Économétrie — TD 4

Les tests d'hypothèses économétriques

Pierre Beaucoral

Rappel de cours

Les hypothèses des estimations MCO

- L'estimateur des Moindres Carrés Ordinaires est le meilleur estimateur linéaire sous certaines hypothèses

> **Note**
>
> - On dit qu'il est `BLUE` (*Best Linear Unbiased Estimator*)
> - Néanmoins il est sensible aux observations extrêmes

- Ces hypothèses concernent les termes d'erreurs (ε):
  - Normalité des résidus: $\varepsilon \rightsquigarrow N(0, \sigma^2_\varepsilon)$
  - Espérance nulle: $E(\varepsilon_i) = 0$
  - Homoscédasticité: $V(\varepsilon_i) = \sigma^2 = \text{constante}$
  - Indépendance sérielle: $Cov(\varepsilon_i, \varepsilon_j) = 0, \quad \forall i \neq j$

    (Absence de corrélation entre les résidus)
  - Orthogonalité des résidus (ou exogénéité): $Cov(x_i, \varepsilon_i) = 0$

Précisions sémantiques

- On distingue les propriétés sur petits échantillons et grands échantillons. Les propriétés sur petits échantillons:
  - L'estimateur est sans biais si $E(\hat{\beta}) = \beta$.
  - L'estimateur est à variance minimale si $Var(\hat{\beta}) \leq Var(\tilde{\beta})$

    avec $\tilde{\beta}$ un autre estimateur sans biais de β.
  - L'estimateur est efficace s'il remplit ces deux propriétés
- Sur grands échantillons:
  - L'estimateur est convergent si la variance de β tend vers 0 quand N tend vers l'infini:

    $\lim_{N\to\infty} Var(\beta) = 0$

Hypothèse et propriétés des estimateurs

Propriété / Hypothèse	Problème si non respectée	Test(s) associé(s)	Méthode(s) de correction
Absence de biais (orthogonalité)	Biais dans les estimations ; non-convergence	—	**Instrumentation** (variables instrumentales)
Efficience (sphéricité des erreurs)	Estimation non efficace (mais pas de biais)	- Homoscédasticité : tests de Breusch–Pagan, White- Absence d'autocorrélation sérielle : tests de Durbin–Watson, Breusch–Godfrey	- Correction de White (robust std. errors)- HAC (Newey–West)

- Ce TD se concentre uniquement sur le problème d'éfficience

Les tests d'hypothèses

- **Objectif :**
  - Voir si le modèle est **économétriquement** correct.
- **Comment ?**
  - En vérifiant que les erreurs respectent les hypothèses des MCO et que l'estimateur est efficace (**BLUE**).
  - En particulier, les tests se concentrent sur cinq hypothèses :
    1. **Normalité des résidus** : $\varepsilon\ N(0, \sigma_\varepsilon^2)$
    2. **Espérance nulle** : $E(\varepsilon_i) = 0$
    3. **Homoscédasticité** : $Var(\varepsilon_i) = \sigma^2$ (constante)
    4. **Indépendance sérielle** : $Cov(\varepsilon_i, \varepsilon_j) = 0$ pour tout $i \neq j$
    5. **Orthogonalité des résidus** : $Cov(x_i, \varepsilon_i) = 0$
- **Remarque :**
  - H2 est respectée par construction de l'estimateur MCO.
  - H5 fait l'objet d'un traitement particulier (cf. Semestre 2).

Les tests d'hypothèses

Hypothèse	Test(s) associé(s)	Traité ?
H1 : Normalité	Test de Bera–Jarque	
H3 : Homoscédasticité	Test de Goldfeld–Quandt	
	Test de Breusch–Pagan	
	Test de White	
H4 : Indépendance sérielle	Test de Durbin–Watson	
	Test de Breusch–Godfrey	

La normalité des erreurs

Le test de Bera–Jarque

- La normalité des écarts aléatoires est utile pour mettre en œuvre les **tests de sphéricité**.

- Le test utilisé est celui de **Bera–Jarque**.
- Ce test repose sur deux indicateurs :
  - **Skewness** η : mesure l'**asymétrie** de la distribution (η doit être $= 0$).
  - **Kurtosis** υ : représente l'aplatissement de la distribution (υ doit être $= 3$).

Le test de Bera–Jarque

- La statistique BJ calculée est :

  BJ = N [^{2} / 6 + (− 3)^2 / 24] → suit une loi ^{2}(2)

- Hypothèses testées :
  - **H :** BJ = 0 → la distribution suit une loi normale
  - **H :** BJ 0 → la distribution ne suit pas une loi normale

- **Règle de décision :**
  Si BJ > ^{2}(2) th (6 au seuil de 5 %),
  on rejette H .

Le test de Bera–Jarque (dans EViews)

- Pour administrer le test via l'interface graphique :
  1. Ouvrir la fenêtre de l'équation.
  2. Aller dans **View** → **Residual Diagnostic**.
  3. Choisir **Histogram – Normality Test** pour lancer le test de BJ.
  4. Lire et interpréter les résultats.

> Tip
>
> **Remarque :** la procédure est identique pour les autres tests de sphéricité, à l'exception de l'étape 3.

Homoscédasticité

- L'homoscédasticité suppose que la variance est constante : $Var(\varepsilon_i) = \sigma^2$.
  L'hétéroscédasticité apparaît généralement lorsque la taille des erreurs est proportionnelle aux valeurs prises par une variable explicative.
- Si cette hypothèse n'est pas respectée :
  - l'estimateur MCO reste **sans biais**,
  - mais il **n'est plus à variance minimale** (moins efficace).
- Tests pour vérifier cette hypothèse :
  - **Goldfeld–Quandt** (non présenté)
  - **Breusch–Pagan**
  - **White**

Test de Breusch–Pagan

- **Logique :** vérifier si la variance des résidus dépend des variables explicatives.
- Modèle estimé :
  - $Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \varepsilon_i$
  - $\varepsilon_i = Y_i – \beta_0 – \beta_1 X_i – \beta_2 Z_i$
- **Hypothèses :**
  - H_0 : $Var(\varepsilon_i) = \sigma^2$ (la variance ne dépend pas des variables explicatives)
  - H_1 : $Var(\varepsilon_i) = \sigma_i^2 = \theta_0 + \theta_1 X_i + \theta_2 Z_i + \omega_i$
    ($\theta_1 et \theta_2 \neq 0 \rightarrow$ variance liée aux variables explicatives)

> Tip
>
> Remarque : la variance des erreurs est approximée par les résidus au carré : $Var(\varepsilon_i) \approx \varepsilon_i^2$.

Test de Breusch–Pagan : procédure

1. Estimer le modèle par MCO.
2. Calculer les résidus au carré : ε_i^2.
3. Régression de test :

   $\varepsilon_i^2 = \theta_0 + \theta_1 X_i + \theta_2 Z_i + \omega_i$.
4. Examiner le pouvoir explicatif via le R² de cette équation :

   - Statistique $BP = N \times R^2 \rightarrow$ suit une loi $\chi^2(K{-}1)$, K = nombre de paramètres.
   - Règle : si $BP \geq \chi_t^2 h$ **rejet de** H_0.

> **Note**
>
> Intuition : sous homoscédasticité, R² → 0, donc X et Z n'expliquent pas la variance des résidus.

Test de White

- Même logique et démarche que Breusch–Pagan, mais avec une représentation **plus flexible** de l'hétéroscédasticité.
- **Hypothèses :**
  - $H_0 : Var(\varepsilon_i) = \sigma^2$
  - $H_1 : Var(\varepsilon_i) = \sigma_i^2 = \theta_0 + \theta_1 X_i + \theta_2 Z_i + \theta_3 X_i^2 + \theta_4 X_i Z_i + \theta_5 Z_i^2 + \omega_i$

    (les coefficients   sont conjointement   0)
- Équation de test : $\varepsilon_i^2 = \theta_0 + \theta_1 X_i + \theta_2 Z_i + \theta_3 X_i^2 + \theta_4 X_i Z_i + \theta_5 Z_i^2 + \omega_i$.

Statistiques du test de White

- Statistique principale :
  - $W = N \times R^2 \to loi\chi^2(K–1)$, K = nombre de paramètres (ici 6).
- Version petits échantillons (F-test) :
  - $W = ((SCR_r–SCR_nr)/SCR_r) \times (N–k)/(k–1) \to F(k–1, N–k)$
    * SCR_r: somme des carrés des résidus en régressant ˆ ^{2} sur constante seule.
    * SCR_nr: idem mais sur l'équation de test.
    * k : nombre de paramètres sous H (ici 3).
- **Règle de décision :** si $W \geq \chi^2_th \Rightarrow rejetdeH_0$.

Mise en œuvre sous EViews

- Les tests de Breusch–Pagan et de White sont directement programmés :
  - **View → Residual Diagnostic → Heteroskedasticity Tests**
    * **Breusch–Pagan–Godfrey** : test de Breusch–Pagan
    * **White** : test de White

Indépendance sérielle

- L'indépendance sérielle est nécessaire pour garantir que l'estimateur des MCO soit **efficace** (variance minimale).
- Indépendance sérielle = **absence d'autocorrélation des erreurs**:

  $Cov(\varepsilon_t, \varepsilon_s) = 0, \forall t \neq s$
- Ce problème concerne surtout les **séries temporelles**.
- Tests usuels :
  - **Durbin–Watson**
  - **Breusch–Godfrey**

Test de Durbin–Watson

- Premier test développé, avec des conditions restrictives :
  - Il faut une **constante** dans le modèle.
  - Le nombre d'observations doit être **supérieur à 15**.
  - La variable expliquée **retardée** ne doit pas être introduite dans le modèle.
  - Pas de données manquantes.
  - On ne peut tester que l'autocorrélation issue d'un processus AR(1) : $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$.
- Ce test a servi de base à de nombreux autres tests d'autocorrélation.

Hypothèses du test

- $H_0 : Y_t = \beta_0 + \beta_1 X_t + \beta_2 Z_t + \varepsilon_t$
- $H_1 : Y_t = \beta_0 + \beta_1 X_t + \beta_2 Z_t + \varepsilon_t$
  avec $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$.
- Sous H , l'écart aléatoire est **corrélé dans le temps**.
- Statistique de Durbin–Watson :

  $DW = \Sigma_{t=2}^{T}(\varepsilon_t – \varepsilon_{t-1})^2 / \Sigma_{t=1}^{T}\varepsilon_t^2 \approx 2(1–\rho)$
- Cette statistique est directement fournie par EViews dans le tableau de régression.

Interprétation

- La statistique DW ne suit pas une loi standard : $0 \leq DW \leq 4$.
- Les auteurs ont tabulé des valeurs critiques : $D_L et D_U (D_L < D_U)$.

Lecture :

Zone	Interprétation
[0, D_L]	Rejet H : autocorrélation positive
(D_L, D_U)	Zone d'incertitude

Zone	Interprétation
[D_U, 2]	Acceptation H
[2, 4 – D_U]	Acceptation H
(4 – D_U, 4 – D_L)	Zone d'incertitude
[4 – D_L, 4]	Rejet H : autocorrélation négative

- Test imparfait en raison de la **zone de doute** et des conditions restrictives.
- Dans la table DW, les colonnes dépendent du nombre de paramètres du modèle hors constante.

Test de Breusch–Godfrey

- Breusch et Godfrey ont développé un test de maximum de vraisemblance plus **flexible** :
  - permet de tester des processus autorégressifs d'ordre   1.
- Exemple d'un processus d'ordre 2 :
  - $H_0 : Y_t = \beta_0 + \beta_1 X_t + \beta_2 Z_t + \varepsilon_t$
  - $H_1 : Y_t = \beta_0 + \beta_1 X_t + \beta_2 Z_t + \varepsilon_t$
    avec $\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + v_t$.
- Équation de test : $\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \theta_1 X_t + \theta_2 Z_t + \omega_t$.
- Statistique : $BG = T \times R^2 \to$ suit une loi $\chi^2(t)$, où t est l'ordre du processus autorégressif (ici 2).
- **Règle de décision :**
  Rejeter H_0 si $BG \geq \chi^2_t h$.
- Procédure sous EViews :
  - **View → Residual Diagnostic → Serial correlation LM test**.
  - Choisir le nombre de retards (lags) à tester.

Correction des écarts aléatoires

- Les tests de **sphéricité** permettent de vérifier si les résidus sont :
  - **hétéroscédastiques**,
  - et/ou **autocorrélés dans le temps**.
- Dans les deux cas, il faut appliquer une **correction** pour améliorer l'**efficience** de l'estimateur.
  (Il existe de nombreuses méthodes, selon le type de problème rencontré.)
- **Sous EViews** (cf. Araujo et al., 2007) :
  - Menu : **Estimate → Options → Coefficient covariance matrix**.
  - En cas d'**hétéroscédasticité** :
    - choisir **White**.
  - En cas d'**autocorrélation sérielle** et/ou d'hétéroscédasticité :
    - choisir **HAC** (Newey–West).

Questions – Réponses (TD4)

Importez le fichier de travail sur les compagnies aériennes.

Estimez l'équation suivante par les MCO :

$log(Pass_i) = \beta_0 + \beta_1 Fatal_Passagers_i + \beta_2 NonFatal_Passagers_i + \beta_3 Low_cost_i \ + \beta_4 Public_i +$
$\beta_5 Inter_i + \beta_6 Age_i + \beta_7 Trafic_nat_i + \beta_8 Trafic_dest_i + \varepsilon_i$

Homoscédasticité

· Qu'est-ce que l'homoscédasticité et quel problème induit son non-respect pour les MCO ?

Afficher la réponse

Homoscédasticité = la variance de l'erreur est **constante** pour toutes les valeurs des régressseurs :

$\mathrm{Var}(u_i \mid X) = \sigma^2$ pour tout i.

Si cette hypothèse est violée (hétéroscédasticité) :

- Les estimateurs MCO $\hat{\beta}$ restent **sans biais** et **consistants** si $E[u \mid X] = 0$ tient, mais ils **ne sont plus efficaces** (plus BLUE) : il existe de meilleurs estimateurs (GLS/WLS).
- Les **écarts-types MCO "classiques"** sont **faussés** **tests t/F et IC** peuvent être trompeurs (trop optimistes ou trop prudents).
- Conséquence pratique majeure : **mauvaise inférence**.

Que faire ?

- Utiliser des **erreurs-types robustes à l'hétéroscédasticité** (HC0–HC3/"White").

Homoscédasticité

· A l'aide des tests de Goldfeld et Quandt, de Breusch-Pagan-Koenker et de White, que peut-on conclure quant à l'homoscédasticité du terme d'erreurs ?

Afficher la réponse

Correction(s)

· En fonction des résultats des divers tests, proposez une correction le cas échéant.

Afficher la réponse

Correction(s)

· Vos conclusions quant à l'effet des accidents mortels et non mortels sont-elles modifiées ?

Afficher la réponse