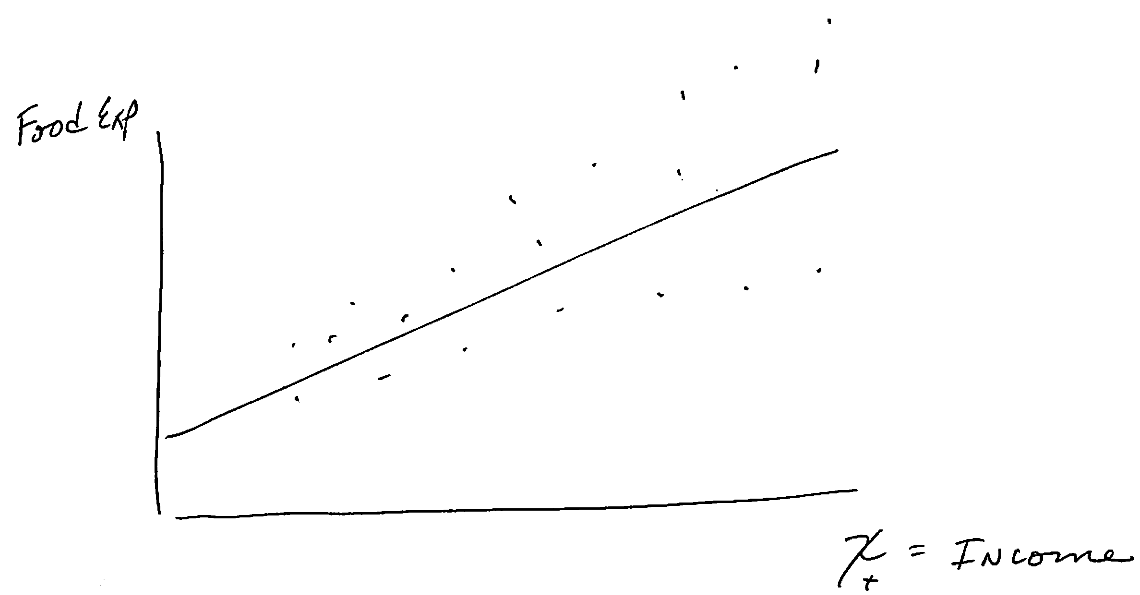
Structures

Heteroscedastic

Problem is that we have more precise information about some obs than others.



There is less variation in food exp among poor than among rich.

LS ignores this (nonsample) info.

1) LS is UNBIASED, but NOT BLUE. a more eff. estimator would use the info that info on poor people is better.

2) std errors, tvals, C.I. ~~are~~ based on OLS are incorrect.

Strategy

1) Use OLS, but modify your estimator of the Var-Cov matrix so that it yields consistent std errors, t-vals, C.I. and F tests.

OLS is unbiased. OLS cov is goofed up. strategy I says use OLS with a non-goofed up cov matrix.

White's Estimator

user let $W = \text{diag}[\hat{e}_t^2]$ $\quad \hat{e}_t = y_t - x_t' \underset{\sim}{b}$.

$$\text{White} = (X'X)^{-1} X' W_T X (X'X)^{-1}$$

instead of $\hat{\sigma}^2 (X'X)^{-1}$ to est LS cov.

$$\text{plim}(\text{white}) = \text{Cov}(\underset{\sim}{b})$$

Food Exp:

$$\hat{y}_t = \underset{\substack{(23.704) \\ (22.139)}}{40.768} + \underset{\substack{(.0382) \\ (.0305)}}{.1283}\, x_t$$

white

Incorrect "usual" OLS estimate

Strategy 2

Model the variance as a function of explanatory vars. - Reweight the OBS based on this model -

WLS - GLS.

$$y_t = \beta_1 + \beta_2 x_{t2} + e_t \qquad e_t \sim (0, \sigma_t^2)$$

If σ_t are known:

~~Est~~ Reweight OBS. using $\sqrt{\sigma_t^2} = \sigma_t$

$$y_t/\sigma_t = \frac{1}{\sigma_t}\beta_1 + \beta_2 x_{t2}/\sigma_t + e_t/\sigma_t$$

$$\operatorname{var}\left(\frac{e_t}{\sigma_t}\right) = \frac{1}{\sigma_t^2}\operatorname{var}(e_t) = 1$$

Satisfies G-M assumptions ∴ OLS of reweighted model is BLUE.

Requires knowledge of σ_t.

If σ_t known $\Rightarrow$ GLS - BLUE

If σ_t unknown $\Rightarrow$ Estimate it

FGLS (NOT BLUE but asympt more eff than OLS)

Consistent OLS cov using SAS.

```
Proc model;
  Parms b1 - bK;
  y = b1 + b2*x2 + b3*x3 + . . . + bK*xK;
  fit y / GMM;
  Instruments x2 x3 . . . xK;
```

This is not exactly the same as White's estimator but it is asymp. equivalent.

2) Model Het as a function of exogenous variables

$$\sigma_t^2 = \gamma_1 + \gamma_2 z_{t2} + \cdots + \gamma_s z_{ts}$$

$$\sigma_t = \gamma_1 + \gamma_2 z_{t2} + \cdots + \gamma_s z_{ts}$$

$$\sigma_t^2 = \exp\{\gamma_1 + \gamma_2 z_{t2} + \cdots + \gamma_s z_{ts}$$

$z_{t2}, \ldots, z_{ts}$ are exogenous vars thought to cause heteroscedasticity. $\gamma_1, \ldots, \gamma_s$ are unknown parameters.

Essentially, the γ's are estimated and used to estimate $\hat{\sigma}_t$ which is used to reweight the data.

$$\mathrm{Var}(e_t/\sigma_t) = 1$$

Estimation of Heteroskedastic Models

$$y_t = \underset{\sim}{x}_t \underset{\sim}{\beta} + u_t \qquad t = 1, 2, \ldots, n$$

$\underset{\sim}{x}_t \quad 1 \times k$

$\underset{\sim}{\beta} \quad k \times 1$

$$E(u_t \mid x_t) = 0$$
$$Var(u_t \mid x_t) = \sigma_t^2$$
$$Cov(u_t, u_s) = 0 \qquad t \neq s.$$

$$\Rightarrow \underset{\sim}{\Omega} = diag(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$$

3 General Models

(1) $\sigma_t^2 = \underset{\sim}{z}_t \underset{\sim}{\alpha}$

$$z_t = \{1 \quad z_{t1} \quad \cdots \quad z_{tr}\}$$

Note: $\cancel{k = r+1}$

$$\sigma_t^2 = \alpha_0 + \alpha_1 z_{t1} + \cdots + \alpha_r z_{tr}$$

Homoskedastic if $\alpha_1 = \alpha_2 = \cdots = \alpha_r = 0$

$$\sigma_t^2 = \alpha_0$$

(2) $\sigma_t^2 = \{z_t \alpha\}^2 \Rightarrow \sigma_t = \underset{\sim}{z}_t \hat{\alpha}$

(3) $\sigma_t^2 = \exp\{z_t \alpha\}$

$= \exp\{\alpha_0 + \alpha_1 z_{t1} + \cdots + \alpha_r z_{tr}\}$

$= \sigma^2 \exp\{z_{t1}\alpha_1 + \cdots + \alpha_r z_{tr}\}$

Model 1

let $D_1 = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_m^2 \end{bmatrix}$ $D^{-1} = \begin{bmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & 1/\sigma_m^2 \end{bmatrix}$

$$\hat{\beta}_{GLS} = (X^T D_1^{-1} X)^{-1} X^T D_1^{-1} \underset{\sim}{y}$$

If $\underset{\sim}{\alpha}$ known then

$$= \left(\sum_{t=1}^{n} (\underset{\sim}{z}_t \underset{\sim}{\alpha})^{-1} x_t^T x_t \right)^{-1} \sum_{t=1}^{n} (\underset{\sim}{z}_t \underset{\sim}{\alpha})^{-1} x_t^T y_t$$

If $\underset{\sim}{\alpha}$ unknown then you'll have to estimate it.

$$\sigma_t^2 = \underset{\sim}{z}_t \underset{\sim}{\alpha}$$

$$\hat{e}_t^2 - \hat{e}_t^2 = \underset{\sim}{z}_t \underset{\sim}{\alpha} - \sigma_t^2$$

$$\hat{e}_t^2 = \underset{\sim}{z}_t \underset{\sim}{\alpha} + (\hat{e}_t^2 - \sigma_t^2)$$

$$\hat{e}_t^2 = \underset{\sim}{z}_t \underset{\sim}{\alpha} + v_t$$

$E(v_t) \neq 0$ But $\operatorname{plim}(v_t) = 0$

(a) LS $\hat{\alpha} = \left(\sum_{t=1}^{n} (z_t^T z_t) \right)^{-1} \sum_{t=1}^{m} z_t^T \hat{e}_t^2$

$\operatorname{plim}(\hat{\alpha}) = \alpha$ $\therefore$ $\hat{D}_1$ leads to FGLS

Note: $Var(v_t) = 2\sigma_t^4$ (not const).

$\Rightarrow$ we can obtain a more efficient estimate of α than LS.

Let $V_1 = diag(2\sigma_1^4 \; 2\sigma_2^4 \; \cdots \; 2\sigma_n^4)$

$$\hat{\hat{\alpha}} = \left(\sum_{t=1}^{n} (2\hat{\sigma}_t^4)^{-1} z_t^T z_t\right)^{-1} \sum_{t=1}^{n} (2\hat{\sigma}_t^4)^{-1} z_t^T \hat{u}_t^2$$

Note: '2' drops out.

$$\hat{\sigma}_t^2 = (\underset{\sim}{z}_t \hat{\underset{\sim}{\alpha}})$$

$$= \left[\sum_{t=1}^{n} (\underset{\sim}{z}_t \hat{\underset{\sim}{\alpha}})^{-1/2} z_t^T z_t\right]^{-1} \sum_{t=1}^{n} (\underset{\sim}{z}_t \hat{\underset{\sim}{\alpha}})^{-1/2} z_t^T \hat{u}_t^2$$

Model 3

$$\sigma_t^2 = \exp\{\underset{\sim}{z}_t \underset{\sim}{\alpha}\}$$

$$= \exp\{\alpha_0\} \exp\{z_t^* \alpha^*\}$$

$$= \sigma^2 \exp\{z_t^* \alpha^*\}$$

$$z_t^* = \{z_{t1} \; \cdots \; z_{tr}\}$$

$$\alpha^* = (\alpha_1, \ldots, \alpha_r)^T$$

Let
$$D_3 = \sigma^2 \Delta = \sigma^2 \begin{bmatrix} \exp\{z_1^\circ \alpha^\circ\} & & \\ & \ddots & \\ & & \exp\{z_n^\circ \alpha^\circ\} \end{bmatrix}$$

$$\hat{\beta}_{GLS} = (X^T D_3^{-1} X)^{-1} X^T D_3^{-1} y$$
$$= \left(\sum_{t=1}^{n} (\exp\{z_t^\circ \alpha^\circ\})^{-1} X_t^T X_t \right)^{-1}$$
$$\sum_{t=1}^{n} \exp\{z_t^\circ \alpha^\circ\}^{-1} X_t^T y_t$$

<u>Est α</u>

$$\ln \sigma_t^2 = \alpha_0 + \alpha_1 z_{t1} + \cdots + \alpha_r z_{tr}$$

$$\ln(\hat{u}_t^2) = z_t \alpha + v_i$$

$$v_t = \ln(\hat{u}_t^2) - \ln(\sigma_t^2)$$
$$v_t = \ln\left(\frac{\hat{u}_t^2}{\sigma_t^2}\right)$$

If $u_t \sim N(0, \sigma_t^2)$, $\frac{u_t}{\sigma_t} \sim N(0,1)$

$$v_t = \ln\left(\frac{u_t}{\sigma_t}\right)^2 \sim \ln(\chi_1^2)$$

$E(V_t) = -1.2704$

$Var(V_t) = 4.9348$

Two implications

1) LS estimate of α_0
Biased By -1.2704

2) usual OLS est of α
works (since V_t homoscedastic.

$$\hat{\alpha} = (Z^T Z)^{-1} Z^T \underset{\sim}{g}$$

$$\underset{\sim}{g} = \begin{pmatrix} \ln(\hat{u}_1^2) \\ \ln(\hat{u}_2^2) \\ \vdots \\ \ln(\hat{u}_n^2) \end{pmatrix}$$

$$\hat{\hat{\alpha}} = \hat{\alpha} + \begin{pmatrix} 1.2704 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad p\lim(\hat{\hat{\alpha}}) = \alpha$$

Structures

Heteroscedasticity

$$\sigma_i^2 = \exp\{\gamma_1 + \gamma_2 z_{i2} + \cdots + \gamma_s z_{is}\}$$

$$= \sigma^2 \exp\{\gamma_2 z_{i2} + \cdots + \gamma_s z_{is}\}$$

where $\exp\{\gamma_1\} = \sigma^2$

Estimate $\gamma_2, \gamma_3, \ldots, \gamma_s$

$$\hat{\sigma}_i^2 = \sigma^2 \exp\{\hat{\gamma}_2 z_{i2} + \cdots + \hat{\gamma}_s z_{is}\}$$

Form $\hat{\Omega} = \begin{bmatrix} \hat{\sigma}_1^2 & 0 & & 0 \\ 0 & \hat{\sigma}_2^2 & \ddots & 0 \\ 0 & & 0 & \hat{\sigma}_n^2 \end{bmatrix}$

Do GLS.

Asympt. more efficient than OLS with HCCME.