



**HAL**  
open science

# Une synthèse des tests de cointégration sur données de panel

Christophe Hurlin, Valérie Mignon

► **To cite this version:**

Christophe Hurlin, Valérie Mignon. Une synthèse des tests de cointégration sur données de panel. 2006. halshs-00070887

**HAL Id: halshs-00070887**

**<https://shs.hal.science/halshs-00070887v1>**

Preprint submitted on 22 May 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



*Laboratoire d'Economie d'Orléans*

## **Document de Recherche**

**n° 2006-12**

**"Une synthèse des tests de cointégration sur données de panel"**

**Christophe HURLIN**  
**(en collaboration avec Valérie MIGNON)**

# Une synthèse des tests de cointégration sur données de panel

Christophe Hurlin\* et Valérie Mignon†

24 avril 2006

## Résumé

L'objet de ce papier est de dresser un panorama complet de la littérature relative aux tests de cointégration sur données de panel. Après un exposé des concepts spécifiques à la cointégration en panel, sont ainsi présentés les tests de l'hypothèse nulle d'absence de cointégration (tests de Pedroni (1995, 1997, 1999, 2003), Kao (1999), Bai et Ng (2001) et Groen et Kleibergen (2003)) ainsi que le test de McCoskey et Kao (1998) reposant sur l'hypothèse nulle de cointégration. Quelques éléments relatifs à l'inférence et l'estimation de systèmes cointégrés sont également fournis.

**Mots-clé** : données de panel non stationnaires, racine unitaire, cointégration.

## Panel data cointegration tests : A survey

The purpose of this paper is to provide a complete overview of panel data cointegration tests. We present the main tests based on the null hypothesis of no cointegration (tests of Pedroni (1995, 1997, 1999, 2003), Kao (1999), Bai and Ng (2001) and Groen and Kleibergen (2003)) and the McCoskey and Kao (1998) test based on the null hypothesis of cointegration. Some developments relating to inference and estimation of cointegrated systems are also provided.

**Keywords** : non-stationary panel data, unit root, cointegration.

**J.E.L. Classification** : C23, C33.

---

\*LEO, Université d'Orléans. Faculté de Droit, d'Economie et de Gestion. Rue de Blois - BP6739. 45067 Orléans Cedex 2. Email : christophe.hurlin@univ-orleans.fr.

†Université Paris X - Nanterre et CEPII. 200 Avenue de la République, 92001 Nanterre Cedex. Tel : 01.40.97.58.60. Email : Valerie.Mignon@u-paris10.fr

# 1 Introduction

Depuis les travaux pionniers de Levin et Lin (1992), la littérature relative à l'économétrie des données de panel non stationnaires — et, en particulier, aux tests de cointégration<sup>1</sup> — connaît un développement considérable<sup>2</sup>. Outre l'abondance des développements théoriques, on recense également de très nombreux travaux empiriques sur données de panel ayant recours aux tests de racine unitaire, dans un contexte univarié, et/ou aux tests de cointégration dans un contexte multivarié. Un grand nombre de ces travaux empiriques porte sur la vérification de l'hypothèse de parité des pouvoirs d'achat (PPA) dans le domaine des changes. Ainsi, si l'on s'intéresse à la version absolue de la PPA, le taux de change réel  $q_{it}$  du pays  $i$  s'écrit :

$$q_{it} = e_{it} + p_t^* - p_{it} \quad (1)$$

où  $e_{it}$  est le taux de change nominal,  $p_t^*$  est l'indice des prix à la consommation étranger et  $p_{it}$  est l'indice des prix à la consommation du pays  $i$ ; toutes les variables étant en logarithmes. Tester la version absolue de la PPA revient à tester l'hypothèse nulle de racine unitaire dans la série  $q_{it}$ . Il est également possible de tester la version relative de la PPA en estimant l'équation :

$$q_{it} = e_{it} + \alpha p_t^* - \beta p_{it} \quad (2)$$

Tester la forme relative de la PPA revient alors à effectuer un test de cointégration entre le change nominal et les indices de prix. Notons que l'on peut également utiliser la relation (2) pour tester la version absolue de la PPA en testant l'égalité à 1 du vecteur de cointégration. Le nombre de travaux empiriques visant à tester la PPA sur données de panel est impressionnant et l'on peut citer, de façon non exhaustive, les études de Frankel et Rose (1995), Wei et Parsley (1995), Bernard et Jones (1996), Canzoneri, Cumby et Diba (1996), Chinn et Johnston (1996), Jorion et Sweeney (1996), McDonald (1996), Obstfeld et Taylor (1996), Oh (1996), Rogoff (1996), Taylor (1996), Wu (1996), Papell (1997), Anthony et McDonald (1998), O'Connell (1998), Alberola *et al.* (1999), Bayoumi et McDonald (1999), Pedroni (1999, 2001, 2003), McDonald et Nagayasu (2000), Chortareas et Driver (2000), Papell et Theodoridis (2001), Camarero, Ordóñez et Tamarit (2002), Camarero et Tamarit (2002), Papell (2002), Chang (2002) ou encore Drine et Rault (2003)<sup>3</sup>.

---

<sup>1</sup>On rappelle que des séries intégrées d'ordre  $d$  (où  $d$  est un entier positif), ce que l'on note  $I(d)$ , sont cointégrées s'il existe une combinaison linéaire de ces séries intégrée d'ordre  $d - b$ , avec  $0 < b \leq d$ .

<sup>2</sup>Pour une revue de la littérature, voir notamment Banerjee (1999), Phillips et Moon (2000), Baltagi et Kao (2000) ou encore Hurlin et Mignon (2006).

<sup>3</sup>Notons que, même si la très grande majorité des études sur données de panel porte sur l'analyse des taux de change, il existe cependant des travaux dans d'autres domaines. A titre d'exemple, McCoskey et Selden (1998) et Gerdrham et Löthgren (1998) se sont attachés à l'étude des dépenses de santé et Breitung et Mayer (1994) à l'analyse des salaires. Islam (1995), Evans et Karas (1966), Sala-i-Martin (1996), Lee, Pesaran et Smith (1997),

Comment expliquer un tel succès des études sur données de panel ? En d'autres termes, quel est l'apport lié à l'ajout de la dimension individuelle à la dimension temporelle usuelle ? Il est bien connu que les tests de cointégration sur séries temporelles sont relativement peu puissants en petit échantillon pour distinguer des séries non stationnaires de séries stationnaires mais fortement persistantes. Or, c'est typiquement le cas lorsque l'on étudie des séries de taux de change sur la période post Bretton-Woods, pour lesquelles on dispose généralement de 20 ou 30 ans de données. L'étendue de la période d'étude étant plus importante que la fréquence des données (voir notamment Shiller et Perron, 1985 ; Perron, 1989, 1991 et Pierce et Snell, 1995), une solution consisterait — lorsque cela est possible — à accroître cette période. Mais, d'une part, cela est rarement possible et, d'autre part, une telle procédure augmente le risque de faire face à des ruptures structurelles. Une alternative consiste alors à accroître le nombre de données en incluant l'information relative à des pays différents et à mener ainsi une analyse multi-pays. Il est en effet naturel de penser que les propriétés de long terme des séries — telle que l'existence de la PPA — ont une forte probabilité d'être communes à plusieurs pays. Le recours aux données de panel permet d'accroître la puissance des tests de cointégration en petit échantillon en augmentant le nombre de données disponibles. Ainsi, le fait que la PPA soit fréquemment rejetée sur données temporelles, alors qu'elle tend à être réhabilitée sur données de panel, est généralement interprété comme le reflet de la faible puissance des tests sur séries temporelles en petit échantillon. Par ailleurs, un autre avantage issu de l'ajout de la dimension individuelle à la dimension temporelle provient du fait que les distributions asymptotiques des tests de racine unitaire sur données de panel sont asymptotiquement normales, alors qu'elles sont non standard lorsque seule la dimension temporelle est prise en compte.

Notre objet est ici de rendre compte des développements relatifs aux tests de cointégration sur données de panel. Ceux-ci constituent pour l'essentiel des extensions multivariées des tests de racine unitaire développés en univarié. Sans rentrer dans les détails<sup>4</sup> et afin de fixer les idées, mentionnons que, de façon générale, les tests de racine unitaire sur données de panel sont basés sur la régression suivante<sup>5</sup> :

$$\Delta y_{it} = \alpha_i + \rho_i y_{it-1} + \varepsilon_{it} \tag{3}$$

---

Pedroni (1998), Cechetti *et al.* (1999) et McCoskey et Kao (1999b) ont étudié les problèmes de croissance et de convergence. La dynamique de l'épargne et de l'investissement a fait l'objet de travaux de Coakley, Kulasi et Smith (1996) et Moon et Phillips (1998) ou encore Banerjee et Zanghieri (2003) sur le paradoxe de Feldstein-Horioka. Mentionnons également les travaux de Funk (1998), McCoskey et Kao (1999a) et Kao, Chiang et Chen (1999) sur les activités de R&D au niveau international.

<sup>4</sup>Voir les références citées à la note 2.

<sup>5</sup>Il est également possible d'introduire un trend individuel spécifique, de même qu'un terme représentant des effets temporels communs.

où  $i = 1, \dots, N$  désigne l'individu,  $t = 1, \dots, T$  et  $\varepsilon_{it} \sim iid(0, \sigma_i^2)$ . L'hypothèse nulle testée est celle de racine unitaire, soit  $\rho_i = 0 \forall i$ . Les différents tests se distinguent alors par le degré d'hétérogénéité introduit sous l'hypothèse alternative. Les premiers tests de racine unitaire sur données de panel sont dus à Quah (1992, 1994) et Levin et Lin (1992, 1993). Les tests proposés par Quah, valables lorsque  $N$  et  $T$  tendent vers l'infini au même taux, ne prennent pas en compte la possibilité de dynamiques hétérogènes entre les individus, l'existence d'effets spécifiques individuels et la présence d'autocorrélation dans les séries résiduelles. Levin et Lin (1992, 1993) ont proposé un test plus général autorisant la présence d'effets spécifiques individuels et d'hétérogénéité entre les individus. Ce test suppose également que  $N$  et  $T$  tendent vers l'infini, mais  $T$  augmente plus rapidement de sorte que  $N/T \rightarrow 0$ . Sous l'hypothèse alternative, le coefficient autorégressif est supposé homogène entre les individus, soit  $\rho_i = \rho < 0 \forall i$ . Cette hypothèse étant fortement restrictive, Im, Pesaran et Shin (2003) ont développé une procédure de test alternative basée sur la moyenne des statistiques de test de racine unitaire autorisant la présence de corrélation sérielle résiduelle et de dynamiques hétérogènes<sup>6</sup>. Souhaitant étendre l'approche des tests de racine unitaire au cas multivarié, divers auteurs (voir *infra*) se sont attachés à développer des tests de cointégration sur données de panel.

L'objet de ce papier est ainsi de dresser un panorama aussi complet que possible des développements relatifs aux tests de cointégration sur données de panel. A cette fin, la section 2 commence par présenter un certain nombre de considérations spécifiques à la cointégration en panel. La section 3 traite du cas des régressions fallacieuses en décrivant les conséquences de l'absence de cointégration. Les deux sections suivantes ont pour objet de présenter de façon détaillée les tests existants. La section 4 est ainsi consacrée aux tests reposant sur l'hypothèse nulle d'absence de cointégration, alors que la section 5 traite du cas de l'hypothèse nulle de cointégration. Quelques éléments relatifs à l'inférence et à l'estimation de systèmes cointégrés sont donnés dans la section 6. Enfin, la section 7 conclut le papier.

## 2 Notions de cointégration en panel

Avant de présenter les tests de l'hypothèse de cointégration, il convient de préciser un certain nombre de notions spécifiques à la cointégration en panel. Parmi celles-ci figurent les notions de (i) cointégration intra et inter-individuelles, (ii) d'hétérogénéité ou d'homogénéité des paramètres des relations de cointégration et de la représentation à correction d'erreur (VECM, *Vector Error Correction Model*), (iii) de dépendances inter-individuelles et (iv) de convergence.

---

<sup>6</sup>Voir également Maddala et Wu (1999).

Pour cela, nous considérerons une **représentation VECM dans un système complet**. Une telle représentation autorise notamment l'existence de dynamiques inter-individuelles : le taux de croissance d'une variable endogène observée pour un individu donné peut ainsi dépendre des autres variables endogènes du système observées pour ce même individu (que ce soit au travers des relations de cointégration et/ou de la dynamique de court terme) mais aussi de l'ensemble des variables endogènes observées pour les autres individus du panel.

Formellement, soit  $x_{it} = \left( x_{it}^{(1)}, \dots, x_{it}^{(k)} \right)'$  un vecteur de  $k$  variables<sup>7</sup>  $I(1)$  observées pour un individu  $i = 1, \dots, N$  à la période  $t = 1, \dots, T$ . On note  $x_t = (x'_{1t}, \dots, x'_{Nt})'$  le vecteur empilé de dimension  $(Nk, 1)$  des variables individuelles. On appelle représentation VECM dans un système complet (*full system VECM*) le système suivant (Groen et Kleibergen, 2003) :

$$\Delta x_t = \underset{(Nk,1)}{\delta} + \underset{(Nk,1)}{\theta} t + \underset{(Nk,Nk)}{\Pi} x_{t-1} + \underset{(Nk,Nkp)}{\Gamma} W_t + \underset{(Nk,1)}{\varepsilon_t} \quad (4)$$

Commentons brièvement les différents termes de ce modèle. Les deux premiers éléments correspondent aux composantes déterministes de la représentation VECM. Le vecteur  $\delta = (\delta'_1, \dots, \delta'_N)'$  regroupe les  $Nk$  effets individuels (supposés fixes) associés à chaque individu et à chaque variable endogène. De la même façon, le vecteur  $\theta$  regroupe les coefficients du trend déterministe que l'on suppose *a priori* différents pour chaque variable considérée et pour chaque individu. Le troisième terme du modèle (4) correspond bien évidemment au terme définissant à la fois les forces de rappel et les éventuelles relations de cointégration sur lesquelles nous reviendrons en détail. La matrice  $\Pi$  peut être partitionnée de la façon suivante :

$$\underset{(Nk,Nk)}{\Pi} = \begin{pmatrix} \Pi_{11} & \dots & \Pi_{1N} \\ \dots & \Pi_{ij} & \dots \\ \Pi_{N1} & \dots & \Pi_{NN} \end{pmatrix} \quad (5)$$

où les matrices  $\Pi_{ij}$  sont de dimension  $(k, k)$ . Le quatrième terme  $\Gamma W_t$  de la relation (4) regroupe l'ensemble des éléments liés à la dynamique d'ajustement de court terme. Pour simplifier, on suppose que pour purger toute l'auto-corrélation des résidus du modèle (4), il est nécessaire d'introduire exactement le même nombre  $p$  d'éléments différenciés retardés pour les tous les individus du panel. Ainsi, on pose :

$$\underset{(Nkp,1)}{W_t} = (\Delta x'_{1,t-1}, \dots, \Delta x'_{1,t-p}, \dots, \Delta x'_{N,t-1}, \dots, \Delta x'_{N,t-p}) \quad (6)$$

$$\underset{(Nk,Nkp)}{\Gamma} = \begin{pmatrix} \Gamma_{11} & \dots & \Gamma_{1N} \\ \dots & \Gamma_{ij} & \dots \\ \Gamma_{N1} & \dots & \Gamma_{NN} \end{pmatrix} \quad (7)$$

---

<sup>7</sup>On peut étendre l'analyse à un vecteur  $x_{it}$  incluant à la fois des variables  $I(0)$  et  $I(1)$ , dès lors qu'au moins deux variables sont intégrées d'ordre 1.

où les matrices  $\Gamma_{ij}$  sont de dimension  $(k, kp)$  et où les vecteurs individuels  $\Delta x_{i,t-p}$  de dimension  $(k, 1)$  contiennent les différences premières retardées des  $k$  variables endogènes pour l'individu  $i$ . Enfin, le vecteur  $\varepsilon_t = (\varepsilon'_{1t}, \dots, \varepsilon'_{Nt})'$  comporte les  $N$  vecteurs de résidus individuels  $\varepsilon_{it} = (\varepsilon_{it}^{(1)}, \dots, \varepsilon_{it}^{(k)})'$  associés aux  $k$  résidus de chaque VECM individuel. On suppose que  $\varepsilon_t$  est *i.i.d.*  $(0, \Omega)$ . L'éventuelle dépendance inter-individuelle de ces résidus  $\varepsilon_{it}$  implique que la matrice  $\Omega$  peut être non bloc diagonale :

$$\underset{(Nk, Nk)}{\Omega} = \begin{pmatrix} \Omega_{11} & \dots & \Omega_{1N} \\ \dots & \Omega_{ij} & \dots \\ \Omega_{N1} & \dots & \Omega_{NN} \end{pmatrix} \quad (8)$$

où les matrices  $\Omega_{ij}$  sont de dimension  $(k, k)$ . Rappelons que, pour un individu  $i$ , si la matrice  $\Omega_{ii}$  est non diagonale, cela traduit l'existence de corrélations instantanées des chocs de la forme réduite, identiques à celles que l'on considère lorsque l'on spécifie un VECM en séries temporelles.

Dans ce cadre, évoquons à présent la première dimension spécifique de l'analyse de panel, à savoir la distinction entre relations de cointégration intra et inter-individuelles.

## 2.1 Relations de cointégration intra et inter-individuelles

Dans une telle représentation, on peut envisager deux types de relations de cointégration. On peut tout d'abord envisager l'existence de relations de cointégration parmi les variables du vecteur  $x_{it}$ , que l'on pourrait qualifier de **relations de cointégration intra-individuelles**. Tout comme en séries temporelles, on dit qu'il existe une ou plusieurs relations de cointégration dans le vecteur  $x_{it}$  si et seulement si il existe une ou plusieurs combinaisons linéaires des variables  $x_{it}^{(j)}$  qui soient stationnaires. Formellement, pour l'individu  $i$ , il existe  $r_i$  relations de cointégration si et seulement si :

$$\underset{(r_i, k_i)}{\beta'_i} x_{it} = \underset{(r_i, 1)}{\mu_{it}} \sim I(0) \quad (9)$$

où  $\beta_i$  désigne la matrice contenant les vecteurs de cointégration pour l'individu  $i$ . Ces  $r_i$  relations de cointégration ne font intervenir que des variables propres à l'individu  $i$  de référence. Par exemple, l'existence d'une telle relation de cointégration intra-individuelle entre le PIB et la consommation dans un panel international impliquerait que pour chaque pays il existe une combinaison linéaire du PIB et de la consommation du pays considéré qui soit stationnaire. Comme nous le verrons par la suite, la question se pose ensuite de l'homogénéité ou au contraire de l'hétérogénéité des paramètres de cette relation de cointégration pour tous les pays de l'échantillon, *i.e.* la question de savoir si  $\beta_i = \beta, \forall i = 1, \dots, N$ . Mais, on peut parallèlement



supposer qu'il existe une ou plusieurs relations de cointégration faisant intervenir les PIB ou les consommations de différents pays. On parle alors de relations de cointégration entre individus ou de **relations de cointégration inter-individuelles** (*cross unit cointegration*).

Formellement, si l'on souhaite tester (ou estimer) à la fois des relations de cointégration entre des variables différentes observées pour un même individu et des relations de cointégration faisant intervenir des variables observées sur plusieurs individus, la solution consiste à considérer le vecteur empilé  $x_t$  des variables individuelles. Pour simplifier, supposons que l'on considère le même nombre de variables endogènes pour tous les individus, *i.e.*  $k_i = k$ . On dit qu'il existe au total  $r$  relations de cointégration intra et/ou inter individuelles si et seulement si :

$$\begin{matrix} \beta' \\ (r, Nk) \end{matrix} \begin{matrix} x_t \\ (Nk, 1) \end{matrix} = \begin{matrix} \mu_t \\ (r, 1) \end{matrix} \sim I(0) \quad (10)$$

La matrice  $\Pi$  de la représentation VECM (équation 4) se décompose alors sous la forme usuelle :

$$\begin{matrix} \Pi \\ (Nk, Nk) \end{matrix} = \begin{matrix} \alpha \\ (Nk, r) \end{matrix} \begin{matrix} \beta' \\ (r, Nk) \end{matrix} \quad (11)$$

où la matrice  $\alpha$  désigne la matrice des coefficients d'ajustement aux relations de long terme. L'existence d'au moins une relation de cointégration inter-individuelle implique qu'une des colonnes de la matrice  $\beta$  possède des éléments non nuls associés à des éléments  $x_{it}^{(k)}$  et  $x_{jt}^{(k')}$  pour  $i \neq j$ .

Naturellement, l'approche qui consiste à tester la cointégration en prenant en compte les éventuelles relations inter-individuelles apparaît comme la plus séduisante dans de nombreuses problématiques économiques. Une telle approche permettrait en effet par exemple de définir des tendances stochastiques au niveau international, régional, *etc.*, suivant le panel considéré. Malheureusement, une telle approche en système "complet" n'est possible sur le plan technique que pour des panels de dimension individuelle relativement modérée. Par exemple, l'estimation d'un simple modèle VAR d'ordre 1 en système complet implique d'estimer  $k^2 N^2$  paramètres à partir de  $NT$  observations. Dès lors, une telle approche n'est techniquement possible que si l'on considère un panel où la dimension temporelle  $T$  est largement supérieure à la dimension individuelle  $N$ , ce qui n'est généralement pas le cas dans les panels macroéconomiques. Ainsi, pour ces raisons, rares sont les modèles proposés qui considèrent un système complet (*full system representation*) autorisant d'éventuelles relations de cointégration inter-individuelles. Une exception notable étant le modèle VECM proposé par Groen et Kleibergen (2003).

## 2.2 Hétérogénéité / homogénéité des relations de cointégration

Une autre problématique spécifique à la cointégration en panel est celle de l'homogénéité ou de l'hétérogénéité des paramètres de la relation de long terme et plus généralement de la représentation VECM. Commençons par évoquer le problème de l'homogénéité de la relation de long terme. Admettons pour cela que l'on exclut l'existence de relations linéaires entre les variables  $I(1)$  appartenant à des individus différents, et que, par conséquent, l'on exclut l'existence de relations de cointégration inter-individuelles. Formellement, cela revient à imposer la nullité des matrices  $\Pi_{ij}$  pour  $i \neq j$ . Supposons également que pour chaque individu parmi les  $k$  variables endogènes  $I(1)$  il existe  $r_i$  relations de cointégration intra-individuelles. La matrice  $\Pi$  de la représentation VECM est alors bloc diagonale et peut s'écrire sous la forme :

$$\Pi_{(Nk, Nk)} = \begin{pmatrix} \Pi_{11} & \dots & \mathbf{0} \\ \dots & \Pi_{ii} & \dots \\ \mathbf{0} & \dots & \Pi_{NN} \end{pmatrix} = \begin{pmatrix} \alpha_1 & \beta'_1 & \dots & \mathbf{0} \\ (k, r_1) & (r_1, k) & & (k, k) \\ \dots & \alpha_i & \beta'_i & \dots \\ & (k, r_i) & (r_i, k) & \\ \mathbf{0} & \dots & \alpha_N & \beta'_N \\ (k, k) & & (k, r_N) & (r_N, k) \end{pmatrix} \quad (12)$$

où  $\beta_i$  désigne la matrice des vecteurs de cointégration de l'individu  $i$  et  $\alpha_i$  est la matrice des coefficients d'ajustement à ces relations de long terme pour l'individu  $i$ . Si  $\beta_i = \beta, \forall i = 1, \dots, N$  on parle alors de **relations de cointégration homogènes**, dans le cas contraire on parle de **relations de cointégration hétérogènes** (Phillips et Moon, 1999). L'hypothèse d'une relation de cointégration homogène est parfois l'expression d'une contrainte théorique, comme par exemple celle issue de la parité des pouvoirs d'achat. Dans ce cas, la relation de long terme est la même pour tous les pays, tandis que les forces de rappel à cette relation de long terme, *i.e.* les  $\alpha_i$ , peuvent être différentes d'un pays à l'autre. En revanche, dans d'autres cas, cette hypothèse peut être plus discutable comme dans le cas de l'estimation d'une fonction de production de long terme sur un panel international<sup>8</sup> (Canning et Pedroni, 1999) et cela même si les ajustements à cette structure productive homogène de long terme sont eux supposés hétérogènes.

Plusieurs remarques s'imposent ici. Tout d'abord, l'hypothèse de relation de cointégration homogène implique naturellement que les vecteurs de cointégration aient la même dimension pour tous les individus, *i.e.*  $\dim(\beta_i) = (k, r), \forall i = 1, \dots, N$ . Autrement dit, un modèle de cointégration homogène<sup>9</sup> implique que les rangs de cointégration soient les mêmes pour tous les individus, *i.e.*  $r_i = r, \forall i = 1, \dots, N$ . En revanche, dans un modèle de cointégration hétérogène,

<sup>8</sup>Rien ne justifie *a priori* que tous les pays tendent vers la même relation technologique à long terme. Cette hypothèse mériterait à tout le moins d'être testée.

<sup>9</sup>Par abus de langage, nous parlons ici de modèle de cointégration homogène, mais rappelons que l'homogénéité des relations de cointégration n'implique pas l'homogénéité de la représentation VECM et donc du panel.

rien n'impose *a priori* que le rang de cointégration pour l'individu  $i$  soit identique à celui de l'individu  $j$ . On peut donc avoir par exemple des individus pour lesquels il existe des relations de cointégration parmi les  $k$  variables endogènes et d'autres pour lesquels il n'en existe pas. A l'extrême limite, on pourrait concevoir un modèle de panel de cointégration hétérogène où n'existeraient des relations de cointégration parmi les  $k$  variables que pour un seul individu. Toutefois, en l'état actuel de la recherche, les tests de cointégration et les représentations VECM autorisant des relations de cointégration hétérogènes (Groen et Kleibergen, 2003) imposent généralement, eux aussi, une **contrainte d'homogénéité sur le rang de cointégration**. Ainsi, dans le cas d'un panel à relation de cointégration hétérogène, on suppose que tous les individus ont le même nombre de relations de cointégration mais que les paramètres des vecteurs de cointégration peuvent être différents d'un individu à l'autre, *i.e.*  $r_i = r$  et  $\exists (i, j) | \beta_i \neq \beta_j$ .

La deuxième remarque porte sur l'estimation des paramètres d'une relation de cointégration hétérogène. Phillips et Moon (1999) montrent que dans ce cas, l'estimateur *pooled* obtenu en estimant (à tort) un modèle totalement homogène converge à la vitesse  $\sqrt{N}$  vers une quantité définie comme un coefficient moyen de long terme (*long run average coefficient*). Les propriétés asymptotiques de l'estimateur *pooled* dans le cas d'une cointégration hétérogène (tant dans la vitesse de convergence que la forme du coefficient moyen de long terme) sont donc similaires à celles que l'on obtient dans le cas d'une régression fallacieuse. Notons que ce coefficient moyen de long terme vers lequel converge l'estimateur *pooled* ne correspond pas (sauf cas particulier) à la moyenne des  $N$  vecteurs individuels de cointégration  $\bar{\beta} = (1/N) \sum_{i=1}^N \beta_i$ .

Au-delà de cette discussion sur l'homogénéité et/ou l'hétérogénéité des paramètres des relations de long terme, se pose naturellement le problème de l'hétérogénéité des autres éléments de la représentation VECM (équation 4). Comme nous l'avons évoqué, il est généralement opportun dans un modèle de cointégration homogène de supposer que les forces de rappel sont elles hétérogènes, *i.e.*  $\alpha_i \neq \alpha_j, \forall i \neq j$  : si tous les individus présentent la ou les mêmes relations de long terme, la dynamique d'ajustement à ces relations communes peut être plus ou moins rapide suivant les individus. De la même façon, les paramètres  $\Gamma_{ii}$  associés aux termes de la dynamique de court terme peuvent être hétérogènes<sup>10</sup>, *i.e.*  $\Gamma_{ii} \neq \Gamma_{jj}$ . Dès lors qu'au moins la relation de cointégration est homogène, on peut en effet supposer à la fois que les paramètres  $\alpha_i$  et  $\Gamma_{ii}$  sont hétérogènes. Il existe ainsi toute une palette de spécifications possibles de l'hétérogénéité allant du VECM totalement homogène comme l'avaient proposé Holtz-Eakin, Newey

<sup>10</sup>L'absence de dépendances inter-individuelles dans la dynamique de court terme se traduit par l'hypothèse  $\Gamma_{ij} = 0, \forall i \neq j$ .

et Rosen (1988) au modèle VECM où seules les relations de long terme sont homogènes (Groen et Kleibergen, 2003). Enfin, les mêmes choix s’offrent quant à la spécification des composantes déterministes : effets individuels  $\delta_i$  et coefficients  $\theta$  associés aux trends déterministes. Naturellement, non seulement la prise en compte de ces composantes (ou leur exclusion) modifie la distribution asymptotique des statistiques de test de l’hypothèse de cointégration (comme en séries temporelles dans le cas des tests de Johansen, 1991), mais en outre l’hypothèse d’homogénéité ou d’hétérogénéité de ces composantes importe.

### 2.3 Dépendances inter-individuelles

Tout comme dans le cas des tests de racine unitaire, les premiers tests de cointégration proposés en panel (en particulier les tests de Pedroni, 1995, 1997 ; voir *infra*) excluaient *a priori* non seulement l’existence de relations de cointégration inter-individuelles, mais plus généralement l’existence de toute dépendance entre les individus. Naturellement cette hypothèse d’indépendance entre les individus est peu crédible dans de très nombreuses applications empiriques, et notamment dans le cas de panels de variables macroéconomiques. La question se pose alors de savoir comment spécifier l’existence de dépendances entre les individus ?

La première façon évidente de spécifier ces dépendances serait d’envisager explicitement des relations de cointégration inter-individuelles dans un système complet. Si l’on exclut cette solution qui pose plusieurs problèmes techniques, les dépendances entre les individus du panel peuvent être spécifiées de différentes façons. Groen et Kleibergen (2003) envisagent ainsi la possibilité d’une corrélation des résidus des relations de cointégration pour différents individus sans spécifier de forme particulière pour cette relation. Dans ce cas, cela se traduit par une matrice de variance covariance  $\Omega$  dans la représentation VECM (équation 4) non diagonale. Une telle spécification correspond alors à une notion de dépendance contemporaine ou de causalité instantanée inter-individuelle.

Une autre spécification possible des liens entre individus dans les relations de cointégration consiste à utiliser des représentations à facteurs communs (Bai et Ng, 2001 ; Dees, di Mauro, Pesaran et Smith, 2005). Afin de mieux comprendre les enjeux associés à cette question de la spécification des dépendances inter-individuelles, considérons la représentation à facteurs communs du vecteur de variables endogènes  $x_{it} = \left( x_{it}^{(1)}, \dots, x_{it}^{(k)} \right)'$  proposée par Dees, di Mauro, Pesaran et Smith (2005) et Breitung et Pesaran (2005). On admet que  $x_{it}$  peut s’écrire sous la

forme suivante :

$$x_{i,t} = \begin{matrix} \tilde{\delta}_i & + & \tilde{\theta}_i & t & + & \lambda_i & d_t & + & \gamma_i & f_t & + & w_{it} \\ (k,1) & & (k,1) & (1,1) & & (k,s) & (s,1) & & (k,m) & (m,1) & & (k,1) \end{matrix} \quad (13)$$

$$\begin{matrix} \Delta f_t & = & \Lambda(L) & \varepsilon_t & & \varepsilon_t & i.i.d. & (0, I_m) \\ (m,1) & & (m,m) & (m,1) \end{matrix} \quad (14)$$

$$\begin{matrix} \Delta w_{it} & = & \Psi_i(L) & v_{it} & & v_{it} & i.i.d. & (0, I_k) \\ (k,1) & & (k,k) & (k,1) \end{matrix} \quad (15)$$

où  $L$  désigne l'opérateur retard,  $\Lambda(L) = \sum_{j=0}^{\infty} \Lambda^j L^j$  et  $\Psi_i(L) = \sum_{j=0}^{\infty} \Psi_i^j L^j$ , pour  $i = 1, \dots, N$ . Pour un individu  $i$  donné, les  $k$  variables non stationnaires du vecteur  $x_{it}$  dépendent (i) d'une composante déterministe<sup>11</sup> (effet individuel et trend déterministe), (ii) d'un ensemble de  $s$  facteurs communs observables  $d_t$  (variables macroéconomiques, autres composantes déterministes comme des dummies par exemple, *etc.*), (iii) d'un ensemble de  $m$  facteurs communs inobservables  $f_t$  et (iv) de  $k$  composantes stochastiques propres à l'individu de référence, notées  $w_{it}$ . On admet que les facteurs communs inobservables et les composantes individuelles sont intégrés d'ordre un et que les conditions d'application de Beveridge et Nelson sont vérifiées. Si l'on suppose en particulier que  $w_{i0} = f_0 = 0$  et que  $\varepsilon_0 = v_{i0} = 0$ , on obtient alors :

$$f_t = \Lambda(1) \sum_{s=1}^t \varepsilon_s + \Lambda^*(L) \varepsilon_t = \Lambda(1) s_t + \Lambda^*(L) \varepsilon_t \quad (16)$$

$$w_{it} = \Psi_i(1) \sum_{s=1}^t v_{is} + \Psi_i^*(L) v_{it} = \Psi_i(1) s_{it} + \Psi_i^*(L) v_{it} \quad (17)$$

où les polynômes matriciels  $\Lambda^*(L) = [\Lambda(L) - \Lambda(1)] / (1 - L)$  et  $\Psi_i^*(L) = [\Psi_i(L) - \Psi_i(1)] / (1 - L)$  sont tels que les racines de  $\det[\Lambda^*(L)]$  et de  $\det[\Psi_i^*(L)]$  sont toutes supérieures à l'unité en module. Dès lors, le vecteur des  $k$  variables endogènes  $x_{it}$  peut être exprimé comme la somme de six composantes : (i) une composante déterministe, (ii) un ensemble de facteurs communs observables, (iii)  $m$  tendances stochastiques  $s_t$  communes à l'ensemble des individus du panel, (iv)  $k$  tendances stochastiques  $s_{it}$  spécifiques à l'individu  $i$ , (v) les composantes stationnaires communes  $\Lambda^*(L) \varepsilon_t$  et (vi) les composantes stationnaires spécifiques  $\Lambda^*(L) \varepsilon_t$  :

$$\begin{aligned} x_{i,t} &= \begin{matrix} \delta_i & + & \theta_i & t & + & \lambda_i & d_t & + & \gamma_i & \Lambda(1) & s_t & + & \Psi_i(1) & s_{it} \\ (k,1) & & (k,1) & (1,1) & & (k,s) & (s,1) & & (k,m) & (m,m) & (m,1) & & (k,k) & (k,1) \end{matrix} \\ &+ \begin{matrix} \gamma_i & \Lambda^*(L) & \varepsilon_t & + & \Psi_i^*(L) & v_{it} \\ (k,m) & (m,m) & (m,1) & & (k,k) & (k,1) \end{matrix} \end{aligned} \quad (18)$$

Dans ce contexte, il convient de bien préciser la notion de tendances ‘‘communes’’. Remarquons qu'en l'absence de tendances stochastiques communes aux  $N$  individus, *i.e.*  $\Lambda(1) = 0$ , les  $k$  variables endogènes observées pour l'individu de référence  $i$  sont cointégrées si la matrice propre à cet individu  $\Psi_i(1)$  est non nulle et admet un rang strictement inférieur à  $k$ . Dans ce

<sup>11</sup>A noter que les effets individuels  $\tilde{\delta}_i$  et les trends  $\tilde{\theta}_i t$  appartiennent à l'espace de cointégration s'il existe et doivent donc être distingués des effets individuels et des trends de l'équation (4).

cas, il existe  $\text{rang}[\Psi_i(1)]$  tendances stochastiques “communes” aux  $k$  variables, mais spécifiques à l’individu  $i$  du panel. Dit autrement, parmi les  $k$  tendances stochastiques  $s_{it}$  spécifiques à l’individu  $i$ , il existe  $\text{rang}[\Psi_i(1)]$  tendances indépendantes. Chaque individu peut dans ce cas avoir un nombre de tendances communes différent suivant le rang de  $\Psi_i(1)$ . Parallèlement, on dit qu’il existe  $m$  tendances communes (aux  $N$  individus) qui ne sont autres que les composantes tendanciennes  $\Lambda(1) s_t$  des  $m$  facteurs communs non stationnaires.

Dans une telle spécification, le lien entre deux individus  $i$  et  $j \neq i$  du panel peut provenir de différentes sources. Cette dépendance peut tout d’abord provenir d’un ensemble de facteurs communs observables  $d_t$  et/ou de la composante déterministe du modèle. Bien évidemment, en pratique, cela pose le problème de l’identification *ex-ante* de ces facteurs communs. Dans leur application, Dees, di Mauro, Pesaran et Smith (2005) considèrent ainsi un ensemble de variables macroéconomiques comme le revenu et l’inflation des pays n’appartenant pas à l’ensemble géographique figurant dans le panel, le prix du pétrole, *etc.* La dépendance inter-individuelle peut par ailleurs provenir d’un ensemble de tendances stochastiques communes à l’ensemble des  $N$  individus du panel. Plus précisément, le lien entre les variables endogènes  $x_{it}$  et  $x_{jt}$  peut provenir d’un ensemble de  $n \leq m$  tendances stochastiques communes<sup>12</sup> parmi les  $m$  composantes non stationnaires  $\Lambda(1) s_t$  des facteurs communs inobservables, où  $n = \text{rang}[\Lambda(1)]$ . Par exemple, ces  $n$  tendances stochastiques communes peuvent s’interpréter comme des tendances stochastiques internationales ou régionales indépendantes dans le cas d’un panel macroéconomique. Enfin, la dépendance entre individus peut être spécifiée tout simplement par l’intermédiaire de la composante stationnaire des facteurs communs, *i.e.*  $\Lambda^*(L) \varepsilon_t$ . Chaque vecteur de variables  $x_{it}$  dépendant d’un ensemble de  $m$  composantes stationnaires communes, on retrouve une dépendance entre les individus du panel.

## 2.4 Notions de Convergence

Comme nous l’avons déjà mentionné dans le cas de la littérature sur les tests de racine unitaire (Hurlin et Mignon, 2006), l’étude des propriétés asymptotiques des estimateurs des paramètres d’une relation de cointégration (ou d’une régression fallacieuse) en panel pose le problème de la notion de convergence utilisée. Rappelons qu’en économétrie de panel, il existe plusieurs concepts de distribution asymptotique (Phillips et Moon, 1999). On distingue tout d’abord la convergence dite séquentielle, pour laquelle on suppose que les dimensions convergent dans un certain ordre. On raisonne dans un premier temps à  $N$  fixe (ou  $T$ ) et l’on fait tendre

<sup>12</sup> “Communes” à la fois au sens usuel du terme en séries temporelles, *i.e.* communes à l’ensemble des  $m$  variables non stationnaires  $s_t$ , mais aussi “communes” au sens spécifique au panel, *i.e.* communes à l’ensemble des  $N$  individus.

$T$  (ou  $N$ ) vers l'infini, puis l'on fait tendre  $N$  (ou  $T$ ) vers l'infini. C'est souvent l'approche la plus simple pour dériver les lois asymptotiques, mais il convient d'être prudent avec ce type de résultats car ils peuvent parfois être différents lorsque l'on renverse l'ordre de convergence ou lorsque les deux dimensions tendent simultanément vers l'infini. Dans la seconde approche, dite convergence le long d'une diagonale, on fait tendre  $N$  et  $T$  vers l'infini simultanément sous l'hypothèse que  $T = T(N)$  est une fonction monotone de  $N$ . Par exemple, on peut supposer que le ratio  $T/N$  converge vers une constante  $c$  non nulle. Bien évidemment, l'inconvénient de cette approche réside, d'une part, dans le caractère spécifique de la fonction  $T(N)$  qui peut conditionner les résultats et, d'autre part, dans le fait que les dimensions  $(N, T)$  d'un panel peuvent ne pas correspondre, et de loin, à la relation  $T = T(N)$ . La troisième approche, la plus générale mais aussi la plus difficile à utiliser, consiste à supposer que  $N$  et  $T$  tendent vers l'infini sans supposer une quelconque relation entre  $T$  et  $N$ .

La prise en compte de ces différentes notions spécifiques aux panels nous permet à présent de décrire les conséquences de l'absence de cointégration, *i.e.* de la régression fallacieuse.

### 3 Régressions fallacieuses en panel

Une des questions essentielles qui se pose en panel, tout comme en séries temporelles, est celle des conséquences d'une régression impliquant des variables non stationnaires et non cointégrées. En d'autres termes, que devient le phénomène de régression fallacieuse dans un modèle de régression de panel ?

Il convient tout d'abord de faire un bref rappel des conséquences de ce phénomène particulier qu'est la régression fallacieuse dans le cadre standard des séries temporelles (pour plus de détails voir Hamilton, 1994). La régression fallacieuse mise en évidence à partir d'expériences de Monte Carlo par Granger et Newbold (1974), et plus tard expliquée théoriquement par Phillips (1986), apparaît lorsque l'on régresse par les moindres carrés ordinaires (MCO) une variable non stationnaire sur un ensemble de variables non stationnaires et que ces variables sont non cointégrées. La première conséquence est que l'estimateur MCO est alors non convergent (ou fallacieux) au sens où il converge vers une variable aléatoire et non vers une constante (qui devrait logiquement être nulle si les variables sont indépendantes). La seconde conséquence principale est que la distribution de la statistique de Student associée au test de nullité du paramètre du modèle de régression est divergente avec la dimension  $T$ . C'est précisément cette seconde propriété qui implique que l'on ait trop souvent tendance, dans une régression falla-

cieuse, à rejeter à tort la nullité du paramètre liant deux marches aux hasard indépendantes et non cointégrées (Granger et Newbold, 1974).

Or, la régression fallacieuse en panel n'implique pas exactement les mêmes conséquences que celles que l'on connaît en séries temporelles. Afin de bien comprendre ces différences, considérons deux marches au hasard  $y_{it}$  et  $x_{it}$  non cointégrées<sup>13</sup> telles que :

$$\begin{pmatrix} \Delta y_{it} \\ \Delta x_{it} \end{pmatrix} = \begin{pmatrix} \mu_{it} \\ \varepsilon_{it} \end{pmatrix} = w_{it} \quad \forall i = 1, \dots, N, \forall t = 1, \dots, T \quad (19)$$

avec  $y_{i0} = x_{i0} = 0$  et où  $w_{it}$  *i.i.d.*  $(0, \Gamma_0)$ ,  $\Gamma_0$  désignant la matrice de variance covariance des innovations. Pour simplifier, on suppose qu'il n'existe pas de dépendance temporelle entre l'innovation du processus  $x_{it}$  et les valeurs passées de l'innovation du processus  $y_{it}$ .

$$\Gamma_j = E[w_t w'_{t-j}] = \begin{cases} \Gamma_0 & \text{si } j = 0 \\ 0 & \text{sinon} \end{cases} \quad \text{avec} \quad \Gamma_0 = \begin{pmatrix} \gamma_\mu^2 & \gamma_{\mu\varepsilon} \\ \gamma_{\mu\varepsilon} & \gamma_\varepsilon^2 \end{pmatrix} \quad (20)$$

Le paramètre de covariance  $\gamma_{\mu\varepsilon}$  traduit la dépendance instantannée qui peut exister, pour un individu donné, entre les innovations des marches au hasard  $x_{it}$  et  $y_{it}$ . Si  $\gamma_{\mu\varepsilon} = 0$ , les innovations des deux marches au hasard sont indépendantes (comme dans le cas des expériences de Granger et Newbold, 1974). On suppose en outre que les processus d'innovation  $w_{it}$  sont indépendants dans la dimension inter-individuelle.

Admettons qu'un économètre souhaite à présent régresser la variable  $y_{it}$  sur la variable  $x_{it}$  dans un modèle de panel à effets individuels fixes :

$$y_{it} = \alpha_i + \beta x_{it} + e_{it} \quad (21)$$

L'estimateur MCO du paramètre  $\beta$  (estimateur *within* ou LSDV) est alors défini par la quantité :

$$\hat{\beta} = \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y}_i)(x_{it} - \bar{x}_i)}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2} \quad (22)$$

Etudions tout d'abord la convergence de  $\hat{\beta}$ . Kao (1999) et Phillips et Moon (1999) montrent que dans le cas d'une convergence séquentielle<sup>14</sup>, l'estimateur  $\hat{\beta}$  possède les propriétés asymptotiques suivantes :

$$\hat{\beta} \xrightarrow[(T, N \rightarrow \infty)_{seq}]{p} \frac{\gamma_{\mu\varepsilon}}{\gamma_\varepsilon^2} \quad (23)$$

$$\sqrt{N} \left( \hat{\beta} - \frac{\gamma_{\mu\varepsilon}}{\gamma_\varepsilon^2} \right) \xrightarrow[(T, N \rightarrow \infty)_{seq}]{l} N \left( 0, \sigma_{\hat{\beta}}^2 \right) \quad (24)$$

<sup>13</sup>On exclut ici à la fois la cointégration intra et inter-individuelles.

<sup>14</sup>Phillips et Moon (1999) démontrent que ces résultats demeurent valides dans le cas d'une convergence simultanée.



où  $\sigma_{\hat{\beta}}^2 = 2\gamma_{\mu}^2/5\gamma_{\varepsilon}^2$  et où  $(T, N \rightarrow \infty)_{seq}$  désigne un mode de convergence séquentiel où  $T$ , puis  $N$  convergent successivement vers l'infini. On observe que ces deux résultats sont très différents de ceux obtenus en séries temporelles dans le cas d'une régression fallacieuse et appellent ainsi plusieurs commentaires.

La première différence de taille avec le cas des séries temporelles réside dans le fait que l'estimateur  $\hat{\beta}$  en panel converge vers sa vraie valeur  $\gamma_{\mu\varepsilon}/\gamma_{\varepsilon}^2$ . Dans le cas où l'on régresse en panel deux marches au hasard totalement indépendantes ( $\gamma_{\mu\varepsilon} = 0$ ), l'estimateur *within* du paramètre liant  $y_{it}$  et  $x_{it}$  converge ainsi en probabilité vers 0 :

$$\hat{\beta} \xrightarrow[(T, N \rightarrow \infty)_{seq}]{p} 0 \quad (25)$$

De plus, dans ce cas,  $\hat{\beta}$  est  $O_p(N^{-\frac{1}{2}})$ . Ainsi, tout le paradoxe réside dans le fait que, dans le cas d'une régression fallacieuse en panel, on retrouve un résultat de convergence similaire à celui que l'on obtiendrait en séries temporelles en régressant deux séries stationnaires : si les deux séries sont indépendantes, l'estimateur converge en probabilité vers 0. Rappelons que dans le cas d'une régression fallacieuse en séries temporelles (équation 21) pour un individu  $i$  donné, lorsque  $\gamma_{\mu\varepsilon} = 0$ , l'estimateur MCO noté  $\hat{\beta}_i$  converge vers une variable aléatoire telle que :

$$\hat{\beta}_i \xrightarrow[T \rightarrow \infty]{l} \left( \frac{\gamma_{\mu}}{\gamma_{\varepsilon}} \right) \left( \frac{\int_0^1 W_1(r) W_2(r) dr - \int_0^1 W_1(r) dr \int_0^1 W_2(r) dr}{\int_0^1 [W_1(r)]^2 dr - \left[ \int_0^1 W_1(r) dr \right]^2} \right) \quad (26)$$

où  $W_1(r)$  et  $W_2(r)$  désignent deux mouvements Browniens indépendants. Dans ce cas, l'estimateur des MCO est non convergent (ou fallacieux) parce que le bruit correspondant au résidu de la régression est aussi fort que le signal issu de l'explicative  $x_{it}$ , puisque les deux variables sont  $I(1)$ . A l'inverse, le fait de considérer un panel d'individus permet, dans le cas d'une régression fallacieuse, d'identifier un coefficient de régression de long terme<sup>15</sup> (*long-run average regression coefficient*, Phillips et Moon, 1999). Ce résultat *a priori* surprenant provient tout simplement du fait que dans un modèle de panel, le bruit associé aux résidus  $e_{it}$  est atténué par rapport au signal contenu dans  $x_{it}$  puisque ces résidus sont considérés pour un grand nombre d'individus supposés indépendants. Le fait de considérer  $N$  séries de résidus indépendants permet donc en quelque sorte d'atténuer le bruit relativement au signal contenu dans les variables explicatives : ceci illustre une nouvelle fois l'importance de l'hypothèse d'indépendance inter-individuelle dans

<sup>15</sup>La notion de long terme provient du fait que Phillips et Moon (1999), tout comme Kao (1999), autorisent des dépendances temporelles entre  $\mu_{it}$  et  $\varepsilon_{it}$ , *i.e.*  $\Gamma_j \neq 0$  pour  $j \neq 0$ . Dans ce cas, les paramètres  $\gamma_{\varepsilon}^2$  et  $\gamma_{\mu\varepsilon}$  apparaissant dans la distribution asymptotique de  $\hat{\beta}$  correspondent aux paramètres de la matrice de long terme  $\Omega = \sum_{j=-\infty}^{\infty} \Gamma_j$ .

un panel<sup>16</sup>.

La deuxième remarque porte sur le fait que l'on retrouve à nouveau (comme dans la littérature sur les tests de racine unitaire) une distribution asymptotique normale en lieu et place des distributions asymptotiques non standard mises en évidence dans le cas des séries temporelles (Phillips, 1986). Ce résultat s'explique par le fait que l'estimateur *within* peut s'écrire comme le rapport de deux moyennes. Ces moyennes sont elles-mêmes des moyennes de  $N$  éléments indépendants qui convergent, lorsque  $T$  tend vers l'infini, vers des variables aléatoires admettant des moments d'ordre deux finis et identiques. Dès lors, l'utilisation d'un théorème central limite permet de démontrer que cette statistique converge vers une loi normale lorsque  $N$  tend à son tour vers l'infini. Considérons par exemple le cas d'une régression fallacieuse en panel sans constante ( $\alpha_i = 0$ ) de deux marches aléatoires totalement indépendantes ( $\gamma_{\mu\varepsilon} = 0$ ). Lorsque  $T$  tend vers l'infini pour une dimension  $N$  fixée, on montre que :

$$\widehat{\beta} = \frac{\sum_{i=1}^N \sum_{t=1}^T y_{it} x_{it}}{\sum_{i=1}^N \sum_{t=1}^T x_{it}^2} \xrightarrow{T \rightarrow \infty} \begin{pmatrix} \gamma_{\mu} \\ \gamma_{\varepsilon} \end{pmatrix} \left( \frac{\frac{1}{N} \sum_{i=1}^N \int_0^1 W_{i,1}(r) W_{i,2}(r) dr}{\frac{1}{N} \sum_{i=1}^N \int_0^1 [W_{i,1}(r)]^2 dr} \right) \quad (27)$$

Ce qui peut se réécrire sous la forme :

$$\sqrt{N} \widehat{\beta} = \xrightarrow{T \rightarrow \infty} \begin{pmatrix} \gamma_{\mu} \\ \gamma_{\varepsilon} \end{pmatrix} \left[ \sqrt{N} \frac{1}{N} \sum_{i=1}^N \int_0^1 W_{i,1}(r) W_{i,2}(r) dr \right] \left[ \frac{1}{N} \sum_{i=1}^N \int_0^1 [W_{i,1}(r)]^2 dr \right]^{-1} \quad (28)$$

où  $W_{i,1}(r)$  et  $W_{i,2}(r)$  désignent des mouvements Browniens standards indépendants,  $\forall i = 1, \dots, N$ . Par application de la loi forte des grands nombres, on admet que :

$$\frac{1}{N} \sum_{i=1}^N \int_0^1 [W_{i,1}(r)]^2 dr \xrightarrow[N \rightarrow \infty]{p} \frac{1}{2} \quad (29)$$

Dès lors, sachant que les éléments  $\int_0^1 W_{i,1}(r) W_{i,2}(r) dr$  sont indépendants pour  $i = 1, \dots, N$ , que  $E \left[ \int_0^1 W_{i,1}(r) W_{i,2}(r) dr \right] = 0$  et  $Var \left[ \int_0^1 W_{i,1}(r) W_{i,2}(r) dr \right] = 1/5$ , l'application du théorème central limite de Lindberg-Levy permet de montrer que le numérateur de  $\sqrt{N} \widehat{\beta}$  converge vers une loi normale lorsque  $N$  tend vers l'infini.

$$N^{-\frac{1}{2}} \sum_{i=1}^N \int_0^1 W_{i,1}(r) W_{i,2}(r) dr \xrightarrow[N \rightarrow \infty]{l} N(0, 1/5) \quad (30)$$

Ainsi, lorsque  $T$  puis  $N$  tendent vers l'infini, l'estimateur  $\widehat{\beta}$  converge vers une loi normale à la vitesse  $\sqrt{N}$ , alors que dans le cas des séries temporelles, l'estimateur MCO convergeait vers une loi non standard.

$$\sqrt{N} \widehat{\beta} \xrightarrow[(T, N \rightarrow \infty)_{seq}]{l} N \left( 0, \frac{2 \gamma_{\mu}^2}{5 \gamma_{\varepsilon}^2} \right) \quad (31)$$

<sup>16</sup>A notre connaissance, les propriétés asymptotiques de l'estimateur *within* dans le cas d'une régression fallacieuse avec dépendances inter-individuelles n'ont pas encore été étudiées.

On retrouve alors un cas particulier du résultat général de Kao (1999) ou Phillips et Moon (1999).

Pour autant, s'il existe de grandes différences quant aux propriétés asymptotiques de l'estimateur  $\widehat{\beta}$  entre la régression fallacieuse en panel et sur séries temporelles, le comportement asymptotique des principales statistiques de test usuelles est en revanche relativement similaire. Kao (1999) retrouve le résultat selon lequel la statistique de Student associée à la nullité du paramètre de la régression fallacieuse diverge avec la dimension temporelle :

$$T^{-\frac{1}{2}}t_{\beta} - T^{-\frac{1}{2}} \frac{\gamma_{\mu\varepsilon}}{s \gamma_{\varepsilon}^2} \sqrt{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2} \xrightarrow{(T, N \rightarrow \infty)_{seq}} N \left( 0, \frac{2}{5} \right) \quad (32)$$

$$s^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \left[ (y_{it} - \bar{y}_i) - \widehat{\beta} (x_{it} - \bar{x}_i) \right]^2 \quad (33)$$

Si les innovations des deux marches aléatoires sont indépendantes ( $\gamma_{\mu\varepsilon} = 0$ ), ce résultat se résume à :

$$T^{-\frac{1}{2}}t_{\beta} \xrightarrow{(T, N \rightarrow \infty)_{seq}} N \left( 0, \frac{2}{5} \right) \quad \text{ou} \quad t_{\beta} = O_p \left( T^{\frac{1}{2}} \right) \quad (34)$$

Ainsi, le fait de considérer un modèle de régression de panel ne change donc rien aux conséquences de la régression fallacieuse quant à l'inférence. On retrouve exactement le même type de phénomène que celui mis en évidence par Phillips (1986) pour le cas des séries temporelles, *i.e.* une divergence de la distribution de la statistique de Student à la vitesse  $\sqrt{T}$ . La seule différence étant que la distribution de la statistique standardisée  $T^{-\frac{1}{2}}t_{\beta}$  est asymptotiquement normale dans le cas du panel alors qu'elle est non standard dans le cas des séries temporelles. Ce résultat s'explique encore une fois par le fait qu'en panel la statistique de Student peut s'écrire comme une moyenne de  $N$  éléments indépendants qui convergent lorsque  $T$  tend vers l'infini vers des variables aléatoires admettant des moments d'ordre deux finis. L'utilisation d'un théorème central limite permet de démontrer que cette statistique converge vers une loi normale lorsque  $N$  tend à son tour vers l'infini. Mais au delà de ce point particulier, on retrouve la conséquence la plus importante des régressions fallacieuse selon laquelle plus la dimension  $T$  augmente, plus l'on risque de rejeter à tort l'hypothèse nulle d'une nullité du coefficient liant deux variables, indépendantes, non stationnaires et non cointégrées, dès lors que l'on se base à tort sur la statistique de Student usuelle  $t_{\beta}$ .

En conclusion, dans un modèle de régression de panel, si l'on souhaite uniquement estimer les paramètres d'une relation entre des variables non stationnaires, on peut dire finalement que peu importe que ces variables soient cointégrées ou non. Dans les deux cas, *i.e.* cointégration

ou régression fallacieuse, les estimateurs usuels convergent en probabilité vers la vraie valeur du paramètre. C'est là une différence fondamentale avec le cadre des régressions fallacieuses en séries temporelles. En revanche, dès lors que l'on souhaite faire de l'inférence sur l'un ou plusieurs des paramètres du modèle, il convient d'être très prudent quant à l'existence ou non d'une relation de cointégration, puisqu'en présence de régression fallacieuse, les statistiques usuelles de tests ont des distributions divergentes, tout comme c'était le cas en séries temporelles. Ces éléments nous conduisent à présenter les divers tests de cointégration en panel, en commençant par les tests basés sur l'hypothèse nulle d'absence de cointégration.

## 4 Hypothèse nulle d'absence de cointégration

Les tests d'absence de cointégration sur données de panel proposés par Pedroni (1995, 1997, 1999, 2003), Kao (1999) et Bai et Ng (2001) sont des tests résiduels analogues aux tests proposés par Engle et Granger (1987) dans le cadre des séries temporelles. Larsson *et al.* (1998) et Groen et Kleibergen (2003) se sont quant à eux inspirés des travaux de Johansen (1991, 1995) afin de proposer des tests basés sur le rapport de vraisemblance dans un système où *a priori* le nombre de relations de cointégration n'est pas connu. Notons que pour tous ces tests, on suppose la convergence séquentielle :  $T \rightarrow \infty$  suivi de  $N \rightarrow \infty$ .

### 4.1 Les tests de Pedroni

Pedroni (1995, 1997) a proposé divers tests visant à appréhender l'hypothèse nulle d'absence de cointégration intra-individuelle à la fois pour des panels homogènes et hétérogènes. Les valeurs critiques figurant dans ces travaux étant relatives à la présence d'un seul régresseur dans les relations de cointégration, Pedroni (1999, 2003) propose une extension au cas où les relations de cointégration comprennent plus de deux variables. Tout comme les tests de racine unitaire de Im, Pesaran et Shin (2003), les tests de Pedroni prennent en compte l'hétérogénéité par le biais de paramètres qui peuvent différer entre les individus. Une telle hétérogénéité peut se situer à la fois au niveau des relations de cointégration, et au niveau de la dynamique de court terme. Ainsi, sous l'hypothèse alternative, il existe une relation de cointégration pour chaque individu, et cette relation de cointégration n'est pas nécessairement la même pour chacun des individus du panel. La prise en compte d'une telle hétérogénéité constitue un avantage indéniable puisqu'en pratique, il est rare que les vecteurs de cointégration soient identiques d'un individu à l'autre du panel. Dans ces conditions, imposer de manière erronée une homogénéité des vecteurs de cointégration aurait pour conséquence un non rejet de l'hypothèse nulle d'absence de cointégration, alors même que les variables sont cointégrées (voir Pedroni, 1998).

La mise en oeuvre des tests nécessite d'estimer en premier lieu la relation de long terme :

$$y_{it} = \alpha_i + \delta_i t + \beta_{1i} x_{1,it} + \beta_{2i} x_{2,it} + \dots + \beta_{Mi} x_{M,it} + \varepsilon_{it} \quad (35)$$

où  $i = 1, \dots, N$  désigne l'individu,  $t = 1, \dots, T$  et  $m = 1, \dots, M$ .

Sur les sept tests proposés par Pedroni, quatre sont basés sur la dimension *within* (intra) et trois sur la dimension *between* (inter). Les deux catégories de tests reposent sur l'hypothèse nulle d'absence de cointégration :  $\rho_i = 1 \forall i$ ,  $\rho_i$  désignant le terme autorégressif des résidus estimés sous l'hypothèse alternative ( $\hat{\varepsilon}_{it} = \rho_i \hat{\varepsilon}_{it-1} + u_{it}$ ). La distinction entre les deux catégories de tests se situe au niveau de la spécification de l'hypothèse alternative :

- Pour les tests basés sur la dimension intra, l'hypothèse alternative s'écrit :  $\rho_i = \rho < 1 \forall i$ .
- Pour les tests basés sur la dimension inter, l'hypothèse alternative s'écrit :  $\rho_i < 1 \forall i$ .

On constate ainsi que le test basé sur la dimension inter est plus général au sens où il autorise la présence d'hétérogénéité entre les individus sous l'hypothèse alternative.

Les 7 tests développés par Pedroni sont les suivants :

- Tests basés sur la dimension *within* (*panel cointegration statistics*) :
- Test non paramétrique de type rapport de variance (panel  $\nu$ -statistic) :

$$T^2 N^{3/2} Z_{\hat{\nu}N,T} \equiv T^2 N^{3/2} \left( \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{\varepsilon}_{it-1}^2 \right)^{-1} \quad (36)$$

- Test non paramétrique du type de la statistique *rho* de Phillips-Perron (panel  $\rho$ -statistic) :

$$TN^{1/2} Z_{\hat{\rho}N,T^{-1}} \equiv TN^{1/2} \left( \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{\varepsilon}_{it-1}^2 \right)^{-1} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \left( \hat{\varepsilon}_{it-1} \Delta \hat{\varepsilon}_{it} - \hat{\lambda}_i \right) \quad (37)$$

- Test non paramétrique du type de la statistique *t* de Phillips-Perron (panel  $t$ -statistic) :

$$Z_{tN,T} \equiv \left( \tilde{\sigma}_{N,T}^2 \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{\varepsilon}_{it-1}^2 \right)^{-1/2} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \left( \hat{\varepsilon}_{it-1} \Delta \hat{\varepsilon}_{it} - \hat{\lambda}_i \right) \quad (38)$$

- Test paramétrique du type de la statistique *t* de Dickey-Fuller Augmenté (panel  $t$ -statistic)<sup>17</sup> :

$$Z_{tN,T}^* \equiv \left( \tilde{s}_{N,T}^{*2} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{\varepsilon}_{it-1}^{*2} \right)^{-1/2} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{\varepsilon}_{it-1}^* \Delta \hat{\varepsilon}_{it}^* \quad (39)$$

- Tests basés sur la dimension *between* (*group mean panel cointegration statistics*) :
- Test non paramétrique du type de la statistique *rho* de Phillips-Perron (group  $\rho$ -statistic) :

<sup>17</sup>Ce test peut être vu comme un analogue du test de racine unitaire de Levin et Lin (voir Levin et Lin, 1992, 1993, et Levin *et al.*, 2002) sur les résidus estimés de la relation de long terme.

$$TN^{-1/2}\tilde{Z}_{\hat{\rho}_{N,T-1}} \equiv TN^{-1/2} \sum_{i=1}^N \left( \sum_{t=1}^T \hat{\varepsilon}_{it-1}^2 \right)^{-1} \sum_{t=1}^T \left( \hat{\varepsilon}_{it-1} \Delta \hat{\varepsilon}_{it} - \hat{\lambda}_i \right) \quad (40)$$

– Test non paramétrique du type de la statistique  $t$  de Phillips-Perron (group  $t$ -statistic) :

$$N^{-1/2}\tilde{Z}_{tN,T} \equiv N^{-1/2} \sum_{i=1}^N \left( \hat{\sigma}_i^2 \sum_{t=1}^T \hat{\varepsilon}_{it-1}^2 \right)^{-1/2} \sum_{t=1}^T \left( \hat{\varepsilon}_{it-1} \Delta \hat{\varepsilon}_{it} - \hat{\lambda}_i \right) \quad (41)$$

– Test paramétrique du type de la statistique  $t$  de Dickey-Fuller Augmenté (group  $t$ -statistic)<sup>18</sup> :

$$N^{-1/2}\tilde{Z}_{tN,T}^* \equiv N^{-1/2} \sum_{i=1}^N \left( \sum_{t=1}^T \hat{s}_i^{*2} \hat{\varepsilon}_{it-1}^{*2} \right)^{-1/2} \sum_{t=1}^T \hat{\varepsilon}_{it-1}^* \Delta \hat{\varepsilon}_{it}^* \quad (42)$$

avec :

$$\hat{\lambda}_i = T^{-1} \sum_{s=1}^{k_i} \left( 1 - \frac{s}{k_i + 1} \right) \sum_{t=s+1}^T \hat{\mu}_{it} \hat{\mu}_{it-s}, \quad \hat{s}_i^2 \equiv T^{-1} \sum_{t=1}^T \hat{\mu}_{it}^2 \quad (43)$$

$$\hat{\sigma}_i^2 = \hat{s}_i^2 + 2\hat{\lambda}_i, \quad \hat{\sigma}_{N,T}^2 \equiv N^{-1} \sum_{i=1}^N \hat{L}_{11i}^{-2} \hat{\sigma}_i^2, \quad \hat{s}_i^{*2} \equiv T^{-1} \sum_{t=1}^T \hat{\mu}_{it}^{*2} \quad (44)$$

$$\hat{s}_{N,T}^{*2} \equiv N^{-1} \sum_{i=1}^N \hat{s}_i^{*2}, \quad \hat{L}_{11i}^2 = T^{-1} \sum_{t=1}^T \hat{\eta}_{it}^2 + 2T^{-1} \sum_{s=1}^{k_i} \left( 1 - \frac{s}{k_i + 1} \right) \sum_{t=s+1}^T \hat{\eta}_{it} \hat{\eta}_{it-s} \quad (45)$$

et :  $\hat{\mu}_{it} = \hat{\varepsilon}_{it} - \hat{\rho}_i \hat{\varepsilon}_{it-1}$ ,  $\hat{\mu}_{it}^* = \hat{\varepsilon}_{it} - \hat{\rho}_i \hat{\varepsilon}_{it-1} - \sum_{k=1}^{K_i} \hat{\rho}_{ik} \Delta \hat{\varepsilon}_{it-k}$ ,  $\hat{\eta}_{it} = \Delta y_{it} - \sum_{m=1}^M \hat{b}_{mi} \Delta x_{mit}$ .

Dans ces sept tests, les statistiques sont construites sur la base des résidus des relations de cointégration (35) et d'un certain nombre d'estimateurs de paramètres de nuisance. A titre d'exemple, le paramètre de nuisance  $\hat{L}_{11i}^2$  correspond à la variance conditionnelle de long terme individuelle des résidus. Notons enfin que le nombre de retards retenus dans les régressions de type ADF (Dickey-Fuller Augmenté) peut varier entre les individus.

Afin de mettre en oeuvre les différents tests, Pedroni suggère une procédure en 5 étapes<sup>19</sup> :

- Etape 1. On estime la relation de long terme (35) et on récupère les résidus estimés  $\hat{\varepsilon}_{it}$ .
- Etape 2. Pour chaque individu, on différencie la série  $y_{it}$  et on calcule les résidus issus de la régression suivante :  $\Delta y_{it} = b_{1i} \Delta x_{1it} + \dots + b_{Mi} \Delta x_{Mit} + \eta_{it}$ .
- Etape 3. On estime la variance de long terme  $\hat{L}_{11i}^2$  de  $\hat{\eta}_{it}$ .
- Etape 4. En utilisant les résidus estimés  $\hat{\varepsilon}_{it}$ , on choisit la régression appropriée :

<sup>18</sup>Ce test peut être vu comme un analogue du test de racine unitaire  $t$ -bar de Im, Pesaran et Shin (2003) sur les résidus estimés de la relation de long terme.

<sup>19</sup>Notons que pour les tests reposant sur la dimension *between*, les étapes 2 et 3 ne sont pas nécessaires, dans la mesure où ceux-ci ne nécessitent pas le calcul du paramètre de nuisance  $\hat{L}_{11i}^2$ .

- Pour les tests non paramétriques, à l’exception des tests de type  $t$  de Dickey-Fuller Augmenté, on estime la relation  $\hat{\varepsilon}_{it} = \hat{\rho}_i \hat{\varepsilon}_{it-1} + \hat{u}_{it}$ . On calcule la variance de long terme de  $\hat{u}_{it}$ , notée  $\hat{\sigma}_i^2$ . On en déduit alors  $\hat{\lambda}_i = \frac{1}{2} (\hat{\sigma}_i^2 - \hat{s}_i^2)$  où  $\hat{s}_i^2$  désigne la variance de  $\hat{u}_{it}$ .
- Pour les tests paramétriques, on estime la relation  $\hat{\varepsilon}_{it} = \hat{\rho}_i \hat{\varepsilon}_{it-1} + \sum_{k=1}^{K_i} \hat{\rho}_{ik} \Delta \hat{\varepsilon}_{it-k} + \hat{u}_{it}^*$ . On calcule la variance de  $\hat{u}_{it}^*$ , notée  $\hat{s}_i^{*2}$ .
- Etape 5. A l’aide des calculs réalisés dans les étapes précédentes, il est possible de construire l’une des sept statistiques présentées.

Pedroni (1995, 1997) a montré que, sous des normalisations appropriées basées sur des fonctions de mouvements browniens, chacune des 7 statistiques suit une loi normale centrée réduite pour  $T$  et  $N$  suffisamment importants :

$$\frac{\varkappa_{N,T} - \mu\sqrt{N}}{\sqrt{v}} \sim N(0, 1) \quad (46)$$

où  $\varkappa_{N,T}$  désigne une des sept statistiques normalisées. Pedroni (1999) a tabulé les valeurs des moments  $\mu$  et  $v$  nécessaires à la définition de telles normalisations en fonction du nombre de régresseurs et de la présence ou non d’une constante et d’un trend dans les relations de long terme. A partir de ces valeurs, il est alors possible de calculer les valeurs critiques relatives à chacun des sept tests (voir Pedroni, 1999, tableau 2).

Les simulations effectuées par Pedroni (1997) montrent que, pour des valeurs de  $T$  supérieures à 100, les sept statistiques donnent des résultats comparables en termes de puissance. Pour des échantillons de petite taille ( $T < 20$ ), le test le plus puissant est le test basé sur la dimension inter similaire au test ADF (*group  $t$ -statistic*).

## 4.2 Les tests de Kao

Kao (1999) a également proposé des tests de l’hypothèse nulle d’absence de cointégration : test de type Dickey-Fuller et test de type Dickey-Fuller Augmenté. Contrairement aux tests de Pedroni, Kao considère le cas particulier où les vecteurs de cointégration sont supposés homogènes entre les individus. En d’autres termes, ces tests ne permettent pas de tenir compte de l’hétérogénéité sous l’hypothèse alternative et ne sont par ailleurs valables que pour un système bivarié (*i.e.* lorsqu’un seul régresseur est présent dans la relation de cointégration).

Soit  $w_{it} = (\mu_{it}, \nu_{it})'$  un processus bivarié de moyenne nulle et de matrice de variance-covariance de long terme :

$$\Omega = \begin{pmatrix} \sigma_{0\mu}^2 & \sigma_{0\mu\nu} \\ \sigma_{0\mu\nu} & \sigma_{0\nu}^2 \end{pmatrix} \quad (47)$$

et soit  $\Sigma$  la matrice telle que :  $\Sigma = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E(w_{it}w'_{it}) = \begin{pmatrix} \sigma_\mu^2 & \sigma_{\mu\nu} \\ \sigma_{\mu\nu} & \sigma_\nu^2 \end{pmatrix}$ .

Kao (1999) considère le modèle suivant<sup>20</sup> avec effets individuels :

$$y_{it} = \alpha_i + \beta x_{it} + \varepsilon_{it} \quad (48)$$

pour  $i = 1, \dots, N$  et  $t = 1, \dots, T$ , avec  $y_{it} = \sum_{s=1}^t \mu_{is}$  et  $x_{it} = \sum_{s=1}^t \nu_{is}$ .

Le premier test proposé est un test de type Dickey-Fuller appliqué aux résidus estimés  $\hat{\varepsilon}_{it}$  de (48) :

$$\hat{\varepsilon}_{it} = \rho \hat{\varepsilon}_{it-1} + u_{it} \quad (49)$$

et consistant à tester l'hypothèse nulle  $\rho = 1$ . L'estimateur MCO  $\hat{\rho}$  de  $\rho$  est donné par :

$$\hat{\rho} = \frac{\sum_{i=1}^N \sum_{t=2}^T \hat{\varepsilon}_{it} \hat{\varepsilon}_{it-1}}{\sum_{i=1}^N \sum_{t=2}^T \hat{\varepsilon}_{it-1}^2} \quad (50)$$

et la  $t$ -statistique correspondante s'écrit :

$$t_\rho = \frac{(\hat{\rho} - 1) \sqrt{\sum_{i=1}^N \sum_{t=2}^T \hat{\varepsilon}_{it-1}^{*2}}}{s_e} \quad (51)$$

où  $s_e^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=2}^T (\hat{\varepsilon}_{it}^* - \hat{\rho} \hat{\varepsilon}_{it-1}^*)^2$ ,  $\hat{\varepsilon}_{it}^* = y_{it}^* - \hat{\alpha}_i^* - \hat{\beta}^* x_{it}^*$ ,  $y_{it}^* = y_{it} - \sigma_{0\mu\nu} \sigma_{0\nu}^{-2} x_{it}$  et  $x_{it}^* = \sigma_{0\nu}^{-1} x_{it}$ .

En posant  $\sigma_u^2 = \sigma_\mu^2 - \sigma_{\mu\nu} \sigma_\nu^{-2}$  et  $\sigma_{0u}^2 = \sigma_{0\mu}^2 - \sigma_{0\mu\nu} \sigma_{0\nu}^{-2}$ , et en notant  $\hat{\sigma}_u^2$  et  $\hat{\sigma}_{0u}^2$  des estimateurs convergents de  $\sigma_u^2$  et  $\sigma_{0u}^2$ , Kao (1999) propose les quatre statistiques suivantes de type Dickey-Fuller :

$$DF_\rho^* = \frac{\sqrt{NT}(\hat{\rho} - 1) + \frac{3\sqrt{N}\hat{\sigma}_u^2}{\hat{\sigma}_{0u}^2}}{\sqrt{3 + \frac{36\hat{\sigma}_u^4}{5\hat{\sigma}_{0u}^4}}} \quad (52)$$

$$DF_t^* = \frac{t_\rho + \frac{\sqrt{6N}\hat{\sigma}_u}{2\hat{\sigma}_{0u}}}{\sqrt{\frac{\hat{\sigma}_{0u}^2}{2\hat{\sigma}_u^2} + \frac{3\hat{\sigma}_u^2}{10\hat{\sigma}_{0u}^2}}} \quad (53)$$

$$DF_\rho = \frac{\sqrt{NT}(\hat{\rho} - 1) + 3\sqrt{N}}{\sqrt{51/5}} \quad (54)$$

$$DF_t = \sqrt{\frac{5t_\rho}{4}} + \sqrt{\frac{15N}{8}} \quad (55)$$

En considérant un mode de convergence séquentiel où  $T \rightarrow \infty$  suivi de  $N \rightarrow \infty$ , Kao montre que ces quatre statistiques suivent une loi normale centrée réduite sous l'hypothèse nulle

<sup>20</sup>Notons que Kao (1999) a également considéré le cas où un trend déterministe est présent.



d'absence de cointégration. Les statistiques  $DF_\rho^*$  et  $DF_t^*$  prennent en compte la possibilité de corrélations entre les innovations des régresseurs  $I(1)$  et les erreurs (endogénéité de second ordre). A l'inverse, les statistiques  $DF_\rho$  et  $DF_t$  reposent sur l'hypothèse d'exogénéité forte des régresseurs et des erreurs puisque, dans ce cas, on a en effet  $\sigma_\mu^2 = \sigma_{0\mu}^2 = \sigma_u^2 = \sigma_{0u}^2$  (voir notamment Kao et Chen, 1995).

Le second test proposé par Kao (1999) est un test de type ADF basé sur la régression suivante :

$$\hat{\varepsilon}_{it} = \rho \hat{\varepsilon}_{it-1} + \sum_{j=1}^p \varphi_j \Delta \hat{\varepsilon}_{it-j} + u_{itp} \quad (56)$$

La statistique de test est donnée par :

$$ADF = \frac{t_{ADF} + \frac{\sqrt{6N}\hat{\sigma}_u}{2\hat{\sigma}_{0\mu}}}{\sqrt{\frac{\hat{\sigma}_{0u}^2}{2\hat{\sigma}_u^2} + \frac{3\hat{\sigma}_u^2}{10\hat{\sigma}_{0u}^2}}} \quad (57)$$

où  $t_{ADF}$  est la  $t$ -statistique de  $\rho$  dans (56). Sous l'hypothèse nulle d'absence de cointégration, cette statistique suit une loi normale centrée réduite.

Afin d'évaluer la performance des cinq tests proposés, Kao (1999) a mené des simulations de Monte Carlo. Celles-ci font globalement ressortir que pour de faibles valeurs de  $T$  ( $T = 10$ ), les écarts à la taille théorique sont importants pour les 5 tests et disparaissent lorsque  $T$  augmente à 25 (quelle que soit la valeur de  $N$ ). De manière générale, les tests  $DF_\rho^*$  et  $DF_t^*$  conduisent à de meilleurs résultats que les 3 autres tests en termes de taille. Concernant la puissance, les tests ne sont pas très performants pour  $T = 10$  et de faibles valeurs de  $N$ , ce qui est conforme à ce que l'on attendait. Pour  $T = 25$ , le test  $DF_\rho^*$  domine le test  $DF_t^*$  ainsi que le test  $ADF$ , quelle que soit la valeur de  $N$ . Les simulations de Kao (1999) mettent donc globalement en avant que les tests  $DF_\rho^*$  et  $DF_t^*$  sont plus performants que les tests  $DF_\rho$ ,  $DF_t$  et  $ADF$  en termes de taille et de puissance.

### 4.3 Le test de Bai et Ng

Bai et Ng (2004) ont proposé un test de racine unitaire<sup>21</sup> qu'ils ont étendu au cas d'un test résiduel afin de tester l'hypothèse nulle d'absence de cointégration. La méthodologie générale sous-jacente au test repose sur la décomposition de chaque série du panel en la somme d'une composante déterministe, d'une composante commune (ensemble de facteurs communs) et d'une composante individuelle (idiosyncratique)<sup>22</sup>. La série ainsi définie est alors non stationnaire si un ou plusieurs facteurs communs est non stationnaire, ou si la composante individuelle est non stationnaire, ou encore si les deux composantes sont non stationnaires. C'est précisément

<sup>21</sup>Pour un exposé détaillé de ce test, voir Hurlin et Mignon (2006).

<sup>22</sup>Cette procédure est dénommée PANIC (*Panel Analysis of Nonstationarity in the Idiosyncratic and Common components*) par les auteurs.

la possibilité qu'un ou plusieurs facteurs communs soit intégré qui permet au test de Bai et Ng de tenir compte de la présence possible de corrélations entre les individus<sup>23</sup>.

Soient  $\tilde{e}_{it}$  les résidus estimés de la relation :

$$y_{it} = c_i + \lambda_i' \hat{F}_t + e_{it} \quad (58)$$

si l'on considère seulement une constante, ou de la relation :

$$y_{it} = c_i + \beta_i t + \lambda_i' \hat{F}_t + e_{it} \quad (59)$$

si l'on rajoute un trend déterministe.  $\hat{F}_t$  désigne l'estimation du vecteur ( $r \times 1$ ) de facteurs communs et l'erreur  $e_{it}$  est la composante individuelle; la composante déterministe étant soit une constante ( $c_i$ ), soit un trend ( $c_i + \beta_i t$ ).

On note  $DF_e^c(i)$  et  $DF_e^r(i)$  les tests de Dickey-Fuller appliqués aux résidus estimés respectivement des modèles (58) et (59). Supposons que  $\bar{r}$  facteurs soient intégrés d'ordre 1, avec  $\bar{r} < r$ ,  $r$  désignant le nombre total de facteurs. Tester l'hypothèse nulle de racine unitaire dans  $e_{it}$  revient alors à tester l'hypothèse nulle d'absence de cointégration entre  $y_{it}$  et les  $\bar{r}$  facteurs. Dans ces conditions, Bai et Ng (2004) montrent que  $DF_e^c(i)$  et  $DF_e^r(i)$  ont les mêmes distributions asymptotiques que les tests de cointégration proposés par Phillips et Ouliaris (1990) avec  $\bar{r}$  régresseurs, auxquels on rajoute une constante et un trend déterministe dans la relation de cointégration, respectivement. Les simulations menées par Bai et Ng (2004) mettent en évidence que la taille du test est proche de la taille théorique de 5%.

#### 4.4 Les tests de Groen et Kleibergen

Groen et Kleibergen (2003) ont suggéré l'utilisation de statistiques du rapport de vraisemblance afin de développer un test du rang de cointégration dans des modèles à correction d'erreur vectoriels individuels, à la fois pour des vecteurs de cointégration homogènes et hétérogènes<sup>24</sup>. Comme nous l'avons évoqué dans la première partie, Groen et Kleibergen autorisent la présence de relations de cointégration intra-individuelles hétérogènes, mais supposent toutefois que le rang de cointégration est le même pour tous les individus du panel. La différence principale avec les tests présentés précédemment étant que ce rang de cointégration homogène est a priori

<sup>23</sup>Ce point est très important dans l'application des tests de racine unitaire et de cointégration sur données de panel (voir Hurlin et Mignon, 2006, pour des explications détaillées). En effet, à l'exception du test de Bai et Ng (2001), les tests de racine unitaire supposent l'absence de corrélation (et donc de relation de cointégration) entre les individus du panel :  $y_{it}$  et  $y_{jt}$  ne sont pas cointégrés pour  $i \neq j$ . Dans un tel contexte, les tests exhibent une taille et une puissance satisfaisantes. Cependant, en présence de cointégration entre les individus, la taille des tests devient très importante, surtout quand  $T$  augmente. En d'autres termes, les tests rejettent trop souvent l'hypothèse nulle d'absence de cointégration (voir les simulations effectuées par Banerjee *et al.*, 2000, 2003).

<sup>24</sup>Il convient de noter que Larsson *et al.* (1998) ont proposé un test du rang de cointégration dans des panels hétérogènes basé sur la moyenne des statistiques de la trace introduites par Johansen (1991, 1995). Ce test souffre néanmoins d'importants problèmes de taille et nécessite une valeur très élevée de  $T$ .

considéré comme inconnu chez Groen et Kleibergen (2003). Notons qu’au contraire, l’approche résiduelle à la Engle et Granger (1987) adoptée dans les tests de Pedroni (1995, 1997), de Kao (1999) ou de Bai et Ng (2004) impose que l’on ne considère qu’une seule relation de cointégration et que l’ensemble des variables intervenant dans cette relation soit a priori connu. Le test de Groen et Kleibergen est basé sur le test de la trace proposé par Johansen (1991, 1995). Soit  $\overline{LR}(r|k)$  la moyenne des  $N$  statistiques de la trace individuelles :

$$\overline{LR}(r|k) = \frac{1}{N} \sum_{i=1}^N LR_i(r|k) \quad (60)$$

où  $LR_i(r|k)$  désigne la  $i^{\text{ème}}$  statistique du rapport de vraisemblance de Johansen,  $r$  étant le rang de cointégration. On rappelle que le test proposé par Johansen a pour objet de tester l’hypothèse nulle de  $r$  relations de cointégration, contre l’hypothèse alternative de  $k$  relations de cointégration, avec  $r = 0, 1, \dots, k - 1$ <sup>25</sup>. On peut alors construire la statistique suivante :

$$\frac{\overline{LR}(r|k) - E[\overline{LR}(r|k)]}{V[\overline{LR}(r|k)]} \quad (61)$$

qui suit une loi normale centrée réduite pour une convergence séquentielle où  $T \rightarrow \infty$  suivi de  $N \rightarrow \infty$ .

## 5 Hypothèse nulle de cointégration : le test de McCoskey et Kao

McCoskey et Kao (1998) ont quant à eux proposé un test de l’hypothèse nulle de cointégration dans des panels hétérogènes. Il s’agit d’un test résiduel du multiplicateur de Lagrange que l’on peut rapprocher du test de Shin (1994) élaboré dans le cadre des séries temporelles. Les auteurs considèrent l’équation de long terme<sup>26</sup> :

$$y_{it} = \alpha_i + \beta_{1i}x_{1,it} + \beta_{2i}x_{2,it} + \dots + \beta_{Mi}x_{M,it} + \varepsilon_{it} \quad (62)$$

avec  $i = 1, \dots, N$ ,  $t = 1, \dots, T$  et  $m = 1, \dots, M$ , et supposent que  $\varepsilon_{it}$  peut être décomposé en deux éléments :

$$\varepsilon_{it} = \theta \sum_{j=1}^t u_{ij} + u_{it} \quad (63)$$

et :

$$x_{it} = x_{it-1} + e_{it} \quad (64)$$

<sup>25</sup>Le test peut alors s’effectuer de manière séquentielle, partant de  $r = 0$  jusqu’à  $r = k - 1$ .

<sup>26</sup>Il est également possible d’introduire un trend spécifique dans la relation de long terme (voir McCoskey et Kao, 1998).

Tester l'hypothèse nulle de cointégration revient à tester  $\theta = 0$ . Sous l'hypothèse nulle, la relation (62) est une relation de cointégration. La matrice de variance-covariance de long terme de  $\omega_{it} = (u_{it}, e'_{it})'$  est donnée par :

$$\Omega = \begin{pmatrix} \varpi_1^2 & \varpi_{12} \\ \varpi_{21} & \Omega_{22} \end{pmatrix} \quad (65)$$

La statistique de test proposée par McCoskey et Kao (1998) s'écrit alors :

$$LM = \frac{\frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} \sum_{t=1}^T S_{it}^2}{\hat{\omega}_{1,2}^2} \quad (66)$$

où  $\hat{\omega}_{1,2}^2$  est un estimateur convergent de  $\varpi_{1,2}^2$ , avec  $\varpi_{1,2}^2 = \varpi_1^2 - \varpi_{12}\Omega_{22}^{-1}\varpi_{21}$  et :

$$S_{it} = \sum_{k=1}^t \left( y_{ik}^+ - \alpha_i - \hat{\beta}_i^{+'} x_{ik} \right) \quad (67)$$

avec  $y_{ik}^+ = y_{ik} - \hat{\omega}_{12}\hat{\Omega}_{22}^{-1}e_{ik}$ .  $\hat{\beta}_i^+$  désigne l'estimateur FM-OLS (*Fully Modified Ordinary Least Squares*)<sup>27</sup> de  $\beta_i$ . Notons qu'un tel estimateur est asymptotiquement sans biais et permet de prendre en compte une possible corrélation sérielle des résidus de la relation (62) ainsi qu'une endogénéité dite de second ordre des régresseurs. Cette endogénéité se traduit par une corrélation des résidus de la relation de cointégration avec les innovations des régresseurs  $I(1)$ . Afin de mettre en oeuvre leur test, McCoskey et Kao (1998) définissent une statistique  $LM$  modifiée donnée par :

$$LM^+ = \frac{\sqrt{N}(LM - \mu_v)}{\sigma_v} \quad (68)$$

où  $\mu_v$  et  $\sigma_v$  sont la moyenne et l'écart type de fonctions de mouvements browniens (voir Harris et Inder, 1994). Sous l'hypothèse nulle de cointégration ( $\theta = 0$ ), la statistique  $LM^+$  suit une loi normale centrée réduite. La distribution asymptotique de  $LM^+$  ne dépend pas de paramètres de nuisance et est robuste à l'hétéroscédasticité.

McCoskey et Kao (1999a, 2001) ont effectué des simulations de Monte Carlo afin de comparer la taille et la puissance de divers tests de cointégration dans le cas de panels hétérogènes. Ils considèrent le test de l'hypothèse nulle de cointégration de McCoskey et Kao (1998) et deux tests de l'hypothèse nulle d'absence de cointégration : un test basé sur une moyenne des statistiques ADF (voir Wu et Yin, 1999) et un test basé sur une moyenne des statistiques de Phillips et Ouliaris (1990). Leurs résultats montrent que le test de McCoskey et Kao (1998) conduit à de meilleurs résultats en termes de taille et de puissance.

<sup>27</sup>Voir *infra*.

## 6 Estimation et inférence : quelques remarques

Pour estimer des systèmes de variables cointégrées, tout comme pour effectuer des tests sur les vecteurs de cointégration, il est nécessaire d'utiliser une méthode d'estimation efficace. Diverses techniques existent, comme par exemple la méthode FM-OLS (*Fully Modified Ordinary Least Squares*) initialement proposée par Phillips et Hansen (1990) ou la méthode des moindres carrés dynamiques (*Dynamic Ordinary Least Squares*, DOLS) de Saikkonen (1991) et Stock et Watson (1993). Dans le cas des données de panel, Kao et Chiang (2000) ont montré que ces deux techniques conduisaient à des estimateurs asymptotiquement distribués selon une loi normale de moyenne nulle. Des résultats similaires sont obtenus par Pedroni (1996) et Phillips et Moon (1999) pour la méthode FM-OLS.

Dans un tel contexte, Pedroni (1996) a proposé une extension des tests de racine unitaire afin de tester des contraintes sur les coefficients estimés de la relation de cointégration, c'est-à-dire sur les vecteurs de cointégration. On sait en effet que, bien que les estimateurs MCO des vecteurs de cointégration sont super-convergens, leur distribution est asymptotiquement biaisée et dépend de paramètres de nuisance associés à la présence de corrélation sérielle dans les données (voir Kao et Chen, 1995 ; Pedroni, 1996 et Kao et Chiang, 2000). De tels problèmes, existant dans le cas traditionnel univarié, se posent également pour les données de panel et ont tendance à être encore plus marqués en présence d'hétérogénéité (voir notamment Kao et Chen, 1995). Les tests proposés par Pedroni (1996) reposent sur la méthode FM-OLS<sup>28</sup> qui présente l'avantage de donner des résultats plus robustes que la méthode usuelle des MCO lorsque les échantillons sont de petite taille ; les simulations de Pedroni (1996) montrant que les tests de racine unitaire basés sur la méthode des MCO sont biaisés en petit échantillon et tendent à rejeter trop fréquemment l'hypothèse nulle (voir également Phillips et Moon, 1999). En outre, les distributions asymptotiques des estimateurs basés sur la méthode FM-OLS sont non biaisés et ne dépendent pas des paramètres de nuisance. Notons que les tests proposés par Pedroni (1996) s'avèrent particulièrement utiles pour tester l'hypothèse de PPA dans sa version absolue puisqu'ils permettent de tester l'égalité à 1 du vecteur de cointégration pour tous les pays<sup>29</sup>.

Pour conclure, mentionnons que Kao et Chiang (2000) se sont intéressés aux propriétés en échantillon fini des estimateurs des MCO, FM-OLS et DOLS. Leur étude met en avant que l'estimateur des MCO souffre d'un important problème de biais et que l'estimateur FM-OLS ne permet pas d'améliorer de façon substantielle l'estimateur des MCO. Ils concluent alors en termes de supériorité de l'estimateur DOLS dans l'estimation des relations de cointégration sur

---

<sup>28</sup>Cette procédure permet de tenir compte des problèmes d'endogénéité des régresseurs et des propriétés d'autocorrélation et d'hétéroscédasticité des résidus.

<sup>29</sup>Pour une application, voir Pedroni (1996) et Basher et Mohsin (2003). Dans les deux études, les auteurs rejettent la version absolue de la PPA.

données de panel.

## 7 Conclusion

Ce papier a proposé une revue de la littérature méthodologique des principaux tests de cointégration sur données de panel. Cette littérature a connu un grand nombre d’extensions depuis les travaux pionniers de Levin et Lin (1992) et fait aujourd’hui l’objet de très nombreuses recherches et de multiples applications au niveau empirique. A la lumière des avancées réalisées en économétrie des séries temporelles non stationnaires, l’économétrie des données de panel non stationnaires doit encore résoudre un certain nombre de problèmes, notamment au niveau de l’estimation et de l’inférence, et doit s’attacher à tenir compte de différents phénomènes — courants dans la dimension temporelle — tels que la présence de mémoire longue, les effets de type ARCH, l’existence de non linéarités, *etc.*

## Références

- [1] Alberola E., Cervero S.G., Lopez H. et Ubide A. (1999), “Global equilibrium exchange rates : Euro, Dollar, “ins”, “outs” and other major currencies in a panel cointegration framework”, *Working Paper of the International Monetary Fund* 99/175.
- [2] Anthony M. et McDonald R. (1998), “On the mean-reverting properties of target zone exchange rates : some evidence for the ERM”, *European Economic Review*, 42, pp. 1493-1523.
- [3] Bai J. et Ng S. (2004), “A PANIC attack on unit roots and cointegration”, *Econometrica*, 72, pp. 1127-1177.
- [4] Baltagi B.H. et Kao C. (2000), “Nonstationary panels, cointegration in panels and dynamic panels : a survey”, in *Advances in Econometrics*, 15, edited by B. Baltagi et C. Kao, pp. 7-51, Elsevier Science.
- [5] Banerjee A. (1999), “Panel data unit roots and cointegration : an overview”, *Oxford Bulletin of Economics and Statistics*, 61, pp. 607-630.
- [6] Banerjee A., Marcellino M. et Osbat C. (2000), “Some cautions on the use of panel methods for integrated series of macro-economic data”, *Working Paper* 170, IGIER.
- [7] Banerjee A., Marcellino M. et Osbat C. (2003), “Testing for PPP : should we use panel methods?”, *mimeo.*
- [8] Banerjee A. et Zanghieri P. (2003), “A new look at the Feldstein-Horioka puzzle using an integrated panel”, *Working Paper* CEPPII, 2003-22.

- [9] Basher S.A. et Mohsin M. (2003), "PPP tests in cointegrated panels : evidence from Asian developing countries", *Working Paper*, Department of Economics, University of Tennessee.
- [10] Bayoumi T. et MacDonald R. (1999), "Deviations of exchange rates from purchasing power parity : a story featuring two monetary unions", *Staff Papers* 46, 1, IMF.
- [11] Bernard A. et Jones C. (1996), "Productivity across industries and countries : time series theory and evidence", *Review of Economics and Statistics*, 78, pp. 135-146.
- [12] Breitung J. et Mayer W. (1994), "Testing for unit roots in panel data : are wages on different bargaining levels cointegrated?", *Applied Economics*, 26, pp. 353-361.
- [13] Breitung, J. et Pesaran, M.H. (2005), "Unit roots and cointegration in panels", *Deutsche Bank Discussion Paper* 42.
- [14] Camarero M., Ordóñez J. et Tamarit C. (2002), "The Euro-Dollar exchange rate : is it fundamental?", *Working Paper* 16/2002, European Economy Group.
- [15] Camarero M. et Tamarit C. (2002), "A panel cointegration approach to the estimation of the peseta real exchange rate", *Journal of Macroeconomics*, 24, pp. 371-393.
- [16] Canning, D. et Pedroni, P (1999), "Infrastructure and long-run economic growth", *CAER II Discussion Paper* 57, Harvard Institute for International Development.
- [17] Canzoneri M.B., Cumby R.E. et Diba B.T. (1996), "Relative labor productivity and the real exchange rate in the long run : evidence for a panel of OECD countries", *NBER Working Paper* 5676.
- [18] Cecchetti S.G., Mark N.C. et Sonora R. (1999), "Price level convergence among United States cities : lessons from the European Central Bank", *Mimeo*.
- [19] Chang Y. (2002), "Nonlinear IV unit root tests in panels with cross-sectional dependency", *Journal of Econometrics*, 110, pp. 261-292.
- [20] Chinn M.D. et Johnston L. (1996), "Real exchange rate levels, productivity and demand shocks : evidence from a panel of 14 countries", *Working Paper*, Department of Economics, University of California, Santa Cruz.
- [21] Chortareas G.E. et Driver R.L. (2000), "PPP and the real exchange rate-real interest rate differential puzzle revisited : evidence from nonstationary panel data", *Working Paper*, Bank of England.
- [22] Coakley J., Kulasi F. et Smith R. (1996), "Current account solvency and the Feldstein-Horioka puzzle", *Economic Journal*, 106, pp. 620-627.
- [23] Dees, S., di Mauro, F, Pesaran, M.H. et Smith, L.V. (2005), "Exploring the international linkages of the euro area : A Global VAR analysis", *CESifo Working Paper* 1425.

- [24] Drine I. et Rault C. (2003), “A re-examination of the purchasing power parity using non-stationary dynamic panel methods : a comparative approach for developing and developed countries”, *Working Paper* 570, William Davidson Institute, University of Michigan Business School.
- [25] Engle R.F. et Granger C.W.J. (1987), “Cointegration and error-correction : representation, estimation and testing”, *Econometrica*, 64, pp. 813-836.
- [26] Evans P. et Karras G. (1996), “Convergence revisited”, *Journal of Monetary Economics*, 37, pp. 249-265.
- [27] Fisher R.A. (1932), *Statistical methods for research workers*, Oliver&Boyd, Edinburgh.
- [28] Frankel J.A. et Rose A.K. (1995), “Empirical research on nominal exchange rates”, in *Handbook of International Economics*, Vol. 3, edited by G.M. Grossman et K. Rogoff, pp. 1689-1729, Elsevier Science.
- [29] Funk M. (1998), “Trade and international R&D spillovers among OECD countries”, *Working Paper*, Department of Economics, St. Louis University.
- [30] Gerdtham U.G. et Löthgren M. (1998), “On stationarity and cointegration of international health expenditure and GDP”, *Working Paper in Economics and Finance* 232, Stockholm School of Economics.
- [31] Granger, C.W.J., et Newbold P. (1974), “Spurious regressions in econometrics”, *Journal of Econometrics*, 2, pp. 111-120.
- [32] Groen, J.J.J. et Kleibergen, F. (2003), “Likelihood-based cointegration analysis in panels of vector error correction models”, *Journal of Business and Economic Statistics*, 21(2), pp. 295-318.
- [33] Hamilton, J.D. (1994), *Time Series Analysis*, Princeton.
- [34] Harris D. et Inder B. (1994), “A test of the null hypothesis of cointegration”, in *Nonstationary time series analysis and cointegration*, Hargreaves C. (ed), Oxford University Press.
- [35] Holtz-Eakin D., Newey W. and Rosen H.S. (1988), “Estimating vector autoregressions with panel data”, *Econometrica*, 56, 1371-1396.
- [36] Hurlin C. et Mignon V. (2006), “Une synthèse des tests de racine unitaire sur données de panel”, à paraître dans *Economie et Prévision*.
- [37] Im K.S., Pesaran M.H. et Shin Y. (2003), “Testing for unit roots in heterogeneous panels”, *Journal of Econometrics*, 115, pp. 53-74.



- [38] Islam N. (1995), "Growth empirics : a panel data approach", *Quarterly Journal of Economics*, 110, pp. 1127-1170.
- [39] Johansen S. (1991), "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models", *Econometrica*, 59, pp. 1551-1580.
- [40] Johansen S. (1995), *Likelihood-based inference in cointegrated vector autoregressive models*, Oxford University Press.
- [41] Jorion P. et Sweeney R. (1996), "Mean reversion in real exchange rates : evidence and implications for forecasting", *Journal of International Money and Finance*, pp. 535-550.
- [42] Kao C. (1999), "Spurious regression and residual-based tests for cointegration in panel data", *Journal of Econometrics*, 90, pp. 1-44.
- [43] Kao C. et Chen B. (1995), "On the estimation and inference for cointegration in panel data when the cross section and time series dimensions are comparable", *Manuscript*, Center for Policy Research, Syracuse University, New York.
- [44] Kao C. et Chiang M.H. (2000), "On the estimation and inference of a cointegrated regression in panel data", *Advances in Econometrics*, 15, edited by B. Baltagi et C. Kao, pp. 179-222, Elsevier Science.
- [45] Kao C., Chiang M.H. et Chen B. (1999), "International R&D spillovers : an application of estimation and inference in panel cointegration", *Oxford Bulletin of Economics and Statistics*, 61, pp. 691-709.
- [46] Larsson R., Lyhagen J. et Löthgren M. (1998), "Likelihood-based cointegration tests in heterogenous panels", *Working Paper in Economics and Finance* 250, Stockholm School of Economics.
- [47] Lee K.E., Pesaran M.H. et Smith R. (1997), "Growth and convergence in a multi-country empirical stochastic Solow model", *Journal of Applied Econometrics*, 12, pp. 357-392.
- [48] Levin A. et Lin C.F. (1992), "Unit root tests in panel data : Asymptotic and finite sample properties", *Discussion Paper* 56, Department of Economics, University of California at San Diego.
- [49] Levin A. et Lin C.F. (1993), "Unit root tests in panel data : new results", *Discussion Paper* 92-93, Department of Economics, University of California at San Diego.
- [50] Levin A., Lin C.F. et Chu J. (2002), "Unit root tests in panel data : Asymptotic and finite sample properties", *Journal of Econometrics*, 108, pp. 1-24.
- [51] Maddala G. et Wu S. (1999), "A comparative study of unit root tests and a new simple test", *Oxford Bulletin of Economics and Statistics*, 61, pp. 631-652.

- [52] McCoskey S. et Kao C. (1998), “A residual-based test of the null of cointegration in panel data”, *Econometric Reviews*, 17, pp. 57-84.
- [53] McCoskey S. et Kao C. (1999a), “Comparing panel data cointegration tests with an application to the “twin deficits” problem”, *Working Paper*, Center for Policy Research, Syracuse University, New York.
- [54] McCoskey S. et Kao C. (1999b), “Testing the stability of a production function with urbanization as a shift factor : an application of non-stationary panel data techniques”, *Oxford Bulletin of Economics and Statistics*, 61, pp. 671-690.
- [55] McCoskey S. et Kao C. (2001), “A Monte Carlo comparison of tests for cointegration in panel data”, *Journal of Propagations in Probability and Statistics*, 1, pp. 165-198.
- [56] McCoskey S. et Selden T. (1998), “Health care expenditures and GDP : panel data unit root test results”, *Journal of Health Economics*, 17, pp. 369-376.
- [57] McDonald R. (1996), “Panel unit root tests and real exchange rates”, *Economic Letters*, 50, pp. 7-11.
- [58] McDonald R. et Nagayasu J. (2000), “The long-run relationship between real exchange rate and real interest differentials : a panel study”, *IMF Staff Papers* 47, pp. 116-128.
- [59] Moon H.R. et Phillips P.C.B. (1998), “A reinterpretation of the Feldstein-Horioka regressions from a nonstationary panel viewpoint”, *Working Paper*, Yale University.
- [60] Obstfeld M. et Taylor (1996), “International capital-market integration over the long run : the great depression as a watershed”, *manuscript*, U.C. Berkeley.
- [61] O’Connell P. (1998), “The overvaluation of purchasing power parity”, *Journal of International Economics*, 44, pp. 1-19.
- [62] Oh K. (1996), “Purchasing power parity and unit root tests using panel data”, *Journal of International Money and Finance*, 15, pp. 405-418.
- [63] Papell D. (1997), “Searching for stationarity : purchasing power parity under the current float”, *Journal of International Economics*, 43, pp. 313-332.
- [64] Papell D. (2002), “The great appreciation, the great depreciation, and the purchasing power parity hypothesis”, *Journal of International Economics*, 57, pp. 51-82.
- [65] Papell D. et Theodoridis H. (2001), “The choice of numeraire currency in panel tests of purchasing power parity”, *Journal of Money, Credit and Banking*, 33, pp. 790-803.
- [66] Pedroni P. (1995), “Panel cointegration, asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis”, *Working Paper in Economics*, 92-013, Indiana University.

- [67] Pedroni P. (1996), “Fully modified OLS for heterogeneous cointegrated panels and the case of purchasing power parity”, *Working Paper in Economics*, Indiana University.
- [68] Pedroni P. (1997), “Panel cointegration, asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis : new results”, *Working Paper in Economics*, Indiana University.
- [69] Pedroni P. (1998), “Testing for convergence to common steady states in nonstationary heterogeneous panels”, *Working Paper in Economics*, Indiana University.
- [70] Pedroni P. (1999), “Critical values for cointegration tests in heterogenous panels with multiple regressors”, *Oxford Bulletin of Economics and Statistics*, S1, 61, pp. 653-670.
- [71] Pedroni P. (2001), “Panel cointegration, asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis”, *Working Paper in Economics*, Indiana University.
- [72] Pedroni P. (2003), “Panel cointegration. Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis”, à paraître dans *Econometric Theory*.
- [73] Perron P. (1989), “Testing for a random walk : a simulation experiment of power when the sampling interval is varied”, in *Advances in Econometrics and Modeling*, edited by B. Jaj, Kluwer Academic Publishers, pp. 47-68.
- [74] Perron P. (1991), “Test consistency with varying sampling frequency”, *Econometric Theory*, 7, pp. 341-368.
- [75] Phillips, P.C.B. (1986), “Understanding spurious regressions in econometrics”, *Journal of Econometrics*, 33, pp. 311-340.
- [76] Phillips P.C.B. et Hansen B.E. (1990), “Statistical inference in instrumental variables regression with I(1) processes”, *Review of Economic Studies*, 57, pp. 99-125.
- [77] Phillips P.C.B. et Moon H.R. (1999), “Linear regression limit theory for nonstationary panel data”, *Econometrica*, 67, pp. 1057-1111.
- [78] Phillips P.C.B. et Moon H.R. (2000), “Nonstationary panel data analysis : an overview of some recent developments”, *Econometric Reviews*, 19, pp. 263-286.
- [79] Phillips P.C.B. et Ouliaris S. (1990), “Asymptotic properties of residual based tests for cointegration”, *Econometrica*, 58, pp. 165-193.
- [80] Pierce R.G. et Snell A.J. (1995), “Temporal aggregation and the power of tests for a unit root”, *Journal of Econometrics*, 65, pp. 333-345.

- [81] Quah D. (1992), “The relative importance of permanent and transitory components : identification and some theoretical bounds”, *Econometrica*, 60, pp. 107-118.
- [82] Quah D. (1994), “Exploiting cross-section variation for unit root inference in dynamic data”, *Economics Letters*, 44, pp. 9-19.
- [83] Rogoff K. (1996), “The purchasing power parity puzzle”, *Journal of Economic Literature*, 34, pp. 647-668.
- [84] Saikkonen P. (1991), “Asymptotically efficient estimation of cointegrating regressions”, *Econometric Theory*, 58, pp. 1-21.
- [85] Sala-i-Martin X. (1996), “The classical approach to convergence analysis”, *Economic Journal*, 106, pp. 1019-1036.
- [86] Shiller R. et Perron P. (1985), “Testing the random walk hypothesis : power versus frequency of observations”, *Economics Letters*, 18, pp. 381-386.
- [87] Shin Y. (1994), “A residual-based test of the null of cointegration against the alternative of no cointegration”, *Econometric Theory*, 19, pp. 91-115.
- [88] Stock J. et Watson M. (1993), “A simple estimator of cointegrating vectors in higher order integrated systems”, *Econometrica*, 61, pp. 783-820.
- [89] Taylor A. (1996), “International capital mobility in history : purchasing power parity in the long run”, *manuscript*, Northwestern University.
- [90] Wei S.J. et Parsley D.C. (1995), “Purchasing power parity during the floating rate period : exchange rate volatility, trade barriers, and other culprits”, *Working Paper*, Kennedy School of Government.
- [91] Wu S. et Yin Y. (1999), “Tests for cointegration in heterogeneous panel : A Monte Carlo study”, *Working Paper*, Department of Economics, State University of New York.
- [92] Wu Y. (1996), “Are real exchange rates nonstationary? Evidence from a panel-data test”, *Journal of Money, Credit and Banking*, 28, pp. 54-63.