

Ecole Doctorale Edocif

Séminaire Méthodologique

L'Econométrie des Données de Panel

Modèles Linéaires Simples

Christophe HURLIN



Figure 0.1:

Présentation

Le but de ce séminaire est de proposer une initiation, tant sur le plan théorique que sur le plan appliqué, à l'économétrie des données de panel. Sur le plan théorique, le séminaire débutera par une présentation des problèmes de spécifications de base en économétrie de panel et par les méthodes d'estimation traditionnelles. Cette première partie insistera en particulier sur les stratégies de tests de spécification du modèle. La seconde partie du séminaire sera consacrée à l'étude des panels dynamiques et proposera une introduction aux concepts de non stationnarité stochastique en panel et notamment à la spécification des tests de non stationnarité et de cointégration. Ces concepts théoriques seront appliqués à différentes problématiques économiques, comme par exemple la problématique de la convergence des PIB par tête, que ce soit au travers de commentaires de travaux empiriques ou par la programmation d'exemples illustratifs sous un logiciel d'économétrie (TSP ou SAS à préciser)

Table des Matières

Introduction	2
Chapitre 1 : Modèles Linéaires Simples	5
Introduction	6
1. Tests de spécification ou tests d'homogénéité	8
1.1. L'exemple d'une fonction de production	8
1.2. Procédure de tests de spécification	10
1.2.1. <i>Procédure générale</i>	10
1.2.2. <i>Construction des statistiques de test</i>	13
1.3. Application	17
2. Modèles à effets individuels	20
2.1. Empilement par pays	21
2.1.1. <i>Exemple</i>	22
2.2. Empilement par dates	23
2.2.1. <i>Exemple</i>	23
3. Modèles à effets fixes	25
3.1. Estimateur Within ou LSDV	26
3.2. Application	28
3.3. Ecriture vectorielle	31
4. Modèles à effets aléatoires	33
4.1. Modèle à variance composée	34
4.2. Estimateurs du modèle à effets aléatoires	35
4.3. Estimateur des Moindres Carrés Généralisés	37
4.4. Application	42
5. Tests de spécification des effets individuels	44
5.1. Corrélation des effets individuels et des variables explicatives	45
5.2. Test de spécification d'Hausman	49
6.3. Application	51
6. Modèles à coefficients fixes et aléatoires	53
6.1. Modèle MFR de Hsiao (1989)	53
6.2. Méthode d'estimation des paramètres	56
6.3. Application	57

Annexes	62
<i>A.1. Equivalence entre les estimateurs Within et pooled : $\psi = 1$</i>	62
<i>A.2. Programme TSP : Modèle MFR</i>	63
Bibliographie	67

Chapitre I

Modèles Linéaires Simples

Introduction

Dans ce premier chapitre, nous nous restreindrons à l'étude des modèles linéaires simples sur données de panel, ces derniers étant définis par opposition aux modèles dynamiques faisant intervenir des variables endogènes retardées. Dans toute la suite de ce chapitre, on considère des processus strictement stationnaires au sens de la stationnarité du second ordre¹.

L'objectif des cinq premières sections de ce chapitre est de faire en sorte que le lecteur puisse interpréter, de façon exhaustive et relativement approfondie, les résultats de base que donnent les principaux logiciels d'économétrie lorsque l'on envisage des modèles de panel. Nous prendrons ici comme référence le logiciel TSP 4.3, mais il est bien entendu évident que ces résultats de base sont sensiblement identiques si l'on considère d'autres logiciels comme Eviews, SAS ou Rats.

Quelles sont les connaissances minimales nécessaires à l'économètre appliqué pour pouvoir interpréter un tableau de résultats d'estimation de panel ? Prenons comme exemple la commande *panel* de TSP. A l'issue de cette procédure, l'utilisateur dispose des réalisations des estimateurs *Pooled*, *Between*, des estimateurs du modèle à effets individuels fixes (*Within*), du modèle à effets individuels aléatoires (*Error Component Model*), des résultats de trois tests de Fischer, d'un estimateur de la variance des effets individuels, d'un estimateur de la variance totale, de l'estimateur d'un paramètre de pondération et de la statistique du test d'Hausman. Voilà ainsi résumés tous les éléments que nous nous proposons d'étudier tout au long des cinq premières sections.

Ainsi, la première section est consacrée aux tests de spécification qui correspondent en fait aux trois tests de Fischer du fichier de résultat des estimations TSP. Il s'agit ici de déterminer la manière dont doit être spécifié un modèle de panel, si tant que l'hypothèse de panel puisse être acceptée. Nous verrons alors que toute l'analyse repose alors sur la notion d'homogénéité des paramètres du modèle envisagé. Les tests présentés visent ainsi à porter un diagnostic sur l'éventuelle nécessité d'intégrer une dimension hétérogène et sur la manière dont cette hétérogénéité doit être spécifiée. Une des manières simples pour ne pas supposer l'existence d'un modèle totalement identique consiste ainsi à introduire des effets individuels sous la forme de constantes spécifiques, propres à chaque individu du panel. Nous proposerons ainsi une procédure générale de tests de spécification des modèles linéaires simples.

¹La définition de la stationnarité fera l'objet du chapitre suivant.

Dans une seconde section, nous présenterons le modèle à effets individuels. Ce modèle suppose l'existence de coefficients identiques pour tous les individus et de constantes spécifiques. Ainsi, la relation économique mise en évidence à travers ce type de modélisation n'est censée différer pour tous les individus qu'au niveau des constantes introduites dans le modèle. Nous insisterons alors sur la construction du modèle et en particulier sur sa représentation sous forme vectorielle. Il s'agit ici de présenter les deux principales méthodes de construction des échantillons de données, à savoir la méthode d'empilement par date et la méthode d'empilement par individus.

La troisième section est consacré au modèle à effets individuels fixes. On suppose dans cette section que les effets individuels sont des paramètres déterministes. Nous serons alors amenés à présenter la construction et les propriétés de l'estimateur *Within* des coefficients du modèle. Une application sur les relations entre le nombre de grèves dans le secteur industriel et les déterminants macroéconomiques sera proposée. Nous nous appuierons sur les résultats et les différents tests proposés sous TSP.

Une quatrième section sera consacrée au modèle à effets individuels aléatoires. On suppose dans cette section que les effets individuels ne sont plus des paramètres, mais des variables aléatoires possédant une distribution commune pour tous les individus. Nous commencerons par étudier la structure des résidus de ce modèle à variance composée. Nous étudierons ensuite les différents estimateurs des coefficients du modèle à effets aléatoires et plus particulièrement l'estimateur des Moindres Carrés Généralisés.

Enfin, la cinquième section sera consacrée aux tests de spécification des effets individuels. La question est alors de savoir quel modèle, parmi les modèles à effets aléatoires et effets fixes, doit être retenu. Ceci nous conduira à présenter le test d'Hausman (1978) qui constitue le test standard de spécification des effets individuels.

Nous concluons enfin en introduisant une sixième section consacrée à la présentation d'un type de modèle de plus en plus utilisé en économétrie appliqué : le modèle avec coefficients fixes et coefficients aléatoires. Ce modèle, introduit notamment par Hsiao (1989), permet de réaliser différents exercices de modélisation économique particulièrement intéressants.

1. Tests de spécification ou tests d'homogénéité

Lorsque l'on considère un échantillon de données de panel, la toute première chose qu'il convient de vérifier est la *spécification homogène ou hétérogène du processus générateur de données*. Sur le plan économétrique, cela revient à tester l'égalité des coefficients du modèle étudié dans la dimension individuelle. Sur le plan économique, les tests de spécification reviennent à déterminer si l'on est en droit de supposer que le modèle théorique étudié est parfaitement identique pour tous les pays, ou au contraire s'il existe des spécificités propres à chaque pays.

Nous commencerons par présenter un petit exemple illustratif construit à partir d'une fonction de production agrégée. Dans une seconde section, nous généraliserons la procédure de tests de spécification en introduisant les notions de variances intra-classes (*Within*) et inter-classes (*Between*). Enfin, nous proposerons une application de ces tests de spécification à l'estimation d'une relation entre le nombre de grèves et un ensemble de variables explicatives.

1.1. L'exemple d'une fonction de production

Prenons l'exemple d'une fonction de production. Supposons que l'on dispose d'un échantillon de données de *PIB* et de facteurs (travail et capital) sur une durée de T périodes pour un ensemble de N pays. Si l'on note $y_{i,t}$ le logarithme du *PIB*, $k_{i,t}$ le logarithme du stock de capital privé et $n_{i,t}$ le logarithme de l'emploi et que l'on suppose une fonction de production de type Cobb Douglass, le modèle général s'écrit sous la forme $\forall i \in [1, N], \forall t \in [1, T]$:

$$y_{i,t} = \alpha_i + \beta_i k_{i,t} + \gamma_i n_{i,t} + \varepsilon_{i,t}$$

Les innovations $\varepsilon_{i,t}$ sont supposées être *i.i.d.* de moyenne nulle et de variance égale à $\sigma_\varepsilon^2, \forall i \in [1, N]$.

Dans un premier temps, la phase de test de spécification, revient sur le plan économique, à déterminer s'il est en droit de supposer une fonction de production totalement identique pour tous les pays (*modèle pooled*). Dans ce cas, les élasticités de l'emploi et du capital sont identiques pour tous les pays ($\beta_i = \beta, \gamma_i = \gamma, \forall i \in [1, N]$) et le niveau moyen de la productivité totale des facteurs, représenté ici par les constantes α_i , est lui aussi identique pour tous les pays. Le modèle s'écrit alors sous la forme :

$$y_{i,t} = \alpha + \beta k_{i,t} + \gamma n_{i,t} + \varepsilon_{i,t}$$

Toutefois, lorsque l'on travaille sur des séries agrégées, il est relativement peu probable que la fonction de production macroéconomique soit strictement identique pour tous les pays étudiés. Si l'hypothèse d'homogénéité totale est rejetée, il convient alors de tester si les élasticités des différents facteurs sont identiques. Si ce n'est pas le cas, il n'existe a priori aucune structure de production commune entre les pays. *Dans ce cas, l'utilisation des données de panel ne se justifie pas et peut même conduire à des biais d'estimation.* On doit donc estimer les fonctions de production pays par pays.

Si en revanche, il s'avère qu'il existe bien une relation identique entre la production et les facteurs pour tous les pays, la source d'hétérogénéité du modèle peut alors provenir des constantes α_i . Dans notre exemple, ces constantes représentent la moyenne de la productivité totale des facteurs de production (résidu de Solow). Or, rien ne garantit que les pays étudiés possèdent le même niveau de productivité structurelle. Au contraire, il se peut parfaitement que des *facteurs a-temporels ou structurels*, comme la position géographique, le climat, l'éloignement par rapport aux grands axes commerciaux etc... aient pu conduire à des différences structurelles de niveau de productivité entre les pays. Il convient donc de tester l'hypothèse d'une constante commune à tous les pays. Si cette hypothèse est rejetée, on obtient alors un modèle avec *effets individuels* qui s'écrit sous la forme :

$$y_{i,t} = \alpha_i + \beta k_{i,t} + \gamma n_{i,t} + \varepsilon_{i,t}$$

Dans ce cas, le niveau moyen de la productivité totale des facteurs, déterminé par $E(\alpha_i + \varepsilon_{i,t}) = \alpha_i$, varie selon les pays, même si la structure de production (les élasticités des deux facteurs travail et capital) est identique.

Cet exemple, que nous allons à présent généraliser, montre bien que *la phase de test de spécification revient à déterminer si le processus générateur de données peut être considéré comme homogène, c'est à dire unique pour tous les individus, ou si au contraire il apparaît totalement hétérogène, auquel cas l'utilisation des techniques de panel ne peut se justifier.* Entre ces deux cas extrêmes, il convient de précisément identifier la source d'hétérogénéité pour bien spécifier le modèle.

1.2. Procédure de tests de spécification

On considère un échantillon de T observations de N processus individuels $\{y_{i,t}, t \in \mathbb{Z}, i \in \mathbb{N}\}$ et $\{x_{i,t}, t \in \mathbb{Z}, i \in \mathbb{N}\}$. Par la suite, on notera $\{y_{i,t}\}$ et $\{x_{i,t}\}$ ces deux processus. On suppose que le processus $\{y_{i,t}\}$ est défini de façon générale par le relation linéaire suivante, $\forall i \in \mathbb{N}, \forall t \in \mathbb{Z}$:

$$y_{i,t} = \alpha_i + \beta_i' x_{i,t} + \varepsilon_{i,t} \quad (1.1)$$

où $\alpha_i \in \mathbb{R}, \beta_i = (\beta_{1,i}, \beta_{2,i}, \dots, \beta_{K,i})'$ est un vecteur de dimension $(K, 1)$. On considère ainsi un vecteur de K variables explicatives :

$$x_{i,t} = (x_{1,i,t}, x_{2,i,t}, \dots, x_{K,i,t})' \quad (1.2)$$

Les innovations $\varepsilon_{i,t}$ sont supposées être *i.i.d.* de moyenne nulle et de variance égale à $\sigma_\varepsilon^2, \forall i \in [1, N]$. Ainsi on suppose que les paramètres α_i et β_i du modèle (1.1) peuvent différer dans la dimension individuelle², mais l'on suppose qu'ils sont constants dans le temps.

1.2.1. Procédure générale

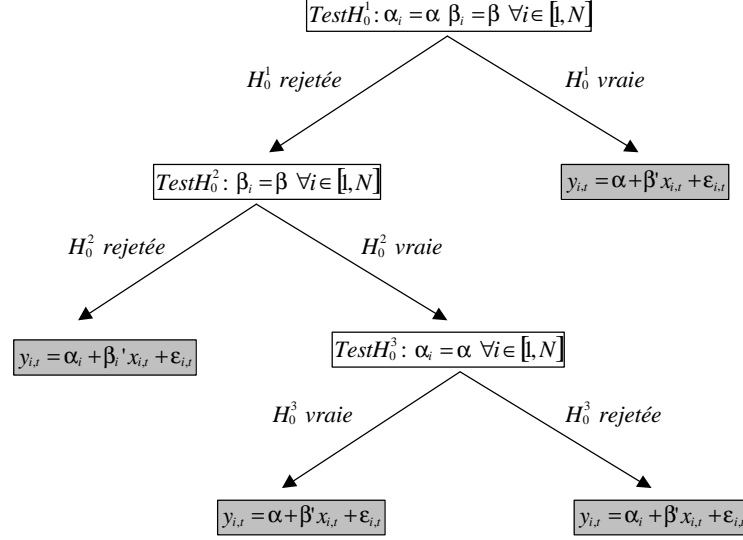
Si l'on considère le modèle (1.1), plusieurs configurations sont alors possibles :

1. Les N constantes α_i et les N vecteurs de paramètres β_i sont identiques : $\alpha_i = \alpha, \beta_i = \beta \forall i \in [1, N]$. On qualifie alors le panel de panel *homogène*.
2. Les N constantes α_i et les N vecteurs de paramètres β_i sont différents selon les individus. On a donc N modèles différents, on rejette la structure de panel.
3. Les N constantes α_i sont identiques, $\alpha_i = \alpha \forall i \in [1, N]$, tandis que les vecteurs de paramètres β_i diffèrent selon les individus. Dans ce cas, tous les coefficients du modèle, à l'exception des constantes, sont différents selon les individus. On a donc N modèles différents.
4. Les N vecteurs de paramètres β_i sont identiques, $\beta_i = \beta \forall i \in [1, N]$, tandis que les constantes α_i diffèrent selon les individus. On obtient *un modèle à effets individuels*.

Pour discriminer ces différentes configurations et pour s'assurer du bien fondé de la structure de panel, il convient d'adopter une procédure de tests d'homogénéité emboîtés. La procédure générale de test présentée dans Hsiao (1986) est décrite sur la figure (1.1).

²A l'exception de la variance des innovations que l'on supposera identique pour tous les individus.

Figure 1.1: Procédure Générale de Tests d'Homogénéité



Dans une première étape, on teste l'hypothèse d'une structure parfaitement homogène (constantes et coefficients identiques) :

$$H_0^1 : \beta_i = \beta \quad \alpha_i = \alpha \quad \forall i \in [1, N]$$

$$H_a^1 : \exists (i, j) \in [1, N] / \beta_i \neq \beta_j \text{ ou } \alpha_i \neq \alpha_j$$

On utilise alors une statistique de Fischer pour tester ces $(K + 1)(N - 1)$ restrictions linéaires³. Si l'on suppose que les résidus $\varepsilon_{i,t}$ sont indépendamment distribués dans les dimensions i et t , suivant une loi normale d'espérance nulle et de variance finie σ_ε^2 , cette statistique suit une distribution de Fischer avec $(K + 1)(N - 1)$ et $NT - N(K + 1)$ degrés de liberté. Les conclusions de ce test sont les suivantes. Si l'on accepte l'hypothèse nulle H_1^0 d'homogénéité, on obtient alors un modèle de *pooled totalement homogène*.

$$y_{i,t} = \alpha + \beta' x_{i,t} + \varepsilon_{i,t} \quad (1.3)$$

Si en revanche, on rejette l'hypothèse nulle, on passe à une seconde étape qui consiste à déterminer si l'hétérogénéité provient des coefficients β_i .

³Dans notre modèle, chaque vecteur β_i comprend K paramètres. Pour les N individus du panel, on obtient KN paramètres. L'égalité des N vecteurs β_i revient donc à imposer $KN - K$ restrictions. De la même façon, l'égalité des N constantes revient à imposer $N - 1$ restrictions. Au total, l'hypothèse H_0^1 revient à imposer $(K + 1)(N - 1)$ restrictions linéaires.

La seconde étape consiste à tester l'égalité pour tous les individus des K composantes des vecteurs β_i .

$$H_0^2 : \beta_i = \beta \forall i \in [1, N]$$

$$H_a^2 : \exists (i, j) \in [1, N] / \beta_i \neq \beta_j$$

Sous l'hypothèse nulle, on n'impose ici aucune restriction sur les constantes individuelles α_i . De la même façon, on construit une statistique de Fischer pour tester ces $(N - 1)K$ restrictions linéaires. Toujours sous l'hypothèse d'indépendance et de normalité des résidus, cette statistique suit une loi de Fischer avec $(N - 1)K$ et $NT - N(K + 1)$ degrés de liberté. *Si l'on rejette l'hypothèse nulle H_2^0 d'homogénéité des coefficients β_i , on rejette alors la structure de panel*, puisque au mieux seules les constantes α_i peuvent être identiques entre les individus :

$$y_{i,t} = \alpha + \beta_i' x_{i,t} + \varepsilon_{i,t} \quad (1.4)$$

On estime alors les paramètres vectoriels β_i en utilisant les modèles différents pays par pays. *Si en revanche, on accepte l'hypothèse nulle H_2^0 d'homogénéité des coefficients β_i , on retient la structure de panel et l'on cherche alors à déterminer dans une troisième étape si les constantes α_i ont une dimension individuelle.*

La troisième étape de la procédure consiste à tester l'égalité des N constantes individuelles α_i , sous l'hypothèse de coefficients β_i communs à tous les individus :

$$H_0^3 : \alpha_i = \alpha \forall i \in [1, N]$$

$$H_a^3 : \exists (i, j) \in [1, N] / \alpha_i \neq \alpha_j$$

Sous l'hypothèse nulle, on impose $\beta_i = \beta$. Sous l'hypothèse d'indépendance et de normalité des résidus, on construit une statistique de Fischer pour tester ces $N - 1$ restrictions linéaires. Cette statistique suit une loi de Fischer avec $(N - 1)K$ et $N(T - 1) - K$ degrés de liberté. *Si l'on rejette l'hypothèse nulle H_3^0 d'homogénéité des constantes α_i , on obtient alors un modèle de panel avec effets individuels :*

$$y_{i,t} = \alpha_i + \beta' x_{i,t} + \varepsilon_{i,t} \quad (1.5)$$

Dans le cas où l'on accepte l'hypothèse nulle H_3^0 , on retrouve alors une structure de panel totalement homogène (*modèle pooled*). Le test H_3^0 ne sert alors qu'à confirmer ou infirmer les conclusions du tests H_1^0 , étant donné que le fait de réduire le nombre de restrictions linéaires permet d'accroître la puissance du test du Fischer.

1.2.2. Construction des statistiques de test

Nous allons à présent, présenter les méthodes de construction des différents tests de Fischer utilisés dans cette procédure. On considère le modèle (1.1) et l'on suppose que les résidus $\varepsilon_{i,t}$ sont indépendamment distribués dans les dimensions i et t , suivant une loi normale d'espérance nulle et de variance finie σ_ε^2 . On suppose que la variance σ_ε^2 est connue.

Test d'homogénéité globale On considère tout d'abord, le test de l'hypothèse d'homogénéité totale H_0^1 :

$$H_0^1 : \beta_i = \beta \quad \alpha_i = \alpha \quad \forall i \in [1, N]$$

Soit F_1 la statistique de Fischer associée à ce test. Ce test revient à imposer $(K+1)(N-1)$ restrictions linéaires sur les coefficients du modèle (1.1). De plus, sous l'hypothèse alternative H_a^2 , il existe au plus NK coefficients différents pour les composantes des N vecteurs β_i (de dimension K) et N constantes individuelles. On dispose donc de $NT - N(K+1)$ degrés de liberté.

Définition 1.1. La statistique de Fischer F_1 associée au test d'homogénéité totale H_0^1 dans le modèle (1.1) :

$$H_0^1 : \beta_i = \underset{(K,1)}{\beta} \quad \alpha_i = \alpha \quad \forall i \in [1, N]$$

s'écrit sous la forme suivante et suit un Fischer avec $(N-1)(K+1)$ et $NT - N(K+1)$ degrés de liberté :

$$F_1 = \frac{(SCR_{1,c} - SCR_1) / [(N-1)(K+1)]}{SCR_1 / [NT - N(K+1)]} \quad (1.6)$$

où SCR_1 désigne la somme des carrés des résidus du modèle (1.1) et $SCR_{1,c}$ la somme des carrés des résidus du modèle contraint :

$$y_{i,t} = \alpha + \beta' x_{i,t} + \varepsilon_{i,t}$$

Ainsi, si la réalisation de la statistique de Fischer pour l'échantillon considéré est supérieure au seuil théorique à $\alpha\%$, on rejette l'hypothèse nulle d'homogénéité. Reste à présent à déterminer la formule générale des sommes de carrés des résidus des modèles contraint et non contraint. Considérons tout d'abord le modèle non contraint :

$$y_{i,t} = \alpha_i + \beta_i' x_{i,t} + \varepsilon_{i,t}$$

Les estimateurs $\hat{\beta}_i$ et $\hat{\alpha}_i$ des paramètres individuels sont obtenus équation par équation pour chaque individu. Soit $SCR_{1,i}$ la somme des carrés des résidus obtenue pour

chaque équation. La somme des carrés des résidus du modèle (1.1) non contraint est alors tout simplement définie comme la somme des N somme des carrés des résidus obtenues pour les N équations individuelles.

$$SCR_1 = \sum_{i=1}^N SCR_{1,i} = \sum_{i=1}^N \left[S_{yy,i} - S'_{xy,i} S_{xx,i}^{-1} S_{xy,i} \right] \quad (1.7)$$

où les sommes $S_{k,i}$ sont définies de la façon suivante $\forall i \in [1, N]$:

$$S_{yy,i} = \sum_{t=1}^T (y_{i,t} - \bar{y}_i)^2 \quad (1.8)$$

$$S_{xx,i} = \sum_{t=1}^T (x_{i,t} - \bar{x}_i) (x_{i,t} - \bar{x}_i)' \quad (1.9)$$

$$S_{xy,i} = \sum_{t=1}^T (x_{i,t} - \bar{x}_i) (y_{i,t} - \bar{y}_i)' \quad (1.10)$$

Les moyennes \bar{x}_i et \bar{y}_i sont définies, pour chaque individu, dans la dimension temporelle de façon traditionnelle par $\forall i \in [1, N]$:

$$\bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{i,t}$$

$$\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{i,t}$$

L'expression de SCR_1 fait ainsi apparaître la somme des variances individuelles des résidus obtenues à partir des N régressions individuelles. L'expression $S_{yy,i} - S'_{xy,i} S_{xx,i}^{-1} S_{xy,i}$ correspond ainsi (à une transformée linéaire près) à la variance intra-groupe (*Within variance*) des résidus.

Le modèle (1.1) contraint sous l'hypothèse H_0^1 s'écrit :

$$y_{i,t} = \alpha + \beta' x_{i,t} + \varepsilon_{i,t} \quad (1.11)$$

On dispose ainsi d'un échantillon de TN observations pour identifier les paramètres communs α et β de cette relation. On applique alors les Moindres Carrés Ordinaires sur les données empilées (modèle *pooled*). La somme des carrés s'écrit sous la forme :

$$SCR_{1,c} = S_{yy} - S'_{xy} S_{xx}^{-1} S_{xy} \quad (1.12)$$

où les sommes S_k sont définies de la façon suivante :

$$S_{yy} = \sum_{i=1}^N \sum_{t=1}^T (y_{i,t} - \bar{y}_i)^2 \quad (1.13)$$

$$S_{xx,i} = \sum_{t=1}^T \sum_{i=1}^N (x_{i,t} - \bar{x}_i) (x_{i,t} - \bar{x}_i)' \quad (1.14)$$

$$S_{xy,i} = \sum_{t=1}^T \sum_{i=1}^N (x_{i,t} - \bar{x}_i) (y_{i,t} - \bar{y}_i)' \quad (1.15)$$

Les moyennes \bar{x} et \bar{y} sont alors définies sur l'ensemble des TN observations :

$$\bar{x} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T x_{i,t}$$

$$\bar{y} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T y_{i,t}$$

L'expression $SCR_{1,c}$ correspond à une transformée de la variance totale des résidus (total variance) obtenus à partir de l'estimation d'un modèle unique sur les NT données empilées.

Test d'homogénéité des coefficients β_i

Considérons à présent le test de l'hypothèse d'homogénéité des coefficients β_i , notée H_0^2 :

$$H_0^2 : \beta_i = \beta \quad \forall i \in [1, N]$$

Soit F_2 la statistique de Fischer associée à ce test. Sous l'hypothèse nulle, on n'impose aucune restriction sur les constantes individuelles α_i . Toujours sous l'hypothèse d'indépendance et de normalité des résidus, on construit une statistique de Fischer pour tester ces $(N-1)K$ restrictions linéaires. Sous l'hypothèse alternative H_a^2 , on retrouve le modèle (1.1) et $NT - N(K+1)$ degrés de liberté.

Definition 1.2. La statistique de Fischer F_2 associée au test d'homogénéité totale H_0^2 dans le modèle (1.1) :

$$H_0^2 : \beta_i = \beta \quad \forall i \in [1, N]$$

(K,1)

s'écrit sous la forme suivante et suit un Fischer avec $(N-1)K$ et $NT - N(K+1)$ degrés de liberté :

$$F_2 = \frac{(SCR_{1,c'} - SCR_1) / [(N-1)K]}{SCR_1 / [NT - N(K+1)]} \quad (1.16)$$

où SCR_1 désigne la somme des carrés des résidus du modèle (1.1) et $SCR_{1,c'}$ la somme des carrés des résidus du modèle contraint (modèle à effets individuels) :

$$y_{i,t} = \alpha_i + \beta' x_{i,t} + \varepsilon_{i,t}$$

La somme des carrés des résidus du modèle non contraint, SCR_1 , est donnée par l'équation (1.7). Sous l'hypothèse H_0^2 , la somme des carrés des résidus dans le modèle à effets individuels est donnée par :

$$SCR_{1,c'} = \sum_{i=1}^N S_{yy,i} - \left(\sum_{i=1}^N S_{xy,i} \right)' \left(\sum_{i=1}^N S_{xx,i} \right)^{-1} \left(\sum_{i=1}^N S_{xy,i} \right) \quad (1.17)$$

où les sommes $S_{k,i}$ ont été définies par les équations (1.8), (1.9) et (1.10). Cette écriture signifie que dans le modèle à effets individuels, les estimateurs (*Within*) des paramètres β_i et α_i sont obtenus en centrant les variables sur leurs moyennes individuelles respectives. Nous reviendrons par la suite sur cette propriété.

Test d'homogénéité des constantes α_i

Considérons enfin le dernier test d'homogénéité des constantes α_i , notée H_0^3 :

$$H_0^3 : \alpha_i = \alpha \quad \forall i \in [1, N]$$

Soit F_3 la statistique de Fischer associée à ce test. Sous l'hypothèse nulle, on impose l'égalité des paramètres β_i . Sous l'hypothèse d'indépendance et de normalité des résidus, on construit une statistique de Fischer pour tester ces $N - 1$ restrictions linéaires. Sous l'hypothèse alternative H_a^3 , les coefficients β_i sont tous égaux, mais les constantes diffèrent selon les individus. On a donc $NT - N - K$ degrés de liberté.

Definition 1.3. La statistique de Fischer F_3 associée au test d'homogénéité totale H_0^3 dans le modèle (1.1) :

$$H_0^3 : \alpha_i = \alpha \quad \forall i \in [1, N]$$

s'écrit sous la forme suivante et suit un Fischer avec $N - 1$ et $N(T - 1) - K$ degrés de liberté :

$$F_3 = \frac{(SCR_{1,c} - SCR_{1,c'}) / (N - 1)}{SCR_{1,c'} / [N(T - 1) - K]} \quad (1.18)$$

où $SCR_{1,c'}$ désigne la somme des carrés des résidus du modèle (1.1) sous l'hypothèse $\beta_i = \beta$ (modèle à effets individuels) et $SCR_{1,c}$ la somme des carrés des résidus du modèle contraint (modèle de pooled) :

$$y_{i,t} = \alpha + \beta' x_{i,t} + \varepsilon_{i,t}$$

Les sommes des carrés des résidus $SCR_{1,c'}$ et $SCR_{1,c}$ ont été respectivement définies par les équations (1.17) et (1.12).

Il est en outre possible de tester la constance dans le temps des différents paramètres du panel suivant une procédure sensiblement identique (voir Hsiao 1986). Mais cette problématique relève plus de la notion traditionnelle de stabilité des coefficients dans le temps que de la pure application des techniques économétriques de données de panel.

1.3. Application

Considérons à présent une application simple de ces tests d'homogénéité à partir de données de panel relatives au nombre de grèves dans le secteur industriel⁴. Les données annuelles couvrent 17 pays⁵ de l'OCDE et sont disponibles de 1951 à 1985. Soit $s_{i,t}$ le nombre de jours chômés pour cause de grève, pour 1000 salariés du secteur industriel, du pays i observé à la date t . Nous cherchons à relier cette variable d'une part au taux de chômage de l'économie, noté $u_{i,t}$, et d'autre au niveau de l'inflation, notée $p_{i,t}$, selon la relation linéaire suivante :

$$s_{i,t} = \alpha_i + \beta_i u_{i,t} + \gamma_i p_{i,t} + \varepsilon_{i,t} \quad \forall i = 1, \dots, 17 \quad (1.19)$$

Appliquons la stratégie de test d'homogénéité à ce modèle. Pour ce faire, nous utiliserons directement les résultats des tests programmés sous TSP (version 4.3A ou versions ultérieures). Les commandes TSP sont donc les suivantes :

```
load(file='strikes.wks');
panel (id=i,time=t,byid) srt u p;
```

La première ligne du programme sert tout simplement à lire les données du fichier strike.wks. La seconde ligne du programme permet d'obtenir l'ensemble des estimateurs de base (pooled, between, effets fixes et effets aléatoires) ainsi que les résultats des principaux tests de spécification (tests d'homogénéité et test d'Hausman). L'option $id = i$ indique le nom de l'indicatrice pays et l'option $time = t$ indique le nom de l'indicatrice temporelle. L'option $byid$ est nécessaire pour afficher l'ensemble des résultats des tests de spécification⁶. Les variables sont respectivement nommées srt , u et p .

Sur la figure (1.2) sont reportés les résultats d'estimation de cette procédure TSP. Nous allons nous concentrer plus particulièrement sur l'analyse des résultats des tests d'homogénéité. TSP propose, avec l'option $byid$, les trois tests de Fischer présentés précédemment.

On observe tout d'abord *le panel est cylindré (balanced), c'est à dire qu'il comporte le même nombre de points dans la dimension temporelle pour tous les individus*. On a ici $N = 17$ et $T = 35$, soit 595 observations. Commençons tout d'abord par le test

⁴Les données ont été collectées par Bruce Western et proviennent du site <http://lib.stat.cmu.edu/datasets/>

⁵Le panel comporte initialement 18 pays, mais pour le pays 3, les données ne sont disponibles que jusqu'en 1980. C'est pourquoi afin de travailler sur un panel cylindré nous avons choisi de retirer ce pays de notre échantillon.

⁶Pour plus de détails sur l'instruction *panel* sous TSP, se reporter au manuel de programmation du logiciel.

Figure 1.2: Résultats des Tests de Spécification sous TSP 4.3A

```

PANEL DATA ESTIMATION
=====
Balanced data: NI= 17, T= 35, NOB= 595
TOTAL (plain OLS) Estimates:
Dependent variable: SRT

Mean of dependent variable = 305.076      Std. error of regression = 557.258
Std. dev. of dependent var. = 571.637      R-squared = .052874
Sum of squared residuals = .183838E+09     Adjusted R-squared = .049674
Variance of residuals = 310536.

Variable      Estimated      Standard
Coefficient    Error      t-statistic
U             27.4379      7.53997    3.63899
P             18.6136      4.98653    3.73277
C             95.0787      43.1414    2.20389

F test of A,B=Ai,Bi: F(48,544) = 3.8320, P-value = [.0000]
Critical F value for diffuse prior (Leamer, p.114) = 7.6418

BETWEEN (OLS on means) Estimates:
Dependent variable: SRT

Mean of dependent variable = 305.076      Std. error of regression = 189.663
Std. dev. of dependent var. = 278.196      R-squared = .593303
Sum of squared residuals = 503607.         Adjusted R-squared = .535203
Variance of residuals = 35972.0

Variable      Estimated      Standard
Coefficient    Error      t-statistic
U             80.9542      23.1650    3.49467
P             59.3882      32.6068    1.82134
C            -341.546     197.643   -1.72810

WITHIN (fixed effects) Estimates:
Dependent variable: SRT

Sum of squared residuals = .146958E+09     R-squared = .242875
Variance of residuals = 255136.             Adjusted R-squared = .219215
Std. error of regression = 505.110

Variable      Estimated      Standard
Coefficient    Error      t-statistic
U            -21.5968      9.19158    -2.34963
P             16.2729      4.75658     3.42113

F test of Ai,B=Ai,Bi: F(32,544) = 1.1845, P-value = [.2266]
Critical F value for diffuse prior (Leamer, p.114) = 6.9699

F test of A,B=Ai,B: F(16,576) = 9.0342, P-value = [.0000]
Critical F value for diffuse prior (Leamer, p.114) = 6.7476

Variance Components (random effects) Estimates:
VWITH (variance of Uit) = 0.25514E+06
VBET (variance of Ai) = 55401.
(computed from small sample formula)
THETA (0=WITHIN, 1=TOTAL) = 0.11628

Dependent variable: SRT

Sum of squared residuals = .152560E+09     R-squared = .214013
Variance of residuals = 264862.             Adjusted R-squared = .189450
Std. error of regression = 514.647

Variable      Estimated      Standard
Coefficient    Error      t-statistic
U            -12.0814      8.78623    -1.37504
P             16.4247      4.72670     3.47489
C            248.622     71.9106     3.45737

Hausman test of H0:RE vs. FE: CHISQ(2) = 13.924, P-value = [.0009]

```

de l'hypothèse d'homogénéité totale (test H_0^1 dans nos notations). Ce dernier est noté sous la forme **F test of $A, B = A_i, B_i$** dans le fichier résultat de TSP. La lettre A désigne ici les constantes, tandis que la lettre B désigne le vecteur des coefficients des variables explicatives. Dans le cadre de notre échantillon, la réalisation de la statistique de Fischer associée au test H_0^1 , notée F_1 , est de 3.8320. Le logiciel indique en outre le nombre de degré de liberté de cette statistique. Nous avons vu précédemment que F_1 suivait un Fischer avec $(N - 1)(K + 1)$ et $NT - N(K + 1)$ degrés de liberté. Compte tenu des dimensions de notre panel et du nombre de variables explicatives ($K = 2$), on doit donc comparer la valeur de cette réalisation au seuil d'un Fischer $F(48, 544)$. Le logiciel nous donne directement la *pvalue* associée à ce test. En l'occurrence ici, cette *pvalue* est très largement inférieure au seuil de 5%, donc pour ce seuil, on rejette l'hypothèse nulle H_0 d'égalité des constantes α_i et des coefficients β_i et γ_i .

Il convient alors de tester l'hypothèse H_0^2 , d'égalité des coefficients β_i et γ_i (coefficients associés aux variables explicatives) entre les pays. Ce test est noté sous la forme **F test of $A_i, B = A_i, B_i$** dans le fichier résultat de TSP. Dans le cadre de notre échantillon, la réalisation de la statistique de Fischer associée au test H_0^2 , notée F_2 , est de 1.1845. Cette valeur est à comparer au seuil d'un Fischer avec $(N - 1)K$ et $NT - N(K + 1)$ degrés de liberté, c'est à dire ici un $F(32, 544)$. La *pvalue* indique ici que jusqu'au seuil de 25%, l'hypothèse nulle ne peut pas être rejetée. A 5%, on confirme donc ici la structure de panel, puisque l'on est en droit de supposer qu'il existe des coefficients communs pour tous les pays entre le volume des grèves et les variables explicatives que sont le chômage et l'inflation.

Reste enfin à tester l'hypothèse H_0^3 de constantes individuelles α_i . Ce test est noté sous la forme **F test of $A_i, B = A_i, B$** dans le fichier résultat de TSP. Dans le cadre de notre échantillon, la réalisation de la statistique de Fischer associée au test H_0^3 , notée F_3 , est de 9.0342. Cette valeur est à comparer au seuil d'un Fischer avec $N - 1$ et $N(T - 1) - K$ degrés de liberté, c'est à dire ici un $F(16, 576)$. La *pvalue* est très largement inférieure au seuil de 5%. Pour ce seuil, on rejette l'hypothèse nulle H_0 d'égalité des constantes α_i . Il est nécessaire d'introduire ici des effets individuels. La spécification finale de notre modèle est donc :

$$s_{i,t} = \alpha_i + \beta u_{i,t} + \gamma p_{i,t} + \varepsilon_{i,t} \quad \forall i = 1, \dots, 17 \quad (1.20)$$

Reste à présent à étudier les différentes méthodes d'estimation des modèles incluant des constantes individuelles.

2. Modèles à effets individuels

Nous allons à présent nous à des modèles de panel hétérogènes, où la seule source d'hétérogénéité provient des constantes individuelles. On suppose ainsi que les coefficients des différentes variables stochastiques explicatives sont identiques pour tous les individus du panel ($\beta_i = \beta$). On suppose en outre que ces coefficients sont des constantes déterministes. Les constantes individuelles α_i , quant à elles, diffèrent selon les individus.

Hypothèse (H1) *On suppose que les N vecteurs de paramètres β_i sont identiques, $\beta_i = \beta \in \mathbb{R}, \forall i \in [1, N]$, tandis que les constantes α_i peuvent différer selon les individus. En particulier, il existe au moins un couple $(j, i) \in [1, N]^2$ tel que $\alpha_j \neq \alpha_i$*

Sous l'hypothèse (H1), le modèle (1.1), s'écrit sous la forme suivante :

$$y_{i,t} = \alpha_i + \beta' x_{i,t} + \varepsilon_{i,t} \quad (2.1)$$

où $\alpha_i \in \mathbb{R}$ et $\beta' = (\beta_1 \beta_2 \dots \beta_K) \in \mathbb{R}^K$ est un vecteur de constantes. Les innovations $\varepsilon_{i,t}$ sont supposées être *i.i.d.* de moyenne nulle, de variance égale à $\sigma_\varepsilon^2, \forall i \in [1, N]$ et sont supposées non corrélées que ce soit dans la dimension individuelle ou dans la dimension temporelle.

Dès lors, dans ce contexte, on doit distinguer deux cas : le cas où les paramètres α_i sont des constantes déterministes (modèle à effets fixes) et le cas où les paramètres α_i sont des réalisations d'un variable aléatoire d'espérance et de variance finie (modèle à effets aléatoires). Nous allons donc successivement envisager ces deux types de modèle (sections 3 et 4). Toutefois, avant de présenter ces deux modèles, nous commencerons tout d'abord par introduire les différentes méthodes d'empilement de données de panel qui autorisent une écriture vectorielle du modèle à effet individuel.

En effet, il existe deux possibilité d'écriture vectorielle du modèle (2.1). Autrement dit, il existe deux façons d'empiler les données :

1. *Empilement par individus : pour une variable donnée, les T réalisations historiques de chaque individu sont stockées dans un vecteur colonne, et les N vecteurs colonnes ainsi obtenus sont ensuite empilés à la suite des uns des autres dans l'ordre des individus.*

2. *Empilement par dates* : pour une variable donnée, les N réalisations individuelles pour une date donnée sont stockées dans un vecteur colonne, et les T vecteurs colonnes ainsi obtenus pour toutes les dates sont ensuite empilés à la suite des uns des autres.

Il est très important de noter que les principaux logiciels d'économétrie (notamment *TSP*) optent généralement pour une méthode d'empilement par pays. Si le logiciel n'empile pas lui même les données, les séries utilisées dans le cadre des applications doivent, de façon impérative, être ordonnées sous la forme préconisée par les concepteurs du logiciel.

2.1. Empilement par pays

Considérons tout d'abord la méthode d'empilement par pays. On considère un échantillon de N individus sur T périodes, et un modèle avec K variables explicatives. On pose :

$$y_i = \begin{pmatrix} y_{i,1} \\ y_{i,2} \\ \dots \\ y_{i,T} \end{pmatrix}_{(T,1)} \quad X_i = \begin{pmatrix} x_{1,i,1} & x_{2,i,1} & \dots & x_{K,i,1} \\ x_{1,i,2} & x_{2,i,2} & \dots & x_{K,i,2} \\ \dots & \dots & \dots & \dots \\ x_{1,i,T} & x_{2,i,T} & \dots & x_{K,i,T} \end{pmatrix}_{(T,K)} \quad \varepsilon_i = \begin{pmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \\ \dots \\ \varepsilon_{i,T} \end{pmatrix}_{(T,1)}$$

On définit en outre un vecteur unitaire, noté e , tel que :

$$e_{(T,1)} = \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix}$$

Definition 2.1. Dans le cas de l'empilement par pays, pour chaque individu $\forall i \in [1, N]$, le modèle (2.1) peut s'écrire sous la forme :

$$y_i = e\alpha_i + X_i\beta + \varepsilon_i \quad \forall i = 1, \dots, N \quad (2.2)$$

C'est principalement cette expression du modèle (2.1) que l'on utilisera par la suite pour étudier les estimateurs du modèle linéaire simple. On peut toutefois écrire le modèle de façon totalement vectorielle en empilant les vecteurs y_i et les matrices X_i . Pour cela on pose :

$$Y_{(TN,1)} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{pmatrix} \quad X_{(TN,K)} = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_N \end{pmatrix} \quad \varepsilon_{(TN,1)} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_N \end{pmatrix}$$

On définit 0_T le vecteur nul de dimension $(T, 1)$.

$$\underset{(TN,N)}{\tilde{e}} = \begin{pmatrix} e & 0_T & \dots & 0_T \\ 0_T & e & \dots & 0_T \\ \dots & \dots & \dots & 0_T \\ 0_T & 0_T & \dots & e \end{pmatrix} \quad \underset{(N,1)}{\tilde{\alpha}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_N \end{pmatrix}$$

On obtient alors la représentation vectorielle suivante :

$$Y = \tilde{e}\tilde{\alpha} + X\beta + \varepsilon \quad (2.3)$$

2.1.1. Exemple

Considérons un panel de 2 pays ($N = 2$), dont les données annuelles sont disponibles sur 3 ans ($T = 3$). On cherche à estimer une fonction de production log-linéaire (Cobb-Douglass), supposée commune à ces deux pays. Soit $y_{i,t}$ le logarithme du PIB, $k_{i,t}$ le logarithme du stock de capital privé et $n_{i,t}$ le logarithme de l'emploi. On suppose que cette fonction de production s'écrit sous la forme :

$$y_{i,t} = \alpha_i + \beta_k k_{i,t} + \beta_n n_{i,t} + \varepsilon_{i,t} \quad \forall i \in [1, 2], \forall t \in [1, T]$$

où β_k et β_n désignent respectivement les élasticités (supposées communes aux deux pays) de la production par rapport au capital et à l'emploi. L'effet individuel α_i mesure ici les spécificités a-temporelles de la productivité totale des facteurs qui correspond dans cette régression, aux résidus $\varepsilon_{i,t}$. Ce modèle s'écrit sous forme vectorielle de la façon suivante :

$$\begin{pmatrix} y_{i,1} \\ y_{i,2} \\ y_{i,3} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \alpha_i + \begin{pmatrix} k_{i,1} & n_{i,1} \\ k_{i,2} & n_{i,2} \\ k_{i,3} & n_{i,3} \end{pmatrix} \begin{pmatrix} \beta_k \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \\ \varepsilon_{i,3} \end{pmatrix}$$

$$\iff y_i = e\alpha_i + X_i\beta + \varepsilon_i \quad \forall i = 1, \dots, N$$

L'écriture vectorielle complète du modèle est la suivante :

$$\begin{pmatrix} y_{1,1} \\ y_{1,2} \\ y_{1,3} \\ y_{2,1} \\ y_{2,2} \\ y_{2,3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} k_{1,1} & n_{1,1} \\ k_{1,2} & n_{1,2} \\ k_{1,3} & n_{1,3} \\ k_{2,1} & n_{2,1} \\ k_{2,2} & n_{2,2} \\ k_{2,3} & n_{2,3} \end{pmatrix} \begin{pmatrix} \beta_k \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,1} \\ \varepsilon_{1,2} \\ \varepsilon_{1,3} \\ \varepsilon_{2,1} \\ \varepsilon_{2,2} \\ \varepsilon_{2,3} \end{pmatrix}$$

$$\iff Y = \tilde{e}\tilde{\alpha} + X\beta + \varepsilon$$

2.2. Empilement par dates

De façon parallèle, il est possible de représenter de façon symétrique le modèle (2.1) en utilisant *la méthode d'empilement par dates*. Pour ce faire, il suffit de poser les définitions suivantes :

$$\underset{(N,1)}{y_t} = \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ \dots \\ y_{N,t} \end{pmatrix} \quad \underset{(N,K)}{X_t} = \begin{pmatrix} x_{1,1,t} & x_{2,1,t} & \dots & x_{K,1,t} \\ x_{1,2,t} & x_{2,2,t} & \dots & x_{K,2,t} \\ \dots & \dots & \dots & \dots \\ x_{1,N,t} & x_{2,N,t} & \dots & x_{K,N,t} \end{pmatrix} \quad \underset{(T,1)}{\varepsilon_t} = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \dots \\ \varepsilon_{N,t} \end{pmatrix}$$

Definition 2.2. Dans le cas de l'empilement par dates, pour chaque date $t \in [1, T]$, le modèle (2.1) peut s'écrire sous la forme :

$$y_t = \tilde{\alpha} + X_t \beta + \varepsilon_t \quad \forall t = 1, \dots, T \quad (2.4)$$

De la même façon, il est possible d'exprimer le modèle (2.1) sous une forme vectorielle complète. On pose :

$$\underset{(TN,1)}{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_T \end{pmatrix} \quad \underset{(TN,K)}{X} = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_T \end{pmatrix} \quad \underset{(TN,1)}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_T \end{pmatrix}$$

Soit I_N la matrice identité de dimension (N, N) . On définit la matrice suivante :

$$\underset{(TN,N)}{\tilde{e}} = \begin{pmatrix} I_N \\ I_N \\ \dots \\ I_N \end{pmatrix}$$

Il vient alors :

$$Y = \tilde{e} \tilde{\alpha} + X \beta + \varepsilon \quad (2.5)$$

2.2.1. Exemple

Reprenons l'exemple présenté précédemment. Lorsque les données sont empilées par dates, ce modèle s'écrit sous forme vectorielle de la façon suivante :

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} k_{1,t} & n_{1,t} \\ k_{2,t} & n_{2,t} \end{pmatrix} \begin{pmatrix} \beta_k \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}$$

$$\iff y_t = \tilde{\alpha} + X_t \beta + \varepsilon_t \quad \forall t = 1, \dots, T$$

L'écriture vectorielle complète du modèle est la suivante :

$$\begin{pmatrix} y_{1,1} \\ y_{2,1} \\ y_{1,2} \\ y_{2,2} \\ y_{1,3} \\ y_{2,3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} k_{1,1} & n_{1,1} \\ k_{2,1} & n_{2,1} \\ k_{1,2} & n_{1,2} \\ k_{2,2} & n_{2,2} \\ k_{1,3} & n_{1,3} \\ k_{2,3} & n_{2,3} \end{pmatrix} \begin{pmatrix} \beta_k \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,1} \\ \varepsilon_{2,1} \\ \varepsilon_{1,2} \\ \varepsilon_{2,2} \\ \varepsilon_{1,3} \\ \varepsilon_{2,3} \end{pmatrix}$$

$$\Longleftrightarrow Y = \widetilde{e}\widetilde{\alpha} + X\beta + \varepsilon$$

Maintenant que nous avons présenté l'écriture vectorielle des modèles de panel à effets individuels, nous allons à présent étudier les propriétés des estimateurs des paramètres de ces modèles suivant la spécification des effets individuels. Nous distinguerons pour cela deux cas : le cas des effets fixes et le cas des effets aléatoires.

3. Modèle à effets fixes

On fait maintenant l'hypothèse que *les effets individuels* α_i *sont représentés par des constantes* (d'où l'appellation modèle à *effets fixes*). Nous allons déterminer la forme générale des estimateurs des paramètres α_i et β dans ce modèle à effets fixes. On considère donc le modèle (1.1) sous l'hypothèse (H1) :

$$y_{i,t} = \alpha_i + \beta' x_{i,t} + \varepsilon_{i,t} \quad \forall i \in [1, N], \forall t \in [1, T] \quad (3.1)$$

où $\alpha_i \in \mathbb{R}$, $\beta' = (\beta_1 \beta_2 \dots \beta_K) \in \mathbb{R}^K$. Tous les paramètres du modèle sont des constantes et l'on suppose pour simplifier qu'il n'existe pas d'effet temporel. On définit le vecteur ε_i tel que :

$$\varepsilon_i = \begin{pmatrix} \varepsilon_{i,1} & \varepsilon_{i,2} & \dots & \varepsilon_{i,T} \end{pmatrix}'_{(T,1)}$$

Pour étudier les propriétés des estimateurs du modèle à effets fixes, nous allons faire une hypothèse supplémentaire sur la nature du processus des résidus $\varepsilon_{i,t}$. Cette hypothèse constitue tout simplement *la généralisation dans la dimension de panel de la définition d'un bruit blanc*.

Hypothèse (H2) *On suppose que les résidus $\varepsilon_{i,t}$ sont i.i.d. et satisfont les conditions suivantes, $\forall i \in [1, N], \forall t \in [1, T]$*

- $E(\varepsilon_{i,t}) = 0$
- $E(\varepsilon_{i,t} \varepsilon_{i,s}) = \begin{cases} \sigma_\varepsilon^2 & t = s \\ 0 & \forall t \neq s \end{cases}$, ce qui implique que $E(\varepsilon_i \varepsilon_i') = \sigma_\varepsilon^2 I_T$ où I_t désigne la matrice identité (T, T) .
- $E(\varepsilon_{i,t} \varepsilon_{j,s}) = 0, \forall j \neq i, \forall (t, s)$

La première condition impose tout simplement que l'espérance des résidus du modèle (3.1) soit nulle. La seconde condition, standard en économétrie des séries temporelles, impose que le processus $\varepsilon_{i,t}$ soit un processus "sans mémoire" (dans la dimension temporelle). Pour chaque individu, il n'existe ainsi aucune corrélation entre le niveau présent du processus $\varepsilon_{i,t}$ et les réalisations passées. Seule la variance du processus $\varepsilon_{i,t}$ est non nulle. L'introduction d'une dimension individuelle, nous oblige ici à définir une seconde contrainte qui est que tous les processus individuels $\varepsilon_{i,t}$ ont la même variance σ_ε^2 quel que soit l'individu considéré. Autrement dit, la matrice de variance covariance du processus ε_i est proportionnelle, à un scalaire près, à la matrice identité. Enfin, la troisième condition stipule qu'il n'existe aucune corrélation entre les processus d'innovation pour deux individus distincts et cela quelle que soit la date considérée.

3.1. Estimateur Within ou LSDV

Considérons l'écriture vectorielle du modèle obtenue par un empilement des données par pays.

$$y_i = \underset{(T,1)}{e} \underset{(T,1)}{\alpha_i} + \underset{(T,K)(K,1)}{X_i} \underset{(T,1)}{\beta} + \underset{(T,1)}{\varepsilon_i} \quad \forall i = 1, \dots, N \quad (3.2)$$

L'estimateur des Moindres Carrés Ordinaires (*MCO*) des paramètres α_i et β dans le modèle à effets fixes est appelé estimateur *Within*, ou *estimateur à effets fixes* ou estimateur *LSDV* (*Least Square Dummy Variable*). Comme nous l'avons vu, le terme *Within* s'explique par le fait cet estimateur tient compte de la variance intra groupe de la variable endogène. La troisième appellation *LSDV* tient au fait que cet estimateur conduit à introduire des variables dummies, qui dans la spécification (3.2) correspondent aux vecteurs colonnes de e .

On peut démontrer que sous l'hypothèse (*H2*), l'estimateur des Moindres Carrés Ordinaires (*MCO*) des paramètres α_i et β du modèle (3.2) est le meilleur estimateur linéaire sans biais (*BLUE*⁷). Si l'on suppose que σ_ε est connu, cet estimateur est obtenu en minimisant la variance des résidus empiriques, notée S , par rapport aux seuls paramètres α_i et β .

$$\min_{\{\alpha_i, \beta\}_{i=1}^N} S = \sum_{i=1}^N \varepsilon_i' \varepsilon_i = \sum_{i=1}^N (y_i - e\alpha_i - X_i\beta)' (y_i - e\alpha_i - X_i\beta)$$

La première condition nécessaire de ce programme, nous donne immédiatement l'expression de l'estimateur de la constante α_i , $\forall i \in [1, N]$:

$$\hat{\alpha}_i = \bar{y}_i - \hat{\beta}_{LSDV}' \bar{x}_i \quad (3.3)$$

où les termes \bar{y}_i et \bar{x}_i désignent les moyennes individuelles des variables endogènes et exogènes.

$$\bar{y}_i = \underset{(1,1)}{\frac{1}{T}} \sum_{t=1}^T y_{i,t} \quad \bar{x}_i = \underset{(K,1)}{\frac{1}{T}} \sum_{t=1}^T x_{i,t} \quad \forall i \in [1, N] \quad (3.4)$$

Ces moyennes sont donc calculées de façon traditionnelle dans la dimension temporelle, et cela pour chaque individu de l'échantillon. C'est pourquoi elles sont indicées en i .

A partir de la seconde condition nécessaire du programme de minimisation, on montre que l'estimateur du paramètre vectoriel β est donné par la relation suivante :

$$\hat{\beta}_{LSDV} = \left[\sum_{i=1}^N \sum_{t=1}^T (x_{i,t} - \bar{x}_i) (x_{i,t} - \bar{x}_i)' \right]^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T (x_{i,t} - \bar{x}_i) (y_{i,t} - \bar{y}_i) \right] \quad (3.5)$$

⁷ *Best Linear Unbiased Estimator*

Proposition 3.1. *L'estimateur *Within* ou *LSDV* de β , obtenu dans le modèle à effets individuels fixes (3.2) est identique à l'estimateur des MCO obtenu à partir d'un modèle transformé où les variables expliquées et explicatives sont centrées sur leur moyennes individuelles respectives :*

$$(y_{i,t} - \bar{y}_i) = \hat{\beta}' (x_{i,t} - \bar{x}_i) + \varepsilon_{i,t} \quad \forall i \in [1, N], \forall t \in [1, T] \quad (3.6)$$

Les réalisations des estimateurs des constantes α_i sont alors déduits de la relation :

$$\hat{\alpha}_i = \bar{y}_i - \hat{\beta}' \bar{x}_i$$

En effet, on voit d'après l'équation (3.5), que l'estimateur *Within* des coefficients de β peut être obtenu en centrant les différentes (variables endogène et exogènes) sur les moyennes individuelles respectives. Ainsi, on peut obtenir le même estimateur en utilisant le modèle transformé suivant $\forall i \in [1, N], \forall t \in [1, T]$:

$$\tilde{y}_{i,t} = \beta' \tilde{x}_{i,t} + \varepsilon_{i,t}$$

avec $\tilde{y}_{i,t} = (y_{i,t} - \bar{y}_i)$ et $\tilde{x}_{i,t} = (x_{i,t} - \bar{x}_i)$. En effet, dans ce cas on obtient :

$$\begin{aligned} \hat{\beta} &= \left[\sum_{i=1}^N \sum_{t=1}^T \tilde{x}_{i,t} \tilde{x}_{i,t}' \right]^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T \tilde{x}_{i,t} \tilde{y}_{i,t} \right] \\ &= \left[\sum_{i=1}^N \sum_{t=1}^T (x_{i,t} - \bar{x}_i) (x_{i,t} - \bar{x}_i)' \right]^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T (x_{i,t} - \bar{x}_i) (y_{i,t} - \bar{y}_i) \right] \end{aligned}$$

On reconnaît ici l'expression générale de l'estimateur *Within* $\hat{\beta}_{LSDV}$. Reste alors à obtenir les réalisations des estimateurs des constantes individuelles selon la formule :

$$\begin{aligned} \hat{\alpha}_1 &= \bar{y}_1 - \hat{\beta}' \bar{x}_1 \\ \hat{\alpha}_2 &= \bar{y}_2 - \hat{\beta}' \bar{x}_2 \\ &\dots \\ \hat{\alpha}_N &= \bar{y}_N - \hat{\beta}' \bar{x}_N \end{aligned}$$

Le fait qu'il soit équivalent d'estimer les paramètres du modèle (3.2) directement à partir de cette spécification incluant des dummies individuelles ou à partir d'un modèle transformé où les variables sont centrées sur leurs moyennes individuelles respectives est particulièrement important. Cela illustre la notion de variance intra-classes qui est fondamentale dans la construction de l'estimateur *Within*. Nous allons à présent, proposer une illustration de cette propriété de l'estimateur *Within*.

3.2. Application

Considérons à nouveau l'exemple du modèle de grèves dans le secteur industriel. Nous avons établi précédemment que le modèle de panel comportait des effets individuels et pouvait s'écrire sous la forme:

$$s_{i,t} = \alpha_i + \beta u_{i,t} + \gamma p_{i,t} + \varepsilon_{i,t} \quad \forall i = 1, \dots, 17 \quad (3.7)$$

où $s_{i,t}$ désigne le nombre de jours chômés pour cause de grève, pour 1000 salariés du secteur industriel, $u_{i,t}$ désigne le taux de chômage de l'économie et $p_{i,t}$ le taux d'inflation. On suppose que l'on peut spécifier les effets individuels α_i sous la forme d'effets fixes. Dans le programme ci-dessous, on cherche à comparer l'estimateur Within et l'estimateur des MCO du modèle transformé :

$$(s_{i,t} - \bar{s}_i) = \beta (u_{i,t} - \bar{u}_i) + \gamma (p_{i,t} - \bar{p}_i) + \varepsilon_{i,t} \quad (3.8)$$

L'instruction **panel** de TSP avec les options **notot**, **nobet** et **novar** permet d'obtenir directement et uniquement les résultats de l'estimation *Within*. Pour construire les estimateurs du modèle transformé, on doit au préalable centrer les variables sur leur moyennes individuelles respectives. Pour réaliser cette opération on construit une boucle sur les $N = 17$ individus, avec un compteur j et à chaque étape on sélectionne un pays dont le code pays (variable i) est égal à la valeur du compteur j (instruction **select i = j**). Pour le sous échantillon sélectionné, l'instruction **msd** [*Nom de la Variable*] permet alors de calculer différents moments de la variable considérée sans les afficher (si l'on ajoute l'option **noprint**). En particulier, la moyenne empirique est stockée dans la variable réservée **@mean**. Dès lors, pour chaque individu, on construit de nouvelles variables en centrant sur les moyennes individuelles correspondantes. Enfin, il ne reste plus qu'à effectuer la régression par les MCO sur le modèle transformé (instruction **ls** ou **ols**)

```
load(file='strikes.wks');
?---- Estimation Within ----
panel (id=i,time=t,notot,nobet,novar) srt u p;
?---- Centrer les Données ---
do j=1 to 17;
    select (i=j);
    msd(noprint) srt;
    srtc=srt-@mean;
    msd(noprint) u;
    uc=u-@mean;
    msd(noprint) p;
```

```

pc=p-@mean;
enddo;
?---- Estimation Modèle Transformé----
ls srtc uc pc;
?---- Affichage des Effets Fixes (Instruction Panel) ----
print @fixed;

```

Les résultats de ce programme sont reproduits sur la figure (3.1). On vérifie que l'estimateur *Within* est totalement équivalent à l'estimateur des *MCO* appliqué sur un modèle transformé où les variables endogènes et exogènes ont été centrées sur leurs moyennes individuelles respectives. Dans les deux cas, les réalisations des estimateurs des coefficients des deux variables explicatives (respectivement $u_{i,t}$ et $p_{i,t}$) sont strictement identiques (respectivement -21.5968 et 16.2729).

Dans le cas où l'on utilise la commande **panel** de TSP, on peut afficher les estimations des effets fixes en utilisant la commande **print @fixed**. Dans le cas où l'on utilise le modèle transformé, il convient alors de reconstruire les estimateurs des effets fixes selon la formule (3.3) en utilisant les moyennes individuelles des variables (**cf.** programme ci-dessus) et les réalisations des estimateurs des coefficients γ et β . Essayons à présent de commenter les estimations ainsi obtenues. Il s'avère ainsi que le pays 9 a, de façon structurelle, le plus de journées chômées pour cause de grève. Pour ce pays, différents facteurs historiques (législation sur le droit de grève) ou sociologique (représentation des syndicats par exemple) expliquent que toutes choses égales par ailleurs (chômage et inflation), la fréquence des grève soit particulièrement élevée. A l'inverse, les pays 15 et 2 ont des effets fixes négatifs. Pour un même niveau de chômage et d'inflation, ces deux pays auront le nombre de jours chômés le plus faible de tout notre échantillon. Il faut bien comprendre que *les estimations des effets individuels ne peuvent s'analyser qu'en niveau relatif (c'est à dire en comparant les différentes réalisations individuelles) et non en niveau absolu.*

Figure 3.1: Comparaison des Estimateurs Whithin et OLS sur Modèle Transformé

```

PANEL DATA ESTIMATION
=====

Balanced data:  NI=   17, T=   35, NOB=   595

WITHIN (fixed effects) Estimates:

Dependent variable: SRT

Sum of squared residuals = .146958E+09          R-squared = .242875
Variance of residuals = 255136.                Adjusted R-squared = .219215
Std. error of regression = 505.110

Variable      Estimated      Standard
Coefficient      Error      t-statistic
U             -21.5968      9.19158     -2.34963
P              16.2729      4.75658      3.42113

Equation      1
=====

Method of estimation = Ordinary Least Squares

Dependent variable: SRTC
Number of observations: 595

Mean of dependent variable = -.159906E-07
Std. dev. of dependent var. = 503.791
Sum of squared residuals = .146958E+09
Variance of residuals = 247822.
Std. error of regression = 497.817
R-squared = .025220
Adjusted R-squared = .023576
Durbin-Watson statistic = 1.98516
F-statistic (zero slopes) = 15.3421
Schwarz Bayes. Info. Crit. = 12.4386
Log of likelihood function = -4538.36

Variable      Estimated      Standard
Coefficient      Error      t-statistic
UC             -21.5968      9.05887     -2.38405
PC              16.2729      4.68791      3.47125

      1          2          3          4
@FIXED      355.49000      -2.58225      801.98611      211.05529

      5          6          7          8
@FIXED      387.83512      339.06634      57.54147      582.61959

      9         10         11         12
@FIXED      1012.66380      114.22273      27.49515      164.23965

      13        14        15        16
@FIXED      14.43655      14.93684      -44.85353      284.45232

      17
@FIXED      495.87113

```

3.3. Ecriture vectorielle

Une façon équivalente d'obtenir l'estimateur *Within* et de montrer l'équivalence avec le modèle transformé où les variables sont centrées, consiste à travailler directement avec l'écriture vectorielle (3.2). Par la suite, nous adopterons souvent cette notation vectorielle qui permet d'alléger particulièrement les notations du modèle et des estimateurs correspondants. On considère donc le modèle vectoriel suivant :

$$y_i = e\alpha_i + X_i\beta + \varepsilon_i \quad \forall i = 1, \dots, N \quad (3.9)$$

Nous allons à présent introduire un opérateur matriciel qui permet de centrer les variables sur leurs moyennes individuelles respectives.

Definition 3.2. Le modèle (3.9) à effet individuels fixes peut s'écrire sous la forme vectorielle suivante $\forall i = 1, \dots, N$:

$$Qy_i = QX_i\beta + Q\varepsilon_i$$

où la matrice Q de dimension (T, T) est telle que :

$$Q = I_T - \frac{1}{T}ee' \quad (3.10)$$

L'estimateur *Within* du vecteur β est alors défini de la façon suivante :

$$\hat{\beta}_{LSDV} = \left[\sum_{i=1}^N X_i' Q X_i \right]^{-1} \left[\sum_{i=1}^N X_i' Q y_i \right] \quad (3.11)$$

En effet, on peut montrer que les processus transformés Qy_i et QX_i correspondent aux variables centrées sur leurs moyennes respectives. En effet, on montre que :

$$Qy_i = \left(I_T - \frac{1}{T}ee' \right) y_i = y_i - e \left(\frac{1}{T}e'y_i \right) = \begin{pmatrix} y_{i,1} \\ y_{i,2} \\ \dots \\ y_{i,T} \end{pmatrix} - \left(\frac{1}{T} \sum_{t=1}^T y_{i,t} \right) \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix}$$

De la même façon, on peut montrer que :

$$QX_i = X_i - \frac{1}{T}ee'X_i$$

$$QX_i = \begin{pmatrix} x_{1,i,1} & x_{2,i,1} & \dots & x_{K,i,1} \\ x_{1,i,2} & x_{2,i,2} & \dots & x_{K,i,2} \\ \dots & \dots & \dots & \dots \\ x_{1,i,T} & x_{2,i,T} & \dots & x_{K,i,T} \end{pmatrix} - \frac{1}{T} \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} \left(\sum_{t=1}^T x_{1,i,t} \quad \sum_{t=1}^T x_{2,i,t} \quad \dots \quad \sum_{t=1}^T x_{K,i,t} \right)$$

Dès lors, afin d'éliminer les effets individuels de l'équation (3.9), il suffit de prémultiplier les membres de cette équation par Q . Ainsi $\forall i = 1, \dots, N$

$$Qy_i = Qe\alpha_i + QX_i\beta + Q\varepsilon_i$$

$$\Longleftrightarrow Qy_i = QX_i\beta + Q\varepsilon_i$$

En appliquant les *MCO* à cette expression on obtient donc une écriture équivalente de $\hat{\beta}_{LSDV}$:

$$\hat{\beta}_{LSDV} = \left[\sum_{i=1}^N X_i' Q X_i \right]^{-1} \left[\sum_{i=1}^N X_i' Q y_i \right] \quad (3.12)$$

4. Modèle à effets aléatoires

Dans la pratique standard de l'analyse économétrique, on suppose qu'il existe un grand nombre de facteurs qui peuvent affecter la valeur de la variable expliquée et qui pourtant ne sont pas introduits explicitement sous la forme de variables explicatives. Ces facteurs sont alors approximés par la structure des résidus. Le problème se pose de la façon similaire en économétrie de panel. La seule différence tient au fait que *trois types de facteurs omis peuvent être envisagés*. Il y a tout d'abord les facteurs qui affectent la variable endogène différemment suivant la période et l'individu considéré. Il peut en outre exister des facteurs qui affectent de façon identique l'ensemble des individus, mais dont l'influence dépend de la période considérée (effets temporel). Enfin, d'autres facteurs peuvent au contraire refléter des différences entre les individus de type structurelles, c'est à dire indépendantes du temps (effets individuel).

Dès lors le résidu, noté $\varepsilon_{i,t}$, d'un modèle de panel peut être décomposé en trois principales composantes de la façon suivante (Hsiao 1986) $\forall i \in [1, N], \forall t \in [1, T]$:

$$\varepsilon_{i,t} = \alpha_i + \lambda_t + v_{i,t} \quad (4.1)$$

Les variables α_i désignent ici les *effets individuels* qui représentent l'ensemble des spécificités structurelles ou a-temporelles de la variable endogène, qui diffèrent selon les individus. *On suppose ici que ces effets sont aléatoires*. Les variables aléatoires λ_t représentent quant à elle les *effets temporels* strictement identiques pour tous les individus. Enfin, le processus stochastique $v_{i,t}$ désigne la composante du résidu total $\varepsilon_{i,t}$ orthogonale aux effets individuels et aux effets temporels. Généralement, on est conduit à faire un certain nombre d'hypothèses techniques sur cette structure de résidus.

Hypothèses (H3) *On suppose que les résidus $\varepsilon_{i,t} = \alpha_i + \lambda_t + v_{i,t}$ sont i.i.d. et satisfont les conditions suivantes, $\forall i \in [1, N], \forall t \in [1, T]$*

- $E(\alpha_i) = E(\lambda_t) = E(v_{i,t}) = 0$
- $E(\alpha_i \lambda_t) = E(\lambda_t v_{i,t}) = E(\alpha_i v_{i,t}) = 0$
- $E(\alpha_i \alpha_j) = \begin{cases} \sigma_\alpha^2 & i = j \\ 0 & \forall i \neq j \end{cases}$
- $E(\lambda_t \lambda_s) = \begin{cases} \sigma_\lambda^2 & t = s \\ 0 & \forall t \neq s \end{cases}$
- $E(v_{i,t} v_{j,s}) = \begin{cases} \sigma_v^2 & t = s, i = j \\ 0 & \forall t \neq s, \forall i \neq j \end{cases}$
- $E(\alpha_i x'_{i,t}) = E(\lambda_t x'_{i,t}) = E(v_{i,t} x'_{i,t}) = 0$

Sous ces hypothèses, la variance de la variable endogène $y_{i,t}$ conditionnellement aux variables explicatives $x_{i,t}$ est alors égale à $\sigma_y^2 = \sigma_\alpha^2 + \sigma_\lambda^2 + \sigma_v^2$. Les variances σ_α^2 , σ_λ^2 et σ_v^2 correspondent aux différentes composantes de la variance totale. C'est pourquoi, le modèle à effets aléatoires est aussi appelé *modèle à erreurs composés* (*error component model*).

Dans une première étape, nous allons présenter le modèle simple à effets individuels aléatoires. Nous supposons ainsi, pour simplifier l'analyse, qu'il n'existe pas d'effet temporel. Nous proposerons alors une écriture vectorielle du modèle à effets aléatoires. Dans une seconde étape nous étudierons les estimateurs des différents coefficients de ce modèle.

4.1. Modèle à variance composée

On se limite au cas où il n'existe pas d'effet temporel ($\lambda_t = 0, \forall t$). On considère donc le modèle suivant $\forall i \in [1, N], \forall t \in [1, T]$:

$$y_{i,t} = \mu + \beta' x_{i,t} + \varepsilon_{i,t} \quad (4.2)$$

$$\varepsilon_{i,t} = \alpha_i + v_{i,t} \quad (4.3)$$

où $\beta' = (\beta_1 \beta_2 \dots \beta_K)$ est un vecteur de constantes et le processus $\{\varepsilon_{i,t}\}$ satisfait les hypothèses (H_3). La constante μ désigne ici l'espérance inconditionnelle du processus $y_{i,t}$ puisque on suppose que $E(\alpha_i) = 0, \forall i \in [1, N]$.

Proposition 4.1. *Le modèle à effets individuels aléatoires s'écrit vectoriellement sous la forme :*

$$\underset{(T,1)}{y_i} = \underset{(T,K+1)(K+1,1)}{\tilde{X}_i} \underset{(T,1)}{\gamma} + \underset{(T,1)}{\varepsilon_i} \quad \forall i = 1, \dots, N \quad (4.4)$$

avec $\varepsilon_i = \alpha_i e + v_i$, $\tilde{X}_i = (e, X_i)$ et $\gamma' = (\mu, \beta')$. Sous les hypothèses (H_3), la matrice de variance covariance de ε_i , notée V , est alors définie par :

$$V = E(\varepsilon_i \varepsilon_i') = E[(\alpha_i e + v_i)(\alpha_i e + v_i)'] = \sigma_\alpha^2 e e' + \sigma_v^2 I_T \quad (4.5)$$

L'inverse de cette matrice de variance covariance est définie comme :

$$V^{-1} = \frac{1}{\sigma_v^2} \left[I_T - \left(\frac{\sigma_\alpha^2}{\sigma_v^2 + T \sigma_\alpha^2} \right) e e' \right] \quad (4.6)$$

Donnons ici une démonstration possible de ce résultat⁸. On pose $\tilde{V} = V^{-1}$. Par définition on a $\tilde{V} V = V \tilde{V} = I_T$. On en déduit que :

$$V \tilde{V} = \sigma_\alpha^2 e e' \tilde{V} + \sigma_v^2 \tilde{V} = I_T \iff \sigma_v^2 \tilde{V} = I_T - \sigma_\alpha^2 e e' \tilde{V}$$

⁸Pour une démonstration plus élégante voir Graybill (1969), Nerlove (1971), Wallace et Hussain (1969)

Reste alors à déterminer l'expression de $ee'\tilde{V}$. Raisonnons par une méthode de coefficient indéterminé. On suppose que $ee'\tilde{V}$ peut s'écrire sous la forme $\theta ee'$ où $\theta \in \mathbb{R}^+$. Considérons l'expression de $ee'\tilde{V}V$ en substituant $ee'\tilde{V}$ par son expression, il vient :

$$\begin{aligned} ee'\tilde{V}V &= \theta ee'V \\ &= \theta ee' (\sigma_\alpha^2 ee' + \sigma_v^2 I_T) \\ &= \theta \sigma_\alpha^2 [e (e'e) e'] + \theta \sigma_v^2 (ee') \end{aligned}$$

Par définition de \tilde{V} , on a $ee'\tilde{V}V = ee'$. Sachant que $e (e'e) e' = T ee'$, on obtient alors la relation suivante :

$$\begin{aligned} ee'\tilde{V}V &= ee' \\ \iff \theta ee' (\sigma_\alpha^2 T + \sigma_v^2) &= ee' \end{aligned}$$

On en déduit immédiatement la valeur du paramètre θ :

$$\theta = \frac{1}{\sigma_\alpha^2 T + \sigma_v^2} \in \mathbb{R}^+$$

On obtient ainsi :

$$ee'\tilde{V} = \left(\frac{1}{\sigma_\alpha^2 T + \sigma_v^2} \right) ee'$$

Dès lors, en reprenant l'expression de \tilde{V} , il vient :

$$\sigma_v^2 \tilde{V} = I_T - \left(\frac{\sigma_\alpha^2}{\sigma_\alpha^2 T + \sigma_v^2} \right) ee'$$

Ainsi, l'inverse de la matrice de variance covariance des ε_i s'écrit sous la forme :

$$V^{-1} = \frac{1}{\sigma_v^2} \left[I_T - \left(\frac{\sigma_\alpha^2}{\sigma_v^2 + T\sigma_\alpha^2} \right) ee' \right]$$

L'écriture vectorielle du modèle à effets aléatoires de la proposition (4.1) nous sera particulièrement dans la section suivante pour construire les estimateurs des différents paramètres de ce modèle.

4.2. Estimateurs du modèle à effets aléatoires

Nous allons à présent nous intéresser aux propriétés des estimateurs du modèle aléatoire. Pour bien comprendre ces propriétés, il est nécessaire de bien appréhender les conséquences de la structure des résidus.

Remark 1. Dans le modèle de la proposition (4.1), le fait que les effets individuels α_i constituent l'une des composantes des résidus ε_i du modèle (4.4), induit une corrélation entre le niveau de ces résidus lorsque l'on considère un individu donné. En revanche, sous (H_3) , il n'existe pas de corrélation de ces résidus dans la dimension inter individuelle.

Quelles sont alors les conséquences de cette structure des résidus ? Si l'on raisonne en termes de biais d'estimation, il n'y en a aucune dès lors que l'on centre les variables du modèle sur leur moyenne respective. En effet, dans ce cas la composante individuelle des résidus disparaît et les corrélations inter-individuels des résidus du modèle transformé disparaissent par la même occasion. Pour le vecteur β , on retrouve alors la forme générale de l'estimateur du modèle à effets fixes. Pour preuve, appliquons l'opérateur Q au modèle à effet aléatoire :

$$Qy_i = Qe\mu + QX_i\beta + Qe\alpha_i + Qv_i$$

Or, on sait par définition de l'application que $Qe = (I_T - T^{-1}ee')e = 0$. Dès lors, on obtient le modèle $\forall i \in [1, N]$:

$$Qy_i = Qe\mu + QX_i\beta + Qv_i \quad (4.7)$$

Dans ce modèle, sous (H_3) , l'estimateur $\hat{\beta}_{LSDV}$ du vecteur β est *non biaisé et convergent*⁹. L'estimateur de la constante μ est alors construit sous la forme :

$$\hat{\mu} = \bar{y} - \hat{\beta}'_{LSDV} \bar{x} \quad (4.8)$$

avec :

$$\bar{y}_{(1,1)} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T y_{i,t} \quad \bar{x}_{(K,1)} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T x_{i,t} \quad (4.9)$$

Ainsi, il n'existe pas de problème de biais lorsque l'on envisage les techniques d'estimation utilisées dans les modèles à effets fixes, c'est à dire lorsque l'on centre les variables sur leurs moyennes individuelles respectives. Toutefois, l'estimateur $\hat{\beta}_{LSDV}$ ainsi obtenu n'est pas un estimateur à variance minimale : c'est pas l'estimateur *BLUE*.

Proposition 4.2. Dans un modèle à effets aléatoires, l'estimateur *Within* $\hat{\beta}_{LSDV}$, obtenu sous l'hypothèse d'effets fixes, est un estimateur sans biais et convergent du vecteur de paramètres β . Toutefois, ce n'est pas l'estimateur *BLUE*. Un estimateur *BLUE* est alors donné par l'estimateur de Moindres Carrés Généralisés (MCG).

⁹ On a en effet :

$$\hat{\beta}_{LSDV} \xrightarrow[NT \rightarrow \infty]{p} \beta$$

où l'indice p désigne la convergence en probabilité (voir section précédente pour une démonstration).

Nous ne donnerons pas ici de démonstration de cette propriété. Si l'on résume, *quelle que soit la nature des effets individuels (fixes ou aléatoires), l'estimateur Within est un estimateur sans biais et convergent. Toutefois, cette estimateur n'est pas un estimateur à variance minimale lorsque les effets individuels sont aléatoires. Un estimateur BLUE possible du vecteur β est alors donné par l'estimateur de Moindres Carrés Généralisés (MCG)*. Cette propriété sera particulièrement intéressante lorsqu'il sera nécessaire de discriminer les deux modèles et de *construire les tests de spécification* appropriés (test d'Hausman 1978, en particulier).

4.3. Estimateur des Moindres Carrés Généralisés

Nous avons vu que dans un modèle à effets aléatoires, un estimateur *BLUE* peut être construit à partir de l'estimateur des Moindres Carrés Généralisés (MCG). Considérons le modèle vectoriel de la proposition (4.1) :

$$y_i = \tilde{X}_i \gamma + \varepsilon_i \quad \forall i = 1, \dots, N \quad (4.10)$$

avec $\varepsilon_i = \alpha_i e + v_i$, $\tilde{X}_i = (e, X_i)$ et $\gamma' = (\mu, \beta')$. Soit V la matrice de variance covariance du vecteur des résidus ε_i : $V = E(\varepsilon_i \varepsilon_i')$. On suppose, dans un premier temps, que V est connue. Le problème est que, du fait de la structure du modèle, *la matrice V n'est pas diagonale en raison de la présence de corrélations intra-individuelles des résidus*. Tout comme en séries temporelles, c'est pourquoi on applique alors les MCG.

Definition 4.3. *Si la matrice V est connue, l'estimateur des MCG du vecteur γ , noté $\hat{\gamma}_{MCG}$, est alors défini par la relation :*

$$\hat{\gamma}_{MCG} = \left[\sum_{i=1}^N \tilde{X}_i' V^{-1} \tilde{X}_i \right]^{-1} \left[\sum_{i=1}^N \tilde{X}_i' V^{-1} y_i \right] \quad (4.11)$$

Cette définition de l'estimateur $\hat{\gamma}_{MCG}$ correspond à la définition générale (identique à celle utilisée en série temporelle, cf. Bourbonnais 2000) des Moindres Carrés Généralisés. *Bien entendu, lorsque la matrice V n'est pas connue, l'estimateur $\hat{\gamma}_{MCG}$ est obtenu en deux étapes*. La première étape consiste à appliquer l'estimateur *Within* pour obtenir une première estimation sans biais et convergente des paramètres β et μ . A partir de ces estimations, on construit les séries de résidus individuels ε_i et l'on construit un estimateur \hat{V} de la matrice de variance covariance V . La seconde étape consiste alors à *appliquer l'estimateur des MCG selon la formule (4.11) en utilisant l'estimateur \hat{V} de V* .

Il existe une autre façon d'obtenir l'estimateur des MCG. En effet, dans la pratique, *la procédure en deux étapes, utilisée dans les principaux logiciels d'économétrie, n'est*

pas exactement similaire à celle que nous venons de décrire¹⁰. L'estimateur des MCG est construit à partir d'une relation existant entre cet estimateur et les estimateurs *Between* et *Within* (Maddala 1971).

Sous les hypothèses (H_3), nous avons vu (proposition 4.1) que l'inverse de la matrice de variance covariance V pouvait s'écrire sous la forme :

$$V^{-1} = \frac{1}{\sigma_v^2} \left[I_T - \left(\frac{\sigma_\alpha^2}{\sigma_v^2 + T\sigma_\alpha^2} \right) ee' \right]$$

Definition 4.4. On pose que la matrice V^{-1} peut être exprimée sous la forme alternative suivante :

$$V^{-1} = \frac{1}{\sigma_v^2} \left[Q + \psi \frac{1}{T} ee' \right] \quad (4.12)$$

avec $Q = (I_T - ee'/T)$ et où le paramètre ψ est défini par la relation :

$$\psi = \left(\frac{\sigma_v^2}{\sigma_v^2 + T\sigma_\alpha^2} \right) \quad (4.13)$$

Reprenons alors l'expression de l'estimateur des MCG (4.11) et substituons cette nouvelle expression de V^{-1} . On obtient alors :

$$\begin{aligned} \hat{\gamma}_{MCG} &= \left[\sum_{i=1}^N \tilde{X}_i' \left(Q + \psi \frac{1}{T} ee' \right) \tilde{X}_i \right]^{-1} \left[\sum_{i=1}^N \tilde{X}_i' \left(Q + \psi \frac{1}{T} ee' \right) y_i \right] \\ \iff \hat{\gamma}_{MCG} &= \left[\sum_{i=1}^N \tilde{X}_i' Q \tilde{X}_i + \psi \frac{1}{T} \sum_{i=1}^N \tilde{X}_i' ee' \tilde{X}_i \right]^{-1} \left[\sum_{i=1}^N \tilde{X}_i' Q y_i + \psi \frac{1}{T} \sum_{i=1}^N \tilde{X}_i' ee' y_i \right] \end{aligned}$$

Sachant que $\tilde{X}_i = (e \ X_i)$ et $\gamma' = (\mu \ \beta')$, on peut montrer que ces différents termes peuvent se réécrire sous la forme suivante :

$$\begin{bmatrix} \hat{\mu}_{MCG} \\ \hat{\beta}_{MCG} \end{bmatrix} = \begin{bmatrix} \psi NT & \psi T \sum_{i=1}^N \bar{x}_i' \\ \psi T \sum_{i=1}^N \bar{x}_i & \sum_{i=1}^N X_i' Q X_i + \psi T \sum_{i=1}^N \bar{x}_i \bar{x}_i' \end{bmatrix}^{-1} \begin{bmatrix} \psi NT \bar{y} \\ \sum_{i=1}^N X_i' Q y_i + \psi T \sum_{i=1}^N \bar{x}_i \bar{y}_i \end{bmatrix}$$

A partir de la formule de l'inverse d'une matrice partitionnée, on montre que :

$$\hat{\beta}_{MCG} = \left[\frac{1}{T} \sum_{i=1}^N X_i' Q X_i + \psi \sum_{i=1}^N (\bar{x}_i - \bar{x}) (\bar{x}_i - \bar{x})' \right]^{-1} \left[\frac{1}{T} \sum_{i=1}^N X_i' Q y_i + \psi \sum_{i=1}^N (\bar{x}_i - \bar{x}) (\bar{y}_i - \bar{y})' \right] \quad (4.14)$$

Introduisons à présent l'estimateur *Between*, qui est construit à partir des N moyennes individuelles des variables endogènes et exogènes, centrées sur la moyenne totale (du fait de la présence d'une constante).

¹⁰Mais sur le fond, les deux procédures sont strictement identiques. Seules leurs mises en oeuvre pratique et l'expression utilisée pour l'estimateur des MCG diffèrent.

Definition 4.5. *L'estimateur inter-classe Between, est l'estimateur des Moindres Carrés Ordinaires obtenu dans la spécification $\forall i \in [1, N]$:*

$$\bar{y}_i = c + \beta' \bar{x}_i + \varepsilon_i \quad (4.15)$$

Soit $\hat{\beta}_{BE}$ l'estimateur Between du vecteur β :

$$\hat{\beta}_{BE} = \left[\sum_{i=1}^N (\bar{x}_i - \bar{x}) (\bar{x}_i - \bar{x})' \right]^{-1} \left[\sum_{i=1}^N (\bar{x}_i - \bar{x}) (\bar{y}_i - \bar{y})' \right] \quad (4.16)$$

A partir de l'équation (4.14), on montre immédiatement que l'estimateur des MCG du modèle à effets aléatoires est une moyenne pondérée de l'estimateur *Between* $\hat{\beta}_{BE}$ (équation 4.16) et de l'estimateur *Within* $\hat{\beta}_{LSDV}$ (équation 3.11).

Proposition 4.6. *Sous les hypothèses (H_3), l'estimateur des MCG des coefficients β du modèle à effets aléatoires, noté $\hat{\beta}_{MCG}$, est une moyenne pondérée des estimateurs *Between* $\hat{\beta}_{BE}$ et *Within* $\hat{\beta}_{LSDV}$:*

$$\hat{\beta}_{MCG} = \Delta \hat{\beta}_{BE} + (I_K - \Delta) \hat{\beta}_{LSDV} \quad (4.17)$$

où les poids déterminés par la matrice Δ sont définis par la relation :

$$\Delta = \psi T \left[\sum_{i=1}^N X_i' Q X_i + \psi T \sum_{i=1}^N (\bar{x}_i - \bar{x}) (\bar{x}_i - \bar{x})' \right]^{-1} \left[\sum_{i=1}^N (\bar{x}_i - \bar{x}) (\bar{x}_i - \bar{x})' \right] \quad (4.18)$$

Essayons à présent d'interpréter ces pondérations et en particulier l'influence du paramètre ψ . Revenons à l'équation (4.14) de l'estimateur des MCG :

$$\hat{\beta}_{MCG} = \left[\frac{1}{T} \sum_{i=1}^N X_i' Q X_i + \psi \sum_{i=1}^N (\bar{x}_i - \bar{x}) (\bar{x}_i - \bar{x})' \right]^{-1} \left[\frac{1}{T} \sum_{i=1}^N X_i' Q y_i + \psi \sum_{i=1}^N (\bar{x}_i - \bar{x}) (\bar{y}_i - \bar{y})' \right]$$

On peut montrer que si le paramètre ψ est égal à 1, l'estimateur des MCG correspond exactement à l'estimateur *des MCO du modèle pooled* (voir annexe A.1). Dans le cas, où $\psi = 0$, l'estimateur des MCG correspond à l'estimateur du modèle à effets fixes :

$$\hat{\beta}_{MCG} = \hat{\beta}_{LSDV} = \left[\sum_{i=1}^N X_i' Q X_i \right]^{-1} \left[\sum_{i=1}^N X_i' Q y_i \right]$$

Remark 2. *Ainsi, le paramètre ψ qui intervient dans la construction de la matrice de poids Δ mesure le poids qui doit être attribué à la variance inter classe (*Between*). Dans le cas, où $\psi = 0$ on retrouve l'estimateur *Within* fondé uniquement sur la variance intra individuelle.*

$$\hat{\beta}_{MCG} \xrightarrow[\psi \rightarrow 0]{} \hat{\beta}_{LSDV} \quad (4.19)$$

Dans le cas complémentaire, $\psi = 1$, on retrouve l'estimateur du modèle *pooled* qui tient compte de la variance totale c'est à dire à la fois de la variance intra individuelle et inter individuelle.

$$\hat{\beta}_{MCG} \xrightarrow{\psi \rightarrow 1} \hat{\beta}_{pooled} \quad (4.20)$$

Si l'on reprend la définition du paramètre ψ , on montre que :

$$\psi = \left(\frac{\sigma_v^2}{\sigma_v^2 + T\sigma_\alpha^2} \right) \quad \lim_{T \rightarrow \infty} \psi = 0 \quad (4.21)$$

Dès lors, à N fixé, lorsque la dimension temporelle du panel tend vers l'infini les estimateurs *MCG* et *Within* sont confondus :

$$\hat{\beta}_{MCG} \xrightarrow{T \rightarrow \infty} \hat{\beta}_{LSDV}$$

Cette propriété sera particulièrement utile pour comprendre la construction du test de spécification d'Hausman.

On constate ainsi, que lorsque $\psi \in]0, 1[$, la structure avec effets individuels aléatoires constitue une solution intermédiaire entre le modèle sans effet individuel (totalement homogène) et le modèle avec effets fixes (totalement hétérogène).

Remark 3. *L'introduction d'effets individuels aléatoires permet d'obtenir une spécification intermédiaire entre le modèle sans effet individuel et le modèle avec effets fixes. L'hypothèse d'une distribution commune des effets individuels permet une d'adopter une structure ni totalement homogène ni totalement hétérogène, dans les effets individuels ont en commun une distribution identique.*

On peut enfin démontrer que la différence entre la matrice de variance de l'estimateur *MCG* et matrice de variance de l'estimateur *Within* est égale à une matrice définie positive (Rao 1972, Thiel 1971). En effet :

$$\text{var} \left(\hat{\beta}_{MCG} \right) = \sigma_v^2 \left[\sum_{i=1}^N X_i' Q X_i + \psi T \sum_{i=1}^N (\bar{x}_i - \bar{x}) (\bar{x}_i - \bar{x})' \right]^{-1} \quad (4.22)$$

On suppose que les quantités $(1/NT) \sum_{i=1}^N X_i' Q X_i$ et $(1/NT) \sum_{i=1}^N X_i' X_i$ convergent vers des matrices définies positives. Puisque le paramètre $\psi > 0$, on vérifie que l'estimateur des *MCG* (estimateur *BLUE*) a toujours une variance inférieure à celle de l'estimateur *Within*.

$$\text{var} \left(\hat{\beta}_{MCG} \right) \leq \text{var} \left(\hat{\beta}_{LSDV} \right)$$

Ce résultat confirme le fait que *sous les hypothèses (H_3), dans un modèle à effets aléatoires, l'estimateur MCG est toujours préférable à l'estimateur Within*. Sachant que le paramètre ψ tend vers 0 lorsque $T \rightarrow \infty$, on en déduit immédiatement que la différence disparaît pour des panels de dimension temporelle infinie :

$$\hat{\beta}_{MCG} \xrightarrow{T \rightarrow \infty} \hat{\beta}_{LSDV}$$

Dans le cas où la matrice de variance covariance V n'est pas connue, il est nécessaire comme nous l'avons dit précédemment de procéder en deux étapes. Dans une première étape on estime les variances des différentes composantes des résidus. Pour cela on estime le modèle à effets fixes $\hat{\beta}_{LSDV}$, et l'on construit les estimateurs des variances intra et inter individuelles :

Definition 4.7. *Les estimateurs des variances des composantes intra classe et inter classe des résidus du modèle à effets aléatoires sont respectivement donnés par :*

$$\hat{\sigma}_v^2 = \frac{\sum_{i=1}^N \sum_{t=1}^T \left[(y_{i,t} - \bar{y}_i) - \hat{\beta}'_{LSDV} (x_{i,t} - \bar{x}_i) \right]^2}{N(T-1) - K} \quad (4.23)$$

$$\hat{\sigma}_\alpha^2 = \frac{\sum_{i=1}^N \left(\bar{y}_i - \hat{\beta}'_{LSDV} \bar{x}_i - \bar{\mu} \right)^2}{N - K - 1} - \frac{1}{T} \hat{\sigma}_v^2 \quad (4.24)$$

On en déduit alors l'estimateur $\hat{\psi}$ du paramètre ψ selon :

$$\hat{\psi} = \left(\frac{\hat{\sigma}_v^2}{\hat{\sigma}_v^2 + T \hat{\sigma}_\alpha^2} \right) \quad (4.25)$$

A l'aide des réalisations de l'estimateur *Between* on construit la matrice de poids Δ . Dans une seconde étape, on applique directement la formule (4.11) des MCG ou l'on utilise la somme pondérée (4.17).

Proposition 4.8. *L'estimateur des MCG construit en deux étapes converge asymptotiquement ($NT \rightarrow \infty$) vers l'estimateur des MCG obtenu dans le cas où la matrice de variance covariance V est connue. Même dans le cas de petits échantillons ($T \geq 3$ et $N - (K + 1) \geq 9$), la procédure en deux étapes reste plus efficace que l'estimateur *Within*, dans le sens où la différence entre la matrices de variances covariances des MCG et celle de l'estimateur reste une matrice définie positive.*

Nous ne proposerons pas ici de démonstration de cette proposition (cf. Fuller and Battese 1974, Taylor 1980). Mais cela confirme l'idée que quelle que soit la taille du panel, l'estimateur des MCG est plus efficace que l'estimateur *Within*. Proposons à présent une application de ces différents résultats.

4.4. Une application

Reprenons l'application proposée à la section précédente. On considère le modèle suivant :

$$s_{i,t} = c + \beta u_{i,t} + \gamma p_{i,t} + v_{i,t} \quad \forall i = 1, \dots, 17 \quad (4.26)$$

$$\varepsilon_{i,t} = \alpha_i + v_{i,t} \quad (4.27)$$

où $s_{i,t}$ désigne le nombre de jours chômés pour cause de grève, pour 1000 salariés du secteur industriel, $u_{i,t}$ désigne le taux de chômage de l'économie et $p_{i,t}$ le taux d'inflation. On suppose ici que l'on peut spécifier les effets individuels α_i sous la forme d'effets aléatoires et que les résidus $\varepsilon_{i,t} = v_{i,t} + \alpha_i$ satisfont les hypothèses (H_3). Si l'on ne désire que les réalisations de l'estimateur des MCG du modèle aléatoire, à l'exclusion de tout autre estimateur, le programme TSP est alors le suivant :

```
load(file='strikes.wks');
?---- Estimation Modèle à Effets Aléatoires ----
panel (id=i,time=t,nowhit,nobet,nowithin) srt u p;
```

Les résultats de ce programme, reproduits sur la figure (4.1) sont bien entendu totalement identiques à ceux de la figure (1.2).

Figure 4.1: Estimation Modèle à Effets Aléatoires

```
PANEL DATA ESTIMATION
=====

Balanced data: NI= 17, T= 35, NOB= 595

Variance Components (random effects) Estimates:

V WITH (variance of Uif) = 0.25514E+06
VBET (variance of Ai) = 55401.
(computed from small sample formula)
THETA (0=WITHIN, 1=TOTAL) = 0.11628

Dependent variable: SRT

Sum of squared residuals = .152560E+09      R-squared = .214013
Variance of residuals = 264862.             Adjusted R-squared = .189450
Std. error of regression = 514.647

      Estimated Standard
Variable Coefficient Error t-statistic
U      -12.0814    8.78623   -1.37504
P       16.4247    4.72670    3.47489
C       248.622    71.9106    3.45737
```

Contrairement au cas de l'estimateur *Within* où il n'y avait que deux coefficients (en dehors des effets fixes), les résultats de l'estimation du modèle à effets aléatoires font apparaître 3 réalisations d'estimateurs des coefficients du chômage (u), de l'inflation (p) et d'une constante (c). Le coefficient associé à cette constante correspond à l'estimateur de la moyenne des effets individuels.

$$E(\alpha_i) = c \quad (4.28)$$

On peut vérifier d'ailleurs que cette réalisation est approximativement égale à la moyenne des réalisations des estimateurs de effets individuels dans le cas du modèle à effets fixes (figure 3.1).

On dispose en outre de trois informations supplémentaires :

1. une réalisation de l'estimateur de la variance intra classe $\hat{\sigma}_v^2$.
2. une réalisation de l'estimateur de la variance inter classe $\hat{\sigma}_\alpha^2$ ou variance des effets individuels.
3. une réalisation de l'estimateur $\hat{\psi}$ du paramètre ψ .

On peut vérifier numériquement que ces estimations satisfont les conditions de la définition (4.7) puisque :

$$\hat{\sigma}_v^2 = \frac{\sum_{i=1}^N \sum_{t=1}^T [(y_{i,t} - \bar{y}_i) - \hat{\beta}'_{LSDV} (x_{i,t} - \bar{x}_i)]^2}{N(T-1) - K} = 0.2551\text{E} + 06$$

$$\hat{\sigma}_\alpha^2 = \frac{\sum_{i=1}^N \bar{y}_i - \hat{\beta}'_{LSDV} \bar{x}_i - \bar{\mu}}{N - K - 1} - \frac{1}{T} \hat{\sigma}_v^2 = 55\,401$$

On en déduit alors une estimation du paramètre ψ :

$$\hat{\psi} = \left(\frac{\hat{\sigma}_v^2}{\hat{\sigma}_v^2 + T\hat{\sigma}_\alpha^2} \right) = \left(\frac{0.2551\text{E} + 06}{0.2551\text{E} + 06 + 37 \times 55\,401} \right) \simeq 0.11628$$

Rappelons que si le paramètre ψ tend vers 0, l'estimateur des MCG converge vers l'estimateur *Within* et que si au contraire ψ tend vers 1, l'estimateur des MCG converge vers l'estimateur du modèle *pooled*. Dans le cas de notre échantillon, on observe que l'estimateur des MCG est plus "près" de l'estimateur *Within* que de celui du modèle totalement homogène. Ceci signifie que dans notre panel, la variance temporelle pour chaque pays (intra classe) domine la variance internationale (inter classe).

5. Tests de spécification des effets individuels

En présence d'un modèle à effets individuels, la question qui se pose immédiatement est de savoir comment ces effets individuels doivent être spécifiés : doit-on adopter l'hypothèse d'effets fixes ou au contraire l'hypothèse d'effets aléatoires ? Nous avons vu dans la section précédente que l'estimateur des MCG, utilisé dans le cas du modèle à effets aléatoires, est asymptotiquement (sous l'hypothèse $T \rightarrow \infty$) identique à l'estimateur *Within*. Toutefois, pour *des panels de dimension temporelle réduite (typiquement le cas des panels macroéconomiques)*, il peut exister de fortes différences entre les réalisations des deux estimateurs (Hausman 1978). Dès lors, au delà de l'interprétation économique, le choix de la spécification, et par là même de la méthode d'estimation, est particulièrement important pour ce type de panels.

Une des principaux problèmes qui peut se poser dans le cadre du modèle à effets aléatoires provient de l'éventuelle corrélation entre les variables explicatives $x_{i,t}$ et les effets individuels α_i . Sur le plan économique, cette corrélation traduit l'influence des spécificités individuelles structurelles (ou a-temporelles) sur la détermination du niveau des variables explicatives. Reprenons l'exemple de la fonction de production. Supposons que l'on dispose d'un échantillon de données de *PIB* et de facteurs (travail et capital) sur une durée de T périodes pour un ensemble de N pays. Si l'on note $y_{i,t}$ le logarithme du *PIB*, $k_{i,t}$ le logarithme du stock de capital privé et $n_{i,t}$ le logarithme de l'emploi et que l'on suppose une fonction de production de type Cobb Douglass, le modèle général s'écrit sous la forme $\forall i \in [1, N], \forall t \in [1, T]$:

$$y_{i,t} = \beta_i k_{i,t} + \gamma_i n_{i,t} + \alpha_i + v_{i,t}$$

Les innovations $v_{i,t}$ sont supposées être *i.i.d.* de moyenne nulle et de variance égale à σ_v^2 , $\forall i \in [1, N]$. On suppose que les effets individuels sont aléatoires et qu'il existe une corrélation entre ces effets et le niveau des variables explicatives :

$$E(\alpha_i k_{i,t}) > 0 \quad E(\alpha_i n_{i,t}) > 0$$

Rappelons enfin que dans ce modèle, le résidu total $\varepsilon_{i,t} = \alpha_i + v_{i,t}$ s'apparente à la productivité totale des facteurs ou au résidu de Solow. Dès lors, *ces deux corrélations traduisent tout simplement le fait que plus un pays est structurellement productif (plus son effet individuel est élevé), plus le niveau de facteur utilisé sera lui même important*. Il y a tout lieu de penser qu'une telle corrélation puisse être observée sur données historiques.

Sur le plan plus technique, la présence d'une corrélation entre les effets individuels et les variables explicatives viole la sixième condition des hypothèses (H_3) puisque :

$$y_{i,t} = \mu + \beta' x_{i,t} + \varepsilon_{i,t} \quad (5.1)$$

$$\varepsilon_{i,t} = \alpha_i + v_{i,t} \quad (5.2)$$

$$E(\alpha_i x'_{i,t}) \neq 0 \quad (5.3)$$

En présence d'une telle corrélation, nous allons montrer que l'estimateur des MCG est un estimateur biaisé des paramètres du vecteur β , alors l'estimateur *Within* ne l'est pas. Ce biais s'assimile tout simplement à un biais d'endogénéité, puisque les résidus totaux $\varepsilon_{i,t} = \alpha_i + v_{i,t}$ sont corrélés avec les variables explicatives. *Toute la stratégie de test de spécification des effets individuels repose alors sur la comparaison des deux estimateurs, dont la divergence traduit la présence d'une corrélation et donc la nécessaire adoption du modèle à effets fixes et de l'estimateur Within. Dans le cas contraire où les deux estimateurs donnent des résultats sensiblement identiques, les hypothèses (H_3) sont satisfaites et l'on peut spécifier le modèle avec des effets individuels aléatoires. L'estimateur des MCG est alors l'estimateur BLUE.*

5.1. Corrélation des effets individuels et des variables explicatives

Considérons le modèle à effets aléatoires suivant $\forall i \in [1, N], \forall t \in [1, T]$:

$$y_{i,t} = \mu + \beta' x_{i,t} + \varepsilon_{i,t}$$

$$\varepsilon_{i,t} = \alpha_i + v_{i,t}$$

où $\beta' = (\beta_1 \beta_2 \dots \beta_K)$ est un vecteur de constantes. On suppose ici que les effets individuels aléatoires α_i sont corrélés aux variables explicatives :

$$E(\alpha_i | X_i) \neq 0$$

Cette corrélation peut être exprimée sous différentes formes. Nous retiendrons la formulation de Mundlak (1978a) qui postule une approximation linéaire de cette espérance.

Definition 5.1. *Dans la formulation de Mundlak (1978a), on suppose ainsi que les effets individuels peuvent s'écrire sous la forme de la somme d'une combinaison linéaire des moyennes individuelles des variables explicatives et d'une composante orthogonale i.i.d. α_i^* .*

$$\alpha_i = \bar{x}_i' a + \alpha_i^* \quad (5.4)$$

avec $a \in \mathbb{R}^K$. Sous cette hypothèse, le modèle devient :

$$y_{i,t} = \mu + \beta' x_{i,t} + \bar{x}_i' a + \varepsilon_{i,t} \quad (5.5)$$

$$\varepsilon_{i,t} = \alpha_i^* + v_{i,t} \quad (5.6)$$

On obtient ainsi une *structure de résidus* $\varepsilon_{i,t} + \bar{x}_i' a$ faisant apparaître trois termes : la composante orthogonale des effets individuels α_i^* , la projection linéaire¹¹ des effets individuels sur l'espace des variables explicatives $\bar{x}_i' a$, et un terme d'erreur *i.i.d.* dans les dimensions temporelle et individuelle, $v_{i,t}$. On suppose que les résidus globaux $\varepsilon_{i,t}$ satisfont les hypothèses (H_4) :

Hypothèses (H_4) On suppose que les résidus $\varepsilon_{i,t} = \alpha_i^* + v_{i,t}$ satisfont les conditions suivantes, $\forall i \in [1, N], \forall t \in [1, T]$

- $E(\alpha_i^*) = E(v_{i,t}) = 0$
- $E(\alpha_i^* v_{i,t}) = 0$
- $E(\alpha_i^* \alpha_j^*) = \begin{cases} \sigma_{\alpha^*}^2 & i = j \\ 0 & \forall i \neq j \end{cases}$
- $E(v_{i,t} v_{j,s}) = \begin{cases} \sigma_v^2 & t = s, i = j \\ 0 & \forall t \neq s, \forall i \neq j \end{cases}$
- $E(v_{i,t} x_{i,t}') = E(\alpha_i^* x_{i,t}') = 0$

De la même façon que pour le modèle à effets aléatoires standard, on construit la matrice de variance covariance des résidus globaux $\varepsilon_{i,t}$, son inverse et l'on construit le modèle vectoriel.

Le modèle à effets individuels aléatoires, sous les hypothèses (H_4), s'écrit vectoriellement sous la forme :

$$\underset{(T,1)}{y_i} = \underset{(T,K+1)(K+1,1)}{\widetilde{X}_i} \underset{(T,K)}{\gamma} + \underset{(T,K)}{e \bar{x}_i'} \underset{(K,1)}{a} + \underset{(T,1)}{\varepsilon_i} \quad \forall i = 1, \dots, N \quad (5.7)$$

avec $\varepsilon_i = \alpha_i^* e + v_i$, $\widetilde{X}_i = (e, X_i)$ et $\gamma' = (\mu, \beta')$. Sous les hypothèses (H_4), la matrice de variance covariance de ε_i , notée V , est alors définie par :

$$E(\varepsilon_i \varepsilon_j') = E[(\alpha_i^* e + v_i)(\alpha_j^* e + v_j)'] = \begin{cases} \sigma_{\alpha^*}^2 e e' + \sigma_v^2 I_T = V^* & i = j \\ 0 & i \neq j \end{cases} \quad (5.8)$$

L'inverse de cette matrice de variance covariance est définie comme :

$$V^{*-1} = \frac{1}{\sigma_v^2} \left[I_T - \left(\frac{\sigma_{\alpha^*}^2}{\sigma_v^2 + T \sigma_{\alpha^*}^2} \right) e e' \right] \quad (5.9)$$

¹¹La projection linéaire étant adaptée au schéma de corrélation proposé.

Si l'on applique les Moindres Carrés Généralisés au modèle (5.7), en tenant compte de la matrice de variance covariance V^* qui possède les bonnes propriétés, on obtiendra alors des estimateurs convergents des différents paramètres :

$$\hat{\mu}_{MCG}^* = \bar{y} - \bar{x}'\hat{\beta}_{BE} \quad (5.10)$$

$$\hat{a}_{MCG}^* = \hat{\beta}_{BE} - \hat{\beta}_{LSDV} \quad (5.11)$$

$$\hat{\beta}_{MCG}^* = \Delta\hat{\beta}_{BE} + (I_K - \Delta)\hat{\beta}_{LSDV} \quad (5.12)$$

Nous ne présenterons pas ici la forme générale de la matrice de poids Δ . Toutefois, en reprenant la formule générale de l'estimateur des MCG, on peut montrer que l'estimateur $\hat{\beta}_{MCG}^*$ est un estimateur non biaisé du paramètre β :

$$\hat{\beta}_{MCG} \xrightarrow{T \rightarrow \infty} \hat{\beta}_{LSDV}$$

$$\hat{\beta}_{MCG} \xrightarrow{NT \rightarrow \infty} \beta$$

Mais un problème se pose lorsque l'on applique, à tort, l'estimateur des MCG, au modèle initial, c'est à dire au modèle dans lequel on n'a pas pris le soin de décomposer les effets individuels en deux composantes, dont une est strictement orthogonale aux variables explicatives. Ainsi, supposons à présent que le modèle (5.7) soit le bon, et que l'on applique à tort les Moindres Carrés Généralisés à la spécification suivante :

$$y_{i,t} = \mu + \beta'x_{i,t} + \varepsilon_{i,t} \quad (5.13)$$

$$\varepsilon_{i,t} = \alpha_i + v_{i,t} \quad (5.14)$$

Soit $\hat{\beta}_{MCG}$ l'estimateur des MCG appliqué à ce modèle, sous l'hypothèse que les données sont générées par le modèle de (5.7). On part de la relation suivante :

$$\hat{\beta}_{MCG} = \Delta\hat{\beta}_{BE} + (I_K - \Delta)\hat{\beta}_{LSDV}$$

Dans ce contexte, on peut montrer que l'estimateur Between $\hat{\beta}_{BE}$ converge vers $\beta + a$, puisque le vecteur a relie les moyennes individuelles x_i aux effets individuels α_i . Dès lors, il peut apparaître un biais dans l'estimateur $\hat{\beta}_{MCG}$ à taille T finie.

Proposition 5.2. *Lorsque les effets individuels et les variables explicatives sont corrélées, lorsque les résidus satisfont les hypothèses (H_4) et le modèle (5.7), l'application à tort des MCG au modèle (5.13)*

$$y_{i,t} = \mu + \beta'x_{i,t} + \varepsilon_{i,t}$$

$$\varepsilon_{i,t} = \alpha_i + v_{i,t}$$

conduit à un biais semi asymptotique dans l'estimateur du vecteur β à taille T finie :

$$\widehat{\beta}_{MCG} \xrightarrow[N \rightarrow \infty]{p} \beta + \Delta a \quad (5.15)$$

Ce biais disparaît asymptotiquement :

$$\widehat{\beta}_{MCG} \xrightarrow[T \rightarrow \infty]{p} \beta \quad (5.16)$$

La démonstration de ce résultat est très simple. Admettons que les deux résultats suivants :

$$\begin{aligned} \widehat{\beta}_{BE} &\xrightarrow[NT \rightarrow \infty]{p} \beta + a \\ \widehat{\beta}_{LSDV} &\xrightarrow[NT \rightarrow \infty]{p} \beta \end{aligned}$$

Dès lors, à taille T d'échantillon finie, on montre que facilement l'existence d'un biais Δa dans l'estimateur des MCG de β :

$$\begin{aligned} \text{plim}_{N \rightarrow \infty} \widehat{\beta}_{MCG} &= \Delta \text{plim}_{N \rightarrow \infty} \widehat{\beta}_{BE} + (I_K - \Delta) \text{plim}_{N \rightarrow \infty} \widehat{\beta}_{LSDV} \\ &= \Delta (\beta + a) + (I_K - \Delta) \beta \\ &= \beta + \Delta a \end{aligned}$$

On sait cependant que lorsque la taille d'échantillon T tend vers l'infini, la matrice de poids Δ tend vers 0, et l'estimateur des MCG converge vers l'estimateur *Within*. Le biais disparaît :

$$\text{plim}_{T \rightarrow \infty} \widehat{\beta}_{MCG} = \text{plim}_{T \rightarrow \infty} \widehat{\beta}_{LSDV} = \beta$$

Remark 4. Ce résultat signifie qu'en présence d'une corrélation entre les effets individuels et les variables explicatives, l'estimateur des MCG est biaisé à taille d'échantillon T finie, et ce même si le nombre d'individus N tend vers l'infini. Par opposition l'estimateur *Within*, dont la construction permet la suppression des effets individuels, est asymptotiquement non biaisé.

Tout le problème repose sur le fait, que si les effets individuels et les variables explicatives ne sont pas corrélées, l'estimateur des MCG est dans ce cas l'estimateur BLUE, c'est à dire l'estimateur convergent, non biaisé à variance minimale. Ainsi, le problème de la spécification des effets individuels est un problème particulièrement important de l'économétrie de panel appliquée. Pour s'en convaincre, il suffit de considérer les résultats de l'estimation de notre modèle sur les grèves (figure 1.2), dans lesquels les estimateurs des MCG et les estimateurs *Within* donnent des résultats très différents sur le plan quantitatifs, notamment en ce qui concerne la variable du chômage.

5.2. Test de spécification d'Hausman

Le test de spécification d'Hausman (1978) est un test général qui peut être appliqué à des nombreux problèmes de spécification en économétrie. Mais son application la plus répandue est celle des tests de spécification des effets individuels en panel. Il sert ainsi à discriminer les effets fixes et aléatoires.

L'idée générale du test d'Hausman est la fois simple et géniale. Supposons que l'on cherche à tester la présence éventuelle d'une corrélation ou d'un défaut de spécification. Admettons que l'on dispose de deux types d'estimateurs pour les paramètres du modèle étudié. Le premier estimateur est supposé être l'estimateur non biaisé à variance minimale sous l'hypothèse nulle de spécification correcte du modèle (absence de corrélation). En revanche, sous l'hypothèse alternative de mauvaise spécification, cet estimateur est supposé être biaisé. On suppose que le second estimateur est non biaisé dans les deux cas. Dès lors, il suffit de comparer une distance, pondérée par une matrice de variance covariance, entre les deux estimateurs pour pouvoir déterminer si la spécification est correcte ou non. Si la distance est statistiquement nulle, la spécification est correcte, on choisit le premier estimateur. Si la distance est importante, le modèle est mal spécifié.

L'application technique de ce principe suppose tout de même que l'on construise la matrice de variance covariance de l'écart entre les deux estimateurs. De façon générale, il devrait alors apparaître des termes de covariance entre les deux estimateurs. Afin de les éliminer, on considère le lemme suivant :

Lemme 5.3. *On considère deux estimateurs $\hat{\beta}_1$ et $\hat{\beta}_2$, d'un vecteur de paramètres $\beta \in \mathbb{R}^K$, que l'on suppose convergent et asymptotiquement normalement distribués. On suppose que $\hat{\beta}_1$ atteint la borne asymptotique de Cramer Rao. Pour un échantillon de taille N , les quantités $\sqrt{N}(\hat{\beta}_1 - \beta)$ et $\sqrt{N}(\hat{\beta}_2 - \beta)$ sont asymptotiquement distribuées selon des lois normales de matrice de variance covariance respectives V_0 et V_1 . Sous ces hypothèses, les distributions asymptotiques de $\sqrt{N}(\hat{\beta}_1 - \beta)$ et la différence $\sqrt{N}(\hat{\beta}_1 - \hat{\beta}_2)$ ne sont pas corrélées, ce qui implique :*

$$\text{var}(\hat{\beta}_1 - \hat{\beta}_2) = \text{var}(\hat{\beta}_1) - \text{var}(\hat{\beta}_2) \quad (5.17)$$

Dès lors, en appliquant ce lemme, Hausman préconise de fonder le test de spécification sur la statistique suivante :

$$H = (\hat{\beta}_1 - \hat{\beta}_2)' [\text{var}(\hat{\beta}_1 - \hat{\beta}_2)]^{-1} (\hat{\beta}_1 - \hat{\beta}_2) \quad (5.18)$$

Sous l'hypothèse nulle de spécification correcte, cette statistique est asymptotiquement distribuée selon un chi deux à K degrés de liberté.

Appliquons à présent le test d'Hausman au problème de la spécification des effets individuels dans un panel.

Definition 5.4. *L'hypothèse testée concerne la corrélation des effets individuels et des variables explicatives :*

$$H_0 : E(\alpha_i | X_i) = 0 \quad (5.19)$$

$$H_1 : E(\alpha_i | X_i) \neq 0 \quad (5.20)$$

Ce test peut être interpréter comme un test de spécification. Sous H_0 , le modèle peut être spécifié avec des effets individuels aléatoires et l'on doit alors retenir l'estimateur des MCG (estimateur BLUE). Sous l'hypothèse alternative H_1 , le modèle doit être spécifié avec des effets individuels fixes et l'on doit alors retenir l'estimateur Within (estimateur non biaisé).

On peut ici appliquer le lemme (5.3) en considérant l'estimateur des MCG (estimateur $\hat{\beta}_1$) et l'estimateur *Within* (estimateur $\hat{\beta}_2$). En effet, nous avons vu que sous H_0 , dans un modèle à effets aléatoires, l'estimateur des MCG $\hat{\beta}_{MCG}$ est l'estimateur BLUE. Sous l'hypothèse de normalité des résidus, les deux estimateurs $\hat{\beta}_{MCG}$ et $\hat{\beta}_{LSDV}$ sont convergents et asymptotiquement distribués selon une loi normale. On en déduit que :

$$var(\hat{\beta}_{MCG} - \hat{\beta}_{LSDV}) = var(\hat{\beta}_{MCG}) - var(\hat{\beta}_{LSDV}) \quad (5.21)$$

Dès lors, sous l'hypothèse nulle H_0 , la distance entre les deux estimateur est nulle puisque tous deux convergent vers la vraie valeur β . En revanche, sous l'hypothèse alternative, la présence de la corrélation entre les effets individuels et les variables explicatives conduit à un biais de l'estimateur des MCG et cette distance devient alors importante.

Definition 5.5. *La statistique du test d'Hausman appliqué au test de la spécification des effets individuels est la suivante :*

$$H = (\hat{\beta}_{MCG} - \hat{\beta}_{LSDV})' \left[var(\hat{\beta}_{MCG} - \hat{\beta}_{LSDV}) \right]^{-1} (\hat{\beta}_{MCG} - \hat{\beta}_{LSDV}) \quad (5.22)$$

Sous l'hypothèse nulle H_0 , la statistique H suit asymptotiquement (N tend vers l'infini) un chi deux à K degrés de liberté.

Ainsi, si la réalisation de la statistique H est supérieure au seuil à $\alpha\%$, on rejette l'hypothèse nulle et l'on privilégie l'adoption d'effets individuels fixes et l'utilisation de l'estimateur Within non biaisé.

Il faut noter que la statistique du test de Hausman est dégénérée lorsque l'on considère que la dimension T tend vers l'infini. En effet, dans ce cas, l'estimateur des MCG converge vers l'estimateur *Within*¹², ce qui implique que le numérateur et le dénominateur de la statistique d'Hausman tendent vers 0. Toutefois, on sait que lorsque T tend vers l'infini, les modèles à effets aléatoires et à effets fixes ne peuvent être distingués et sont parfaitement similaires. Dès lors, la question de la spécification de ces effets importe peu.

5.3. Application

Reprenons l'application sur les journées de grèves proposée à la section précédente. Si l'on désire obtenir les réalisations des estimateurs MCG et *Within*, ainsi que la statistique du test d'Hausman, à l'exclusion des estimateurs *Between* et *Pooled*, le programme TSP est alors le suivant :

```
load(file='strikes.wks');
?---- Test Hausman, Within et VarComp----
panel (id=i,time=t,nowhit,nobet) srt u p;
```

Les résultats de ce programme, reproduits sur la figure (5.1), sont bien entendu totalement identiques à ceux de la figure (1.2).

Pour l'échantillon considéré, la réalisation de la statistique du test d'Hausman est de 13,924. Etant donné que le modèle comporte deux variables explicatives ($K = 2$), cette statistique suit un chi deux à deux degrés de liberté. A 95%, le seuil est de 5.99, donc ici on rejette l'hypothèse nulle d'absence de corrélation entre les effets individuels et les variables explicatives. Le chômage et l'inflation sont corrélés aux spécificités structurelles et a-temporelles du volume des jours chômés pour cause de grève. Plus un pays, toutes choses égales par ailleurs, connaît de jours de grèves, plus son niveau de chômage et d'inflation sont élevés (ou faibles suivant le signe de la corrélation). Ainsi, *on doit ici privilégier l'adoption d'un modèle à effets fixes et retenir l'estimateur Within*.

On peut vérifier la validité de cette conclusion, en comparant de façon heuristique les réalisations des estimateurs des modèles à effets fixes et à effets aléatoires. On

¹² Contrairement au cas où N tend vers l'infini où l'on a :

$$\lim_{N \rightarrow \infty} \hat{\beta}_{MCG} = \lim_{N \rightarrow \infty} \hat{\beta}_{LSDV} = \beta$$

dans le cas où la dimension T tend vers l'infini, on a :

$$\lim_{N \rightarrow \infty} \hat{\beta}_{MCG} = \hat{\beta}_{LSDV}$$

Dès lors, les deux estimateurs tendent à être identiques.

Figure 5.1: Résultats Test d'Hausman (1978)

```

PANEL DATA ESTIMATION
=====

Balanced data: NI= 17, T= 35, NOB= 595

WITHIN (fixed effects) Estimates:

Dependent variable: SRT

Sum of squared residuals= .146958E+09      R-squared = .242875
Variance of residuals = 255136.           Adjusted R-squared = .219215
Std. error of regression = 505.110

      Estimated Standard
Variable Coefficient Error t-statistic
U      -21.5968    9.19158  -2.34963
P       16.2729    4.75658   3.42113

Variance Components (random effects) Estimates:

V WITH (variance of Uit) = 0.25514E+06
V BET (variance of Ai) = 55401.
(computed from small sample formula)
THETA (0=WITHIN, 1=TOTAL) = 0.11628

Dependent variable: SRT

Sum of squared residuals= .152560E+09      R-squared = .214013
Variance of residuals = 264862.           Adjusted R-squared = .189450
Std. error of regression = 514.647

      Estimated Standard
Variable Coefficient Error t-statistic
U      -12.0814    8.78623  -1.37504
P       16.4247    4.72670   3.47489
C       248.622    71.9106   3.45737

Hausman test of H0:RE vs. FE: CHISQ(2) = 13.924, P-value = [.0009]

```

observe que pour l'échantillon étudié, les coefficients estimés notamment pour le chômage sont sensiblement différents dans les deux cas (-21 pour l'estimateur *Within* et -12 pour l'estimateur des MCG). Rappelons, que sous l'hypothèse nulle d'absence de corrélation ces deux estimateurs devraient converger vers la même valeur. Cette observation confirme donc notre diagnostic quant à la présence d'une corrélation entre les effets individuels et les deux variables explicatives du modèle. On doit ainsi privilégier les résultats de l'estimateur *Within*.

6. Modèle à coefficients fixes et aléatoires

On peut envisager plusieurs variantes du modèle linéaire standard. La plupart de ces variantes sont construites en levant progressivement les différentes hypothèses (H_3) sur la structure des résidus. On peut d'abord introduire une autocorrélation des résidus (dans la dimension temporelle). On utilise alors des méthodes d'estimation proches de celles de Cochrane Orcut (1949), utilisée en séries temporelles. Une seconde variante possible, et par ailleurs souvent utilisée, consiste à supposer une structure de variance covariance arbitraire pour les résidus (méthode de Chamberlain 1982 et 1984). Il est dans ce cas possible de traiter des problèmes d'autocorrélation et d'hétéroscédasticité. Mais nous allons ici nous concentrer sur un autre type de modèle linéaire : *les modèles avec un mélange de coefficients fixes et aléatoires*. En particulier, nous allons étudier le *modèle MFR (Mixed Fixed and Random Coefficient)* proposé par Hsiao (1989).

6.1. Modèle MFR de Hsiao (1989)

Commençons par présenter la structure générale des modèle à effets individuels fixes et à coefficients aléatoires.

Definition 6.1. *On considère le modèle MFR suivant :*

$$y_{i,t} = \alpha_i + \beta_i' x_{i,t} + v_{i,t} \quad \forall i \in [1, N], \forall t \in [1, T] \quad (6.1)$$

où $\alpha_i \in \mathbb{R}, \beta_i' = (\beta_{i,1}, \beta_{i,2}, \dots, \beta_{i,K}) \in \mathbb{R}^K$. On suppose ici que les effets individuels α_i sont fixes et que les coefficients β_i des exogènes sont distribués selon une distribution commune de moyenne $E(\beta_i) = \beta$ et de matrice de variance covariance :

$$E[(\beta_i - \beta)(\beta_i - \beta)'] = \underset{(K,K)}{\Delta} = \begin{pmatrix} \sigma_{\beta_1}^2 & \sigma_{\beta_1, \beta_2} & \dots & \sigma_{\beta_1, \beta_K} \\ \sigma_{\beta_2, \beta_1} & \sigma_{\beta_2}^2 & \dots & \sigma_{\beta_2, \beta_K} \\ \dots & \dots & \dots & \dots \\ \sigma_{\beta_K, \beta_1} & \sigma_{\beta_K, \beta_2} & \dots & \sigma_{\beta_K}^2 \end{pmatrix} \quad (6.2)$$

Le modèle (6.1) peut s'écrire sous la forme vectorielle suivante :

$$y_i = e\alpha_i + X_i\beta_i + v_i \quad \forall i \in [1, N] \quad (6.3)$$

où e désigne un vecteur unitaire de dimension $(T, 1)$ et où l'on suppose que les résidus $v_{i,t}$ satisfont les hypothèses (H_1), avec en particulier $E(v_i v_i') = \sigma_{v_i}^2 I_T$. L'hypothèse selon laquelle le paramètre β_i est un coefficient aléatoire implique une nouvelle structure des résidus qui peut s'exprimer sous la forme suivante :

$$y_i = e\alpha_i + X_i\beta + \varepsilon_i \quad (6.4)$$

avec $\forall i \in [1, N]$

$$\varepsilon_i = X_i \tilde{\beta}_i + v_i \quad (6.5)$$

où les coefficients aléatoires $\tilde{\beta}_i = \beta_i - \beta$, sont tels que $E(\tilde{\beta}_i) = 0$, $E(X_i \tilde{\beta}_i) = 0$ et $E(\tilde{\beta}_i \tilde{\beta}_i') = \Delta$. A partir de l'équation (6.4), on obtient ainsi un modèle où *les coefficients α_i et β sont fixes, mais où la structure des résidus ne satisfait pas les hypothèses standards*. Pour le montrer, établissons la forme de la matrice de variance covariance des résidus individuels, noté Φ_i , est alors définie par¹³ :

$$\Phi_i = E(\varepsilon_i \varepsilon_i') = E \left[(X_i \tilde{\beta}_i + v_i) (X_i \tilde{\beta}_i + v_i)' \right] = X_i E(\tilde{\beta}_i \tilde{\beta}_i') X_i' + E(v_i v_i')$$

On montre ainsi que :

$$\Phi_i = X_i \Delta X_i' + \sigma_{v_i}^2 I_T \quad (6.6)$$

La matrice de variance covariance est ainsi conditionnelle aux individus et de plus n'est pas diagonale. On doit donc *appliquer une méthode de Moindres Carrés Généralisés sur un modèle transformé excluant les effets individuels*.

Supposons que l'on connaisse la matrice de variance covariance Φ_i et que celle-ci soit une matrice définie positive, on cherche alors à construire un modèle transformé où les résidus satisfont les hypothèses (H_1) . Pour cela, on introduit une transformée P_i telle que :

$$\Phi_i^{-1} = P_i' P_i \quad \forall i \in [1, N] \quad (6.7)$$

On définit alors les variables transformées suivantes :

$$X_i^* = P_i X_i \quad y_i^* = P_i y_i \quad e_i^* = P_i e \quad \varepsilon_i^* = P_i \varepsilon_i \quad (6.8)$$

Dès lors, le modèle (6.4), s'écrit sous la forme :

$$P_i y_i = P_i e \alpha_i + P_i X_i \beta + P_i \varepsilon_i$$

ou encore

$$y_i^* = \alpha_i e_i^* + X_i^* \beta + \varepsilon_i^* \quad (6.9)$$

Dans ce modèle transformé, les résidus ε_i^* satisfont les hypothèses H_1 puisque :

$$E[\varepsilon_i^* (\varepsilon_i^*)'] = P_i E(\varepsilon_i \varepsilon_i') P_i' = P_i (P_i' P_i)^{-1} P_i' = I_T \quad (6.10)$$

Reste enfin à appliquer un opérateur permettant d'éliminer les effets individuels transformés. Pour cela on construit une application Q_i^* telle $\forall i \in [1, N]$:

$$Q_i^* e_i^* = Q_i^* P_i e = 0$$

¹³Si X_i est non stochastique.

On montre alors que Q_i^* est telle :

$$Q_i^* = I_T - P_i e (e' \Phi_i^{-1} e)^{-1} e P_i' \quad (6.11)$$

La preuve est la suivante :

$$\begin{aligned} Q_i^* P_i e &= \left[I_T - P_i e (e' \Phi_i^{-1} e)^{-1} e P_i' \right] P_i e \\ &= P_i e - P_i e (e' \Phi_i^{-1} e)^{-1} e P_i' P_i e \\ &= P_i e - P_i e (e' P_i' P_i e)^{-1} e P_i' P_i e \\ &= P_i e - P_i e \\ &= 0 \end{aligned}$$

Appliquons à présent cette seconde transformation au modèle transformé (6.9), il vient :

$$\begin{aligned} Q_i^* y_i^* &= Q_i^* e_i^* \alpha_i + Q_i^* X_i^* \beta + Q_i^* \varepsilon_i^* \\ \iff Q_i^* y_i^* &= Q_i^* X_i^* \beta + Q_i^* \varepsilon_i^* \end{aligned} \quad (6.12)$$

A partir de cette expression, on peut finalement proposer un estimateur de la moyenne des coefficients β_i , c'est un estimateur du paramètre β . Soit $\hat{\beta}_{MFR}$ cet estimateur :

$$\hat{\beta}_{MFR} = \left[\sum_{i=1}^N (X_i^*)' Q_i^* X_i^* \right]^{-1} \left[\sum_{i=1}^N (X_i^*)' Q_i^* y_i^* \right] \quad (6.13)$$

En développant cette expression on obtient :

$$\begin{aligned} \hat{\beta}_{MFR} &= \left\{ \sum_{i=1}^N X_i' P_i' \left[I_T - P_i e (e' \Phi_i^{-1} e)^{-1} e P_i' \right] P_i X_i \right\}^{-1} \\ &\quad \left\{ \sum_{i=1}^N X_i' P_i' \left[I_T - P_i e (e' \Phi_i^{-1} e)^{-1} e P_i' \right] P_i y_i \right\} \end{aligned}$$

Proposition 6.2. *L'estimateur BLUE de la moyenne des coefficients aléatoires β_i , $i = 1, \dots, N$ du modèle (6.1) à coefficients fixes et aléatoires (MFR) :*

$$\begin{aligned} \hat{\beta}_{MFR} &= \left[\sum_{i=1}^N X_i' \Phi_i^{-1} X_i - \sum_{i=1}^N X_i' \Phi_i^{-1} e (e' \Phi_i^{-1} e)^{-1} e \Phi_i^{-1} X_i \right]^{-1} \\ &\quad \left[\sum_{i=1}^N X_i' \Phi_i^{-1} y_i - \sum_{i=1}^N x_i' \Phi_i^{-1} e (e' \Phi_i^{-1} e)^{-1} e \Phi_i^{-1} y_i \right] \end{aligned} \quad (6.14)$$

Un estimateur convergent de la matrice de variance covariance de ces paramètres est alors défini par (Swamy 1970) :

$$\hat{\Delta}_{(K,K)} = \frac{1}{N-1} \sum_{i=1}^N \left[\left(\hat{\beta}_i - \frac{1}{N} \sum_{i=1}^N \hat{\beta}_i \right) \left(\hat{\beta}_i - \frac{1}{N} \sum_{i=1}^N \hat{\beta}_i \right)' \right] \quad (6.15)$$

où les estimateurs $\hat{\beta}_i, \forall i \in [1, N]$, sont les estimateurs des MCO obtenus équation par équation pour chaque individu.

On retrouve ici la formule générale de l'estimateur de la moyenne des coefficients aléatoires β_i dans un modèle *MFR*. Cette formule diffère légèrement de celle proposée dans l'article de Hsiao (1989)¹⁴, dans le sens où nous nous sommes limités ici à un modèle avec uniquement des effets individuels fixes. Dans l'article de Hsiao, l'auteur introduit non pas des effets individuels, mais des variables dichotomiques individuelles différentes selon les mois de l'année et selon les individus.

6.2. Procédure d'estimation des paramètres du modèle MFR

La procédure générale permettant d'obtenir les estimateurs de l'espérance et de la variance des coefficients aléatoires du modèle de Hsiao est la suivante :

1. **Etape 1** : on estime le modèle (6.1) pour chaque individu, en séries temporelles. Soient $\hat{\beta}_i$ et $\hat{\sigma}_{v_i}^2$ les estimateurs des MCO des paramètres β_i et de la variance des résidus $\forall i \in [1, N]$:

$$y_{i,t} = \hat{\alpha}_i + \hat{\beta}_i' x_{i,t} + \hat{v}_{i,t} \quad (6.16)$$

$$\hat{\sigma}_{v_i}^2 = \frac{1}{T-K} \sum_{t=1}^T \hat{v}_{i,t}^2 \quad (6.17)$$

2. **Etape 2** : on construit l'estimateur de Swamy (1970) de la matrice de variance covariance des paramètres β_i , noté $\hat{\Delta}$, selon la formule :

$$\hat{\Delta}_{(K,K)} = \frac{1}{N-1} \sum_{i=1}^N \left[\left(\hat{\beta}_i - \frac{1}{N} \sum_{i=1}^N \hat{\beta}_i \right) \left(\hat{\beta}_i - \frac{1}{N} \sum_{i=1}^N \hat{\beta}_i \right)' \right] \quad (6.18)$$

3. **Etape 3** : on construit l'estimateur de la moyenne β des paramètres aléatoires β_i de la façon suivante :

$$\hat{\beta}_{MFR} = \left[\sum_{i=1}^N X_i' \hat{\Phi}_i^{-1} X_i - \sum_{i=1}^N X_i' \hat{\Phi}_i^{-1} e \left(e' \hat{\Phi}_i^{-1} e \right)^{-1} e \hat{\Phi}_i^{-1} X_i \right]^{-1} \left[\sum_{i=1}^N X_i' \hat{\Phi}_i^{-1} y_i - \sum_{i=1}^N x_i' \hat{\Phi}_i^{-1} e \left(e' \hat{\Phi}_i^{-1} e \right)^{-1} e \hat{\Phi}_i^{-1} y_i \right] \quad (6.19)$$

où e désigne le vecteur unitaire de dimension $(T, 1)$ et la matrice $\hat{\Phi}_i, \forall i \in [1, N]$ est définie par :

$$\hat{\Phi}_i = X_i \hat{\Delta} X_i' + \hat{\sigma}_{v_i}^2 I_T \quad (6.20)$$

¹⁴Formule (3.3) page 578.

A la fin de ces trois étapes on dispose d'un estimateur $\hat{\beta}_{MFR}$ du vecteur des moyennes des paramètres β_i et d'un estimateur $\hat{\Delta}$ de la matrice de variance covariance associée. On peut alors construire les statistiques de Student.

6.3. Une application

Considérons une application de la technique d'estimation d'un modèle *MFR* à la problématique de la convergence proposée dans Gaulier, Hurlin et Jean-Pierre (2000). On cherche à estimer une équation de convergence conditionnelle du type :

$$\Delta(y_{i,t} - \bar{y}_t) = \alpha_i + \rho_i (y_{i,t-1} - \bar{y}_{t-1}) + \varepsilon_{i,t} \quad (6.21)$$

avec $\forall t \in [1, T]$

$$\bar{y}_t = \frac{1}{N} \sum_{i=1}^N y_{i,t} \quad (6.22)$$

Ce modèle de convergence est construit à partir de l'écart du PIB par tête $y_{i,t}$ à la moyenne internationale \bar{y}_t des PIB par tête. On cherche à déterminer si le taux de croissance de cet écart est lié négativement au niveau de l'écart observé à la période précédente. Si pour tous les pays, les constantes α_i sont nulles et *les coefficients ρ_i sont négatifs cela signifie que l'on observe un phénomène de convergence des PIB par tête.* Dans ce cas, un écart positif du PIB par tête à la moyenne internationale induit une diminution de la croissance de cet écart à la date suivante. Autrement dit les écarts des PIB par tête de tous les pays à la moyenne internationale tendent à se resserrer et à par là même à disparaître. On dit alors qu'il y a *convergence absolue des niveaux de PIB par tête.*

$$\Delta(y_{i,t} - \bar{y}_t) = \rho_i (y_{i,t-1} - \bar{y}_{t-1}) + \varepsilon_{i,t}$$

ce qui implique que :

$$\lim_{p \rightarrow \infty} (y_{i,t+p} - \bar{y}_{t+p}) = 0 \quad \forall i \in [1, N] \quad (6.23)$$

Mais de tels phénomènes sont rarement observés sur des échantillons internationaux, c'est pourquoi généralement on introduit des effets individuels α_i dans la spécifications des modèles. Ces effets individuels permettent d'envisager la notion de *convergence conditionnelle*. Supposons en effet que les constantes α_i soient non nulles, dans ce cas on montre que :

$$\lim_{p \rightarrow \infty} (y_{i,t+p} - \bar{y}_{t+p}) = \alpha_i \quad \forall i \in [1, N] \quad (6.24)$$

Les PIB par tête des différents pays convergent ainsi vers des sentiers réguliers parallèles mais non confondues : les taux de croissance sont identiques mais les niveaux demeurent différents.

Dans cette étude, nous montrons à l'aide de tests de non stationnarité en panel (Evans et Karras 1996), qu'il existe un processus de convergence conditionnel sur un échantillon de 27 pays de l'OCDE (1960-1990). Pour cet échantillon, nous proposons alors *une estimation par la méthode MFR de la moyenne de la vitesse de convergence des PIB par tête*. On considère donc le modèle suivant :

$$\Delta(y_{i,t} - \bar{y}_t) = \alpha_i + \rho_i(y_{i,t-1} - \bar{y}_{t-1}) + \gamma_i \Delta(y_{i,t-1} - \bar{y}_{t-1}) + \varepsilon_{i,t} \quad (6.25)$$

Tout comme en séries temporelles, on introduit ici un terme de type *ADF* $\Delta(y_{i,t-1} - \bar{y}_{t-1})$ de façon à blanchir la structure des résidus. On suppose que les effets individuels α_i sont fixes et que les paramètres ρ_i et γ_i sont des coefficients aléatoires. On a donc une structure de type *MFR* identique à celle présentée précédemment avec $N = 27$, $T = 31$ et $K = 2$. On s'intéresse ici plus particulièrement aux vitesses de convergence ρ_i dont la distribution, supposée commune, à une espérance ρ et une variance σ_ρ^2 . On pose ainsi :

$$\beta_i = \begin{pmatrix} \rho_i \\ \gamma_i \end{pmatrix}_{(2,1)} \quad \forall i \in [1, N] \quad (6.26)$$

$$E(\beta_i) = \beta = \begin{pmatrix} \rho \\ \gamma \end{pmatrix} \quad E[(\beta_i - \beta)(\beta_i - \beta)'] = \Delta = \begin{pmatrix} \sigma_\rho^2 & \sigma_{\rho\gamma} \\ \sigma_{\rho\gamma} & \sigma_\gamma^2 \end{pmatrix} \quad (6.27)$$

Faisons abstraction ici des problèmes liés à la structure dynamique du modèle (*cf.* Hsiao 1989 pour une application avec endogène retardée). Le problème consiste à estimer les paramètres ρ , γ , σ_ρ^2 , σ_γ^2 et $\sigma_{\rho\gamma}$. Le programme TSP qui reprend les 3 étapes de la procédure d'estimation des paramètres du modèle *MFR* est fourni en annexe (A.2).

Tout d'abord, commençons par présenter les résultats des estimations des 27 équations individuelles (*étape 1*). Les résultats figurent dans le tableau (6.1).

Ces résultats ont été stockés dans un fichier Excel grâce aux commandes suivantes :

```
smp1 1 N;
unmake ls_alpha alpha_i;
unmake ls_rho rho_i gam_i;
unmake sigls sig_i;
write(file='dea.xls') alpha_i rho_i gam_i sig_i;
```

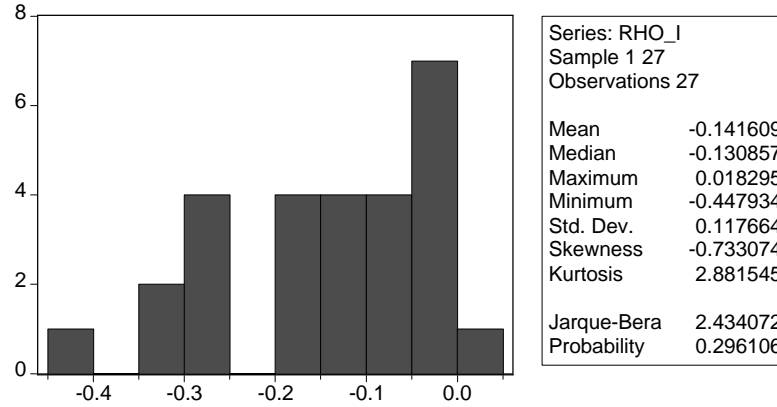
Le fichier Excel a ensuite été importé dans un Workfile sous *Eviews*, ce qui nous permet d'utiliser l'interface graphique, relativement pratique, de ce logiciel pour présenter les résultats. On peut en particulier étudier la distribution des vitesses de convergence

Figure 6.1: Résultat des Estimations Individuelles

Obs	Alpha_l	Rho_l	Gam_l	Sig_l
1	0,13376594	-0,3052974	0,02715862	0,00032369
2	-0,0893829	-0,1493948	0,47335094	0,00137522
3	0,02164255	-0,0482918	0,15210068	0,00047172
4	0,02181379	-0,0713512	0,08023112	0,00058703
5	0,05501872	0,01829461	0,13963646	0,00152846
6	0,03200307	-0,2798161	-0,1547206	0,00014765
7	0,03007381	-0,1644535	-0,2663595	0,00021004
8	-0,0968696	-0,1697266	0,22403599	0,00123583
9	0,01394913	-0,0702855	0,01103817	0,00068135
10	0,0468125	-0,2748415	0,30987847	0,00069222
11	0,00018522	-0,0081573	0,08217304	0,00013755
12	0,03381594	-0,1308573	0,17844738	0,00024927
13	-0,06906	-0,1506083	-0,1468081	0,00074661
14	0,04178007	-0,2517085	0,35072318	0,00173777
15	-0,0555896	-0,176831	0,27316603	0,0007863
16	0,0280554	-0,3224786	0,18986426	0,0003754
17	0,08583123	-0,2657224	0,20064856	0,00104416
18	-0,0022626	-0,0062728	0,19309539	0,00019094
19	0,02389198	-0,0916293	0,51222789	0,00038848
20	-0,0432805	-0,0772415	0,14977539	0,00113732
21	-0,020066	-0,1200527	0,52000779	0,00032102
22	0,00564093	-0,0314211	0,32853037	0,00029894
23	0,01507706	-0,0447503	0,23266649	0,00039859
24	-0,5000802	-0,4479339	0,33779788	0,00123425
25	0,01952289	-0,1447489	0,15997526	0,0006339
26	-0,0074825	-0,0072746	-0,250526	0,00059937
27	-0,0061079	-0,0305862	0,18348967	0,00073227

individuelles ρ_i . À partir de l'histogramme (6.2), on constate que la moyenne simple des réalisations des estimateurs des MCO $\hat{\rho}_i$ est de -0.1416 , et que la dispersion est relativement importante puisque l'écart type des $\hat{\rho}_i$ est de 0.117 . De plus, les statistiques de la Kurtosis, de la Skewness et de Jarque-Bera semble indiquer que la distribution n'est pas normale. En particulier, il semble que la distribution des $\hat{\rho}_i$ soit non symétrique (Kurtosis) et décalée vers la droite par rapport à la moyenne.

Cette première analyse est confirmée par l'examen de l'estimateur à noyau de la densité des $\hat{\rho}_i$ et ce quel que soit le *bandwidth parameter* choisi, comme le montre la figure (6.3). On observe que l'estimateur de la distribution des $\hat{\rho}_i$ n'est pas symétrique autour de la moyenne et tend à être décalé vers la droite. Économiquement, cela signifie que la probabilité que les réalisations individuelles des vitesses de convergence soient supérieures à la moyenne -0.1416 est elle-même supérieure à la probabilité que ces vitesses de convergence soient inférieures à la moyenne. On observe en outre qu'il existe une probabilité non nulle, mais faible, que les paramètres ρ_i soient positifs ou

Figure 6.2: Histogramme des $\hat{\rho}_i$ 

nuls, ce qui traduirait une hypothèse de non convergence. Il serait possible de calculer cette probabilité.

Pour autant jusqu'à présent *nous nous sommes contentés d'étudier la distribution des $\hat{\rho}_i$ sans tenir compte de la précision des estimations individuelles*, c'est à dire des variances des résidus σ_v^2 . C'est pourquoi, nous avons construit l'estimateur MFR construit à partir d'une correction faisant intervenir la matrice de variance des différentes composantes des résidus. Ainsi, dans le programme, nous avons construit l'estimateur de Swamy (1970) de cette matrice de variance covariance (*étape 2 de la procédure générale*). Les résultats obtenus sont les suivants :

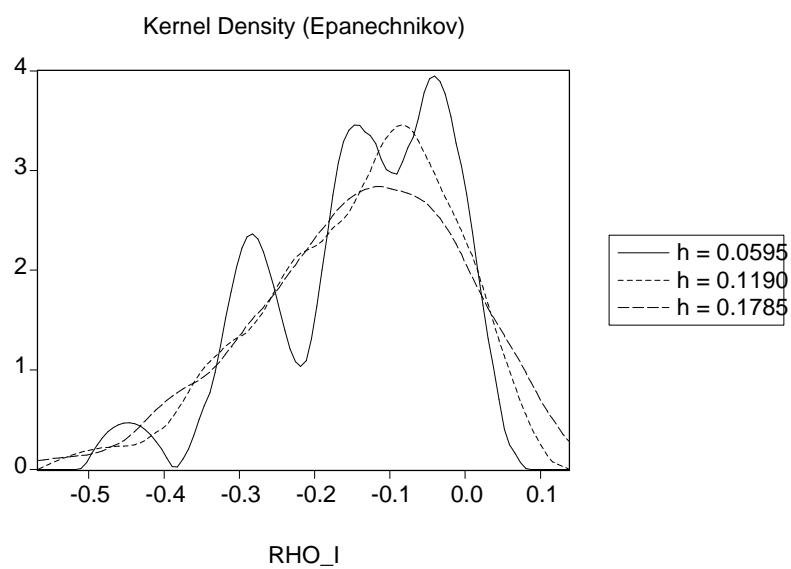
$$\hat{\Delta} = \begin{pmatrix} \hat{\sigma}_\rho^2 & \hat{\sigma}_{\rho\gamma} \\ \hat{\sigma}_{\rho\gamma} & \hat{\sigma}_\gamma^2 \end{pmatrix} = \begin{pmatrix} 0.013845 & -0.0024814 \\ -0.0024814 & 0.042026 \end{pmatrix} \quad (6.28)$$

Bien entendu, on retrouve l'écart type des paramètres ρ_i que nous avons calculé précédemment puisque $\hat{\sigma}_\rho = \sqrt{0.013845} = .11766$. Mais nous tenons compte ici des éventuelles covariances pouvant exister entre les coefficients γ_i et ρ_i .

Enfin, dans une dernière étape, nous avons construit l'estimateur MFR associé aux moyennes des paramètres γ_i et ρ_i (*étape 3 de la procédure générale*). Les résultats obtenus sont les suivants :

$$\hat{\beta} = \begin{pmatrix} \hat{\rho} \\ \hat{\gamma} \end{pmatrix} = \begin{pmatrix} -0.11395 \\ 0.15865 \end{pmatrix} \quad (6.29)$$

On constate ainsi que la réalisation de l'estimateur MFR de la moyenne $\hat{\rho}$ est inférieure à la simple moyenne des $\hat{\beta}_i$ individuels. *Cela signifie que l'estimateur MFR peut être envisagé comme une moyenne pondérée des estimations individuelles $\hat{\beta}_i$, dont*

Figure 6.3: Estimateur Kernel de la Densité des $\hat{\rho}_i$ 

les pondérations sont fonctions de la précisions des estimations individuelles (de la variance des résidus).

A. Annexes

A.1. Equivalence entre les estimateurs Within et pooled dans le cas $\psi = 1$

Revenons à l'équation (4.14) de l'estimateur des MCG :

$$\hat{\beta}_{MCG} = \left[\frac{1}{T} \sum_{i=1}^N X_i' Q X_i + \psi \sum_{i=1}^N (\bar{x}_i - \bar{x}) (\bar{x}_i - \bar{x})' \right]^{-1} \left[\frac{1}{T} \sum_{i=1}^N X_i' Q y_i + \psi \sum_{i=1}^N (\bar{x}_i - \bar{x}) (\bar{y}_i - \bar{y})' \right] \quad (\text{A.1})$$

On suppose pour simplifier qu'il n'existe qu'une variable explicative ($K = 1$) et l'on suppose que $\psi = 1$. L'estimateur des MCG s'écrit alors sous la forme :

$$\hat{\beta}_{MCG} = \left[\frac{1}{T} \sum_{i=1}^N X_i' Q X_i + \sum_{i=1}^N (\bar{x}_i - \bar{x})^2 \right]^{-1} \left[\frac{1}{T} \sum_{i=1}^N X_i' Q y_i + \sum_{i=1}^N (\bar{x}_i - \bar{x}) (\bar{y}_i - \bar{y}) \right] \quad (\text{A.2})$$

Considérons tout d'abord le dénominateur de cette expression :

$$D = \frac{1}{T} \sum_{i=1}^N X_i' Q X_i + \sum_{i=1}^N (\bar{x}_i - \bar{x})^2 = \frac{1}{T} \sum_{i=1}^N \sum_{t=1}^T (x_{i,t} - \bar{x}_i)^2 + \sum_{i=1}^N (\bar{x}_i - \bar{x})^2$$

Développons cette expression, il vient :

$$D = \sum_{i=1}^N \left[\frac{1}{T} \sum_{t=1}^T (x_{i,t} - \bar{x}_i)^2 + (\bar{x}_i - \bar{x})^2 \right] = \sum_{i=1}^N \left[\frac{1}{T} \sum_{t=1}^T x_{i,t}^2 - 2\bar{x}_i \frac{1}{T} \sum_{t=1}^T x_{i,t} + 2\bar{x}_i^2 - 2\bar{x}_i \bar{x} + \bar{x}^2 \right]$$

Rappelons que :

$$\sum_{t=1}^T x_{i,t} = T\bar{x}_i$$

On obtient ainsi :

$$D = \sum_{i=1}^N \left[\frac{1}{T} \sum_{t=1}^T x_{i,t}^2 - 2\bar{x}_i \bar{x} + \bar{x}^2 \right] = \sum_{i=1}^N \sum_{t=1}^T (x_{i,t} - \bar{x})^2$$

De la même façon, on montre que le numérateur peut s'écrire sous la forme :

$$N = \sum_{i=1}^N \sum_{t=1}^T (x_{i,t} - \bar{x}) (y_{i,t} - \bar{y})$$

Ainsi, on retrouve l'expression du l'estimateur du modèle *pooled* :

$$\hat{\beta}_{MCG} = \hat{\beta}_{pooled} = \left[\sum_{i=1}^N \sum_{t=1}^T (x_{i,t} - \bar{x})^2 \right]^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T (x_{i,t} - \bar{x}) (y_{i,t} - \bar{y}) \right]$$

Le même résultat peut être démontré dans le cas général ($K \geq 1$).

A.2. Programme TSP : Modèle MFR

Le programme TSP ci-dessous reprend les 3 étapes de la procédure d'estimation des paramètres du modèle *MFR* appliqué à la problématique de convergence de Hurlin, Gaulier et Jean Pierre (2000). Toutes les lignes de commandes en italique doivent être modifiées dans le cas d'une autre application.

```
?*****
?**** ESTIMATEUR MFR *****
?**** HSIAO (1989) *****
?*****
?--- Options -----
OPTIONS(LIMWARN=0,LIMERR=10);
SUPRES SMPL;
?-----

?*****
?**** Paramètres du Modèle ****
?*****
set N=27; ? Nombre d'Individus N
set l=31; ? Nombre de Périodes T
set K=2; ? Nombre de Variables Explicatives K
set nobs=N*l; ? Nombre Total d'Observations NT
set t0=60; ? Date de Début d'Echantillon
set t1=90; ? Date de Fin d'Echantillon
set retard=2; ? Nombre de Retards (Modèle Dynamique)
smpl 1 nobs;

?*****
?**** Variables *****
?*****
?— Base de données ———
load(file="ocde.wks");
?— Données Centrées sur la Moyenne Internationale —
y=log(y);
do j=t0 to t1;
    select (t=j);
    msd(noprint) Y;
    YB=Y-@MEAN;
```

```

enddo;
?— Construction des Séries —
smpl 1 nobs;
y=yb;
select i=i(-1);
dy=y-y(-1);
?*****

?-----
?---- ETAPE 1 : Estimations pays par pays ----
?-----

mform(nrow=N,ncol=1) ls_alpha=0;
mform(nrow=N,ncol=K) ls_rho=0;
mform(nrow=N,ncol=1) sigls=0;
do j=1 to N;
    ?--- Régressions Individuelles ---
    smpl 1 nobs;
    select (i=j) & (i=i(-2));
    ls(silent) dy c y(-1) dy(-1);
    ?--- Construction des Vecteurs des Coefficients ---
    set ls_alpha(j,1)=@coef(1);
    do p=1 to K;
        set pf=p+1;
        set ls_rho(j,p)=@coef(pf);
    enddo;
    set sigls(j,1)=@s2;
enddo;
print ls_alpha ls_rho sigls;

?-----
?---- ETAPE 2 : Swamy (1970) ----
?---- Estimateur de la variance du parametre rho ----
?-----

?--- Calcul de la Moyenne des Coefficients Alatoires ---
mform(nrow=N,ncol=1) en=1;
mat sum_rho=((en')*ls_rho)/N;
?--- Construction du Vecteur des Moyennes ---
if K=1; then; do;

```



```

        mat vsum_rho=sum_rho*en;
    enddo;
    if K>1; then; do;
        mat vsum_rho=sum_rho(1,1)*en;
        do j=2 to K;
            mat sum_rhoi=sum_rho(1,j)*en;
            mmake vsum_rho vsum_rho sum_rhoi;
        enddo;
    enddo;
    ?--- Construction du l'estimateur de Delta ---
    mat Delta=(1/(N-1))*((ls_rho-vsum_rho)')*(ls_rho-vsum_rho);

    ?-----
    ?-- ETAPE 3 : Construction de Beta ---
    ?-- Construction des Phi_i ---
    ?-----

    smpl 1 nob;
    set l1=l-retard;
    mform(nrow=l1,ncol=l1,type=diag) unite=1;
    mform(nrow=l1,ncol=1) z=1;
    mform(nrow=K,ncol=K) T1=0;
    mat T2=T1;
    mform(nrow=K,ncol=1) T3=0;
    mat T4=T3;

    do j=1 to N;
        smpl 1 nob;
        ?*****
        ?*** Construction des séries ***
        ?*****
        select (i=j)ℓj(i=i(-2));
        y1=y(-1);
        dy1=dy(-1);
        mmake x y1 dy1;
        mmake yy dy;
        ?-- Construction de la matrice (Phi)-1
        mform(nrow=l1,ncol=l1) Phi=0;
        mat Phi=x*Delta*x'+sigls(j,1)*unite;
    enddo;

```

```

    mat Phi=(Phi)'';
    ?-- Premier Terme :  T1 ---
    mat t1_i=(x')*Phi*x;
    mat T1=T1+t1_i;
    ?-- Deuxime Terme :  T2 ---
    mat t2_i=(x')*Phi*z*(((z')*Phi*z)')*(z')*Phi*x;
    mat T2=T2+t2_i;
    ?-- Troisième Terme :  T3 ---
    mat t3_i=(x')*Phi*yy;
    mat T3=T3+t3_i;

    ?-- Quatrième Terme :  T4 ---
    mat t4_i=(x')*Phi*z*(((z')*Phi*z)')*(z')*Phi*yy ;
    mat T4=T4+t4_i;
    smpl 1 nobs;
enddo;

?-----
?-- Estimateur de rho_bar -----
?-----
mat rhols_m=(sum_rho)';
mat rho_mfr=((T1-T2)')*(T3-T4);
print rhols_m rho_mfr;
print delta;

?-----
?--- Calcul des T-stats ---
?-----
if K=1; then; do;
    set T_stat=rho_mfr/sqrt(Delta);
enddo;
if K>1; then; do;
    mform(nrow=K,ncol=1) T_stat=0;
    do j=1 to K;
        set T_stat(j,1)=rho_mfr(j,1)/sqrt(Delta(j,j));
    enddo;
enddo;
print t_stat;

```

Bibliographie

Ahn, S.C., et Schmidt, P., (1995), "Efficient Estimation of Models for Dynamic Panel Data", *Journal of Econometrics*, 68, 5-27.

Bernard, A., et Jones, C., (1996), "Productivity Across Industries and Countries : Times Series Theory and Evidence", *The Review of Economics and Statistics*, 135-146.

Bernard A.B., et Durlauf S.N., (1995), "Convergence of International Output", *Journal of Applied Econometrics*, 10, 97-108.

Blundell, R., et Schmidt, R., (1991), "Conditions Initiales et Estimation Efficace dans les Modèles Dynamiques sur Données de Panel : Une Application au Comportement d'Investissement des Entreprises", *Annales d'Economie et de Statistique*, 20/21, 109-123.

Gaulier G, Hurlin C et Jean-Pierre P. (1999), "Testing Convergence : A Panel Data Approach", *Annales d'Economie et de Statistiques*, 55-56, 411-427.

Harris, R.D.F. et Tzavalis, E., (1996), "Inference for Unit Root in Dynamic Panels", *Working Paper, University of Exeter, UK*.

Hausman J.A., (1978) "Specification Tests in Econometrics", *Econometrica*, 46, 1251-1271

Hsiao, C., (1986), "Analysis of Panel Data", *Econometric society Monographs N0 11*. Cambridge University Press.

Hsiao C. (1989), "Modelling Ontario Regional Electricity System Demand Using a Mixed Fixed and Random Coefficient Approach", *Regional Science and Urban Economics*, 19, 565-587

Im, K.S., Pesaran, M.H., et Shin, Y., (1995), "Testing for Unit Roots in Heterogeneous Panels", *DAE, Working Paper 9526*, University of Cambridge.

Islam, N., (1995), "Growth Empirics : A Panel Data Approach", *Quarterly Journal of Economics*.

Levin, A., et Lin, C-F., (1992), "Unit Root Test in Panel Data: Asymptotic and Finite-Sample Properties", *Discussion Paper 92-23, Department of Economics*, University of California, San Diego.

Quah, D., (1994), "Exploiting Cross-Section Variations for Unit Root Inference in Dynamic Data", *Economics Letters*, 44, 9-19.

Mundlak Y. (1978), "On the Pooling of Time Series and Cross Section Data", *Econometrica*, 46, 69-85

Nickell S. (1981), "Biases in Dynamic Models with Fixed Effects", *Econometrica*, 49, 1399-1416

Sevestre, P., et Trognon, A., (1992), "Linear Dynamic Models", in *Matyas and P. Sevestre (eds)*, Ch. 6., 94-116.

Swamy P.A. (1970), "Efficient Inference in a Random Coefficient Regression Model", *Econometrica*, 38, 311-323

Taylor W.E. (1980), "Small Sample Consideration in Estimation from Panel Data", *Journal of Econometrics*, 13, 203-223