Chapitre 4 : Non respect des hypothèses du modèle de régression multiple

Partie 4-1 : Hétéroscédasticité

A. Fadhuile (adelaide.fadhuile@univ-grenoble-alpes.fr)
Univ Grenoble Alpes

Année 2023-2024

Cadre général du Chapitre 4

- Considérons le modèle : $y = X\beta + u$, avec les hypothèses suivantes vérifiées
  - H_1 : $E[u] = 0, \forall\ t$ l'espérance mathématique de l'erreur est nulle
  - H_2 : $V[u] = E(uu^T) = \sigma_u^2 I_N$ la variance de l'erreur est constante
  - H_3 : matrice $\mathbf{X}$ est non-aléatoire
  - H_4 : le modéle est correctement spécifié
  - H_5 : la matrice $\mathbf{X}$ est de plein rang : $k + 1 < T$
- Quel est l'impact du non respect de $H2$?
  - en termes de biais : a. biaisé ; b. sans biais ; c. Je ne sais pas
  - de variance : a. minimale ; b. biaisée ; c. Je ne sais pas

- Partie 4-1 : Hétéroscédasticité **H2**
- Partie 4-2 : Autocorrélation **H2**
- Partie 4-3 : Modèles à erreurs sur les variables **H3**

Plan du cours

2 – Introduction

2.1 – Définition

- H_2 : $V[u] = E(uu^T) = \sigma_u^2 I_N$ la variance de l'erreur est constante
- Homoscédasticité et Hétéroscédasticité
  - Homoscédasticité
    - chaque écart aléatoire est constant ou
    - variance des perturbations est identique $\forall\, N$

    $\Rightarrow$ traduit des **comportements homogènes pour toutes les observations**.
  - Hétéroscédasticité
    - perturbations sont différentes pour toutes les observations

    $\Rightarrow$ reflet effet taille $\Rightarrow$ principalement sur des coupes transversales
- Autocorrélation **Abordé en Partie 4-2**.
  - corrélation entre les perturbations sont supposées non nulles

  $\Rightarrow$ dépendance temporelle entre les observations

$\Rightarrow$ Avec des données réelles, les 2 sont svt combinés

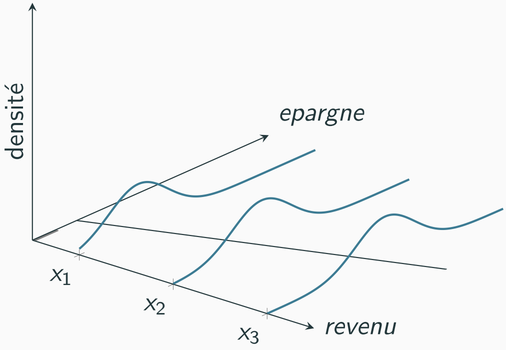
2 – Introduction

2.2 – Illustration graphique

- $V[u] = E(uu^T) = \sigma_u^2 I_N$ la variance de l'erreur est constante
  - homo (égal) "spread" (scédastique)
    - chaque écart aléatoire est constant
    - variance des perturbations est identique $\forall N$

    $\Rightarrow$ traduit des **comportements homogènes pour toutes les observations**.
- Exemple d'une fonction d'épargne : $epargne_i = \beta_0 + \beta_1 revenu_i + \epsilon_i$
  - une augmentation du revenu augmente le niveau d'épargne : la variance du niveau d'épargne serait identique pour tout le niveau de revenu.

Figure 1: Homoscédasticité

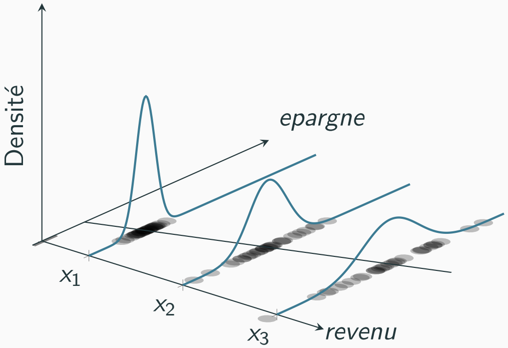


2 – Introduction

2.2 – Illustration graphique

- $V[u] = E(uu^T) = \sigma_u^2 \Omega \neq \sigma_u^2 I_N$ la variance de l'erreur **N'est PAS** constante
  - hétéro (non égal) "spread" (scédastique)
    - perturbations sont différentes pour toutes les observations
    - $\Rightarrow$ reflet effet taille $\Rightarrow$ principalement sur des coupes transversales
- Exemple d'une fonction d'épargne : $epargne_i = \beta_0 + \beta_1 revenu_i + \epsilon_i$
  - une augmentation du revenu augmente le niveau d'épargne : la variance du niveau d'épargne augmente avec le niveau de revenu.

Figure 2: Heteroscédasticité



2 – Introduction

2.3 – Contexte

- Dans quelle situation peut-on rencontrer des problèmes d'hétéroscédasticité ?
  - le plus souvent : avec des données en coupe transversale
  - grande hétérogénéité des valeurs prises par les variables explicatives : effet taille.
- Comment expliquer cela ?
  - Mauvaise spécification, i.e. H4 non respectée ;
  - Présence de valeurs extrêmes ;
  - Même variable expliquée est répétée pour des valeurs différentes d'une variable explicative ;
  - Observations représentent des moyennes calculées sur des échantillon de taille différentes ;
  - Taille des erreurs est liée, de façon proportionnelle, aux valeurs prises par une variable explicative dans le modèle.

2 – Introduction

2.4 – Exemple : Consommation de cigarettes

smoke.gdt /Users/afadhuile/Dropbox/Enseignements/Econometrie/L3...

ID #	Nom de variable	Description
0	const	
1	educ	years of schooling
2	cigpric	state cigarette price, cents per pack
3	white	=1 if white
4	age	in years
5	income	annual income, $
6	cigs	cigs. smoked per day
7	restaurn	=1 if state restaurant smoking restrictions
8	lincome	log(income)
9	agesq	age^2
10	lcigpric	log(cigprice)

Non daté : étendue complète 1 – 807

- Analyse de la consommation de cigarettes
  - Données Wooldridge "smoke.gdt"
  - Nature de la base de données : · · · · · · · · ·
  - Nombre d'observations : · · · · · · · · ·

2 – Introduction

2.4 – Exemple : Consommation de cigarettes

smoke.gdt /Users/afadhuile/Dropbox/Enseignements/Econometrie/L3...

ID #	Nom de variable	Description
0	const	
1	educ	years of schooling
2	cigpric	state cigarette price, cents per pack
3	white	=1 if white
4	age	in years
5	income	annual income, $
6	cigs	cigs. smoked per day
7	restaurn	=1 if state restaurant smoking restrictions
8	lincome	log(income)
9	agesq	age^2
10	lcigpric	log(cigprice)

Non daté : étendue complète 1 – 807

- Analyse de la consommation de cigarettes
  - Données Wooldridge "smoke.gdt"
  - Nature de la base de données : · · · · · · · · ·
  - Nombre d'observations : · · · · · · · · ·
  - Description des variables

2 – Introduction

2.4 – Exemple : Consommation de cigarettes

```
Modèle 1: MCO, utilisant les observations 1-807
Variable dépendante: cigs

              coefficient    éc. type     t de Student   p. critique
  ---------------------------------------------------------------
  const       -3,63987      24,0787        -0,1512       0,8799
  lincome      0,880269      0,727784       1,210        0,2268
  lcigpric    -0,750854      5,77334       -0,1301       0,8966
  educ        -0,501498      0,167077      -3,002        0,0028     ***
  age          0,770694      0,160122       4,813        1,78e-06   ***
  agesq       -0,00902280    0,00174303    -5,176        2,86e-07   ***
  restaurn    -2,82508       1,11179       -2,541        0,0112     **

Moyenne var. dép.     8,686493   Éc. type var. dép.    13,72152
Somme carrés résidus  143750,7   Éc. type régression   13,40479
R2                    0,052737   R2 ajusté             0,045632
F(6, 800)             7,423062   P. critique (F)       9,50e-08
Log de vraisemblance -3236,227   Critère d'Akaike      6486,454
Critère de Schwarz    6519,307   Hannan-Quinn          6499,069
```

- Commentaires
  - Quel est le modèle économique ? économétrique ?
  - description de la regression
  - significativité ?
  - discussion : ppé MCO en termes de biais et de variance ?

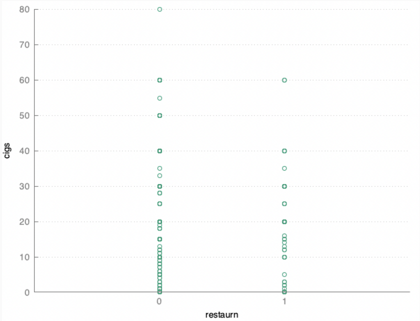
2 – Introduction

2.5 – Comment identifier la présence d'hétéroscédasticité

- Par les statistiques descriptives
  - Graphiques
  - Statistiques Descriptives dans les selon la présence de restrictions de conso ds les restaurants vs sans.
- Après avoir estimé le modèle par les MCO :
  - Par l'analyse graphique
    - Analyse graphique des résidus au carré
    - Graphique des résidus au carré vs y prédit
  - Par la réalisation de test pour détecter l'hétéroscédasticité.
    - notamment : Test de Breuch Pagan, Test de White

2 – Introduction

2.5 – Comment identifier la présence d'hétéroscédasticité



- Que constatez-vous en termes de dispersion ?

$\Rightarrow$ suggère une possible hétéroscédasticité.

2 – Introduction

2.5 – Comment identifier la présence d'hétéroscédasticité

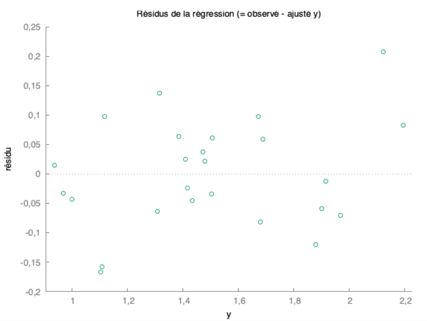
- Comparaison des écarts-types conditionnels s'il y a des restrictions ou pas.

Variable	Moyenne	Médiane	E.T.	Min	Max	N
cigs si restaurn=1	6,60	0,00	11,9	0,00	60,0	199,0
cigs si restaurn=0	9,37	0,00	14,2	0,00	80,0	80,0

- Les deux écarts types conditionnels 14,2 et 11,9 sont différents $\Rightarrow$ suggère une possible hétéroscédasticité.
- Méthode informelle car **nous ne savons pas si cette différence est statistiquement significative.**

2 – Introduction

2.5 – Comment identifier la présence d'hétéroscédasticité



- Que constatez-vous en termes de dispersion ?

$\Rightarrow$ suggère une possible hétéroscédasticité.

Plan du cours

3 – Définition

3.1 – Implications de H2

- Si H_2 est respectée, la variance de l'erreur est constante : $V[u] = E(uu^T) = \sigma_u^2 I_N$

$$V(u) = E[uu^T] = \begin{bmatrix} \sigma_u^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_u^2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_u^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_u^2 \end{bmatrix} = \sigma_u^2 \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} = \sigma_u^2 I_N$$

- H_2 n'est pas respectée : $V[u] = E(uu^T) = \sigma_u^2 \Omega \neq \sigma_u^2 I_N$

$$\text{V(u)=} \begin{bmatrix} \omega_{11} & 0 & 0 & \cdots & 0 \\ 0 & \omega_{22} & 0 & \cdots & 0 \\ 0 & 0 & \omega_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \omega_{NN} \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} 1 & \omega_{12} & \omega_{13} & \cdots & \omega_{1N} \\ \omega_{21} & 1 & \omega_{31} & \cdots & \omega_{2N} \\ \omega_{31} & \omega_{32} & 1 & \cdots & \omega_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \omega_{N1} & \omega_{N1} & \omega_{N1} & \cdots & 1 \end{bmatrix}$$

$\Downarrow$
Hétéroscédasticité

$\Downarrow$
Autocorrélation (partie 4-2)

3 – Définition

3.2 – Conséquences pour les MCO

- Rappel

$$V[\hat{\beta}] = (X^TX)^{-1}X^T \underbrace{E[uu^T]}_{V(u)} X(X^TX)^{-1}$$

- Si H2 est respectée

$$V[\hat{\beta}] = (X^TX)^{-1}X^T \underbrace{E[uu^T]}_{\sigma_u^2 I_N} X(X^TX)^{-1} = \sigma_u^2(X^TX)^{-1}$$

- **Si H2 n'est pas respectée** : $V[u] = E(uu^T) = \sigma_u^2\Omega \neq \sigma_u^2 I_N$

$$V[\hat{\beta}] = \sigma_u^2(X^TX)^{-1}X^T\Omega X(X^TX)^{-1} \text{ [sous } H_2\text{]}$$

- Seule H_2 n'est pas respectée : l'estimateur des MCO reste sans biais et convergent
- MAIS sa matrice des variances covariances n'est plus valide.

Plan du cours

4 – Détection de l'hétéroscédasticité

4.1 – Test Breusch-Pagan, 1979.

- **Hypothèses** :
  - Homoscédasticité : $H_0 : \sigma_i^2 = \sigma_u^2$ contre $H_1 : \sigma_i^2 = h(Z_i a)$, avec $i = 1, ..., N$

  $\Leftrightarrow$ $H_0 : a_1 = a_2 = ... = a_p$ contre $H_1 : \exists a_j \neq 0$ avec $j = 1, ..., p$
- Les étapes :
  1. Estimation du modèle par les MCO
  2. prédiction de $\widehat{u}$
  3. Régression auxiliaire

$$\frac{\widehat{u}_i^2}{\widehat{\sigma}_u^2} = a_0 + a_1 Z_{1i} + a_2 Z_{2i} + ... + a_p Z_{pi} + \eta_i (\text{eq}^\circ\ 2)$$

  - La **statistique de test** est la moitié de la SCE de la régression auxiliaire (eq° 2) :

$$Q_{BP} = SCE/2$$

- **Règle de décision** :
  - Si $Q_{BP} > \chi^2(p)$: Rejeter de H_0 $\rightarrow$ perturbations hétéroscédastiques.
  - Si $Q_{BP} < \chi^2(p)$: Ne pas rejeter H_0 $\rightarrow$ perturbations homoscédastiques.

4 – Détection de l'hétéroscédasticité

4.1 – Test Breusch-Pagan, 1979.

```
Test de Breusch-Pagan pour l'hétéroscédasticité
MCO, utilisant les observations 1-807
Variable dépendante: uhat^2 standardisé

              coefficient    éc. type      t de Student   p. critique
  -----------------------------------------------------------------
  const       -3,57214      3,66303        -0,9752       0,3298
  lincome      0,138318     0,110716        1,249        0,2119
  lcigpric     0,342317     0,878285        0,3898       0,6968
  educ        -0,0133848    0,0254171      -0,5266       0,5986
  age          0,109008     0,0243590       4,475        8,75e-06   ***
  agesq       -0,00120580   0,000265164    -4,547        6,27e-06   ***
  restaurn    -0,399604     0,169135       -2,363        0,0184     **

  Somme des carrés expliquée = 138,52

Statistique de test: LM = 69,260025,
avec p. critique = P(Khi-deux(6) > 69,260025) = 0,000000
```

•

4 – Détection de l'hétéroscédasticité

4.2 – Test de White, 1980

- Généralisation du test de BP, sans forme particulière sur l'hétéroscédasticité.
- Homoscédasticité : $H_0 : \sigma_i^2 = \sigma_u^2 \; \forall$ i vs. $H_1 : \sigma_i^2 \neq \sigma_u^2$ pour au moins un seul $i \neq j$
-
  1. Estimation du modèle par les MCO
  2. Estimation de $\hat{u}_i^2$
  3. Régression auxiliaire

$$\widehat{u_i^2} = a_0 + \sum_{j=1}^{k} a_j x_{ij} + \sum_{j=1}^{k} \sum_{l \geq j}^{k} a_{jl} x_{ij} x_{il} + \epsilon_i$$

- La **statistique de test**, notée Q_W :

$$Q_W = NR^2 \approx \chi^2 \left(\frac{(k+1)(k+2)}{2} - 1 \right)$$

- Règle de décision :
  - Si $Q_W > \chi^2_{th}$, les perturbations sont hétéroscédastiques
  - Si $Q_W < \chi^2_{th}$, les perturbations sont homoscédastiques

4 – Détection de l’hétéroscédasticité

4.2 – Test de White, 1980

Test de White pour l'hétéroscédasticité
MCO, utilisant les observations 1-807
Variable dépendante: uhat^2
Omis pour cause de multicollnéarité parfaite : sq_age

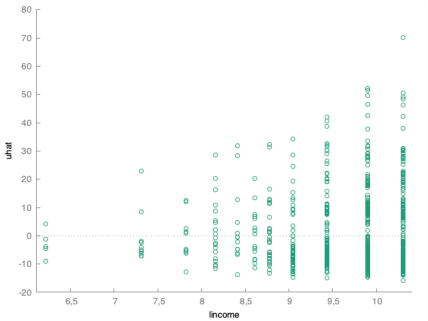
	coefficient	éc. type	t de Student	p. critique	
const	29374,9	20559,1	1,429	0,1535	
lincome	-1049,63	963,437	-1,089	0,2763	
lcigpric	-10340,7	9754,56	-1,060	0,2894	
educ	-117,471	251,285	-0,4675	0,6403	
age	-264,147	235,762	-1,120	0,2629	
agesq	3,46860	3,19465	1,086	0,2779	
restaurn	-2868,18	2986,78	-0,9603	0,3372	
sq_lincome	-3,94118	17,0712	-0,2309	0,8175	
X2_X3	329,889	239,242	1,379	0,1683	
X2_X4	-9,59186	8,04708	-1,192	0,2336	
X2_X5	-3,35456	6,68220	-0,5020	0,6158	
X2_X6	0,0267043	0,0730249	0,3657	0,7147	
X2_X7	-59,8870	49,6905	-1,205	0,2285	
sq_lcigpric	668,534	1204,32	0,5551	0,5790	
X3_X4	32,9139	59,0625	0,5573	0,5775	
X3_X5	62,8818	55,2901	1,137	0,2558	
X3_X6	-0,622372	0,594730	-1,046	0,2957	
X3_X7	862,155	720,622	1,196	0,2319	
sq_educ	-0,290343	1,28760	-0,2255	0,8217	
X4_X5	3,61705	1,72466	2,097	0,0363	**
X4_X6	-0,0355583	0,0176645	-2,013	0,0445	**
X4_X7	-2,89650	10,6571	-0,2718	0,7859	
X5_X6	-0,0191111	0,0286551	-0,6669	0,5050	
X5_X7	-4,93321	10,8403	-0,4551	0,6492	
sq_agesq	0,000117773	0,000145840	0,8076	0,4196	
X6_X7	0,0384459	0,120459	0,3192	0,7497	

R2 non-ajusté = 0,064650

Statistique de test: TR^2 = 52,172447,
avec p. critique = P(Khi-deux(25) > 52,172447) = 0,001140

4 – Détection de l'hétéroscédasticité

4.2 – Test de White, 1980



•

4 – Détection de l'hétéroscédasticité

4.3 – Application 2 : Estimation de la demande de travail

- N= 569 entreprises observées en 1996.
- La base de données comprend les variables suivantes :
  - $labor_i$ est l'effectif total de la firme,
  - $capital_i$, le stock de capital (en millions de francs belges),
  - $wage_i$, le coût salarial (en millions de francs belges) et
  - $output_i$, la valeur ajoutée (en millions francs belges).
- On considère le modèle de demande de travail :

$$\log labor_i = b_0 + b_1 \log wage_i + b_2 \log output_i + b_3 \log capital_i + u_i$$

- Remarque: modèle log-log: Coef estimés st des **élasticités**
- Données de firmes: possibilité d'hétéroscédasticité

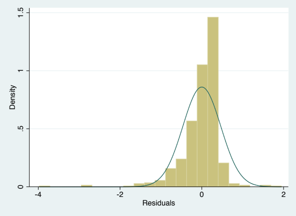
4 – Détection de l'hétéroscédasticité

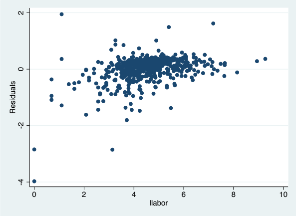
4.3 – Application 2 : Estimation de la demande de travail

Source	SS	df	MS
Model	656.747036	3	218.915679
Residual	122.338811	565	.216528869
Total	779.085847	568	1.37163001

Number of obs	=	569
F(3, 565)	=	1011.02
Prob > F	=	0.0000
R-squared	=	0.8430
Adj R-squared	=	0.8421
Root MSE	=	.46533

llabor	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
loutput	.9900474	.0264103	37.49	0.000	.938173	1.041922
lwage	-.9277642	.0714046	-12.99	0.000	-1.068015	-.7875133
lcapital	-.0036975	.0187697	-0.20	0.844	-.0405644	.0331695
_cons	-.4480909	.0932397	-4.81	0.000	-.6312297	-.2649522





4 – Détection de l'hétéroscédasticité

4.3 – Application 2 : Estimation de la demande de travail

- Application du Test de BP
  - Ici, $R^2 = 0.01358$,
  - $p = 3$ et
  - $SCR = 9621.824$,
    d'où $SCE = 132.46$
  - Et: $Q_{BP} = 132.46/2 = 66.23$
  - **Bilan** $Q_{BP} > \chi^2(3) = 7.81$, on rejette H_0, perturbations sont hétéroscédastiques. Inférence statistique fondée sur la matrice des variances-covariances des MCO n'est pas valide.
- Application du Test de White
  - $k = 3$ $Q_W = 58.54 > \chi^2_9 = 16.92$, les perturbations sont hétéroscédastiques.

Plan du cours

5 – Comment corriger l'hétéroscédasticité?

5.1 – Inférence statistique robuste

- Définition : **Un "Ecart-type Robuste"** "résiste" à une hétéroscédasticité de forme inconnue.
- Pour une inférence statistique correcte $\Rightarrow$ estimateur convergent de la matrice des variances-covariances des MCO
- Correction de White pour les estimations.
  - La "Bonne" matrice des variances-covariances des MCO proposée par *White* est donnée par :

$$\hat{V}(\widehat{\beta}_{MCO}) = (X^T X)^{-1} \widehat{V} (X^T X)^{-1}$$

  Avec $\widehat{V} = \sum_{i=1}^{N} \widehat{u}_i^2 (X_i X_i^T)^{-1}$ (Correction White).

  - Dans quelle mesure les écarts-types estimés remettent en question la significativité des coefficients estimés (ici, les élasticités) ?

5 – Comment corriger l'hétéroscédasticité?

5.2 – Matrice des variances covariances

5.2.1 – Demande de travail (suite)

- Par les MCO la matrice des variances covariances est donnée par :

$$
\begin{aligned}
\hat{V}(\widehat{b}_{MCO}) &= \hat{\sigma}_u^2 (X^T X)^{-1} \\
&= \begin{pmatrix} 0{,}0087 & & & \\ 0{,}0011 & 0{,}0051 & & \\ -0{,}0019 & -0{,}0006 & 0{,}0007 & \\ 0{,}0003 & 4{,}06\text{E-}05 & -0{,}0004 & 0{,}0004 \end{pmatrix}
\end{aligned}
$$

5 – Comment corriger l'hétéroscédasticité?

5.2 – Matrice des variances covariances

5.2.1 – *Demande de travail (suite)*

- La matrice des variances-covariances Robuste des MCO proposée par *White* est donnée par : $\hat{V}(\widehat{b}_{MCO}) = (X^TX)^{-1}\widehat{V}(X^TX)^{-1}$ Avec $\widehat{V} = \sum_{i=1}^{N} \widehat{u}_i^2 (X_i X_i^T)^{-1}$ (Correction White).

$$
\begin{aligned}
\hat{V}(\widehat{b}_{MCO,\text{White}}) &= (X^TX)^{-1}\widehat{V}(X^TX)^{-1} \\
&= \begin{pmatrix} 0{,}0178 & & & \\ 0{,}0032 & 0{,}007 & & \\ -0{,}0047 & -0{,}0007 & 0{,}002 & \\ 0{,}0015 & -0{,}0003 & -0{,}0015 & 0{,}001 \end{pmatrix}
\end{aligned}
$$

$\Rightarrow$ Ecarts-types robustes à l'hétéroscédasticité

5 – Comment corriger l'hétéroscédasticité?

5.2 – Matrice des variances covariances

5.2.2 – Demande de cigarettes (suite)

- MCO : $\hat{V}(\widehat{\beta}_{MCO}) = \sigma_u^2 (X^T X)^{-1}$

const	lincome	lcigpric	educ	age	agesq	restaurn	
425749,	-2306,46	-98114,2	238,148	-108,271	105,053	2894,44	const
	388,950	-176,539	-226,899	-229,090	0,250753	-391,904	lincome
		24476,2	-730,077	-446,066	0,00536611	-675,462	lcigpric
			204,986	-218,400	0,0308140	-406,245	educ
				188,275	-0,201740	319,422	age
					0,00223100	-0,0266133	agesq
						907,690	restaurn

- Ecarts-types robustes : $\hat{V}(\widehat{\beta}_{MCO}) = (X^T X)^{-1} \widehat{V} (X^T X)^{-1}$, avec $\widehat{V} = \sum_{i=1}^{N} \widehat{u}_i^2 (X_i X_i^T)^{-1}$ (Correction White).

const	lincome	lcigpric	educ	age	agesq	restaurn	
650,511	-264,247	-149,814	-0,240230	-0,489138	0,00510753	382,269	const
	0,352149	0,00873226	-0,0292544	-0,0187503	0,000192373	-0,0768362	lincome
		361,101	0,0481493	0,0775259	-0,000858126	-0,783484	lcigpric
			0,0261429	-0,000808724	1.70E+00	-0,0124365	educ
				0,0189566	-0,000197847	-0,00248534	age
					2.12E-01	2.60E+00	agesq
						100,732	restaurn

5 – Comment corriger l'hétéroscédasticité?

5.3 – Intervalles de confiance

5.3.1 – *Demande de cigarettes (suite)*

- MCO

Variable	Coefficient	intervalle de confiance de 95%	
const	−3,63987	−50,9047	43,6249
lincome	0,880269	−0,548322	2,30886
lcigpric	−0,750854	−12,0835	10,5818
educ	−0,501498	−0,829460	−0,173537
age	0,770694	0,456384	1,08500
agesq	−0,00902280	−0,0124443	−0,00560134
restaurn	−2,82508	−5,00746	−0,642708

- Ecarts-types robustes

Variable	Coefficient	intervalle de confiance de 95%	
const	−3,63987	−53,7047	46,4250
lincome	0,880269	−0,284577	2,04512
lcigpric	−0,750854	−12,5465	11,0447
educ	−0,501498	−0,818880	−0,184116
age	0,770694	0,500431	1,04096
agesq	−0,00902280	−0,0118805	−0,00616510
restaurn	−2,82508	−4,79519	−0,854984

5 – Comment corriger l'hétéroscédasticité?

5.4 – Les Moindres carrés (quasi) généralisés

- Considérons une matrice Ω telle que :

$$\Omega = \Phi^T \Phi$$

  où Ω est une matrice non-singulière /inversible inconnue.
- On peut transformer le modèle de la façon suivante:

$$\Phi y = \Phi X \beta + \Phi \epsilon$$

$$y^* = X^* \beta + \epsilon^*$$

- On peut montrer que $E(\epsilon^*|X^*) = E(X^*) = 0$ et que $V(\epsilon^*|X^*) = \Phi V(\epsilon|X)\Phi^T = \sigma_u^2 \Phi \Omega \Phi^T = \sigma_u^2 I_T$
  Donc $\Phi\epsilon = \epsilon^*$ respecte les hyp de GM.
  L'application des MCO a ce modèle transformé donne l'estimateur des Moindres Carrés Généralisés (MCG)

5 – Comment corriger l'hétéroscédasticité?

5.4 – Les Moindres carrés (quasi) généralisés

- L'estimateur des Moindres Carrés Généralisés (MCG) est équivalent à l'estimateur des MCO appliqué au modèle transformé:

$$\Omega^{-1/2}Y = \Omega^{-1/2}X\beta + \Omega^{-1/2}u$$

- Tout l'enjeu consiste alors à **estimer correctement la matrice Ω** :
  - On parlera alors de Moindres Carrés Quasi Généralisés (MCQG, Feasible Generalized Least Square)
  - On estimera un estimateur de Ω, noté $\widehat{\Omega}$.
  - Estimation du modèle suivant par les MCQG:

$$\widehat{\Omega}^{-1/2}Y = \widehat{\Omega}^{-1/2}X\beta + \widehat{\Omega}^{-1/2}u$$

5 – Comment corriger l'hétéroscédasticité?

5.4 – Les Moindres carrés (quasi) généralisés

- Dans le cas où la matrice des variances covariances n'est pas scalaire (i.e. les perturbations sont autocorrélées et/ou hétéroscédastiques) :
- Le meilleur estimateur linéaire sans biais de β est l'estimateur des Moindres Carrés Généralisés (MCG),

$$\hat{\beta} = (X^T\Omega^{-1}X)^{-1}X^T\Omega^{-1}y$$

  Avec comme matrice des variances covariances:

$$V[\hat{\beta}] = \sigma_u^2(X^T\Omega^{-1}X)^{-1}$$

- Une fois Ω estimé on parlera de MCQG, avec :

$$\hat{\beta} = (X^T\widehat{\Omega}^{-1}X)^{-1}X^T\widehat{\Omega}^{-1}y$$

$$V[\hat{\beta}] = \sigma_u^2(X^T\widehat{\Omega}^{-1}X)^{-1}$$

Plan du cours

6 – Propriétés MCO/MCG

6.1 – Estimateur de la variance des perturbations

- variance des perturbations $\widehat{\sigma}^2_{u,MCO}$

$$E(\hat{\sigma}^2_{u,MCO}) = \frac{\widehat{u}^T\widehat{u}}{T-(k+1)}$$

- Où : $\widehat{u}^T\widehat{u} = u^T(M^TM)u = u^TMu$, car :

$$\widehat{\beta} = (X^TX)^{-1}X^TY \Leftrightarrow \widehat{y} = X(X^TX)^{-1}X^Ty = Py \Leftrightarrow y - \widehat{y} = y - Py$$
$$\Leftrightarrow \widehat{u} = (I_T - P)y \Leftrightarrow \widehat{u} = My$$

- Donc

$$\begin{aligned} E(\hat{\sigma}^2_{u,MCO}) &= E\left[\frac{u^TMu}{T-(k+1)}\right] = \frac{1}{T-(k+1)}E(u^TMu) \\ &= \frac{1}{T-(k+1)}E\left(Tr(Muu^T)\right) = \frac{1}{T-(k+1)}Tr(M\,\underbrace{E(uu^T)}_{\sigma_u^2\Omega}) \end{aligned}$$

6 – Propriétés MCO/MCG

6.1 – Estimateur de la variance des perturbations

$$
\begin{aligned}
E(\hat{\sigma}^2_{u,MCO}) &= \frac{1}{T-(k+1)} Tr\left(M\sigma^2_u\Omega\right) = \frac{\sigma^2_u}{T-(k+1)} Tr\left(M + M\Omega - M\right) \\
&= \frac{\sigma^2_u}{T-(k+1)} Tr(M) + \frac{\sigma^2_u}{T-(k+1)} Tr\left(M\Omega - M\right) \\
&= \frac{(T-(k+1))\sigma^2_u}{T-(k+1)} + \frac{\sigma^2_u}{T-(k+1)} Tr\left(M(\Omega - I)\right) \\
E(\hat{\sigma}^2_{u,MCO}) &= \sigma^2_u + \sigma^2_u \underbrace{\frac{Tr\left(M(\Omega - I)\right)}{T-(k+1)}}_{\neq 0}
\end{aligned}
$$

Bilan : L'estimateur de la variance des perturbations $\widehat{\sigma}^2_{u,MCO}$ est biaisé.

6 – Propriétés MCO/MCG

6.1 – Estimateur de la variance des perturbations

On sait que

$$\widehat{\sigma}^2_{u,MCG} = \frac{\widehat{u}^{*T}\widehat{u}^*}{T-(k+1)}$$

où : $u^* = \Omega^{-1/2}u$ et $\widehat{u}^* = M^* u^*$
D'où :

$$\begin{aligned}
\widehat{u}^{*T}\widehat{u}^* &= u^{*T}\underbrace{M^{*T}M^*}_{M^*}u^* \\
&= u^T\Omega^{-1/2}M^*\Omega^{-1/2}u, \\
&\text{Avec } \Omega^{-1/2} \text{ symétrique}
\end{aligned}$$

6 – Propriétés MCO/MCG

6.1 – Estimateur de la variance des perturbations

$$
\begin{aligned}
E(\sigma^2_{u,MCG}) &= E\left(\frac{(u^T\Omega^{-1/2}M^*\Omega^{-1/2}u)}{T-(k+1)}\right) \\
&= \frac{1}{T-(k+1)}E\left(\mathrm{Tr}\left(\Omega^{-1/2}M^*uu^T\Omega^{-1/2}\right)\right) \\
&= \frac{1}{T-(k+1)}\mathrm{Tr}\left(M^*\Omega^{-1/2}E(uu^T)\Omega^{-1/2}\right) \\
&= \frac{1}{T-(k+1)}\mathrm{Tr}\left(\Omega^{-1/2}\sigma^2_u\Omega\Omega^{-1/2}M^*\right) \\
&= \frac{\sigma^2_u}{T-(k+1)}\mathrm{Tr}\left(M^*\right) \\
&= \sigma^2_u, \text{ car } \mathrm{Tr}(M^*) = T-(k+1)
\end{aligned}
$$

6 – Propriétés MCO/MCG

6.2 – Moindres Carrés Pondérés (MCP)

- Si la forme de l'hétéroscédastité est connue, l'estimateur des MCQG, peut-être appliqué.
- Supposons que la variance de l'erreur puisse s'écrire en fonction d'une autre variable z_t, telle que:

$$\mathrm{var}(u_t) = \sigma^2 z_t^2$$

- Pour supprimer l'hétéroscédasticité, il suffit de normaliser les variables par z_t

$$\frac{y_t}{z_t} = \beta_1 \frac{1}{z_t} + \beta_2 \frac{x_{2t}}{z_t} + \beta_3 \frac{x_{3t}}{z_t} + v_t$$

où $v_t = \dfrac{u_t}{z_t}$ est un terme d'erreur.

- Donc $\mathrm{var}(u_t) = \sigma^2 z_t^2$, $\mathsf{var}(v_t) = \mathsf{var}\left(\dfrac{u_t}{z_t}\right) = \dfrac{\mathsf{var}(u_t)}{z_t^2} = \dfrac{\sigma^2 z_t^2}{z_t^2} = \sigma^2$ pour tout z_t.

6 – Propriétés MCO/MCG

6.2 – Moindres Carrés Pondérés (MCP)

6.2.1 – *Les étapes de l'estimations*

- Forme de l'hétéroscédasticité imposée : l'heteroscedasticite multiplicative : telle que $\sigma_i = \sigma_u^2 \exp(Z_i a)$
  - Estimation du modèle par les MCO;
  - Calcul de $\log \widehat{u}_i^2$ à partir des résidus $\widehat{u}_i$ des MCO.
  - Estimation du modèle : $\log \widehat{u}_i^2 = \log \sigma_u^2 + Z_i a + \epsilon_i$.
  - Calcul de $\widehat{\sigma}_i^2 = \exp(Z_i \widehat{a})$ d'où : $\widehat{\sigma}_i = \sqrt{\exp(Z_i \widehat{a})}$
  - Par exemple, estimation du modèle suivant par les MCO, soit les MCQG :

$$\frac{1}{\widehat{\sigma}_i} \log \mathit{labor} = \widetilde{b}_0 + \widetilde{b}_1 \frac{1}{\widehat{\sigma}_i} \log \mathit{wage} + \widetilde{b}_2 \frac{1}{\widehat{\sigma}_i} \log \mathit{output} + \widetilde{b}_3 \frac{1}{\widehat{\sigma}_i} \log \mathit{capital} + \frac{1}{\widehat{\sigma}_i} u_i$$

6 – Propriétés MCO/MCG

6.2 – Moindres Carrés Pondérés (MCP)

6.2.2 – Demande cigarettes (suite)

```
Modèle 5: MCP, utilisant les observations 1-807
Variable dépendante: cigs
Variable de pondération: we

             coefficient    éc. type     t de Student  p. critique
  -----------------------------------------------------------------
  const       5,63534      17,8031        0,3165      0,7517
  lincome     1,29524       0,437012      2,964       0,0031     ***
  lcigpric   -2,94028       4,46014      -0,6592      0,5099
  educ       -0,463446      0,120159     -3,857       0,0001     ***
  age         0,481947      0,0968082     4,978       7,86e-07   ***
  agesq      -0,00562721    0,000939479  -5,990       3,17e-09   ***
  restaurn   -3,46107       0,795505     -4,351       1,53e-05   ***

Statistiques basées sur les données pondérées:

Somme carrés résidus  1993,833   Éc. type régression   1,578699
R2                    0,113409   R2 ajusté             0,106760
F(6, 800)             17,05551   P. critique (F)       1,32e-18
Log de vraisemblance -1510,045   Critère d'Akaike      3034,091
Critère de Schwarz    3066,944   Hannan-Quinn          3046,706

Statistiques basées sur les données initiales:

Moyenne var. dép.     8,686493   Éc. type var. dép.    13,72152
Somme carrés résidus  144812,7   Éc. type régression   13,45421
```

- Que dire des coefficients estimés ?

6 – Propriétés MCO/MCG

6.2 – Moindres Carrés Pondérés (MCP)

6.2.3 – *Demande de travail (suite)*

Dependent Variable: LOGL
Method: Least Squares
Sample: 1 570
Included observations: 569
Weighting series: 1/SIGMAWEIGHT

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.466471	0.090673	-5.144570	0.0000
LOGWAGE	-0.855579	0.071876	-11.90347	0.0000
LOGOUTPUT	1.034611	0.027306	37.88991	0.0000
LOGCAPITAL	-0.056863	0.021576	-2.635526	0.0086
	Weighted Statistics			
R-squared	0.905087	Mean dependent var		4.526752
Adjusted R-squared	0.904583	S.D. dependent var		1.529756
S.E. of regression	0.472536	Akaike info criterion		1.345599
Sum squared resid	126.1589	Schwarz criterion		1.376136
Log likelihood	-378.8229	F-statistic		1074.479
Durbin-Watson stat	1.958670	Prob(F-statistic)		0.000000
	Unweighted Statistics			
R-squared	0.840367	Mean dependent var		4.488665
Adjusted R-squared	0.839519	S.D. dependent var		1.171166
S.E. of regression	0.469170	Sum squared resid		124.3678
Durbin-Watson stat	2.015780			

6 – Propriétés MCO/MCG

6.2 – Moindres Carrés Pondérés (MCP)

- Quelle comparaison entre MCO et MCP?
- Résultats proches des coefficients estimés proches car :
    - l'estimateur des MCO est sans biais
    - et que l'estimateur des MCP est sans biais.
- l'élasticité du coût de capital par rapport à la demande de travail, devient significativement différente de zéro.
- Il n'est pas possible de comparer directement les R^2, car :
    - le second modèle est un modèle transformé.

Plan du cours

7 – Conclusion

7.1 – Conclusion: Ecart-types robustes, MCO ou MCQG?

- Si les erreurs sont normalement distribuées et homoscédastiques
  - statistique de Student provenant des MCO habituels suit exactement une loi de Student, indépendamment de la taille de l'échantillon.
- À l'inverse, l'approche robuste requiert de **grands échantillons**

  $\Rightarrow$ pour obtenir une statistique qui suive une loi de Student
- Si petit échantillon
  - les statistiques de Student calculées à partir des écarts-types robustes peuvent avoir une distribution très éloignée de la distribution de Student

  $\Rightarrow$ l'inférence statistique peu fiable
- Si grand échantillon

  $\Rightarrow$ Reporter écarts-types robustes à l'hétéroscédasticité et ceux des MCO.

7 – Conclusion

7.2 – Quelle solution si H_2 n'est pas respectée

- Ecarts-types Robustes

  ou

- MCP/MCQG