# Section 03 : Extensions au modèle linéaire simple (Séance 5)

## GSF-6053: Économétrie Financière

Simon-Pierre Boucher[1]

[1]Département de finance, assurance et immobilier
Faculté des sciences de l'administration
Université Laval

8 février 2022

# Références

**Obligatoires:**

- **Notes de cours:** Section 3 (Professeure: Marie-Hélène Gagnon)
- **Woolridge:** chapitres 3, 8, 12.

**Complémentaires:**

- **Gujarati et Porter:** chapitres 10, 11, 12, 13 et appendice C.
- **Greene:** chapitres 2, 3, 4, 5, 9, 14, 20 C et D

# Plan de la séance

Multicolinéarité

Problèmes de spécification

Erreurs non-sphériques

Les Moindres carrés généralisés

Multicolinéarité

Problèmes de spécification

Erreurs non-sphériques

Les Moindres carrés généralisés

# Multicolinéarité

- Une des hypothèses du modèle linéaire classique est qu'il n'y a pas de colinéarité parfaite entre les variables explicatives dans la matrice X.
- Supposons le modèle suivant:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki} + \epsilon_i$$

## Multicollinéarité parfaite (ou exacte)

$$\lambda_1 X_{1i} + \lambda_2 X_{2i} + ... + \lambda_k X_{ki} = 0$$

## Multicollinéarité imparfaite

$$\lambda_1 X_{1i} + \lambda_2 X_{2i} + ... + \lambda_k X_{ki} + v_i = 0$$

# Multicolinéarité

- Dans le cas de la multicolinéarité parfaite
  - Alors la matrice $(X'X)$ n'est plus de plein rang colonne et n'est pas inversible, car une des colonnes peut être écrite en fonction linéaire des autres.
  - Il survient alors un problème majeur au niveau de l'identification des paramètres.
- Dans le cas de la multicolinéarité imparfaite
  - Les régresseurs ne sont pas parfaitement corrélés, mais fortement corrélés à un choc près.

# Multicolinéarité

## Les conséquences de la multicolinéarité

- La multicolinéarité imparfaite ne viole pas les hypothèses du théorène de Gauss-Markov.
  - Les estimateurs des MCO en cas de multicolinéarité gardent la propriété BLUE.
- Les variances et les erreurs standard des estimations des coefficients de régression vont augmenter.
  - Cela signifie des t-statistiques plus faibles.
- L'ajustement global de l'équation de régression ne sera pas affecté par la multicollinéarité.
  - Cela signifie également que la prévision et la prédiction ne seront pas affectées.
- Les coefficients de régression seront sensibles aux spécifications.
  - Les coefficients de régression peuvent changer considérablement lorsque des variables sont ajoutées ou supprimées.

# Multicolinéarité

## La détection de la multicolinéarité

- Coefficients de corrélation élevés
  - Les corrélations par paires entre les variables indépendantes peuvent être élevées (en valeur absolue).
  - Règle générale : si la corrélation est $> 0.8$, il peut y avoir une forte multicollinéarité.
- $R^2$ élevé avec des valeurs de la t-statistiques faibles
  - Il est possible que les coefficients de régression individuels ne soient pas significatifs mais que l'ajustement global de l'équation soit élevé.
- Facteurs d'inflation de la variance (VIF) élevés
  - VIF quantifie dans quel mesure la multicollinéarité a augmenté la variance d'un coefficient estimé.
  - Il examine dans quelle mesure une variable explicative peut être expliquée par toutes les autres variables explicatives de l'équation.

# Multicolinéarité

## Remèdes contre la multicolinéarité

- Ne rien faire
- Abandon d'une variable redondante
  - Si une variable est redondante, elle n'aurait jamais dû être incluse dans le modèle en premier lieu. Ainsi, l'abandonner ne fait que corriger une erreur de spécification.
  - Utilisez la théorie économique pour guider votre choix de la variable à supprimer.
- Transformer les variables multicollinéaires
  - Vous pouvez réduire la multicollinéarité en respécifiant le modèle, par exemple, en créant une combinaison des variables multicollinéaires.
- Augmenter la taille de l'échantillon
  - L'augmentation de la taille de l'échantillon améliore la précision d'un estimateur et réduit les effets négatifs de la multicollinéarité.
  - En général, l'ajout de données n'est pas possible.

Multicolinéarité

Problèmes de spécification

Erreurs non-sphériques

Les Moindres carrés généralisés

# Problèmes de spécification

- Le modèle des OLS postule que nous avons la forme vraie fonctionnelle qui définit la relation entre Y et des régresseurs X.
- Comment savoir si nous avons les bons X?
- Quels sont les impacts de se tromper dans le choix de X?
- De façon générale, nous avons les guides suivants pour choisir les variables explicatives :
  - Le modèle doit être admissible et possible à tester.
  - Le modèle doit être cohérent avec la théorie économique et financière.
  - Choisir des régresseurs exogènes au terme d'erreur.
  - Les paramètres estimés doivent être stables empiriquement.
  - Les résidus de régression. Sinon, on peut revisiter la spécification de X ou si elle est correcte corriger pour la non-sphéricité des erreurs.

# Problèmes de spécification

## Omission d'une variable explicative importante

- On suppose le vrai modèle suivant:

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + u_t$$

- Cependant, vous avez effectuer l'estimation d'un modèle en omettant la variable indépendante $X_2$.

$$Y_t = \beta_1 X_{1t} + u_t$$

- Étant donné que $X_2$ n'est pas inclus dans la régression, nous allons uniquement obtenir un solution pour estimateur de $\beta_1$

$$\hat{\beta}_1 = X_1' Y (X_1' X_1)^{-1} = \frac{X_1' Y}{X_1' X_1}$$

# Problèmes de spécification

## Omission d'une variable explicative importante

- Afin de voir les effets de l'omission de la variable $X_2$, nous allons substituer dans l'équation de l'estimateur $\hat{\beta}_1$, l'équation du vrai modèle.

$$
\begin{aligned}
\hat{\beta}_1 = \frac{X_1'Y}{X_1'X_1} &= \frac{X_1'(\beta_1 X_{1t} + \beta_2 X_{2t} + u_t)}{X_1'X_1} \\
&= \frac{\beta_1 X_1'X_1 + \beta_2 X_1'X_2 + X_1'u}{X_1'X_1} \\
&= \frac{\beta_1 X_1'X_1}{X_1'X_1} + \frac{\beta_2 X_1'X_2}{X_1'X_1} + \frac{X_1'u}{X_1'X_1} \\
&= \beta_1 + \frac{\beta_2 X_1'X_2}{X_1'X_1} + \frac{X_1'u}{X_1'X_1}
\end{aligned}
$$

# Problèmes de spécification

## Omission d'une variable explicative importante

- Nous allons maintenant prendre l'espérance de chaque coté de l'équation de $\hat{\beta}_1$

$$E(\hat{\beta}_1) = \beta_1 + E\left(\frac{\beta_2 X_1' X_2}{X_1' X_1}\right) + E\left(\frac{X_1' u}{X_1' X_1}\right)$$

- Sachant que $E\left(\frac{X_1' u}{X_1' X_1}\right) = 0$ et que $\frac{X_1' X_2}{X_1' X_1}$ est le coefficient de régression de $X_2$ sur $X_1$, que nous représenterons par $b_{21}$.

$$E(\hat{\beta}_1) = \beta_1 + \beta_2 \hat{b}_{21}$$

- $\hat{\beta}_1$ sera un estimateur biaisé de $\beta_1$: Le biais dépend du coefficient de la variable omise et du coefficient de la régression de la variable omise sur les variables incluses

# Problèmes de spécification

## Inclusion d'une variable non pertinente

- On suppose le vrai modèle suivant:

$$Y_t = \beta_1 X_{1t} + u_t$$

- Cependant, vous avez effectuer l'estimation d'un modèle en ajoutant la variable indépendante $X_2$, qui est non pertinente au modèle..

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + u_t$$

- Nous avons donc les régresseurs $X_1$ et $X_2$ dans notre modèle.
- On peut représenter les deux régresseurs dans la matrice X

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

# Problèmes de spécification

## Inclusion d'une variable non pertinente

- Nous avons donc les régresseurs $X_1$ et $X_2$ dans notre modèle.
- On peut représenter les deux régresseurs dans la matrice X

$$X = \begin{pmatrix} X_1 & X_2 \end{pmatrix}$$

- Sachant que dans le vrai modèle nous avons uniquement $_1$ comme régresseur, on peut écrire le régresseur du vrai modèle comme suit

$$X \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} X_1 & X_2 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

# Problèmes de spécification

## Inclusion d'une variable non pertinente

- Maintenant, nous allons substituer Y dans l'équation de l'estimateur $\hat{\beta}$ , par le vrai modèle.

$$\begin{aligned}
\hat{\beta} &= (X'X)^{-1}X'Y \\
&= (X'X)^{-1}X'(\beta_1 X_1 + u) \\
&= (X'X)^{-1}X'(\beta_1 X \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} + u) \\
&= (X'X)^{-1}X'\beta_1 X \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (X'X)^{-1}X'u \\
&= \beta_1 \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (X'X)^{-1}X'u
\end{aligned}$$

# Problèmes de spécification

## Inclusion d'une variable non pertinente

- On applique l'espérance de chaque coté de l'équation de l'estimateur $\hat{\beta}$

$$E(\hat{\beta}) = \beta_1 \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (X'X)^{-1}X'E(u)$$

Sachant $E(u) = 0$

$$E(\hat{\beta}) = \beta_1 \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ 0 \end{pmatrix} = \beta_1$$

# Problèmes de spécification

## Inclusion d'une variable non pertinente

- L'estimateur de $\hat{\beta}$ restera un estimateur sans biais.
- Pour cette raison, la littérature a souvent suggéré d'inclure plus de variables dans le doute pour ne pas en omettre.
- Il n'est donc pas recommandé d'ajouter des variables sans avoir un indice économique sérieux qu'elles doivent être incluses dans le modèle.

Multicolinéarité

Problèmes de spécification

Erreurs non-sphériques

Les Moindres carrés généralisés

# Erreurs non-sphériques

- Le modèle linéaire comporte des hypothèses très strictes sur la matrice de variance covariance des erreurs du modèle.
- Celle-ci est diagonale et de forme

$$E(uu') = V = \sigma^2 I_t$$

- Cela implique que les erreurs du modèle sont dites homoscédastiques, c'est-à-dire que sa variance est constante dans le temps, ne dépend pas de variables exogènes et n'est pas autocorrélée.
- C'est une hypothèse distributionnelle très forte surtout en coupe transversale.
- Raisons pour rencontrer des erreurs non-sphériques sont nombreuses:
  - Meilleure collection des données dans le temps (l'erreur devrait donc diminuer)
  - Présence d'observations aberrantes ou d'outliers
  - Mauvaise spécification du modèle

# Erreurs non-sphériques

**Estimateurs pour des formes de matrices variance covariance des erreurs plus générales:**

- Soit le modèle suivant :

$$Y = X\beta + u$$

Avec

$$E(uu') = V = \sigma^2\Omega$$

Où $\Omega$ est une matrice symétrique et inversible, mais n'est pas égale à la matrice d'identité.

# Erreurs non-sphériques

## Matrice $\Omega$

- Cette formulation permet l'hétéroscédasticité des erreurs de plusieurs formes.
  - Hétéroscédasticité groupée de différentes formes.
  - Corrélation sérielle des erreurs.
  - Variance qui varie en fonction de variables exogènes.
  - Effets arch.

# Erreurs non-sphériques

## Implication si $\Omega \neq I$

- Vérifions si l'estimateur des OLS reste sans biais et convergent lorsque la variance du terme d'erreur n'est plus homoscédastique.
- Biais de l'estimateur OLS si l'hypothèse que $\Omega = I$ n'est pas respecté.

$$\begin{aligned}
\hat{\beta}_{OLS} &= (X'X)^{-1}X'Y \\
&= (X'X)^{-1}X'[X\beta + u] \\
&= (X'X)^{-1}X'X\beta + (X'X)^{-1}X'u \\
&= \beta + (X'X)^{-1}X'u
\end{aligned}$$

# Erreurs non-sphériques

## Implication si $\Omega \neq I$

$$\begin{aligned} E(\hat{\beta}_{OLS}) &= \beta + (X'X)^{-1}X'E(u) \\ &= \beta \end{aligned}$$

- On voit que comme la matrice de variance covariance des erreurs n'intervient pas dans le calcul de l'espérance de l'estimateur OLS.
- L'estimateur des OLS est toujours sans biais.

# Erreurs non-sphériques

## Implication si $\Omega \neq I$

- Variance de l'estimateur OLS.

$$
\begin{aligned}
V(\hat{\beta}) &= E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'] \\
&= E[(X'X)^{-1}X'uu'X(X'X)^{-1}] \\
&= (X'X)^{-1}X'E(uu')X(X'X)^{-1} \\
&= \sigma^2(X'X)^{-1}X'\Omega X(X'X)^{-1}
\end{aligned}
$$

- Cet estimateur est problématique, car on ne peut pas le comparer à l'estimateur des OLS.
  - On ne peut pas dire s'il est plus grand, plus petit
- On ne peut donc pas appliquer la preuve du théorème de Gauss-Marcov ni regarder la borne de Cramer-Rao.

Multicolinéarité

Problèmes de spécification

Erreurs non-sphériques

Les Moindres carrés généralisés

# Les Moindres carrés généralisés

- Transformation du modèle contenant des erreurs sphériques, pour le rendre optimal et compatible avec les hypothèses des MCO.
- Pour tous $\Omega$ inversible, il existe une matrice P tel que

$$P'P = PP' = \Omega$$

- Si on inverse les deux cotés de la denière équation

$$\Omega^{-1} = [PP']^{-1} = [P']^{-1}P^{-1}$$

où P n'est pas aléatoire

# Les Moindres carrés généralisés

## Modèle transformé:

$$P^{-1}Y = P^{-1}X\beta + P^{-1}u$$

**Variance des erreurs:**

$$\begin{aligned}
E[(P^{-1}u)(P^{-1}u)'] &= E(P^{-1}uu'P^{-1\prime}) \\
&= P^{-1}E(uu')P^{-1} \\
&= P^{-1}\sigma^2\Omega P^{-1\prime} \\
&= \sigma^2 P^{-1}\Omega P^{-1\prime} \\
&= \sigma^2 P^{-1}PP'P^{-1\prime} \\
&= \sigma^2 I
\end{aligned}$$

- Donc, l'estimateur OLS appliqué au modèle transformé sera optimal!

# Les Moindres carrés généralisés

Modèle transformé:

**Estimateur GLS**

$$\begin{aligned}\hat{\beta}_{GLS} &= [(P^{-1}X)'(P^{-1}X]^{-1}(P^{-1}X)'(P^{-1}Y)\\ &= [X'P^{-1\prime}P^{-1}X]^{-1}X'P^{-1\prime}P^{-1}Y\\ &= [X'\Omega^{-1}X]^{-1}X'\Omega^{-1}Y\end{aligned}$$

# Les Moindres carrés généralisés

## Modèle transformé:

**Estimateur FGLS**

- Si $\Omega$ (ou $P$) n'est pas connu, on le remplace par un estimé convergent $\hat{\Omega}(\hat{P})$.

$$\begin{aligned}\hat{\beta}_{FGLS} &= [(\hat{P}^{-1}X)'(\hat{P}^{-1}X]^{-1}(\hat{P}^{-1}X)'(\hat{P}^{-1}Y) \\ &= [X'\hat{\Omega}^{-1}X]^{-1}X'\hat{\Omega}^{-1}Y \\ &= PLIM\hat{\beta}_{FGLS} = \beta\end{aligned}$$

- Ce résultat demande que $\hat{\Omega}$ soit convergent.
- Cet estimateur n'est pas optimal au sensde Cramer-Rao, il est optimal asymptotiquement.

# Les Moindres carrés généralisés

## Modèle transformé:

**Estimateur convergent**

- Un estimateur est convergent (consistent en anglais) si lorsque la taille de l'échantillon augmente vers l'infini, l'estimateur se concentre (converge) sur la vraie valeur du paramètre.
- Les conditions suffisantes pour la convergence en probabilité sont :

$$\lim_{T\to\infty} E(\hat{\beta}_T) = \beta$$

$$\lim_{T\to\infty} Var(\hat{\beta}_T = 0$$

- La convergence est une propriété qui fait intervenir la loi des grands nombres et le théorème de limite centrale qui sont des résultats statistiques asymptotiques.

# Les Moindres carrés généralisés

## Modèle transformé:

**Espérance estimateur GLS**

- Si $\Omega$ est connu et non estimé.

$$\begin{aligned}
\hat{\beta}_{GLS} &= [X'\Omega^{-1}X]^{-1}X'\Omega^{-1}Y \\
&= [X'\Omega^{-1}X]^{-1}X'\Omega^{-1}[X\beta + u] \\
&= [X'\Omega^{-1}X]^{-1}X'\Omega^{-1}X\beta + [X'\Omega^{-1}X]^{-1}X'\Omega^{-1}u \\
&= \beta + [X'\Omega^{-1}X]^{-1}X'\Omega^{-1}u
\end{aligned}$$

En espérance:

$$\begin{aligned}
E(\hat{\beta}_{GLS}) &= \beta + [X'\Omega^{-1}X]^{-1}X'\Omega^{-1}E(u) \\
&= \beta
\end{aligned}$$

# Les Moindres carrés généralisés

## Modèle transformé:

**Variance estimateur GLS**

- Si $\Omega$ est connu et non estimé.

$$
\begin{aligned}
V(\hat{\beta}_{GLS}) &= E[(\hat{\beta}-\beta)(\hat{\beta}-\beta)'] \\
&= E[(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}uu'\Omega^{-1}X(X'\Omega^{-1}X)^{-1}] \\
&= (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}E(uu')\Omega^{-1}X(X'\Omega^{-1}X)^{-1} \\
&= \sigma^2(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\Omega\Omega^{-1}X(X'\Omega^{-1}X)^{-1} \\
&= \sigma^2(X\Omega^{-1}X)^{-1}
\end{aligned}
$$

# Les Moindres carrés généralisés

## Variances des différents estimateurs

- La variance des OLS sur le modèle standard sans hétéroscédasticité :

$$V(\hat{\beta}) = \sigma^2(X'X)^{-1}$$

- La variance des OLS avec hétéroscédasticité sur le modèle non transformé :

$$V(\hat{\beta}_{OLS})_{HE} = \sigma^2(X'X)^{-1}X'\Omega X(X'X)^{-1}$$

- La variance du modèle GLS transformé par la matrice P:

$$V(\hat{\beta}_{GSL}) = \sigma^2(X'\Omega^{-1}X)^{-1}$$

# Les Moindres carrés généralisés

## Variances des différents estimateurs

- Les éléments sur la diagonale de $V(\hat{\beta}_{OLS})_{HE}$: sont plus grand que ceux de $V(\hat{\beta}_{GSL})$
- Ce n'est pas le même modèle qui est estimé, de même que les coefficients associés sont différents.
- Il n'est donc pas approprié de comparer les deux séries de résultats entre eux.
- Lorsqu'il y a hétéroscédasticité la variance des OLS standards $V(\hat{\beta}) = \sigma^2(X'X)^{-1}$ est **FAUSSE** et elle ne peut pas être comparée aux deux autres.
- On ne sait pas si elle est plus petite ou plus grande, elle est simplement fausse.

# Les Moindres carrés généralisés

## Si on connait $\Omega$

- Il est préférable de prendre lestimateur GLS.

## Si on ne connait pas $\Omega$

- Il est préférable de prendre lestimateur GLS et d'estimer $\Omega$.
- Le résultat de l'estimateur FGLS tient, mais en limite seulement.