Chapter 16: Heteroskedasticity

Chapter 16 Outline

• Review
  o Regression Model
  o Standard Ordinary Least Squares (OLS) Premises
  o Estimation Procedures Embedded within the Ordinary Least Squares (OLS) Estimation Procedure

• What Is Heteroskedasticity?
• Heteroskedasticity and the Ordinary Least Squares (OLS) Estimation Procedure: The Consequences
  o The Mathematics
    ▪ Ordinary Least Squares (OLS) Estimation Procedure for the Coefficient Value
    ▪ Ordinary Least Squares (OLS) Estimation Procedure for the Variance of the Coefficient Estimate’s Probability Distribution
  o Our Suspicions
  o Confirming Our Suspicions

• Accounting for Heteroskedasticity: An Example
• Justifying the Generalized Least Squares (GLS) Estimation Procedure

• Robust Standard Errors

Chapter 16 Prep Questions

1. What are the standard ordinary least squares (OLS) premises?

2. In Chapter 6 we showed that the ordinary least squares (OLS) estimation procedure for the coefficient value was unbiased; that is, we showed that $\text{Mean}[b_x] = \beta$. Review the algebra. What role, if any, did the first standard ordinary least squares (OLS) premise, the error term equal variance premise, play?

3. In Chapter 6 we showed that the variance of the coefficient estimate’s probability distribution equals the variance of the error term’s probability distribution divided by the sum of the squared $x$ deviations; that is, we showed that $\text{Var}[b_x] = \frac{\text{Var}[e]}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$. Review the algebra. What role, if any, did the first standard ordinary least squares (OLS) premise, the error term equal variance premise, play?
4. Consider the following data describing Internet use in 1992:

**1992 Internet Data:** Cross section data of Internet use and gross domestic product for 29 countries in 1992.

<table>
<thead>
<tr>
<th>Country,</th>
<th>Name of country t</th>
</tr>
</thead>
<tbody>
<tr>
<td>GdpPC,</td>
<td>Per capita GDP (1,000’s of real “international” dollars) in nation t</td>
</tr>
<tr>
<td>LogUsersInternet,</td>
<td>Log of Internet users per 1,000 people in nation t</td>
</tr>
<tr>
<td>Year,</td>
<td>Year</td>
</tr>
</tbody>
</table>

Focus on the following model:

\[
\text{LogUsersInternet}_t = \beta_{\text{const}} + \beta_{\text{GDP}} \text{GdpPC}_t + e_t
\]

a. What is your theory concerning how per capita GDP should affect Internet use? What does your theory suggest about the sign of the \( \beta_{\text{GDP}} \) coefficient?

b. Run the appropriate regression. Do the data support your theory?

[Link to MIT-Internet-1992.wf1 goes here.]

c. What are the appropriate null and alternative hypotheses?

d. After running the regression, plot a scatter diagram of those residuals and per capita GDP.

**Getting Started in EViews**

- Run the ordinary least squares (OLS) regression.
- The residuals from the regression we just ran are automatically stored by EViews as the variable *resid*.
- In the Workfile window: First Click *gdppc*; then, hold down the <Ctrl> key and click *resid*.
- In the Workfile window: Double click on a highlighted variable
- In the Workfile window: Click **Open Group**
- In the Group window: Click **View** and then **Graph**
- In the Graph Options window: Click **Scatter**

e. Based on the scatter diagram, what do you conclude about the variance of the residuals as per capita GDP increases?

5. Again, consider the following model:

\[
\text{LogUsersInternet}_t = \beta_{\text{const}} + \beta_{\text{GDP}} \text{GdpPC}_t + e_t
\]

Assume that the variance of the error term’s probability distribution is proportional to each nation’s per capita GDP:

\[
\text{Var}[e_t] = V \times \text{GdpPC}_t \quad \text{where } V \text{ is a constant}
\]
Now, divide both sides of the equation that specifies the model by the square root of per capita GDP, $\sqrt{GdpPC_t}$. Let

$$\epsilon_t = \frac{e_t}{\sqrt{GdpPC_t}}$$

What is the variance of $\epsilon_t$’s probability distribution?

**Review**

**Regression Model**

We begin by reviewing the basic regression model:

$$y_t = \beta_{Const} + \beta_x x_t + e_t$$

- $y_t =$ Dependent variable
- $x_t =$ Explanatory variable
- $e_t =$ Error term

$t = 1, 2, \ldots, T$

$T =$ Sample size

The error term is a random variable that represents random influences: $\text{Mean}[e_t] = 0$

**The Standard Ordinary Least Squares (OLS) Premises**

We shall now focus our attention on the standard ordinary least squares (OLS) regression premises:

1. **Error Term Equal Variance Premise:** The variance of the error term’s probability distribution for each observation is the same; all the variances equal $\text{Var}[e]$:

   $$\text{Var}[e_1] = \text{Var}[e_2] = \ldots = \text{Var}[e_T] = \text{Var}[e]$$

2. **Error Term/Error Term Independence Premise:** The error terms are independent: $\text{Cov}[e_i, e_j] = 0$.

   Knowing the value of the error term from one observation does not help us predict the value of the error term for any other observation.

3. **Explanatory Variable/Error Term Independence Premise:** The explanatory variables, the $x_t$’s, and the error terms, the $e_t$’s, are not correlated.

   Knowing the value of an observation’s explanatory variable does not help us predict the value of that observation’s error term.
Estimation Procedures Embedded within the Ordinary Least Squares (OLS) Estimation Procedure

The ordinary least squares (OLS) estimation procedure includes three important estimation procedures. A procedure to estimate the:

- Values of the regression parameters, $\beta_x$ and $\beta_{\text{Const}}$:
  \[
  b_x = \frac{\sum_{t=1}^{T} (y_t - \bar{y})(x_t - \bar{x})}{\sum_{t=1}^{T} (x_t - \bar{x})^2} \quad \text{and} \quad b_{\text{Const}} = \bar{y} - b_x \bar{x}
  \]

- Variance of the error term’s probability distribution, \( \text{Var}[e] \):
  \[
  \text{EstVar}[e] = \frac{\text{SSR}}{\text{Degrees of Freedom}}
  \]

- Variance of the coefficient estimate’s probability distribution, \( \text{Var}[b_x] \):
  \[
  \text{EstVar}[b_x] = \frac{\text{EstVar}[e]}{\sum_{t=1}^{T} (x_t - \bar{x})^2}
  \]

When the standard ordinary least squares (OLS) regression premises are met:

- Each estimation procedure is unbiased; that is, each estimation procedure does not systematically underestimate or overestimate the actual value.
- The ordinary least squares (OLS) estimation procedure for the coefficient value is the best linear unbiased estimation procedure (BLUE).

Crucial Point: When the ordinary least squares (OLS) estimation procedure performs its calculations, it implicitly assumes that the standard ordinary least squares (OLS) regression premises are satisfied.

In this chapter we shall focus on the first standard ordinary least squares (OLS) premise, the error term equal variance premise. We begin by examining precisely what the premise means. Subsequently, we investigate what problems do and do not emerge when the premise is violated and finally what can be done to address the problems that do arise.
What Is Heteroskedasticity?

Heteroskedasticity refers to the variances of the error terms’ probability distributions. The syllable “hetero” means different; the syllable “skedasticity” refers to the spread of the distribution. Heteroskedasticity means that the spread of the error term’s probability distribution differs from observation to observation.

Recall the error term equal variance premise:

- **Error Term Equal Variance Premise:** The variance of the error term’s probability distribution for each observation is the same; all the variances equal \( \text{Var}[e] \):
  \[
  \text{Var}[e_1] = \text{Var}[e_2] = \ldots = \text{Var}[e_T] = \text{Var}[e]
  \]

The presence of heteroskedasticity violates the error term equal variance premise.

We begin by illustrating the effect of heteroskedasticity on the error terms. Consider the three students in Professor Lord’s class who must take a quiz every Tuesday morning:

<table>
<thead>
<tr>
<th>Student 1</th>
<th>Student 2</th>
<th>Student 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_1 = \beta_{\text{Const}} + \beta_0 x_1 + e_1 )</td>
<td>( y_2 = \beta_{\text{Const}} + \beta_0 x_2 + e_2 )</td>
<td>( y_3 = \beta_{\text{Const}} + \beta_0 x_3 + e_3 )</td>
</tr>
<tr>
<td>( y_1 = \text{Student 1’s score} )</td>
<td>( y_2 = \text{Student 2’s score} )</td>
<td>( y_3 = \text{Student 3’s score} )</td>
</tr>
<tr>
<td>( x_1 = \text{Student 1’s studying} )</td>
<td>( x_2 = \text{Student 2’s studying} )</td>
<td>( x_3 = \text{Student 3’s studying} )</td>
</tr>
<tr>
<td>( e_1 = \text{Student 1’s error term} )</td>
<td>( e_2 = \text{Student 2’s error term} )</td>
<td>( e_3 = \text{Student 3’s error term} )</td>
</tr>
</tbody>
</table>

The error terms, the \( e_i \)’s, represent random influences; that is, the error terms have no systematic effort on the dependent variable \( y_i \). Consequently, the mean of each error term’s probability distribution, \( \text{Mean}[e_i] \), equals 0. In other words, if the experiment were repeated many, many times, the error term would average out to be 0. When the distribution is symmetric, half the time the error term would be positive leading to a higher than normal value of \( y_i \) and half the time it would be negative leading to a lower than normal value of \( y_i \). We shall use a simulation to illustrate this.
Econometrics Lab 16.1: The Error Terms and Heteroskedasticity

The list labeled Heter is the “heteroskedasticity factor.” Initially, Heter is specified as 0, meaning that no heteroskedasticity is present. Click Start and then Continue a few times to note that the distribution of each student’s error terms is illustrated in the three histograms at the top of the window. Also, the mean and variance of each student’s error terms are computed. Next, uncheck the Pause checkbox and click Continue; after many, many repetitions of the experiment click Stop. The mean of each student’s error term is approximately 0 indicating that the error terms truly represent random influences; the error terms have no systematic affect on a student’s quiz score. Furthermore, the spreads of each student’s error term distribution appear to be nearly identical; the variance of each student’s error term is approximately the same. Hence, the error term equal variance premise is satisfied.

Figure 16.1: Heteroskedasticity Simulation

The mean and variance of each student’s error term are computed. Next, uncheck the Pause checkbox and click Continue; after many, many repetitions of the experiment click Stop. The mean of each student’s error term is approximately 0 indicating that the error terms truly represent random influences; the error terms have no systematic affect on a student’s quiz score. Furthermore, the spreads of each student’s error term distribution appear to be nearly identical; the variance of each student’s error term is approximately the same. Hence, the error term equal variance premise is satisfied.

Figure 16.2: Error Term Probability Distributions –Error Term Equal Variance Premise Satisfied
Next, change the value of the Heter. When a positive value is specified, the distribution spread increases as we move from Student 1 to Student 2 to Student 3; when a negative value is specified, the spread decreases. Specify 1 instead of 0. Note that when you do this, the title of the variance list changes to Mid Err Var. This occurs because heteroskedasticity is now present and the variances differ from student to student. The list specifies the middle student’s, Student 2’s, variance. By default Student 2’s variance is 500. Now, click Start and then after many, many repetitions of the experiment click Stop. The distribution spreads of each student’s error terms are not identical.

The error term equal variance premise is now violated. What might cause this discrepancy? Suppose that Student 1 tries to get a broad understanding of the material and hence reads all the assignment albeit quickly. On the other hand, Student 3 guesses on what material will be covered on the quiz and spends his/her time thoroughly studying only that material. When Student 3 guesses right he/she will do very well on the quiz, but when he/she guesses wrong he/she will do very poorly. Hence, we would expect Student 3’s quiz scores to be more volatile than Student 1’s. This volatility is reflected by the variance of the error terms. The variance of Student 3’s error term distribution would be greater than Student 1’s.
Heteroskedasticity and the Ordinary Least Squares (OLS) Estimation Procedure: The Consequences

The Mathematics

Now, let us explore the consequences of heteroskedasticity. We shall focus on two of the three estimation procedures embedded within the ordinary least squares (OLS) estimation procedure, the procedures to estimate the:

• value of the coefficient.
• variance of the coefficient estimate’s probability distribution.

Question: Are these estimation procedures still unbiased when heteroskedasticity is present?

Ordinary Least Squares (OLS) Estimation Procedure for the Coefficient Value

Begin by focusing on the coefficient value. Previously, we showed that the estimation procedure for the coefficient value was unbiased by

• applying the arithmetic of means;
and
• recognizing that the means of the error terms’ probability distributions equal 0 (since the error terms represent random influences).

Let us quickly review. First, recall the arithmetic of means:

- Mean of a constant plus a variable: Mean[c + x] = c + Mean[x]
- Mean of a constant times a variable: Mean[cx] = c Mean[x]
- Mean of the sum of two variables: Mean[x + y] = Mean[x] + Mean[y]

To keep the mathematics straightforward, we focused on a sample size of 3:

Equation for Coefficient Estimate:

\[ b_x = \beta_x + \frac{\sum_{t=1}^{T} (x_t - \bar{x})e_t}{\sum_{t=1}^{T} (x_t - \bar{x})^2} = \beta_x + \frac{(x_1 - \bar{x})e_1 + (x_2 - \bar{x})e_2 + (x_3 - \bar{x})e_3}{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2} \]
Now, some algebra:

\[ \text{Mean}[\beta_x] = \text{Mean} \left[ \beta_x + \frac{(x_1 - \bar{x})e_1 + (x_2 - \bar{x})e_2 + (x_3 - \bar{x})e_3}{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2} \right] \]

Applying \( \text{Mean}[c + x] = c + \text{Mean}[x] \)

\[ = \beta_x + \text{Mean} \left[ \frac{1}{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2} \right] \left[ (x_1 - \bar{x})e_1 + (x_2 - \bar{x})e_2 + (x_3 - \bar{x})e_3 \right] \]

Rewriting the fraction as a product

\[ = \beta_x + \frac{1}{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2} \text{Mean} \left[ (x_1 - \bar{x})e_1 + (x_2 - \bar{x})e_2 + (x_3 - \bar{x})e_3 \right] \]

Applying \( \text{Mean}[cx] = c \text{Mean}[x] \)

\[ = \beta_x + \frac{1}{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2} \text{Mean} \left[ (x_1 - \bar{x})e_1 + (x_2 - \bar{x})e_2 + (x_3 - \bar{x})e_3 \right] \]

Applying \( \text{Mean}[x + y] = \text{Mean}[x] + \text{Mean}[y] \)

\[ = \beta_x + \frac{1}{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2} \left[ \text{Mean}[(x_1 - \bar{x})e_1] + \text{Mean}[(x_2 - \bar{x})e_2] + \text{Mean}[(x_3 - \bar{x})e_3] \right] \]

Applying \( \text{Mean}[cx] = c \text{Mean}[x] \)

\[ = \beta_x + \frac{1}{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2} \left[ (x_1 - \bar{x})\text{Mean}[e_1] + (x_2 - \bar{x})\text{Mean}[e_2] + (x_3 - \bar{x})\text{Mean}[e_3] \right] \]

Since \( \text{Mean}[e_1] = \text{Mean}[e_2] = \text{Mean}[e_3] = 0 \)

\[ = \beta_x \]

What is the critical point here? We have not relied on the error term equal variance premise to show that the estimation procedure for the coefficient value is unbiased. Consequently, we suspect that the estimation procedure for the coefficient value should still be unbiased in the presence of heteroskedasticity.
Ordinary Least Squares (OLS) Estimation Procedure for the Variance of the Coefficient Estimate’s Probability Distribution

Next, consider the estimation procedure for the variance of the coefficient estimate’s probability distribution used by the ordinary least squares (OLS) estimation procedure. The strategy involves two steps:

- First, we used the adjusted variance to estimate the variance of the error term’s probability distribution: $\text{EstVar}[e] = \frac{SSR}{\text{Degrees of Freedom}}$ estimates $\text{Var}[e]$.

- Second, we applied the equation relating the variance of the coefficient estimates probability distribution and the variance of the error term’s probability distribution:

$\text{Var}[b_x] = \frac{\text{Var}[e]}{\sum (x_i - \bar{x})^2}$

**Step 1:** Estimate the variance of the error term’s probability distribution from the available information – data from the first quiz

\[ \text{EstVar}[e] = \frac{SSR}{\text{Degrees of Freedom}} \]

\[ \text{EstVar}[b_x] = \frac{\text{EstVar}[e]}{\sum (x_i - \bar{x})^2} \]

**Step 2:** Apply the relationship between the variances of coefficient estimate’s and error term’s probability distributions

\[ \text{Var}[b_x] = \frac{\text{Var}[e]}{\sum (x_i - \bar{x})^2} \]

This strategy is grounded on the premise that the variance of each error term’s probability distribution is the same, the error term equal variance premise:

$\text{Var}[e_1] = \text{Var}[e_2] = \ldots = \text{Var}[e_T] = \text{Var}[e]$.

Unfortunately, when heteroskedasticity is present, the error term equal variance premise is violated because there is not a single $\text{Var}[e]$. The variance differs from observation to observation. When heteroskedasticity is present the strategy used by the ordinary least squares (OLS) estimation procedure to estimate the coefficient estimate’s probability distribution is based on a faulty premise. The ordinary least squares (OLS) estimation procedure is trying to estimate something that does not exist, a single $\text{Var}[e]$. Consequently, we should be suspicious of the procedure.
Our Suspicions

So, where do we stand? We suspect that when heteroskedasticity is present the ordinary least squares (OLS) estimation procedure for the

- coefficient value will still be unbiased.
- variance of the coefficient estimate’s probability distribution may be biased.

Confirming Our Suspicions

We shall use the Econometrics Lab to address the two suspicions.

Econometrics Lab 16.2: Heteroskedasticity and the Ordinary Least Squares (OLS) Estimation Procedure

[Link to MIT-Lab 16.2 goes here.]

Figure 16.4: Heteroskedasticity Simulation
The simulation allows us to address the two critical questions:

- **Question 1:** Is the estimation procedure for the coefficient’s value unbiased; that is, does the mean of the coefficient estimate’s probability distribution equal the actual coefficient value? The relative frequency interpretation of probability allows us to address this question by using the simulation. After many, many repetitions the distribution of the estimated values mirrors the probability distribution. Therefore, we need only compare the mean of the estimated coefficient values with the actual coefficient values. If the two are equal after many, many repetitions, the estimation procedure is unbiased.

<table>
<thead>
<tr>
<th>Relative Frequency</th>
<th>Probability of the coefficient estimate’s probability distribution</th>
<th>Actual coefficient value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Down</td>
<td>= or ≠</td>
</tr>
<tr>
<td>After many, many repetitions</td>
<td>Down</td>
<td></td>
</tr>
<tr>
<td>Mean (average) of the estimated coefficient values after many, many repetitions</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Question 2:** Is the estimation procedure for the variance of the coefficient’s probability distribution unbiased? Again, the relative frequency interpretation of probability allows us to address this question by using the simulation. We need only compare the variance of the estimated coefficient values and estimates for the variance after many, many repetitions. If the two are equal, the estimation procedure is unbiased.

<table>
<thead>
<tr>
<th>Relative Frequency</th>
<th>Probability of the coefficient estimate’s probability distribution</th>
<th>Mean (average) of the variance estimates after many, many repetitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance</td>
<td>Down</td>
<td>= or ≠</td>
</tr>
<tr>
<td>After many, many repetitions</td>
<td>Down</td>
<td></td>
</tr>
<tr>
<td>Variance of the estimated coefficient values after many, many repetitions</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 16.5: *Heteroskedasticity Factor List*

Note that the “Heter” list now appears in this simulation. This list allows us to investigate the effect of heteroskedasticity. Initially, 0 is specified as a benchmark, meaning that no heteroskedasticity is present. Click Start and then after many, many repetitions click Stop. The simulation results appear in Table 16.1. In the absence of heteroskedasticity both estimation procedures are unbiased:

- The estimation procedure for the coefficient value is unbiased. The mean (average) of the coefficient estimates equals the actual coefficient value; both equal 2.
- The estimation procedure for the variance of the coefficient estimates probability distribution is unbiased. The mean (average) of the variance estimates equals the actual variance of the coefficient estimates; both equal 2.5.

When the standard ordinary least squares (OLS) premises are satisfied both estimation procedures are unbiased.

Next, we shall investigate the effect of heteroskedasticity by selecting 1.0 from the “Heter” list. Heteroskedasticity is now present. Click Start and then after many, many repetitions click Stop.

- The estimation procedure for the coefficient value is unbiased. The mean (average) of the coefficient estimates equals the actual coefficient value; both equal 2.
- The estimation procedure for the variance of the coefficient estimates probability distribution is biased. The mean (average) of the estimated variances equals 2.9, while the actual variance equals 3.6.
Is OLS estimation procedure for the coefficient’s value unbiased?

Actual coefficient value

↓

Mean (Average) of the Estimated Values, $b_x$, from All Repetitions

<table>
<thead>
<tr>
<th>Heter Factor of $\beta_x$</th>
<th>Actual Value of $\beta_x$</th>
<th>All Repetitions</th>
<th>Variance of the Estimated Coefficient Values, $b_x$, from All Repetitions</th>
<th>Average of Estimated Variances, $\text{EstVar}[b_x]$, from All Repetitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.0</td>
<td>$\approx$2.0</td>
<td>$\approx$2.5</td>
<td>$\approx$2.5</td>
</tr>
<tr>
<td>1</td>
<td>2.0</td>
<td>$\approx$2.0</td>
<td>$\approx$3.6</td>
<td>$\approx$2.9</td>
</tr>
</tbody>
</table>

Table 16.1: Heteroskedasticity Simulation Results

The simulation results confirm our suspicions. When heteroskedasticity is present there is some good news, but also some bad news:

- **Good news:** The ordinary least squares (OLS) estimation procedure for the coefficient value is still unbiased.
- **Bad news:** The ordinary least squares (OLS) estimation procedure for the variance of the coefficient estimate’s probability distribution is biased.

When the estimation procedure for the variance of the coefficient estimate’s probability distribution is biased, all calculations based on the estimate of the variance will be flawed also; that is, the standard errors, $t$-statistics, and tail probabilities appearing on the ordinary least squares (OLS) regression printout are unreliable. Consequently, we shall use an example to explore how we can account for the presence of heteroskedasticity.
Accounting for Heteroskedasticity: An Example

We can account for heteroskedasticity by applying the following steps:

- **Step 1:** Apply the Ordinary Least Squares (OLS) Estimation Procedure.
  - Estimate the model’s parameters with the ordinary least squares (OLS) estimation procedure.

- **Step 2:** Consider the Possibility of Heteroskedasticity.
  - Ask whether there is reason to suspect that heteroskedasticity may be present.
  - Use the ordinary least squares (OLS) regression results to “get a sense” of whether heteroskedasticity is a problem by examining the residuals.
  - If the presence of heteroskedasticity is suspected, formulate a model to explain it.
  - Use the Breusch-Pagan-Godfrey approach by estimating an artificial regression to test for the presence of heteroskedasticity.

- **Step 3:** Apply the Generalized Least Squares (GLS) Estimation Procedure.
  - Apply the model of heteroskedasticity and algebraically manipulate the original model to derive a new, tweaked model in which the error terms do not suffer from heteroskedasticity.
  - Use the ordinary least squares (OLS) estimation procedure to estimate the parameters of the tweaked model.

We shall illustrate this approach by considering the effect of per capita GDP on Internet use.

**Theory:** Higher per capita GDP increases Internet use.

**Project:** Assess the effect of GDP on Internet use.

To assess the theory we construct a simplified model with a single explanatory variable, per capita GDP. Previously, we showed that several other factors proved important in explaining Internet use. We include only per capita GDP here for pedagogical reasons: to provide a simple illustration of how we can account for the presence of heteroskedasticity.

\[
\text{LogUsersInternet}_t = \beta_{\text{Const}} + \beta_{\text{GDP}} \text{GdpPC}_t + \epsilon_t
\]

where 
- \( \text{LogUsersInternet}_t \) = Log of Internet users per 1,000 people in nation \( t \)
- \( \text{GdpPC}_t \) = Per capita GDP (1,000’s of real “international” dollars) in nation \( t \)

The theory suggests that the model’s coefficient, \( \beta_{\text{GDP}} \), is positive. To keep the exposition clear, we shall use data from a single year, 1992, to test this theory:

<table>
<thead>
<tr>
<th>Countryᵢ</th>
<th>Name of countryᵢ</th>
</tr>
</thead>
<tbody>
<tr>
<td>GdpPCᵢ</td>
<td>Per capita GDP (1,000’s of real “international” dollars) in nationᵢ</td>
</tr>
<tr>
<td>LogUsersInternetᵢ</td>
<td>Log of Internet users per 1,000 people in countryᵢ</td>
</tr>
<tr>
<td>Yearᵢ</td>
<td>Year</td>
</tr>
</tbody>
</table>

We now follow the steps outlined above.

**Step 1:** Apply the Ordinary Least Squares (OLS) Estimation Procedure. Using statistical software, we run a regression with the log of Internet use as the dependent variable and the per capita GDP as the explanatory variable:

[Link to MIT-Internet-1992.wf1 goes here.]

**Ordinary Least Squares (OLS)**

<table>
<thead>
<tr>
<th>Dependent Variable: LogUsersInternet</th>
<th>Explanatory Variable(s):</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
</tr>
<tr>
<td>GdpPC</td>
<td>0.100772</td>
</tr>
<tr>
<td>Const</td>
<td>-0.486907</td>
</tr>
</tbody>
</table>

Number of Observations 29

Estimated Equation: \( \text{EstLogUsersInternet} = -0.487 + 0.101 \text{GdpPC} \)

Interpretation of Estimates:

\( b_{\text{GDP}} = 10.1 \): a $1,000 increase in real per capita GDP results in a 10.1 percent increase in Internet users.

Critical Result: The GdpPC coefficient estimate equals .101. The positive sign of the coefficient estimate suggests that higher per capita GDP increases Internet use. This evidence supports the theory.

Table 16.2: Internet Regression Results
Since the evidence appears to support the theory, we construct the null and alternative hypotheses:

\[ H_0: \beta_{\text{GDP}} = 0 \quad \text{Per capita GDP does not affect Internet use} \]
\[ H_1: \beta_{\text{GDP}} > 0 \quad \text{Higher per capita GDP increases Internet use} \]

As always, the null hypothesis challenges the evidence; the alternative hypothesis is consistent with the evidence. Next, we calculate Prob[Results IF H₀ True].

Prob[Results IF H₀ True]: What is the probability that the \( \text{GdpPC} \) estimate from one repetition of the experiment will be .101 or more, if H₀ were true (that is, if the per capita GDP has no effect on the Internet use, if \( \beta_{\text{GDP}} \) actually equals 0)?

We now apply the properties of the coefficient estimate’s probability distribution:

- **OLS estimation**
- **If H₀**
- **Standard error**
- **Number of observations**
- **Number of parameters**

\[
\text{Mean}[b_{\text{GDP}}] = \beta_{\text{GDP}} = 0 \\
\text{SE}[b_{\text{GDP}}] = .0326 \\
\text{DF} = 29 - 2 = 27
\]

**Econometrics Lab 16.3: Calculating Prob[Results IF H₀ True]**

To emphasize that the Prob[Results IF H₀ True] depends on the standard error we shall use the Econometrics Lab to calculate the probability. The following information has been entered in the lab:

- Mean = 0
- Value = .101
- Standard Error = .0326
- Degrees of Freedom = 27

Click Calculate.

\[ \text{Prob[Results IF H₀ True]} = .0023. \]

We use the standard error provided by the ordinary least squares (OLS) regression results to compute the Prob[Results IF H₀ True].

**We can also calculate Prob[Results IF H₀ True] using the tails probability reported in the regression printout. Since this is a one-tailed test, we divide the tails probability by 2:**

\[ \text{Prob[Results IF H₀ True]} = \frac{.0046}{2} = .0023 \]

Based on the 1 percent significance level, we would reject that null hypothesis. We would reject the hypothesis that per capita GDP has no effect on Internet use.
There may a problem with this, however. The equation used by the ordinary least squares (OLS) estimation procedure to estimate the variance of the coefficient estimate’s probability distribution assumes that the error term equal variance premise is satisfied. Our simulation revealed that when heteroskedasticity is present and the error term equal variance premise is violated, the ordinary least squares (OLS) estimation procedure estimating the variance of the coefficient estimate’s probability distribution was flawed. Recall that the standard error equals the square root of the estimated variance. Consequently, if heteroskedasticity is present, we may have entered the wrong value for the standard error into the Econometrics Lab when we calculated \( \text{Prob}[\text{Results IF } H_0 \text{ True}] \). When heteroskedasticity is present the ordinary least squares (OLS) estimation procedure bases its computations on a faulty premise, resulting in flawed standard errors, \( t \)-Statistics, and tails probabilities. Consequently, we should move on to the next step.

**Step 2:** Consider the Possibility of Heteroskedasticity.
We must pose the following question:

**Question:** Is there reason to believe that heteroskedasticity could possibly be present?

Intuition leads us to suspect that the answer is yes. When the per capita GDP is low, individuals have little to spend on any goods other than the basic necessities. In particular, individuals have little to spend on Internet use and consequently Internet use will be low. This will be true for all countries in which per capita GDP is low. On the other hand, when per capita GDP is high, individuals can afford to purchase more goods. Naturally, consumer tastes vary from nation to nation. In some high per capita GDP nations, individuals will opt to spend much on Internet use while in other nations individuals will spend little. A scatter diagram of per capita GDP and Internet use appears to confirm our intuition:
As per capita GDP rises we observe a greater variance for the log of Internet use per 1,000 persons. In nations with low levels of per capita GDP (less than $15,000), the log varies between about 0 and 1.6. Whereas in nations with high level of per capita GDP (more than $15,000), the log varies between about 0 and 3.20. What does this suggest about the error term in our model:

\[ \text{LogUsersInternet}_t = \beta_{\text{Const}} + \beta_{\text{GDP}} \text{GdpPC}_t + e_t \]

Two nations with virtually the same level of per capita GDP have quite different rates of Internet use. The error term in the model would capture these differences. Consequently, as per capita GDP increases we would expect the variance of the error term’s probability distribution to increase.

Of course we can never observe the error terms themselves. We can, however, think of the residuals as the estimated error terms:

\[ y_t = \beta_{\text{Const}} + \beta_x x_t + e_t \]
\[ \text{Res}_t = y_t - \text{Est}_t \]
\[ e_t = y_t - (\beta_{\text{Const}} + \beta_x x_t) \]
\[ \text{Res}_t = y_t - (b_{\text{Const}} + b_x x_t) \]

Since the residuals are observable we can plot a scatter diagram of the residuals, the estimated errors, and per capita GDP to illustrate how they are related:
Figure 16.7: Scatter Diagram: GdpPC versus Residuals

Getting Started in EViews

- Run the ordinary least squares (OLS) regression.
- The residuals from the regression we just ran are automatically stored by EViews as the variable resid.
- In the Workfile window: First Click gdppe; then, hold down the <Ctrl> key and click resid.
- In the Workfile window: Double click on a highlighted variable
- In the Workfile window: Click Open Group
- In the Group window: Click View and then Graph
- In the Graph Options window: Click Scatter

Our suspicions appear to be borne out. The residuals in nations with high per capita GDP are more spread out than in nations with low per capita GDP. It appears that heteroskedasticity could be present. The error term equal variance premise may be violated. Consequently, we must be suspicious of the standard errors and probabilities appearing in the regression printout; the ordinary least squares (OLS) estimation procedure is calculating these values based on what could be an invalid premise, the error term equal variance premise.
Since the scatter diagram suggests that our fears may be warranted, we now test the heteroskedasticity more formally. While there are several different approaches, we shall focus on the Breusch-Pagan-Godfrey test which utilizes an artificial regression based on the following model:

**Heteroskedasticity Model**: \((e_t - \text{Mean}[e_t])^2 = \alpha_{\text{Const}} + \alpha_{\text{GDP}}GdpPC_t + v_t\)

The model suggests that as \(GdpPC_t\) increases, the squared deviation of the error term from its mean increases. Based on the scatter diagram appearing in Figure 16.7, we suspect that \(\alpha_{\text{GDP}}\) is positive:

Theory: \(\alpha_{\text{GDP}} > 0\)

We can simplify this model by recognizing that the error term represents random influences; hence, the mean of its probability distribution equals 0; hence,

\((e_t - \text{Mean}[e_t])^2 = e_t^2\)

So, our model becomes

**Heteroskedasticity Model**: \(e_t^2 = \alpha_{\text{Const}} + \alpha_{\text{GDP}}GdpPC_t + v_t\)

Of course we can never observe the error terms themselves. We can, however, think of the residuals as the estimates of the error terms. We substitute the residual squared for the error term squared:

**Heteroskedasticity Model**: \(ResSqr_t = \alpha_{\text{Const}} + \alpha_{\text{GDP}}GdpPC_t + v_t\)

where \(ResSqr_t = \text{Square of the residual for nation } t\)
We must generate the new variable \(ResSqr\). To do so, run the original regression so that the statistical software calculates the residuals. Then, after gaining access to the residuals, square them to generate the new residuals squared variable, \(ResSqr\). Then, estimate the model’s parameters:

\[
\text{Ordinary Least Squares (OLS)}
\]

**Dependent Variable:** \(ResSqr\)

**Explanatory Variable(s):**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>SE</th>
<th>(t)-Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>(GdpPC)</td>
<td>0.086100</td>
<td>0.031863</td>
<td>2.702189</td>
<td>0.0118</td>
</tr>
<tr>
<td>(Const)</td>
<td>–0.702317</td>
<td>0.617108</td>
<td>–1.138078</td>
<td>0.2651</td>
</tr>
</tbody>
</table>

Number of Observations 29

**Critical Result:** The \(GdpPC\) coefficient estimate equals .086. The positive sign of the coefficient estimate suggests that higher per capita GDP increases the squared deviation of the error term from its mean. This evidence supports the theory that heteroskedasticity is present.

Table 16.3: *Breusch-Pagan-Godfrey Results*

Next, formulate the null and alternative hypotheses for the artificial regression model:

\[
\begin{align*}
H_0 &: \alpha_{GDP} = 0 \quad \text{Per capita GDP does not affect the squared deviation of the residual} \\
H_1 &: \alpha_{GDP} > 0 \quad \text{Higher per capita GDP increases the squared deviation of the residual}
\end{align*}
\]

and compute \(Prob[\text{Results IF } H_0 \text{ True}]\) from the tails probability reported in the regression printout:

\[
Prob[\text{Results IF } H_0 \text{ True}] = \frac{.0118}{2} \approx .0059
\]
Statistical software can perform the Breusch-Pagan-Godfrey test automatically:

**Getting Started in EViews**

- Run the ordinary least squares (OLS) regression.
- In the equation window, click View, Residual Diagnostics, and Heteroskedasticity Tests
- The Breusch-Pagan-Godfrey test is the default
- After checking the explanatory variables (the regressors), click OK.

We reject the null hypothesis at the traditional significance levels of 1, 5, and 10 percent. Our formal test reinforces our suspicion that heteroskedasticity is present. Furthermore, note that the estimate of the constant is not statistically significantly different from 0 even at the 10 percent significance level. We shall exploit this to simplify the mathematics that follow. We assume that the variance of the error term’s probability distribution is directly proportional to per capita GDP:

\[ \text{Var}[e_t] = V \times \text{GdpPC}_t \]

where \( V \) equals a constant.

As per capita GDP increases the variance of the error term’s probability increases just as the scatter diagram suggests. We shall now use this specification of error term variance to tweak the original model to “eliminate” heteroskedasticity. This approach is called the generalized least square (GLS) estimation procedure.

**Step 3:** Apply the Generalized Least Squares (GLS) Estimation Procedure.

**Strategy:** Algebraically manipulate the original model so that the problem of heteroskedasticity is eliminated in the new model. That is, tweak the original model so that variance of each nation’s error term’s probability distribution is the same. We can accomplish this with just a little algebra.

Based on our scatter diagram and the Breusch-Pagan-Godfrey test, we assume that the variance of the error term’s probability distribution is proportional to per capita GDP:

\[ \text{Var}[e_t] = V \times \text{GdpPC}_t \]

where \( V \) equals a constant.

We begin with our original model:

\[ \text{Original Model: } \log \text{UsersInternet}_t = \beta_{\text{Const}} + \beta_{\text{GDP}} \text{GdpPC}_t + e_t \]

Now, divide both sides of the equation by \( \sqrt{\text{GdpPC}_t} \). (For the moment, do not worry about why we divide by \( \sqrt{\text{GdpPC}_t} \); we shall justify that shortly.)
Tweaked Model: \[ \frac{\text{LogUsersInternet}_i}{\sqrt{\text{GdpPC}_i}} = \frac{\beta_{\text{Const}}}{\sqrt{\text{GdpPC}_i}} + \beta_{\text{GDP}} \frac{\text{GdpPC}_i}{\sqrt{\text{GdpPC}_i}} + \frac{\epsilon_i}{\sqrt{\text{GdpPC}_i}} \]

\[ = \frac{\beta_{\text{Const}}}{\sqrt{\text{GdpPC}_i}} + \beta_{\text{GDP}} \sqrt{\text{GdpPC}_i} + \epsilon_i \]

where \( \epsilon_i = \frac{\epsilon_i}{\sqrt{\text{GdpPC}_i}} \)

To understand why we divided by \( \sqrt{\text{GdpPC}_i} \), focus on the error term, \( \epsilon_i \), in the tweaked model. What does the variance of this error term’s probability distribution equal?

\[ \text{Var}[\epsilon_i] = \text{Var}\left[\frac{\epsilon_i}{\sqrt{\text{GdpPC}_i}}\right] \]

\[ \downarrow \]

Arithmetic of variances: \( \text{Var}[cx] = c^2 \text{Var}[x] \)

\[ = \frac{1}{\sqrt{\text{GdpPC}_i}} \text{Var}[\epsilon_i] \]

\[ \downarrow \]

\( \text{Var}[\epsilon_i] = V \times GpcPC_i \), where \( V \) equals a constant

\[ = \frac{1}{\sqrt{\text{GdpPC}_i}} V \times GpcPC_i \]

\[ = V \]

We divided the original model by \( \sqrt{\text{GdpPC}_i} \) so that the variance of the error term’s probability distribution in the tweaked model equals \( V \) for each observation. Consequently, the error term equal variance premise is satisfied in the tweaked model. Therefore, the ordinary least squares (OLS) estimation procedure computations of the estimates for the variance of the error term’s probability distribution will not be flawed in the tweaked model.

The dependent and explanatory variables in the new tweaked model are:

Tweaked Dependent Variable: \( \text{AdjLogUsersInternet}_i = \frac{\text{LogUsersInternet}_i}{\sqrt{\text{GdpPC}_i}} \)

Tweaked Explanatory Variables:

\( \text{AdjConst}_i = \frac{1}{\sqrt{\text{GdpPC}_i}} \)

\( \text{AdjGdpPC}_i = \frac{1}{\sqrt{\text{GdpPC}_i}} \)

NB: The tweaked model does not include a constant.
When we run a regression with the new dependent and new explanatory variables, we should now have an unbiased procedure to estimate the variance of the coefficient estimate’s probability distribution. Be careful to eliminate the constant when running the tweaked regression:

[Link to MIT-Internet-1992.wf1 goes here.]

**Ordinary Least Squares (OLS)**

**Dependent Variable:** AdjLogUsersInternet  
**Explanatory Variable(s):**  
- AdjGdpPC: 0.113716, 0.026012, 4.371628, 0.0002  
- AdjConst: -0.726980, 0.450615, -1.613306, 0.1183  

Number of Observations: 29

**Estimated Equation:**  
EstAdjLogUsersInternet = -.727AdjConst + .114GdpPC

**Interpretation of Estimates:**  
\( b_{GDP} = 11.4 \): a $1,000 increase in real per capita GDP results in a 11.4 percent increase in Internet users.

Table 16.4: Tweaked Internet Regression Results

Now, let us compare the tweaked regression for \( \beta_{GDP} \) and the original one:

<table>
<thead>
<tr>
<th>( \beta_{GDP} )</th>
<th>Coefficient Estimate</th>
<th>Standard Error</th>
<th>( t )-Statistic</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ordinary Least Squares (OLS)</td>
<td>.101</td>
<td>.033</td>
<td>3.09</td>
<td>.0046</td>
</tr>
<tr>
<td>Generalized Least Squares (GLS)</td>
<td>.114</td>
<td>.026</td>
<td>4.37</td>
<td>.0002</td>
</tr>
</tbody>
</table>

Table 16.5: Comparison of Internet Regression Results

The most striking differences are the standard error, the \( t \)-Statistic, and Prob values. This is hardly surprising. The ordinary least squares (OLS) regression calculations are based on the error term equal variance premise. Our analysis suggests that this premise is violated ordinary least squares (OLS) regression, however. Consequently, the standard Error, \( t \)-Statistic, and Prob. calculations, will be flawed when we use the ordinary least squares (OLS) estimation procedure. The general least squares (GLS) regression corrects for this.
Recall the purpose of our analysis in the first place: Assess the effect of per capita GDP on Internet use. Recall our theory and associated hypotheses:

**Theory:** Higher per capita GDP increases Internet use.

H₀: \( \beta_{GDP} = 0 \) Per capita GDP does not affect Internet use

H₁: \( \beta_{GDP} > 0 \) Higher per capita GDP increases Internet use

We see that the value of the tails probability decreases from .0046 to .0002. Since a one-tailed test is appropriate, the Prob[Results IF H₀ True] declines from .0023 to .0001. Accounting for heteroskedasticity has an impact on the analysis.

**Justifying the Generalized Least Squares (GLS) Estimation Procedure**

We shall now use a simulation to illustrate that the generalized least squares (GLS) estimation procedure for the

- value of the coefficient estimate is unbiased.
- variance of the coefficient estimate’s probability distribution is unbiased.
- value of the coefficient estimate is the best linear unbiased estimation procedure (BLUE).

*Econometrics Lab 16.4: Generalized Least Squares (GLS) Estimation Procedure*

[Link to MIT-Lab 16.4 goes here.]
A new drop down box now appears in which we can specify the estimation procedure: either ordinary least squares (OLS) or generalized least squares (GLS). Initially, OLS is specified indicating that the ordinary least squares (OLS) estimation procedure is being used. Also, note that by default .0 is specified in the Heter list which means that no heteroskedasticity is present. Recall our previous simulations illustrating that the ordinary least squares (OLS) estimation procedure to estimate the coefficient value and the ordinary least squares (OLS) procedure to estimate the variance of the coefficient estimate’s probability distribution were both unbiased when no heteroskedasticity is present. To review this, click Start and then after many, many repetitions click Stop.
Next, introduce heteroskedasticity by selecting 1.0 from the “Heter” list. Recall that while the ordinary least squares (OLS) estimation procedure for the coefficient’s value was still unbiased, the ordinary least squares (OLS) estimation procedure for the variance of the coefficient estimate’s probability distribution was biased. To review this, click Start and then after many, many repetitions click Stop.

<table>
<thead>
<tr>
<th>Heter Factor</th>
<th>Estim Proc</th>
<th>Actual Value of $\beta_x$</th>
<th>Mean (Average) of the Estimated Values, $b_x$, from All Repetitions</th>
<th>Variance of the Estimated Coefficient Values, $b_x$, from All Repetitions</th>
<th>Average of Estimated Variances, $\text{EstVar}[b_x]$, from All Repetitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>OLS</td>
<td>2.0</td>
<td>≈2.0</td>
<td>≈2.5</td>
<td>≈2.5</td>
</tr>
<tr>
<td>1</td>
<td>OLS</td>
<td>2.0</td>
<td>≈2.0</td>
<td>≈3.6</td>
<td>≈2.9</td>
</tr>
<tr>
<td>1</td>
<td>GLS</td>
<td>2.0</td>
<td>≈2.0</td>
<td>≈2.3</td>
<td>≈2.3</td>
</tr>
</tbody>
</table>

Table 16.6: Heteroskedasticity Simulation Results

Finally, select the generalized least squares (GLS) estimation procedure instead of the ordinary least squares (OLS) estimation procedure. Click Start and then after many, many repetitions click Stop. The generalized least squares (GLS) results are reported in the last row of Table 16.6. When heteroskedasticity is present and the generalized least squares (GLS) estimation procedure is used, the variance of the estimated coefficient values from each repetition of the experiment equals the average of the estimated variances. This suggests that the generalized least squares (GLS) procedure indeed provides an unbiased estimation procedure for the variance. Also, note that when heteroskedasticity is present, the variance of the estimated values resulting from generalized least squares (GLS) is less than ordinary least squares (OLS), 2.3 versus 2.9. What does this suggest? The lower variance suggests that the generalized least squares (GLS) procedure provides more reliable estimates when heteroskedasticity is present. In fact, it can be shown that the generalized least squares (GLS) procedure is indeed the best linear unbiased estimation (BLUE) procedure when heteroskedasticity is present.
Let us summarize:

Is the estimation procedure:

- an unbiased estimation procedure for the coefficient’s value?
  - Yes
- variance of the coefficient estimate’s probability distribution?
  - Yes
- for the coefficient value the best linear unbiased estimation procedure (BLUE)?
  - Yes

<table>
<thead>
<tr>
<th>Std Premises</th>
<th>Heteroskedasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS</td>
<td>OLS</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>GLS</td>
<td>Yes</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Robust Standard Errors**

We have seen that two issues emerge when heteroskedasticity is present:

- The standard error calculations made by the ordinary least squares (OLS) estimation procedure are flawed.
- While the ordinary least squares (OLS) for the coefficient value is unbiased, it is not the best linear unbiased estimation procedure (BLUE).

Robust standard errors address the first issue and are particularly appropriate when the sample size is large. White standard errors constitute one such approach. We shall not provide a rigorous justification of this approach, the mathematics is too complex. We shall, however, provide the motivation by taking a few liberties. Begin by reviewing our derivation of the variance of the coefficient estimate’s probability distribution, $\text{Var}[b_x]$, presented in Chapter 6:

$$
\beta_x = \beta + \frac{\sum_{i=1}^{T} (x_i - \bar{x})e_i}{\sum_{i=1}^{T} (x_i - \bar{x})^2} = \beta + \frac{(x_1 - \bar{x})e_1 + (x_2 - \bar{x})e_2 + (x_3 - \bar{x})e_3}{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2}
$$

Next, a little algebra:

$$
\text{Var}[b_x] = \text{Var}\left[ \beta + \frac{(x_1 - \bar{x})e_1 + (x_2 - \bar{x})e_2 + (x_3 - \bar{x})e_3}{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2} \right]
$$

Applying $\text{Var}[c + x] = \text{Var}[x]$

$$
= \text{Var}\left[ \frac{(x_1 - \bar{x})e_1 + (x_2 - \bar{x})e_2 + (x_3 - \bar{x})e_3}{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2} \right]
$$

Rewriting the fraction as a product

$$
= \text{Var}\left[ \frac{1}{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2} \right] \text{Var}\left[ (x_1 - \bar{x})e_1 + (x_2 - \bar{x})e_2 + (x_3 - \bar{x})e_3 \right]
$$

Applying $\text{Var}[cx] = c^2\text{Var}[x]$

$$
= \frac{1}{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2} \cdot \text{Var}\left[ (x_1 - \bar{x})e_1 + (x_2 - \bar{x})e_2 + (x_3 - \bar{x})e_3 \right]
$$

Error Term/Error Term Independence Premise
The error terms are independent: \( \text{Var}[x + y] = \text{Var}[x] + \text{Var}[y] \)

\[
\begin{align*}
\text{Var}[(x_i - \bar{x})e_1] + \text{Var}[(x_2 - \bar{x})e_2] + \text{Var}[(x_3 - \bar{x})e_3] = \\
\frac{1}{[(x_i - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2]^2} \left[ (x_i - \bar{x})^2 \text{Var}[e_1] + (x_2 - \bar{x})^2 \text{Var}[e_2] + (x_3 - \bar{x})^2 \text{Var}[e_3] \right]
\end{align*}
\]

Applying \( \text{Var}[cx] = c^2 \text{Var}[x] \)

\[
\begin{align*}
\frac{1}{[(x_i - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2]^2} \left[ (x_i - \bar{x})^2 \text{Var}[e_1] + (x_2 - \bar{x})^2 \text{Var}[e_2] + (x_3 - \bar{x})^2 \text{Var}[e_3] \right]
\end{align*}
\]

Generalizing

\[
\frac{\sum_{i=1}^{y} (x_i - \bar{x})^2 \text{Var}[e_i]}{\left[ \sum_{i=1}^{y} (x_i - \bar{x})^2 \right]^2}
\]

Focus on \( \text{Var}[e_i] \) and recall that the variance equals the average of the squared deviations from the mean:

\( \text{Var}[e_i] = \text{Average of } (e_i - \text{Mean}[e_i])^2 \)

Since error terms represent random influences, the mean equals 0 and since we are considering a single error term, the average is \( e_i^2 \):

\( \text{Var}[e_i] = e_i^2 \)

While the error terms are not observable, we can think of the residuals as the estimated error term. Consequently, we shall use \( \text{Res}_i^2 \) to estimate \( e_i^2 \):

\( e_i^2 \to \text{Res}_i^2 \)

Applying this to the equation for the variance of the coefficient estimate’s probability distribution:
The White robust standard error is the square root of the estimated variance.² Statistical software makes it easy to compute robust standard errors:

[Link to MIT-Internet-1992.wf1 goes here.]

**Getting Started in EViews**

- Run the ordinary least squares (OLS) regression.
- In the equation window, click Estimate and Options
- In the Coefficient covariance matrix box select White from the drop down list.
- Click OK.

**Dependent Variable:** LogUsersInternet  
**Explanatory Variable(s):**  

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>SE</th>
<th>t-Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>GdpPC</td>
<td>0.100772</td>
<td>0.032428</td>
<td>3.107552</td>
<td>0.0044</td>
</tr>
<tr>
<td>Const</td>
<td>-0.486907</td>
<td>0.525871</td>
<td>-0.925906</td>
<td>0.3627</td>
</tr>
</tbody>
</table>

White heteroskedasticity-consistent SEs

**Number of Observations** 29

**Estimated Equation:**  
\[ \text{EstLogUsersInternet} = -0.487 + 0.101 \text{GdpPC} \]

**Table 16.7:** Internet Regression Results – Robust Standard Errors
Recall that to keep the algebra straightforward we assume that the explanatory variables are constants. By doing so, we can apply the arithmetic of means easily. Our results are unaffected by this assumption.

While it is beyond the scope of this textbook, it can be shown that while this estimation procedure is biased, the magnitude of the bias diminishes and approaches zero as the sample size approaches infinity.