Université PARIS-XII — Faculté de Sciences économiques et de gestion
Master *Sciences économiques*
Mention *Économie quantitative et expertise internationale*
Économétrie III — *F. LEGENDRE* — année universitaire 2005-06

Économétrie des données de panel
Résumé de quelques points difficiles du cours

1 L'hypothèse d'erreurs composées

Le modèle, linéaire, à erreurs composées s'écrit (si les données sont relatives à N individus observés pendant T périodes) :

$$\underline{y} = \mathrm{X}\underline{a} + \underline{u}$$

où $\underline{y}$ est un vecteur de taille NT×1 correspondant à la variable à expliquer, X une matrice de taille NT×k correspondant aux k variables explicatives du modèle, $\underline{a}$ le vecteur des coefficients (inconnus, à estimer) et $\underline{u}$ un vecteur correspondant aux termes d'erreurs.

L'hypothèse d'erreurs composées revient à introduire un effet aléatoire individuel ; c'est-à-dire à spécifier les termes d'erreurs de la façon suivante :

$$u_{it} = \mu_i + v_{it}$$

où μ_i est l'effet aléatoire propre à l'individu i ($i = 1,\ldots,\mathrm{N}$). La matrice de variance-covariance de $\underline{u}$, $\mathrm{V}(\underline{u})$, n'est alors pas de la forme $\sigma^2 \mathrm{I}_{\mathrm{NT}}$ (où I_{NT} est la matrice identité de taille NT×NT).

2 Forme de la matrice de variance-covariance de $\underline{u}$, $\mathrm{V}(\underline{u})$

Pour expliciter la forme de $\mathrm{V}(\underline{u})$, il faut, d'abord, s'accorder sur la façon dont sont rangées les observations. On retient généralement l'ordre individus/périodes. Ainsi, le vecteur $\underline{z}$ se bloc-décompose de la façon suivante :

$$\underline{z} = \begin{pmatrix} \underline{z}_1 \\ \underline{z}_2 \\ \vdots \\ \underline{z}_i \\ \vdots \\ \underline{z}_{\mathrm{N}} \end{pmatrix}$$

où $\underline{z}_i$ est le vecteur de taille T×1 regroupant toutes les observations relatives à l'individu i. Il faut, ensuite, détailler les hypothèses retenues sur les termes d'erreurs. On suppose ainsi :

$$\begin{aligned}
\mathrm{E}(\mu_i) &= 0 \quad \forall\, i \\
\mathrm{E}(v_{it}) &= 0 \quad \forall\, i,t \\
\mathrm{V}(\mu_i) &= \sigma^2_\mu \quad \forall\, i \\
\mathrm{V}(v_{it}) &= \sigma^2_v \quad \forall\, i,t \\
\mathrm{Cov}(\mu_i,\mu_j) &= 0 \quad \forall\, i \neq j \\
\mathrm{Cov}(v_{it},v_{jt'}) &= 0 \quad \forall\, i \neq j, t \neq t' \\
\mathrm{Cov}(\mu_i,v_{jt}) &= 0 \quad \forall\, i,j,t
\end{aligned}$$

Ces hypothèses permettent de remarquer que :

$$\mathrm{V}(\underline{u}_i) = \Sigma \;\; \forall\, i \quad \text{et} \quad \mathrm{E}(\underline{u}_i \underline{u}'_j) = 0_{\mathrm{T}} \;\; \forall\, i \neq j$$

où 0_{T} est la matrice de taille T×T dont tous les éléments sont nuls. Ainsi, la matrice de variance-covariance de $\underline{u}$ est de la forme :

$$\mathrm{V}(\underline{u}) = \begin{pmatrix} \Sigma & 0_{\mathrm{T}} & \cdots & 0_{\mathrm{T}} \\ 0_{\mathrm{T}} & \Sigma & \cdots & 0_{\mathrm{T}} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{\mathrm{T}} & 0_{\mathrm{T}} & \cdots & \Sigma \end{pmatrix} = \mathrm{I}_{\mathrm{N}} \otimes \Sigma$$

Il reste maintenant à détailler la structure de Σ :

$$\Sigma = \begin{pmatrix} \sigma^2_\mu + \sigma^2_v & \sigma^2_\mu & \cdots & \sigma^2_\mu \\ \sigma^2_\mu & \sigma^2_\mu + \sigma^2_v & \cdots & \sigma^2_\mu \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2_\mu & \sigma^2_\mu & \cdots & \sigma^2_\mu + \sigma^2_v \end{pmatrix} = \sigma^2_v \mathrm{W}_{\mathrm{T}} + (\mathrm{T}\sigma^2_\mu + \sigma^2_v)\mathrm{B}_{\mathrm{T}}$$

où W_{T} et B_{T} sont des matrices de taille T×T définies de la façon suivante :

$$\mathrm{W}_{\mathrm{T}} = \mathrm{I}_{\mathrm{T}} - \frac{1}{\mathrm{T}}\mathrm{J}_{\mathrm{T}} \quad \text{et} \quad \mathrm{B}_{\mathrm{T}} = \frac{1}{\mathrm{T}}\mathrm{J}_{\mathrm{T}}$$

avec J_{T} la matrice de taille T×T dont tous les éléments sont égaux à 1. En notant $\underline{e}_{\mathrm{T}}$ le vecteur de taille T×1 dont touts les éléments sont égaux à 1, la matrice J_{T} s'écrit sous la forme $\underline{e}_{\mathrm{T}}\underline{e}'_{\mathrm{T}}$. En définissant les matrices W et B (de taille NT×NT), $\mathrm{W} = \mathrm{I}_{\mathrm{N}} \otimes \mathrm{W}_{\mathrm{T}}$ et $\mathrm{B} = \mathrm{I}_{\mathrm{N}} \otimes \mathrm{B}_{\mathrm{T}}$, on obtient finalement :

$$\mathrm{V}(\underline{u}) = \sigma^2_v \mathrm{W} + (\mathrm{T}\sigma^2_\mu + \sigma^2_v)\mathrm{B}$$

3 Interprétation des matrices W et B

Ces deux matrices correspondent aux transformations suivantes. Le vecteur $W\underline{z}$ est de terme $z_{it} - \overline{z}_i$ où $\overline{z}_i$ est la moyenne de $\underline{z}$ pour l'individu i ($\overline{z}_i = \left[\sum_{t=1}^{T} z_{it}\right]/T$). Le vecteur $B\underline{z}$ est de terme $\overline{z}_i$. Mais, de plus, ces deux matrices sont idempotentes et symétriques ; elles représentent donc des projections orthogonales, « ennemies » en outre :

$$W^2 = W, \;\; W' = W, \;\; B^2 = B, \;\; B' = B, \;\; I_{NT} = W + B \quad \text{et} \quad W'B = 0_{NT}$$

On a ainsi :

$$\underline{z} = W\underline{z} + B\underline{z} \quad \text{et} \quad \underline{z}'\underline{z} = \underline{z}'W\underline{z} + \underline{z}'B\underline{z}$$

Cette dernière égalité est très importante, elle correspond, pour un vecteur centré, à la décomposition de la variabilité suivante :

$$\frac{\sum_{i=1}^{N}\sum_{t=1}^{T} z_{it}^2}{NT} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(z_{it} - \overline{z}_i)^2}{NT} + \frac{\sum_{i=1}^{N}\overline{z}_i^2}{N}$$

La variabilité totale se décompose en une variabilité intra-individuelle et une variabilité inter-individuelle. W reçoit ainsi le nom de within (intra en français) et B le nom de between (inter en français).

4 L'estimateur du maximum de vraisemblance

A cause de la forme particulière de la matrice de variance-covariance de $\underline{u}$, $V(\underline{u})$, la méthode d'estimation des moindres carrés ordinaires ne jouit plus de ses bonnes propriétés, comme dans le cas habituel. Elle reste convergente, mais ne produit pas les estimateurs asymptotiquement efficaces. Il convient alors de rechercher les estimateurs du maximum de vraisemblance. La vraisemblance du modèle, si les termes d'erreurs sont gaussiens, est de la forme suivante :

$$\mathscr{L}(\underline{a}, \sigma_\mu^2, \sigma_\nu^2) = (2\pi)^{-NT/2}|\Omega|^{-1/2}\exp\left\{-\frac{1}{2}(\underline{y} - X\underline{a})'\Omega^{-1}(\underline{y} - X\underline{a})\right\}$$

Il s'avère plus judicieux de paramétrer Ω de la façon suivante :

$$\Omega = \lambda_1 W + \lambda_2 B \quad \text{avec} \quad \lambda_1 = \sigma_\nu^2 \;\; \text{et} \;\; \lambda_2 = (T\sigma_\mu^2 + \sigma_\nu^2)$$

On a alors :

$$|\Omega| = \lambda_1^{N(T-1)}\lambda_2^{N} \quad \text{et} \quad \Omega^{-1} = \lambda_1^{-1}W + \lambda_2^{-1}B$$

La log-vraisemblance se réécrit ainsi (à une constante près) :

$$-\frac{1}{2}N(T-1)\log\lambda_1 - \frac{1}{2}N\log\lambda_2 - \frac{1}{2\lambda_1}(\underline{y} - X\underline{a})'W(\underline{y} - X\underline{a}) - \frac{1}{2\lambda_2}(\underline{y} - X\underline{a})'B(\underline{y} - X\underline{a})$$

Les estimateurs du maximum de vraisemblance, notés $\widehat{\underline{a}}_{MV}$, $\widehat{\lambda}_{1\,MV}$ et $\widehat{\lambda}_{2\,MV}$ sont les solutions du système suivant (non linéaire) :

$$\widehat{\underline{a}}_{MV} = \left(\frac{1}{\widehat{\lambda}_{1\,MV}}X'WX + \frac{1}{\widehat{\lambda}_{2\,MV}}X'BX\right)^{-1}\left(\frac{1}{\widehat{\lambda}_{1\,MV}}X'W\underline{y} + \frac{1}{\widehat{\lambda}_{2\,MV}}X'B\underline{y}\right)$$

$$\widehat{\lambda}_{1\,MV} = \frac{(\underline{y} - X\widehat{\underline{a}}_{MV})'W(\underline{y} - X\widehat{\underline{a}}_{MV})}{N(T-1)}$$

$$\widehat{\lambda}_{1\,MV} = \frac{(\underline{y} - X\widehat{\underline{a}}_{MV})'B(\underline{y} - X\widehat{\underline{a}}_{MV})}{N}$$

Ce système peut être résolu soit analytiquement (si le nombre de variables explicatives n'est pas trop élevé) soit numériquement. Bien sûr, c'est compliqué et il est permis de rechercher des estimateurs qui soient plus simples à mettre en œuvre.

5 L'estimateur des moindres carrés généralisés

Comme nous l'avons déjà vu, l'inverse de la matrice de variance-covariance du terme d'erreurs est

$$\Omega^{-1} = \lambda_1^{-1}W + \lambda_2^{-1}B$$

Cela nous suggère de pré-multiplier le modèle par la matrice suivante :

$$\lambda_1^{-1/2}W + \lambda_2^{-1/2}B$$

On a alors

$$V[(\lambda_1^{-1/2}W + \lambda_2^{-1/2}B)\underline{u}] = E[(\lambda_1^{-1/2}W + \lambda_2^{-1/2}B)\underline{u}][(\lambda_1^{-1/2}W + \lambda_2^{-1/2}B)\underline{u}]' =$$

$$[(\lambda_1^{-1/2}W + \lambda_2^{-1/2}B)]V(\underline{u})[(\lambda_1^{-1/2}W + \lambda_2^{-1/2}B)] = I_{NT}$$

Cette transformation est cependant excessive. Il suffit d'utiliser la transformation suivante :

$$W + \frac{\lambda_2^{-1/2}}{\lambda_1^{-1/2}}B = W + \sqrt{\theta}B \quad \text{avec} \quad \theta = \frac{\lambda_1}{\lambda_2}$$

On a en effet

$$V[(W+\sqrt{\theta}B)\underline{u}] = \lambda_1 I_{NT}$$

La matrice de variance-covariance est bien de la forme $\sigma^2 I_{NT}$. Cette dernière transformation est très simple ; elle prend la forme suivante :

$$\widetilde{z}_{it} = z_{it} - \overline{z}_i + \sqrt{\theta}\,\overline{z}_i = z_{it} - (1-\sqrt{\theta})\overline{z}_i$$

L'on voit donc que si θ est égal à 1 — c'est-à-dire si $\lambda_1 = \lambda_2$ c'est-à-dire si $\sigma^2_\mu = 0$ c'est-à-dire en absence d'un effet individuel — alors $\widetilde{z}_{it} = z_{it}$.

6 L'estimateur des moindres carrés quasi-généralisés

La mise en œuvre des MCQG est détaillée dans la fiche de travail n°4. La matrice de variance-covariance de l'estimateur des MCO n'est pas égale, sous l'hypothèse d'erreurs composées, à $\sigma^2(X'X)^{-1}$. La mise en œuvre des MCQG est beaucoup plus motivée par le fait que les MCO ne permettent pas d'effectuer des tests que par le gain en efficacité.

La première étape des MCQG repose sur l'obtention d'un estimateur convergent de θ. On peut prendre

$$\hat{\theta} = \frac{SCR_W/[NT-N-k]}{SCR_B/(N-k)}$$

comme cela est justifié dans la fiche de travail n°4.

Dans une seconde étape, il faut transformer les données en utilisant $\hat{\theta}$:

$$z^\star_{it} = z_{it} - (1-\sqrt{\hat{\theta}})\overline{z}_i$$

Le modèle devient ainsi

$$\underline{y}^\star = X^\star \underline{a} + \underline{u}^\star$$

L'estimateur des MCQG du vecteur $\underline{a}$ est l'estimateur des MCO de ce modèle :

$$\underline{\widehat{a}}_{MCQG} = (X^{\star\prime}X^\star)^{-1}X^{\star\prime}\underline{y}^\star$$