

# Econométrie des données de panel

Guillaume Horny\*

\*Banque de France

Master 2 MASERATI

# Chapitre 3

# Plan

- 1 Introduction
- 2 Estimateurs de classe  $\lambda$
- 3 Modèle à erreurs corrélées
- 4 Repérer le problème
- 5 Les solutions au problème

# Plan

- 1 Introduction
- 2 Estimateurs de classe  $\lambda$
- 3 Modèle à erreurs corrélées
- 4 Repérer le problème
- 5 Les solutions au problème

# Variable aléatoire ou paramètre ?

Modèle de base :

$$y_{it} = x'_{it}\beta + \alpha_i + \epsilon_{it},$$

On a vu que les  $\alpha_i$  peuvent être considérés comme des paramètres (modèle à effets fixes, chapitre 1) ou bien comme les réalisations 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  $\lambda$** .

# Plan

- 1 Introduction
- 2 Estimateurs de classe  $\lambda$**
- 3 Modèle à erreurs corrélées
- 4 Repérer le problème
- 5 Les solutions au problème

# Estimateurs de classe $\lambda$

Un estimateur de classe  $\lambda$  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,
- $\lambda$  est un scalaire,
- $B = X_1(X_1'X_1)^{-1}X_1'$ .

Ecriture de  $B$  (1/1)

En effet, lorsque  $X_1$  est la matrice des indicatrices d'individus :

$$X_1 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$



Ecriture de  $B$  (2/2)

$$(X_1'X_1)^{-1} = \begin{pmatrix} T & 0 & \dots & 0 \\ 0 & T & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & T \end{pmatrix}^{-1} = \frac{1}{T}I_N.$$

$$X_1(X_1'X_1)^{-1}X_1' = \frac{1}{T}X_1X_1' = \frac{1}{T} \begin{pmatrix} 1 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 1 & \dots & 1 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & \dots & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 1 \end{pmatrix}$$

## Ecriture de $B$ (3/3)

On a :

$$BY = \begin{pmatrix} \frac{1}{T} \sum_{t=1}^T y_{1t} \\ \vdots \\ \frac{1}{T} \sum_{t=1}^T y_{1t} \\ \vdots \\ \frac{1}{T} \sum_{t=1}^T y_{Nt} \\ \vdots \\ \frac{1}{T} \sum_{t=1}^T y_{Nt} \end{pmatrix}$$

$\Rightarrow$  il s'agit de la transformation *between* de  $Y$ . De même, on peut montrer que  $BX_2$  est la transformation *between* de  $X_2$ .

## Régression *between* (1/3)

L'estimateur *between* est l'estimateur OLS appliqué au modèle :

$$BY = BX\beta + B\epsilon.$$

En effet :

$$\begin{aligned}\hat{\beta}_{OLS} &= \left( (BX)' BX \right)^{-1} (BX)' BY \\ &= \left( X' B' BX \right)^{-1} X' B' BY \\ &= \left( X' BX \right)^{-1} X' BY,\end{aligned}$$

car  $B$  est carrée, symétrique et idempotente ( $B = B^2$ ).

## Régression *between* (2/3)

On a vu dans le chapitre 2 :

*“ L’estimateur between est l’estimateur OLS de la régression de  $y_i$  sur une constante et les moyennes intertemporelles  $x_i$ . ( $N$  observations!). ”*

Or,  $BY$  n’est pas de dimension  $(N \times 1)$  mais  $(NT \times 1)$ ?!?

On peut montrer que :

- la régression de  $BY$  sur  $BX$  et une constante, où  $BY$  est de dimension  $(NT \times 1)$ ,
- celle du vecteur des  $y_i$  (de dimension  $(N \times 1)$ ), sur  $x_i$  et une constante, conduisent au même estimateur  $\hat{\beta}$ .

## Régression *between* (3/3)

Par contre, comme le vecteur des résidus n'est pas de même dimension, l'estimateur de  $\text{Var}(\beta) = E(\epsilon\epsilon')$  ne sera pas identique. Faire la régression sur un vecteur ( $N \times 1$ ) est un moyen simple de corriger pour la corrélation des erreurs (répéter  $T$  fois une observation n'apporte pas d'information!).

## Ecriture de $W$

Ainsi :

$$MY = \left( I_{NT} - \frac{1}{T} X_1 X_1' \right) Y = Y - \begin{pmatrix} \frac{1}{T} \sum_{t=1}^T y_{1t} \\ \vdots \\ \frac{1}{T} \sum_{t=1}^T y_{1t} \\ \vdots \\ \frac{1}{T} \sum_{t=1}^T y_{Nt} \\ \vdots \\ \frac{1}{T} \sum_{t=1}^T y_{Nt} \end{pmatrix}.$$

$MY$  est donc la transformée *within* de  $Y$ . De même,  $MX_2$  est la transformée *within* de  $X_2$ .

Retour sur l'estimateurs de classe  $\lambda$ 

Estimateur de classe  $\lambda$  :

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

On a :

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

## Estimateurs de classe $\lambda$ : le cas *between*

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

$$\begin{aligned}\hat{\beta}(\lambda) &= [X'(\lambda B)X]^{-1} X'(\lambda B)Y \\ &= [X'BX]^{-1} X'BY.\end{aligned}$$

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 $\lambda$ : 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 :

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

où  $\hat{\lambda}$  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  $\lambda$ .

# Plan

- 1 Introduction
- 2 Estimateurs de classe  $\lambda$
- 3 Modèle à erreurs corrélées**
- 4 Repérer le problème
- 5 Les solutions au problème

# 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 problème

On repart du modèle à effets aléatoires :

$$y_{it} = x'_{it}\beta + \alpha_i + w_{it}.$$

On suppose cette fois-ci que les  $\alpha_i$  sont corrélés avec les variables explicatives :

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

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

## Rappel : loi des espérances itérées (1/2)

La loi des espérances itérées nous dit que  $E_X [E(Y|X)] = E(Y)$ . Si  $E(Y|X) = 0$ , alors  $E_X [E(Y|X)] = E_X [0] = 0$ . D'où  $E(Y|X) = E(Y)$ . Il s'agit d'un résultat très spécifique, pas d'une généralité!

**Intuitivement** : Supposons que  $X$  soit une variable aléatoire discrète prenant les valeurs  $c_1$  et  $c_2$  avec probabilités  $p_1$  et  $p_2$ . Alors

$$\begin{aligned} E_X [E(Y|X)] &= p_1 E(Y|X = c_1) + p_2 E(Y|X = c_2) \\ &= E(Y). \end{aligned}$$

- Si  $E(Y|X = c_1) = E(Y|X = c_2) = k$ , alors leur moyenne vaudra  $k$  et donc  $E(Y) = k$ .
- À l'inverse,  $E(Y) = k \not\Rightarrow E(Y|X) = k$

## Rappel : loi des espérances itérées (2/2)

On n'a de manière générale  $E(Y|X) = E(Y)$  que si  $Y$  est sans corrélation avec les  $X$ , ce qui est le cas lorsque les variables sont indépendantes, mais pas seulement (cf exemple du chapitre 1 où  $\text{corr}(X, Y) = 0$  mais où  $X$  et  $Y$  ne sont pas indépendantes) !

# Le problème (1/2)

Les autres hypothèses sont valides, à quelques ajustements près :

- ① H1b :  $E(w_{it} | x_{i1}, \dots, x_{iT}) = 0, t = 1, \dots, T.$
- ② H2b :  $E(\alpha_j) = 0$
- ③ H3 :
  - ▶  $E(\alpha_j^2 | x_{i1}, \dots, x_{iT}) = \sigma_\alpha^2,$
  - ▶  $E \left[ (w_{i1}, \dots, w_{iT})(w_{i1}, \dots, w_{iT})' | x_{i1}, \dots, x_{iT}, \alpha_j \right] = \sigma_w^2 I_T.$

On retrouve donc toujours la structure si spécifique de la matrice de variance.

## Le problème (2/2)

La spécificité du modèle à erreurs corrélées est que :

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

Les caractéristiques inobservables affectent ici l'espérance conditionnelle de  $y_{it}$  ainsi que sa variance (cette dernière par H3). Les inobservables introduisent à la fois des écarts dans les moyennes individuelles et une surdispersion.



## Les conséquences du problème (1/4)

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

D'où :

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

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

## Les conséquences du problème (2/4)

Poursuivons les calculs :

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

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

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

## Les conséquences du problème (3/4)

Comme  $w$  est supposé sans corrélation avec les  $X$  :

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

Ainsi :

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

Au final, on a :

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

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

## Les conséquences du problème (4/4)

De tous les estimateurs de classe  $\lambda$  (estimateur du modèle empilé, *within*, FGLS...), seul l'estimateur *within* est sans biais et convergent lorsque les effets individuels  $\alpha_i$  sont corrélés avec les régresseurs.

Intuitivement, en éliminant la source de la corrélation, c'est-à-dire les  $\alpha_i$ , on a éliminé la corrélation entre l'erreur composée et les variables explicatives. La convergence est par ailleurs également vérifiée pour l'estimateur du modèle en **différence première**, qui n'est pas un estimateur de classe  $\lambda$ .

Convergence de  $\hat{\beta}$ 

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

# Plan

- 1 Introduction
- 2 Estimateurs de classe  $\lambda$
- 3 Modèle à erreurs corrélées
- 4 Repérer le problème**
- 5 Les solutions au problème

## Le test d'Hausman

L'objectif est de tester une hypothèse d'absence de corrélation. 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  $\alpha$ . 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 sont très différentes, au moins l'une des estimations n'est pas convergente et on rejette l'hypothèse d'exogénéité des  $X$  par rapport à  $\alpha$ .

# Statistique de test

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

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

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

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

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

## Exemple de test d'Hausman (R)

```

R> data("Gasoline", package = "plm")
R> wi <- plm(lgaspcar ~ lincomep + lrpmpg + lcarpcap, \\
  data = Gasoline, model = "within")
R> re <- plm(lgaspcar ~ lincomep + lrpmpg + lcarpcap, \\
  data = Gasoline, model = "random")
R> phtest(wi, re)

```

## Hausman Test

```

data:  lgaspcar ~ lincomep + lrpmpg + lcarpcap
chisq = 302.8037, df = 3, p-value < 2.2e-16
alternative hypothesis: one model is inconsistent

```



# Plan

- 1 Introduction
- 2 Estimateurs de classe  $\lambda$
- 3 Modèle à erreurs corrélées
- 4 Repérer le problème
- 5 Les solutions au problème**

# Les solutions au problème

La présence de corrélation entre les  $\alpha_j$  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
- avoir recours à des techniques de variables instrumentales

On va voir brièvement ici l'idée générale des techniques de **variables instrumentales**. Les méthodes de variables instrumentales pouvant être utilisées dans des contextes très différents, on va se limiter à survol rapide de ce domaine très riche.

## Principe général des variables instrumentales

$$Y = X\beta + \epsilon, E(X' \epsilon) \neq 0.$$

Les erreurs sont corrélées avec les explicatives et les estimateurs habituels sont biaisés. L'idée est de trouver des variables  $Z$ , que l'on va appeler "instruments" telles que :

- $E(Z' \epsilon) = 0$ , les instruments sont sans corrélation avec le terme d'erreur. On dira alors qu'ils sont **valides**,
- $E(Z' X) \neq 0$ , les instruments sont corrélés avec les variables endogènes. S'ils ne le sont pas, on parlera d'instruments **faibles**.

Chacune de ces deux propriétés peut être testée (test de Sargan, de Stock et Yogo...). Dans la pratique, trouver des variables  $Z$  satisfaisant ces deux propriétés est difficile et demande souvent de tester des centaines d'ensembles de variables candidates...

## Mise en oeuvre (1/2)

Des dizaines de possibilités existent : doubles moindres carrés (2SLS), Hausman et Taylor (1981), Arellano et Bond (1991), Blundell et Bond (1998)...

Les 2SLS restent la méthode la plus intuitive, constituée de deux étapes :

- dans une première étape, on régresse les variables explicatives endogènes sur les instruments et les variables explicatives exogènes. On calcule les valeurs prédites des endogènes,
- dans une seconde étape, on régresse  $Y$  sur les valeurs prédites des endogènes et les variables exogènes.

Intuitivement, on remplace les variables problématiques par des valeurs prédites avec les instruments. Si les instruments sont valides, les prévisions seront sans corrélation avec  $\epsilon$ . Si les instruments sont faibles, on peut montrer que les résultats de deuxième étape seront peu précis et peu robustes (vraisemblance “plate” pour le modèle de la deuxième étape).



## Mise en oeuvre (2/2)

Les subtilités sont nombreuses (correction de la matrice de variance lors de la deuxième étape, gestion des variables incluses/exclues, plusieurs tests possibles d'instruments faibles...). Dans le cas des variables instrumentales, le mieux est de ne jamais se lancer dans la programmation manuelle de ce type d'estimateurs et de toujours utiliser des **routines préprogrammées** et déjà testées. A ma connaissance, Stata est le seul logiciel à proposer des fonctions pour un grand nombre de modèles, qui fournissent de surcroît tous les critères et tests appropriés (routines `ivreg2` et `xtivreg2`, à installer manuellement).