

## Ch: 8 Heteroskedasticity

$$\text{Var}(u | x_1, \dots, x_k) \neq \sigma^2$$

but depends on  $x_j$

$$\text{price} = \beta_0 + \beta_1 \text{sq. ft.} + \beta_2 \text{lot size} + \beta_3 \text{bdrms} + u$$

(e.g. quality)

$\text{Var}(u | \text{lot size, sq. ft., bdrms})$  depends on sq. ft.

Under ass<sup>n</sup>s for unbiasedness of  $\hat{\beta}_j \rightarrow$  unbiased for  $\beta_j$ .

Usual std. errors & test statistics  $\rightarrow$  no longer valid.

# Heteroskedasticity Robust Inference

Heteroskedasticity of unknown form.

" robust std. errors available in case of large samples.

Simple regression:

$$se(\hat{\beta}_1) = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 \hat{u}_i^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

→  $SST_x$

$$\hat{u} = y - \hat{y}$$

↳  $\hat{\beta}_0 + \hat{\beta}_1 x$

Does  $\hat{u}$  change with robust SEs? No.

# Multiple regression

$$\text{score} = \beta_0 + \beta_1 \text{ class size}$$

$$+ \beta_2 \text{ PC stud. exp.}$$

$$+ \beta_3 \text{ \% minority}$$

$$+ u$$

$$se(\hat{\beta}_j) = \frac{\sqrt{\sum_{i=1}^n \hat{\epsilon}_{ij}^2}}{\sqrt{SST_j (1 - R_j^2)}}$$

$\hat{\beta}_j$

$$\sum_{i=1}^n (\epsilon_{ij} - \bar{\epsilon}_j)^2$$

$$\sum_{i=1}^n (\text{class}_i - \overline{\text{class}})^2$$

$R^2$  from reg. of  $\epsilon_j$  on all other  $\epsilon$ 's.

from reg. of class. on PC stud. exp. & \% minority

$\hat{\epsilon}_{ij}$  : residual from reg. of  $\epsilon_j$  on other  $\epsilon$ 's for obs.  $i$   
 ← class on PC stud. exp. & \% minority

# Testing

can test whether

$\hat{u}^2$  depends on  $x_1 \dots x_k$   
 $x_1^2, x_2^2 \dots$   
 $x_1 * x_2, x_1 * x_3 \dots$   
 $\hat{y}$

## Weighted least squares

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + u$$

$$\text{var}(u | x_1, \dots, x_k) = \sigma^2 h$$



Suppose we know the heterosk. f<sub>u</sub>: h

e.g.  $\exp(\beta_1 x_1 + \beta_2 x_2)$

$$\sqrt{x_1 \cdot x_2}$$

Estimate transformed model

$$\frac{y}{\sqrt{h}} = \beta_0 \cdot \frac{1}{\sqrt{h}} + \beta_1 \frac{x_1}{\sqrt{h}} + \dots + \beta_k \frac{x_k}{\sqrt{h}} + \frac{u}{\sqrt{h}}$$

$$\text{var}\left(\frac{u}{\sqrt{h}} \mid x_1, \dots, x_k\right) = \frac{\sigma^2 h}{h} = \sigma^2$$

Pooled cross sections → indep. cross-sections  
pooled over time

Panel data →  
Same units observed  
over time

Ch: 13

Pooling <sup>Indep.</sup> Cross - Sections over Time

Data: random samples from a pop. at diff.  
pts. in time

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \delta_0 d_2 + \delta_1 d_3 + \dots + \delta_{T-2} d_T + u$$

Convert nominal values (e.g. wages) to real values.

$$\begin{array}{r}
 \$350 \\
 - \$200 \\
 \hline
 \$150
 \end{array}$$

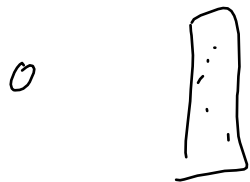
Diff. in T  
 - Diff. in C

$$\begin{array}{r}
 \$250 \\
 \boxed{\$130 \text{ less}} \\
 - \$100 \\
 \hline
 \$20
 \end{array}$$

How incinerator affects property values?

$$= \$20 - \$150$$

$$= -\$130$$



0

0

0