

Une Synthèse des Tests de Cointégration sur Données de Panel

Christophe Hurlin* et Valérie Mignon†

30 novembre 2006

Résumé

L'objet de ce papier est de proposer une revue de la littérature méthodologique 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, Kao, Bai et Ng, Groen et Kleibergen) ainsi que le test de McCoskey et Kao reposant sur l'hypothèse nulle de cointégration. Quelques éléments relatifs à la comparaison des tests par le biais de résultats de simulations de Monte Carlo, 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 an overview of panel data cointegration tests. We present the main tests based on the null hypothesis of no cointegration (tests of Pedroni, Kao, Bai and Ng, Groen and Kleibergen) and the McCoskey and Kao test based on the null hypothesis of cointegration. Some developments relating to the comparison of the size and power of tests, inference and estimation of cointegrated systems are also provided.

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

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

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 cointégration¹ — connaît un développement considérable². 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 exemple flagrant est sans conteste celui des travaux empiriques portant sur la

*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.

†EconomiX, 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

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

²Pour une revue de la littérature, voir notamment Banerjee (1999), Phillips et Moon (2000), Baltagi et Kao (2000) ou encore Hurlin et Mignon (2005).

vérification de l'hypothèse de parité des pouvoirs d'achat³ (PPA) dans le domaine des changes (voir notamment McDonald, 1996 ; Oh, 1996 ; ou encore Rogoff, 1996)⁴.

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 ? Outre l'argument classique relatif à la faible puissance des tests de cointégration sur séries temporelles en petit échantillon, l'ajout de la dimension individuelle permet d'accroître le nombre de données en incluant l'information relative à des pays différents et mener ainsi une analyse multi-pays⁵. 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. 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 considérée. Enfin, les nouvelles générations de tests de racine unitaire et de cointégration en panel permettent de prendre en compte les éventuelles interdépendances entre les individus du panel sous différentes formes. Le panel n'est plus alors une simple collection d'individus indépendants, mais une structure qui peut être soumise par exemple à l'influence de facteurs communs observables ou non observables. Dès lors, dans ce contexte, le fait d'utiliser un modèle de panel en lieu et place de séries temporelles ne relève plus d'un choix, mais relève d'avantage d'une nécessité liée à la prise en compte et à la modélisation de ces facteurs communs.

Dans cet article, notre objectif est de rendre compte des développements relatifs aux tests de cointégration sur données de panel. Comme dans le cas des séries temporelles, les tests de cointégration en panel peuvent tout d'abord être classés en deux catégories : d'une part, les tests de cointégration *à la* Engle et Granger (1987) dans lesquels le rang de l'espace de cointégration et la forme de la relation de cointégration sont connus (Pedroni, 1995, 1997 ; Kao, 1999 ; Bai et Ng, 2004). Ces tests reviennent alors à tester l'hypothèse de racine unitaire des résidus de cointégration et sont donc par conséquent extrêmement proches dans leur structure des tests de racine unitaire en panel de première et de deuxième générations. D'autre part, sont aujourd'hui développés des

³Rappelons que tester la forme relative de la PPA revient à effectuer un test de cointégration entre le change nominal et les indices de prix.

⁴Notons 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 Gerdtham 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 (1996), Sala-i-Martin (1996), Lee, Pesaran et Smith (1997), 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.

⁵Augmenter le nombre de données en ajoutant la dimension individuelle et non en augmentant la dimension temporelle réduit en outre le risque de faire face à des ruptures structurelles.

tests de cointégration multiples à la Johansen (1991) dans lesquels le rang de cointégration est *a priori* inconnu (Larsson *et al.* 2001 ; Groen et Kleibergen, 2003 ; Breitung, 2005). Il existe en outre dans cette littérature un deuxième clivage similaire à celui que l'on avait pu observer en séries temporelles entre les tests admettant pour hypothèse nulle l'absence de relation de cointégration et les tests à la Shin (1994) reposant au contraire sur l'hypothèse nulle de cointégration (McCoskey et Kao, 1998).

Mais, à ces deux oppositions méthodologiques "traditionnelles", s'ajoutent à l'analyse de la cointégration de nouvelles dimensions propres au panel. Premièrement, émerge ainsi une opposition entre les notions de cointégration inter-individuelle et de cointégration intra-individuelle n'impliquant que des variables propres à un individu donné. Deuxièmement, se pose en panel la question de l'homogénéité / hétérogénéité de la relation de cointégration, ce qui revient à poser la question de l'homogénéité de la situation de long terme des pays ou des individus du panel. Plus généralement, cette problématique s'inscrit dans le cadre de la spécification homogène / hétérogène de la représentation à correction d'erreur (dynamique de court terme, vitesse d'ajustement à la relation de cointégration, *etc.*). On distinguera ainsi les tests en panel suivant la forme hétérogène ou homogène retenue de la représentation à correction d'erreur (VECM). Troisièmement, l'analyse de la cointégration en panel soulève le problème de la prise en compte et de la spécification des dépendances inter-individuelles que nous avons déjà évoqué dans le cadre des tests de racine unitaire de seconde génération (Hurlin et Mignon, 2005).

Enfin, la dimension de panel apporte non seulement un enrichissement à l'analyse de la cointégration, mais aussi une modification des conséquences statistiques de la situation de non cointégration entre variables non stationnaires, c'est-à-dire la situation de régression fallacieuse (Phillips et Moon, 1999).

Ainsi, cet enrichissement de l'analyse lié au passage en panel nécessite que soit posé un certain nombre de définitions, de notions ou de concepts propres à la cointégration et la régression fallacieuse en panel. Ce sera précisément l'objet des deux premières sections de cet article. Les deux sections suivantes visent à 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 la comparaison des tests, 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.

Pour cela, considérons une **représentation VECM dans un système complet**. Une telle représentation autorise l'existence de dynamiques inter-individuelles : le taux de croissance d'une variable endogène observée pour un individu donné peut 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 $x_{it} = (x_{it}^{(1)}, \dots, x_{it}^{(k)})'$ un vecteur de k variables⁶ $I(1)$ 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) :

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

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 θ regroupe les coefficients des tendances déterministes que l'on suppose *a priori* différents pour chaque variable considérée et pour chaque individu. Le troisième terme du modèle (1) 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. On suppose que la matrice Π peut être partitionnée de la façon suivante :

$$\underset{(Nk,Nk)}{\Pi} = \begin{pmatrix} \Pi_{11} & \dots & \Pi_{1N} \\ \dots & \Pi_{ij} & \dots \\ \Pi_{N1} & \dots & \Pi_{NN} \end{pmatrix} \quad (2)$$

où les matrices Π_{ij} sont de dimension (k, k) . Le quatrième terme ΓW_t de la relation (1) regroupe l'ensemble des éléments liés à la dynamique d'ajustement de court terme. Pour simplifier, on suppose que pour purger toute l'autocorrélation des résidus du modèle (1), il est nécessaire d'introduire le même nombre p d'éléments différenciés retardés pour tous les individus du panel. Ainsi,

⁶On 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.

on pose :

$$W_t = (\Delta x'_{1,t-1}, \dots, \Delta x'_{1,t-p}, \dots, \Delta x'_{N,t-1}, \dots, \Delta x'_{N,t-p}) \quad (3)$$

$$\Gamma_{(Nk, Nkp)} = \begin{pmatrix} \Gamma_{11} & \dots & \Gamma_{1N} \\ \dots & \Gamma_{ij} & \dots \\ \Gamma_{N1} & \dots & \Gamma_{NN} \end{pmatrix} \quad (4)$$

où les matrices Γ_{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 les résidus ε_t sont *i.i.d.* $(0, \Omega)$. L'éventuelle dépendance inter-individuelle de ces résidus implique que la matrice Ω peut être non bloc diagonale :

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

où les matrices Ω_{ij} sont de dimension (k, k) . Rappelons que, pour un individu i , si la matrice Ω_{ii} est non diagonale, cela traduit l'existence de corrélations instantanées des chocs de la forme réduite, identiques à celles que l'on considère dans 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 intra-individuelles si et seulement si :

$$\beta'_i x_{it} = \mu_{it} \sim I(0) \quad (6)$$

où β_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 implique que pour un pays donné, il existe une combinaison linéaire du PIB et de la consommation qui soit stationnaire. Comme nous le verrons, 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 :

$$\begin{matrix} \beta' & x_t \\ (r, Nk) & (Nk, 1) \end{matrix} = \begin{matrix} \mu_t \\ (r, 1) \end{matrix} \sim I(0) \quad (7)$$

La matrice Π de la représentation VECM (équation 1) se décompose alors sous la forme usuelle :

$$\begin{matrix} \Pi \\ (Nk, Nk) \end{matrix} = \begin{matrix} \alpha & \beta' \\ (Nk, r) & (r, Nk) \end{matrix} \quad (8)$$

où la matrice α 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 β 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-individuelles⁷.

⁷Une exception notable est le modèle VECM proposé par Groen et Kleibergen (2003). Cette approche semble toutefois difficile à mettre en oeuvre dans la mesure où elle nécessite des panels de taille très importante.

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 nous intéresser à la question de l'homogénéité de la relation de long terme avant, dans un second temps, de nous intéresser à la question de l'homogénéité de la représentation VECM dans son ensemble.

2.2.1 La relation de long terme

Admettons 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 Π_{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 Π de la représentation VECM est alors bloc diagonale et peut s'écrire sous la forme :

$$\Pi_{(Nk, Nk)} = \begin{pmatrix} \Pi_{11} & \dots & 0 \\ (k, k) & & (k, k) \\ \dots & \Pi_{ii} & \dots \\ & (k, k) & \\ 0 & \dots & \Pi_{NN} \\ (k, k) & & (k, k) \end{pmatrix} = \begin{pmatrix} \alpha_1 & \beta'_1 & \dots & 0 \\ (k, r_1) & (r_1, k) & & (k, k) \\ \dots & \alpha_i & \beta'_i & \dots \\ & (k, r_i) & (r_i, k) & \\ 0 & \dots & \alpha_N & \beta'_N \\ (k, k) & & (k, r_N) & (r_N, k) \end{pmatrix} \quad (9)$$

où β_i désigne la matrice des vecteurs de cointégration de l'individu i et α_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'homogénéité de la relation de cointégration est parfois l'expression d'une contrainte théorique, comme par exemple dans le cas 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 α_i , peuvent être différentes d'un pays à l'autre. En revanche, dans d'autres cas, cette hypothèse peut être plus discutable comme par exemple lors de l'estimation d'une fonction de production de long terme sur un panel international⁸ (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 homogène⁹ implique que les rangs de cointégration soient les mêmes pour tous les individus, *i.e.*

⁸Rien 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.

⁹Par 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.

$r_i = r, \forall i = 1, \dots, N$. En revanche, dans un modèle de cointégration hétérogène, 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 $\bar{\beta} = (1/N) \sum_{i=1}^N \beta_i$.

2.2.2 La représentation VECM

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 1). 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 différente suivant les individus. De la même façon, les paramètres Γ_{ii} associés aux termes de la dynamique de court terme peuvent être hétérogènes¹⁰, *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 α_i et Γ_{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 et Rosen (1988) au modèle VECM où seules les relations de

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

long terme sont homogènes (Groen et Kleibergen, 2003). Enfin, les mêmes choix s’offrent quant à la spécification des composantes déterministes : effets individuels δ_i et coefficients θ associés aux tendances 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 dans une représentation VECM.

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, cette dépendance se traduit par une matrice de variance-covariance Ω dans la représentation VECM (équation 1) 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 :

$$x_{i,t} = \begin{matrix} \tilde{\delta}_i & + & \tilde{\theta}_i & t & + & \lambda_i & d_t & + & \gamma_i & f_t & + & w_{it} \\ (k,1) & & (k,1) & (1,1) & & (k,s) & (s,1) & & (k,m) & (m,1) & & (k,1) \end{matrix} \quad (10)$$

$$\Delta f_t = \begin{matrix} \Lambda(L) & \varepsilon_t & \varepsilon_t \text{ i.i.d. } (0, I_m) \\ (m,m) & (m,1) & \end{matrix} \quad (11)$$

$$\Delta w_{it} = \begin{matrix} \Psi_i(L) & v_{it} \\ (k,1) & (k,k) \end{matrix} \begin{matrix} (k,1) \\ (k,1) \end{matrix} v_{it} \text{ i.i.d. } (0, I_k) \quad (12)$$

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 déterministe¹¹ (effet individuel et tendance déterministe), (ii) d'un ensemble de s facteurs communs observables d_t (variables macroéconomiques, autres composantes déterministes comme des variables indicatrices 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} .

Admettons que les facteurs communs inobservables f_t et les composantes individuelles w_{it} sont intégrés d'ordre un. Si l'on suppose en particulier que $w_{i0} = f_0 = 0$ et que $\varepsilon_0 = v_{i0} = 0$, on obtient alors :

$$f_t = \Lambda(1) \sum_{s=1}^t \varepsilon_s + \Lambda^*(L) \varepsilon_t = \Lambda(1) s_t + \Lambda^*(L) \varepsilon_t \quad (13)$$

$$w_{it} = \Psi_i(1) \sum_{s=1}^t v_{is} + \Psi_i^*(L) v_{it} = \Psi_i(1) s_{it} + \Psi_i^*(L) v_{it} \quad (14)$$

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 $\Lambda^*(L) \varepsilon_t$:

$$\begin{aligned} x_{i,t} &= \begin{matrix} \delta_i & \theta_i & t & \lambda_i & d_t & \gamma_i & \Lambda(1) & s_t & \Psi_i(1) & s_{it} \\ (k,1) & (k,1) & (1,1) & (k,s) & (s,1) & (k,m) & (m,m) & (m,1) & (k,k) & (k,1) \end{matrix} \\ &+ \begin{matrix} \gamma_i & \Lambda^*(L) & \varepsilon_t & \Psi_i^*(L) & v_{it} \\ (k,m) & (m,m) & (m,1) & (k,k) & (k,1) \end{matrix} \end{aligned} \quad (15)$$

Dès lors, dans une telle représentation VECM, il convient de bien préciser la notion de tendances "communes". On doit en effet distinguer, d'une part, les tendances stochastiques communes aux différents individus du panel et, d'autre part, les tendances stochastiques communes observées parmi les variables endogènes du système pour un même individu (qui correspondent aux tendances communes d'une représentation VECM en séries temporelles). En effet, dans la représentation (15), il existe tout d'abord m tendances stochastiques "communes" (aux N individus) qui correspondent aux composantes tendanciennes $\Lambda(1) s_t$ des m facteurs communs non stationnaires.

¹¹A 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 (1).

De plus, 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 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 stochastiques indépendantes. Chaque individu peut dans ce cas avoir un nombre de tendances communes différent suivant le rang de $\Psi_i(1)$.

Dans une telle spécification, le lien entre deux individus i et $j \neq i$ du panel peut provenir de trois sources. Premièrement, cette dépendance peut 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.*

Deuxièmement, la dépendance inter-individuelle peut être spécifiée sous la forme 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 communes¹² 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, troisièmement, la dépendance entre individus peut être spécifiée tout simplement au travers d'une corrélation des composantes stationnaires 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.

Naturellement, c'est le choix du test de cointégration qui conditionne le schéma de dépendance (ou d'indépendance) inter-individuelle retenu même si peu de tests de cointégration, encore aujourd'hui, intègrent les dépendances inter-individuelles.

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, 2005), 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

¹² “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.

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 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 $T = T(N)$ postulée. 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 . Dans les tests que nous présenterons par la suite, la convergence séquentielle considérée est la suivante : $T \rightarrow \infty$ suivi de $N \rightarrow \infty$.

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, initialement mise en avant par Yule (1926) au travers du concept de “corrélations dénuées de sens”, a été illustrée à partir d’expériences de Monte Carlo par Granger et Newbold (1974), et plus tard expliquée théoriquement par Phillips (1986). Elle 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 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 fallacieuse, à rejeter à tort la nullité du paramètre liant deux marches aléatoires 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 aléatoires y_{it} et x_{it} non cointégrées¹³ telles que :

$$\begin{pmatrix} \Delta y_{it} \\ \Delta x_{it} \end{pmatrix} = \begin{pmatrix} \mu_{it} \\ \varepsilon_{it} \end{pmatrix} = w_{it} \quad \forall i = 1, \dots, N, \forall t = 1, \dots, T \quad (16)$$

avec $y_{i0} = x_{i0} = 0$ et où w_{it} *i.i.d.* $(0, \Gamma_0)$, Γ_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} .

$$\Gamma_j = E[w_t w_{t-j}'] = \begin{cases} \Gamma_0 & \text{si } j = 0 \\ 0 & \text{sinon} \end{cases} \quad \text{avec} \quad \Gamma_0 = \begin{pmatrix} \gamma_\mu^2 & \gamma_{\mu\varepsilon} \\ \gamma_{\mu\varepsilon} & \gamma_\varepsilon^2 \end{pmatrix} \quad (17)$$

Le paramètre de covariance $\gamma_{\mu\varepsilon}$ traduit la dépendance instantanée qui peut exister, pour un individu donné, entre les innovations des marches aléatoires x_{it} et y_{it} . Si $\gamma_{\mu\varepsilon} = 0$, les innovations des deux marches aléatoires 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 :

$$y_{it} = \alpha_i + \beta x_{it} + e_{it} \quad (18)$$

L'estimateur MCO du paramètre β (estimateur *within* ou LSDV, *Least Squares Dummy Variables*) est alors défini par la quantité :

$$\hat{\beta} = \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y}_i)(x_{it} - \bar{x}_i)}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2} \quad (19)$$

Etudions tout d'abord la convergence de $\hat{\beta}$. Kao (1999) et Phillips et Moon (1999) montrent que dans le cas d'une convergence séquentielle¹⁴, l'estimateur $\hat{\beta}$ possède les propriétés asymptotiques

¹³On exclut ici à la fois la cointégration intra et inter-individuelles.

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

suivantes :

$$\widehat{\beta} \xrightarrow[(T, N \rightarrow \infty)_{seq}]{p} \frac{\gamma_{\mu\varepsilon}}{\gamma_{\varepsilon}^2} \quad (20)$$

$$\sqrt{N} \left(\widehat{\beta} - \frac{\gamma_{\mu\varepsilon}}{\gamma_{\varepsilon}^2} \right) \xrightarrow[(T, N \rightarrow \infty)_{seq}]{l} N \left(0, \sigma_{\widehat{\beta}}^2 \right) \quad (21)$$

où $\sigma_{\widehat{\beta}}^2 = 2\gamma_{\mu}^2/5\gamma_{\varepsilon}^2$ 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 $\widehat{\beta}$ en panel converge vers sa vraie valeur $\gamma_{\mu\varepsilon}/\gamma_{\varepsilon}^2$. Dans le cas où l'on régresse en panel deux marches aléatoires totalement indépendantes ($\gamma_{\mu\varepsilon} = 0$), l'estimateur *within* du paramètre liant y_{it} et x_{it} converge ainsi en probabilité vers 0 :

$$\widehat{\beta} \xrightarrow[(T, N \rightarrow \infty)_{seq}]{p} 0 \quad (22)$$

De plus, dans ce cas, $\widehat{\beta}$ est $O_p \left(N^{-\frac{1}{2}} \right)$. 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.

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 (voir encadré).

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, montrant que le fait de

considérer un modèle de régression de panel ne change rien aux conséquences de la régression fallacieuse quant à l'inférence. Ainsi, 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_β .

Encadré - Régression Fallacieuse en Panel

Considérons 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 :

$$\widehat{\beta} = \frac{\sum_{i=1}^N \sum_{t=1}^T y_{it} x_{it}}{\sum_{i=1}^N \sum_{t=1}^T x_{it}^2} \xrightarrow{T \rightarrow \infty} \left(\frac{\gamma_\mu}{\gamma_\varepsilon} \right) \left(\frac{\frac{1}{N} \sum_{i=1}^N \int_0^1 W_{i,1}(r) W_{i,2}(r) dr}{\frac{1}{N} \sum_{i=1}^N \int_0^1 [W_{i,1}(r)]^2 dr} \right) \quad (23)$$

Ce qui peut se réécrire sous la forme :

$$\sqrt{N} \widehat{\beta} = \frac{l}{T \rightarrow \infty} \left(\frac{\gamma_\mu}{\gamma_\varepsilon} \right) \left[\sqrt{N} \frac{1}{N} \sum_{i=1}^N \int_0^1 W_{i,1}(r) W_{i,2}(r) dr \right] \left[\frac{1}{N} \sum_{i=1}^N \int_0^1 [W_{i,1}(r)]^2 dr \right]^{-1} \quad (24)$$

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 :

$$\frac{1}{N} \sum_{i=1}^N \int_0^1 [W_{i,1}(r)]^2 dr \xrightarrow[N \rightarrow \infty]{p} \frac{1}{2} \quad (25)$$

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.

$$N^{-1/2} \sum_{i=1}^N \int_0^1 W_{i,1}(r) W_{i,2}(r) dr \xrightarrow[N \rightarrow \infty]{l} N(0, 1/5) \quad (26)$$

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.

$$\sqrt{N} \widehat{\beta} \xrightarrow[(T, N \rightarrow \infty)_{seq}]{l} N \left(0, \frac{2 \gamma_\mu^2}{5 \gamma_\varepsilon^2} \right) \quad (27)$$

On retrouve alors un cas particulier du résultat général de Kao (1999) ou Phillips et Moon (1999). En outre, Kao (1999) retrouve le résultat usuel selon lequel :

$$T^{-\frac{1}{2}} t_\beta - T^{-\frac{1}{2}} \frac{\gamma_{\mu\varepsilon}}{s \gamma_\varepsilon^2} \sqrt{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2} \xrightarrow[(T, N \rightarrow \infty)_{seq}]{l} N \left(0, \frac{2}{5} \right) \quad (28)$$

$$s^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \left[(y_{it} - \bar{y}_i) - \hat{\beta} (x_{it} - \bar{x}_i) \right]^2 \quad (29)$$

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

$$T^{-\frac{1}{2}} t_{\beta} \xrightarrow[(T, N \rightarrow \infty)_{seq}]{l} N \left(0, \frac{2}{5} \right) \quad \text{ou} \quad t_{\beta} = O_p \left(T^{\frac{1}{2}} \right) \quad (30)$$

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, 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.

4 Hypothèse nulle d'absence de cointégration

Les tests d'absence de cointégration sur données de panel proposés par Pedroni (1995, 1997, 1999, 2004), 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.* (2001) 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, 2004) 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. Ainsi, sous l'hypothèse alternative, il existe une

relation de cointégration pour chaque individu, et les paramètres de cette relation de relation ne sont pas nécessairement les mêmes 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 :

$$y_{it} = \alpha_i + \delta_i t + \beta_{1i} x_{1,it} + \beta_{2i} x_{2,it} + \dots + \beta_{Mi} x_{M,it} + \varepsilon_{it} \quad (31)$$

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 : $\rho_i = 1 \forall i$, ρ_i désignant le terme autorégressif des résidus estimés sous l'hypothèse alternative tels que :

$$\hat{\varepsilon}_{it} = \rho_i \hat{\varepsilon}_{it-1} + u_{it} \quad (32)$$

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 sept 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 ν -statistic) :

$$T^2 N^{3/2} Z_{\hat{\nu}, N, T} \equiv T^2 N^{3/2} \left(\sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{\varepsilon}_{it-1}^2 \right)^{-1} \quad (33)$$

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

$$TN^{1/2} Z_{\hat{\rho}, N, T-1} \equiv TN^{1/2} \left(\sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{\varepsilon}_{it-1}^2 \right)^{-1} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \left(\hat{\varepsilon}_{it-1} \Delta \hat{\varepsilon}_{it} - \hat{\lambda}_i \right) \quad (34)$$

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

$$Z_{t, N, T} \equiv \left(\hat{\sigma}_{N, T}^2 \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{\varepsilon}_{it-1}^2 \right)^{-1/2} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \left(\hat{\varepsilon}_{it-1} \Delta \hat{\varepsilon}_{it} - \hat{\lambda}_i \right) \quad (35)$$

- Test paramétrique du type de la statistique t de Dickey-Fuller Augmenté (panel t -statistic)¹⁵ :

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

- Tests basés sur la dimension *between* (group mean panel cointegration statistics) :

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

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

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

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

- Test paramétrique du type de la statistique t de Dickey-Fuller Augmenté (group t -statistic)¹⁶ :

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

avec :

$$\hat{\lambda}_i = T^{-1} \sum_{s=1}^{k_i} \left(1 - \frac{s}{k_i + 1} \right) \sum_{t=s+1}^T \hat{\mu}_{it} \hat{\mu}_{it-s}, \quad \hat{s}_i^2 \equiv T^{-1} \sum_{t=1}^T \hat{\mu}_{it}^2 \quad (40)$$

$$\hat{\sigma}_i^2 = \hat{s}_i^2 + 2\hat{\lambda}_i, \quad \hat{\sigma}_{N,T}^2 \equiv N^{-1} \sum_{i=1}^N \hat{L}_{11i}^{-2} \hat{\sigma}_i^2, \quad \hat{s}_i^{*2} \equiv T^{-1} \sum_{t=1}^T \hat{\mu}_{it}^{*2} \quad (41)$$

$$\hat{s}_{N,T}^{*2} \equiv N^{-1} \sum_{i=1}^N \hat{s}_i^{*2}, \quad \hat{L}_{11i}^2 = T^{-1} \sum_{t=1}^T \hat{\eta}_{it}^2 + 2T^{-1} \sum_{s=1}^{k_i} \left(1 - \frac{s}{k_i + 1} \right) \sum_{t=s+1}^T \hat{\eta}_{it} \hat{\eta}_{it-s} \quad (42)$$

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

Dans ces sept tests, les statistiques sont construites sur la base des résidus des relations de cointégration (31) 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 cinq étapes¹⁷ :

¹⁵Ce 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.

¹⁶Ce 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.

¹⁷Notons 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 .

- Etape 1. On estime la relation de long terme (31) 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 :
 - 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 sept statistiques suit une loi normale centrée réduite pour T et N suffisamment importants :

$$\frac{\varkappa_{N,T} - \mu\sqrt{N}}{\sqrt{v}} \xrightarrow[N,T \rightarrow \infty]{l} N(0, 1) \quad (43)$$

où $\varkappa_{N,T}$ désigne une des sept statistiques normalisées. Pedroni (1999) a tabulé les valeurs des moments μ 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'une tendance 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).

Afin de comparer les différents tests entre eux, synthétisons les résultats issus des simulations effectuées par Pedroni (1997, 2004). Concernant tout d'abord la taille des tests, les conclusions peuvent être résumées comme suit :

- Dans le cas où N est fixe ($N = 20$) et où T varie, les tests du type t -statistics exhibent une taille relativement importante (en excès de l'ordre de 10% par rapport à la taille théorique pour $T = 40$), alors que les autres tests sont caractérisés par une taille inférieure à la taille théorique.
- Dans le cas où T est fixe ($T = 250$) et N varie, les tests panel ν -statistic et panel ρ -statistic affichent une taille au dessus de la taille théorique pour de faibles valeurs de N , alors que la taille des tests du type t -statistics est très proche de la taille théorique, même pour de faibles valeurs de la dimension individuelle.
- Lorsque N et T varient à des taux différents ($N = T^{1/2}$ à $N = T^{5/6}$), le test panel ρ -statistic converge vers la taille théorique très rapidement pour $N = T^{3/4}$ et plus lentement pour

$N = T^{1/2}$. Pour ce qui est du test group ρ -statistic, les meilleurs résultats sont obtenus pour $N = T^{5/6}$.

Concernant à présent la puissance des tests, les simulations menées par Pedroni montrent que :

- Lorsque les résidus de la relation de cointégration suivent un processus $AR(1)$ avec un coefficient autorégressif égal à 0,9, N est fixe ($N = 20$) et T varie, la puissance de tous les tests augmente rapidement lorsque T augmente ; le test atteignant la puissance maximale le plus lentement étant le test group ρ -statistic. Les autres tests atteignent la puissance de 100% pour une valeur de T égale à 50.
- Lorsque le coefficient autorégressif augmente et prend la valeur 0,95, les résultats précédents restent globalement valables, même si la taille T nécessaire pour atteindre la puissance maximale est désormais plus élevée. A titre d'exemple, le test panel ν -statistic atteint une puissance égale à 100% pour $T = 90$.
- Enfin, dans le cas où le coefficient autorégressif est très proche de 1 (0,99), c'est-à-dire que l'hypothèse alternative est très proche de l'hypothèse nulle, la taille de l'échantillon nécessaire pour atteindre une puissance égale à 100% est nettement plus importante : $T = 350$ pour le test panel ν -statistic et $T = 500$ pour le test panel ρ -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 :

$$\Omega = \begin{pmatrix} \sigma_{0\mu}^2 & \sigma_{0\mu\nu} \\ \sigma_{0\mu\nu} & \sigma_{0\nu}^2 \end{pmatrix} \quad (44)$$

et soit Σ la matrice telle que :

$$\Sigma = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E(w_{it}w'_{it}) = \begin{pmatrix} \sigma_{\mu}^2 & \sigma_{\mu\nu} \\ \sigma_{\mu\nu} & \sigma_{\nu}^2 \end{pmatrix} \quad (45)$$

Kao (1999) considère le modèle suivant¹⁸ avec effets individuels :

$$y_{it} = \alpha_i + \beta x_{it} + \varepsilon_{it} \quad (46)$$

¹⁸Notons que Kao (1999) a également considéré le cas où une tendance déterministe est présente.

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 (46) :

$$\hat{\varepsilon}_{it} = \rho \hat{\varepsilon}_{it-1} + u_{it} \quad (47)$$

qui consiste à tester l'hypothèse nulle $\rho = 1$. Soit $\hat{\rho}$ l'estimateur MCO de ρ et soit t_ρ la t -statistique correspondante définie par :

$$t_\rho = \frac{(\hat{\rho} - 1) \sqrt{\sum_{i=1}^N \sum_{t=2}^T \hat{\varepsilon}_{it-1}^2}}{s_e} \quad (48)$$

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

$$DF_\rho^* = \frac{\sqrt{NT}(\hat{\rho} - 1) + \frac{3\sqrt{N}\hat{\sigma}_u^2}{\hat{\sigma}_{0u}^2}}{\sqrt{3 + \frac{36\hat{\sigma}_u^4}{5\hat{\sigma}_{0u}^4}}} \quad (49)$$

$$DF_t^* = \frac{t_\rho + \frac{\sqrt{6N}\hat{\sigma}_u}{2\hat{\sigma}_{0u}}}{\sqrt{\frac{\hat{\sigma}_{0u}^2}{2\hat{\sigma}_u^2} + \frac{3\hat{\sigma}_u^2}{10\hat{\sigma}_{0u}^2}}} \quad (50)$$

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

$$DF_t = \sqrt{\frac{5t_\rho}{4}} + \sqrt{\frac{15N}{8}} \quad (52)$$

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 d'absence de cointégration. Les statistiques DF_ρ^* 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_ρ 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 :

$$\hat{\varepsilon}_{it} = \rho \hat{\varepsilon}_{it-1} + \sum_{j=1}^p \varphi_j \Delta \hat{\varepsilon}_{it-j} + u_{itp} \quad (53)$$

La statistique de test est donnée par :

$$ADF = \frac{t_{ADF} + \frac{\sqrt{6N}\hat{\sigma}_u}{2\hat{\sigma}_{0\mu}}}{\sqrt{\frac{\hat{\sigma}_{0u}^2}{2\hat{\sigma}_u^2} + \frac{3\hat{\sigma}_u^2}{10\hat{\sigma}_{0u}^2}}} \quad (54)$$

où t_{ADF} est la t -statistique de ρ dans (53). 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_ρ^* 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_ρ^* 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_ρ^* et DF_t^* sont plus performants que les tests DF_ρ , 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 unitaire¹⁹ 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)²⁰. 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 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 individus²¹.

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

$$y_{it} = c_i + \lambda_i' \hat{F}_t + e_{it} \quad (55)$$

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

$$y_{it} = c_i + \beta_i t + \lambda_i' \hat{F}_t + e_{it} \quad (56)$$

si l'on rajoute une tendance déterministe. \hat{F}_t désigne l'estimation du vecteur ($r \times 1$) de facteurs

¹⁹Pour un exposé détaillé de ce test, voir Hurlin et Mignon (2005).

²⁰Cette procédure est dénommée PANIC (*Panel Analysis of Nonstationarity in the Idiosyncratic and Common components*) par les auteurs.

²¹Ce point est très important dans l'application des tests de racine unitaire et de cointégration sur données de panel (voir Hurlin et Mignon, 2005, 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).

communs et l'erreur e_{it} est la composante individuelle ; la composante déterministe étant soit une constante (c_i), soit une tendance ($c_i + \beta_i t$).

On note $DF_e^c(i)$ et $DF_e^\tau(i)$ les tests de Dickey-Fuller appliqués aux résidus estimés respectivement des modèles (55) et (56). 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 $DF_e^\tau(i)$ 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 une tendance 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 Une note sur les tests de cointégration multiple

Les tests que nous avons précédemment présentés reposent sur l'existence d'au plus une seule relation de cointégration pour chaque individu du panel et supposent que l'ensemble des variables intervenant dans cette relation est *a priori* connu. Il est cependant possible, à l'instar des travaux de Johansen (1991) en séries temporelles, de recourir à une approche vectorielle multiple de la cointégration en panel. Une telle approche permet de détecter l'existence de plusieurs relations de cointégration (cointégration multiple), sans supposer que le rang de cointégration est *a priori* connu. Dans cette catégorie de tests figurent notamment ceux proposés par Larsson *et al.* (2001), Groen et Kleibergen (2003) ou encore Breitung (2005).

Groen et Kleibergen (2003) ont ainsi 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 hétérogènes²². Il convient de souligner que, même si la présence de relations de cointégration intra-individuelles hétérogènes est autorisée, le rang de cointégration est supposé être le même pour tous les individus du panel. 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 :

$$\overline{LR}(r|k) = \frac{1}{N} \sum_{i=1}^N LR_i(r|k) \quad (57)$$

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

²²Il convient de noter que Larsson *et al.* (2001) 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 .

nulle de r relations de cointégration, contre l'hypothèse alternative de k relations de cointégration, avec $r = 0, 1, \dots, k - 1$ ²³. On peut alors construire la statistique suivante :

$$\frac{\overline{LR}(r|k) - E[\overline{LR}(r|k)]}{V[\overline{LR}(r|k)]} \quad (58)$$

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

L'intérêt de ces tests est qu'ils permettent de détecter des relations de cointégration multiples au sein du panel considéré. Toutefois, comme le souligne Banerjee (1999), des recherches futures dans ce domaine sont nécessaires car divers résultats ne sont pas clairement établis. Tel est notamment le cas pour le test de Larsson *et al.* (2001) dont le théorème principal est supposé et non démontré.

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 terme²⁴ :

$$y_{it} = \alpha_i + x'_{it}\beta_i + \varepsilon_{it} \quad (59)$$

avec $\beta_i = (\beta_{1i}, \beta_{2i}, \dots, \beta_{Mi})'$, $x_{it} = (x_{1,it}, x_{2,it}, \dots, x_{M,it})'$, $i = 1, \dots, N$, $t = 1, \dots, T$ et M désigne le nombre de régresseurs. On suppose que ε_{it} peut être décomposé en deux éléments :

$$\varepsilon_{it} = \theta \sum_{j=1}^t u_{ij} + u_{it} \quad (60)$$

où u_{it} est *i.i.d.* $N(0, \sigma_u^2)$ et :

$$x_{it} = x_{it-1} + e_{it} \quad (61)$$

Tester l'hypothèse nulle de cointégration revient à tester $\theta = 0$. Sous l'hypothèse nulle, la relation (59) est une relation de cointégration. La matrice de variance-covariance de long terme de $\omega_{it} = (u_{it}, e'_{it})'$ est donnée par :

$$\Omega = \begin{pmatrix} \varpi_1^2 & \varpi_{12} \\ \varpi_{21} & \Omega_{22} \end{pmatrix} \quad (62)$$

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

$$LM = \frac{\frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} \sum_{t=1}^T S_{it}^2}{\hat{\omega}_{1,2}^2} \quad (63)$$

²³Le test peut alors s'effectuer de manière séquentielle, partant de $r = 0$ jusqu'à $r = k - 1$.

²⁴Il est également possible d'introduire une tendance spécifique dans la relation de long terme (voir McCoskey et Kao, 1998).

où $\hat{\varpi}_{1,2}^2$ 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 :

$$S_{it} = \sum_{k=1}^t \left(y_{ik}^+ - \alpha_i - \hat{\beta}_i^{+'} x_{ik} \right) \quad (64)$$

avec $y_{ik}^+ = y_{ik} - \hat{\varpi}_{12}\hat{\Omega}_{22}^{-1}e_{ik}$. $\hat{\beta}_i^+$ désigne l'estimateur FM-OLS (*Fully Modified Ordinary Least Squares*)²⁵ de β_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 (59) 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 :

$$LM^+ = \frac{\sqrt{N}(LM - \mu_v)}{\sigma_v} \quad (65)$$

où μ_v et σ_v sont la moyenne et l'écart type de fonctions de mouvements browniens²⁶. 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.

6 Simulations, estimation et inférence

6.1 Simulations : quelques résultats

Il existe assez peu de travaux proposant une comparaison des divers tests de cointégration en termes de taille et de puissance. Nous nous proposons ici de résumer les principaux résultats issus des simulations de Monte Carlo menées par Gutierrez (2003). Ce dernier compare la puissance des tests de Pedroni (1999), Kao (1999) et de Larsson *et al.* (2001). Il considère diverses valeurs pour la dimension individuelle N et la dimension temporelle T (10, 25, 50 et 100) et fait également varier le pourcentage relatif au nombre de relations qui sont cointégrées (de 0 à 100% relations de cointégration). Les résultats obtenus par Gutierrez (2003) peuvent être synthétisés comme suit :

²⁵ Voir *infra*.

²⁶ Ces constantes peuvent être obtenues par simulations de Monte Carlo. Elles dépendent uniquement du nombre de régresseurs M et de la présence ou non de termes constants et/ou de tendances dans la relation de cointégration. Elles prennent les valeurs suivantes : $\mu_v = 0,1162$ et $\sigma_v^2 = 0,0109$ (voir McCoskey et Kao, 1998).

- La puissance de tous les tests croît avec la dimension temporelle, pour une valeur fixe de N .
- Lorsque la dimension temporelle est fixée et faible ($T = 10$) et que N augmente, les tests de Kao sont plus puissants que les tests de Pedroni. Pour ces derniers, la puissance reste relativement faible, même lorsque la dimension individuelle est égale à 100.
- Les tests de Kao dominent ceux de Pedroni en termes de puissance lorsque la dimension temporelle est faible, et les tests de Pedroni sont plus puissants que ceux de Kao lorsque T augmente. Ainsi, plus la taille de l'échantillon augmente, plus les tests de Pedroni dominent ceux de Kao.
- Concernant le test de cointégration multiple de Larsson *et al.* (2001), la puissance et la taille sont proches pour de faibles valeurs de T . La puissance du test tend à augmenter lorsque la dimension temporelle croît, mais les tests de Kao et de Pedroni restent toujours plus puissants.

De façon générale, Gutierrez (2003) montre que la puissance de tous les tests tend à augmenter lorsque N et T augmentent et lorsque le nombre de relations de cointégration au sein du panel croît.

6.2 Estimation et inférence

Bien que les estimateurs MCO des vecteurs de cointégration sont super-convergeants, 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).

Pour estimer des systèmes de variables cointégrées, tout comme pour effectuer des tests sur les vecteurs de cointégration, il est en conséquence 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.

6.2.1 La méthode FM-OLS

Cette procédure, étudiée notamment par Pedroni (1996), permet de tenir compte des problèmes d'endogénéité du second ordre des régresseurs (engendrée par la corrélation entre le résidu de

cointégration et les innovations des variables $I(1)$ présentes dans la relation de cointégration) et des propriétés d'autocorrélation et d'hétéroscédasticité des résidus. Considérons le système cointégré suivant²⁷ :

$$y_{it} = \alpha_i + \beta x_{it} + \varepsilon_{it} \quad (66)$$

avec :

$$x_{it} = x_{it-1} + u_{it} \quad (67)$$

pour $i = 1, \dots, N$. Le vecteur des erreurs $(\varepsilon_{it}, u_{it})'$ est stationnaire de matrice de variance-covariance asymptotique Ω_i , que l'on peut décomposer comme suit :

$$\Omega_i = \Omega_i^0 + \Gamma_i + \Gamma_i' \quad (68)$$

où Ω_i^0 désigne la covariance contemporaine et Γ_i est une somme pondérée d'autocovariances. β est le vecteur de cointégration, le vecteur $(y_{it}, x_{it})'$ étant intégré d'ordre 1.

L'estimateur FM-OMS $\hat{\beta}_{FM}$ de β est donné par (voir Pedroni, 1996) :

$$\hat{\beta}_{FM} = \left(\sum_{i=1}^N \hat{L}_{22i}^{-2} \sum_{t=1}^T (x_{it} - \bar{x}_i)^2 \right)^{-1} \sum_{i=1}^N \hat{L}_{11i}^{-1} \hat{L}_{22i}^{-1} \left(\sum_{t=1}^T (x_{it} - \bar{x}_i) y_{it}^* - T \hat{\gamma}_i \right) \quad (69)$$

où :

$$y_{it}^* = (y_{it} - \bar{y}_i) - \frac{\hat{L}_{21i}}{\hat{L}_{22i}} \Delta x_{it} + \frac{\hat{L}_{11i} - \hat{L}_{22i}}{\hat{L}_{22i}} \beta (x_{it} - \bar{x}_i) \quad (70)$$

\hat{L}_i est la décomposition triangulaire inférieure de l'estimateur $\hat{\Omega}_i$ de la matrice de variance-covariance asymptotique Ω_i , \hat{L}_i étant normalisé de telle sorte que $\hat{L}_{i22} = \hat{\Omega}_{i22}^{-1/2}$, et le paramètre d'ajustement pour la corrélation sérielle $\hat{\gamma}_i$ s'écrit :

$$\hat{\gamma}_i = \hat{\Gamma}_{21i} + \hat{\Omega}_{21i}^0 - \frac{\hat{L}_{21i}}{\hat{L}_{22i}} \left(\hat{\Gamma}_{22i} + \hat{\Omega}_{22i}^0 \right) \quad (71)$$

Ainsi que l'a montré Pedroni (1996), l'estimateur (69) est distribué selon une loi normale.

Il convient par ailleurs de noter que 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. Ces tests reposent sur la méthode FM-OLS 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, comme nous l'avons précédemment mentionné, les distributions asymptotiques des estimateurs basés sur la méthode FM-OLS sont non biaisés et ne dépendent pas des paramètres de nuisance.

²⁷Dans le cas général, x_i peut désigner un vecteur de M régresseurs qui ne sont pas cointégrés entre eux. A des fins de simplification, on considère ici que x_i est univarié, les résultats pouvant être étendus sans difficulté au cas multivarié.

6.2.2 La méthode DOLS

L'approche DOLS a été initialement suggérée par Saikkonen (1991) dans le cas des séries temporelles, puis adaptée par Kao et Chiang (2000) et Mark et Sul (2003) au cas de données de panel. Cette technique consiste à inclure des valeurs avancées et retardées de Δx_{it} dans la relation de cointégration, afin d'éliminer la corrélation entre les variables explicatives et le terme d'erreur :

$$y_{it} = \alpha_i + \beta x_{it} + \sum_{k=-\infty}^{\infty} \delta_{ik} \Delta x_{it+k} + \varepsilon_{it} \quad (72)$$

En pratique, la somme infinie est bien évidemment tronquée à une valeur relativement faible du nombre de retards/avances et l'estimateur DOLS est obtenu en estimant la régression (72) par les MCO. L'estimateur DOLS a la même distribution asymptotique que l'estimateur FM-OLS.

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 données de panel.

7 Conclusion

Ce papier a proposé une revue de la littérature méthodologique consacrée aux 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] Bai, J. et Ng, S. (2004), "A PANIC attack on unit roots and cointegration", *Econometrica*, 72, pp. 1127-1177.
- [2] 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.

- [3] Banerjee, A. (1999), "Panel data unit roots and cointegration : an overview", *Oxford Bulletin of Economics and Statistics*, 61, pp. 607-630.
- [4] 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, IGER.
- [5] Banerjee, A., Marcellino, M. et Osbat, C. (2003), "Testing for PPP : should we use panel methods?", *mimeo*.
- [6] Banerjee, A. et Zanghieri, P. (2003), "A new look at the Feldstein-Horioka puzzle using an integrated panel", *Working Paper* CEPII, 2003-22.
- [7] 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.
- [8] Breitung, J. (2005), "A parametric approach to the estimation of cointegration vectors in panel data", *Econometric Reviews*, pp. 151-174
- [9] 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.
- [10] Breitung, J. et Pesaran, M.H. (2005), "Unit roots and cointegration in panels", *Deutsche Bank Discussion Paper* 42.
- [11] Canning, D. et Pedroni, P (1999), "Infrastructure and long-run economic growth", *CAER II Discussion Paper* 57, Harvard Institute for International Development.
- [12] Cechetti, S.G., Mark, N.C. et Sonora, R. (1999), "Price level convergence among United States cities : lessons from the European Central Bank", *Mimeo*.
- [13] Coakley, J., Kulasi, F. et Smith, R. (1996), "Current account solvency and the Feldstein-Horioka puzzle", *Economic Journal*, 106, pp. 620-627.
- [14] 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.
- [15] Engle, R.F. et Granger, C.W.J. (1987), "Cointegration and error-correction : representation, estimation and testing", *Econometrica*, 64, pp. 813-836.
- [16] Evans, P. et Karras, G. (1996), "Convergence revisited", *Journal of Monetary Economics*, 37, pp. 249-265.
- [17] Fisher, R.A. (1932), *Statistical methods for research workers*, Oliver&Boyd, Edinburgh.
- [18] Funk, M. (1998), "Trade and international R&D spillovers among OECD countries", *Working Paper*, Department of Economics, St. Louis University.
- [19] 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.

- [20] Granger, C.W.J. et Newbold, P. (1974), “Spurious regressions in econometrics”, *Journal of Econometrics*, 2, pp. 111-120.
- [21] 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.
- [22] Gutierrez, L. (2003), “On the power of panel cointegration tests : a Monte Carlo comparison”, *Economics Letters*, 80, pp. 105-111.
- [23] Hamilton, J.D. (1994), *Time Series Analysis*, Princeton.
- [24] Holtz-Eakin, D., Newey, W. and Rosen, H.S. (1988), “Estimating vector autoregressions with panel data”, *Econometrica*, 56, 1371-1396.
- [25] Hurlin, C. et Mignon, V. (2005), “Une synthèse des tests de racine unitaire sur données de panel”, *Economie et Prévision*, 169-170-171, pp. 253-294
- [26] Im, K.S., Pesaran, M.H. et Shin, Y. (2003), “Testing for unit roots in heterogeneous panels”, *Journal of Econometrics*, 115, pp. 53-74.
- [27] Islam, N. (1995), “Growth empirics : a panel data approach”, *Quarterly Journal of Economics*, 110, pp. 1127-1170.
- [28] Johansen, S. (1991), “Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models”, *Econometrica*, 59, pp. 1551-1580.
- [29] Johansen, S. (1995), *Likelihood-based inference in cointegrated vector autoregressive models*, Oxford University Press.
- [30] Kao, C. (1999), “Spurious regression and residual-based tests for cointegration in panel data”, *Journal of Econometrics*, 90, pp. 1-44.
- [31] 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.
- [32] 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.
- [33] 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.
- [34] Larsson, R., Lyhagen, J. et Löthgren, M. (2001), “Likelihood-based cointegration tests in heterogeneous panels”, *Econometrics Journal*, 4, pp. 109-142.

- [35] 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.
- [36] 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.
- [37] 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.
- [38] 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.
- [39] 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.
- [40] Mark, N.C. et Sul, D. (2003), "Cointegration vector estimation by panel DOLS and long-run money demand", *Oxford Bulletin of Economics and Statistics*, 65, pp. 655-680.
- [41] McCoskey, S. et Kao, C. (1998), "A residual-based test of the null of cointegration in panel data", *Econometric Reviews*, 17, pp. 57-84.
- [42] 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.
- [43] 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.
- [44] 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.
- [45] 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.
- [46] McDonald, R. (1996), "Panel unit root tests and real exchange rates", *Economic Letters*, 50, pp. 7-11.
- [47] 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.
- [48] Oh, K. (1996), "Purchasing power parity and unit root tests using panel data", *Journal of International Money and Finance*, 15, pp. 405-418.
- [49] 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.

- [50] Pedroni, P. (1996), “Fully modified OLS for heterogeneous cointegrated panels and the case of purchasing power parity”, *Working Paper in Economics*, Indiana University.
- [51] 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.
- [52] Pedroni, P. (1998), “Testing for convergence to common steady states in nonstationary heterogeneous panels”, *Working Paper in Economics*, Indiana University.
- [53] 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.
- [54] 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.
- [55] Pedroni, P. (2004), “Panel cointegration. Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis”, *Econometric Theory*, 20(3), pp. 597-625.
- [56] Phillips, P.C.B. (1986), “Understanding spurious regressions in econometrics”, *Journal of Econometrics*, 33, pp. 311-340.
- [57] 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.
- [58] Phillips, P.C.B. et Moon, H.R. (1999), “Linear regression limit theory for nonstationary panel data”, *Econometrica*, 67, pp. 1057-1111.
- [59] 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.
- [60] Phillips, P.C.B. et Ouliaris, S. (1990), “Asymptotic properties of residual based tests for cointegration”, *Econometrica*, 58, pp. 165-193.
- [61] Quah, D. (1992), “The relative importance of permanent and transitory components : identification and some theoretical bounds”, *Econometrica*, 60, pp. 107-118.
- [62] Quah, D. (1994), “Exploiting cross-section variation for unit root inference in dynamic data”, *Economics Letters*, 44, pp. 9-19.
- [63] Rogoff, K. (1996), “The purchasing power parity puzzle”, *Journal of Economic Literature*, 34, pp. 647-668.
- [64] Saikkonen, P. (1991), “Asymptotically efficient estimation of cointegrating regressions”, *Econometric Theory*, 58, pp. 1-21.

- [65] Sala-i-Martin, X. (1996), "The classical approach to convergence analysis", *Economic Journal*, 106, pp. 1019-1036.
- [66] Shin, Y. (1994), "A residual-based test of the null of cointegration against the alternative of no cointegration", *Econometric Theory*, 19, pp. 91-115.
- [67] Stock, J. et Watson, M. (1993), "A simple estimator of cointegrating vectors in higher order integrated systems", *Econometrica*, 61, pp. 783-820.
- [68] 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.
- [69] Yule, G.U. (1926), "Why do we sometimes get nonsense correlation between time series?", *Journal of the Royal Statistical Society*, 89, pp. 1-69.