Heteroskedasticity

# Chapter 8

Prepared by Vera Tabakova, East Carolina University

# Chapter 8: Heteroskedasticity

- 8.1 The Nature of Heteroskedasticity
- 8.2 Using the Least Squares Estimator
- 8.3 The Generalized Least Squares Estimator
- 8.4 Detecting Heteroskedasticity

# 8.1 The Nature of Heteroskedasticity

$$E(y) = \beta_1 + \beta_2 x \quad (8.1)$$

$$e_i = y_i - E(y_i) = y_i - \beta_1 - \beta_2 x_i \quad (8.2)$$

$$y_i = \beta_1 + \beta_2 x_i + e_i \quad (8.3)$$

# 8.1 The Nature of Heteroskedasticity

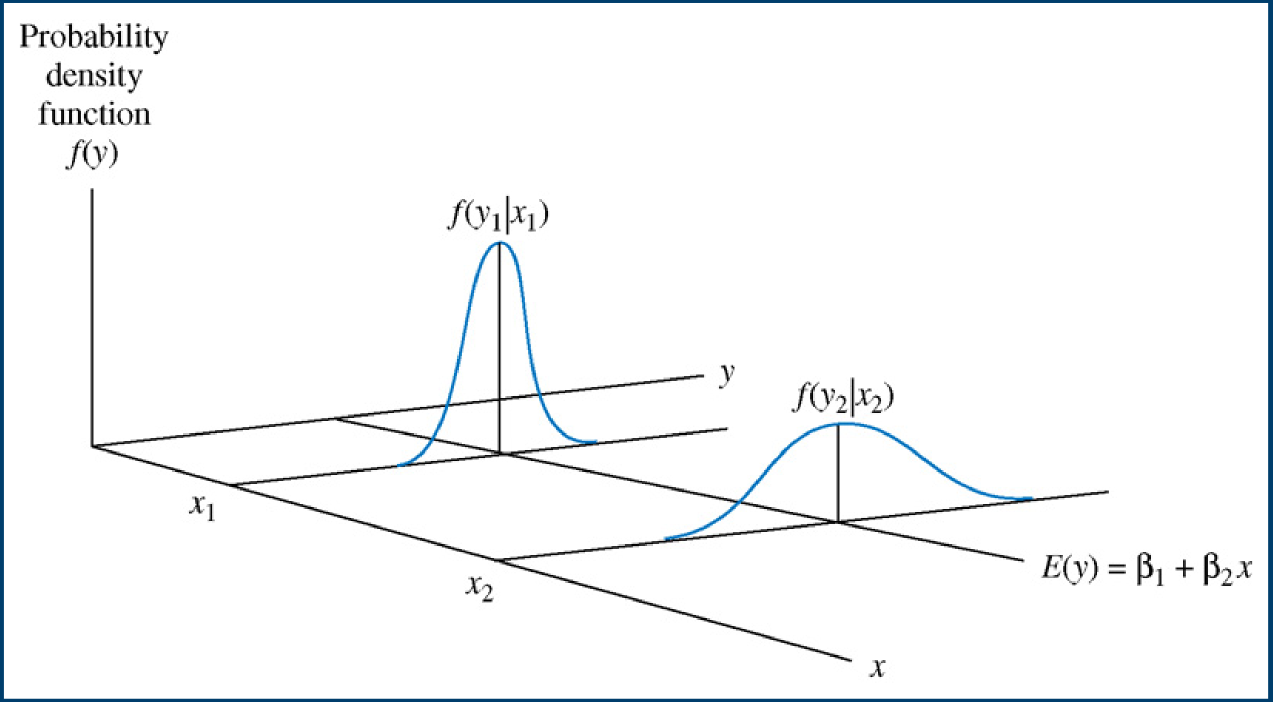


Figure 8.1 Heteroskedastic Errors

# 8.1 The Nature of Heteroskedasticity

$$E(e_i)=0 \qquad \text{var}(e_i)=\sigma^2 \qquad \text{cov}(e_i,e_j)=0$$

$$\text{var}(y_i)=\text{var}(e_i)=h(x_i) \tag{8.4}$$

$$\hat{y}_i = 83.42 + 10.21x_i$$

$$\hat{e}_i = y_i - 83.42 - 10.21x_i$$

# 8.1 The Nature of Heteroskedasticity


Figure 8.2 Least Squares Estimated Expenditure Function and Observed Data Points

# 8.2 Using the Least Squares Estimator

The existence of heteroskedasticity implies:

- The least squares estimator is still a linear and unbiased estimator, but it is no longer best. There is another estimator with a smaller variance.
- The standard errors usually computed for the least squares estimator are incorrect. Confidence intervals and hypothesis tests that use these standard errors may be misleading.

# 8.2 Using the Least Squares Estimator

$$y_i = \beta_1 + \beta_2 x_i + e_i \qquad \text{var}(e_i) = \sigma^2 \tag{8.5}$$

$$\text{var}(b_2) = \frac{\sigma^2}{\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{8.6}$$

$$y_i = \beta_1 + \beta_2 x_i + e_i \qquad \text{var}(e_i) = \sigma_i^2 \tag{8.7}$$

# 8.2 Using the Least Squares Estimator

$$\operatorname{var}(b_2)=\sum_{i=1}^{N} w_i^2\sigma_i^2=\frac{\sum_{i=1}^{N}\left[(x_i-\bar{x})^2\sigma_i^2\right]}{\left[\sum_{i=1}^{N}(x_i-\bar{x})^2\right]^2} \tag{8.8}$$

$$\widehat{\operatorname{var}}(b_2)=\sum_{i=1}^{N} w_i^2\hat{e}_i^2=\frac{\sum_{i=1}^{N}\left[(x_i-\bar{x})^2\hat{e}_i^2\right]}{\left[\sum_{i=1}^{N}(x_i-\bar{x})^2\right]^2} \tag{8.9}$$

# 8.2 Using the Least Squares Estimator

$$\hat{y}_i = 83.42 + 10.21x_i$$

(27.46) (1.81) (White se)

(43.41) (2.09) (incorrect se)

White: $b_2 \pm t_c \text{se}(b_2) = 10.21 \pm 2.024 \times 1.81 = [6.55,\ 13.87]$

Incorrect: $b_2 \pm t_c \text{se}(b_2) = 10.21 \pm 2.024 \times 2.09 = [5.97,\ 14.45]$

# 8.3 The Generalized Least Squares Estimator

$$y_i = \beta_1 + \beta_2 x_i + e_i$$

$$E(e_i) = 0 \qquad \text{var}(e_i) = \sigma_i^2 \qquad \text{cov}(e_i, e_j) = 0 \qquad (8.10)$$

# 8.3.1 Transforming the Model

$$\operatorname{var}(e_i) = \sigma_i^2 = \sigma^2 x_i \qquad (8.11)$$

$$\frac{y_i}{\sqrt{x_i}} = \beta_1\left(\frac{1}{\sqrt{x_i}}\right) + \beta_2\left(\frac{x_i}{\sqrt{x_i}}\right) + \frac{e_i}{\sqrt{x_i}} \qquad (8.12)$$

$$y_i^* = \frac{y_i}{\sqrt{x_i}} \quad x_{i1}^* = \frac{1}{\sqrt{x_i}} \quad x_{i2}^* = \frac{x_i}{\sqrt{x_i}} = \sqrt{x_i} \quad e_i^* = \frac{e_i}{\sqrt{x_i}} \qquad (8.13)$$

# 8.3.1 Transforming the Model

$$y_i^* = \beta_1 x_{i1}^* + \beta_2 x_{i2}^* + e_i^* \tag{8.14}$$

$$\text{var}(e_i^*) = \text{var}\left(\frac{e_i}{\sqrt{x_i}}\right) = \frac{1}{x_i}\text{var}(e_i) = \frac{1}{x_i}\sigma^2 x_i = \sigma^2 \tag{8.15}$$

# 8.3.1 Transforming the Model

To obtain the best linear unbiased estimator for a model with heteroskedasticity of the type specified in equation (8.11):

1. Calculate the transformed variables given in (8.13).

2. Use least squares to estimate the transformed model given in (8.14).

# 8.3.1 Transforming the Model

The generalized least squares estimator is as a **weighted least squares** estimator. Minimizing the sum of squared transformed errors that is given by:

$$\sum_{i=1}^{N} e_i^{*2} = \sum_{i=1}^{N} \frac{e_i^2}{x_i} = \sum_{i=1}^{N} (x_i^{-1/2} e_i)^2$$

When $\sqrt{x_i}$ is small, the data contain more information about the regression function and the observations are weighted heavily.

When $\sqrt{x_i}$ is large, the data contain less information and the observations are weighted lightly.

# 8.3.1 Transforming the Model

$$\hat{y}_i = 78.68 + 10.45 x_i \qquad (8.16)$$
$$(\text{se}) \quad (23.79) \quad (1.39)$$

$$\hat{\beta}_2 \pm t_c \text{se}(\beta_2) = 10.451 \pm 2.024 \times 1.386 = [7.65, 13.26]$$

# 8.3.2 Estimating the Variance Function

$$\text{var}(e_i) = \sigma_i^2 = \sigma^2 x_i^\gamma \qquad (8.17)$$

$$\ln(\sigma_i^2) = \ln(\sigma^2) + \gamma \ln(x_i)$$

$$\sigma_i^2 = \exp\left(\ln(\sigma^2) + \gamma \ln(x_i)\right)$$

$$= \exp(\alpha_1 + \alpha_2 z_i) \qquad (8.18)$$

# 8.3.2 Estimating the Variance Function

$$\sigma_i^2 = \exp(\alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_s z_{iS}) \qquad (8.19)$$

$$\ln(\sigma_i^2) = \alpha_1 + \alpha_2 z_i \qquad (8.20)$$

$$y_i = E(y_i) + e_i = \beta_1 + \beta_2 x_i + e_i$$

# 8.3.2 Estimating the Variance Function

$$\ln(\hat{e}_i^2) = \ln(\sigma_i^2) + v_i = \alpha_1 + \alpha_2 z_i + v_i \qquad (8.21)$$

$$\ln(\hat{\sigma}_i^2) = .9378 + 2.329 z_i$$

$$\hat{\sigma}_i^2 = \exp(\tilde{\alpha}_1 + \alpha_1 z_i)$$

$$\left(\frac{y_i}{\sigma_i}\right) = \beta_1\left(\frac{1}{\sigma_i}\right) + \beta_2\left(\frac{x_i}{\sigma_i}\right) + \left(\frac{e_i}{\sigma_i}\right)$$

# 8.3.2 Estimating the Variance Function

$$\operatorname{var}\left(\frac{e_i}{\sigma_i}\right) = \left(\frac{1}{\sigma_i^2}\right)\operatorname{var}(e_i) = \left(\frac{1}{\sigma_i^2}\right)\sigma_i^2 = 1 \qquad (8.22)$$

$$y_i^* = \left(\frac{y_i}{\sigma_i}\right) \qquad x_{i1}^* = \left(\frac{1}{\sigma_i}\right) \qquad x_{i2}^* = \left(\frac{x_i}{\sigma_i}\right) \qquad (8.23)$$

$$y_i^* = \beta_1 x_{i1}^* + \beta_2 x_{i2}^* + e_i^* \qquad (8.24)$$

# 8.3.2 Estimating the Variance Function

$$y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_k x_{iK} + e_i \quad (8.25)$$

$$\text{var}(e_i) = \sigma_i^2 = \exp(\alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_s z_{iS}) \quad (8.26)$$

# 8.3.2 Estimating the Variance Function

The steps for obtaining a feasible generalized least squares estimator for $\beta_1, \beta_2, \ldots, \beta_K$ are:

- 1. Estimate (8.25) by least squares and compute the squares of the least squares residuals $\hat{e}_i^2$.
- 2. Estimate $\alpha_1, \alpha_2, \ldots, \alpha_S$ by applying least squares to the equation

$$\ln \hat{e}_i^2 = \alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS} + v_i$$

# 8.3.2 Estimating the Variance Function

3. Compute variance estimates $\hat{\sigma}_i^2 = \exp(\hat{\alpha}_1 + \alpha_2 z_{i2} + \cdots + \hat{\alpha}_S z_{iS})$

4. Compute the transformed observations defined by (8.23), including $x_{i3}^*, \ldots, x_{iK}^*$ if $K > 2$.

5. Apply least squares to (8.24), or to an extended version of (8.24) if $K > 2$ .

$$\hat{y}_i = 76.05 + 10.63x \qquad (8.27)$$
$$\text{(se)} \quad (9.71) \quad (.97)$$

# 8.3.3 A Heteroskedastic Partition

$$\widehat{WAGE} = -9.914 + 1.234EDUC + .133EXPER + 1.524METRO \quad (8.28)$$
$$\text{(se)} \quad (1.08) \quad (.070) \quad (.015) \quad (.431)$$

$$WAGE_{Mi} = \beta_{M1} + \beta_2 EDUC_{Mi} + \beta_3 EXPER_{Mi} + e_{Mi} \qquad i = 1, 2, \ldots, N_M \quad (8.29a)$$

$$WAGE_{Ri} = \beta_{R1} + \beta_2 EDUC_{Ri} + \beta_3 EXPER_{Ri} + e_{Ri} \qquad i = 1, 2, \ldots, N_R \quad (8.29b)$$

$$b_{M1} = -9.914 + 1.524 = -8.39$$

# 8.3.3 A Heteroskedastic Partition

$$\operatorname{var}(e_{Mi}) = \sigma_M^2 \qquad \operatorname{var}(e_{Ri}) = \sigma_R^2 \tag{8.30}$$

$$\hat{\sigma}_M^2 = 31.824 \qquad \sigma_R^2 = 15.243$$

$$b_{M1} = -9.052 \qquad b_{M2} = 1.282 \qquad b_{M3} = .1346$$

$$b_{R1} = -6.166 \qquad b_{R2} = .956 \qquad b_{R3} = .1260$$

# 8.3.3 A Heteroskedastic Partition

$$\left(\frac{WAGE_{Mi}}{\sigma_M}\right)=\beta_{M1}\left(\frac{1}{\sigma_M}\right)+\beta_2\left(\frac{EDUC_{Mi}}{\sigma_M}\right)+\beta_3\left(\frac{EXPER_{Mi}}{\sigma_M}\right)+\left(\frac{e_{Mi}}{\sigma_M}\right) \quad i=1,2,\ldots,N_M \tag{8.31a}$$

$$\left(\frac{WAGE_{Ri}}{\sigma_R}\right)=\beta_{R1}\left(\frac{1}{\sigma_R}\right)+\beta_2\left(\frac{EDUC_{Ri}}{\sigma_R}\right)+\beta_3\left(\frac{EXPER_{Ri}}{\sigma_R}\right)+\left(\frac{e_{Ri}}{\sigma_R}\right) \quad i=1,2,\ldots,N_R \tag{8.31b}$$

# 8.3.3 A Heteroskedastic Partition

Feasible generalized least squares:

1. Obtain estimated $\hat{\sigma}_M$ and $\hat{\sigma}_R$ by applying least squares separately to the metropolitan and rural observations.

2. $\hat{\sigma}_i = \begin{cases} \hat{\sigma}_M & \text{when } METRO_i = 1 \\ \hat{\sigma}_R & \text{when } METRO_i = 0 \end{cases}$

3. Apply least squares to the transformed model

$$\left(\frac{WAGE_i}{\hat{\sigma}_i}\right) = \beta_{R1}\left(\frac{1}{\sigma_i}\right) + \beta_2\left(\frac{EDUC_i}{\hat{\sigma}_i}\right) + \beta_3\left(\frac{EXPER_i}{\sigma_i}\right) + \delta\left(\frac{METRO_i}{\hat{\sigma}_i}\right) + \left(\frac{e_i}{\sigma_i}\right) \quad (8.32)$$

# 8.3.3 A Heteroskedastic Partition

$$\widehat{WAGE} = -9.398 + 1.196EDUC + .132EXPER + 1.539METRO \qquad (8.33)$$
$$\text{(se)} \quad (1.02) \quad (.069) \quad (.015) \quad (.346)$$

# 8.3.3 A Heteroskedastic Partition

**Remark:** To implement the generalized least squares estimators described in this Section for three alternative heteroskedastic specifications, an assumption about the form of the heteroskedasticity is required. Using least squares with White standard errors avoids the need to make an assumption about the form of heteroskedasticity, but does not realize the potential efficiency gains from generalized least squares.

# 8.4 Detecting Heteroskedasticity

## 8.4.1 Residual Plots

- Estimate the model using least squares and plot the least squares residuals.
- With more than one explanatory variable, plot the least squares residuals against each explanatory variable, or against $\hat{y}_i$, to see if those residuals vary in a systematic way relative to the specified variable.

## 8.4.2 The Goldfeld-Quandt Test

$$F = \frac{\hat{\sigma}_M^2 / \sigma_M^2}{\hat{\sigma}_R^2 / \sigma_R^2} \square F_{(N_M - K_M, N_R - K_R)} \qquad (8.34)$$

$$H_0 : \sigma_M^2 = \sigma_R^2 \quad \text{against} \quad H_0 : \sigma_M^2 \neq \sigma_R^2 \qquad (8.35)$$

$$F = \frac{\hat{\sigma}_M^2}{\hat{\sigma}_R^2} = \frac{31.824}{15.243} = 2.09$$

# 8.4 Detecting Heteroskedasticity

## 8.4.2 The Goldfeld-Quandt Test

$$\hat{\sigma}_1^2 = 3574.8$$

$$\hat{\sigma}_2^2 = 12{,}921.9$$

$$F = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2} = \frac{12{,}921.9}{3574.8} = 3.61$$

# 8.4 Detecting Heteroskedasticity

## 8.4.3 Testing the Variance Function

$$y_i = E(y_i) + e_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} + e_i \quad (8.36)$$

$$\text{var}(y_i) = \sigma_i^2 = E(e_i^2) = h(\alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS}) \quad (8.37)$$

$$h(\alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS}) = \exp(\alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS})$$

$$h(\alpha_1 + \alpha_2 z_i) = \exp\left(\ln(\sigma^2) + \gamma \ln(x_i)\right)$$

# 8.4 Detecting Heteroskedasticity

## 8.4.3 Testing the Variance Function

$$h(\alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS}) = \alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS} \quad (8.38)$$

$$h(\alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS}) = h(\alpha_1)$$

$$H_0 : \alpha_2 = \alpha_3 = \cdots = \alpha_S = 0$$

$$H_1 : \text{ not all the } \alpha_s \text{ in } H_0 \text{ are zero} \quad (8.39)$$

# 8.4 Detecting Heteroskedasticity

## 8.4.3 Testing the Variance Function

$$\text{var}(y_i) = \sigma_i^2 = E(e_i^2) = \alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS} \quad (8.40)$$

$$e_i^2 = E(e_i^2) + v_i = \alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS} + v_i \quad (8.41)$$

$$\hat{e}_i^2 = \alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS} + v_i \quad (8.42)$$

$$\chi^2 = N \times R^2 \sim \chi^2_{(S-1)} \quad (8.43)$$

# 8.4 Detecting Heteroskedasticity

## *8.4.3a The White Test*

$$E(y_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$$

$$z_2 = x_2 \qquad z_3 = x_3 \qquad z_4 = x_2^2 \qquad z_5 = x_3^2$$

# 8.4 Detecting Heteroskedasticity

## *8.4.3b Testing the Food Expenditure Example*

$$SST = 4,610,749,441 \qquad SSE = 3,759,556,169$$

$$R^2 = 1 - \frac{SSE}{SST} = .1846$$

$$\chi^2 = N \times R^2 = 40 \times .1846 = 7.38$$

$$\chi^2 = N \times R^2 = 40 \times .18888 = 7.555 \qquad p\text{-value} = .023$$

# Keywords

- Breusch-Pagan test
- generalized least squares
- Goldfeld-Quandt test
- heteroskedastic partition
- heteroskedasticity
- heteroskedasticity-consistent standard errors
- homoskedasticity
- Lagrange multiplier test
- mean function
- residual plot
- transformed model
- variance function
- weighted least squares
- White test

# Chapter 8 Appendices

- Appendix 8A Properties of the Least Squares Estimator
- Appendix 8B Variance Function Tests for Heteroskedasticity

# Appendix 8A
# Properties of the Least Squares Estimator

$$y_i = \beta_1 + \beta_2 x_i + e_i$$

$$E(e_i) = 0 \quad \text{var}(e_i) = \sigma_i^2 \quad \text{cov}(e_i, e_j) = 0 \quad (i \neq j)$$

$$b_2 = \beta_2 + \sum w_i e_i \qquad \text{(8A.1)}$$

$$w_i = \frac{x_i - \bar{x}}{\sum \left(x_i - \bar{x}\right)^2}$$

# Appendix 8A
# Properties of the Least Squares Estimator

$$E(b_2) = E(\beta_2) + E\left(\sum w_i e_i\right)$$

$$= \beta_2 + \sum w_i E(e_i) = \beta_2$$

# Appendix 8A
# Properties of the Least Squares Estimator

$$\begin{aligned}
\operatorname{var}\left(b_2\right) &= \operatorname{var}\left(\sum w_i e_i\right) \\
&= \sum w_i^2 \operatorname{var}\left(e_i\right) + \sum_{i \neq j}\sum w_i w_j \operatorname{cov}\left(e_i, e_j\right) \\
&= \sum w_i^2 \sigma_i^2 \\
&= \frac{\sum\left[(x_i - \overline{x})^2 \sigma_i^2\right]}{\left[\sum (x_i - \overline{x})^2\right]^2}
\end{aligned} \tag{8A.2}$$

# Appendix 8A
# Properties of the Least Squares Estimator

$$\text{var}(b_2) \neq \frac{\sigma^2}{\sum\left(x_i - \bar{x}\right)^2} \qquad (8A.3)$$

# Appendix 8B
# Variance Function Tests for Heteroskedasticity

$$\hat{e}_i^2 = \alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS} + v_i \qquad (8B.1)$$

$$F = \frac{(SST - SSE)/(S-1)}{SSE/(N-S)} \qquad (8B.2)$$

$$SST = \sum_{i=1}^{N} \left( \hat{e}_i^2 - \overline{\hat{e}^2} \right)^2 \quad \text{and} \quad SSE = \sum_{i=1}^{N} \hat{v}_i^2$$

# Appendix 8B
# Variance Function Tests for Heteroskedasticity

$$\chi^2 = (S-1) \times F = \frac{SST - SSE}{SSE/(N-S)} \sim \chi^2_{(S-1)} \qquad (8B.3)$$

$$\widehat{\text{var}}(e_i^2) = \widehat{\text{var}}(v_i) = \frac{SSE}{N-S} \qquad (8B.4)$$

$$\chi^2 = \frac{SST - SSE}{\widehat{\text{var}}(e_i^2)} \qquad (8B.5)$$

# Appendix 8B
# Variance Function Tests for Heteroskedasticity

$$\chi^2 = \frac{SST - SSE}{2\hat{\sigma}_e^4} \qquad (8B.6)$$

$$\operatorname{var}\left(\frac{e_i^2}{\sigma_e^2}\right) = 2 \qquad \frac{1}{\sigma_e^4}\operatorname{var}(e_i^2) = 2 \qquad \operatorname{var}(e_i^2) = 2\sigma_e^4$$

$$\widehat{\operatorname{var}}(e_i^2) = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{e}_i^2 - \overline{\hat{e}^2}\right)^2 = \frac{SST}{N} \qquad (8B.7)$$

# Appendix 8B
# Variance Function Tests for Heteroskedasticity

$$\chi^2 = \frac{SST - SSE}{SST / N}$$

$$= N \times \left(1 - \frac{SSE}{SST}\right) \qquad (8B.8)$$

$$= N \times R^2$$