

# Econometria

## Tema 6: Heteroscedasticitat

Ramon Alemany

Grau Estadística UB-UPC

**Curs 2017-18**

# Presentació

- 1 Bibliografia
- 2 Definició i causes
- 3 Estimació amb heteroscedasticitat: MQO versus MQG
- 4 Detecció de l'heteroscedasticitat

# Bibliografia

- GREENE, W. (1999)  
**Análisis econométrico. 3a Ed.**  
*Capítol 12*
- WOOLDRIDGE, J. (2009)  
**Introducción a la Econometría. Un enfoque moderno. 4a Ed.**  
*Capítol 8*
- STOCK, J. & WATSON, M. (2012)  
**Introducción a la Econometría. 3a Ed.**  
*Capítol 18*

# Definició i causes

## 1. Definició d'Heteroscedasticitat

- El terme de pertorbació és heteroscedàstic quan la variància del terme de pertorbació no és constant.
- Els elements de la diagonal principal de la matriu de variàncies i covariàncies no són iguals entre sí.

$$\text{Var}(U_i) = \sigma_i^2 = \sigma^2 \delta_i$$
$$\text{Var}(U) = E(UU') = \sigma^2 \Omega = \sigma^2 \begin{bmatrix} \delta_1 & 0 & \dots & 0 \\ 0 & \delta_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \delta_N \end{bmatrix}$$

# Definició i causes

## 2. Causes de l'heterosedasticitat

- Teòriques: treballem amb dades de tall transversal amb unitats mostrals que presenten un comportament molt heterogeni.
- Mostrals: relacionades amb la manera com recollim la informació (valors mitjans o agregats per a submostres de la mostra global).
- Espúries: incompliment d'altres supòsits bàsics com l'omissió de variables rellevants o la presència de canvis estructurals.

# Definició i causes

## Dades agregades: N observacions en M grups

$$\Omega = \begin{bmatrix} N_1 & 0 & \dots & 0 \\ 0 & N_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & N_m \end{bmatrix}$$

$$P = \begin{bmatrix} \sqrt{N_1} & 0 & \dots & 0 \\ 0 & \sqrt{N_2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \sqrt{N_m} \end{bmatrix} \quad T = P^{-1} = \begin{bmatrix} \frac{1}{\sqrt{N_1}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{N_2}} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{N_m}} \end{bmatrix}$$

# Definició i causes

## Dades en mitjanes: N observacions resumides en M grups

$$\Omega = \begin{bmatrix} \frac{1}{N_1} & 0 & \dots & 0 \\ 0 & \frac{1}{N_2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \frac{1}{N_m} \end{bmatrix}$$

$$P = \begin{bmatrix} \frac{1}{\sqrt{N_1}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{N_2}} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{N_m}} \end{bmatrix} \quad T = P^{-1} = \begin{bmatrix} \sqrt{N_1} & 0 & \dots & 0 \\ 0 & \sqrt{N_2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \sqrt{N_m} \end{bmatrix}$$

# Estimació amb heteroscedasticitat: MQO versus MQG

## 3. Estimació amb heteroscedasticitat

### a) Estimadors MQO

- Propietats
- Relació asimptòtica entre l'estimació MQO de la variància dels estimadors i la variància dels estimadors.
- L'estimador robust de White

### b) Estimadors MQG

- $\Omega$  coneguda
- $\Omega$  desconeguda



# Estimació amb heteroscedasticitat: MQO versus MQG

## Estimadors MQO

### Propietats $\hat{\beta}_{\text{MQO}}$

- No esbiaixats
- Consistents
- INEFICIENTS

$$\text{Var}(\hat{\beta}_{\text{MQO}}) = \sigma_u^2 (X'X)^{-1} X' \Omega X (X'X)^{-1} \geq \sigma_u^2 (X'X)^{-1}$$

# Estimació amb heteroscedasticitat: MQO versus MQG

## Estimadors MQO

Així doncs, l'habitual estimador  $\widehat{\text{Var}}(\hat{\beta}) = \hat{\sigma}_u^2(X'X)^{-1}$  és **esbiaixat**. Per tant, els contrastos  $t$  i  $F$  que es fonamenten en aquest estimador ja no seran vàlids.

## Estimador Robust de White

L'estimador MQO es distribueix asimptòticament (per mostres grans) com una normal:

$$\hat{\beta}_{\text{MQO}} \overset{A}{\sim} N(\beta, \text{Var}(\hat{\beta}))$$

i la variança dels estimadors es pot estimar consistentment com:

$$\begin{aligned}\widehat{\text{Var}}(\hat{\beta}_{\text{MQO}}) &= \hat{\sigma}_u^2(X'X)^{-1}X'\hat{\Sigma}X(X'X)^{-1} \\ \hat{\Sigma} &= \text{diag}(e_i^2)\end{aligned}$$

# Estimació amb heteroscedasticitat: MQO versus MQG

## Estimadors MQG

### Amb $\Omega$ coneguda

- 1a alternativa: transformació del model

$$\begin{array}{lcl} Y = X\beta + U & \longrightarrow & Y^* = X^*\beta + U^* \\ U \sim N(0, \sigma^2\Omega) & & U^* \sim N(0, \sigma^2 I) \end{array} \quad \begin{array}{l} Y^* = TY \\ X^* = TX \\ U^* = TU \end{array}$$

$$\hat{\beta}_{\text{MQG}} = (X^{*'}X^*)^{-1}(X^{*'}Y^*)$$

- 2a alternativa: estimació directa

$$\hat{\beta}_{\text{MQG}} = (X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}Y)$$

# Estimació amb heteroscedasticitat: MQO versus MQG

## Estimadors MQG: Transformació del model

$$Y = X\beta + U$$

$$\sigma_i^2 = \text{Var}(U_i) = \sigma^2 h_i$$

$$\Omega = \begin{bmatrix} h_1 & 0 & \dots & 0 \\ 0 & h_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & h_N \end{bmatrix} \quad P = \begin{bmatrix} \sqrt{h_1} & 0 & \dots & 0 \\ 0 & \sqrt{h_2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \sqrt{h_N} \end{bmatrix}$$

$$T = P^{-1} = \begin{bmatrix} 1/\sqrt{h_1} & 0 & \dots & 0 \\ 0 & 1/\sqrt{h_2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1/\sqrt{h_N} \end{bmatrix} \quad TY = \begin{bmatrix} Y_1/\sqrt{h_1} \\ Y_2/\sqrt{h_2} \\ \vdots \\ Y_N/\sqrt{h_N} \end{bmatrix}$$

$$TX = \begin{bmatrix} 1/\sqrt{h_1} & x_{21}/\sqrt{h_1} & \dots & x_{k1}/\sqrt{h_1} \\ 1/\sqrt{h_2} & x_{22}/\sqrt{h_2} & \dots & x_{k2}/\sqrt{h_2} \\ \vdots & \vdots & \dots & \vdots \\ 1/\sqrt{h_N} & x_{2N}/\sqrt{h_N} & \dots & x_{kN}/\sqrt{h_N} \end{bmatrix}$$

# Estimació amb heteroscedasticitat: MQO versus MQG

## Estimadors MQG: Transformació del model

$$\frac{y_i}{\sqrt{h_i}} = \beta_0 \frac{1}{\sqrt{h_i}} + \beta_1 \frac{x_{1i}}{\sqrt{h_i}} + \dots + \beta_k \frac{x_{ki}}{\sqrt{h_i}} + \frac{u_i}{\sqrt{h_i}}$$
$$E \left[ \left( \frac{u_i}{\sqrt{h_i}} \right)^2 \right] = E \left[ \frac{u_i^2}{h_i} \right] = \sigma^2 \frac{h_i}{h_i} = \sigma^2$$

### Supòsit 1:

Heteroscedasticitat provocada per una variable aliena al model

$$\text{Var}(U_i) = \sigma_i^2 = \lambda^2 Z_i^2 \quad h_i = Z_i^2 \quad \text{weight} = \frac{1}{Z_i^2}$$

$$\frac{y_i}{Z_i} = \beta_0 \frac{1}{Z_i} + \beta_1 \frac{x_{1i}}{Z_i} + \dots + \beta_k \frac{x_{ki}}{Z_i} + \frac{u_i}{Z_i}$$

# Estimació amb heteroscedasticitat: MQO versus MQG

## Estimadors MQG: Transformació del model

### Supòsit 2:

Heteroscedasticitat provocada pel quadrat d'una variable explicativa del model

$$\text{Var}(U_i) = \sigma_i^2 = \lambda^2 x_{1i}^2 \quad h_i = x_{1i}^2 \quad \text{weight} = \frac{1}{x_{1i}^2}$$

$$\frac{y_i}{x_{1i}} = \beta_0 \frac{1}{x_{1i}} + \beta_1 \frac{x_{1i}}{x_{1i}} + \dots + \beta_k \frac{x_{ki}}{x_{1i}} + \frac{u_i}{x_{1i}}$$

$$\frac{y_i}{x_{1i}} = \beta_0 \frac{1}{x_{1i}} + \beta_1 + \dots + \beta_k \frac{x_{ki}}{x_{1i}} + \frac{u_i}{x_{1i}}$$

# Estimació amb heteroscedasticitat: MQO versus MQG

## Estimadors MQG: Transformació del model

### Supòsit 3:

Heteroscedasticitat provocada per una variable explicativa del model

$$\text{Var}(U_i) = \sigma_i^2 = \lambda^2 x_{1i} \quad h_i = x_{1i} \quad \text{weight} = \frac{1}{x_{1i}}$$

$$\frac{y_i}{\sqrt{x_{1i}}} = \beta_0 \frac{1}{\sqrt{x_{1i}}} + \beta_1 \frac{x_{1i}}{\sqrt{x_{1i}}} + \dots + \beta_k \frac{x_{ki}}{\sqrt{x_{1i}}} + \frac{u_i}{\sqrt{x_{1i}}}$$

# Estimació amb heteroscedasticitat: MQO versus MQG

## Estimadors MQG

### Amb $\Omega$ desconeguda

$$\hat{\beta}_{\text{MQGF}} \implies \hat{\Omega} = ? \implies \text{Var}(U_i) = \sigma_i^2 = \sigma^2 \delta_i$$



N paràmetres desconeguts



Supòsit: Esquemes de dependència funcional de la variància

$$\text{Var}(U_i) = \sigma_i^2 = \alpha' Z_i \implies Z_i \text{ variables (explicatives) - contrastos}$$



# Estimació amb heteroscedasticitat: MQO versus MQG

## Estimadors MQGF

### Supòsit 4:

Heteroscedasticitat multiplicativa

$$\text{Var}(U_i | X_i) = \sigma^2 \exp(\delta_0 + \delta_1 x_{1i} + \dots \delta_k x_{ki}) = \sigma^2 h(x)$$

$$U_i^2 = \sigma^2 \exp(\delta_0 + \delta_1 x_{1i} + \dots \delta_k x_{ki}) \cdot V_i$$

amb  $\left\{ E(V) = 1 \right\}$

$$\ln(U_i^2) = \alpha_0 + \delta_1 x_{1i} + \dots \delta_k x_{ki} + e_i$$

amb  $\left\{ (\alpha_0 = \ln \sigma^2 + \delta_0) \quad E(e) = 0 \right\}$

# Estimació amb heteroscedasticitat: MQO versus MQG

## Estimadors MQGF

Usant els residus del model ( $\hat{u}_i^2$ ):

$$\ln(\hat{u}_i^2) = \alpha_0 + \delta_1 x_{1i} + \dots \delta_k x_{ki} + e_i$$

amb  $\left\{ (\alpha_0 = \ln \sigma^2 + \delta_0) \quad E(e) = 0 \right\}$

El valor ajustat de la regressió auxiliar serà:

$$\hat{g}_i = \widehat{\ln(\hat{u}_i^2)} \implies \hat{h}_i = \exp(\hat{g}_i) \quad \text{weight} = \frac{1}{\hat{h}_i}$$

El model transformat serà:

$$\frac{y_i}{\sqrt{\hat{h}_i}} = \beta_0 \frac{1}{\sqrt{\hat{h}_i}} + \beta_1 \frac{x_{1i}}{\sqrt{\hat{h}_i}} + \dots + \beta_k \frac{x_{ki}}{\sqrt{\hat{h}_i}} + \frac{u_i}{\sqrt{\hat{h}_i}}$$

# Estimació amb heteroscedasticitat: MQO versus MQG

## Estimadors MQGF: Propietats

L'estimador Mínim Quadrat Generalitzat Factible (MQGF) és **esbiaixat** però **consistent** i **asimptòticament més eficient** que el de MQO.

Així, per mostres grans l'estimació MQGF és una alternativa millor que el MQO quan existeix evidència d'heteroscedasticitat que “infla” la variància dels estimadors.

# Detecció de l'heteroscedasticitat

## 3.Detecció de l'heteroscedasticitat

1. Mètodes gràfics
2. Contrastos
  - Contrast de White
  - Contrast de Breusch-Pagan
  - Contrast de Goldfeld-Quandt

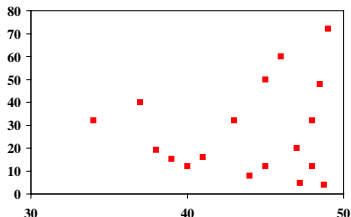
Els estimadors  $\hat{\beta}_{MQO}$  són estimadors consistents de  $\beta$  en presència d'heteroscedasticitat i, per tant, els residus MQO són una bona aproximació del terme de pertorbació

# Detecció de l'heteroscedasticitat

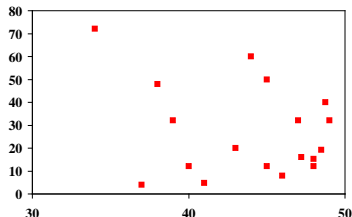
## Mètodes gràfics

### Mètodes gràfics

Consisteix en fer un gràfic de dispersió entre els residus (en valor absolut o elevats al quadrat) i cadascuna de les variables explicatives i observar si hi ha algun tipus de relació.



Proporcionalitat directa



Proporcionalitat inversa

# Detecció de l'heteroscedasticitat

## Contrastos d'heteroscedasticitat

### Contrastos d'heteroscedasticitat

- Tenen dues finalitats:
  - La principal consisteix a identificar la presència o no d'heteroscedasticitat al model analitzat
  - Però quan  $\Omega$  és desconeguda, també poden ajudar a establir esquemes de dependència funcional de la variància.  
$$\text{Var}(U_i) = \sigma_i^2 = \alpha'Z_i$$
- Tots els contrastos amb que treballarem es basen en analitzar el comportament dels residus MQO i les seves  $H_0$  i  $H_A$  es poden resumir com:
  - $H_0$ : homoscedasticitat  $\text{Var}(U_i) = \sigma^2$
  - $H_A$ : heteroscedasticitat  $\text{Var}(U_i) = \sigma_i^2 = \sigma^2\delta_i$

# Detecció de l'heteroscedasticitat

## Contrast de White

### Contrast de White

$$\begin{cases} H_0: \text{homoscedasticitat} & \text{Var}(U_i) = \sigma^2 \\ H_A: \text{heteroscedasticitat} & \text{Var}(U_i) = \sigma_i^2 = \sigma^2 \delta_i \end{cases}$$

### Fases del contrast:

1. Estimar  $Y = X\beta + U$  per MQO

2. Estimar la regressió auxiliar:

$$\begin{aligned} \hat{u}_i^2 = & \alpha_0 + \alpha_1 x_{1i} + \dots + \alpha_k x_{ki} + \\ & + \beta_1 x_{1i} x_{2i} + \dots + \beta_k x_{(k-1)i} x_{ki} + \\ & + \gamma_1 x_{1i}^2 + \dots + \gamma_k x_{ki}^2 + v_i \end{aligned}$$

3. Es calcula l'estadístic de prova:

$$WH = NR^2 \sim \chi_{P-1}^2$$

# Detecció de l'heteroscedasticitat

## Contrast de White

### Característiques i limitacions

- És el més general dels contrastos que veurem (suposa que l'heteroscedasticitat és provocada per totes les variables explicatives).
- Només és vàlid asimptòticament.
- És poc potent davant hipòtesis alternatives més concretes.
- En cas de Rebutjar  $H_0$ , no ofereix cap “guia” i tampoc serveix com a ajuda per establir esquemes de dependència funcional de la variància.



# Detecció de l'heteroscedasticitat

## Contrast de White

### Característiques i limitacions

- Si una o més de les variables explicatives són variables fictícies, aleshores cal anar amb compte amb l'especificació de la regressió auxiliar.

Suposem que  $x_3$  és una variable fictícia. En aquest cas, la variable  $x_3^2$  és el mateix que la variable  $x_3$ . Si incloem totes dues a la regressió auxiliar, llavors hi haurà multicol·linealitat perfecta. Per tant, s'ha d'excloure  $x_3^2$  de la regressió auxiliar.

- El mateix pot succeir quan alguna de les variables explicatives s'especifica al quadrat o al cub en el model per recollir possibles no linealitats.

# Detecció de l'heteroscedasticitat

## Contrast de White

### Característiques i limitacions

- Si tenim un gran nombre de variables explicatives en el model aleshores el nombre d'explicatives en la regressió auxiliar podria superar el nombre d'observacions.

En aquest cas, s'han d'excloure algunes variables de la regressió auxiliar. Es podrien excloure els termes lineals i/o els termes de productes creuats, però sempre s'han de mantenir els termes al quadrat de la regressió auxiliar.

# Detecció de l'heteroscedasticitat

## Contrast de Breusch-Pagan

### Contrast de Breusch-Pagan

En aquest cas la hipotesi alternativa per a l'heteroscedasticitat s'especifica com a funció d'un conjunt de variables.

$$\text{Heteroscedasticitat} \longrightarrow \text{Var}(U_i) = \sigma_i^2 = \sigma^2 f(\gamma_0 + Z_i \gamma_i)$$

Per tant les hipòtesis del contrast són ara:

$$\begin{cases} H_0: \text{homoscedasticitat} & \text{Var}(U_i) = \sigma^2 f(\gamma_0) \\ H_A: \text{heteroscedasticitat} & \text{Var}(U_i) = \sigma_i^2 = \sigma^2 f(\gamma_0 + Z_i \gamma_i) \end{cases}$$

El que serà equivalent a contrastar la significació estadística dels paràmetres  $\gamma$ .

# Detecció de l'heteroscedasticitat

## Contrast de Breusch-Pagan

### Fases del contrast:

1. Estimar  $Y = X\beta + U$  per MQO i obtenir  $\hat{u}_i$ .
2. Calcular el residu normalitzat  $\hat{u}_i^2 / \hat{\sigma}^2$  amb  $\hat{\sigma}^2 = \hat{u}'\hat{u} / N$
3. Estimar la regressió auxiliar:

$$\hat{u}_i^2 / \hat{\sigma}^2 = \gamma_0 + \gamma_1 Z_{1i} + \gamma_2 Z_{2i} + \dots + \gamma_p Z_{pi} + v_i$$

4. Es calcula l'estadístic de prova:

$$BP = \frac{SQR}{2} \sim \chi_p^2$$

# Detecció de l'heteroscedasticitat

## Contrast de Breusch-Pagan

### Característiques i limitacions

- Ofereix informació per a establir esquemes de dependència funcional de la variància.
- Només és vàlid asimptòticament.
- Pot no detectar heteroscedasticitat, tot i havent-hi, si les  $Z_i$  especificades a la regressió auxiliar no són les variables causants.

# Detecció de l'heteroscedasticitat

## Contrast de Goldfeld-Quandt

### Contrast de Goldfeld-Quandt

El contrast de Goldfeld-Quandt suposa que l'heteroscedasticitat és provocada per una variable d'entre les explicatives del model.

$$\text{Heteroscedasticitat} \longrightarrow \text{Var}(U_i) = \sigma_i^2 = \sigma^2 f(\omega_i) \quad \omega_i \in X$$

Per tant les hipòtesis del contrast són ara:

$$\left\{ \begin{array}{l} H_0: \text{homoscedasticitat} \quad \text{Var}(U_i) = \sigma^2 \\ H_A: \text{heteroscedasticitat} \quad \text{Var}(U_i) = \sigma_i^2 = \sigma^2 f(\omega_i) \end{array} \right.$$

# Detecció de l'heteroscedasticitat

## Contrast de Goldfeld-Quandt

### Fases del contrast:

1. Ordenar les observacions d' $Y$  i  $X$  en funció dels valors de  $\omega_i$ .
2. S'eliminen les  $P$  observacions centrals (amb  $P=N/3$  o  $P=N/4$ ).
3. S'estima el model amb les:
  - $(N - P)/2$  primeres observacions  $Y^1 = X^1\beta^1 + U^1$   $SQE^1 = \sum(\hat{u}^1)^2$
  - $(N - P)/2$  darreres observacions  $Y^3 = X^3\beta^3 + U^3$   $SQE^3 = \sum(\hat{u}^3)^2$
4. Es calcula l'estadístic de prova:

$$GQ = \frac{SQE^3}{SQE^1} \sim F_{\frac{N-P}{2}-k, \frac{N-P}{2}-k}$$

# Detecció de l'heteroscedasticitat

## Contrast de Goldfeld-Quandt

### Característiques i limitacions

- Subjectivitat per fixar el valor de  $P$ .
- En cas de No Rebutjar  $H_0$ , no podem afirmar que hi hagi homoscedasticitat amb tota seguretat, ja que pot haver-hi altres variables diferents d' $\omega_i$  que podrien causar-la.
- En cas d'haver-hi una proporcionalitat inversa, caldrà definir l'estadístic de prova intercanviant numerador i denominador.



# Econometria

## Tema 6: Heteroscedasticitat

Ramon Alemany

Grau Estadística UB-UPC

**Curs 2017-18**