Ch: 8 Heteroskedasticity

$Var(u \mid x_1 \ldots x_k) \neq \sigma^2$

but depends on x_j

$price = \beta_0 + \beta_1 \, sq.ft. + \beta_2 \, lotsize + \beta_3 \, bdrms + u$

(e.g. quality)

$var(u \mid lotsize, sq.ft., bdrms)$ depends on sq.ft.

Under assn.s for unbiasedness of $\hat{\beta}_j \longrightarrow$ unbiased for β_j.

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 errors & test statistics $\longrightarrow$ no longer valid.

Heteroskedasticity Robust Inference

Heteroskedasticity of unknown form.

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

Simple regression:

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

$$\hat{u} = y - \underbrace{\hat{y}}_{\hat{\beta}_0 + \hat{\beta}_1 x}$$

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

Multiple regression

score $= \beta_0 + \beta_1$ class size

$+ \beta_2$ PC stud. exp.

$+ \beta_3$ % minority

$+ u$

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

$\hat{\beta}_1$

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

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

R^2 from reg. of x_j on all other x's.

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

$\hat{r}_{ij}$: residual from reg. of x_j on other x's for obs. i

" class on PC stud. exp. & % minority

Testing

can test whether

$\hat{u}^2$ depends on $x_1 \dots x_k$

$x_1^2, x_2^2 \dots$

$x_1 \times x_2, \; x_1 \times x_3 \dots$

$\hat{y}$

Weighted least squares

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

$$\mathrm{var}(u \mid x_1 \dots x_k) = \sigma^2 h$$

Suppose we know the heterosk. fn h

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

$\sqrt{x_1 \cdot x_2}$

Estimate transformed model

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

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

Pooled cross sections $\longrightarrow$ indep. cross-sections
pooled over time

Panel data $\rightarrow$ same units observed
over time

Ch: 13 Pooling Indep. Cross-Sections Over Time

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

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \delta_0 d2 + \delta_1 d3 + \ldots + \delta_{T-2} dT + u$$

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

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

Diff. in T

− Diff. in C

$250

$\boxed{\$130 \text{ less}}$

$$\begin{array}{r} \$120 \\ - \$100 \\ \hline \$20 \end{array}$$

How incinerator affects property values?

$= \$20 - \$150$

$= -\$130$