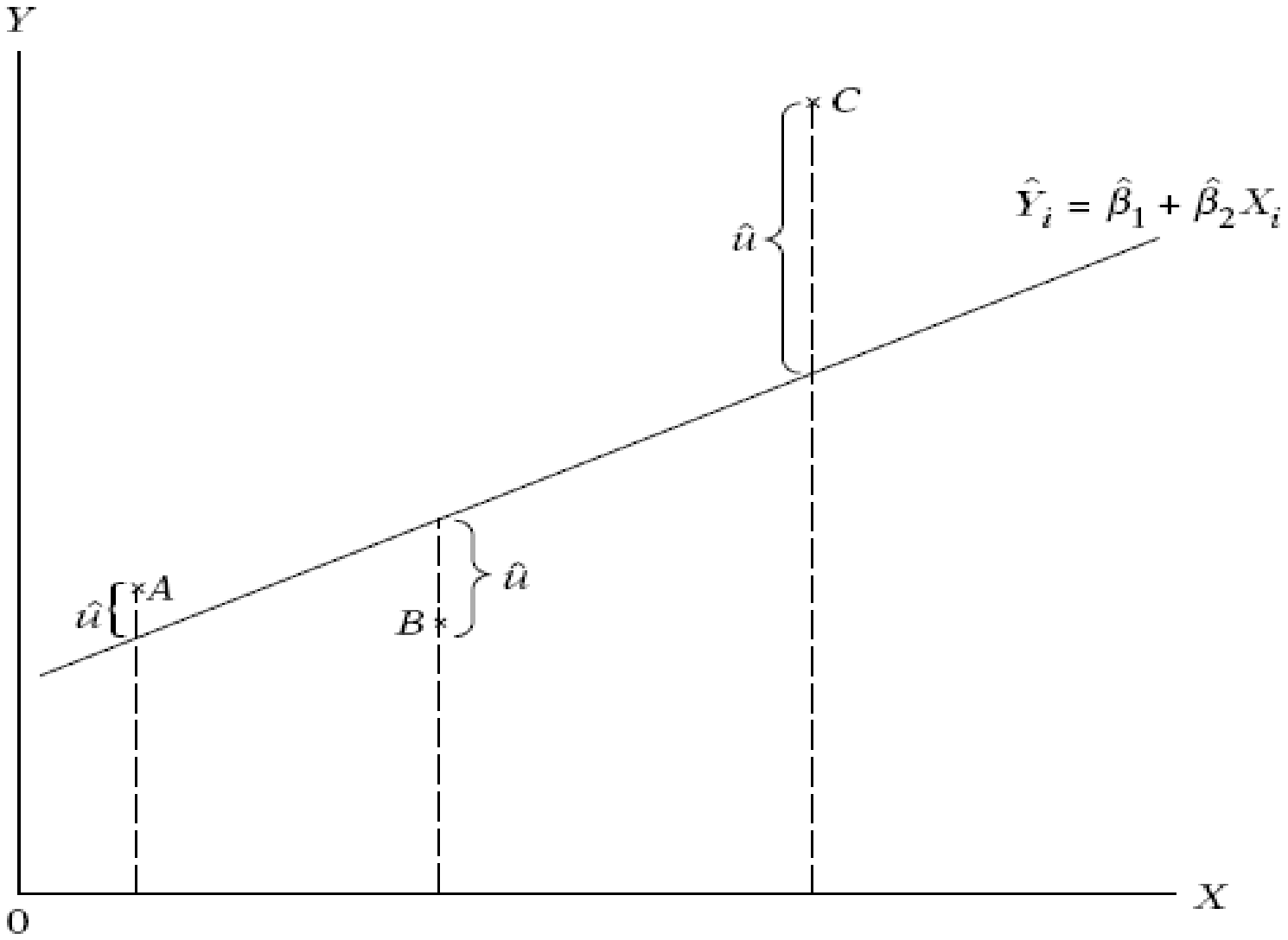


FIGURE 11.7 Hypothetical scattergram.

4.2 HETEROSKEDASTICITY

- As we have seen, both $\hat{\beta}^*_2$ and $\hat{\beta}_2$ are (linear) unbiased estimators: In repeated sampling, on the average, $\boldsymbol{\hat{\beta}^*_2}$ and $\boldsymbol{\hat{\beta}_2}$ will equal the true $\boldsymbol{\beta_2}$**;** that is, they are both unbiased estimators. But we know that it is $\boldsymbol{\hat{\beta}^*_2}$ that is efficient, that is, has the smallest variance. What happens to our confidence interval, hypotheses testing, and other procedures if we continue to use the OLS estimator $\boldsymbol{\hat{\beta}_2}$**?** We distinguish two cases.

OLS Estimation Allowing for Heteroscedasticity

- Suppose we use $\boldsymbol{\hat{\beta}_2}$ and use the variance formula given in (11.2.2), which takes into account heteroscedasticity explicitly. Using this variance, and assuming σ^2_i are known, can we establish confidence intervals and test hypotheses with the usual t and F

4.2 HETEROSKEDASTICITY

OLS Estimation Disregarding Heteroscedasticity

- The situation can become serious if we not only use $\hat{\beta}_2$ but also continue to use the usual (homoscedastic) variance formula given in (11.2.3) even if heteroscedasticity is present or suspected, whatever conclusions we draw or inferences we make may ***be very misleading.***

DETECTION OF HETEROSCEDASTICITY

As with multicollinearity, the important practical question is: How does one know that heteroscedasticity is present in a specific situation?

- Some of the informal and formal methods are used for detecting heteroscedasticity. Most of these methods are based on the ***examination of the OLS residuals*** $\hat{U}_i$ since they are the ones we observe, and ***not the disturbances*** U_i.
- One hopes that they are good estimates of U_i, a hope that may be fulfilled if the sample size is fairly large.

4.2 HETEROSKEDASTICITY

Informal Methods

- **Nature of the Problem.** Very often the nature of the problem under consideration suggests whether heteroscedasticity is likely to be encountered.
- In cross-sectional data involving ***heterogeneous*** units, heteroscedasticity may be the rule rather than the exception.
- Thus, in a cross-sectional analysis involving the investment expenditure in relation to sales, rate of interest, etc., heteroscedasticity is generally expected if small-, medium-, and

4.2 HETEROSKEDASTICITY

Graphical Method

- If there is no a priori or empirical information about the nature of heteroscedasticity, in practice one can do the regression analysis on the assumption that there is no heteroscedasticity and then do an examination of the residual squared $u\hat{}^2_i$ to see if they exhibit any systematic pattern.
- Although $u\hat{}^2_i$ are not the same thing as u^2_i , ***they can be used as proxies*** especially if the sample size is sufficiently large. An examination of the $u\hat{}^2_i$ may reveal patterns

4.2 HETEROSKEDASTICITY

see that there is no systematic pattern between the two variables, suggesting that perhaps no heteroscedasticity is present in the data. Figure 11.8*b* to *e*, however, exhibits definite patterns. For instance, Figure 11.8*c* suggests a linear relationship, whereas Figure 11.8*d* and *e* indicates a quadratic relationship between $\hat{u}^2_i$ and $\hat{Y}_i$. Using such knowledge, albeit informal, one may

4.2 HETEROSKEDASTICITY

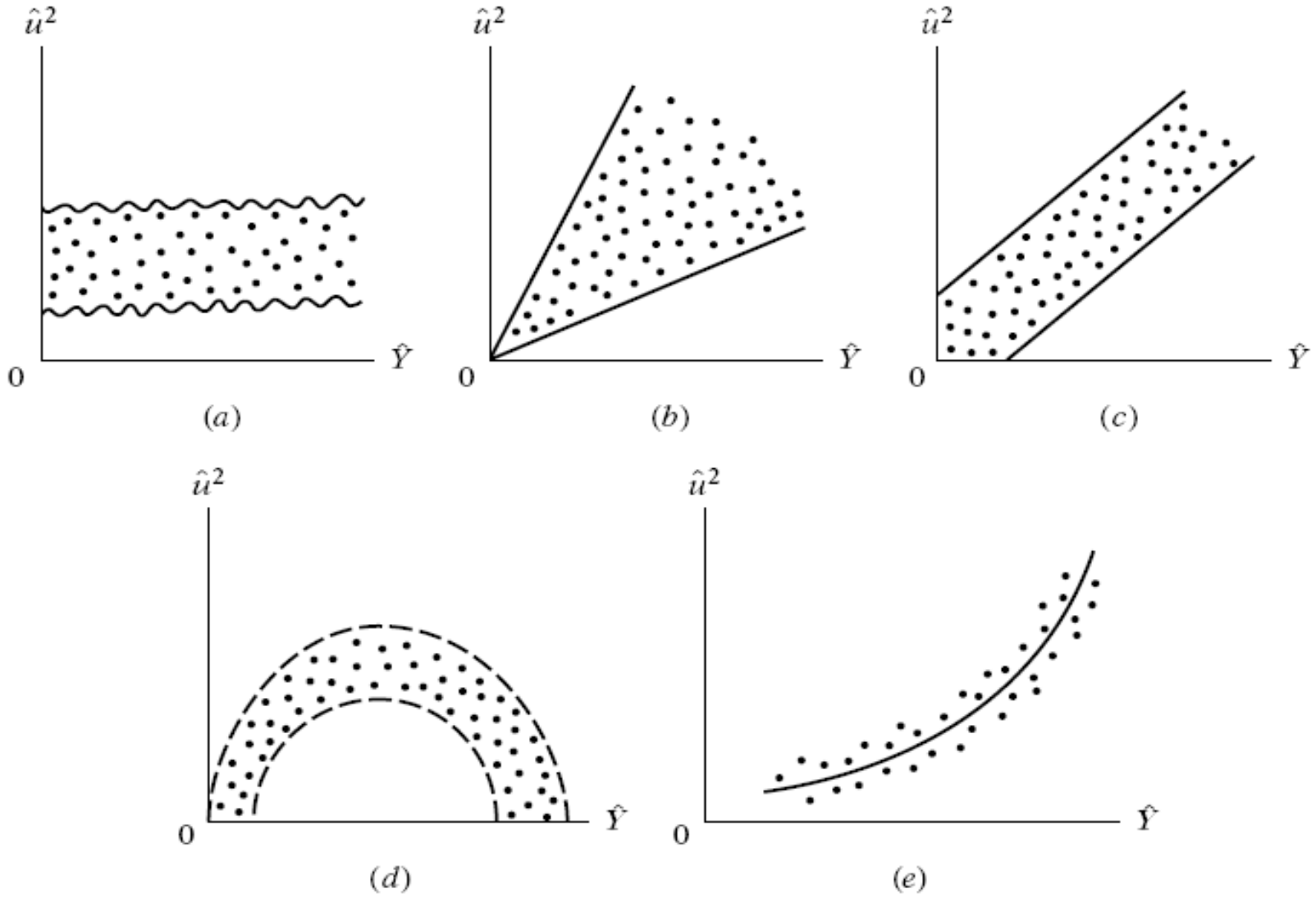


FIGURE 11.8 Hypothetical patterns of estimated squared residuals.

4.2 HETEROSKEDASTICITY

- Instead of plotting $\hat{u}^2_i$ against $\hat{Y}_i$, one may plot them against one of the explanatory variables X_i.
- A pattern such as that shown in Figure 11.9*c*, for instance, suggests that the variance of the disturbance term is linearly related to the *X* variable. Thus, if in the regression of savings on income one finds a pattern such as that shown in Figure 11.9*c*, it suggests that the heteroscedastic variance may be *proportional* to the value of the income variable. This knowledge may help us in transforming our

4.2 HETEROSKEDASTICITY

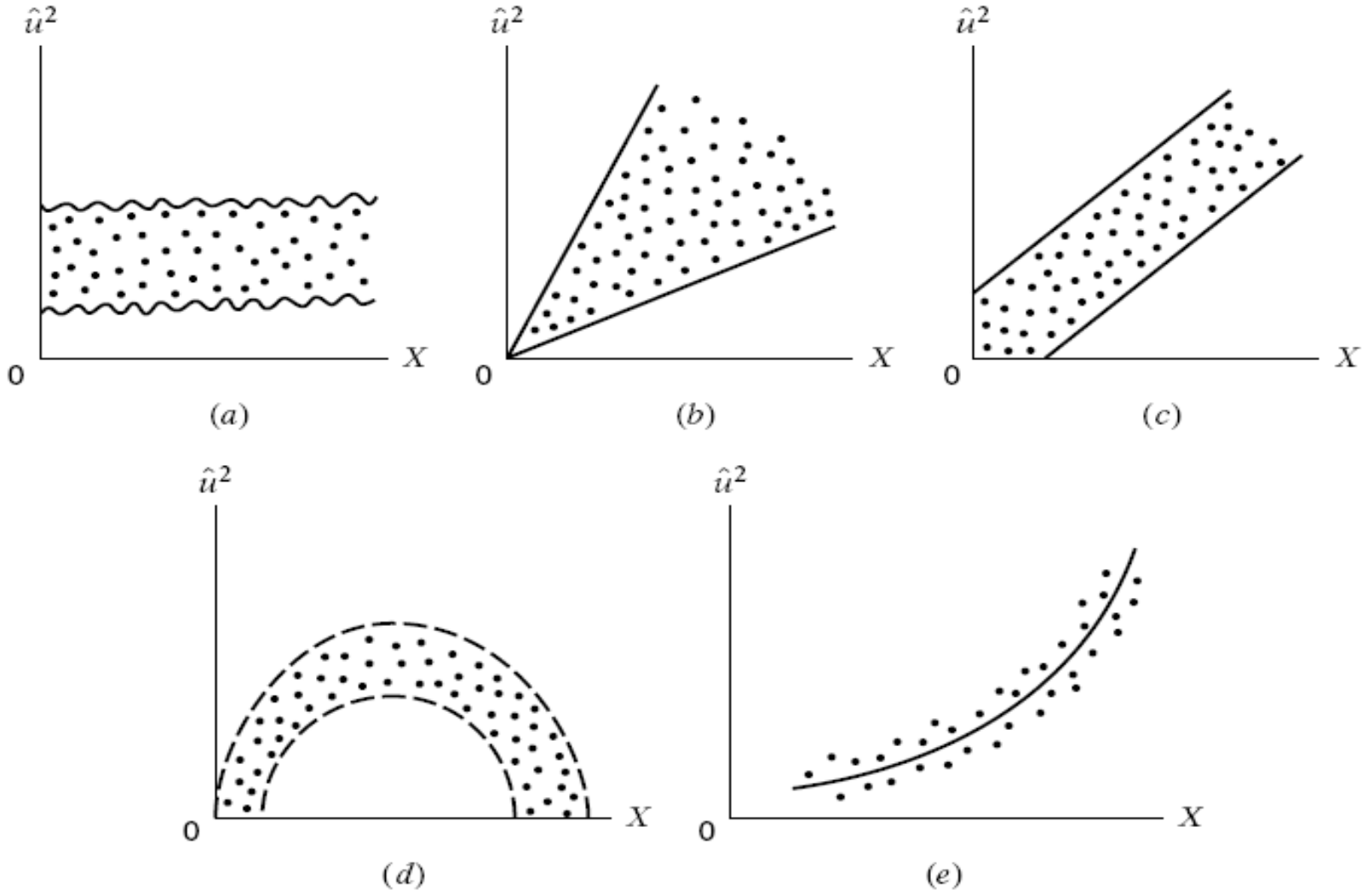


FIGURE 11.9 Scattergram of estimated squared residuals against X.

4.2 HETEROSKEDASTICITY

Formal Methods

1. Park Test. Park suggests that σ^2_i is some function of the explanatory variable X_i. The functional form he suggested was $\boldsymbol{\sigma^2_i = \sigma^2 X^{\beta}_i e^{vi}}$ or

$$\boldsymbol{\ln \sigma^2_i = \ln \sigma^2 + \beta \ln X_i + v_i} \qquad \textbf{(11.5.1)}$$

- where v_i is the stochastic disturbance term. Since σ^2_i is generally not known, Park suggests using $\hat{u}^2_i$ as a proxy and running the following regression:

$$\boldsymbol{\ln \hat{u}^2_i = \ln \sigma^2 + \beta \ln X_i + v_i = \alpha + \beta \ln X_i + v_i}$$

(11.5.2)

- If $\boldsymbol{\beta}$ turns out to be ***statistically significant***, it would suggest that ***heteroscedasticity*** is present in the data. If it turns out to be insignificant, we may accept the assumption of homoscedasticity. The Park test is a two stage procedure.

4.2 HETEROSKEDASTICITY

- Although empirically appealing, the Park test has some problems. Goldfeld and Quandt have argued that the error term v_i entering into (11.5.2) may not satisfy the OLS assumptions and may itself be heteroscedastic.
- **EXAMPLE 11.1 Relationship between compensation and productivity**

use the data given in Table 11.1 to run the following regression:

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

- where Y = average compensation in thousands of dollars, X = average productivity in thousands of dollars, and $i = i^{\text{th}}$ employment size of the establishment. The results of the regression were as follows:

$$\boldsymbol{\hat{Y}_i = 1992.3452 + 0.2329 X_i}$$

4.2 HETEROSKEDASTICITY

- The residuals obtained from regression (11.5.3) were regressed on *Xi* as suggested in Eq. (11.5.2), giving the following results:

$$\ln \hat{u}_i^2 = 35.817 - 2.8099 \ln X_i$$

$$\text{se} = (38.319)\ (4.216)$$

$$t = (0.934)\quad (-0.667)\qquad R^2 = 0.0595 \qquad (11.5.4)$$

- Obviously, there is no statistically significant relationship between the two variables. Following the Park test, one may conclude that

4.2 HETEROSKEDASTICITY

2. Goldfeld-Quandt Test.

This popular method is applicable if one assumes that the heteroscedastic variance, σ^2_i , is positively related to *one* of the explanatory variables in the regression model. For simplicity, consider the usual two-variable model: $\boldsymbol{Y_i = \beta_1 + \beta_2 X_i + U_i}$

Suppose σ^2_i is positively related to X_i as

$$\sigma^2_i = \sigma^2 X_i^2 \qquad (11.5.10)$$

4.2 HETEROSKEDASTICITY

- Assumption (11.5.10) postulates that σ^2_i is proportional to the square of the X variable.
- Such an assumption has been found quite useful by Prais and outhakker in their study of family budgets.
- If (11.5.10) is appropriate, **it would mean σ^2_i would be larger, the larger the values of X_i.**
- If that turns out to be the case,

4.2 HETEROSKEDASTICITY

Step 1. Order or rank the observations according to the values of Xi, beginning with the lowest X value.

Step 2. Omit c central observations, where c is specified a priori, and divide the remaining $(n - c)$ observations into two groups each of $(n - c) / 2$ observations.

Step 3. Fit separate OLS regressions to the first $(n - c)/2$ observations and the last $(n - c) / 2$ observations, and obtain the respective residual sums of squares RSS_1 and RSS_2, RSS_1 representing the RSS from the regression corresponding to the smaller X_i values (the small variance group) and RSS2 that from the larger Xi values (the large variance group). These RSS each have $(n- c) / 2 - k$ or $(n- c - 2k) / 2$ df where k is the number of parameters to be estimated, including the intercept.

4.2 HETEROSKEDASTICITY

Step 4. Compute the ratio

λ = RSS2/df / RSS1/df

(11.5.11)

- *If U_i are assumed to be normally distributed* (which we usually do), and *if the assumption of homoscedasticity is valid*, then it can be shown that **λ** of (11.5.10) follows the F distribution with numerator and denominator df each of $(n - c - 2k)/2$.

If in an application the computed λ (= F) is greater than the critical F at the chosen level of significance, we can reject the

4.2 HETEROSKEDASTICITY

- To illustrate the Goldfeld–Quandt test, we present in Table 11.3 data on consumption expenditure in relation to income for a cross section of 30 families. Suppose we postulate that consumption expenditure is linearly related to income but that heteroscedasticity is present in the data. We further postulate that the nature of heteroscedasticity is as given in (11.5.10). The necessary reordering of the data for the application of the test is also presented in Table 11.3.
- Dropping the middle 4 observations, the OLS regressions based on the first 13 and the last 13 observations and their associated residual sums of squares are as shown next (standard

4.2 HETEROSKEDASTICITY

- $\hat{Y}i = 3.4094 + 0.6968Xi$
  (8.7049) (0.0744) $r^2 = 0.8887$
  RSS1 = 377.17 df = 11
- Regression based on the last 13 observations:
- $\hat{Y}i = -28.0272 + 0.7941Xi$
  (30.6421) (0.1319) $r^2 = 0.7681$
  RSS2 = 1536.8 df = 11

From these results we obtain

$\lambda = (RSS_2/df)/(RSS_1/df) = (1536.8/11)/(377.17/11)$

$\lambda = 4.07$

- **The critical *F* value for 11 numerator and 11 denominator df at the 5 percent level is 2.82. Since the estimated $F(=\lambda)$ value exceeds the critical value, we may conclude that there is heteroscedasticity in the error variance.** However, if the level of significance is fixed at 1 percent, we may not reject the assumption of homoscedasticity. (Why?)
  Note that the *p* value of the observed λ is 0.014

TABLE 11.3
HYPOTHETICAL DATA ON CONSUMPTION EXPENDITURE *Y*($) AND INCOME *X*($) TO ILLUSTRATE THE GOLDFELD–QUANDT TEST

		Data ranked by *X* values		
Y	*X*	*Y*	*X*	
55	80	55	80	
65	100	70	85	
70	85	75	90	
80	110	65	100	
79	120	74	105	
84	115	80	110	
98	130	84	115	
95	140	79	120	
90	125	90	125	
75	90	98	130	
74	105	95	140	
110	160	108	145	
113	150	113	150	
125	165	110	160	
108	145	125	165	Middle 4 observations
115	180	115	180	
140	225	130	185	
120	200	135	190	
145	240	120	200	
130	185	140	205	
152	220	144	210	
144	210	152	220	
175	245	140	225	
180	260	137	230	
135	190	145	240	
140	205	175	245	
178	265	189	250	
191	270	180	260	
137	230	178	265	
189	250	191	270	

4.2 HETEROSKEDASTICITY

3. Breusch–Pagan–Godfrey Test

The success of the Goldfeld–Quandt test depends not only on the value of c(the number of central observations to be omitted) but also on identifying the correct X variable with which to order the observations.

This limitation of this test can be avoided if we consider the **Breusch–Pagan–Godfrey (BPG) test.**

4.2 HETEROSKEDASTICITY

To illustrate this test, consider the k-variable linear regression model

$$Y_i = \beta_1 + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i \tag{11.5.12}$$

Assume that the error variance σ_i^2 is described as

$$\sigma_i^2 = f(\alpha_1 + \alpha_2 Z_{2i} + \cdots + \alpha_m Z_{mi}) \tag{11.5.13}$$

that is, σ_i^2 is some function of the nonstochastic variables Z's; some or all of the X's can serve as Z's. Specifically, assume that

$$\sigma_i^2 = \alpha_1 + \alpha_2 Z_{2i} + \cdots + \alpha_m Z_{mi} \tag{11.5.14}$$

4.2 HETEROSKEDASTICITY

that is, σ_i^2 is a linear function of the Z's. If $\alpha_2 = \alpha_3 = \cdots = \alpha_m = 0$, $\sigma_i^2 = \alpha_1$, which is a constant. Therefore, to test whether σ_i^2 is homoscedastic, one can test the hypothesis that $\alpha_2 = \alpha_3 = \cdots = \alpha_m = 0$. This is the basic idea behind the Breusch–Pagan test. The actual test procedure is as follows.

4.2 HETEROSKEDASTICITY

Step 1. Estimate (11.5.12) by OLS and obtain the residuals $\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_n$.

Step 2. Obtain $\tilde{\sigma}^2 = \sum \hat{u}_i^2/n$. Recall from Chapter 4 that this is the maximum likelihood (ML) estimator of σ^2. [*Note:* The OLS estimator is $\sum \hat{u}_i^2/(n-k)$.]

Step 3. Construct variables p_i defined as

$$p_i = \hat{u}_i^2/\tilde{\sigma}^2$$

which is simply each residual squared divided by $\tilde{\sigma}^2$.

Step 4. Regress p_i thus constructed on the Z's as

$$p_i = \alpha_1 + \alpha_2 Z_{2i} + \cdots + \alpha_m Z_{mi} + v_i \qquad \textbf{(11.5.15)}$$

where v_i is the residual term of this regression.

Step 5. Obtain the ESS (explained sum of squares) from (11.5.15) and define

$$\Theta = \frac{1}{2}(\text{ESS}) \qquad \textbf{(11.5.16)}$$

4.2 HETEROSKEDASTICITY

Assuming Ui are normally distributed, one can show that if there is homoscedasticity and if the sample size n increa

$$\Theta \underset{\text{asy}}{\sim} \chi^2_{m-1} \qquad (11.5.17)$$

that is, Θ follows the chi-square distribution with (m− 1) degrees of freedom. (Note: asy means asymptotically.)

Therefore, if in an application the computed Θ ($= \chi^2$) exceeds the critical χ^2 value at the chosen level of significance, one can reject the hypothesis of homoscedasticity; otherwise one does not reject it.

The reader may wonder why BPG chose $\frac{1}{2}$ESS as the test statistic. The reasoning is slightly involved and is left for the references.[22]

4.2 HETEROSKEDASTICITY

EXAMPLE 11.5

THE BREUSCH–PAGAN–GODFREY (BPG) TEST

As an example, let us revisit the data (Table 11.3) that were used to illustrate the Goldfeld–Quandt heteroscedasticity test. Regressing Y on X, we obtain the following:

Step 1.

$$\hat{Y}_i = 9.2903 + 0.6378X_i$$

$$\text{se} = (5.2314) \quad (0.0286) \qquad \text{RSS} = 2361.153 \qquad R^2 = 0.9466 \qquad \textbf{(11.5.18)}$$

Step 2.

$$\tilde{\sigma}^2 = \sum \hat{u}_i^2/30 = 2361.153/30 = 78.7051$$

Continued.....

Step 3. Divide the squared residuals $\hat{u}_i$ obtained from regression (11.5.18) by 78.7051 to construct the variable p_i.

Step 4. Assuming that p_i are linearly related to $X_i (= Z_i)$ as per (11.5.14), we obtain the regression

$$\hat{p}_i = -0.7426 + 0.0101 X_i$$

$$\text{se} = (0.7529) \quad (0.0041) \qquad \text{ESS} = 10.4280 \qquad R^2 = 0.18 \qquad (11.5.19)$$

Step 5.

$$\Theta = \frac{1}{2}(\text{ESS}) = 5.2140 \qquad (11.5.20)$$

Continued.....

- Under the assumptions of the BPG test in (11.5.20) asymptotically follows the chi square distribution with 1 df. [Note:There is only one regressor in (11.5.19).] Now from the chi-square table we find that for 1 df the 5 percent critical chi-square value is 3.8414 and the 1 percent critical $\chi 2$ value is 6.6349. Thus, the observed chi-square value of 5.2140 is significant at the 5 percent but not the 1 percent level of significance.
- Therefore, we reach the same conclusion as the Goldfeld–Quandt test. But keep in mind that, strictly speaking, the BPG test is an asymptotic, or large-sample, test and in the present example 30 observations may not constitute a large sample. It should also be pointed out that in small samples the test is sensitive to the assumption that the

4. White's General Heteroscedasticity Test

- Unlike the Goldfeld–Quandt test, which requires reordering the observations with respect to the X variable that supposedly caused heteroscedasticity, or the BPG test, which is sensitive to the normality assumption, **the general test of heteroscedasticity proposed by White does not rely on the normality assumption and is easy to implement.**
- As an illustration of the basic idea, consider the following three-variable regression model (the generalization to the k-variable model is straightforward):

Continued.....

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \qquad (11.5.21)$$

The White test proceeds as follows:

Step 1. Given the data, we estimate (11.5.21) and obtain the residuals, $\hat{u}_i$.

Step 2. We then run the following (*auxiliary*) regression:

$$\hat{u}_i^2 = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + \alpha_4 X_{2i}^2 + \alpha_5 X_{3i}^2 + \alpha_6 X_{2i} X_{3i} + v_i \qquad (11.5.22)^{25}$$

Continued.....

- That is, the squared residuals from the original regression are regressed on the original X variables or regressors, their squared values, and the cross product(s) of the regressors.
- Higher powers of regressors can also be introduced.
- Note that there is a constant term in this equation even though the original regression may or may not contain it. Obtain the R2 from this (auxiliary)

Continued.....

- Step 3.Under the null hypothesis that there is no heteroscedasticity, it can be shown that sample size (n) times the R2 obtained from the auxiliary regression asymptotically follows the chi-square distribution with df equal to the number of regressors (excluding the constant term) in t on. That is,

$$n \cdot R^2 \underset{\text{asy}}{\sim} \chi^2_{\text{df}} \qquad (11.5.23)$$

where df is as defined previously. In our example, there are 5 df since there are 5 regressors in the auxiliary regression.

Step 4. If the chi-square value obtained in (11.5.23) exceeds the critical chi-square value at the chosen level of significance, the conclusion is that there is heteroscedasticity. If it does not exceed the critical chi-square value, there is no h

(11.5.21), $\alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = 0$ (see footnote 25).

at in t

EXAMPLE 11.6

WHITE'S HETEROSCEDASTICITY TEST

From cross-sectional data on 41 countries, Stephen Lewis estimated the following regression model

$$\ln Y_i = \beta_1 + \beta_2 \ln X2_i + \beta_3 \ln X3_i + U_i \quad (11.5.24)$$

Where Y=ratio of trade taxes (import and export taxes) to total government revenue,

X2=ratio of the sum of exports plus imports to GNP, and X3=GNP per capita; and ln stands for natural log. His hypotheses were that Y and X2 would be positively related (the higher the trade volume, the higher the trade tax revenue) and that Y and X3 would be negatively related (as income increases, government finds it is easier to collect direct taxes—e.g., income tax—than rely on trade taxes)

- The empirical results supported the hypotheses. For our purpose, the important point is whether there is heteroscedasticity in the data.
- Since the data are cross-sectional involving a heterogeneity of countries, a priori one would expect heteroscedasticity in the error variance.
- By applying White's heteroscedasticity test to the residuals obtained from regression (11.5.24), the fo

$$\widehat{u}_i^2 = -5.8417 + 2.5629 \ln \text{Trade}_i + 0.6918 \ln \text{GNP}_i$$
$$-0.4081(\ln \text{Trade}_i)^2 - 0.0491(\ln \text{GNP}_i)^2 \qquad (11.5.25)$$
$$+0.0015(\ln \text{Trade}_i)(\ln \text{GNP}_i) \qquad R^2 = 0.1148$$

Note: The standard errors are not given, as they are not pertinent for our purpose here.

Now $n \cdot R^2 = 41(0.1148) = 4.7068$, which has, asymptotically, a chi-square distribution with 5 df (why?). The 5 percent critical chi-square value for 5 df is 11.0705, the 10 percent critical value is 9.2363, and the 25 percent critical value is 6.62568. For all practical purposes, one can conclude, on the basis of the White test, that there is no heteroscedasticity.

REMEDIAL MEASURES

- As we have seen, heteroscedasticity does not destroy the unbiasedness and consistency properties of the OLS estimators, but they are no longer efficient, not even asymptotically (i.e., large sample size). This lack of efficiency makes the usual hypothesis-testing procedure of dubious value.
- Therefore, remedial measures may be called for. There are two approaches to remediation: when σ2i is known and when σ2i is not known.

When σ2i Is Known: The Method of Weighted Least Squares

As we have seen in Section 11.3, if σ2i is known, the most straightforward method of correcting heteroscedasticity is **by means of weighted least squares**, for the estimators thus **obtained are**

SUMMARY AND CONCLUSIONS

1. A critical assumption of the classical linear regression model is that the disturbances Ui have all the same variance, σ2. If this assumption is not satisfied, there is heteroscedasticity.

2. Heteroscedasticity does not destroy the unbiasedness and consistency properties of OLS estimators.

3. But these estimators are no longer minimum variance or efficient. That is, they are not BLUE.

4. The BLUE estimators are provided by the method of weighted least squares, provided the

Continued.....

5. In the presence of heteroscedasticity, the variances of OLS estimators are not provided by the usual OLS formulas. But if we persist in using the usual OLS formulas, the tandFtests based on them can be highly misleading, resulting in erroneous conclusions.

6. Documenting the consequences of heteroscedasticity is easier than detecting it. There are several diagnostic tests available, but one cannot tell for sure

Continued.....

7. Even if heteroscedasticity is suspected and detected, it is not easy to correct the problem. If the sample is large, one can obtain **White's heteroscedasticity corrected standard errors of OLS estimators** and conduct statistical inference based on these standard errors.

8. Otherwise, on the basis of OLS residuals, one can make educated guesses of the likely pattern of heteroscedasticity and transform the original data in such a way

END