Econométrie des données de panel: Modèle à erreurs corrélées

Guillaume Horny*

*Banque de France

M2 SE

Chapitre 4 : Modèle à erreurs corrélées

Plan

1. Introduction
2. Estimateurs de classe λ
3. Modèle à erreurs corrélées
4. Repérer les erreurs corrélées : le test d'Hausman
5. Une solution au problème : l'approche de Mundlak

Plan

1. Introduction
2. Estimateurs de classe λ
3. Modèle à erreurs corrélées
4. Repérer les erreurs corrélées : le test d'Hausman
5. Une solution au problème : l'approche de Mundlak

Variable aléatoire ou paramètre ?

Modèle de base :

$$y_{it} = x_{it}^{'}\beta + \alpha_i + \epsilon_{it},$$

On a vu que les α_i peuvent être considérés comme des paramètres (modèle à effets fixes, chapitre 1) ou bien comme les réalisation d'une variable aléatoire (modèle à effets aléatoires, chapitre 2). Ce choix de modélisation conduit à faire des hypothèses différentes, aboutissant à des estimateurs différents.

On va voir ici que tous les estimateurs vus précédemment appartiennent à une même famille : celle des **estimateurs de classe** λ.

Plan

1. Introduction
2. Estimateurs de classe λ
3. Modèle à erreurs corrélées
4. Repérer les erreurs corrélées : le test d'Hausman
5. Une solution au problème : l'approche de Mundlak

Estimateurs de classe λ

Un estimateur de classe λ est tel que :

$$\widehat{\beta}(\lambda) = \left[X'(W + \lambda B)X\right]^{-1} X'(W + \lambda B)Y,$$

où :

- $W = I_{NT} - X_1(X_1'X_1)^{-1}X_1'$. Cette matrice était notée M_{X_1} dans le chapitre 1, avec X_1 les indicatrices d'individus,
- λ est un scalaire,
- $B = X_1(X_1'X_1)^{-1}X_1'$.

Retour sur l'estimateur de classe λ

Estimateur de classe λ :

$$\widehat{\beta}(\lambda) = \left[X'(W + \lambda B)X\right]^{-1} X'(W + \lambda B)Y.$$

On a :

- $\widehat{\beta}(0) = \widehat{\beta}_{Within}$,
- $\widehat{\beta}(1) = \widehat{\beta}_{OLS}$, car $W = I_{NT} - B$,
- $\widehat{\beta}(\infty) = \widehat{\beta}_{Between}$,
- $\widehat{\beta}\left(\frac{\sigma_w^2}{\sigma_w^2 + T\sigma_\alpha^2}\right) = \widehat{\beta}_{GLS}$,
- $\widehat{\beta}\left(\frac{\widehat{\sigma}_w^2}{\widehat{\sigma}_w^2 + T\widehat{\sigma}_\alpha^2}\right) = \widehat{\beta}_{FGLS}$.

Estimateurs de classe λ : le cas *between*

Lorsque $\lambda \to \infty$, W devient négligeable par rapport à B. D'où :

$$\widehat{\beta}(\lambda) = \left[X'(\lambda B)X\right]^{-1} X'(\lambda B)Y$$
$$= \left[X'BX\right]^{-1} X'BY.$$

L'écriture ci-dessus n'est pas ce qu'on a fait de plus propre mathématiquement. L'idée est juste que la composante intra-individuelle est écrasée par la composante inter-individuelle.

Estimateurs de classe λ : les cas FGLS et GLS

- **Estimateur FGLS**
  On a vu dans le chapitre 2 que l'estimateur FGLS est équivalent à l'estimateur OLS du modèle transformé :

$$y_{it} - \widehat{\lambda} y_{i.} = (1 - \widehat{\lambda})\beta_0 + (x_{it} - \widehat{\lambda} x_{i.})' \beta + v_{it},$$

  où $\widehat{\lambda}$ est un estimateur convergent de $\lambda = 1 - \sigma_w^2 / (T\sigma_\alpha^2 + \sigma_w^2)$.
- **Estimateur GLS**
  L'estimateur GLS est obtenu lorsqu'on dispose de la vraie valeur de λ.

Plan

1. Introduction
2. Estimateurs de classe λ
3. Modèle à erreurs corrélées
4. Repérer les erreurs corrélées : le test d'Hausman
5. Une solution au problème : l'approche de Mundlak

Principe général

On a vu que :

- sous l'hypothèse de modèle à **effet fixe**, c'est-à-dire de corrélation potentielle entre les caractéristiques observables et inobservables, on dispose d'estimateurs basés sur des transformations des données qui sont convergents mais **pas efficaces**
- sous l'hypothèse de modèle à **effet aléatoire**, c'est-à-dire sous l'hypothèse d'absence de corrélation entre les caractéristiques observables et inobservables, on dispose d'un estimateur FGLS convergent et **efficace**.

On cherche maintenant des estimateurs qui, lorsque les caractéristiques inobservables sont corrélées avec les observables, resteront toujours efficaces.

Le modèle à erreurs corrélées

On repart du modèle à effets aléatoires :

$$y_{it} = x_{it}^{'}\beta + \alpha_i + w_{it}.$$

On suppose cette fois-ci que les α_i sont corrélés avec les variables explicatives :

$$E(\alpha_i | x_{i1}, \ldots, x_{iT}) \neq E(\alpha_i).$$

Attention, il ne s'agit en aucune manière d'une conséquence des espérances itérées !

Les hypothèses (1/2)

Les autres hypothèses sont valides, à quelques ajustements près :

1. H1b : $\mathrm{E}(w_{it}|x_{i1},\ldots,x_{iT}) = 0, t = 1,\ldots,T$,
   remplace H1a : $\mathrm{E}(\epsilon_{it}|x_{i1},\ldots,x_{iT},\alpha_i) = 0, t = 1,\ldots,T$.
2. H2b : $\mathrm{E}(\alpha_i) = 0$,
   remplace H2a : $\mathrm{E}(\alpha_i|x_{i1},\ldots,x_{iT}) = \mathrm{E}(\alpha_i) = 0$
3. H3 :
   - $\mathrm{E}(\alpha_i^2|x_{i1},\ldots,x_{iT}) = \sigma_\alpha^2$,
   - $\mathrm{E}\left[(w_{i1},\ldots,w_{iT})(w_{i1},\ldots,w_{iT})'|x_{i1},\ldots,x_{iT},\alpha_i\right] = \sigma_w^2 I_T$.

On retrouve donc toujours la structure si spécifique de la matrice de variance du modèle à effet aléatoire.

Les hypothèses (2/2)

La spécificité du modèle à erreurs corrélées est que :

$$\begin{aligned} E(y_{it}|x_{i1},\ldots,x_{iT}) &= x_{it}^{'}\beta + E(\epsilon_{it}|x_{i1},\ldots,x_{iT}) \\ &= x_{it}^{'}\beta + E(\alpha_i|x_{i1},\ldots,x_{iT}) + E(w_{it}|x_{i1},\ldots,x_{iT}) \\ &= x_{it}^{'}\beta + E(\alpha_i|x_{i1},\ldots,x_{iT}) \\ &\neq x_{it}^{'}\beta \end{aligned}$$

Les caractéristiques inobservables affectent ici l'espérance conditionnelle de y_{it} ainsi que sa variance (cette dernière par H3). Les inobservables introduisent :

- des écarts dans les moyennes individuelles
- une surdispersion

Erreurs corrélées et biais (1/4)

$$\widehat{\beta}(\lambda) = \left[X'(W + \lambda B)X\right]^{-1} X'(W + \lambda B)Y,$$

D'où :

$$\begin{aligned}
E[\widehat{\beta}(\lambda)] =& E\left[\left[X'WX + \lambda X'BX\right]^{-1}(X'W + \lambda X'B)(X\beta + \epsilon)\right] \\
=& E\left[\left[X'WX + \lambda X'BX\right]^{-1}(X'W + \lambda X'B)X\beta\right] \\
&+ E\left[\left[X'WX + \lambda X'BX\right]^{-1}(X'W + \lambda X'B)\epsilon\right] \\
=& \beta + E\left[\left[X'WX + \lambda X'BX\right]^{-1}(X'W + \lambda X'B)\epsilon\right].
\end{aligned}$$

Comme ϵ est corrélé avec X, on ne peut pas écrire $E[f(X)\epsilon] = E[f(X)]E[\epsilon]$.

Erreurs corrélées et biais (2/4)

Poursuivons les calculs :

$$
\begin{aligned}
\mathrm{E}[[X^{'}WX + \lambda X^{'}BX]^{-1}(X^{'}W+\lambda X^{'}B)\epsilon] = \\
& \mathrm{E}\left[[X^{'}WX + \lambda X^{'}BX]^{-1}X^{'}W(\alpha + w)\right] \\
& + \mathrm{E}\left[[X^{'}WX + \lambda X^{'}BX]^{-1}\lambda X^{'}B(\alpha + w)\right].
\end{aligned}
$$

W est l'expression matricielle de la transformation *within*. Le produit $W\alpha$ revient donc à calculer pour chaque observation l'écart entre α_i et sa moyenne pour l'individu i (qui vaut α_i). Ainsi, $W\alpha = 0$. De plus, $\mathrm{E}(w) = 0$. D'où la simplification :

$$
\begin{aligned}
\mathrm{E}[[X^{'}WX + \lambda X^{'}BX]^{-1}(X^{'}W + \lambda X^{'}B)\epsilon] \\
&= \mathrm{E}\left[[X^{'}WX + \lambda X^{'}BX]^{-1}\lambda X^{'}B(\alpha + w)\right] \\
&= \mathrm{E}\left[[X^{'}WX + \lambda X^{'}BX]^{-1}\lambda X^{'}B\alpha\right] \\
&\quad + \mathrm{E}\left[[X^{'}WX + \lambda X^{'}BX]^{-1}\lambda X^{'}Bw\right].
\end{aligned}
$$

Erreurs corrélées et biais (3/4)

Comme w est supposé sans corrélation avec les X :

$$\mathrm{E}\left[[X^{'}WX + \lambda X^{'}BX]^{-1}\lambda X^{'}Bw\right] = 0.$$

Ainsi :

$$\mathrm{E}\left[[X^{'}WX + \lambda X^{'}BX]^{-1}(X^{'}W + \lambda X^{'}B)\epsilon\right] = \mathrm{E}\left[[X^{'}WX + \lambda X^{'}BX]^{-1}\lambda X^{'}B\alpha\right]$$

Au final, on a :

$$\mathrm{E}[\widehat{\beta}(\lambda)] = \beta + \mathrm{E}\left[[X^{'}WX + \lambda X^{'}B]^{-1}\lambda X^{'}B\alpha\right].$$

Pour qu'un estimateur de la famille de classe λ soit sans biais, il faut que le dernier terme soit nul, ce qui arrive lorsque $\lambda = 0$.

Convergence de $\widehat{\beta}$

	Modèle supposé	
	Effets fixes	Effets aléatoires
Estimateur de β	$\text{corr}(\alpha, X_k) \neq 0$	$\text{corr}(\alpha, X_k) = 0$
Empilé	Non-convergent	Convergent
Within	Convergent	Convergent
Différence première	Convergent	Convergent
Between	Non-convergent	Convergent
Effet aléatoire	Non-convergent	Convergent

Quelles options ?

Nous avons potentiellement un problème, il convient :

1. de le détecter ;
2. ou de lui trouver des solutions. On verra que ces solutions sont :
    1. soit des solutions partielles, assez simples à mettre en oeuvre mais qui supposent des hypothèses restrictives ;
    2. soit des solutions générales, dont le domaine de validité est plus large, mais dont la mise en oeuvre est aussi plus... stimulante !

D'où l'intérêt de savoir si le problème se pose vraiment !

Plan

1. Introduction
2. Estimateurs de classe λ
3. Modèle à erreurs corrélées
4. Repérer les erreurs corrélées : le test d'Hausman
5. Une solution au problème : l'approche de Mundlak

Le test d'Hausman

L'objectif est de tester une hypothèse d'absence de corrélation ($H_0 : E(X'\epsilon) = 0$). Selon son acceptation ou son rejet, on va privilégier certains estimateurs plutôt que d'autres.

Jerry Hausman (1946-), actuellement professeur d'Economie au MIT. A étudié le secteur des télécommunication et les questions de concurrence, de régulation et de taxation, entre autres. Sa contribution la plus connue est le test qui porte son nom, publié en 1978.

Le test d'Hausman : l'idée générale

Nous sommes dans un cas de figure où il existe plusieurs manières d'estimer un modèle. Si une des hypothèses du modèle n'est pas vérifiée, certains estimateurs seront convergents tandis que d'autres ne le seront plus. On va mesurer l'écart qu'il y a entre les deux estimations. Deux possibilités :

- les deux estimations sont **similaires**, l'écart est proche de 0, l'hypothèse testée semble raisonnable,
- les deux estimations sont **différentes**, l'écart est significativement différent de 0, l'hypothèse testée ne semble pas satisfaite.

Application à des données de panel

Sous l'hypothèse de corrélation entre les caractéristiques inobservables et les variables explicatives, l'estimateur FGLS n'est plus convergent alors que l'estimateur *within* reste convergent. On peut donc comparer les deux estimations.

- Si elles sont voisines, on peut accepter l'hypothèse d'exogénéité des X par rapport à α. On n'a donc pas de problème d'endogénéité des X par rapport au terme d'erreur composé.
- A l'inverse, si les deux estimations son très différentes, au moins l'une des estimations n'est pas convergente et on rejette l'hypothèse d'exogénéité des X par rapport à α.

Statistique de test

Dans le cas des panels, la statistique de test originelle est :

$$S_H = (\widehat{\beta}_{within} - \widehat{\beta}_{FGLS}) \left[\text{Var}(\widehat{\beta}_{within}) - \text{Var}(\widehat{\beta}_{FGLS}) \right]^{-1} (\widehat{\beta}_{within} - \widehat{\beta}_{FGLS}).$$

Elle suit un χ^2 à $\dim(\beta)$ degrés de libertés. Une statistique alternative a été proposée par Hausman et Taylor en 1981 :

$$S_{HT} = (\widehat{\beta}_{between} - \widehat{\beta}_{within}) \left[\text{Var}(\widehat{\beta}_{between}) + \text{Var}(\widehat{\beta}_{within}) \right]^{-1} (\widehat{\beta}_{between} - \widehat{\beta}_{within}).$$

Elle suit un χ^2 à autant de degrés de libertés qu'il y a de variables explicatives dans le modèle après transformation *within*.

Plan

1. Introduction
2. Estimateurs de classe λ
3. Modèle à erreurs corrélées
4. Repérer les erreurs corrélées : le test d'Hausman
5. Une solution au problème : l'approche de Mundlak

Les solutions au problème

La situation la plus commune est celle où le test d'Hausman rejette H_0. La présence de corrélation entre les α_i et les x_{it} implique une corrélation entre l'erreur composée et les variables explicatives. De ce point de vue, il s'agit d'un problème classique d'endogénéité. Différentes approches ont été proposées :

- "corriger" le modèle pour évacuer cette corrélation (différence première, transformation *within*...), mais les estimateurs ne sont pas efficaces et éliminent les variables constantes dans le temps
- avoir recours à des techniques de variables instrumentales

Une solution partielle : l'approche de Mundlak

Mundlak a montré que si la relation entre l'effet individuel et les variables explicatives est stable dans le temps, on peut décomposer l'effet individuel en un terme qui rend compte de sa corrélation avec les régresseurs et un autre terme qui est sans corrélation avec les régresseurs :

$$\alpha_i = x_{i.}' \pi + \eta_i,$$

où $E(\eta_i | x_{i1}, \ldots, x_{iT}) = 0$.
On retrouve un modèle à erreurs composées standard, dans lequel on a purgé de la corrélation entre les régresseurs et les effets individuels :

$$y_{it} = x_{it}' \beta + x_{i.}' \pi + \eta_i + w_{it},$$

où $E(\eta_i | x_{i1}, \ldots, x_{iT}) = 0$ et $E(w_{it} | x_{i1}, \ldots, x_{iT}) = 0$.

Estimation du modèle de Mundlak (1/2)

On retrouve un modèle à erreurs composées :

$$Y = X\beta + BX\pi + \xi,$$

où $\xi = \eta + w$. On a :

$$\begin{aligned} E(\xi|X) &= 0 \\ Var(\xi|X) &= \begin{pmatrix} \sigma_\eta^2 + \sigma_w^2 & \sigma_\eta^2 & \dots & \sigma_\eta^2 \\ \sigma_\eta^2 & \sigma_\eta^2 + \sigma_w^2 & \dots & \sigma_\eta^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\eta^2 & \sigma_\eta^2 & \dots & \sigma_\eta^2 + \sigma_w^2 \end{pmatrix}. \end{aligned}$$

Le meilleur estimateur sans biais est de type MCG

Estimation du modèle de Mundlak (2/2)

On peut montrer, sous l'hypothèse que la corrélation entre les effets individuels et les régresseurs ne dépend pas de t, que :

$$\widehat{\beta}^{MCG} = \widehat{\beta}^{Within}.$$

L'estimateur de π est :

$$\widehat{\pi} = (X'BX)^{-1}X'BY - (X'WX)^{-1}X'WY,$$

qui est l'écart entre l'estimateur *between* et l'estimateur *within*, c'est-à-dire le biais de l'estimateur MCG induit par la corrélation.