Économétrie

Violations des conditions idéales

Stéphane Adjemian

stephane.adjemian@univ-lemans.fr

Septembre 2024

Les conditions idéales

Dans le chapitre précédent, nous avons supposé que :

- Les erreurs sont centrées, *ie* $\mathbb{E}[\varepsilon] = 0$.
- Le modèle empirique est bien spécifié.
- Les erreurs sont normalement distruées.
- La matrice X est de plein rang colonne.
- Les erreurs sont homoscédastiques et non autocorrélées, *ie* $\mathbb{V}[\varepsilon] = \sigma_\varepsilon^2 I_T$.

$\Rightarrow$ Comment se comporte l'estimateur des MCO si ces conditions ne sont pas réunies ?

Erreurs non centrées, I

Proposition 1

Supposons que $\mathbb{E}\left[\varepsilon\right] = \xi \neq 0$, *alors :*

1. $\mathbb{E}\left[\hat{\mathbf{b}}\right] = \beta + (X'X)^{-1}X'\xi$
2. $\hat{\mathbf{b}} \underset{T\to\infty}{\overset{\textit{proba}}{\longrightarrow}} \beta + Q^{-1} \lim\limits_{T\to\infty} \frac{X'\xi}{T}$

- Les colonnes de X sont linéairement indépendantes $\Rightarrow X'\xi = 0$ ssi $\xi = 0$.
- $\hat{\mathbf{b}}$ est un estimateur sans biais de β ssi $\xi = 0$.
- **Remarque:** A priori ξ est différent de $\mathbf{1} \triangleq \begin{pmatrix} 1 & 1 & \dots & 1 \end{pmatrix}'$.

Preuve de la proposition 1. Supposons que $\varepsilon = \nu + \xi$ où ν est un vecteur gaussien d'espérance nulle et de variance $\sigma_\varepsilon^2 I_T$. En substituant le DGP dans l'espression de l'estimateur des MCO, il vient :

$$\begin{aligned}\hat{\mathbf{b}}_T &= (X'X)^{-1}X'(X\beta + \xi + \nu)\\ &= \beta + (X'X)^{-1}X'\xi + (X'X)^{-1}X'\nu\end{aligned}$$

En prenant l'espérance, on obtient directement :

$$\mathbb{E}\left[\hat{\mathbf{b}}\right] = \beta + (X'X)^{-1}X'\xi$$

puisque le vecteur ν est d'espérance nulle. On a aussi :

$$\hat{\mathbf{b}} \underset{T\to\infty}{\xrightarrow{\text{proba}}} \beta + Q^{-1}\lim_{T\to\infty}\frac{X'\xi}{T} + Q^{-1}\underset{T\to\infty}{\text{plim}}\frac{X'\nu}{T}$$

On peut montrer que le dernier terme est nul. En effet $\mathbb{E}\left[\frac{X'\nu}{T}\right] = 0$ et

$$\mathbb{V}\left[\frac{X'\nu}{T}\right] = \frac{\sigma_\varepsilon^2}{T}\frac{X'X}{T} \underset{T\to\infty}{\longrightarrow} 0 \times Q = 0$$

Ainsi nous avons bien :

$$\hat{\mathbf{b}} \underset{T\to\infty}{\xrightarrow{\text{proba}}} \beta + Q^{-1}\lim_{T\to\infty}\frac{X'\xi}{T}$$

C.Q.F.D.

Erreurs non centrées, II

Proposition 2

Supposons que $\mathbb{E}[\varepsilon] = \xi \neq 0$*, alors :*

1. $\mathbb{E}\left[s^2\right] = \sigma_\varepsilon^2 + \frac{\xi' M \xi}{T-K}$
2. $s^2 \underset{T\to\infty}{\xrightarrow{\text{proba}}} \sigma_\varepsilon^2 + \lim_{T\to\infty} \frac{\xi' M \xi}{T}$

où $M = I - X(X'X)^{-1}X'$.

- $\xi' M \xi = \xi'\xi - \xi' X (X'X)^{-1} X' \xi \geq 0$, car M est semi-définie positive.
- L'estimateur est sans biais ssi $\xi = 0$.
- s^2 sur-estime la variance des erreurs.

Preuve de la proposition 2. Supposons que $\varepsilon = \nu + \xi$ où ν est un vecteur gaussien d'espérance nulle et de variance $\sigma_\varepsilon^2 I_T$. Il vient :

$$\begin{aligned} s^2 &= \frac{\varepsilon' M \varepsilon}{T-K} \\ &= \frac{(\nu+\xi)' M (\nu+\xi)}{T-K} \\ &= \frac{\nu' M \nu}{T-K} + \frac{\xi' M \xi}{T-K} + 2\frac{\nu' M \xi}{T-K} \end{aligned}$$

En prenant l'espérance, on obtient directement (puisque ν est centré et ξ déterministe) :

$$\mathbb{E}\left[s^2\right] = \sigma_\varepsilon^2 + \frac{\xi' M \xi}{T-K}$$

On a aussi :

$$s^2 \xrightarrow[T\to\infty]{\text{proba}} \sigma_\varepsilon^2 + \lim_{T\to\infty} \frac{\xi' M \xi}{T} + 2 \operatorname*{plim}_{T\to\infty} \frac{\nu' M \xi}{T}$$

Si ${}^{\xi' M \xi}/_{T}$ converge vers un nombre fini, ce que nous supposons, alors le dernier terme doit être nul. En effet, nous avons :

$$\mathbb{E}\left[\frac{\nu' M \xi}{T}\right] = 0$$

et

$$\mathbb{V}\left[\frac{\nu' M \xi}{T}\right] = \sigma_\varepsilon^2 \frac{\xi' M \xi}{T^2} \xrightarrow[T\to\infty]{} 0$$

car la convergence vers un nombre fini de ${}^{\xi' M \xi}/_{T}$ implique la convergence vers 0 de

${}^{\xi' M\xi}/_{T^2}$ quand T tend vers l'infini. Ainsi ${}^{\nu' M\xi}/_{T}$ converge bien en probabilité vers 0. C.Q.F.D.

Conclusion. En toute généralité, si les erreurs ne sont pas centrées, il n'est pas possible d'obtenir une estimation sans biais des paramètres. La convergence en probabilité de $\hat{\mathbf{b}}$ vers β, même si l'estimateur est biaisé, est assurée si et seulement si la condition $\lim\limits_{T\to\infty} \frac{X'\xi}{T} = 0$ est satisfaite. La présence d'erreurs non centrées n'affecte pas seulement les estimateurs mais aussi les tests présentés dans le chapitre I.

Erreurs non centrées, III

Proposition 3

On considère le DGP et le modèle empirique partitionnés suivants :

$$\mathbf{y} = X_1\beta_1 + X_2\beta_2 + \varepsilon$$

$$\mathbf{y} = X_1\mathbf{b}_1 + X_2\mathbf{b}_2 + \epsilon$$

avec $\begin{pmatrix} X_1 & X_2 \end{pmatrix} = X$*,* X_1 *une matrice* $T \times K_1$*,* X_2 *une matrice* $T \times K_2$*,* $K = K_1 + K_2$*. Le DGP vérifie l'ensemble des conditions idéales assurant les bonnes propriétés de l'estimateur des MCO, sauf l'espérance des erreurs qui satisfait* $\mathbb{E}\left[\varepsilon\right] = X_1\gamma$ *où* γ *est un vecteur de paramètres* $K_1 \times 1$ *(l'espérance de* ε *est une combinaison linéaire des colonnes de* X_1*). Alors* $\hat{\mathbf{b}}_2$ *est un estimateur sans biais et convergent de* β_2*,* s^2 *est un estimateur sans biais et convergent de* σ^2_ε*, et* $\mathbb{E}\left[\hat{\mathbf{b}}_1\right] = \beta_1 + \gamma$*.*

Erreurs non centrées, IV

- La portée de la proposition 3 peut sembler relativement limitée...
- Mais : Si le modèle contient une constante (disons X_1 se réduit à une colonne de 1) et si $\mathbf{E}[\varepsilon_i] = \mathbf{E}[\varepsilon_j] = c$ pour tout $(i, j) \in \{1, \ldots, T\}^2$ alors l'estimation des paramètres de pentes (les paramètres associés aux variables explicatives non constantes) est sans biais et convergente.

$\Rightarrow$ La non nullité des erreurs n'est pas un problème si le modèle contient une constante.

Il n'est alors pas possible d'identifier la constante et l'espérance des erreurs.

Preuve de la proposition 3. On sait que :

$$\mathbb{E}\left[\hat{\mathbf{b}}\right] = \beta + (X'X)^{-1}X'\xi$$

En substituant l'expression de ξ :

$$\mathbb{E}\left[\hat{\mathbf{b}}\right] = \beta + (X'X)^{-1}X'X_1\gamma$$

En reprenant la partition :

$$\mathbb{E}\left[\hat{\mathbf{b}}\right] = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{pmatrix}^{-1} \begin{pmatrix} X_1'X_1 \\ X_2'X_1 \end{pmatrix} \gamma$$

On peut montrer que :

$$\begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{pmatrix}^{-1} \begin{pmatrix} X_1'X_1 \\ X_2'X_1 \end{pmatrix} = \begin{pmatrix} I \\ 0 \end{pmatrix}$$

En effet :

$$\begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{pmatrix}^{-1} \begin{pmatrix} X_1'X_1 \\ X_2'X_1 \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix}$$

s'écrit de façon équivalente :

$$\begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} X_1'X_1 \\ X_2'X_1 \end{pmatrix}$$

soit :

$$\begin{cases} X_1'X_1A + X_1'X_2B & = X_1'X_1 \\ X_2'X_1A + X_2'X_2B & = X_2'X_1 \end{cases}$$

on doit donc avoir $A = I$ et $B = 0$. L'espérance del'estimateur est donc :

$$\mathbb{E}\left[\hat{\mathbf{b}}\right] = \begin{pmatrix} \beta_1 + \gamma \\ \beta_2 \end{pmatrix}$$

L'estimateur $\hat{\mathbf{b}}_2$ est aussi convergent :

$$\begin{aligned}
\underset{T\to\infty}{\text{plim}}\, \hat{\mathbf{b}} &= \beta + \lim_{T\to\infty} (X'X)^{-1} X' X_1 \gamma + \underset{T\to\infty}{\text{plim}}\, (X'X)^{-1} X'\nu \\
&= \beta + \begin{pmatrix} \gamma I \\ 0 \end{pmatrix} + Q^{-1} \underset{T\to\infty}{\text{plim}} \frac{X'\nu}{T} \\
&= \begin{pmatrix} \beta_1 + \gamma \\ \beta_2 \end{pmatrix}
\end{aligned}$$

car ${}^{X'\nu}/_{T}$ converge en probabilité vers 0. On montre facilement que le biais de s^2 est nul. En effet, le biais (voir la proposition 2) est proportionnel à :

$$\xi' M \xi = \gamma' X_1' M X_1 \gamma$$

Puisque $MX = 0$, a fortiori on doit avoir $MX_1 = 0$ et donc un biais nul. On montre tout aussi facilement que cet estimateur est convergent. C.Q.F.D.

Mauvaise spécification du modèle empirique, I

- Le modèle empirique peut être différent du modèle générateur des données. Par exemple, si des variables explicatives sont omises.

- Supposons que le modèle générateur des données soit :

$$\mathbf{y} = X_1\beta_1 + X_2\beta_2 + \varepsilon$$

  où ε est un vecteur aléatoire normal $T \times 1$ d'espérance nulle et de variance $\sigma_\varepsilon^2 I_T$, $\mathbf{y}$ un vecteur $T \times 1$, X_1 et X_2 respectivement des matrices déterministes $T \times K_1$ et $T \times K_2$, β_1 et β_2 respectivement des vecteurs de paramètres $K_1 \times 1$ et $K_2 \times 1$.

- Le modèle empirique est :

$$\mathbf{y} = X_1\mathbf{b}_1 + \epsilon$$

  $\Rightarrow$ les variables X_2 sont omises.

Mauvaise spécification du modèle empirique, II

- Il n'est évidemment pas possible d'estimer β_2...
- Est-il possible d'estimer β_1 ? σ^2_ε ?
- A priori non, car les variables omises contaminent les erreurs (ϵ) qui n'ont plus les bonnes propriétés.
- En comparant les deux modèles, on comprend que $\epsilon = \varepsilon + X_2\beta_2$. Les propriétés de $\hat{\mathbf{b}}_1$ ou s^2 doivent dépendre de $X_2\beta_2$...

Mauvaise spécification du modèle empirique, III

Proposition 4

Soit le modèle générateur des données :

$$\mathbf{y} = X_1\beta_1 + X_2\beta_2 + \varepsilon$$

satisfaisant toutes les conditions idéales. On estime le modèle :

$$\mathbf{y} = X_1\mathbf{b}_1 + \epsilon$$

par les MCO. On note $\hat{\mathbf{b}}_1$ *et* s^2 *les estimateurs de* $\mathbf{b}_1$ *et* σ^2_ε*, on a :*

1. $\mathbb{E}\left[\mathbf{b}_1\right] = \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2$,
2. $\hat{\mathbf{b}}_1 \underset{T\to\infty}{\xrightarrow{\textit{proba}}} \beta_1 + \lim\limits_{T\to\infty}\left(\frac{X_1'X_1}{T}\right)^{-1}\lim\limits_{T\to\infty}\frac{X_1'X_2}{T}\beta_2$
3. $\mathbb{E}\left[s^2\right] = \sigma^2_\varepsilon + \frac{\beta_2'X_2'M_1X_2\beta_2}{T-K}$
4. $s^2 \underset{T\to\infty}{\xrightarrow{\textit{proba}}} \sigma^2_\varepsilon + \lim\limits_{T\to\infty}\frac{\beta_2'X_2'M_1X_2\beta_2}{T}$

avec $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$.

Mauvaise spécification du modèle empirique, IV

- Comme X_2 est de plein rang colonne, $\nexists\ \beta_2 \neq 0$ tel que $X_2\beta_2 = 0$.

- Les estimateurs ne sont pas biaisés si $\beta_2 = 0$.

- Les estimateurs ne sont pas biaisés si $X_1'X_2 = 0$.

$\Rightarrow$ Si les variables omises ne sont pas corrélées avec les variables incluses dans le modèle empirique, alors les estimateurs ne sont pas biaisés.

- Dans ce cas $X'X$ est une matrice bloc diagonale avec $X_1'X_1$ et $X_2'X_2$ le long de la diagonale. Ainsi :

$$\hat{\mathbf{b}} = \begin{pmatrix} (X_1'X_1)^{-1} & 0 \\ 0 & (X_2'X_2)^{-1} \end{pmatrix} \begin{pmatrix} X_1'\mathbf{y} \\ X_2'\mathbf{y} \end{pmatrix} = \begin{pmatrix} (X_1'X_1)^{-1}X_1'\mathbf{y} \\ (X_2'X_2)^{-1}X_2'\mathbf{y} \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{b}}_1 \\ \hat{\mathbf{b}}_2 \end{pmatrix}$$

  Les estimations du modèle mal spécifié et du modèle bien spécifié sont identiques.

Non normalité des erreurs, I

- On suppose que $\mathbb{E}\left[\varepsilon_t\right] = 0$, $\mathbb{V}\left[\varepsilon_t\right] = \sigma^2_\varepsilon$ et $\varepsilon_t \perp \varepsilon_s$ $(s \neq t)$ pour tout $t = 1, \dots, T$, mais que ε n'est pas normalement distribué.

- $\hat{\mathbf{b}}$ et s^2 sont toujours sans biais, le théorème de Gauss-Markov reste valide (puisque sa démonstration ne fait pas appel à la normalité des erreurs).

- Mais $\hat{\mathbf{b}}$ n'est plus normalement distribué, s^2 ne suit plus une loi du khi-deux.

- Les tests développés dans le chapitre I ne sont plus valides.

- Néamoins, un théorème de la limite centrale peut être utilisé pour établir la normalité asymptotique de $\hat{\mathbf{b}}$, et pour obtenir les distributions asymptotiques des tests présentés dans le chapitre I.

Références

Greene, William H. (2017). *Econometric analysis*. Pearson.

Schmidt, Peter (1976). *Econometrics*. Taylor & Francis.