Chapitre 1, Données de panel_ Notions de cours et Applications :

Etude de l'impact du tourisme sur la croissance économique

I. Notions de Cours : Différentes alternatives de l'analyse pooled

1. Modèle pooled MCO :

C'est un modèle à effet uniforme, où on néglige la nature des données transversales et chronologiques ; c'est un modèle à effet uniforme à $NT-(K+1)$ degrés de liberté d'estimation (MCO) du modèle.

$$y_{it} = \mu + \sum_{k=1}^{K} X_{k\,it}\,\beta_k + \varepsilon_{it} \qquad \forall\, i = 1, \cdots, N \; et \; \forall\, t = 1, \cdots, T$$

2. Modèle de panel à Effet fixe, MEF, (appelé aussi Modèle de la Covariance) : One-Way Fixed Effects Model

Ce modèle est couramment utilisé dans le cadre des données de panel <u>où la dimension individuelle/spatiale est très importante.</u>

C'est un modèle qui tient compte des spécificités des individus (i.e. de l'hétérogénéité des comportements des individus) sous forme d'effets fixes en spécifiant un terme constant, μ_i, spécifique à chaque individu/unité spatiale.

Le modèle MEF est défini par l'équation de comportement de l'individu i observé à l'instant t suivante,

$$(1) \qquad y_{it} = \mu_i + \sum_{k=1}^{K} X_{k\,it}\,\beta_k + \varepsilon_{it} \qquad \forall\, i = 1, \cdots, N \; et \; t = 1, \cdots, T$$

où

$l'indice\ i\ définit\ la\ dimension\ individuelle/cross-section\ des\ unités\ spatiales;$

$L'indice\ t\ définit\ la\ dimension\ temporelle;$

$y_{it}\ vecteur\ (1,1)\ dépendant\ des\ observations\ individuelles-temporelles;$

$X_{it}\ vecteur\ (1,K) des\ régresseurs;$

$\beta\ vecteur(K,1)\ des\ paramètres\ structurels\ inconnus;$

$\mu_i\ l'effet\ spécifique\ spatial/individuel, local\ observ$é et invariant dans le temps;

$\varepsilon_{it} \sim BB(0, \sigma_\varepsilon^2)$

D'hypothèses :

$$E(\varepsilon_{it}) = 0, \quad \forall\, i, t$$
$$E(\varepsilon_{it}\varepsilon_{js}) = \delta_{ts}\delta_{ij}\,\sigma^2 = \begin{cases} \sigma^2 & si\ i = j\ et\ t = s \\ 0 & ailleurs \end{cases}$$
$$avec$$
$$\delta_{ts} = \begin{cases} 1 & t = s \\ 0 & ailleurs \end{cases}$$
$$et$$
$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & ailleurs \end{cases}$$

Remarques:

- Le modèle (1) est connu aussi sous le nom de <u>modèle à variables muettes individuelles (variable dummy)</u> où l'hétérogénéité des comportements, μ_i, est explicitée par N variables binaires/indicatrices (<u>sous condition pratique que N n'est pas trop grand</u>). Il est défini comme suit,

$$y_{it} = \sum_{k=1}^{K} X_{k\,it}\,\beta_k + \sum_{i=1}^{N} \mu_i\, d_{it} + \varepsilon_{it} \qquad \forall\, i = 1, \cdots, N\ et\ t = 1, \cdots, T$$
$$où$$
$$\forall t, \quad d_{it} = \begin{cases} 1 & si\ l'observation\ est\ relative\ à\ l'individu\ \ i \\ 0 & sinon \end{cases}$$
$$alors,\ l'équation\ de\ comportement\ de\ l'individu\ i \text{ observée durant les T périodes, s'écrit comme suit}$$
$$y_{it} = \sum_{k=1}^{K} X_{k\,it}\,\beta_k + \mu_i + \varepsilon_{it} \qquad \forall\ t = 1, \cdots, T.$$

C'est un modèle :

- qui peut être estimé par les MCOs,
- les tests et les intervalles de confiance se calculent de la manière habituelle (en utilisant les erreurs-types robustes à l'hétéroscédasticité),
- difficile à utiliser pour un très grand nombre d'individus.

- Dans la modélisation des donnés de panel, il est possible de tenir compte aussi de l'effet spécifique temporel, λ_t, $\Longrightarrow$ Two-Way Fixed Effects Model défini par :

$$(1') \qquad y_{it} = \mu_i + \lambda_t + \sum_{k=1}^{K} X_{k\,it}\,\beta_k + \varepsilon_{it} \qquad \forall\, i = 1, \cdots, N\ et\ t = 1, \cdots, T$$
$$où$$
$$\mu_i\ effet\ spécifique\ spatial/individuel\,,local\ observé \text{ et invariant dans le temps;}$$
$$\lambda_t\ effet\ spécifique\ temporel, invariant\ dans\ l'espace;$$
$$\varepsilon_{it}\ \rightsquigarrow\ i.i.d.(0, \sigma_\varepsilon^2\,)$$

Rappelons aussi que μ_i, λ_t et ε_{it} sont des composantes de la variables dépendante, y_{it}, qui sont non expliqués par X_{it}.

$\overset{i)}{\rightarrow}$ L'estimation du modèle MEF se fait en trois étapes :

En 1ère étape, on centre le modèle MEF par l'opérateur Within, dit aussi opérateur intra-individuel ou projecteur orthogonal, défini par

$$W_{(NT,NT)} = I_N \otimes \left(I_T - \frac{J_T}{T}\right)$$
$$où$$
$$J_T = S_T S_T^{'} = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$$
$$et \quad S_T = [1]_{(T,1)}, vecteur\ unitaire.$$

Ceci revient à dévier et à transformer chaque variable par rapport à sa moyenne, impliquant ainsi un modèle centré/transformé défini par

$$y_{it} - \overline{y_{\imath .}} = \sum_{k=1}^{K} \beta_k \left(X_{k\,it} - \overline{X_{k\,\imath}}\right) + \varepsilon_{it} - \overline{\varepsilon_{\imath .}} \qquad \forall\ t = 1, \cdots, T.$$
$$\Leftrightarrow\ y_{it}^{W} = \sum_{k=1}^{K} \beta_k X_{k\,it}^{W} + \varepsilon_{it}^{W}$$
$$\Leftrightarrow\ y_{i}^{W} = X_{i}^{W} \beta + \varepsilon_{i}^{W}$$
$$avec\ \ y_{it}^{W} = y_{it} - \overline{y_{\imath .}} \ \text{ et } \ \overline{y_{\imath .}} = \frac{\sum_{t=1}^{T} y_{it}}{T} \text{ définit la moyenne de y pour l'individu i}$$
$$\Leftrightarrow qui\ est\ ré-\ écrit\ de\ façon\ compacte\ suivante:\ Y^W = X^W \beta + \varepsilon^W \qquad où\ Y^W = WY.$$

$\boldsymbol{Dans\ une\ 2ème\ étape},\ \ cette\ régression\ transformée\ par\ W_{(NT,NT)}\ est\ estimée\ par\ OLS$

$$\Longrightarrow\ \widehat{\beta_{(K,1)_{OLS}}} = \widehat{\beta_{(K,1)_W}} = \widehat{\beta_{(K,1)_{intra-\imath}}} = \left(X^{W\,'} X^W\right)^{-1} X^{W'} Y^W \overset{\text{def}}{=} \left(X^{'}WX\right)^{-1} X^{'}W\,Y$$
$$C^{'}est\ \ l^{'}estimateur\ \text{Within}\ \ \widehat{\beta_{(K,1)_W}}\ (sous\ sa\ dénomination\ anglaise);$$
$$il\ est\ meilleur\ estimateur\ linéaire\ sans\ biais.$$

$\boldsymbol{Dans\ une\ 3ème\ étape}, les\ effets\ spécifiques\ spatiaux\ seront\ estimés\ comme\ suit:$

$$\widehat{\mu_{\imath\,OLS}} = \overline{y_{\imath .}} - \sum_{k=1}^{K} \beta_k\ \overline{X_{k\,\imath}} = \frac{1}{T}\sum_{t=1}^{T}\left(y_{it} - \sum_{k=1}^{K} \beta_k\, X_{k\,it}\right) \qquad \forall\ i = 1, \cdots, N$$

Tout ce qui est attribuable aux différences constantes dans le temps entre les individus est exclu de l'estimateur $intra - i$. Il utilise uniquement l'information contenue dans les fluctuations observées pour chaque individu autour de son niveau moyen. Dans un 2ème temps, il est toujours possible de récupérer les effets fixes (et la constante, si elle existe). Cependant, dans de nombreuses estimations, notamment micro-économiques, l'objectif n'est pas d'identifier les effets spécifiques mais d'inférer (de déduire) les déterminants de la variable à expliquer en contrôlant pour l'hétérogénéité inobservée.

Egalement, on observe que le modèle $intra - i$ inclut aussi moins de variables à estimer que le modèle à variables dummy. De même, on doit noter que l'estimateur des MCOs du modèle à variables dummy est strictement identique à celui de l'estimateur $intra - i$.

L'avantage principal de cette procédure d'ajustement-centrée est que l'estimation du vecteur des paramètres, β, implique l'utilisation d'une matrice inverse de régresseurs d'ordre (K, K) (i.e. la matrice $\left(X^{c'}X^{c}\right)^{-1}$) plutôt que $(K+N, K+N)$ qui est connu comme un ralentissement de l'estimation et une détérioration considérable de l'exactitude et de la précision de l'estimation surtout lorsque N est large.

Sous l'hypothèse supplémentaire que les perturbations ε sont normales, l'estimateur within, $\widehat{\beta_{(K,1)_W}}$, est également NORMAL tel que :

$$\widehat{\beta_{(K,1)_W}} \sim \mathcal{N}\left(\beta, \sigma_W^2 (X'WX)^{-1}\right)$$

$$et$$

$$[NT-(N+K)]\frac{\widehat{\sigma_W^2}}{\sigma_W^2} \sim \mathcal{X}^2_{[NT-(N+K)]}$$

$$Rappelons\ que$$

$$V\left(\widehat{\beta_{(K,1)_W}}\right) = \sigma_W^2 (X'^W X^W)^{-1} = \sigma_W^2 (X'WX)^{-1} = \sigma_W^2 \left[\sum_i \sum_t (X_{it} - X_{i.})'(X_{it} - X_{i.})\right]^{-1}$$

$$et$$

$$\widehat{\sigma_{OLS}^2} = \widehat{\sigma_W^2} = \frac{\widehat{\varepsilon_W}'\widehat{\varepsilon_W}}{NT-(N+K)} = \frac{\left(Y^W - X^W \widehat{\beta_{MCO}}\right)'\left(Y^W - X^W \widehat{\beta_{MCO}}\right)}{NT-(N+K)}$$

$$= \frac{\left(WY - WX\widehat{\beta_{MCO}}\right)'\left(WY - WX\widehat{\beta_{MCO}}\right)}{NT-(N+K)}$$

$où\ \sigma_W^2\ est\ estimé\ par\ la\ somme\ des\ carrés\ des\ résidus\ de\ la\ régression\ intra - individuelle$ $divisée\ par\ le\ nbre\ de\ degrés\ de\ liberté\ associés\ à\ l'estimationdu\ modèle\ MEF, soit\ NT-(N+K).$

Aussi, on a

$$* \ E(\varepsilon_{it}^{W}) = E(\varepsilon_{it} - \overline{\varepsilon_{1.}}) = E(\varepsilon_{it}) - E(\overline{\varepsilon_{1.}}) = 0 \quad où \quad \varepsilon_{it} - \overline{\varepsilon_{1.}} = \varepsilon_{it}^{W}$$

$$* \ E\left(\varepsilon_{it}^{W^2}\right) = E[(\varepsilon_{it} - \overline{\varepsilon_{1.}})(\varepsilon_{it} - \overline{\varepsilon_{1.}})']$$

$$= V(\varepsilon_{it}) + V(\overline{\varepsilon_{1.}}) - 2\, Cov(\varepsilon_{it}, \overline{\varepsilon_{1.}})$$

$$= \sigma_{\varepsilon}^2 + \frac{\sigma_{\varepsilon}^2}{T} - \frac{2}{T} Cov(\varepsilon_{it}, \varepsilon_{it})$$

$$= \sigma_{\varepsilon}^2 + \frac{\sigma_{\varepsilon}^2}{T} - \frac{2}{T}\, V(\varepsilon_{it})$$

$$= \sigma_{\varepsilon}^2 + \frac{\sigma_{\varepsilon}^2}{T} - \frac{2}{T} \sigma_{\varepsilon}^2$$

$$\Leftrightarrow \ V(\varepsilon_{it}^{W}) = E\left(\varepsilon_{it}^{W^2}\right) = \sigma_{\varepsilon}^2 \left(1 - \frac{1}{T}\right)$$

$$Variance\ résiduelle\ de\ la\ dimension\ intra - individuelle.$$

3. Modèle Moindres Carrés à variables dummy_ LSDV modèle :

L’équation de comportement (1) de l’individu i observé durant les T périodes est définie

$$(2) \qquad y_i = S_T\, \mu_i + X_i\, \beta + \varepsilon_i \qquad \forall\ i = 1, \cdots, N$$

$$avec$$

$$y_i = [y_{it}]_{(\mathrm{T},1)}; \quad S_T = [1]_{(T,1)}; \quad X_i = \begin{bmatrix} X_{1\,i1} & X_{2\,i1} & \cdots & X_{K\,i1} \\ X_{1\,i2} & X_{2\,i2} & \cdots & X_{K\,i2} \\ \vdots & \vdots & \cdots & \vdots \\ X_{1\,iT} & X_{2\,iT} & \cdots & X_{KiT} \end{bmatrix} = [X_{k\,it}]_{(\mathrm{T,K})}$$

$$\beta_{(K,1)} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_K \end{pmatrix}; \quad \varepsilon_i = [\varepsilon_{it}]_{(\mathrm{T},1)}$$

A un niveau plus agrégé encore où on empile les N individus observés chacun sur les T périodes, on définit la forme compacte du modèle:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{bmatrix} S_T & & \square \\ & \ddots & \\ \square & & S_T \end{bmatrix} \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_N \end{bmatrix} + \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix}$$

$$\Leftrightarrow$$

$$Y_{(NT,1)} = (I_N \otimes S_T)\ \alpha_{(N,1)} + X_{(NT,K)}\ \beta_{(K,1)} + \varepsilon_{(NT,1)}$$

$$\Leftrightarrow$$

$$Y_{(NT,1)} = \ D_{(NT,N)}\ \alpha_{(N,1)} + X_{(NT,K)}\ \beta_{(K,1)} + \varepsilon_{(NT,1)} \qquad (3)$$

$$\Leftrightarrow$$

$$Y_{(NT,1)} = \ [D_{(NT,N)} \quad X_{(NT,K)}] \begin{bmatrix} \alpha_{(N,1)} \\ \beta_{(K,1)} \end{bmatrix}_{(N+K,1)} + \varepsilon_{(NT,1)}$$

$$\Leftrightarrow$$

$$Y_{(NT,1)} = Z_{(NT,\ \ N+K)}\ \delta_{(N+K,1)} + \varepsilon_{(NT,1)} \qquad (4)$$

$$avec$$

$$D_{(NT,N)} matrice\ de\ N\ variables\ indicatrices/muettes\ individuelles,$$

$$D_{(NT,N)} = \begin{bmatrix} d_1 & \cdots d_{i_{(NT,1)}} \cdots & d_N \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \cdots & \cdots & 0 \\ \vdots & & & & \vdots \\ \vdots & & & & 0 \\ 0 & & \ddots & & 1 \\ \vdots & & & & \vdots \\ 0 & & & & 1 \end{bmatrix} = \begin{bmatrix} S_{T_{(T,1)}} & 0_{(T,1)} & \cdots & 0_{(T,1)} \\ 0_{(T,1)} & S_{T_{(T,1)}} & \cdots & 0_{(T,1)} \\ \vdots & & \ddots & \vdots \\ 0_{(T,1)} & \cdots & 0_{(T,1)} & S_{T_{(T,1)}} \end{bmatrix} = \begin{bmatrix} S_T & & \square \\ & \ddots & \\ \square & & S_T \end{bmatrix}$$

$$X_{(NT,K)} = \begin{bmatrix} x_{111} & \cdots & x_{K11} \\ \cdot & \cdots & \cdot \\ x_{11T} & \cdots & x_{K1T} \\ x_{121} & \cdots & x_{K21} \\ \cdot & \cdots & \cdot \\ x_{12T} & \cdots & x_{K2T} \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ x_{1N1} & \cdots & x_{KN1} \\ \cdot & \cdots & \cdot \\ x_{1NT} & \cdots & x_{KNT} \end{bmatrix}$$

$$On\ a\ \alpha_{(N,1)} vecteur\ des\ N\ paramètres\ spécifiques\ spatiaux,$$

$$\beta_{(K,1)} vecteur\ de\ K\ paramètres\ uniformes.$$

$$RQUE: Pour\ plus\ de\ précision\ économétrique, on\ a\ concaténé(i.e. enchaîner):$$

$$*\ Horizontalement\ les\ matrices\ X_{(NT,K)}\ et\ D_{(NT,N)} \Rightarrow\ Z_{(NT,\ \ N+K)} = [D_{(NT,N)} \quad X_{(NT,K)}]$$

$$et$$

$$*\ Verticalement, les\ vecteurs\ des\ paramètres\ \alpha_{(N,1)} et\ \ \beta_{(K,1)} \Rightarrow\ \delta_{(N+K,1)} = \begin{bmatrix} \alpha_{(N,1)} \\ \beta_{(K,1)} \end{bmatrix}_{(N+K,1)}.$$

Comme la matrice $X_{(NT,K)}$ ne comporte pas de vecteur unitaire (car de façon logique, on considère un modèle sans constante globale puisque des constantes individuelles sont déjà prises en compte), la matrice $Z_{(NT,\ \ N+K)} = [I_N \otimes S_T \quad X_{(NT,K)}]$ des variables explicatives est de plein rang colonne (on a indépendance dans les vecteurs colonnes). Ainsi, on remarque que rien ne s'oppose, théoriquement à ce que les coefficients du modèle (4) (i.e. les K coefficients uniformes de β et les N coefficients spécifiques spatiaux de α) soient estimés par les MCOs,

$$\widehat{\delta_{MCO}} = \begin{bmatrix} \widehat{\alpha_{(N,1)}} \\ \widehat{\beta_{(K,1)}} \end{bmatrix}_{(N+K,1)} = (Z'Z)^{-1}Z'Y = \widehat{\delta_{LSDV}} \qquad (5)$$

$$\hat{V}\left(\widehat{\delta_{MCO}}\right) = \widehat{\sigma^2}(Z'Z)^{-1}$$

$$avec$$

$$\widehat{\sigma^2} = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{NT-(N+K)} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{\varepsilon_{it}^2}}{NT-(N+K)}$$

$$et$$

$$\widehat{\varepsilon_{it}} = y_{it} - \hat{\mu}_i - \sum_{k=1}^{K} X_{k\,it}\,\hat{\beta}_k$$

L'estimateur, $\widehat{\delta_{MCO}} = \begin{bmatrix} \widehat{\alpha_{(N,1)}} \\ \widehat{\beta_{(K,1)}} \end{bmatrix}_{(N+K,1)}$, obtenu ainsi directement, est connu comme estimateur LSDV, noté $\widehat{\delta_{LSDV}}$ (Least Square with Dummy Variables), où les estimations des constantes spécifiques, μ_i, sont obtenues directement en régressant, en particulier, la variable Y sur un ensemble de variables binaires/dummy, défini par la matrice $D_{(NT,N)}$, qui permet d'identifier les différents individus.

Sur le plan pratique appliqué, l'estimation des effets spécifiques individuels, les μ_i, ne se fait pas d'une manière directe selon l'expression (5), compte tenu de l'importance de la dimension individuelle lorsqu'on utilise les données de panel (où on peut se heurter à un problème de capacité de calcul) car l'estimation de $\widehat{\delta_{MCO}} = \begin{bmatrix} \widehat{\alpha_{(N,1)}} \\ \widehat{\beta_{(K,1)}} \end{bmatrix}_{(N+K,1)}$ exige l'inversion de la matrice $(Z'Z)_{(N+K,N+K)}$ et pour N large cette opération devient non attractive, non motivante.

De ce fait, il est nécessaire d'utiliser la technique de régression partitionnée ou de manière équivalente appliquer le théorème de Frish-Waugh-Lovell qui donne les mêmes résultats et qui utilise l'inversion d'une matrice d'ordre juste (K, K). On procède, ainsi, préalablement à l'estimation du $(K, 1)$ vecteur des paramètres uniformes, $\beta_{(K,1)}$, en transformant l'expression (3) par la matrice régulière (i.e. idempotente et symétrique) M_D, où

$$M_D = I_{NT} - D(D'D)^{-1}D'$$
$$où$$
$$\begin{cases} M_D = M_D{}' \\ M_D{}^2 = M_D \\ M_D D = 0 \end{cases}$$

Alors, l'approche par le théorème de Frish-Waugh-Lovell est :

$$\xRightarrow{M_D \text{appliquée sur le Modèle } (3):\, Y_{(NT,1)} = D_{(NT,N)}\, \alpha_{(N,1)} + X_{(NT,K)}\, \beta_{(K,1)} + \varepsilon_{(NT,1)}} M_D Y = M_D\, X\beta + M_D \varepsilon \qquad (6)$$
$$\Rightarrow$$
$$\widehat{\beta_{COV}} = \ (X' M_D\, X)^{-1} X' M_D Y$$

Et,

$$\widehat{\alpha_{(N,1)}} = (D'D)^{-1}D'\left(Y - X\,\widehat{\beta_{COV}}\right) = \frac{1}{T}D'\left(Y - X\,\widehat{\beta_{COV}}\right) = \begin{bmatrix} \overline{Y_{1.}} \\ \vdots \\ \overline{Y_{\iota.}} \\ \vdots \\ \overline{Y_{N.}} \end{bmatrix} - \begin{bmatrix} \overline{X_{1,1.}} & \overline{X_{2,1.}} & \overline{X_{K,1.}} \\ \vdots & \vdots & \vdots \\ \overline{X_{1,\iota.}} & \overline{X_{2,\iota.}} & \overline{X_{K,\iota.}} \\ \vdots & \vdots & \vdots \\ \overline{X_{1,N.}} & \overline{X_{2,N.}} & \overline{X_{K,N.}} \end{bmatrix} \begin{bmatrix} \widehat{\beta_1} \\ \vdots \\ \widehat{\beta_K} \end{bmatrix}$$

$$où$$

$$\frac{1}{T}D'Y = [\overline{Y_1} \cdots \overline{Y_N}] \quad et \quad \frac{1}{T}D'X = [\overline{X_1} \cdots \overline{X_N}]'$$

⇢ <u>Identification de la matrice M_D</u> : On a

$$M_D = I_{NT} - D(D'D)^{-1}D' \;\; avec \;\; D_{(NT,N)} = \; I_N \otimes S_T$$
$$\Leftrightarrow$$
$$\begin{aligned} M_D = \; & I_N \otimes I_T - (I_N \otimes S_T)\, [(I_N{}' \otimes S_T{}')(I_N \otimes S_T)]^{-1}(I_N{}' \otimes S_T{}') \\ & = I_N \otimes I_T - (I_N \otimes S_T)\, [I_N \otimes S_T{}' S_T]^{-1}(I_N{}' \otimes S_T{}') \\ & = I_N \otimes I_T - (I_N \otimes S_T)\, [T I_N]^{-1}(I_N{}' \otimes S_T{}') \\ & = I_N \otimes I_T - \left(I_N \otimes \frac{S_T S_T{}'}{T}\right) \end{aligned}$$
$$\Rightarrow$$
$$M_D = I_N \otimes \left(I_T - \frac{J_T}{T}\right) = W_{(NT,NT)}$$
$$\xRightarrow{alors}$$
$$\widehat{\beta_{COV}} = \ (X' M_D\, X)^{-1} X' M_D Y = (X' W\, X)^{-1} X' W Y = \ \widehat{\beta_W}$$

Remarquons que l'estimateur $\widehat{\beta_{COV}}$ du paramètre β dans le cadre du modèle de la covariance, où la spécificité individuelle/spatiale se présente sous forme d'un effet fixe, est égale à l'estimateur utilisant la transformation Within, $W_{(NT,NT)}$, dans le cadre d'un modèle à erreurs composées élémentaires, i.e. Modèle MEF. A ce niveau, l'estimation du paramètre β est définie en terme du théorème de Frish-Waugh-Lovell.

On rappelle l'énoncé du théorème de Frish-Waugh-Lovell :

L'inférence est conditionnelle sur les effets individuels où l'estimation est obtenue en régressant Y sur X et sur variables dummy D.

Les estimations du paramètre β sont numériquement identiques dans les deux procédures (i) et (ii) , i.e.

$$\widehat{\beta_{COV}} = \widehat{\beta_W}$$

$$o\text{ù}$$

$$M_D = I_{NT} - D(D'D)^{-1}D' \quad est\ le\ projecteur\ sur\ la\ vari\acute{e}t\acute{e}\ lin\acute{e}aire, i.e.\ I_{NT}, orthogonale\ \grave{a}\ la\ vari\acute{e}t\acute{e}\ lin\acute{e}aire\ engendr\acute{e}e\ par\ les\ D\ (i.e. la\ partie\ D(D'D)^{-1}D').$$

⇢ On ajoute, aussi, les autres formules associées aux estimateurs $\widehat{\beta_{COV}}$ et $\widehat{\alpha_{(N,1)}}$:

$$SCR_{W\ ou\ M_D} = \hat{\varepsilon}'_{M_D ou\, W}\hat{\varepsilon}_{M_D ou\, W} = Y'\,M_D Y - \widehat{\beta_{COV}}' X' M_D Y$$

$$\hat{\sigma}^2_{M_D ou\, W} = \frac{SCR_{W\ ou\ M_D}}{NT-(N+K)} \quad Variance\ estim\acute{e}e\ du\ terme\ d'erreursidiosyncratiques\ .$$

$$o\text{ù}$$

$$K\ d\acute{e}finit\ le\ nombre\ de\ r\acute{e}gresseurs\ (constante\ et\ variables\ muettes\ exclues);$$

$$\hat{\varepsilon}'_{M_D ou\, W}\hat{\varepsilon}_{M_D ou\, W} \quad est\ la\ somme\ des\ carr\acute{e}s\ des\ r\acute{e}sidus\ de\ la\ r\acute{e}gression\ intra.$$

$$Alors, on\ a$$

$$V\left(\widehat{\beta_{COV}}\right) = \hat{\sigma}^2_{M_D ou\, W}(X'M_D X)^{-1}$$

$$V\left(\widehat{\alpha_{(N,1)}}\right) = \frac{\hat{\sigma}^2_{M_D ou\, W}}{T}\ I_N + \frac{1}{T} D'X\, V\left(\widehat{\beta_{COV}}\right) X'D\,\frac{1}{T}$$

$$En\ particulier\ sur\ le\ mod\grave{e}le\ transform\acute{e}\ par\ M_D, on\ r\acute{e}cup\grave{e}re\ dans\ un\ 2\grave{e}me\ temps\ les\ effets\ fixe$$

$$\widehat{\alpha_\iota} = \overline{y_\iota} - \overline{x_\iota}'\widehat{\beta} = \overline{y_{\iota.}} - \widehat{\beta_1}\,\overline{x_{1,\iota.}} - \cdots - \widehat{\beta_K}\,\overline{x_{K,\iota.}}$$

$$V(\widehat{\alpha_\iota}) = \frac{1}{T}\hat{\sigma}^2_{M_D ou\, W} + \overline{x_\iota}'\,V(\hat{\beta})\,\overline{x_\iota}$$

$$Cov\left(\widehat{\alpha_\iota}, \widehat{\alpha_J}\right) = \overline{x_\iota}'\,V(\hat{\beta})\,\overline{x_J} \qquad o\text{ù}\ \overline{x_{\iota.}} = \left[\overline{x_{1,\iota.}}, \cdots, \overline{x_{K,\iota.}}\right]$$

Remarque:

Notons que la partie/variété linéaire engendrée par la matrice D est au fait que le projecteur orthogonal $B_{(NT,NT)}$:

$$D(D'D)^{-1}D' = (I_N \otimes S_T)\ [(I_N{}' \otimes S_T{}')(I_N \otimes S_T)]^{-1}(I_N{}' \otimes S_T{}')$$
$$= (I_N \otimes S_T)\ [I_N \otimes S_T{}'S_T]^{-1}(I_N{}' \otimes S_T{}')$$
$$= (I_N \otimes S_T)\ [TI_N]^{-1}(I_N{}' \otimes S_T{}')$$
$$= I_N \otimes \frac{S_T S_T{}'}{T}$$
$$= I_N \otimes \frac{J_T}{T} = B_{(NT,NT)}$$

⇢ <u>Propriétés de la matrice</u> $D_{(NT,N)}$:

$$DS_N = (I_N \otimes S_T)S_N = S_N \otimes S_T = S_{NT} \quad (Exhaustivité)$$
$$D'D = (I_N \otimes S_T)'(I_N \otimes S_T) = I_N \otimes S_T{}'S_T = TI_N \quad (Orthogonalité)$$
$$DD' = (I_N \otimes S_T)(I_N \otimes S_T)' = I_N \otimes S_T S_T{}' = I_N \otimes J_T$$
$$\frac{1}{T}D'Y = [\overline{Y_1} \cdots \overline{Y_N}]\ et\ \frac{1}{T}D'X = [\overline{X_1} \cdots \overline{X_N}]'$$
$$où\ \overline{Y_i} = \frac{\sum_{t=1}^{T} y_{it}}{T} \ \text{ la moyenne de y en l'unité spatiale i,}$$
$$\overline{X_i}' = \frac{\sum_{t=1}^{T} x_{it}{}'}{T} \ \text{ est un (K, 1) vecteur des moyennes individuelles des variables explicatives.}$$

4. Spécification du modèle, MEC et procédure d’estimation :

Au niveau du modèle MEC, la spécificité individuelle et temporelle apparaît sous forme stochastique car les agents économiques qu’on essaye de tester et de quantifier leurs comportements, peuvent présenter certaines spécificités non observables.

⇢ <u>Dans sa forme la plus générale</u>, le modèle MEC considère un effet spatial/individuel, u_i, et un effet temporel, v_t,

$$w_{it} = u_i + v_t + \varepsilon_{it}$$

$$Le\ résidu\ comprend\ deux\ termes\ u_i\ et\ v_t\ caractéristiques\ resp.\ de\ l'individu\ i\ et\ du\ temps\ t:$$
$$dans\ les\ variables\ omises\ (supposées\ indépendantes\ de\ X_{it})figure\ une\ (ou\ des)\ variable(s)$$
$$caractéristique(s)de\ l'individu\ i\ et/ou\ du\ temps\ t.$$
$$\varepsilon_{it}\ est\ une\ perturbation\ à\ caractère\ aléatoire\ habituel.$$

En effet, lorsqu’on estime des demandes de facteurs ou une fonction de production sur des données d’entreprises, par exemple, on peut envisager qu’une variable «qualité du management », non mesurable, intervienne dans la relation sous la forme d’un effet individuel/spatial spécifique figurant dans les perturbations. Pareillement, si l’on tente d'examiner, sur des panels d'individus, la rentabilité (en terme de salaire) des études, il est clair qu'un facteur personnel non évaluable, qui a trait aux qualités propres de l'individu, joue dans l'explication recherchée. L'omission, faute de pouvoir le mesurer, de ce facteur personnel dans la liste des variables explicatives conduit probablement à l'existence d'un effet individuel/spatial spécifique dans les perturbations.

L'effet temporel correspond à l'omission, dans la liste des variables explicatives, de variables dont la valeur est identique pour tous les individus en un point donné du temps : il peut s'agir, par exemple, du niveau des prix, des tendances de la conjoncture ou encore, de façon plus générale, du « climat » d'optimisme ou de pessimisme régnant dans le milieu des affaires.

⇢ <u>Ecriture du modèle MEC</u>,

L'équation de comportement de l'individu i observé à l'instant t est définie,

$$(13) \qquad y_{it} = \mu + \sum_{k=1}^{K} X_{k\,it}\,\beta_k + w_{it} \qquad \forall\, i = 1, \cdots, N \quad et \quad t = 1, \cdots, T$$

$$o\text{ù}$$

$$w_{it} = u_i + v_t + \varepsilon_{it}$$

u_i *effet stochastique, spécifique spatial/individuel, local observé* et invariant dans le temps;

v_t *effet stochastique, spécifique temporel, supposé négligeable*;

ε_{it} *est une perturbation à caractère aléatoire habituel*;

$$u_i \sim N(0, \sigma_u^2)$$

$$v_t \sim N(0, \sigma_v^2)$$

$$\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$$

Du fait de la non prise en compte de l'effet spécifique temporel dans les perturbations, on présente ainsi une version simplifiée du modèle MEC. Dans la pratique, cette simplification s'avère le plus souvent justifiée étant donné la faible quantité d'information en terme de variabilité qu'apporte la dimension temporelle.

⇢ <u>Hypothèses sur les composantes du terme d'erreurs,</u>

$$(13.1) \qquad * E(w_{it}) = 0, \quad \forall\, i, t$$

$$* E\left(w_{it} w_{js}\right) = \delta_{ij}\, \sigma_u^2 + \delta_{ij}\delta_{ts}\, \sigma_\varepsilon^2 = \begin{cases} \sigma_u^2 + \sigma_\varepsilon^2 \;\; si\ i = j\ et\ t = s\ (Homoscédasticité) \\ \sigma_u^2 \;\; si\ i = j\ et\ t \neq s\ (Equicorrélation\ des\ erreurs\ dans\ le\ temps \\ i.e.\ autocorrélation\ indépendante\ du\ temps) \\ 0 \;\; ailleurs\ (i.e.\ où\ i \neq j\ [et\ t = s];\ Corrélation \\ contemporaine\ nulle). \end{cases}$$

$$avec$$

$$\delta_{ts} = \begin{cases} 1 & t = s \\ 0 & ailleurs \end{cases}$$

$$et$$

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & ailleurs \end{cases}$$

$$* \;\; Hypothèse\ d'Orthogonalité: E[X_{k\,it} w_{it}] = 0 \;\Rightarrow hypothèse\ cruciale\ (importante\ et\ critique)\ qui\ signifie\ notamment\ l'absence\ de\ corrélation\ entre\ les\ effets\ spécifiques\ et\ les\ régresseurs.$$

<u>Remarque</u> : <u>En résumé</u>, on peut écrire le modèle MEC comme suit,

$$\begin{cases} E(y_{it}) = \sum_{\mathrm{k=1}}^{\mathrm{K}} X_{k\,it}\, \beta_k \\ \text{i.e. pas de prise en compte d'un effet spécifique au premier ordre.} \\ \mathrm{Cov}\left(y_{it} y_{js}\right) = \delta_{ij}\, \sigma_u^2 + \delta_{ij}\delta_{ts}\, \sigma_\varepsilon^2 \\ \text{i.e. la prise en compte d'un effet spécifique au deuxième ordre.} \end{cases}$$

La structure (13.1) conduit à définir pour une même unité spatiale i la matrice de variances-covariances des écarts comme suit :

$$(13.2) \qquad E\left(W_{i_{(T,1)}} W'_{i_{(1,T)}}\right) = E\left[\begin{pmatrix} w_{i1} \\ \vdots \\ w_{iT} \end{pmatrix} \begin{pmatrix} w_{i1} & \cdots & w_{iT} \end{pmatrix}'\right]$$

$$= \begin{bmatrix} \sigma_u^2 + \sigma_\varepsilon^2 & \sigma_u^2 & \cdots & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2 + \sigma_\varepsilon^2 & & \vdots \\ \vdots & & \ddots & \sigma_u^2 \\ \sigma_u^2 & \cdots & \sigma_u^2 & \sigma_u^2 + \sigma_\varepsilon^2 \end{bmatrix}$$

$$= \sigma_u^2\, J_T + \sigma_\varepsilon^2\, \mathrm{I_T}$$

$$= \Sigma_{(T,T)}$$

$$Remarquons\ que\ la\ valeur\ de\ E\left(W_{i_{(T,1)}} W'_{i_{(1,T)}}\right) = \Sigma_{(T,T)}\ \ est\ indépendante\ de\ l'unité\ spatiale\ \ i\ \ considérée.$$

Notons aussi que la structure de cette matrice var-covariances résiduelles constitue la particularité fondamentale du modèle dans la mesure où, en dehors de la diagonale, se trouvent des termes non nuls. Ces derniers matérialisent une autocorrélation particulière qui ne dépend pas du temps séparant deux observations, i.e. une autocorrélation intra-individuelle résiduelle attribuée par la présence de l'erreur spécifique u_i.

En empilant les données pour l'ensemble des observations, individu par individu, on définit la matrice de variances-covariances[1] comme suit,

$$\Omega_{(NT,NT)} = \begin{bmatrix} \Sigma_{(T,T)} & & \square \\ & \ddots & \\ \square & & \Sigma_{(T,T)} \end{bmatrix}$$

$$= \mathrm{I_N} \otimes \Sigma_{(T,T)}$$

$$\Rightarrow \text{L'absence de corrélation entre les individus (i.e. nullité des termes en dehors de la diagonale)}$$
$$\text{rend la structure de cette matrice relativement simple, en bloc diagonale.}$$

Mais, il est encore possible d'écrire cette matrice $\Omega_{(NT,NT)}$ de façon plus « parlante », en faisant apparaître les projecteurs orthogonaux, Between et Within. En effet, on a

[1] On rappelle que la matrice des perturbations est $w_{(NT,1)} = \begin{pmatrix} w_{11} \\ \vdots \\ w_{1T} \\ \vdots \\ w_{N1} \\ \vdots \\ w_{NT} \end{pmatrix}$; avec $w_{i_{(T,1)}} = S_T u_i + \varepsilon_i$

donc, $\Omega_{(NT,NT)} = E(ww') = E\begin{bmatrix} \begin{bmatrix} {w_{11}}^2 & \cdots & w_{11}w_{1T} \\ \vdots & \ddots & \vdots \\ w_{11}w_{1T} & \cdots & {w_{1T}}^2 \end{bmatrix} & \cdots & \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ \vdots & \ddots & \vdots \\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \cdots & \begin{bmatrix} {w_{N1}}^2 & \cdots & w_{N1}w_{NT} \\ \vdots & \ddots & \vdots \\ w_{N1}w_{NT} & \cdots & {w_{NT}}^2 \end{bmatrix} \end{bmatrix} = \mathrm{I_N} \otimes \Sigma_{(T,T)}$

Où, on a :

- Les covariances entre les perturbations de deux unités spatiales différentes sont nulles ;
- Les covariances entre les perturbations d'une même unité spatiale i mais à des dates différentes sont égales à σ_u^2 ;
- La variance du résidu relatif à l'unité spatiale i à l'instant t est égale à $\sigma_u^2 + \sigma_\varepsilon^2$.

$$E\left(W_{i_{(T,1)}}W'_{i_{(1,T)}}\right) = \Sigma_{(T,T)} = \sigma_u^2 J_T + \sigma_\varepsilon^2 \mathrm{I_T} = T\,\sigma_u^2 \frac{J_T}{T} + \sigma_\varepsilon^2 \frac{J_T}{T} - \sigma_\varepsilon^2 \frac{J_T}{T} + \sigma_\varepsilon^2 \mathrm{I_T}$$

$$= \sigma_\varepsilon^2 \left[\left(\mathrm{I_T} - \frac{J_T}{T}\right) + \frac{\mathrm{T}\,\sigma_u^2 + \sigma_\varepsilon^2}{\sigma_\varepsilon^2} \frac{J_T}{T}\right]$$

$$\text{Soit} \quad \Sigma_{(T,T)} = \sigma_\varepsilon^2 \left[\left(\mathrm{I_T} - \frac{J_T}{T}\right) + \frac{1}{\theta^2} \frac{J_T}{T}\right] \text{ avec } \theta^2 = \frac{\sigma_\varepsilon^2}{\mathrm{T}\,\sigma_u^2 + \sigma_\varepsilon^2}$$

Et

$$\Omega_{(NT,NT)} = \mathrm{I_N} \otimes \Sigma_{(T,T)} = \sigma_\varepsilon^2 \left[\mathrm{I_N} \otimes \left(\left(\mathrm{I_T} - \frac{J_T}{T}\right) + \frac{1}{\theta^2} \frac{J_T}{T}\right)\right]$$

$$= \sigma_\varepsilon^2 \left[\left(\mathrm{I_N} \otimes \left(\mathrm{I_T} - \frac{J_T}{T}\right)\right) + \frac{1}{\theta^2}\left(\mathrm{I_N} \otimes \frac{J_T}{T}\right)\right]$$

D'où,

$$\Omega_{(NT,NT)} = \sigma_\varepsilon^2 \left(\mathrm{W} + \frac{1}{\theta^2} \mathrm{B}\right) = \sigma_\varepsilon^2 \mho_{(NT,NT)} \qquad (13.3)$$

$\Rightarrow$ Cette expression de la matrice $\Omega_{(NT,NT)}$ est formée par des projecteurs orthogonaux, W et B, connus et par de simples coefficients inconnus, σ_ε^2 et θ^2, à estimer :
ce qui signifie qu'on a une simplification de l'existence dans le Panel.

Etant donné les propriétés des projecteurs orthogonaux, W et B, l'écriture (13.3) nous permet d'obtenir directement, l'expression de l'inverse de cette matrice $\Omega_{(NT,NT)}$:

$$\Omega^{-1}_{(NT,NT)} = \frac{1}{\sigma_\varepsilon^2}(\mathrm{W} + \theta^2\,\mathrm{B}) = \frac{1}{\sigma_\varepsilon^2}\mho^{-1}_{(NT,NT)} \qquad (13.4)$$

Remarquons que ceci est vérifié aisément en calculant le produit $\Omega\Omega^{-1}$qui n'est autre que l'identité, $\mathrm{I_{NT}}$.

Remarque:

Aussi, l'expression (13.3) implique que la matrice $\Omega_{(NT,NT)}$ apparaît comme combinaison linéaire des projecteurs orthogonaux, W et B, i.e :

$$(13.3) \Longrightarrow \ \Omega_{(NT,NT)} = \ \sigma_\varepsilon^2 \, W + (\ \sigma_\varepsilon^2 + T\ \sigma_u^2)\ B$$
$$\Leftrightarrow$$
$$\Omega_{(NT,NT)} = [W \quad B]\begin{bmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_\varepsilon^2 + T\ \sigma_u^2 \end{bmatrix}\begin{bmatrix} W \\ B \end{bmatrix} = C\ \mathbb{V}\ C'$$
$$Ce\ qui\ définit$$
$$la\ décomposition\ spectrale\ de\ la\ matrice\ var - cov\ \text{r}ésiduelle, \Omega_{(NT,NT)}, \text{où}\ les\ opérateurs$$
$$W\ et\ B\ sont\ ces\ vecteurs\ propres, associés\ aux\ valeurs\ propres\ respectives\ \sigma_\varepsilon^2\ et\ (\ \sigma_\varepsilon^2 + T\ \sigma_u^2).$$
$$avec$$
$$\mathbb{V}\ \ définit\ la\ matrice\ des\ valeurs\ propres$$
$$et$$
$$\begin{cases} \Omega\, W = \ \sigma_\varepsilon^2\ \mathrm{W} \\ \Omega\ \mathrm{B} = (\ \sigma_\varepsilon^2 + T\ \sigma_u^2) B \end{cases}$$

<u>Du fait de l'autocorrélation intra-individuelle</u>, le meilleur estimateur sans biais est celui des MCGs. Via l'expression (13.4), on peut développer son écriture en fonction des variances et covariances Between et Within des variables en jeu :

$$\hat{\beta}_{MCG} = \ (X'\Omega^{-1}\ X)^{-1} X'\Omega^{-1}\ Y = (X'\mho^{-1}\ X)^{-1} X'\mho^{-1}\ Y$$

$$= (X'W\ X + \ \theta^2\ X'B\ X\ )^{-1}\ (X'W\ Y + \ \theta^2\ X'B\ Y)$$
$$\text{avec } \theta^2 = \frac{\sigma_\varepsilon^2}{\mathrm{T}\ \sigma_u^2 + \sigma_\varepsilon^2}$$
$$La\ variance\ de\ cet\ estimateur\ MCG\ \ est\ minimale:$$
$$V(\hat{\beta}_{MCG}) = (X'\Omega^{-1}\ X)^{-1} = \ \sigma_\varepsilon^2\ (X'\mho^{-1}\ X)^{-1} = \ \sigma_\varepsilon^2 [X'\ (\mathrm{W} + \theta^2 \mathrm{B})X]^{-1}$$
$$= \ \sigma_\varepsilon^2\ (X'W\ X + \ \theta^2\ X'B\ X)^{-1}$$

Ainsi, la méthode des MCGs utilise à la fois les dimensions Between et Within (i.e. dimensions inter-individuelle et intra-individuelle) de l'information disponible, prises dans une proportion particulière.

<u>La mise en œuvre de cet estimateur des MCGs est plutôt simple</u> : En effet, comme tout estimateur des MCGs, l'estimateur $\hat{\beta}_{MCG}$ est équivalent à l'estimateur des MCOs appliqués aux données transformés $y_{it} - (1-\theta)\mathrm{B}\ \mathrm{y} = y_{it} - (1-\theta)\mathrm{y}_{\mathrm{i.}}$. En effet, on a

$$\sigma_\varepsilon^2\ \Omega^{-1} = \mathrm{W} + \theta^2 \mathrm{B} = \mathrm{I_{NT}} - \mathrm{B} + \theta^2 \mathrm{B} = \mathrm{I_{NT}} - (1-\theta^2)\mathrm{B}$$
$$= [\mathrm{I_{NT}} - (1-\theta)\mathrm{B}][\mathrm{I_{NT}} - (1-\theta)\mathrm{B}]$$

Egalement, cette mise en œuvre est équivalente à transformer le modèle (13) en le multipliant des deux côtés par la matrice $P = \ \Omega^{-\frac{1}{2}}$ et d'appliquer les MCOs aux données sphériques:

$$Soit\ la\ matrice\ de\ transformation\ P = \Omega^{-\frac{1}{2}} = \mathrm{I}_{\mathrm{NT}} - (1-\theta)\,\mathrm{B}_{(NT,NT)}\ \text{ où }\ \theta = \sqrt{\frac{\sigma_\varepsilon^2}{\mathrm{T}\,\sigma_u^2 + \sigma_\varepsilon^2}}$$

$$(13)\quad y_{it}=\mu+\sum_{k=1}^{K} x_{k\,it}\,\beta_k+w_{it}\ \text{où}\ \ w_{it}=u_i+\varepsilon_{it} \Longrightarrow Y_{(NT,1)} = X_{(NT,K+1)}\beta_{(K+1,1)} + w_{(NT,1)}$$

$$\xrightarrow{Par\ la\ transformation\ P,on\ a} \begin{cases} P\,Y = PX\,\beta + Pw \Longrightarrow le\ modèle\ sphérisé: \tilde{Y} = \tilde{X}\beta + \widetilde{w} \\ \text{où} \\ \tilde{Y} = PY = [y_{it}]_{(NT,1)} - (1-\theta)[\overline{y_{i.}}]_{(NT,1)}\ et\ \beta_{(K+1,1)} = \begin{pmatrix} \mu_{(1,1)} \\ \beta_{(K,1)} \end{pmatrix} \end{cases}$$

$$alors$$

$$\begin{cases} y_{it} - (1-\theta)\overline{y_{i.}} = \theta\mu + \sum_{k=1}^{K} \beta_k[x_{k\,it} - (1-\theta)\overline{x_{ki.}}] + \underbrace{[w_{it} - (1-\theta)\overline{w_{i.}}]}_{=\,\eta_{it}} \qquad (14) \\ \forall\, i = 1,\cdots,N \quad et \quad t = 1,\cdots,T \\ \text{où}\ \ \eta_{it} vérifie\ les\ hypothèses\ classiques\ des\ MCOs. \end{cases}$$

$$avec$$

$$E(\widetilde{w}) = E[Pw] = PE[w] = 0$$

$$E(\widetilde{w}\,\widetilde{w}') = E[Pww'P'] = PE[ww']P' = P\Omega\mathrm{P}' = \sigma_\varepsilon^2 \mathrm{P}\mho\mathrm{P}' = \sigma_\varepsilon^2 \mathrm{I}_{\mathrm{NT}}$$

$$et$$

$$\hat{\beta}_{MCG} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\,Y = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{Y}$$

L’estimateur des MCGs[2] est convergent pour T fini lorsque N tend vers l’infini. Néanmoins, il n’est pas opératoire, i.e. il ne peut pas être mis en œuvre, puisque certains paramètres de la matrice de variances-covariances résiduelle demeurent a priori inconnus, à savoir les variances, σ_u^2 et σ_ε^2 qui interviennent dans la grandeur $\theta^2 = \frac{\sigma_\varepsilon^2}{\mathrm{T}\,\sigma_u^2 + \sigma_\varepsilon^2}$. Autrement dit, θ^2 est généralement inconnue et l’estimateur des MCGs ne peut être ainsi calculé. Dans la pratique, l’estimateur des Moindres Carrés Quasi Généralisé, MCQGs, est utilisé. La méthode d’estimation est réalisée en deux temps. Dans un 1er temps, la perturbation stochastique $\widehat{w_{it}}$ est estimée afin d’obtenir les valeurs du facteur de transformation des données $(1-\theta)$, i.e. on estime dans un 1er temps les variances σ_u^2 et σ_ε^2 pour en déduire une estimation de θ. La substitution de cette estimation $\hat{\theta}$ à la vraie valeur (inconnue) de θ permet d’estimer le modèle (transformé) et de calculer ainsi ce qu’il est convenu d’appeler l’estimateur des MCQGs.

La mise en œuvre pratique de cette méthode d’estimation, MCQG, nécessite donc de disposer d’estimations des variances résiduelles inconnues, σ_u^2 et σ_ε^2 , i.e. d’examiner le problème d’estimation de θ^2. On procède tout d’abord à l’interprétation de ce paramètre θ^2 :

[2] On rappelle que les propriétés de cet estimateur sont les propriétés usuelles des estimateurs des MCGs, i.e. sans biais et efficace.

⇢ Interprétation de $\theta^2 = \frac{\sigma_\varepsilon^2}{T\,\sigma_u^2 + \sigma_\varepsilon^2}$: peut être interprété à partir des composantes Between et Within de la variance de la perturbation (ou de y). A partir de la définition de w_{it} et de ses moments du second ordre, on a en effet,

$$w_{it} = u_i + \varepsilon_{it}$$

$$La\ variance\ Within\ de\ w_{it}\ est\ égale\ à$$

$$V(Ww_{it}) = V(w_{it} - \ \ w_{i.}) = V(\varepsilon_{it} - \varepsilon_{i.}) = E(\varepsilon_{it} - \varepsilon_{i.})^2 = \ \sigma_\varepsilon^2 + \frac{1}{T^2}\mathrm{E}[(\textstyle\sum_t \varepsilon_{it})^2] - \frac{2}{T}\ \mathrm{E}[\varepsilon_{it}\textstyle\sum_t \varepsilon_{it}]$$

$$= \ \sigma_\varepsilon^2\left(1 + \frac{1}{T} - \frac{2}{T}\right) = \ \sigma_\varepsilon^2\left(1 - \frac{1}{T}\right)(= \sigma_{w_W}^2)$$

$$Et, la\ variance\ Between\ de\ w_{it}\ est\ définie\ par$$

$$V(Bw_{it}) = V(\ w_{i.}) = V(u_i) + V(\ \varepsilon_{i.}) = \ \sigma_u^2 + \frac{\sigma_\varepsilon^2}{T}(= \sigma_{w_B}^2)$$

D′où,

$$\theta^2 = \frac{1}{T-1}\,\frac{V(w_{it} - \ w_{i.})}{V(\ w_{i.})}\ \text{dépend, à un facteur}\ \frac{1}{T-1}\ \text{près,}$$

du rapport de la variance Within du résidu à la variance Between du résidu.

⇢ Estimation convergente de θ^2 : (il existe différentes procédures d'estimer ce θ^2, i.e. estimer les variances σ_u^2et σ_ε^2, et on cite deux usuelles)

- A partir de n'importe quel estimateur sans biais (ou convergent) $\hat{\beta}$ de β, on peut obtenir une estimation de w : $\hat{w} = Y - X\hat{\beta}$. Du fait que $w_{it} = u_i + \varepsilon_{it}$, il suffit alors d'effectuer une analyse de la variance sur $\hat{w}$ pour en déduire des estimations de σ_u^2et σ_ε^2. Comme $\hat{\beta}_W$ est asymptotiquement, quand $N\ et\ T \to \infty$, équivalent aux MCGs, on peut conseiller son utilisation, (plutôt que celle de $\hat{\beta}_{MCO}\,ou\,\hat{\beta}_B$ qui conduirait à des estimations moins efficaces asymptotiquement.

Cette démarche aboutit à des estimations convergentes. Mais, dans la pratique, quand T est fini , on peut se heurter à des difficultés, comme par exemple, l'obtention d'un $\widehat{\sigma_u^2}$ négatif ou d'un $\widehat{\theta^2}$ supérieure à l'unité.

On note que :

$$\begin{cases} * \text{ lorsque } \widehat{\theta^2} = 1, \text{i. e. } \widehat{\sigma_u^2} = 0 \xRightarrow{\text{on retrouve}} \text{le modèle MCO simple sur l'ensemble de l'échantillon, i. e.} \\ (13)\ y_{it} = \mu + \sum_{k=1}^{K} X_{k\,it}\,\beta_k + w_{it} \qquad \forall\, i = 1, \cdots, N \quad et \quad t = 1, \cdots, T \\ \text{où les effets spécifiques sont identiques pour les individus et sont compris dans la dérive:} \\ \text{cela revient à ne pas avoir d'effetsspécifiques} \\ * \text{ lorsque } \widehat{\theta^2} = 0 \text{ , i. e. si } \widehat{\sigma_u^2} \text{est grande par rapport à celle des erreurs idiosyncratiques } \widehat{\sigma_\varepsilon^2} \\ \text{où l'essentiel des écarts entre les individus est dû à des effets spécifiques constants dans le temps} \\ \text{qui peuvent donc être modélisés comme des variables muettes} \\ \text{ou bien si T} \to \infty \text{ alors l'information sera essentiellement d'origine intra} - \text{individuelle} \\ \xRightarrow{\text{on retrouve alors}} \text{ le modèle intra, i. e. le modèle (14) implique} \\ y_{it} - \overline{y_{i.}} = \sum_{k=1}^{K} \beta_k[x_{k\,it} - \overline{x_{ki.}}] + \underbrace{[w_{it} - \overline{w_{i.}}]}_{=\eta_{it}} \quad \forall\, i = 1, \cdots, N \quad et \quad t = 1, \cdots, T \end{cases}$$

Par conséquent, le modèle MEC peut donc être interprété comme une moyenne pondérée du modèle MCO et du modèle à effets fixes. Il combine la variabilité intra − individuelle et la variabilité inter − individuelle de manière à minimiser la variance des erreurs.

- Pratiquement, il existe une solution plus simple et la plus souvent retenue, celle proposée par Swamy-Arora (1972) qui consiste à utiliser les variances estimées des perturbations issues de l'estimation du modèle MEC par les méthodes de régression Within d'une part et Between d'autre part (i.e. régressions intra-individuelle et inter-individuelle) et cette procédure Swamy-Arora permet aussi d'obtenir des estimations convergentes et même sans biais de σ_u^2et σ_ε^2.

A partir de la régression Between, on a :

$$BY = BX\beta + Bw$$
$$On\ obtient\ le\ résidu\ estimé:$$
$$\widehat{w}_{B\,(NT,1)} = BY - BX\,\hat{\beta}_B$$

$$Soit\ \hat{\sigma}_{w_B}^2\ défini\ par\ \hat{\sigma}_{w_B}^2 = \frac{\widehat{w}_B'\,\widehat{w}_B}{N-(K+1)}$$
$$avec \begin{cases} N\ moyennes\ individuelles \\ K\ nbre\ de\ paramètres\ à\ estimer\ dans\ le\ modèle\ transformé. \end{cases}$$
$$vérifie\ que$$
$$E\left(\hat{\sigma}_{w_B}^2\right) = \mathrm{T}\,\sigma_u^2 + \sigma_\varepsilon^2 \qquad (on\ peut\ dire\ que\ \hat{\sigma}_{w_B}^2 = \mathrm{T}\,\sigma_u^2 + \sigma_\varepsilon^2)$$

En effet, on a (comme preuve):

$$\widehat{w}_B = M_{BX}\, Bw \quad avec \quad M_{BX} = I_{NT} - BX\,(X'B\,X)^{-1}X'B = \; I_{NT} - P_{BX}$$
$$et$$
$$E[\widehat{w}_B'\widehat{w}_B] = E\big[tr\big(w' \; BM_{BX}Bw \quad\big)\big] = tr[BM_{BX}B\; E(w \quad w' \;)] = tr[BM_{BX}B\;\Omega]$$
$$= tr\left[BM_{BX}B\;\left(\;\sigma_\varepsilon^2\left(\mathrm{W} + \frac{1}{\theta^2}\;\mathrm{B}\right)\right)\right]$$
$$= \frac{\sigma_\varepsilon^2}{\theta^2}\,tr[BM_{BX}B]$$
$$or\; BM_{BX}B = M_{BX}B \;\; et\; tr[M_{BX}B] = tr[B - BX\,(X'B\,X)^{-1}X'B] = tr[B] - tr[BX\,(X'B\,X)^{-1}X'B]$$
$$or, on\; a\; tr[B] = tr\left[I_N \otimes \frac{J_T}{T}\right] = tr(I_N).\,tr\left(\frac{J_T}{T}\right) = N.\frac{T}{T} = N$$
$$de\; \text{même}, tr[BX\,(X'B\,X)^{-1}X'B] = tr[\,(X'B\,X)^{-1}X'BX] = tr\;(I_{K+1}) = K + 1$$
$$Alors, E[\widehat{w}_B'\widehat{w}_B] = \frac{\sigma_\varepsilon^2}{\theta^2}\big(N - (K+1)\big) = (\mathrm{T}\,\sigma_u^2 + \;\sigma_\varepsilon^2)\big(N - (K+1)\big)$$
$$et, on\; retrouve\; ainsi\; l'expression\; E\big(\hat{\sigma}_{w_B}^2\big) = \;\mathrm{T}\,\sigma_u^2 + \;\sigma_\varepsilon^2.$$

Pareillement, <u>à partir de la régression Within</u>, on a :

$$WY = WX\beta + Ww$$
$$le\; \text{résidu}\; estim\text{é}\; obtenu:$$
$$\widehat{w}_{W\,(NT,1)} = WY - WX\,\hat{\beta}_W$$

$$tel\; que\; \hat{\sigma}_{w_W}^2\; \text{dé}fini\; par$$
$$\hat{\sigma}_{w_W}^2 = \frac{\widehat{w}_W'\,\widehat{w}_W}{(NT - N - K)}$$
$$v\text{é}rifie\; que$$
$$E\big(\hat{\sigma}_{w_W}^2\big) = \;\sigma_\varepsilon^2 \qquad (on\; peut\; dire\; que\; \hat{\sigma}_{w_W}^2 = \;\sigma_\varepsilon^2)$$
$$avec$$
$$NT\; nombre\; d'observations\; dans\; tout\; l'\text{é}chantillon,$$
$$N\; moyennes\; individuelles,$$
$$K\; nbre\; de\; param\text{è}tres\; \text{à}\; estimer\; dans\; le\; mod\text{è}le\; transform\text{é}.$$

Et de la même manière, on a (<u>comme preuve</u>):

$$\widehat{w}_W = M_{WX}\, Ww$$
$$E[\widehat{w}'_W \widehat{w}_W] = E\left[tr\left(w'\ WM_{WX}Ww\ \ \right)\right] = tr[WM_{WX}W\ E(w\ \ w'\ )] = tr[WM_{WX}W\ \Omega]$$
$$= \sigma_\varepsilon^2\ tr[M_{WX}W]\ \ où\quad WM_{WX}W = M_{WX}W$$
$$= \sigma_\varepsilon^2\ tr[W - WX\ (X'W\ X)^{-1}X'W]$$
$$or, on\ a\ tr[W] = tr\left[I_N \otimes \left(I_T - \frac{J_T}{T}\right)\right] = N\ tr\left(I_T - \frac{J_T}{T}\right) = N\left[tr(I_T) - tr\left(\frac{J_T}{T}\right)\right] = N(T-1)$$
$$de\ même, tr[WX\ (X'W\ X)^{-1}X'W] = tr[\ (X'W\ X)^{-1}X'WX] = tr\ (I_K) = K$$
$$Alors, E[\widehat{w}'_W \widehat{w}_W] = \sigma_\varepsilon^2(N(T-1) - K)$$
$$et, on\ retrouve\ ainsi\ l'expression\ E\left(\widehat{\sigma}^2_{w_W}\right) = \sigma_\varepsilon^2$$

Ainsi, on obtient à partir des perturbations des estimations Between et Within, des estimations sans biais des variances σ_u^2 et σ_ε^2.

⇢ L’estimateur des Moindres Carrés Quasi Généralisés, MCQGs :

Son expression correspond à celle des MCGs avec pour matrice var-covariance résiduelle, $\Omega_{(NT,NT)}$, une estimation obtenue dans une première étape :

$$\hat{\beta}_{MCQG} = \left(X'\widehat{\Omega}^{-1}\ X\right)^{-1} X'\widehat{\Omega}^{-1}\ Y \qquad (15.1)$$

$où$

$$\widehat{\Omega}_{(NT,NT)} = \hat{\sigma}_\varepsilon^2\left(\mathrm{W} + \frac{1}{\hat{\theta}^2}\ \mathrm{B}\right) = \hat{\sigma}_\varepsilon^2 \widehat{\mho}_{(NT,NT)} \qquad (15.2)$$

avec, pour $\hat{\theta}^2$, l'une des estimations présentées ci-dessus.

On pourra prendre ainsi $\hat{\theta}^2 = \frac{\hat{\sigma}^2_{w_W}}{\hat{\sigma}^2_{w_B}}$

5. Inférence : Tests d’existence de spécificités individuelles. Appelé aussi Tests d’homogénéité des comportements ou Test de Redondance (Redundant Test) _ test de poolabilité.

Dans notre cours, on a supposé l’existence d’une hétérogénéité inter-individuelle des comportements, hétérogénéité prise en compte par l’adjonction d’effets spécifiques individuels, fixes aux « véritables » variables explicatives. Il paraît naturel de tester l’existence d’une telle hétérogénéité. Ceci revient à discerner/distinguer entre le modèle à effet uniforme (à $NT - (K+1)$ degrés de liberté) :

$$(10) \qquad y_{it} = \mu + \sum_{k=1}^{K} X_{k\,it}\,\beta_k + \varepsilon_{it} \qquad \forall\ i = 1, \cdots, N\ et\ \forall\ t = 1, \cdots, T$$

Et, le modèle avec effets spécifiques individuels fixes (à $NT - (N+K)$ degrés de liberté) :,

$$(11) \qquad y_{it} = \mu_i + \sum_{k=1}^{K} X_{k\,it}\,\beta_k + \varepsilon_{it} \qquad \forall\, i = 1, \cdots, N \; et\; \forall\, t = 1, \cdots, T$$

Sous l'hypothèse de normalité des résidus, les tests usuels, le Student-test et le Fisher-test, sont exécutés dans ce contexte (d'existence ou non d'une telle hétérogénéité).

<u>En particulier</u>, si on teste le corps d'hypothèses suivant :

$$\begin{cases} H_0 : \mu_i = \mu_j, \quad \text{pour certaines unités spatiales i et j où } i \neq j \\ H_a : \mu_i \neq \mu_j \end{cases}$$

$$Test \text{ à } une\ seule\ restriction \Rightarrow C'est\ le\ Student - test\ défini\ par\ la\ statistique:$$

$$\mathrm{t}_{\widehat{\widehat{\mu_i}-\widehat{\mu_j}}} = \frac{\widehat{\mu_i} - \widehat{\mu_j}}{\sigma_{\widehat{\widehat{\mu_i}-\widehat{\mu_j}}}} \xrightarrow{s.\ H_0} \mathcal{T}\big(NT - (N+K)\big)$$

$$Décision, rejet\ de\ H_0\ si\ \mathrm{t}_{\widehat{\widehat{\mu_i}-\widehat{\mu_j}}} > \mathrm{t}^{\alpha=5\%}_{(NT-(N+K))}.$$

<u>Le cas général</u> est défini par la question pertinente <u>s'il n'existe pas d'effets spécifiques individuels/spatiaux</u> $\Longrightarrow$ on teste alors, l'hypothèse jointe de base, qui est une hypothèse contraignante celle d'une homogénéité complète des comportements où :

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \cdots = \mu_N = \mu \quad \text{(et le modèle correspondant est le modèle contraint } (10)) \\ H_a : \text{sinon ou bien } \overline{H_0} \end{cases}$$

$$Test\ d'hypothèse\ jointe, H_0, \text{à } plusieurs\ restrictions$$

$$qui\ signifie\ qu'on\ teste\ N-1\ égalités H_0 : \mu_1 = \mu_2;\ \mu_2 = \mu_3, etc$$

$$qui\ se\ fait\ très\ classiquement\ par\ un\ Fisher - test\ défini\ par\ la\ statistique$$

$$\hat{\mathcal{F}} = \frac{(SCR_c - SCR)\big/_{N-1}}{SCR\big/_{NT-(N+K)}} \xrightarrow{s.\ H_0} \mathbb{F}\big(N-1, NT-(N+K)\big)$$

$$avec$$

$$SCR_c\ est\ la\ somme\ des\ carrés\ des\ résidus\ associés \text{ à } l'estimation\ du\ modèle\ (10) ne\ comporta$$

$$aucune\ variable\ indicatrice\ spécifique\ aux\ individus$$

$$et$$

$$SCR\ est\ celle\ associée\ au\ modèle\ non\ contraint, (11), comportant\ ces\ variables.$$

$$Décision\ du\ test:$$

$$Si\ \hat{\mathcal{F}} \geq au\ fractile\ \mathbb{F}^{1-\alpha}_{(N-1,NT-(N+K))}, rejet\ de\ H_0; i.e. le\ modèle\ doit\ inclure\ des\ effets\ spécifiques\ fi$$

$$Si\ \hat{\mathcal{F}} < \mathbb{F}^{1-\alpha}_{(N-1,NT-(N+K))}, accepter\ H_0; i.e. les\ effets\ fixes\ sont\ superflus\ (parasites, surplus).$$

Rappelons que le nombre de degré de liberté employé au numérateur est $(N-1)$ et non N, car tester que N coefficients, μ_i, sont tous égaux revient à tester que $(N-1)$ différences, $\mu_{i+1} - \mu_i$, i $= 1, \cdots,$ N, sont toutes nulles.

Cependant, l'interprétation/ l'explication qu'on donne au rejet de l'hypothèse de base, H_0, doit être réfléchie (prudente). En effet, en pratique, ce rejet doit être compris avant tout comme un signe d'existence d'une hétérogénéité des comportements. La question de savoir si l'inclusion d'effets fixes, spécifiques individuels dans le modèle suffit à en rendre compte totalement, doit ensuite être étudiée soigneusement. Pour ce faire, on peut, par exemple, regrouper les individus/unités spatiales selon la valeur prise par leur effet spécifique (par exemple, ceux pour lesquels cet effet est négatif, resp. positif) et estimer, par la suite, un modèle pour chacun des groupes ainsi constitués (en signe de référence, par exemple, voir Hultberg-Nadiri-Sickles 1999) .

6. Un test d'absence de corrélation entre les effets spécifiques et les variables explicatives : le Hausman-Test (1978).

Le Hausman-Test (1978) répond à la critique de Mundlak en réalisant deux estimations et en comparant les performances de leurs coefficients de pente dans deux hypothèses différentes :

$$\begin{cases} H_0: \mathrm{Cov}(x_{k\,it}, u_i) = 0 \text{ , i. e. absence de corrélation entre régresseurs et effets spécifiques} \\ \text{, i. e. Exogénéité des régresseurs par rapport à l'erreur spécifique du modèle} \\ \overset{\text{Du résultat de ce test dépend la mise en oeuvre de la méthode d'estimation qui sera finalement retenue}}{\Longleftrightarrow} \\ \hat{\beta}_{\mathrm{Within}} = \hat{\beta}_{\mathrm{MCQG}} \Leftrightarrow \hat{\beta}_{\mathrm{MEF}} = \hat{\beta}_{\mathrm{MEC}} \text{, i. e. absence de différence significative entre les estimations} \\ \text{Within et MCQG;} \\ \text{i. e. les variables sont exogènes par rapport à l'erreur spécifique} \Leftrightarrow \text{on privilégie l'usage de} \\ \text{l'estimateur } \widehat{\beta}_{\mathrm{MCQG}} \text{ qui est sans biais et plus efficace que l'estimateur } \widehat{\beta}_{\mathrm{Within}} \\ \Leftrightarrow \\ \text{C'est le modèle à effets aléatoires qui l'emporte} \\ \\ H_a: \mathrm{Cov}(x_{k\,it}, u_i) \neq 0 \text{, i. e. l'un au moins des régresseurs est corrélé avec l'erreur spécifique} \\ \Leftrightarrow \\ \hat{\beta}_{\mathrm{Within}} \neq \hat{\beta}_{\mathrm{MCQG}} \Leftrightarrow \hat{\beta}_{\mathrm{MEF}} \neq \hat{\beta}_{\mathrm{MEC}} \text{ i. e. présence de différence significative entre les estimations} \\ \text{Within et MCQG;} \\ \Leftrightarrow \text{on privilégie l'usage de l'estimateur } \widehat{\beta}_{\mathrm{Within}} \text{ qui guarantit des estimations} \\ \text{sans biais et convergentes des coefficients } \beta_k \\ \Leftrightarrow \\ \text{C'est le modèle à effets fixes qui l'emporte} \end{cases}$$

Formellement, la statistique du Hausman − test (dérivée de celle du Wald − test) est donnée par:

$$H = \left(\hat{\beta}_{\mathrm{Within}} - \hat{\beta}_{\mathrm{MCQG}}\right)' \left[\hat{V}\left(\hat{\beta}_{W}\right) - \hat{V}\left(\hat{\beta}_{\mathrm{MCQG}}\right)\right]^{-1} \left(\hat{\beta}_{\mathrm{Within}} - \hat{\beta}_{\mathrm{MCQG}}\right) \xrightarrow{s.H_0} \mathcal{X}^2(K)$$

où K: degré de liberté, définit par le nbre de régresseurs sans la constante.

Et, la décision est le rejet de H_0 si la Prob. critique < à seuil α% ;
et c'est le modèle MEF qui est privilégié pour le panel étudié.

II. Application : Etude de l'impact du tourisme sur la croissance économique

On mène une étude de l'impact du tourisme sur la croissance économique sur la base de données en moyennes quinquennales[3] observées sur la période 1968-1997, composées de 63 pays. Chaque pays dispose théoriquement de six points d'observations.

Les principales variables d'intérêts sont le taux de croissance du PIB par tête- $GROWTH$-, le nombre de touristes (en log)- $TOURISM$-, le niveau du PIB par tête initial (1988)- LYO-, le taux de scolarisation primaire- $PRIM$-, le taux d'inflation- $INFL$- et l'indicateur de politique d'ouverture- SW-. Sur données de panel, on régresse le modèle de croissance économique:

[3] Quinquennale, adj. qui s'étend sur cinq ans.

```
. xtdescribe

      id:  1, 2, ..., 63                                    n =         63
    tri5:  1, 2, ..., 6                                     T =          6
           Delta(tri5) = 1 unit
           Span(tri5)  = 6 periods
           (id*tri5 uniquely identifies each observation)

Distribution of T_i:   min      5%     25%       50%       75%     95%     max
                         6       6       6         6         6       6       6
```

Freq.	Percent	Cum.	Pattern
63	100.00	100.00	111111
63	100.00		XXXXXX

tri5	Freq.	Percent	Cum.
1	63	16.67	16.67
2	63	16.67	33.33
3	63	16.67	50.00
4	63	16.67	66.67
5	63	16.67	83.33
6	63	16.67	100.00
Total	378	100.00	

Variable	Obs	Mean	Std. Dev.	Min	Max
growth	357	.0098586	.0264009	-.0828881	.1092166
tourism	120	12.80934	1.760858	7.600903	16.8789
lyo	357	6.712912	1.108785	4.416319	9.592634
prim	368	85.00207	29.03906	8.001669	147.3614
infl	324	.4384221	1.658044	-.040074	18.31301
sw	366	.270674	.4258422	0	1

Variable		Mean	Std. Dev.	Min	Max	Observations
growth	overall	.0098586	.0264009	-.0828881	.1092166	N = 357
	between		.0155993	-.02847	.055284	n = 62
	within		.0213847	-.064387	.0999846	T-bar = 5.75806
tourism	overall	12.80934	1.760858	7.600903	16.8789	N = 120
	between		1.784817	8.29405	16.8789	n = 61
	within		.2251897	12.11619	13.50248	T-bar = 1.96721
lyo	overall	6.712912	1.108785	4.416319	9.592634	N = 357
	between		1.106729	4.676627	9.325356	n = 62
	within		.2100614	5.878072	7.630522	T-bar = 5.75806
prim	overall	85.00207	29.03906	8.001669	147.3614	N = 368
	between		26.70497	17.55525	132.2406	n = 62
	within		11.87192	33.90285	138.1902	T-bar = 5.93548
infl	overall	.4384221	1.658044	-.040074	18.31301	N = 324
	between		.9880508	.0376138	5.089035	n = 61
	within		1.38742	-3.898802	14.32677	T-bar = 5.31148
sw	overall	.270674	.4258422	0	1	N = 366
	between		.313667	0	1	n = 61
	within		.2903487	-.4626594	1.104007	T = 6

```
. regress growth tourism lyo prim infl sw, ro

Linear regression                                      Number of obs =     114
                                                       F(  5,   108) =   13.97
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.4468
                                                       Root MSE      =  .02014
```

growth	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
tourism	.0078449	.0014233	5.51	0.000	.0050236	.0106662
lyo	-.0037806	.0020623	-1.83	0.070	-.0078684	.0003072
prim	.0000935	.0001095	0.85	0.395	-.0001237	.0003106
infl	-.0028851	.0008289	-3.48	0.001	-.0045281	-.001242
sw	.0090248	.0046823	1.93	0.057	-.0002563	.0183059
_cons	-.0790599	.0158867	-4.98	0.000	-.1105502	-.0475697

```
Fixed-effects (within) regression               Number of obs      =       114
Group variable: id                              Number of groups   =        58

R-sq:  within  = 0.6368                         Obs per group: min =         1
       between = 0.0018                                        avg =       2.0
       overall = 0.0036                                        max =         2

                                                F(5,51)            =     17.89
corr(u_i, Xb)  = -0.9431                        Prob > F           =    0.0000

------------------------------------------------------------------------------
      growth |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tourism |   .0270759   .0057555     4.70   0.000     .0155213    .0386304
         lyo |  -.0873547   .0135308    -6.46   0.000     -.114519   -.0601904
        infl |  -.0007259   .0008021    -0.90   0.370    -.0023362    .0008845
        prim |   .0006828   .0002603     2.62   0.011     .0001604    .0012053
          sw |   .0107407   .0072223     1.49   0.143    -.0037586    .0252399
       _cons |   .1805315   .1039866     1.74   0.089    -.0282302    .3892933
-------------+----------------------------------------------------------------
     sigma_u |  .07662421
     sigma_e |  .01232982
         rho |  .97476057   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(57, 51) =     4.16              Prob > F = 0.0000
```

```
i.id              _Iid_1-63           (naturally coded; _Iid_1 omitted)
note: _Iid_20 omitted because of collinearity
note: _Iid_21 omitted because of collinearity
note: _Iid_34 omitted because of collinearity
note: _Iid_38 omitted because of collinearity
note: _Iid_53 omitted because of collinearity
```

Source	SS	df	MS
Model	.071390106	62	.001151453
Residual	.00775325	51	.000152025
Total	.079143355	113	.000700384

Number of obs =	114
F(62, 51) =	7.57
Prob > F =	0.0000
R-squared =	0.9020
Adj R-squared =	0.7829
Root MSE =	.01233

growth	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
tourism	.0270759	.0057555	4.70	0.000	.0155213	.0386304
lyo	-.0873547	.0135308	-6.46	0.000	-.114519	-.0601904
prim	.0006828	.0002603	2.62	0.011	.0001604	.0012053
infl	-.0007259	.0008021	-0.90	0.370	-.0023362	.0008845
sw	.0107407	.0072223	1.49	0.143	-.0037586	.0252399
_Iid_2	-.1752464	.0459019	-3.82	0.000	-.2673983	-.0830946
_Iid_3	-.1325757	.0310644	-4.27	0.000	-.1949401	-.0702114
_Iid_4	-.0539857	.0144011	-3.75	0.000	-.0828972	-.0250742
_Iid_5	-.1771766	.0522164	-3.39	0.001	-.2820054	-.0723479
_Iid_6	-.2255026	.0512207	-4.40	0.000	-.3283324	-.1226728
_Iid_7	-.1400042	.034563	-4.05	0.000	-.2093923	-.070616
_Iid_8	-.1168481	.0477145	-2.45	0.018	-.212639	-.0210572
_Iid_9	-.1365772	.0519349	-2.63	0.011	-.2408409	-.0323135
_Iid_10	-.0171402	.0157143	-1.09	0.281	-.0486879	.0144076
_Iid_38	0	(omitted)				
_Iid_39	-.0214307	.0188014	-1.14	0.260	-.059176	.0163146
_Iid_40	-.1424438	.0213035	-6.69	0.000	-.1852124	-.0996752
_Iid_41	-.1598454	.0284195	-5.62	0.000	-.2168999	-.1027909
_Iid_42	-.2631256	.0489437	-5.38	0.000	-.3613842	-.1648669
_Iid_43	-.1914818	.0388336	-4.93	0.000	-.2694435	-.1135201
_Iid_44	-.1523035	.0533613	-2.85	0.006	-.2594309	-.0451762
_Iid_45	-.2341146	.0466565	-5.02	0.000	-.3277815	-.1404478
_Iid_46	-.1628169	.040578	-4.01	0.000	-.2442806	-.0813533
_Iid_47	-.0593807	.0344714	-1.72	0.091	-.1285848	.0098235
_Iid_48	-.091657	.0225798	-4.06	0.000	-.1369878	-.0463262
_Iid_49	-.0829616	.0210966	-3.93	0.000	-.1253148	-.0406084
_Iid_50	-.1707535	.0279428	-6.11	0.000	-.226851	-.1146561
_Iid_51	-.1356847	.0513964	-2.64	0.011	-.2388673	-.0325022
_Iid_52	-.1437048	.0387004	-3.71	0.001	-.2213991	-.0660104
_Iid_53	0	(omitted)				
_Iid_54	-.0950598	.0153366	-6.20	0.000	-.1258492	-.0642703
_Iid_55	-.1534824	.0340743	-4.50	0.000	-.2218894	-.0850755
_Iid_56	-.0997078	.0230139	-4.33	0.000	-.1459103	-.0535054
_Iid_57	-.1735903	.0439687	-3.95	0.000	-.2618611	-.0853195
_Iid_58	-.1148465	.0200697	-5.72	0.000	-.155138	-.0745549
_Iid_59	-.1878816	.0463149	-4.06	0.000	-.2808627	-.0949005
_Iid_60	-.0185726	.0155753	-1.19	0.239	-.0498414	.0126961
_Iid_61	-.028124	.0167502	-1.68	0.099	-.0617515	.0055035
_Iid_62	-.2025895	.0379115	-5.34	0.000	-.2787001	-.1264789
_Iid_63	-.1997571	.0355722	-5.62	0.000	-.2711713	-.1283429
_cons	.3182153	.1310041	2.43	0.019	.0552135	.5812171

```
. testparm _I*

 ( 1)  _Iid_2 = 0
 ( 2)  _Iid_3 = 0
 ( 3)  _Iid_4 = 0
 ( 4)  _Iid_5 = 0
 ( 5)  _Iid_6 = 0
 ( 6)  _Iid_7 = 0
 ( 7)  _Iid_8 = 0
 ( 8)  _Iid_9 = 0
 ( 9)  _Iid_10 = 0
 (10)  _Iid_11 = 0
 (11)  _Iid_12 = 0
 (12)  _Iid_13 = 0
 (13)  _Iid_14 = 0
 (14)  _Iid_15 = 0
 (15)  _Iid_16 = 0
 (16)  _Iid_17 = 0
 (17)  _Iid_18 = 0
 (18)  _Iid_19 = 0
 (19)  _Iid_22 = 0
 (20)  _Iid_23 = 0
 (21)  _Iid_24 = 0
 (22)  _Iid_25 = 0
 (23)  _Iid_26 = 0
 (24)  _Iid_27 = 0
 (25)  _Iid_28 = 0
 (26)  _Iid_29 = 0
 (27)  _Iid_30 = 0
 (28)  _Iid_31 = 0
 (29)  _Iid_32 = 0
 (30)  _Iid_33 = 0
 (31)  _Iid_35 = 0
 (32)  _Iid_36 = 0
 (33)  _Iid_37 = 0
 (34)  _Iid_39 = 0
 (35)  _Iid_40 = 0
 (36)  _Iid_41 = 0
 (37)  _Iid_42 = 0
 (38)  _Iid_43 = 0
 (39)  _Iid_44 = 0
 (40)  _Iid_45 = 0
 (41)  _Iid_46 = 0
 (42)  _Iid_47 = 0
 (43)  _Iid_48 = 0
 (44)  _Iid_49 = 0
 (45)  _Iid_50 = 0
 (46)  _Iid_51 = 0
 (47)  _Iid_52 = 0
 (48)  _Iid_54 = 0
 (49)  _Iid_55 = 0
 (50)  _Iid_56 = 0
 (51)  _Iid_57 = 0
 (52)  _Iid_58 = 0
 (53)  _Iid_59 = 0
 (54)  _Iid_60 = 0
 (55)  _Iid_61 = 0
 (56)  _Iid_62 = 0
 (57)  _Iid_63 = 0

       F( 57,    51) =    4.16
            Prob > F =    0.0000
```

```
Random-effects GLS regression                   Number of obs      =       114
Group variable: id                              Number of groups   =        58

R-sq:  within  = 0.3317                         Obs per group: min =         1
       between = 0.5010                                        avg =       2.0
       overall = 0.4314                                        max =         2

                                                Wald chi2(5)       =     69.81
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

------------------- theta --------------------
  min      5%       median        95%      max
0.3399   0.4723     0.4723     0.4723   0.4723
```

growth	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
tourism	.0091043	.0019265	4.73	0.000	.0053285	.0128802
lyo	-.007757	.0031227	-2.48	0.013	-.0138773	-.0016366
prim	.0001867	.0001291	1.45	0.148	-.0000663	.0004397
infl	-.0023485	.0007261	-3.23	0.001	-.0037716	-.0009254
sw	.0116479	.0057053	2.04	0.041	.0004657	.0228301
_cons	-.0785983	.0207657	-3.79	0.000	-.1192984	-.0378983
sigma_u	.0140329					
sigma_e	.01232982					
rho	.56433332	(fraction of variance due to u_i)				

```
Breusch and Pagan Lagrangian multiplier test for random effects

        growth[id,t] = Xb + u[id] + e[id,t]

        Estimated results:
```

	Var	sd = sqrt(Var)
growth	.0007004	.0264648
e	.000152	.0123298
u	.0001969	.0140329

```
        Test:   Var(u) = 0
                             chibar2(01) =     4.02
                          Prob > chibar2 =   0.0225
```

Variable	OLS	FE	LSDV	RE
tourism	0.0078	0.0271	0.0271	0.0091
	0.0014	0.0058	0.0058	0.0019
	0.0000	0.0000	0.0000	0.0000
lyo	-0.0038	-0.0874	-0.0874	-0.0078
	0.0021	0.0135	0.0135	0.0031
	0.0695	0.0000	0.0000	0.0130
prim	0.0001	0.0007	0.0007	0.0002
	0.0001	0.0003	0.0003	0.0001
	0.3955	0.0114	0.0114	0.1481
infl	-0.0029	-0.0007	-0.0007	-0.0023
	0.0008	0.0008	0.0008	0.0007
	0.0007	0.3697	0.3697	0.0012
sw	0.0090	0.0107	0.0107	0.0116
	0.0047	0.0072	0.0072	0.0057
	0.0566	0.1431	0.1431	0.0412
_Iid_2			-0.1752	
			0.0459	
			0.0004	
_Iid_3			-0.1326	
			0.0311	
			0.0001	
_Iid_4			-0.0540	
			0.0144	
			0.0005	
_Iid_5			-0.1772	
			0.0522	
			0.0013	
_Iid_6			-0.2255	
			0.0512	
			0.0001	
_Iid_7			-0.1400	
			0.0346	
			0.0002	
_Iid_8			-0.1168	
			0.0477	
			0.0178	
_Iid_9			-0.1366	
			0.0519	
			0.0113	
_Iid_10			-0.0171	
			0.0157	
			0.2805	
_Iid_11			-0.2471	
			0.0451	
			0.0000	
_Iid_12			-0.0908	
			0.0199	
			0.0000	
_Iid_13			-0.2450	
			0.0513	
			0.0000	
_Iid_14			-0.1036	
			0.0352	
			0.0049	
_Iid_15			-0.0436	
			0.0173	
			0.0149	
_Iid_16			-0.1188	
			0.0348	
			0.0013	
_Iid_17			-0.1137	
			0.0244	
			0.0000	
_Iid_18			-0.1675	
			0.0316	
			0.0000	
_Iid_19			-0.2531	
			0.0613	
			0.0001	
_Iid_20			(omitted)	
_Iid_21			(omitted)	
_Iid_22			-0.1709	
			0.0429	
			0.0002	

_Iid_23	-0.1904
	0.0421
	0.0000
_Iid_24	-0.1018
	0.0256
	0.0002
_Iid_25	-0.1868
	0.0401
	0.0000
_Iid_26	-0.1632
	0.0329
	0.0000
_Iid_27	-0.2482
	0.0440
	0.0000
_Iid_28	-0.1834
	0.0323
	0.0000
_Iid_29	0.0707
	0.0166
	0.0001
_Iid_30	-0.0948
	0.0200
	0.0000
_Iid_31	-0.1118
	0.0239
	0.0000
_Iid_32	-0.2444
	0.0424
	0.0000
_Iid_33	0.0304
	0.0131
	0.0246
_Iid_34	(omitted)
_Iid_35	-0.2080
	0.0469
	0.0000
_Iid_36	-0.2675
	0.0530
	0.0000
_Iid_37	-0.0670
	0.0184
	0.0006
_Iid_38	(omitted)
_Iid_39	-0.0214
	0.0188
	0.2597
_Iid_40	-0.1424
	0.0213
	0.0000
_Iid_41	-0.1598
	0.0284
	0.0000

_Iid_42			-0.2631	
			0.0489	
			0.0000	
_Iid_43			-0.1915	
			0.0388	
			0.0000	
_Iid_44			-0.1523	
			0.0534	
			0.0062	
_Iid_45			-0.2341	
			0.0467	
			0.0000	
_Iid_46			-0.1628	
			0.0406	
			0.0002	
_Iid_47			-0.0594	
			0.0345	
			0.0910	
_Iid_48			-0.0917	
			0.0226	
			0.0002	
_Iid_49			-0.0830	
			0.0211	
			0.0003	
_Iid_50			-0.1708	
			0.0279	
			0.0000	
_Iid_51			-0.1357	
			0.0514	
			0.0110	
_Iid_52			-0.1437	
			0.0387	
			0.0005	
_Iid_53			(omitted)	
_Iid_54			-0.0951	
			0.0153	
			0.0000	
_Iid_55			-0.1535	
			0.0341	
			0.0000	
_Iid_56			-0.0997	
			0.0230	
			0.0001	
_Iid_57			-0.1736	
			0.0440	
			0.0002	
_Iid_58			-0.1148	
			0.0201	
			0.0000	
_Iid_59			-0.1879	
			0.0463	
			0.0002	
_Iid_60			-0.0186	
			0.0156	
			0.2386	
_Iid_61			-0.0281	
			0.0168	
			0.0993	
_Iid_62			-0.2026	
			0.0379	
			0.0000	
_Iid_63			-0.1998	
			0.0356	
			0.0000	
_cons	-0.0791	0.1805	0.3182	-0.0786
	0.0159	0.1040	0.1310	0.0208
	0.0000	0.0886	0.0187	0.0002
N	114	114	114	114
df_r	108.0000	51.0000	51.0000	
df_m	5.0000	62.0000	62.0000	5.0000
r2	0.4468	0.6368	0.9020	
r2_a	0.4211	0.1953	0.7829	
rmse	0.0201	0.0123	0.0123	0.0146
F	13.9723	17.8855	7.5741	

legend: b/se/p

```
. hausman FE RE, sigmamore

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       FE           RE         Difference          S.E.
-------------+----------------------------------------------------------------
     tourism |    .0270759     .0091043        .0179716        .0065577
         lyo |   -.0873547     -.007757       -.0795977         .015762
        prim |    .0006828     .0001867        .0004962        .0002808
        infl |   -.0007259    -.0023485        .0016226        .0006166
          sw |    .0107407     .0116479       -.0009073        .0064038
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(5) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       35.07
                Prob>chi2 =      0.0000
```

Selma JELASSI.