Ch. 8 Heteroskedasticity

Consequences:

$$\text{var}(u \mid x_1 \ldots x_k) \neq \sigma^2$$

but depends on x_j

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

(e.g. quality)

$$\text{var}(u \mid \text{lotsize}, \text{sq.ft.}, \text{bdrms})$$

dep. on sq.ft.

under ass^{n}s for unbiasedness, $\hat{\beta}_j$ unbiased for β_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 errors & test statistics $\rightarrow$ no longer valid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:

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

$$SST_x = \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Multiple regression

score $= \beta_0 + \beta_1$ class-size $+ \beta_2$ PC stud. exp. $+ \beta_3$ % minority $+ u$

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

e.g. $\hat{\beta}_1$

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

R_j^2 from reg. of x_j (class-size) on all other x's (PC stud. exp. & % minority)

$\hat{r}_{ij}$: residual from reg. of x_j (class-size) on all other x's (PC stud. exp. & % minority) for obs. i

Testing

Model (with ass^{n}s up to those reqd. for unbiasedness of $\hat{\beta}$)

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

Null hyp. of homosk.

$$H_0 : Var(u \mid x_1 \ldots x_k) = \sigma^2$$

or, $$H_0 : E(u^2 \mid x_1 \ldots x_k) = \sigma^2$$

Suppose $$u^2 = \delta_0 + \delta_1 x_1 + \ldots + \delta_k x_k + v$$

$$H_0 : \delta_1 = 0$$
$$\vdots$$
$$\delta_k = 0$$

Can estimate: $$\hat{u}^2 = \delta_0 + \delta_1 x_1 + \ldots + \delta_k x_k + v$$

Joint test : H_0

Can also include quadratics (x_1^2, x_2^2), interactions ($x_1 . x_2$), & fitted values ($\hat{y}$).

Works only in case of large samples.

```
. reg price lotsize sqrft bdrms

      Source |       SS           df       MS      Number of obs   =        88
-------------+----------------------------------   F(3, 84)        =     57.46
       Model |  617130.701         3  205710.234   Prob > F        =    0.0000
    Residual |  300723.805        84   3580.0453   R-squared       =    0.6724
-------------+----------------------------------   Adj R-squared   =    0.6607
       Total |  917854.506        87  10550.0518   Root MSE        =    59.833

------------------------------------------------------------------------------
       price | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
     lotsize |   .0020677   .0006421     3.22   0.002     .0007908    .0033446
       sqrft |   .1227782   .0132374     9.28   0.000     .0964541    .1491022
       bdrms |   13.85252   9.010145     1.54   0.128    -4.065141    31.77018
       _cons |  -21.77031   29.47504    -0.74   0.462    -80.38466    36.84405
------------------------------------------------------------------------------

. predict uhat, res

.
. g uhatsq = uhat^2

.
. reg uhatsq lotsize sqrft bdrms

      Source |       SS           df       MS      Number of obs   =        88
-------------+----------------------------------   F(3, 84)        =      5.34
       Model |   701213780         3   233737927   Prob > F        =    0.0020
    Residual |  3.6775e+09        84  43780003.5   R-squared       =    0.1601
-------------+----------------------------------   Adj R-squared   =    0.1301
       Total |  4.3787e+09        87  50330276.7   Root MSE        =    6616.6

------------------------------------------------------------------------------
      uhatsq | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
     lotsize |   .2015209   .0710091     2.84   0.006     .0603116    .3427302
       sqrft |   1.691037    1.46385     1.16   0.251    -1.219989    4.602063
       bdrms |    1041.76    996.381     1.05   0.299    -939.6526    3023.173
       _cons |  -5522.795   3259.478    -1.69   0.094    -12004.62    959.0348
------------------------------------------------------------------------------

.
. test lotsize sqrft bdrms

 ( 1)  lotsize = 0
 ( 2)  sqrft = 0
 ( 3)  bdrms = 0

       F(  3,    84) =    5.34
            Prob > F =    0.0020
```

Weighted Least Squares

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

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

Transformed model :

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

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

Generalized / Weighted Least Squares (GLS/ WLS).

Feasible / Estimated GLS : h is estimated