

# Une Synthèse des Tests de Racine Unitaire sur Données de Panel

Christophe Hurlin\* et Valérie Mignon†

Janvier 2005

## Résumé

Cet article propose une synthèse de la littérature concernant les tests de racine unitaire en panel. Deux principales évolutions peuvent être mises en évidence dans cette voie de recherche depuis les travaux fondateurs de Levin et Lin (1992). D'une part, on a pu assister depuis la fin des années 90 à une évolution tendant à prendre en compte une hétérogénéité des propriétés dynamiques des séries étudiées, avec notamment les travaux d'Im, Pesaran et Shin (1997) et de Maddala et Wu (1999). D'autre part, un second type de développements récents dans cette littérature tend à introduire une dichotomie entre deux générations de tests : la première génération repose sur une hypothèse d'indépendance entre les individus, ce qui apparaît peu plausible notamment dans le cas de certaines applications macro-économiques. La seconde génération, actuellement en plein développement, intègre diverses formes possibles de dépendances inter-individuelles (Bai et Ng (2001), Phillips et Sul (2003a), Moon et Perron (2004), Choi (2002), Pesaran (2003) et Chang (2002)). Ces deux générations de tests sont présentées dans cette revue de la littérature.

- *Mots-Clé* : Données de panel non stationnaires, racine unitaire, hétérogénéité, dépendances inter-individuelles.
- *J.E.L Classification* : C23, C33

---

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

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

# 1 Introduction

L'étude des séries temporelles non stationnaires est devenue aujourd'hui incontournable dans la pratique économétrique courante. Les travaux empiriques débutent ainsi fréquemment par une analyse de la stationnarité des séries temporelles considérées avec l'application de divers tests de racine unitaire. Dans un contexte multivarié, les études cherchent souvent à mettre en évidence des relations d'équilibre de long terme entre les variables par l'application de tests de cointégration. En revanche, l'analyse des données de panel non stationnaires ne s'est développée que très récemment, depuis les travaux pionniers de Levin et Lin (1992). Elle s'est en particulier développée avec l'utilisation croissante des bases de données macro-économiques présentant une dimension temporelle suffisante (supérieure à vingt ans) pour que cette problématique présente un intérêt appliqué. Les champs d'application des tests de racine unitaire en panel couvrent aujourd'hui l'étude de la parité des pouvoirs d'achat (PPA), les problèmes de croissance et de convergence, les dynamiques de l'épargne et de l'investissement, les activités de recherche-développement au niveau international, *etc.*

## Pourquoi tester la racine unitaire en panel ?

L'ajout de la dimension individuelle à la dimension temporelle usuelle présente un intérêt important pour l'analyse des séries non stationnaires. Les tests de racine unitaire et de cointégration sur données de panel temporelles sont en effet plus puissants que leurs analogues sur séries temporelles individuelles en petit échantillon. A titre d'exemple, lorsque l'on étudie des séries de taux de change, on dispose généralement de 20 ou 30 ans de données. Or, pour de telles tailles d'échantillon, les tests de racine unitaire sont en général très peu puissants pour distinguer des séries non stationnaires et des séries stationnaires mais fortement persistantes (voir Salanié 1999 pour une synthèse). L'étendue de la période d'étude étant plus importante que la fréquence des données (Pierce et Snell, 1995), une solution consisterait — lorsque cela est possible — à accroître cette période. Mais, d'une part, cela est rarement possible et, d'autre part, une telle procédure augmente le risque de faire face à des ruptures structurelles. Si l'on reprend l'exemple des taux de change, la fin du système de Bretton Woods a engendré une rupture structurelle telle qu'il est peu pertinent de travailler sur une longue période mêlant divers régimes de change. Plutôt que d'étendre la période d'étude, une alternative consiste alors à accroître le nombre de données en incluant l'information relative à des pays différents. Il est en effet naturel de penser que les propriétés de long terme des séries, de même que leurs caractéristiques en termes de stationnarité, ont une forte probabilité d'être communes à plusieurs pays. Le recours aux données de panel permet ainsi de travailler sur des échantillons de taille réduite (dans la dimension temporelle) en augmentant le nombre de données disponibles (dans la dimension individuelle), diminuant ainsi la probabilité de faire face à des ruptures structurelles et palliant le problème de la faible puissance des tests en petit échantillon. Ainsi que le notent Baltagi et Kao (2000), l'économétrie des données de panel non stationnaires vise à combiner le "meilleur des deux mondes" : le traitement des séries non stationnaires à l'aide des méthodes des séries temporelles et l'augmentation du nombre de données et de la puissance des tests avec le recours à la dimension individuelle.

## Quelles sont les différences fondamentales entre les tests de racine unitaire en panel et en séries temporelles ?

Il existe deux principales différences entre les deux approches. La première concerne les distributions asymptotiques. Dans le cas des séries temporelles, les statistiques de tests usuels possèdent des distributions asymptotiques non standard (point *a*) et conditionnelles au modèle utilisé pour tester la racine unitaire (point *b*). Ainsi, par exemple, sous l'hypothèse nulle de racine unitaire, la *t*-statistique de Dickey-Fuller admet une distribution asymptotique qui est une fonction de mouvements Browniens. Cette distribution est en outre différente suivant la spécification de la composante déterministe du modèle dans lequel on teste la racine unitaire :

modèle avec une constante, une constante et une tendance, ou ni constante ni tendance. Dans le cadre des modèles de panel, les statistiques des tests de racine unitaire (à l'exception des tests de Fisher) admettent pour loi asymptotique des lois normales (point *a*). Les premiers à mettre en évidence ce résultat furent Levin et Lin (1992). Nous donnerons l'intuition de ce résultat dans le cadre des tests de première génération en évoquant au passage les différentes notions de convergence en panel (Phillips et Moon, 1999). En revanche, tout comme dans le cas des séries temporelles, ces lois asymptotiques demeurent conditionnelles au modèle utilisé pour tester la racine unitaire (point *b*) : les lois asymptotiques, et donc les seuils, ne sont pas les mêmes suivant que l'on considère un modèle avec ou sans constante, avec ou sans tendance. Mais, contrairement aux séries temporelles, les lois asymptotiques restent des lois normales : seules la variance et l'espérance sont modifiées suivant le modèle utilisé. Ainsi, la première différence réside dans le fait que l'économètre appliqué qui utilisait des tables de lois particulières dans le cadre des séries temporelles pour mener à bien ses tests, devra utiliser pour effectuer des tests similaires en panel les tables de la loi normale.

Mais la différence essentielle réside dans le problème de l'hétérogénéité du modèle qui ne se pose pas dans le contexte des séries temporelles. Dans le cas univarié, on se donne un modèle pour tester la présence d'une racine unitaire dans la dynamique d'une variable pour un individu donné. Lorsque l'on passe en panel, la question qui vient immédiatement à l'esprit est la suivante : peut-on considérer un même modèle pour tester la présence d'une racine unitaire dans la dynamique d'une variable observée sur plusieurs individus ? Si l'on répond par l'affirmative à cette question, cela implique l'existence de propriétés dynamiques strictement identiques pour la variable quel que soit le pays considéré : on parle alors de panel homogène. Prenons un exemple : si l'on teste la présence d'une racine unitaire dans la dynamique du chômage des 30 pays membres de l'OCDE à partir d'un même modèle, cela revient à tester l'hypothèse de racine unitaire en imposant, sans doute à tort, que la dynamique du chômage puisse être décrite pour les 30 pays par le même modèle. Le test de racine unitaire ainsi construit en panel peut alors conduire à des résultats fallacieux. Dès lors, de façon générale, la première question centrale des tests de racine unitaire en panel devient celle de la forme de l'hétérogénéité du modèle utilisé pour tester la racine unitaire. Cette question rejoint celle de la spécification même du test de racine unitaire comme nous le verrons par la suite.

Cette notion d'hétérogénéité du modèle est une notion centrale de l'économétrie des panels (Hsiao, 1986, et Sevestre, 2002). La forme la plus simple d'hétérogénéité est celle qui consiste à postuler l'existence de constantes spécifiques à chaque individu. Il s'agit bien entendu du modèle à effets individuels (spécifiés de façon fixe ou aléatoire), qui traduit une hétérogénéité uniquement du niveau moyen mais qui conserve l'hypothèse d'homogénéité des autres paramètres du modèle et en particulier de la racine autorégressive. C'est notamment ce type de modélisation qu'utiliseront les premiers tests de racine unitaire de Levin et Lin (1992). Mais très rapidement, cette conception de l'hétérogénéité limitée aux seuls effets individuels ou aux tendances déterministes est apparue peu plausible dans le cas des applications macro-économiques.

Ainsi, un premier mouvement s'est dessiné dans le contexte des tests de première génération allant vers une prise en compte d'une plus grande hétérogénéité de la dynamique de la série étudiée. Ce mouvement s'inscrit dans le contexte plus large de la littérature du milieu des années 90 consacrée aux panels dynamiques dits hétérogènes (Pesaran et Smith, 1995). Les premiers tests de racine unitaire sur panels hétérogènes ont alors été proposés par Im, Pesaran et Shin (1997) et Maddala et Wu (1999). Ces tests autorisent sous l'hypothèse alternative non seulement une hétérogénéité de la racine autorégressive, mais aussi une hétérogénéité quant à la présence même d'une racine unitaire dans le panel. Par exemple, si l'on raisonne sur un panel international, sous l'hypothèse alternative la variable étudiée peut être non stationnaire dans un premier groupe de pays et stationnaire pour un autre groupe non vide de pays de l'échantillon.

## L'hypothèse d'indépendance inter-individuelle : vers une nouvelle génération de tests

Cette première évolution vers des spécifications hétérogènes se double depuis trois ans d'une seconde évolution allant dans le sens de la prise en compte des corrélations de la variable entre individus. En effet, au delà du problème de l'hétérogénéité des paramètres du modèle, une autre problématique spécifique aux données de panel est aujourd'hui devenue centrale dans la littérature sur les tests de racine unitaire : il s'agit de la prise en compte des éventuelles dépendances inter-individuelles. La question est tout simplement de savoir si l'on autorise la présence d'éventuelles corrélations entre les résidus des différents individus du panel. Selon la réponse à cette question, on peut opposer deux générations de tests comme l'indique le tableau récapitulatif 1.

TAB. 1 – Tests de Racine Unitaire en Panel

<b>Tests 1<sup>ère</sup> Génération</b>	Indépendance entre individus
<i>1- Spécification homogène de la racine autorégressive sous <math>H_1</math></i>	
	Levin et Lin (1992, 1993)
	Levin, Lin et Chu (2002)
	Harris et Tzavalis (1999)
<i>2- Spécification hétérogène de la racine autorégressive</i>	
	Im, Pesaran et Shin (1997, 2002 et 2003)
	Maddala et Wu (1999)
	Choi (1999, 2001)
	Hadri (2000)
<i>3- Test séquentiel</i>	Hénin, Jolivaldt et Nguyen (2001)
<b>Tests 2<sup>ème</sup> Génération</b>	Dépendances entre individus
<i>1- Tests fondés sur des modèles factoriels</i>	
	Bai et Ng (2001)
	Moon et Perron (2004)
	Phillips et Sul (2003a)
	Pesaran (2003)
	Choi (2002)
<i>2- Autres approches</i>	
	O'Connell (1998)
	Chang (2002, 2004)

La première génération de tests repose sur l'hypothèse d'indépendance inter-individuelle des résidus. C'est précisément cette hypothèse d'indépendance qui, comme nous allons le voir, permet d'établir très simplement les distributions statistiques de test et d'obtenir généralement des distributions asymptotiques ou semi-asymptotiques normales. Dans cette perspective, les éventuelles corrélations entre individus constituent des paramètres de nuisance. Cette hypothèse d'indépendance inter-individuelle est particulièrement gênante dans la plupart des applications macro-économiques des tests de racine unitaire, que ce soit dans le domaine des tests de convergence (Phillips et Sul, 2003) ou dans le domaine des tests de la PPA (O'Connell, 1998). Or l'application à tort des tests de première génération dans un contexte avec dépendances inter-

individuelles conduit à des distorsions de taille et à des puissances de tests très faibles (Banerjee, Marcellino et Osbat, 2000, Strauss et Yigit, 2003).

La seconde génération de tests, plus récents, tend à lever cette hypothèse d'indépendance. Ces tests renversent totalement la perspective jusqu'alors adoptée car, plutôt que de considérer les corrélations entre individus comme des paramètres de nuisance, ils proposent d'exploiter ces co-mouvements pour définir de nouvelles statistiques de test. Tout le problème consiste alors à proposer le test permettant la prise en compte la plus générale des différentes formes possibles de dépendances entre individus. Comme le souligne Quah (1994), la modélisation des dépendances inter-individuelles est délicate dans la mesure où il n'existe pas *a priori* d'ordre naturel dans les observations individuelles. C'est pourquoi de nombreux tests sont aujourd'hui proposés. La plupart s'inscrivent dans la lignée des travaux fondateurs de Bai et Ng (2001, 2004) et retiennent le cadre d'un modèle à facteurs communs. Mais si Bai et Ng (2004) proposent deux tests séparés de racine unitaire sur les composantes commune et individuelle de la série, les autres tests ne considèrent uniquement que la composante idiosyncratique. Les approches diffèrent alors suivant la méthode retenue pour extraire de la série brute la composante idiosyncratique inobservable. Ainsi, nous illustrerons successivement ces approches en étudiant les tests de Moon et Perron (2004), Choi (2002) et Pesaran (2003). Nous évoquerons enfin des tests alternatifs (Chang 2002, 2004) dans lesquels aucune forme particulière de dépendance inter-individuelle n'est postulée, et qui ne nécessitent donc pas par conséquent l'identification de facteurs communs.

Dans le cadre de cette revue de la littérature, nous présenterons tout d'abord quatre tests de racine unitaire en panel de première génération afin de bien comprendre les enjeux de l'hypothèse d'indépendance inter-individuelle. Puis nous présenterons cinq approches des tests de deuxième génération. Ces différents tests sont illustrés au travers d'une application sur les séries de PNB réels des pays de l'OCDE.

## 2 Les tests de Levin et Lin

Andrew Levin et Chien-Fu Lin dans une série de contributions (Levin et Lin 1992, Levin et Lin 1993 et Levin, Lin et Chu 2002) ont proposé le premier test de racine unitaire en panel. Leur démarche est directement inspirée de celle des tests de racine unitaire en séries temporelles de Dickey et Fuller (1979). Par conséquent, les auteurs considèrent trois modèles pour tester la racine unitaire<sup>1</sup> suivant la forme que revêt la composante déterministe :

$$\text{Modèle 1 : } \Delta y_{i,t} = \rho y_{i,t-1} + \varepsilon_{i,t} \quad (1)$$

$$\text{Modèle 2 : } \Delta y_{i,t} = \alpha_i + \rho y_{i,t-1} + \varepsilon_{i,t} \quad (2)$$

$$\text{Modèle 3 : } \Delta y_{i,t} = \alpha_i + \beta_i t + \rho y_{i,t-1} + \varepsilon_{i,t} \quad (3)$$

où  $i = 1, \dots, N$  et  $t = 1, \dots, T$  et où les termes d'erreur  $\varepsilon_{i,t}$  sont distribués indépendamment entre les individus  $i$  et suivent un processus ARMA stationnaire et inversible admettant une représentation  $AR(\infty)$  du type :

$$\varepsilon_{i,t} = \sum_{k=1}^{\infty} \theta_{i,k} \varepsilon_{i,t-k} + \mu_{i,t} \quad (4)$$

---

<sup>1</sup>Nous ne présentons ici que les trois modèles retenus dans la version finale de l'article (Levin, Lin et Chu, 2002). Ces trois modèles excluent notamment la présence d'effets temporels communs qui figuraient dans les versions de 1992 et 1993 du papier.

Les processus  $\mu_{i,t}$  pour  $i = 1, \dots, N$  sont *i.i.d.*  $(0, \sigma_{\mu,i}^2)$ . On peut faire deux remarques à ce niveau<sup>2</sup>. La première est que les trois modèles de Levin et Lin supposent l'indépendance des termes d'erreur dans la dimension individuelle. Cette hypothèse, bien que particulièrement forte, est adoptée dans tous les tests de racine unitaire en panel de première génération : c'est en effet cette hypothèse qui, comme nous le verrons par la suite, permet d'utiliser un théorème central limite pour obtenir de façon relativement simple les distributions asymptotiques (normales) des statistiques de tests.

La seconde remarque porte sur la question de l'hétérogénéité du processus générateur des données retenu par les auteurs. Comme nous l'avons mentionné en introduction, il s'agit d'un problème fondamental en économétrie de panel. Dans le cas présent, les trois modèles de Levin et Lin imposent l'hypothèse d'homogénéité de la racine autorégressive ( $\rho_i = \rho_j = \rho, \forall i, j$ ) et par conséquent l'homogénéité de la conclusion quant à la présence d'une racine unitaire dans la dynamique de la variable  $y$  : soit l'on accepte l'hypothèse d'une racine unitaire pour l'ensemble des individus du panel, soit l'on rejette l'hypothèse d'une racine unitaire pour l'ensemble des individus. C'est précisément la principale limite de ce test. En effet, même si dans les modèles 2 et 3, Levin et Lin autorisent la présence d'une hétérogénéité du processus générateur *via* l'introduction d'effets individuels fixes ( $\alpha_i \neq \alpha_j$  pour  $i \neq j$ ) et d'éventuelles tendances déterministes individuelles ( $\beta_i \neq \beta_j$  pour  $i \neq j$ ), il n'en demeure pas moins que le degré de persistance des chocs<sup>3</sup> de  $\varepsilon_{i,t}$  sur la variable  $y_{i,t}$  est supposé être le même pour tous les individus du panel. Dans le cas de panels macro-économiques, on conçoit évidemment que cette hypothèse d'homogénéité pose problème.

A partir de ces trois modèles, Levin et Lin proposent de tester les hypothèses suivantes :

$$\begin{aligned} \text{Modèle 1} \quad & : \quad H_0 : \rho = 0 \\ & \quad H_1 : \rho < 0 \end{aligned}$$

$$\begin{aligned} \text{Modèle 2} \quad & : \quad H_0 : \rho = 0 \text{ et } \alpha_i = 0, \forall i = 1, \dots, N \\ & \quad H_1 : \rho < 0 \text{ et } \alpha_i \in \mathbb{R}, \forall i = 1, \dots, N \end{aligned}$$

$$\begin{aligned} \text{Modèle 3} \quad & : \quad H_0 : \rho = 0 \text{ et } \beta_i = 0, \forall i = 1, \dots, N \\ & \quad H_1 : \rho < 0 \text{ et } \beta_i \in \mathbb{R}, \forall i = 1, \dots, N \end{aligned}$$

Il est important de noter que les hypothèses nulles des tests de Levin et Lin dans les modèles 2 et 3 sont des hypothèses jointes. Dans le modèle 2, l'hypothèse nulle testée est l'hypothèse de racine unitaire pour tous les individus du panel ( $\rho_i = \rho = 0$ ) conjointement à l'hypothèse d'absence d'effets individuels, plus précisément à la nullité de toutes les constantes individuelles ( $\alpha_i = 0$ ). Dans le modèle 3, l'hypothèse nulle consiste en l'hypothèse de racine unitaire et d'absence de composante tendancielle déterministe pour tous les individus du panel ( $\beta_i = 0$ ). On retrouve ainsi exactement la structure des deux tests joints proposés dans le cas des séries temporelles par Dickey et Fuller (1981).

A présent, nous allons donner l'intuition des principaux résultats asymptotiques de Levin et Lin dans le cadre très simple du modèle 1 en l'absence d'autocorrélation des résidus. Nous comprendrons ainsi pourquoi, dans la pratique, les tests de racine unitaire en panel de Levin et Lin doivent être réalisés à partir des seuils de la loi normale, contrairement au cas des séries temporelles. Nous considérerons ensuite le cas général avec autocorrélation des résidus, et présenterons la démarche générale en trois étapes des tests de Levin et Lin.

<sup>2</sup>En ce qui concerne la discussion sur l'éventuel choix du modèle à adopter pour tester la racine unitaire, les arguments développés dans le cadre des séries temporelles (voir Salanié, 1999) peuvent être repris dans le contexte des données de panel.

<sup>3</sup>En revanche, l'écriture hétérogène ( $\theta_{i,k} \neq \theta_{j,k}$ ) de la forme ARMA des processus  $\varepsilon_{i,t}$  permet d'envisager, dans le cas stationnaire, des degrés de persistance différents pour les chocs sur  $\mu_{i,t}$ .

## 2.1 Des lois asymptotiques normales : illustration dans un cas simple

Considérons le modèle 1 de Levin et Lin sans effet individuel, en l'absence d'autocorrélation des résidus ( $\theta_{i,k} = 0$ ).

$$y_{i,t} = \phi y_{i,t-1} + \mu_{i,t}$$

avec  $\phi = 1 + \rho \in \mathbb{R}$  et où, pour simplifier, on suppose que les résidus  $\mu_{i,t}$  pour  $i = 1, \dots, N$  sont *i.i.d.*  $(0, \sigma_\mu^2)$  et indépendamment distribués dans la dimension individuelle. On cherche à tester l'hypothèse nulle  $H_0 : \phi = 1$  contre  $H_1 : \phi < 1$ . Dans ce cas simple, la statistique de test de Levin et Lin, identique à celle du test de Dickey-Fuller, correspond à la t-statistique associée au test  $\phi = 1$ , notée  $t_{\phi=1}$  :

$$t_{\phi=1} = \frac{(\hat{\phi} - 1)}{\sigma_{\hat{\phi}}} = (\hat{\phi} - 1) \left( \frac{1}{s^2} \sum_{i=1}^N \sum_{t=1}^T y_{i,t-1}^2 \right)^{\frac{1}{2}} \quad (5)$$

où  $\hat{\phi}$  désigne l'estimateur des MCO du paramètre  $\phi$  obtenu à partir des  $NT$  observations empilées. Sous l'hypothèse nulle  $H_0$ , le biais de cet estimateur peut s'écrire sous la forme suivante :

$$\hat{\phi} - 1 = \left( \sum_{i=1}^N \sum_{t=1}^T y_{i,t-1}^2 \right)^{-1} \left( \sum_{i=1}^N \sum_{t=1}^T y_{i,t-1} \mu_{i,t} \right) \quad (6)$$

L'estimateur de la variance des résidus  $s$  est défini de façon usuelle par :

$$s^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \hat{\mu}_{i,t}^2 = \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{T} \sum_{t=1}^T \hat{\mu}_{i,t}^2 \right) = \frac{1}{N} \sum_{i=1}^N s_i^2$$

En économétrie de panel, il existe plusieurs concepts de distribution asymptotique. On distingue tout d'abord ce que l'on qualifie parfois de distribution semi-asymptotique ; c'est le cas où l'une des dimensions ( $N$  ou  $T$ ) est fixe et l'autre tend vers l'infini (voir Hsiao, 1986). En ce qui concerne la distribution asymptotique au sens propre (où les deux dimensions tendent vers l'infini), on doit distinguer différents modes de convergence qui ont été clairement définis par Phillips et Moon (1999). La première approche, dite convergence séquentielle, consiste à supposer que les dimensions convergent dans un certain ordre. On raisonne 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<sup>4</sup>. Dans la seconde approche, dite convergence le long d'une diagonale, on fait tendre  $N$  et  $T$  vers l'infini simultanément sous l'hypothèse que  $T = T(N)$  est une fonction monotone de  $N$ . Par exemple, on peut supposer que le ratio  $T/N$  converge vers une constante  $c$  non nulle. Bien évidemment, l'inconvénient de cette approche réside, d'une part, dans le caractère spécifique de la fonction  $T(N)$  qui peut conditionner les résultats et, d'autre part, dans le fait que les dimensions ( $N, T$ ) d'un panel peuvent ne pas correspondre, et de loin, à la relation  $T = T(N)$ . La troisième approche, la plus générale mais aussi la plus difficile à utiliser, consiste à supposer que  $N$  et  $T$  tendent vers l'infini sans supposer une quelconque relation entre  $T$  et  $N$ .

Dans le cas de la littérature sur les tests de racine unitaire en panel, et plus particulièrement dans le cas des tests de première génération, les modes de convergence adoptés sont

<sup>4</sup>Il convient de noter que ce mode de convergence joue un rôle fondamental dans la détermination des lois asymptotiques des tests de racine unitaire. En effet, le fait que les statistiques des tests de racine unitaire en panel sont asymptotiquement distribuées selon des lois normales, impose une convergence, à  $N$  fixé, de  $T$  vers l'infini. Ceci peut paraître quelque peu contradictoire avec l'approche des tests de racine unitaire en panel dans la mesure où l'attrait pour ceux-ci réside notamment dans la faiblesse de la dimension temporelle de l'échantillon disponible.

essentiellement séquentiels et le long d'une diagonale. Nous ne considérerons ici que le cas le plus simple de la distribution asymptotique séquentielle, lorsque  $T$  tend vers l'infini, puis  $N$  tend vers l'infini à son tour.

Raisonnons tout d'abord à  $N$  fixé. On note  $\xrightarrow{d}$  la convergence en loi fonctionnelle et  $\xrightarrow{l}$  la convergence en loi. Sous l'hypothèse nulle de racine unitaire, pour chaque individu du panel les moments empiriques  $\sum_{t=1}^T y_{i,t-1} \mu_{i,t}$  et  $\sum_{t=1}^T y_{i,t}^2$ , notés respectivement  $A_i$  et  $B_i$ , convergent en distribution respectivement vers des fonctions de mouvements Browniens d'où l'on déduit les distributions usuelles des tests de Dickey-Fuller (voir par exemple Hamilton 1994, chapitre 17). Ainsi, sous l'hypothèse nulle  $\phi = 1$  :

$$A_i = \frac{1}{T^2} \sum_{t=1}^T y_{i,t-1}^2 \xrightarrow[T \rightarrow \infty]{d} \sigma_\mu^2 \int_0^1 [W_i(r)]^2 dr \quad \forall i = 1, \dots, N$$

$$B_i = \frac{1}{T} \sum_{t=1}^T y_{i,t-1} \mu_{i,t} \xrightarrow[T \rightarrow \infty]{d} \frac{\sigma_\mu^2}{2} [W_i(1)^2 - 1] \quad \forall i = 1, \dots, N$$

où les variables  $W_i(r)$ , pour  $i = 1, \dots, N$ , désignent des mouvements Browniens standards indépendants. On admet parallèlement que les estimateurs individuels  $s_i^2$ , convergent en probabilité vers la variance de la population des résidus :

$$s_i^2 = \frac{1}{T} \sum_{t=1}^T \hat{\mu}_{i,t}^2 \xrightarrow[T \rightarrow \infty]{p} \sigma_\mu^2 \quad \forall i = 1, \dots, N$$

Le fait de raisonner en panel nous amène à présent à considérer des statistiques de tests (équations 5 et 6) qui s'expriment en fonction des moyennes des moments empiriques individuels  $A_i$  et  $B_i$ . En contrôlant les vitesses de convergence, la statistique de test de Levin et Lin peut ainsi se réécrire sous la forme suivante :

$$t_{\phi=1} = \left[ T\sqrt{N} (\hat{\phi} - 1) \right] \times \frac{1}{s} \times \sqrt{\frac{1}{N} \sum_{i=1}^N A_i} \quad (7)$$

avec

$$T\sqrt{N} (\hat{\phi} - 1) = \left( \frac{1}{N} \sum_{i=1}^N A_i \right)^{-1} \left( \frac{1}{\sqrt{N}} \sum_{i=1}^N B_i \right) \quad (8)$$

C'est à ce niveau qu'intervient l'hypothèse fondamentale d'indépendance des résidus entre individus. En effet, sous cette hypothèse d'indépendance, il convient de déterminer la loi asymptotique de moyennes (sur  $N$ ) de moments empiriques individuels indépendants, à savoir les quantités  $N^{-1} \sum_{i=1}^N A_i$  et  $N^{-1/2} \sum_{i=1}^N B_i$ . Or, dès lors que la dimension  $T$  tend vers l'infini, les moments  $A_i$  pour  $i = 1, \dots, N$  (ou  $B_i$ ) sont (i) indépendamment distribués, (ii) d'espérance et de variance finie, et dans notre cas très simple (iii) identiquement distribués<sup>5</sup>. Ne reste plus alors qu'à faire tendre la dimension  $N$  vers l'infini, et appliquer le théorème central limite de Lindberg-Levy (voir Amemiya, 1985). Pour un  $N$  fixé, lorsque la dimension  $T$  tend vers l'infini on montre que :

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N [B_i - E(B_i)] \xrightarrow[T \rightarrow \infty]{d} \frac{1}{\sqrt{N}} \sum_{i=1}^N \left\{ \frac{\sigma_\mu^2}{2} [W_i(1)^2 - 1] - \frac{\sigma_\mu^2}{2} E[W_i(1)^2 - 1] \right\}$$

<sup>5</sup>Cette dernière condition n'est pas essentielle à l'application du théorème central limite et n'est d'ailleurs plus valide dès lors que l'on lève l'hypothèse d'homoscédasticité inter-individuelle  $\sigma_{\mu,i}^2 = \sigma_\mu^2$ .

Sous nos hypothèses, quand  $T \rightarrow \infty$  les moments  $B_i$  convergent vers une distribution dont l'espérance  $E(B_i)$  est nulle puisque  $E[W_i(1)^2 - 1] = 0$ . Ce résultat provient tout simplement du fait que  $W_i(1)^2$  est distribué selon un  $\chi^2(1)$ . Lorsque  $N$  tend vers l'infini, l'application du théorème central limite de Lindberg-Levy, nous permet de montrer que le terme de droite (qui est une moyenne sur  $N$  de termes *i.i.d.*) converge à son tour vers une loi normale. La variance de cette loi normale correspond tout simplement à la variance de la quantité  $(\sigma_\mu^2/2) [W_i(1)^2 - 1]$ . Sachant que  $Var[W_i(1)^2] = 2$ , on obtient par conséquent :

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \left\{ \frac{\sigma_\mu^2}{2} [W_i(1)^2 - 1] \right\} \xrightarrow[N \rightarrow \infty]{l} N \left( 0, \frac{\sigma_\mu^4}{2} \right)$$

Ainsi, on montre que la quantité  $N^{-1/2} \sum_{i=1}^N [B_i - E(B_i)]$  converge séquentiellement lorsque  $T$  puis  $N$  tendent vers l'infini vers une loi normale.

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N B_i = \frac{1}{\sqrt{N}} \sum_{i=1}^N \left( \frac{1}{T} \sum_{t=1}^T y_{i,t-1} \mu_{i,t} \right) \xrightarrow[(T, N \rightarrow \infty)_{seq}]{l} N \left( 0, \frac{\sigma_\mu^4}{2} \right) \quad (9)$$

où  $(T, N \rightarrow \infty)_{seq}$  désigne la convergence séquentielle lorsque  $T$  puis  $N$  tendent vers l'infini. C'est ainsi par ce type de raisonnement (dans le cas d'une convergence séquentielle) que l'on établit l'existence de distributions asymptotiques normales, contrairement aux distributions asymptotiques non standard que l'on obtenait traditionnellement en séries temporelles lorsque  $T \rightarrow \infty$ .

Parallèlement, on peut établir la convergence en probabilité de la quantité  $N^{-1} \sum_{i=1}^N A_i$  de la même façon. Pour un  $N$  fixé, on a :

$$\frac{1}{N} \sum_{i=1}^N A_i \xrightarrow[T \rightarrow \infty]{p} \frac{1}{N} \sum_{i=1}^N \left[ \sigma_\mu^2 \int_0^1 W_i(r)^2 dr \right]$$

Lorsque  $N$  tend vers l'infini, la loi faible des grands nombres nous donne :

$$\frac{1}{N} \sum_{i=1}^N \left[ \sigma_\mu^2 \int_0^1 W_i(r)^2 dr \right] \xrightarrow[N \rightarrow \infty]{p} \sigma_\mu^2 E \left[ \int_0^1 W_i(r)^2 dr \right]$$

Sachant que cette espérance est égale à  $1/2$ , on en déduit alors la convergence en probabilité séquentielle du moment  $\frac{1}{N} \sum_{i=1}^N A_i$ .

$$\frac{1}{N} \sum_{i=1}^N A_i = \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{T^2} \sum_{t=1}^T y_{i,t-1}^2 \right) \xrightarrow[(T, N \rightarrow \infty)_{seq}]{p} \frac{\sigma_\mu^2}{2} \quad (10)$$

Ainsi, à partir des résultats (9) et (10), on établit immédiatement la distribution asymptotique séquentielle de l'estimateur  $\hat{\phi}$  de la racine autorégressive (équation 8).

$$T\sqrt{N} (\hat{\phi} - 1) = \left( \frac{1}{N} \sum_{i=1}^N A_i \right)^{-1} \left( \frac{1}{\sqrt{N}} \sum_{i=1}^N B_i \right) \xrightarrow[(T, N \rightarrow \infty)_{seq}]{l} \frac{2}{\sigma_\mu^2} \times N \left( 0, \frac{\sigma_\mu^4}{2} \right)$$

On retrouve ainsi le résultat de Levin et Lin (1992) dans le cas du modèle sans effet individuel.

$$T\sqrt{N} (\hat{\phi} - 1) \xrightarrow[(T, N \rightarrow \infty)_{seq}]{l} N(0, 2) \quad (11)$$

Notons la vitesse de convergence  $T\sqrt{N}$  de l'estimateur  $\hat{\phi}$  que l'on retrouvera dans le cadre de l'application que ce soit des tests de première ou de deuxième génération.

Il ne reste plus alors qu'à déterminer la distribution asymptotique séquentielle de la statistique de test de Levin et Lin donnée par l'équation (7). Sachant que l'estimateur  $s^2$  converge séquentiellement en probabilité vers la variance des résidus  $\sigma_\mu^2$ , on a :

$$t_{\phi=1} = \left[ T\sqrt{N} (\hat{\phi} - 1) \right] \times \frac{1}{s} \times \sqrt{\frac{1}{N} \sum_{i=1}^N A_i} \xrightarrow{(T, N \rightarrow \infty)_{seq}} N(0, 2) \times \frac{1}{\sigma_\mu} \times \sqrt{\frac{\sigma_\mu^2}{2}}$$

On vérifie alors que sous  $H_0$  la statistique de test de racine unitaire  $t_{\phi=1}$  de Levin et Lin converge séquentiellement vers une loi normale centrée réduite :

$$\text{Modèle 1 : } t_{\phi=1} \xrightarrow{(T, N \rightarrow \infty)_{seq}} N(0, 1) \quad (12)$$

Ainsi, contrairement au cas des séries temporelles où les statistiques de Dickey-Fuller suivent des distributions non standard, la statistique de Levin et Lin en panel est asymptotiquement normalement distribuée. Le même résultat peut être établi dans le cas d'une convergence le long d'une diagonale. En ce sens, on peut dire que le panel contribue à "réhabiliter" l'usage des tables de la loi normale dans le cadre des tests de racine unitaire. On retrouve de la même façon des distributions normales dans les modèles avec effets individuels et/ou tendance déterministe. Seules les espérances et les variances de ces lois normales changent. Ainsi, Levin et Lin obtiennent les résultats suivants en l'absence d'autocorrélation des résidus :

$$\text{Modèle 2 : } \sqrt{1.25} t_{\phi=1} + \sqrt{1.875N} \xrightarrow{(T, N \rightarrow \infty)_{seq}} N(0, 1)$$

$$\text{Modèle 3 : } \sqrt{\frac{448}{277}} \left( t_{\phi=1} + \sqrt{3.75N} \right) \xrightarrow{(T, N \rightarrow \infty)_{seq}} N(0, 1)$$

Pour finir, citons les travaux de Harris et Tzavalis (1999) qui proposent une extension des tests de Levin et Lin (1992), toujours sous l'hypothèse d'absence d'autocorrélation des résidus, pour des panels de taille  $T$  finie. Toutefois, ces premiers tests ne prennent pas en compte l'éventuelle autocorrélation des résidus. C'est pourquoi à présent, nous allons nous intéresser à la mise en oeuvre des tests de Levin et Lin dans un cas général.

## 2.2 Mise en oeuvre du test de Levin et Lin dans le cas général

Dans le cas général, en présence d'une éventuelle autocorrélation des résidus ( $\theta_{i,k} \neq 0$ ), le test de Levin et Lin est construit à partir de modèles de type Dickey-Fuller Augmentés (ADF) permettant de blanchir les résidus et de se ramener à des distributions connues pour les t-statistiques individuelles.

$$\text{Modèle 1 : } \Delta y_{i,t} = \rho y_{i,t-1} + \sum_{s=1}^{p_i} \gamma_{i,s} \Delta y_{i,t-s} + \mu_{i,t} \quad (13)$$

$$\text{Modèle 2 : } \Delta y_{i,t} = \alpha_i + \rho y_{i,t-1} + \sum_{s=1}^{p_i} \gamma_{i,s} \Delta y_{i,t-s} + \mu_{i,t} \quad (14)$$

$$\text{Modèle 3 : } \Delta y_{i,t} = \alpha_i + \beta_i t + \rho y_{i,t-1} + \sum_{s=1}^{p_i} \gamma_{i,s} \Delta y_{i,t-s} + \mu_{i,t} \quad (15)$$

où  $\mu_{i,t}$  est *i.i.d.*  $(0, \sigma_{\mu,i}^2)$ . Etant donné que l'ordre des retards  $p_i$  permettant de purger l'autocorrélation des résidus est *a priori* inconnu, Levin et Lin proposent une procédure de test en trois étapes applicable dans chacun des modèles 1, 2 et 3. La manière dont Levin et Lin ont construit leur test implique que ces trois étapes ne nécessitent finalement aucune technique d'estimation propre aux données de panel : on peut donc utiliser les commandes de base de l'économétrie des séries temporelles de tout logiciel pour le mettre en oeuvre.

### Etape 1 Construction d'un estimateur de la racine autorégressive

L'objectif de cette première étape consiste à construire un estimateur  $\hat{\rho}$  de la racine autorégressive commune. Pour cela, il convient tout d'abord de déterminer, pour chaque individu, l'ordre des retards optimal  $p_i$ . Puisque à ce niveau, on raisonne sur la base de régressions individuelles indépendantes les unes des autres, les méthodes de sélection du retard  $p_i$  sont les mêmes qu'en séries temporelles. La méthode la plus simple, dite du  $p_{max}$ , consiste à se donner un ordre maximum de retard admissible  $p_{max}$  conditionnellement à la dimension  $T$  des observations et à tester la significativité du dernier retard à partir de la statistique de Student qui admet dans ce cas une distribution standard (voir Salanié, 1999). Il est également possible de recourir à une technique de vraisemblance pénalisée consistant à estimer plusieurs processus pour différentes valeurs de  $p$  et à retenir le modèle minimisant les critères d'information tels que ceux d'Akaike ou de Schwarz.

Une fois que l'on dispose pour tous les individus du panel de l'ordre optimal des retards  $p_i$ ,  $i = 1, \dots, N$ , on peut alors estimer le paramètre  $\rho$ . Levin et Lin n'estiment pas directement cette racine autorégressive à partir des modèles ADF. La raison de ce choix technique tient simplement au fait que la spécification d'un paramètre de retard  $p_i$  différent selon les individus rend malaisée l'application sous les logiciels usuels (Eviews, TSP, Limdep, Rats,...) des estimateurs de panels (estimateur MCO, communément appelé estimateur Within dans ce cas). Ainsi, plutôt que d'estimer directement les modèles (13), (14) ou (15), Levin et Lin estiment de façon équivalente deux régressions auxiliaires individu par individu. Par exemple, si l'on considère le modèle 2, il convient d'estimer pour chaque individu  $i = 1, \dots, N$ , les deux équations suivantes par MCO :

$$\text{Equation auxiliaire n°1 : } \Delta y_{i,t} = \hat{a}_i + \sum_{s=1}^{p_i} \hat{b}_{i,s} \Delta y_{i,t-s} + \hat{e}_{i,t} \quad \forall t = p_i + 2, \dots, T \quad (16)$$

$$\text{Equation auxiliaire n°2 : } y_{i,t-1} = \hat{c}_i + \sum_{s=1}^{p_i} \hat{d}_{i,s} \Delta y_{i,t-s} + \hat{v}_{i,t} \quad \forall t = p_i + 2, \dots, T \quad (17)$$

Dans le cas du modèle 1, les équations auxiliaires sont identiques à la différence près qu'il n'y a pas de constante dans la régression. Dans le cas du modèle 3, il convient au contraire de rajouter une tendance aux régressions (16) et (17). On dispose alors de  $N$  séries de réalisations des résidus individuels  $\{\hat{e}_{i,t}\}_{t=1}^{T_i}$  et  $\{\hat{v}_{i,t}\}_{t=1}^{T_i}$  où  $T_i = T - p_i - 1$  désigne le nombre d'observations disponibles pour l'individu  $i$ . On peut alors construire un estimateur convergent du paramètre  $\rho$  en projetant directement les  $\sum_{i=1}^N T_i$  réalisations des résidus  $\hat{e}_{i,t}$  sur les  $\sum_{i=1}^N T_i$  réalisations des résidus  $\hat{v}_{i,t}$ . Toutefois, afin de contrôler l'hétéroscédasticité inter-individuelle, on construit au préalable des séries de résidus normalisés :

$$\tilde{e}_{i,t} = \frac{\hat{e}_{i,t}}{\hat{\sigma}_{\mu,i}} \quad \tilde{v}_{i,t} = \frac{\hat{v}_{i,t}}{\hat{\sigma}_{\mu,i}} \quad \forall i = 1, \dots, N, \quad \forall t = p_i + 2, \dots, T \quad (18)$$

où l'estimateur de la variance individuelle des résidus  $\hat{\sigma}_{\mu,i}^2$  correspond à l'estimateur standard de la variance des résidus du modèle ADF (équation 13, 14 ou 15 suivant les cas). Pour un individu  $i$  donné, cet estimateur peut être simplement obtenu sans procéder à l'estimation du

modèle ADF, de la façon suivante :

$$\hat{\sigma}_{\mu,i}^2 = \frac{1}{T - p_i - 1} \sum_{t=p_i+2}^T (\hat{e}_{i,t} - \hat{\rho}_i \hat{v}_{i,t})^2 \quad \forall t = p_i + 2, \dots, T \quad (19)$$

où  $\hat{\rho}_i$  désigne l'estimateur des MCO du paramètre  $\rho_i$  dans la régression  $\hat{e}_{i,t} = \hat{\rho}_i \hat{v}_{i,t-1} + \zeta_{i,t}$  pour l'individu  $i$ . Ainsi, à partir des  $\sum_{i=1}^N T_i$  réalisations des résidus normalisés  $\{\tilde{e}_{i,t}\}_{t=1}^{T_i}$  et  $\{\tilde{v}_{i,t}\}_{t=1}^{T_i}$  pour  $i = 1, \dots, N$ , on construit alors un estimateur convergent  $\hat{\rho}$  de la racine autorégressive commune  $\rho$  en appliquant les MCO aux observations empilées dans le modèle :

$$\tilde{e}_{i,t} = \hat{\rho} \tilde{v}_{i,t} + \varepsilon_{i,t} \quad \forall i = 1, \dots, N, \quad \forall t = p_i + 2, \dots, T \quad (20)$$

### Etape 2 Estimation des ratios de variances individuelles

L'objectif de la deuxième étape consiste à estimer la moyenne des  $N$  ratios  $s_i$  de la variance de long terme  $\Omega_i^2$  du modèle sur la variance de court terme des résidus individuels  $\sigma_{\mu,i}^2$ , pour  $i = 1, \dots, N$  :

$$\hat{S}_N = \frac{1}{N} \sum_{i=1}^N \hat{s}_i = \frac{1}{N} \sum_{i=1}^N \frac{\hat{\Omega}_i}{\hat{\sigma}_{\mu,i}} \quad (21)$$

C'est précisément cette moyenne des ratios de variances individuelles qui va nous servir dans la dernière étape à ajuster la moyenne de la distribution de la statistique de Student du test de racine unitaire.

### Etape 3 Construction de la statistique de test de Levin et Lin

A partir de l'estimateur des MCO  $\hat{\rho}$  obtenu à l'étape 1, on construit la statistique associée au test  $H_0 : \rho = 0$ . De façon traditionnelle, on devrait utiliser pour cela la statistique de Student définie par :

$$t_{\rho=0} = \frac{\hat{\rho}}{\hat{\sigma}_{\hat{\rho}}} \quad (22)$$

où  $\hat{\sigma}_{\hat{\rho}}^2$  désigne un estimateur de la variance de  $\hat{\rho}$  défini de façon standard par la quantité :

$$\hat{\sigma}_{\hat{\rho}}^2 = \hat{\sigma}_{\varepsilon}^2 \left[ \sum_{i=1}^N \sum_{t=p_i+2}^T \tilde{v}_{i,t}^2 \right]^{-1}$$

L'estimateur  $\hat{\sigma}_{\varepsilon}^2$  de la variance des perturbations  $\varepsilon_{i,t}$  supposées être homoscédastiques, est défini de façon usuelle par la quantité :

$$\hat{\sigma}_{\varepsilon}^2 = \left( \sum_{i=1}^N T_i \right)^{-1} \left[ \sum_{i=1}^N \sum_{t=p_i+2}^T (\tilde{e}_{i,t} - \hat{\rho} \tilde{v}_{i,t})^2 \right] \quad (23)$$

Levin et Lin proposent d'approximer le nombre total d'observations  $\sum_{i=1}^N T_i$  par la quantité  $N\tilde{T} = N(T - \bar{p} - 1)$ .

Dans le cas du modèle 1 (sans constante), Levin et Lin démontrent que la statistique de test  $t_{\rho=0}$  a une distribution asymptotique normale centrée réduite sous l'hypothèse nulle de racine unitaire  $H_0 : \rho = 0$ . Toutefois, dès lors que la composante déterministe du modèle est non nulle (cas des modèles 2 ou 3), cette statistique de test diverge vers  $-\infty$  sous  $H_0$ . Il est donc nécessaire pour ces deux modèles de construire une statistique corrigée centrée et réduite

permettant de se ramener dans tous les cas sous  $H_0$  à une distribution normale centrée réduite. Ainsi, de façon générale, la statistique de test de racine unitaire de Levin et Lin s'écrit comme une statistique de Student modifiée :

$$\text{Statistiques LL : } t_{\rho=0}^* = \frac{1}{\sigma_{m,\tilde{T}}^*} \left( t_{\rho=0} - N\tilde{T} \times \frac{\hat{S}_N}{\hat{\sigma}_\varepsilon^2} \times \hat{\sigma}_{\hat{\rho}} \times \mu_{m,\tilde{T}}^* \right) \quad \forall m = 1, 2, 3 \quad (24)$$

où pour chaque modèle ( $m = 1, 2$  ou  $3$ ) suivant la dimension  $\tilde{T} = T - \bar{p} - 1$  avec  $\bar{p} = (1/N) \sum_{i=1}^N p_i$ , les auteurs proposent une valeur de la composante d'ajustement de la moyenne  $\mu_{m,\tilde{T}}^*$  et de la composante d'ajustement de la variance  $\sigma_{m,\tilde{T}}^*$ . Ces valeurs sont reportées dans le tableau 2.

TAB. 2 – Facteurs d'Ajustement de la t-Statistique Corrigée

$\tilde{T}$	$q$	$\mu_{1,\tilde{T}}^*$	$\sigma_{1,\tilde{T}}^*$	$\mu_{2,\tilde{T}}^*$	$\sigma_{2,\tilde{T}}^*$	$\mu_{3,\tilde{T}}^*$	$\sigma_{3,\tilde{T}}^*$
25	9	0.004	1.049	-0.554	0.919	-0.703	1.003
30	10	0.003	1.035	-0.546	0.889	-0.674	0.949
35	11	0.002	1.027	-0.541	0.867	-0.653	0.906
40	11	0.002	1.021	-0.537	0.850	-0.637	0.871
45	11	0.001	1.017	-0.533	0.837	-0.624	0.842
50	12	0.001	1.014	-0.531	0.826	-0.614	0.818
60	13	0.001	1.011	-0.527	0.810	-0.598	0.780
70	13	0.000	1.008	-0.524	0.798	-0.587	0.751
80	14	0.000	1.007	-0.521	0.789	-0.578	0.728
90	14	0.000	1.006	-0.520	0.782	-0.571	0.710
100	15	0.000	1.005	-0.518	0.776	-0.566	0.695
250	20	0.000	1.001	-0.509	0.742	-0.533	0.603
$\infty$	—	0.000	1.000	-0.500	0.707	-0.500	0.500

Source : Table 2, page 14, Levin et Lin (2002)

Dans la première colonne figure l'approximation moyenne de la dimension temporelle des observations  $\tilde{T} = (T - \bar{p} - 1)$  où  $\bar{p} = (1/N) \sum_{i=1}^N p_i$ . Pour une taille  $\tilde{T}$  donnée, figure dans la deuxième colonne la valeur du paramètre de troncature  $q_i$  retenu par Levin et Lin dans la procédure d'estimation par noyau de la variance de long terme des résidus (étape 2). Rappelons que les auteurs n'adoptent pas la démarche d'Andrews (1991) appliquée individu par individu, et imposent le même paramètre de troncature pour tous les individus du panel  $q_i = q, \forall i = 1, \dots, N$ . On peut retrouver les valeurs des auteurs en retenant l'entier le plus proche de la valeur  $q = 3.21 \tilde{T}^{1/3}$ . Dans les colonnes suivantes figurent les facteurs d'ajustement de l'espérance et de la variance de la statistique de Student pour les modèles 1, 2 et 3<sup>6</sup>.

Tout d'abord, dans le cas du modèle 1 (sans constante ni tendance), comme nous l'avons précisé, la statistique de Student standard  $t_{\rho=0}$  (équation 22) converge asymptotiquement vers une loi normale centrée réduite : dès lors, il n'y a pas besoin d'apporter une quelconque correction à cette statistique de test pour obtenir une distribution asymptotique normale :

$$\text{Modèle 1 : } \mu_{1,\tilde{T}}^* \xrightarrow{\tilde{T} \rightarrow \infty} 0 \quad \text{et} \quad \sigma_{1,\tilde{T}}^* \xrightarrow{\tilde{T} \rightarrow \infty} 1$$

<sup>6</sup> Il convient de souligner que les résultats figurant dans ce tableau sont obtenus sous l'hypothèse de nullité des  $p_i$ , ce qui ne nous permet donc pas d'appréhender l'influence des retards sur les résultats. De même, le fait de retenir une valeur unique pour le paramètre de troncature  $q$  ne permet pas d'apprécier l'impact du choix de ce paramètre sur les résultats. Nous reviendrons sur ces problèmes par la suite en évoquant notamment la sensibilité du test de Levin et Lin au nombre de retards retenus.

Par conséquent, dans ce modèle, la statistique corrigée  $t_{\rho=0}^*$  converge asymptotiquement vers la statistique de Student non corrigée  $t_{\rho=0}$  ; cette dernière statistique convergeant elle-même vers une loi normale centrée réduite. Toutefois, pour des tailles d'échantillons  $\tilde{T}$  finies, Levin et Lin proposent de corriger la statistique de Student, bien que les termes de correction soient très faibles pour l'espérance et proches de l'unité pour la variance, comme on peut le voir dans les troisième et quatrième colonnes du tableau 1.

En revanche, dans les modèles 2 (avec effets individuels) et 3 (avec tendance) la statistique de Student  $t_{\rho=0}$  diverge négativement. Les termes de correction de l'espérance  $\mu_{2,\tilde{T}}^*$  et  $\mu_{3,\tilde{T}}^*$  sont donc asymptotiquement non nuls et les termes de correction de la variance  $\sigma_{2,\tilde{T}}^*$  et  $\sigma_{3,\tilde{T}}^*$  sont asymptotiquement différents de l'unité contrairement au cas précédent.

Sous l'hypothèse que l'ordre maximal des retards admissibles  $p_{max}$  (étape 1) croît à une vitesse  $T^p$  avec  $0 < p \leq 1/4$  et sous l'hypothèse que le paramètre de troncature  $q_i$  (étape 2) croît à une vitesse<sup>7</sup>  $T^q$  avec  $0 < q < 1$ , Levin et Lin établissent la distribution asymptotique de la statistique corrigée  $t_{\rho=0}^*$ . Ils montrent que, sous l'hypothèse nulle de racine unitaire, cette statistique converge séquentiellement et le long d'une diagonale vers une loi normale centrée réduite quel que soit le modèle retenu :

$$t_{\rho=0}^* \xrightarrow[T, N \rightarrow \infty]{l} N(0, 1) \quad \text{avec } \sqrt{N}/T \rightarrow 0 \quad (25)$$

Ainsi, si la réalisation de la statistique corrigée de Levin et Lin  $t_{\rho=0}^*$  est inférieure au seuil de la loi normale centrée réduite (pour un test non symétrique à 5% de risque de première espèce ce seuil est égal à  $-1.64$ ), on rejette l'hypothèse nulle de racine unitaire pour l'ensemble des individus du panel.

Quel gain de puissance peut-on espérer d'un test de racine unitaire en panel comparative-ment aux tests effectués sur séries temporelles ? Comme le montrent Levin et Lin, pour des échantillons de petite taille  $T$  on peut atteindre des puissances sensiblement plus importantes que celles obtenues dans le cas de tests menés sur séries temporelles individuelles. Tout d'abord, la taille du test en panel est correcte, même si elle est légèrement sous-estimée dans des panels de taille  $N$  modérée. En l'absence d'effets fixes individuels (modèle 1), on sait que pour de petites tailles d'échantillon ( $T \leq 50$ ), la puissance des tests de Dickey-Fuller est relativement faible (27% pour un risque de première espèce de 5% pour  $T = 25$ ). Or, pour une même taille  $T$ , l'application du test de Levin et Lin conduit à des niveaux de puissance beaucoup plus importants et ce même si le panel a une faible dimension individuelle  $N$ . Ainsi, avec seulement 10 individus pour  $T = 25$ , la puissance du test de Levin et Lin est de l'ordre de 99% dans l'expérience simulée par les auteurs. Ces quelques expériences<sup>8</sup> montrent à quel point, pour une taille empirique à peu près similaire, les gains de puissance peuvent être importants dans le passage à des tests de racine unitaire en panel lorsque la dimension temporelle est modérée ( $T < 250$ ). Toutefois, lorsque la dimension  $T$  est suffisamment grande, les tests usuels de racine unitaire sont suffisamment puissants et peuvent être appliqués séparément à chaque individu du panel. Il n'y a alors aucun intérêt (en termes de puissance et de taille) à passer au panel pour étudier la non stationnarité. Des résultats similaires peuvent être obtenus à partir des autres tests de racine unitaire en panel que nous présenterons par la suite. Ainsi, un des arguments principaux qui motivent le passage au panel réside dans les gains de puissance sur de petits échantillons. Enfin, en présence de termes déterministes (effets fixes et/ou tendance temporelle), la puissance du test usuel de Dickey-Fuller est très faible, même pour des séries temporelles relativement longues ( $T < 100$ ) alors qu'elle peut être très élevée pour le test de Levin et Lin dans des panels de taille modérée ( $N = 10$  et  $T = 50$  ou  $N = 25$  et  $T = 25$ ).

<sup>7</sup>Ce qui est le cas avec la fenêtre d'Andrews (1991).

<sup>8</sup>Voir Levin, Lin et Chu (2002) ou Maddala et Wu (1999) pour une discussion plus approfondie sur les gains en puissance des tests de racine unitaire en panel et la comparaison de la puissance des différents tests de panel.

### 3 Les tests d'Im, Pesaran et Shin

Comme nous l'avons évoqué, une des principales limites du test de Levin et Lin réside dans le caractère homogène de la racine autorégressive sous l'hypothèse alternative. Il est peu probable en effet qu'en cas de rejet de l'hypothèse de racine unitaire on puisse accepter l'hypothèse d'une racine autorégressive  $\rho_i$  commune à tous les individus si l'on applique des tests usuels de spécification<sup>9</sup>. Les tests proposés par Im, Pesaran et Shin dans une série de contributions (1997, 2002 et 2003) permettent de répondre à cette critique. En effet, ces auteurs furent les premiers à développer un test autorisant sous l'hypothèse alternative non seulement une hétérogénéité de la racine autorégressive ( $\rho_i \neq \rho_j$ ), mais aussi une hétérogénéité quant à la présence d'une racine unitaire dans le panel.

Im, Pesaran et Shin (IPS par la suite) considèrent un modèle avec effets individuels et sans tendance déterministe (équivalent du modèle 2 chez Levin et Lin). En l'absence d'autocorrélation des résidus, ce modèle s'écrit :

$$\text{Modèle IPS : } \Delta y_{i,t} = \alpha_i + \rho_i y_{i,t-1} + \varepsilon_{i,t} \quad (26)$$

où l'effet individuel  $\alpha_i$  est défini par  $\alpha_i = -\rho_i \gamma_i$  avec  $\gamma_i \in \mathbb{R}$  et où  $\varepsilon_{i,t} \sim N.i.d. (0, \sigma_{\varepsilon,i}^2)$ . Le test d'IPS, tout comme le test de Levin et Lin est un test joint de l'hypothèse nulle de racine unitaire ( $\rho_i = 0$ ) et de l'absence d'effets individuels puisque sous l'hypothèse nulle  $\alpha_i = 0$ .

$$\begin{aligned} \text{Test IPS : } \quad H_0 : \rho_i = 0, \forall i = 1, \dots, N \\ H_1 : \rho_i < 0, \forall i = 1, 2, \dots, N_1 \\ \rho_i = 0, \forall i = N_1 + 1, N_1 + 2, \dots, N \end{aligned}$$

Sous l'hypothèse alternative peuvent coexister deux types d'individus : des individus indicés  $i = 1, \dots, N_1$  pour lesquels la variable  $y_{i,t}$  est stationnaire et des individus indicés  $i = N_1 + 1, \dots, N$  pour lesquels la dynamique de la variable  $y_{i,t}$  admet une racine unitaire. La taille  $N_1$  de l'ensemble des individus stationnaires est *a priori* inconnue mais vérifie  $0 < N_1 \leq N$ , puisque si  $N_1 = 0$  on retrouve alors l'hypothèse nulle. On admet en outre que le ratio  $N_1/N$  vérifie  $\lim_{N \rightarrow \infty} N_1/N = \delta$  avec  $0 < \delta \leq 1$ .

Ainsi, le premier avantage de l'approche d'IPS par rapport à Levin et Lin tient à la prise en compte de l'hétérogénéité de la racine autorégressive sous l'alternative. Mais ce n'est pas le seul avantage. Comme nous allons le voir, les auteurs proposent une statistique de test très simple fondée sur la moyenne des statistiques de Dickey-Fuller ou de Dickey-Fuller Augmentées individuelles.

- Sous l'hypothèse d'absence d'autocorrélation des résidus, IPS dérivent la loi asymptotique de leur statistique moyenne (lorsque  $T$  et  $N$  convergent vers l'infini) mais aussi la loi semi-asymptotique lorsque  $T$  est fixe et  $N$  converge vers l'infini. Dans ce cas, il est en effet possible de dériver la loi exacte de la statistique de test de racine unitaire pour une taille  $T$  quelconque, contrairement à la statistique de Levin et Lin. IPS proposent même dans ce cas des approximations des seuils de rejet à distance finie pour  $T$  et  $N$  fixes.
- En revanche, sous l'hypothèse d'autocorrélation des résidus, on ne peut plus caractériser la loi exacte de la statistique moyenne pour une taille  $T$  donnée : IPS dérivent dans ce cas les lois asymptotiques pour  $T$  et  $N$  tendant vers l'infini (soit de façon séquentielle, soit le long d'une diagonale) et proposent deux statistiques moyennes standardisées. On retrouve encore une fois une distribution normale.

<sup>9</sup> Voir Hurlin (2001) pour les tests d'homogénéité dans le cas d'un modèle linéaire et Phillips et Sul (2003a) pour un test d'homogénéité de la racine autorégressive.

### 3.1 Une moyenne de statistiques individuelles

Considérons le cas le plus simple où les résidus  $\varepsilon_{i,t}$  du modèle (26) sont *N.i.d.*  $(0, \sigma_{\varepsilon,i}^2)$  et où la taille d'échantillon disponible pour l'estimation de chaque équation individuelle est identique<sup>10</sup> pour tous les pays :  $T_i = T_j = T$ . Tout comme Levin et Lin, IPS supposent que les résidus sont indépendants dans la dimension individuelle. Dans ce cas, ils proposent une statistique de test notée  $t\_bar_{NT}$  définie par la moyenne des  $N$  statistiques individuelles de Dickey-Fuller :

$$t\_bar_{NT} = \frac{1}{N} \sum_{i=1}^N t_{iT} \quad (27)$$

où  $t_{iT}$  correspond à la statistique de Student associée<sup>11</sup> à l'hypothèse nulle  $\rho_i = 0$  dans le modèle (26).

Quelle est la loi de cette statistique moyenne  $t\_bar_{NT}$  sous  $H_0$ ? Commençons par évoquer le cas de la distribution asymptotique lorsque les deux dimensions  $N$  et  $T$  tendent vers l'infini. Pour ce faire, nous considérerons le cas le plus simple de la convergence séquentielle lorsque  $T$  puis  $N$  tendent vers l'infini. On sait que sous l'hypothèse nulle de non stationnarité  $\rho_i = 0$ , la statistique du test joint de Dickey-Fuller de l'individu  $i$  converge lorsque  $T$  tend vers l'infini vers une distribution non standard (Dickey et Fuller, 1979) définie par :

$$t_{iT} \xrightarrow[T \rightarrow \infty]{d} \frac{\frac{1}{2} [W_i(r)^2 - 1] - W_i(1) \int_0^1 W_i(r) dr}{\left\{ \int_0^1 W_i(r)^2 dr - \left[ \int_0^1 W_i(r) dr \right]^2 \right\}^{\frac{1}{2}}} \quad \forall i = 1, \dots, N \quad (28)$$

où les variables  $W_i(r)$ , pour  $i = 1, \dots, N$ , désignent des mouvements Browniens standard indépendants. On sait en outre que cette distribution asymptotique admet des moments d'ordre deux finis (Nabeya, 1999), avec  $E(t_{iT}) = -1.533$  et  $Var(t_{iT}) = 0.706$  quand  $T \rightarrow \infty$ . De plus, sous l'hypothèse d'indépendance inter-individuelle des résidus, on peut montrer que les statistiques individuelles  $t_{iT}$  sont indépendantes. Ainsi, en résumé, lorsque la dimension  $T$  tend vers l'infini, les statistiques individuelles  $t_{iT}$  sont (i) identiquement et (ii) indépendamment distribuées selon une distribution admettant (iii) des moments d'ordre deux finis. Or, la statistique de test d'IPS n'est rien d'autre que la moyenne sur  $N$  de ces statistiques individuelles. Il suffit donc, dans un second temps de faire tendre  $N$  vers l'infini et d'appliquer le théorème central limite de Lindberg-Levy afin de montrer que la moyenne de ces  $N$  statistiques individuelles converge vers une loi normale. En définitive, lorsque  $T$  puis  $N$  tendent vers l'infini,  $t\_bar_{NT}$  converge vers une loi normale. On définit alors une variable moyenne standardisée  $Z_{tbar}$  convergeant une loi normale centrée réduite.

$$Z_{tbar} = \frac{\sqrt{N}(t\_bar_{NT} + 1.533)}{\sqrt{0.706}} \xrightarrow[(T, N \rightarrow \infty)_{seq}]{l} N(0, 1) \quad (29)$$

En outre, par ce même raisonnement, IPS peuvent établir la loi exacte (sous l'hypothèse d'absence d'autocorrélation des résidus) de leur statistique pour une taille  $T$  fixe lorsque  $N$  tend vers l'infini. Ce résultat est particulièrement intéressant pour les applications dans le cas des panels macro-économiques où la taille est généralement faible, de l'ordre de 30 points.

<sup>10</sup>Dans le cas contraire, on conserve les mêmes lois asymptotiques et semi-asymptotiques, mais la démonstration de ce résultat requiert l'application du théorème limite de Lyapunov et non plus celui de Lindberg-Levy. En effet, dans ce cas, les statistiques individuelles ne sont plus identiquement distribuées à  $T_i$  fini.

<sup>11</sup>IPS considèrent deux définitions notées  $\tilde{t}_{iT}$  et  $t_{iT}$  de la statistique de Student suivant la définition de l'estimateur de la variance des résidus. Nous considérerons ici la statistique  $t_{iT}$  programmée sous les logiciels usuels.

Contrairement au cas où  $T$  tend vers l'infini, on ne connaît pas la distribution exacte de la statistique individuelle  $t_{iT}$  pour une taille  $T$  donnée. Pour autant, il est possible de montrer que la distribution de la moyenne  $t\_bar_{NT}$  converge vers une loi normale lorsque  $N$  tend vers l'infini en utilisant toujours le théorème central limite de Lindberg-Levy. Pour ce faire, il convient de montrer les points (i), (ii) et (iii) sans connaître la distribution exacte des  $t_{iT}$  pour un  $T$  fixe. L'astuce d'IPS consiste alors à exprimer les statistiques de Dickey-Fuller sous la forme d'un produit de ratios de formes quadratiques définies dans un vecteur normal, puis à appliquer le théorème de Magnus (1986). Ce théorème permet, sous des conditions très générales, de dériver les moments exacts d'un ratio de formes quadratiques définies dans un vecteur normal, ainsi que les conditions d'existence de ces moments. Ainsi, en utilisant une inégalité de Cauchy-Schwarz, on peut dériver très simplement une condition<sup>12</sup> qui garantit que les moments d'ordre deux des statistiques individuelles sont finis : cette condition se ramène tout simplement à l'inégalité  $T \geq 6$ . Dès lors sous cette condition, l'application du théorème central limite<sup>13</sup> nous permet d'établir la loi exacte de la statistique moyenne standardisée, notée  $Z_{tbar}$  :

$$Z_{tbar} = \frac{\sqrt{N} [t\_bar_{NT} - E(t_{iT})]}{\sqrt{Var(t_{iT})}} \xrightarrow{N \rightarrow \infty} N(0, 1) \quad \text{si } T \geq 6 \quad (30)$$

où  $E(t_{iT})$  et  $Var(t_{iT})$  désignent respectivement l'espérance et la variance de la statistique de Dickey-Fuller. Lorsque  $T$  tend vers l'infini, ces moments tendent vers les moments de la distribution asymptotique de Dickey-Fuller et l'on retrouve la statistique standardisée (29). Toutefois, pour des tailles d'échantillon  $T$  plus petites, l'utilisation des moments de la distribution asymptotique de Dickey-Fuller peut conduire à une perte de puissance des tests. C'est pourquoi IPS proposent des valeurs simulées pour ces deux moments pour différentes tailles  $T$  (voir tableau 3). Prenons par exemple le cas d'un panel de taille  $T = 10$  avec  $N = 5$ . On construit la statistique moyenne standardisée  $Z_{tbar}$  à partir de la statistique moyenne  $t\_bar_{NT}$  en utilisant  $E(t_{iT}) = -1.504$  et  $Var(t_{iT}) = 1.069$ . Pour un risque de première espèce de  $\alpha\%$ , on compare alors la réalisation de  $Z_{tbar}$  au seuil  $z_\alpha$  de la loi normale centrée réduite obtenue sous l'hypothèse  $N \rightarrow \infty$ . Pour un risque de 5%, si la réalisation de  $Z_{tbar}$  est inférieure au seuil  $-1.64$  on rejette l'hypothèse nulle de racine unitaire pour l'ensemble des individus du panel.

Enfin, à partir de la loi exacte de la statistique standardisée  $Z_{tbar}$  obtenue sous l'hypothèse  $N \rightarrow \infty$ , IPS proposent des approximations des seuils pour la statistique moyenne  $t\_bar_{NT}$  dans le cas d'échantillons de petite dimension  $N$ . Ces valeurs critiques approximées notées  $c_T(\alpha)$ , pour un niveau de risque de première espèce de  $\alpha\%$ , peuvent être obtenues par la formule suivante :

$$c_T(\alpha) = z_\alpha \sqrt{N^{-1} Var(t_T)} + E(t_T) \quad (31)$$

où  $z_\alpha$  désigne la valeur critique associée à une distribution normale centrée réduite pour un niveau de risque de  $\alpha\%$ . Reprenons l'exemple d'un panel de taille  $T = 10$  avec  $N = 5$ . On peut tester la racine unitaire à partir de la statistique  $Z_{tbar}$ . Toutefois, l'utilisation de la loi semi-asymptotique obtenue sous l'hypothèse  $N \rightarrow \infty$  pour un panel de cette dimension peut conduire à une faible puissance ou à des distorsions de taille. Une autre solution consiste à comparer directement la réalisation de la statistique moyenne  $t\_bar_{NT}$  au seuil approximé

<sup>12</sup>Il convient de préciser que cette condition théorique ne peut être dérivée qu'à partir des statistiques individuelles non corrigées  $\tilde{t}_{iT}$ . Mais la même condition est obtenue par simulation dans le cas des statistiques corrigées  $t_{iT}$ .

<sup>13</sup>Dans le cas d'un panel non cylindré où les tailles  $T_i$  diffèrent selon les individus, on obtient le même résultat asymptotique dès lors que  $T_i \geq 10, \forall i = 1, \dots, N$  avec :

$$Z_{tbar} = \frac{\sqrt{N} \left[ t\_bar_{NT} - N^{-1} \sum_{i=1}^N E(t_{iT}) \right]}{\sqrt{N^{-1} \sum_{i=1}^N Var(t_{iT})}} \xrightarrow{N \rightarrow \infty} N(0, 1)$$

puisque les moments individuels  $E(t_{iT})$  et  $Var(t_{iT})$  varient selon la taille  $T_i$  (tableau 3).

TAB. 3 – Moments des Statistiques Individuelles  $t_{iT}$

$T$	$E(t_{iT})$	$Var(t_{iT})$
6	-1.520	1.745
10	-1.504	1.069
15	-1.514	0.923
20	-1.522	0.851
25	-1.520	0.809
30	-1.526	0.789
40	-1.523	0.770
50	-1.527	0.760
100	-1.532	0.735
500	-1.531	0.715
$\infty$	-1.533	0.706

Source : Table 1, page 60, IPS (2003)

$c_T(\alpha)$  pour une taille  $N = 5$ . Dans notre exemple, pour un risque de première espèce de 5%, en utilisant les moments simulés du tableau 3, on montre que le seuil approximé est égal à  $c_T(\alpha) = -1.64 * \sqrt{5^{-1}} * 1.069 - 1.504 = -2.2623$ . Si la réalisation de la statistique moyenne  $t\_bar_{NT}$  est inférieure à ce seuil, on rejette l'hypothèse nulle. IPS montrent que les seuils approximés sont très proches des seuils de la vraie loi à  $N$  fini de la statistique moyenne  $t\_bar_{NT}$  obtenus par simulation.

En résumé, en l'absence d'autocorrélation des résidus, dès lors que  $T \geq 6$  pour un panel cylindré, la statistique standardisée  $Z_{tbar}$  suit sous l'hypothèse nulle de racine unitaire une loi normale lorsque  $N$  tend vers l'infini. C'est encore une fois l'une des principales différences avec les tests sur séries temporelles, où les distributions asymptotiques des statistiques sont non standard.

### 3.2 Mise en oeuvre du test dans le cas général

A présent plaçons nous dans le cas général où il existe une éventuelle autocorrélation des résidus. On considère un modèle de type Dickey-Fuller Augmenté (ADF) pour chaque individu  $i = 1, \dots, N$  du panel :

$$\text{Modèle IPS : } \Delta y_{i,t} = \alpha_i + \rho_i y_{i,t-1} + \sum_{j=1}^{p_i} \beta_{i,j} \Delta y_{i,t-j} + \varepsilon_{i,t} \quad (32)$$

où l'effet individuel  $\alpha_i$  est défini par  $\alpha_i = -\rho_i \gamma_i$  avec  $\gamma_i \in \mathbb{R}$  et où  $\varepsilon_{i,t}$  est *N.i.d.*  $(0, \sigma_i^2)$ . Comme pour tous les tests de première génération, les résidus sont indépendamment distribués dans la dimension individuelle. On remarque qu'IPS autorisent la présence d'une autocorrélation des résidus d'ordre différent pour chaque individu du panel. Ceci implique que le nombre de termes ADF diffère *a priori* suivant les individus  $p_i \neq p_j$  comme dans le test de Levin et Lin. Les hypothèses nulle et alternative du test sont les mêmes que dans le cas précédent. Pour mener à bien ce test, IPS proposent à nouveau d'utiliser la moyenne des statistiques individuelles ADF :

$$t\_bar_{NT} = \frac{1}{N} \sum_{i=1}^N t_{iT}(p_i, \beta_i) \quad (33)$$

où  $t_{iT}(p_i, \beta_i)$  correspond à la statistique individuelle de Student associée à l'hypothèse nulle  $H_{0,i} : \rho_i = 0$  dans le modèle (32) pour un nombre de retards  $p_i$  et un vecteur de paramètres ADF

$\beta_i = (\beta_{i,1}, \dots, \beta_{i,p_i})'$ . Il s'agit de la statistique ADF standard obtenue à partir d'un modèle avec constante et qui est programmée dans la plupart des logiciels usuels pour un retard  $p_i$  donné. Pour chaque individu, le choix du retard optimal  $p_i$  permettant de purger l'autocorrélation des résidus peut être choisi de la même façon que dans le cas des tests de Levin et Lin.

A partir des  $N$  statistiques ADF individuelles  $t_{iT}(p_i, \beta_i)$ , on construit la statistique standardisée  $Z_{tbar}(p, \beta)$  centrée sur l'espérance de la distribution asymptotique  $\eta$  de la statistique individuelle ADF et réduite par la variance de cette même distribution :

$$Z_{tbar}(p, \beta) = \frac{\sqrt{N} [t_{-bar_{NT}} - E(\eta)]}{\sqrt{Var(\eta)}} \quad (34)$$

où les moments  $E(\eta)$  et  $Var(\eta)$  correspondent à l'espérance et à la variance de la distribution asymptotique (quand  $T \rightarrow \infty$ ) d'une statistique ADF sous l'hypothèse nulle de racine unitaire ( $\rho_i = 0$ ) dans un modèle avec constante. Comme nous l'avons précisé, ces moments sont respectivement égaux à  $E(t_{iT}) = -1.533$  et  $Var(t_{iT}) = 0.706$ . Ainsi, on peut très facilement montrer que la statistique  $Z_{tbar}(p, \beta)$  converge séquentiellement vers une loi normale centrée réduite lorsque  $T$  puis  $N$  tendent vers l'infini.

Il est évident toutefois que cette approche fondée sur la distribution asymptotique peut poser problème dans des panels de petite taille  $T$ . C'est pourquoi IPS proposent une seconde statistique standardisée, notée  $W_{tbar}(p, \beta)$ , qui asymptotiquement possède la même distribution que  $Z_{tbar}(p, \beta)$ , mais qui en outre possède l'avantage d'être beaucoup plus puissante à distance finie. C'est généralement la statistique de test d'IPS que l'on retient<sup>14</sup>, puisque c'est la plus générale (elle tient compte de l'autocorrélation des résidus) et la plus puissante. Cette statistique standardisée est définie de la même façon que  $Z_{tbar}(p, \beta)$  à la différence près que l'on centre et l'on réduit à partir des moments de la statistique ADF obtenue sous l'hypothèse nulle de racine unitaire (comme pour  $Z_{tbar}(p, \beta)$ ) et sous l'hypothèse que les paramètres  $\beta_i$  des termes ADF sont nuls pour tous les individus. Ces moments sont respectivement notés  $E[t_{iT}(p_i, 0) | \rho_i = 0]$  et  $Var[t_{iT}(p_i, 0) | \rho_i = 0]$ . Il convient de noter que ces moments tiennent compte de l'information contenue dans le nombre de retards  $p_i$ . Asymptotiquement, cette statistique d'IPS standardisée converge vers la même distribution que la statistique  $Z_{tbar}(p, \beta)$  :

$$W_{tbar}(p, \beta) = \frac{\sqrt{N} \left[ t_{-bar_{NT}} - N^{-1} \sum_{i=1}^N E[t_{iT}(p_i, 0) | \rho_i = 0] \right]}{\sqrt{N^{-1} \sum_{i=1}^N Var[t_{iT}(p_i, 0) | \rho_i = 0]}} \xrightarrow[T, N \rightarrow \infty]{l} N(0, 1) \quad (35)$$

Les moments  $E[t_{iT}(p_i, 0) | \rho_i = 0]$  et  $Var[t_{iT}(p_i, 0) | \rho_i = 0]$  ont été tabulés pour différents ordres de retards  $p_i$  et différentes tailles  $T$  par les auteurs. Certaines valeurs pour le modèle sans tendance sont reportées dans le tableau 4. Tout comme pour les résultats de la section précédente, ces résultats peuvent être étendus au cas de panels non cylindrés présentant des tailles  $T_i$  différentes selon les individus.

Les simulations menées par IPS montrent qu'en l'absence de corrélation, la statistique  $Z_{tbar}$  (équation 30) conduit à de très bons résultats en termes de taille et de puissance et ce même en petit échantillon ( $T = 10$ ). Lorsque le nombre de retards est correctement choisi ou sur-estimé, la taille du test  $Z_{tbar}$  est très satisfaisante, même pour  $T = 10$ . De plus, la puissance est largement supérieure à celle d'un test de Dickey-Fuller mené sur séries temporelles. Un tel résultat illustre l'apport de la dimension individuelle comparativement à la prise en compte

<sup>14</sup>C'est le cas dans la présentation de Banerjee (1999), mais pas dans celle de Baltagi et Kao (2000) qui privilégient la statistique  $Z_{tbar}(p, \beta)$ . Toutefois, IPS (2003) n'évaluent la puissance de leur test qu'à partir de la statistique  $W_{tbar}(p, \beta)$ .

TAB. 4 – Moments des Statistiques Individuelles  $t_{iT}(p_i, 0)$

T	$p_i = 0$		$p_i = 1$		$p_i = 2$		$p_i = 3$		$p_i = 4$	
	Mean	Var	Mean	Var	Mean	Var	Mean	Var	Mean	Var
10	-1.504	1.069	-1.488	1.255	-1.319	1.421	-1.306	1.759	-1.171	2.080
15	-1.514	0.923	-1.503	1.011	-1.387	1.078	-1.366	1.181	-1.260	1.279
20	-1.522	0.851	-1.516	0.915	-1.428	0.969	-1.413	1.037	-1.329	1.097
25	-1.520	0.809	-1.514	0.861	-1.443	0.905	-1.433	0.952	-1.363	1.005
30	-1.526	0.789	-1.519	0.831	-1.460	0.865	-1.453	0.907	-1.394	0.946
40	-1.523	0.770	-1.520	0.803	-1.476	0.830	-1.471	0.858	-1.428	0.886
50	-1.527	0.760	-1.524	0.781	-1.493	0.798	-1.489	0.819	-1.454	0.842
60	-1.519	0.749	-1.519	0.770	-1.490	0.789	-1.486	0.802	-1.458	0.819
70	-1.524	0.736	-1.522	0.753	-1.498	0.766	-1.495	0.782	-1.470	0.801
100	-1.532	0.735	-1.530	0.745	-1.514	0.754	-1.512	0.761	-1.495	0.771

Source : Im, Pesaran et Shin (2003), Table 3, page 66

de la seule dimension temporelle. Des résultats allant dans le même sens sont obtenus par les auteurs dans le cas d'une autocorrélation des résidus sur la base de la statistique  $W_{tbar}(p, \beta)$ , mais il convient alors que  $T$  et  $N$  soient suffisamment importants.

Par ailleurs, ces simulations montrent que le choix du nombre de retards  $p_i$  dans les régressions ADF individuelles est crucial. On connaît traditionnellement ce problème en séries temporelles (Ng et Perron, 1995 ou Lopez, 1997). Si l'on surestime le nombre de retards, la puissance du test ADF est détériorée, mais le problème est plus fondamental si le nombre de retards est sous évalué : dans ce cas la paramétrisation du modèle ne permet pas de blanchir totalement les résidus, en conséquence de quoi les distributions asymptotiques de Dickey-Fuller ne sont plus valides. Il est donc évident qu'en panel, la puissance de la statistique  $Z_{tbar}(p, \beta)$  d'IPS fondée sur une moyenne de statistiques ADF et standardisée à partir des moments  $E(\eta)$  et  $Var(\eta)$  de la distribution asymptotique de Dickey-Fuller est extrêmement sensible au choix des retards  $p_i$ . Si lorsque  $T$  tend vers l'infini, une ou plusieurs statistiques ADF individuelles ne convergent plus vers  $\eta$  parce que l'on a sous estimé  $p_i$ , la statistique standardisée  $Z_{tbar}(p, \beta)$  n'a plus de sens. Il en va de même pour  $W_{tbar}(p, \beta)$ . A titre d'exemple, les auteurs montrent que lorsque l'on retient de manière erronée un nombre de retards égal à zéro, la taille du test tend vers zéro. Notons que ce problème se pose avec autant d'acuité dans le cas du test de Levin et Lin.

## 4 Le test de Maddala et Wu

Le troisième test de cette première génération est un test non paramétrique de Fisher (1932) initialement appliqué à l'étude de la PPA par Choi<sup>15</sup> (2001) et présenté de façon générale par Maddala et Wu (1999). Le principe est simple et repose sur une combinaison des niveaux de significativité (c'est-à-dire des p-values) de  $N$  tests individuels de racine unitaire indépendants. Soit  $p_i = F_{T_i}(G_i)$  la p-value associée à une statistique de test  $G_i$  de l'hypothèse nulle de racine unitaire pour un individu  $i$  donné où  $F_{T_i}(\cdot)$  désigne la fonction de répartition associée à la statistique individuelle  $G_i$  pour un échantillon de taille  $T_i$ . La statistique de test  $G_i$  peut être choisie comme la t-statistique d'un test ADF, où la statistique de n'importe quel autre test de l'hypothèse nulle de racine unitaire (Phillips et Perron, 1988, Elliott, Rothenberg et Stock, 1996, etc.). Il existe alors de très nombreuses façons de combiner les p-values afin de construire

<sup>15</sup>La version document de travail date de 1999 : Choi I. (1999), "Unit Root Tests for Panel Data", *Manuscript*, Kookmin University, Corée.

un test de racine unitaire en panel (en utilisant le minimum, la somme, *etc.*). Maddala et Wu (1999) retiennent la statistique de test définie par la quantité :

$$P_{MW} = -2 \sum_{i=1}^N \ln(p_i) \quad (36)$$

Si les statistiques individuelles de test sont continues, les p-values sont distribuées selon des lois uniformes sur  $[0, 1]$  et  $\ln(p_i)$  est distribuée selon un  $\chi^2(1)$ ,  $\forall i = 1, \dots, N$ . Maddala et Wu (1999) considèrent le cas où les statistiques individuelles sont indépendantes. Tout comme IPS ou Levin et Lin, ce test est donc un test de première génération reposant sur l'exclusion d'une quelconque relation entre les statistiques individuelles et plus généralement sur l'absence de corrélation inter-individuelle. Sous cette hypothèse, la statistique  $P_{MW}$  suit, sous l'hypothèse nulle de racine unitaire, un  $\chi^2(2N)$  quelle que soit la taille  $N$  de l'échantillon. Pour un risque de première espèce donné, si la réalisation de  $P_{MW}$  est supérieure au seuil d'un  $\chi^2(2N)$  on rejette l'hypothèse nulle de racine unitaire pour les individus du panel.

Pour des valeurs élevées de  $N$ , Choi (2001) suggère d'utiliser la statistique standardisée suivante :

$$Z_{MW} = \frac{\sqrt{N} \{N^{-1}P_{MW} - E[-2 \ln(p_i)]\}}{\sqrt{Var[-2 \ln(p_i)]}} = \frac{1}{2\sqrt{N}} \sum_{i=1}^N [-2 \ln(p_i) - 2] \quad (37)$$

En effet, sous l'hypothèse de continuité de la statistique  $G_i$ , on sait que  $E[-2 \ln(p_i)] = 2$  et  $Var[-2 \ln(p_i)] = 4$ . La statistique de Choi correspond ainsi tout simplement à une statistique moyenne de type  $N^{-1}P_{MW}$  centrée et réduite. Si l'on suppose que les p-values sont *i.i.d.*, l'utilisation du théorème de Lindberg-Levy nous permet de conclure que la statistique réduite  $Z_{MW}$  suit sous  $H_0$  une loi  $N(0, 1)$  lorsque  $N$  tend vers l'infini.

Tout comme le test proposé par IPS, le test de Maddala et Wu (1999) ne retient pas l'hypothèse alternative restrictive du test de Levin et Lin selon laquelle le coefficient autorégressif  $\rho_i$  est le même pour tous les individus. Ce test est donc directement comparable au test IPS et lui est très similaire. Les deux tests reposent sur une combinaison de statistiques individuelles : des statistiques ADF chez IPS et des seuils de significativité chez Maddala et Wu. Les deux tests peuvent être appliqués sur des panels non cylindrés<sup>16</sup> et on peut éventuellement établir soit la loi exacte (chez Maddala et Wu), soit une approximation des seuils (chez IPS) de la statistique de test pour une taille  $N$  fixe.

Quels sont donc les avantages respectifs des deux tests ? Les deux tests peuvent tout d'abord être comparés en termes de taille empirique et de puissance (contrairement au test de Levin et Lin qui n'admet pas la même hypothèse alternative). Maddala et Wu considèrent une série d'expériences dans lesquelles leur statistique de Fisher est construite à partir des p-values des t-statistiques des tests ADF individuels. Il ressort que ce test de Fisher domine légèrement le test d'IPS, puisque pour des tailles empiriques similaires la puissance du test est légèrement augmentée. Lorsque  $T$  est suffisamment grand ( $T = 50$  ou  $100$ ), les puissances sont comparables, mais les distorsions de taille sont moins importantes avec le test de Maddala et Wu. Un autre avantage traditionnel d'une approche non paramétrique comme celle de Maddala et Wu réside dans la robustesse notamment à une erreur de spécification sur la distribution des résidus (supposés normalement distribués chez IPS). En revanche, le principal "inconvenient empirique" de l'approche de Maddala et Wu provient de la nécessité de simuler par Bootstrap

<sup>16</sup> Contrairement à ce qui est avancé dans la synthèse de Baltagi et Kao (2000) fondée sur la version de 1997 du papier d'IPS.

les distributions des statistiques individuelles afin de construire les p-values individuelles<sup>17</sup>. A l'inverse, on a vu que les statistiques standardisées des tests d'IPS peuvent être construites à partir des moments fournis par les auteurs, ce qui en facilite d'autant la mise en oeuvre.

## 5 Le test de stationnarité de Hadri (2000)

Se démarquant des trois tests de racine unitaire de première génération présentés précédemment qui reposaient sur l'hypothèse nulle de non stationnarité, le test de Hadri (2000) est basé sur l'hypothèse nulle de stationnarité. Ce test consiste en une extension du test de stationnarité proposé par Kwiatkowski *et al.* (1992) dans le cadre de l'économétrie des séries temporelles. Il s'agit d'un test du multiplicateur de Lagrange visant à tester l'hypothèse nulle de stationnarité des séries  $y_{i,t}$  (pour  $i = 1, \dots, N$ ) contre l'hypothèse alternative de racine unitaire. Hadri (2000) considère les deux modèles suivants :

$$y_{i,t} = r_{i,t} + \varepsilon_{i,t} \quad (38)$$

et

$$y_{i,t} = r_{i,t} + \beta_i t + \varepsilon_{i,t} \quad (39)$$

où  $r_{i,t}$  est une marche aléatoire :  $r_{i,t} = r_{i,t-1} + u_{i,t}$ ,  $u_{i,t}$  est *i.i.d.*  $(0, \sigma_u^2)$ ,  $u_{i,t}$  et  $\varepsilon_{i,t}$  étant indépendants. L'hypothèse nulle peut alors s'écrire  $\sigma_u^2 = 0$ . Par ailleurs, dans la mesure où les  $\varepsilon_{i,t}$  sont supposés *i.i.d.*, alors, sous l'hypothèse nulle,  $y_{i,t}$  est stationnaire en niveau pour le modèle (38) et stationnaire autour d'une tendance déterministe pour le modèle (39). Le modèle (38) peut encore s'écrire :

$$y_{i,t} = r_{i,0} + e_{i,t} \quad (40)$$

et le modèle (39) :

$$y_{i,t} = r_{i,0} + \beta_i t + e_{i,t} \quad (41)$$

avec  $e_{i,t} = \sum_{j=1}^t u_{i,j} + \varepsilon_{i,t}$ ,  $r_{i,0}$  étant des valeurs initiales jouant le rôle de constantes hétérogènes.

Il convient de remarquer que si  $\sigma_u^2 = 0$ , alors  $e_{i,t} \equiv \varepsilon_{i,t}$  est stationnaire ( $r_{i,t}$  est une constante). Si  $\sigma_u^2 \neq 0$ ,  $e_{i,t}$  est non stationnaire ( $r_{i,t}$  est une marche aléatoire). Plus spécifiquement, Hadri (2000) teste l'hypothèse nulle  $\lambda = 0$  contre l'hypothèse alternative  $\lambda > 0$  où  $\lambda = \sigma_u^2 / \sigma_\varepsilon^2$ . En notant  $\hat{e}_{i,t}$  les résidus estimés de (40) ou (41), la statistique *LM* est donnée par :

$$LM = \frac{1}{\hat{\sigma}_\varepsilon^2} \frac{1}{NT^2} \left( \sum_{i=1}^N \sum_{t=1}^T S_{i,t}^2 \right) \quad (42)$$

où  $S_{i,t}$  désigne la somme partielle des résidus :  $S_{i,t} = \sum_{j=1}^t \hat{e}_{i,j}$  et  $\hat{\sigma}_\varepsilon^2$  est un estimateur convergent de  $\sigma_\varepsilon^2$ . Sous l'hypothèse nulle de stationnarité en niveau (modèle (38)), la statistique de test :

$$Z_\mu = \frac{\sqrt{N} \left\{ LM - E \left[ \int_0^1 V(r)^2 dr \right] \right\}}{\sqrt{V \left[ \int_0^1 V(r)^2 dr \right]}} \quad (43)$$

suit une loi normale centrée réduite, où  $V(r)$  est un pont brownien standard, pour  $T \rightarrow \infty$  suivi de  $N \rightarrow \infty$ . Les cumulants de la fonction caractéristique de  $\int_0^1 V^2$  donnent respectivement

<sup>17</sup>Nous faisons figurer "inconvenient empirique" entre guillemets dans la mesure où il va de soi que cet argument purement empirique ne doit pas constituer un frein à la mise en oeuvre du test de Maddala et Wu et ne constitue en aucun cas un critère de choix entre les tests.

la moyenne et la variance de  $\int_0^1 V(r)^2$  intervenant dans (43) : 1/6 (cumulant d'ordre 1) pour l'espérance et 1/45 (cumulant d'ordre 2) pour la variance (voir Hadri, 2000 pour les détails).

Sous l'hypothèse nulle de stationnarité autour d'une tendance déterministe (modèle (39)), la statistique de test :

$$Z_\tau = \frac{\sqrt{N} \left( LM - E \left[ \int_0^1 V_2(r)^2 dr \right] \right)}{\sqrt{V \left[ \int_0^1 V_2(r)^2 dr \right]}} \quad (44)$$

suit une loi normale centrée réduite, où  $V_2(r) = W(r) + (2r - 3r^2)W(1) + 6r(r - 1) \int_0^1 W(s)ds$  (voir Kwiatkowski *et al.* (1992)). La moyenne et la variance de  $\int_0^1 V_2^2$  sont données par les deux premiers cumulants soit, respectivement, 1/15 et 11/6300.

Pour finir, notons que Hadri (2000) a proposé une extension de son test consistant à relâcher l'hypothèse selon laquelle les erreurs  $\varepsilon_{it}$  sont *iid* afin de tenir compte de la présence de corrélation sérielle.

Afin d'étudier les performances de son test, Hadri (2000) a mené des simulations de Monte Carlo. Celles-ci font globalement ressortir que la précision du test est d'autant plus importante que  $T$  et  $N$  sont suffisamment importants. Plus spécifiquement, la taille du test  $Z_\mu$  est proche de la taille théorique de 5% pour  $T > 10$  et la taille du test  $Z_\tau$  est correcte pour  $T > 25$ . Concernant la puissance, il ressort que celle-ci augmente avec la valeur de  $\lambda$  pour tout  $T$  et  $N$ .

## 6 Comparaison et bilan des tests de première génération

Les simulations de Im *et al.* ont pour objet de comparer les performances des tests de Levin et Lin (LL) et d'IPS ( $Z_{tbar}$ ). Les principaux résultats peuvent s'énoncer comme suit :

- En l'absence de corrélation sérielle, et comparativement au test  $Z_{tbar}$ , le test LL a tendance à rejeter trop fréquemment l'hypothèse nulle lorsque  $N$  augmente. Par ailleurs, pour de faibles valeurs de  $T$ , le test  $Z_{tbar}$  donne de meilleurs résultats que le test LL en termes de puissance.
- En présence de corrélation sérielle, le test LL tend à nouveau à rejeter trop fréquemment l'hypothèse nulle, ceci étant d'autant plus marqué que  $N$  est élevé. Il ressort également que le test  $Z_{tbar}$  est plus puissant que le test LL. De manière générale, lorsque le nombre de retards dans les régressions ADF est correctement choisi ou sur-estimé, les performances du test  $Z_{tbar}$  sont meilleures que celles du test LL.

Les simulations effectuées par Breitung (2000) ont également pour objet de comparer les performances des tests  $Z_{tbar}$  et LL. Celles-ci font ressortir que la puissance de ces deux tests est très sensible à la spécification des termes déterministes et tend à diminuer fortement lorsque des tendances spécifiques individuelles sont incluses.

Selon Maddala et Wu (1999) et Levin *et al.* (2002), les comparaisons directes entre les tests LL et  $Z_{tbar}$  ne sont pas valides car, même si les deux tests reposent sur une hypothèse nulle identique, l'hypothèse alternative est différente. On rappelle en effet que, sous l'hypothèse alternative, le coefficient autorégressif est le même pour tous les individus pour le test LL, alors que ce même coefficient peut différer entre les individus pour le test  $Z_{tbar}$ . Par ailleurs, Maddala et Wu (1999) ont mené des simulations afin de comparer leur test (MW), le test LL et le test  $Z_{tbar}$ . Selon ces auteurs, le test  $Z_{tbar}$  et le test MW sont directement comparables dès lors que le test de racine unitaire utilisé est le test ADF. Leurs simulations mettent en évidence les résultats suivants :

- Pour des tailles d'échantillon égales entre les individus, lorsque le coefficient autorégressif  $\rho_i$  n'est pas trop variable et proche de 1, le test  $Z_{tbar}$  est plus puissant que le test MW. Lorsque, sous l'hypothèse alternative,  $\rho_i$  n'est pas trop proche de 1 (par exemple  $\rho_i = 0,8$ ),

le test MW est plus puissant que le test  $Z_{tbar}$ . Dans tous les cas, les tests  $Z_{tbar}$  et MW sont plus puissants que le test LL.

- Pour des tailles d'échantillon inégales, le test MW est plus performant que les tests  $Z_{tbar}$  et LL.
- En présence de corrélation dans les termes d'erreur entre les individus, le test MW est plus performant en termes de taille que le test  $Z_{tbar}$ , pour  $T$  grand et un nombre d'individus modéré. Lorsque  $T$  et  $N$  sont grands, la taille du test MW est comparable à celle du test  $Z_{tbar}$ .

Maddala et Wu (1999) concluent alors de manière générale que leur test est plus performant que les tests  $Z_{tbar}$  et LL. Les simulations menées par Choi (1999), visant également à comparer les performances des tests  $Z_{tbar}$  et de MW, montrent que la taille des deux tests est proche de la taille théorique de 5% lorsque  $N$  est faible et que la taille du test  $Z_{tbar}$  est plus stable que celle du test MW. En termes de puissance (corrigée par la taille), il ressort que le test MW donne de meilleurs résultats que le test  $Z_{tbar}$ . Enfin, la puissance des deux tests diminue de façon très importante lorsqu'est introduite une tendance déterministe.

Concernant le test de Hadri (2000), on peut simplement noter qu'un de ses avantages réside dans des considérations purement empiriques : les moments de sa distribution asymptotique sont dérivés de manière exacte et n'ont donc pas à être dérivés de simulations.

Au delà de leurs spécificités, ces tests de première génération butent sur deux problèmes identiques : le problème du caractère hétérogène de la racine unitaire et le problème de l'indépendance inter-individuelle. Le premier problème concerne uniquement les tests d'IPS et Maddala et Wu et tient au fait que l'hypothèse alternative autorise la présence d'un sous ensemble d'individus (de taille  $N_1$  chez IPS) dont la variable d'intérêt suit un processus stationnaire. Or, l'ajout ou le retrait d'un pays dans le panel peut dans ce cas modifier radicalement la conclusion des tests en faveur ou en défaveur de l'hypothèse nulle de racine unitaire. C'est par exemple le cas dans l'étude de Choi (2001) sur la validité de la PPA dont les résultats diffèrent suivant l'inclusion ou non du Japon. Ceci indique clairement que l'hypothèse nulle de ces tests, à savoir la présence d'une racine unitaire pour l'ensemble des individus du panel, n'est sans doute pas l'hypothèse qui présente le plus grand intérêt pour les économistes dans de nombreuses problématiques. Dans une démarche proche du "data mining", on aimerait être en mesure d'identifier deux ensembles d'individus : un ensemble avec racine unitaire (absence de convergence des PIB par exemple) et un ensemble pour lesquels la variable est stationnaire (convergence des PIB). C'est le cas sous l'hypothèse alternative d'IPS mais l'on ne connaît pas la taille  $N_1$  du groupe d'individus stationnaires, ni les individus appartenant à cet ensemble. L'identification de ces deux groupes suppose l'adoption d'une démarche séquentielle. Il s'agirait alors d'effectuer un test de racine unitaire avec par exemple le Japon, puis de refaire le test sur le même échantillon sans le Japon. Toutefois, il convient alors de contrôler le risque de première espèce. Cette approche séquentielle est notamment adoptée par Hénin, Jolivaldt et N'Guyen (2001) qui utilisent une statistique minimum définie à un ordre  $n$  par :

$$t^m(n) = \min_{\{i > N-n\}} t_{iT}$$

où  $t_{iT}$  correspond à la statistique ADF pour le pays  $i$  et où les indices  $i$  sont classés selon les réalisations croissantes de  $\hat{t}_{iT}$ . A un ordre  $n$ , on teste la présence d'une racine unitaire pour les  $n$  derniers individus du panel. En supposant que les statistiques ADF sont identiquement et indépendamment distribuées dans la dimension individuelle, on peut alors contrôler le risque de première espèce à un ordre  $n$  de la façon suivante<sup>18</sup> :

$$\Pr(t^m(n) < t_\alpha^{DF}) = 1 - (1 - \alpha)^n \left[ 1 - \Pr(t^m(n) < t_n) \right]^{-n}$$

<sup>18</sup>Cette inégalité est fondée sur l'inégalité de Bonferroni qui joue un rôle important dans les tests de seconde génération en ce qui concerne la détection de sous ensembles d'individus stationnaires ou non dans le panel.

où  $t_\alpha^{DF}$  correspond au seuil à  $\alpha\%$  de la loi de Dickey-Fuller et  $t_n$  correspond à la t-statistique ADF du pays exclu à l'étape  $n - 1$ . Les auteurs proposent de la même façon une statistique de Fisher itérée, ainsi qu'une statistique IPS itérée.

La seconde principale limite de ces tests de racine unitaire réside dans l'hypothèse d'indépendance entre individus : indépendance des résidus chez Levin et Lin ou chez IPS, indépendance des p-values chez Maddala et Wu. Nous avons vu que c'est précisément cette hypothèse qui permet de dériver très simplement les lois asymptotiques normales des statistiques de tests. Dès lors que cette hypothèse est violée, ces lois asymptotiques ne sont plus valides. Or, cette hypothèse d'indépendance pose problème notamment pour des applications macro-économiques. Prenons par exemple le cas des nombreux tests de convergence en panel construits à partir des tests de racine unitaire de première génération (Evans et Karras, 1996, Gaulier, Hurlin et Jean-Pierre, 1999). Ces tests reposent ainsi sur l'hypothèse d'indépendance inter-individuelle des écarts de PIB par tête à la moyenne internationale, éventuellement contrôlés par des effets temporels censés capter les effets de conjoncture internationale. Or rien ne garantit *a priori* que l'introduction des effets temporels ou le fait de centrer les données sur la moyenne internationale permette de purger les éventuelles dépendances inter-individuelles des revenus par tête.

Les auteurs de ces tests de première génération étaient pleinement conscients de cette limite. Dès la première version de leur document de travail, IPS (1997) envisagent ainsi la présence d'une corrélation inter-individuelle de la variable endogène, mais ils limitent leur analyse à l'introduction d'effets temporels dans la perspective d'un modèle à erreurs composées (voir Sevestre, 2002). Le résidu  $\varepsilon_{i,t}$  du modèle (32) se décompose alors en une composante commune strictement identique pour tous les individus  $\theta_t$  et une composante idiosyncratique  $v_{i,t}$ . D'une part, cette formulation des dépendances inter-individuelles est très restrictive puisque tous les pays sont influencés de façon symétrique par le facteur commun. D'autre part, Strauss et Yigit (2003) montrent qu'une telle approche ne fait que repousser le problème en abaissant l'éventuelle corrélation inter-individuelle si la composante idiosyncratique présente elle-même des corrélations entre individus. De même, Maddala et Wu (1999) dans un exercice de simulation envisagent la présence de co-mouvements inter-individuels en considérant une forme très particulière de corrélations retenue par O'Connell (1998) dans l'étude de la PPA. Au delà de la forme spécifique des corrélations, cette approche pose un certain nombre de problèmes techniques puisqu'il convient de tenir compte de ces corrélations dans la simulation par Bootstrap des p-values. Mais, que ce soit chez IPS ou chez Maddala et Wu, les corrélations inter-individuelles restent finalement considérées comme des paramètres de nuisance. Les tests de deuxième génération vont totalement renverser cette perspective.

## 7 L'hypothèse de dépendances inter-individuelles : vers une deuxième génération de tests

La littérature actuelle se développe autour de l'idée qu'il convient de prendre en compte de façon explicite les dépendances entre les individus du panel. Ainsi, ces tests de deuxième génération renversent totalement la perspective jusqu'alors adoptée, car plutôt que de considérer les corrélations entre individus comme des paramètres de nuisance, ils proposent d'exploiter ces co-mouvements pour définir de nouvelles statistiques de test. Contrairement aux approches développées dans le cadre des tests de première génération, les tests de deuxième génération ne considèrent pas nécessairement que les corrélations inter-individuelles de la variable  $y_{i,t}$  sont uniquement dues à une corrélation inter-individuelle des résidus. Ils envisagent notamment le cas où les corrélations de  $y_{i,t}$  proviennent de la présence d'une ou plusieurs composantes communes. Tout le problème consiste alors à proposer le test permettant la prise en compte la plus générale des différentes formes possibles de dépendance entre individus. Il convient en effet d'éviter de

construire un test pour une forme trop particulière de corrélation des résidus, de composante commune ou de cointégration entre individus, qui n’aurait pas de bonnes propriétés en termes de taille ou de puissance pour des formes alternatives.

De nombreux tests sont aujourd’hui développés dans cette perspective. La plupart s’inscrivent dans la lignée du test de Bai et Ng (2001, 2004) fondé sur un modèle à facteurs communs. Mais, dans ce cadre, différentes approches sont proposées. Si Bai et Ng (2001) considèrent deux tests séparés de racine unitaire sur les composantes commune et individuelle de la série, les autres études sont généralement fondées sur un test unique de racine unitaire de la série étudiée ; le point commun étant de tester la racine unitaire uniquement sur la composante idiosyncratique de la série. Les tests diffèrent alors suivant la méthode retenue pour extraire de la série brute la composante idiosyncratique inobservable. Ainsi, nous étudierons successivement les tests de Moon et Perron (2004), Choi (2002) et Pesaran (2003). Nous présenterons enfin le test de Chang (2002) qui lui n’est pas fondé sur un modèle factoriel et qui permet ainsi de considérer une forme générale de dépendance inter-individuelle.

## 7.1 Le test de Bai et Ng (2004)

Bai et Ng (2001, 2004) ont proposé le premier test de l’hypothèse nulle de racine unitaire prenant en compte la présence possible d’une corrélation inter-individuelle des séries testées. Le problème consiste alors à spécifier une forme particulière de ces dépendances. Bai et Ng adoptent une démarche très simple et considèrent un modèle factoriel :

$$y_{i,t} = D_{i,t} + \lambda_i' F_t + e_{i,t} \quad (45)$$

où  $D_{i,t}$  est une fonction polynomiale du temps d’ordre  $t$ ,  $F_t$  est un vecteur de dimension  $(r, 1)$  de facteurs communs et  $\lambda_i$  un vecteur de paramètres. Ainsi, la série individuelle  $y_{i,t}$  se décompose en une composante déterministe hétérogène  $D_{i,t}$ , une composante commune  $\lambda_i' F_t$  et un terme d’erreur  $e_{i,t}$  idiosyncratique. C’est donc la présence des facteurs communs  $F_t$ , par rapport auxquels chaque individu a une élasticité  $\lambda_i$  propre, qui est à l’origine des dépendances inter-individuelles.

Dans ce cas, la variable  $y_{i,t}$  est dite non stationnaire dès lors qu’au moins un des facteurs communs du vecteur  $F_t$  est non stationnaire et/ou le terme idiosyncratique  $e_{i,t}$  est non stationnaire. Rien ne garantit de façon générale que ces deux termes aient les mêmes propriétés dynamiques : l’un peut être stationnaire, l’autre pas, certaines composantes de  $F_t$  peuvent être  $I(0)$  d’autres  $I(1)$ ,  $F_t$  et  $e_{i,t}$  peuvent être intégrées d’ordre différents, *etc.* Or, on sait qu’une série définie par la somme de deux composantes aux propriétés dynamiques différentes a elle-même des propriétés dynamiques très différentes des entités qui la constituent. Ainsi, il peut être très difficile de diagnostiquer la non stationnarité de  $y_{i,t}$  si cette série admet une composante stationnaire importante. C’est pourquoi, plutôt que de tester la non stationnarité directement à partir de la série  $y_{i,t}$ , l’idée de Bai et Ng (2001) consiste à tester séparément la présence d’une racine unitaire dans les composantes commune et individuelle. Cette procédure est dénommée PANIC (*Panel Analysis of Nonstationarity in the Idiosyncratic and Common components*) par les auteurs. Quel est l’avantage de cette procédure au regard du problème des dépendances inter-individuelles ? C’est principalement que la composante idiosyncratique  $e_{i,t}$  peut être considérée comme faiblement corrélée entre individus alors que parallèlement la série totale  $y_{i,t}$  peut présenter de fortes corrélations entre individus. On lève ainsi une des principales critiques adressées aux tests de la première génération, notamment dans le cadre des applications macro-économiques.

Supposons que la composante déterministe  $D_{i,t}$  soit au plus d’ordre 1. Le modèle s’écrit alors sous la forme suivante :

$$y_{i,t} = \alpha_i + \beta_i t + \lambda_i' F_t + e_{i,t}, \quad t = 1, \dots, T \quad (46)$$

$$F_{m,t} = \tau_m F_{m,t-1} + v_{m,t}, \quad m = 1, \dots, r \quad (47)$$

$$e_{i,t} = \rho_i e_{i,t-1} + \varepsilon_{i,t}, \quad i = 1, \dots, N \quad (48)$$

Le  $m^{\text{ème}}$  facteur commun  $F_{m,t}$  est stationnaire si  $\tau_m < 1$ . La composante idiosyncratique  $e_{i,t}$  est stationnaire pour le  $i^{\text{ème}}$  individu si  $\rho_i < 1$ . L'objectif est d'appréhender la stationnarité de  $F_{m,t}$  et  $e_{i,t}$  sachant que ces composantes ne sont pas observées et doivent être estimées. Toute la validité de PANIC repose ainsi sur le fait qu'il soit possible d'obtenir des estimateurs de  $F_{m,t}$  et  $e_{i,t}$  préservant leur degré d'intégration, et ce que  $e_{i,t}$  soit  $I(0)$  ou  $I(1)$ . Pour cela les auteurs adoptent une analyse en composantes principales à partir des données différenciées. On suppose ici que le nombre de facteurs communs  $r$  est connu<sup>19</sup>.

Considérons le cas du modèle sans tendance ( $\beta_i = 0$ ) en différences premières :

$$\Delta y_{i,t} = \lambda'_i f_t + z_{i,t} \quad (49)$$

où  $z_{i,t} = \Delta e_{i,t}$  et où  $f_t = \Delta F_t$  vérifie  $E(f_t) = 0$ . On pose les définitions suivantes :

$$\underset{(N,r)}{\Lambda} = \begin{pmatrix} \lambda'_1 \\ (1,r) \\ \dots \\ \lambda'_N \\ (1,r) \end{pmatrix} \quad \underset{(T-1,r)}{f} = \begin{pmatrix} f'_2 \\ (1,r) \\ f'_3 \\ (1,r) \\ \dots \\ f'_T \\ (1,r) \end{pmatrix} \quad \underset{(T-1,N)}{X} = \begin{pmatrix} \Delta y_{1,2} & \dots & \Delta y_{N,2} \\ \Delta y_{1,3} & & \Delta y_{N,3} \\ \dots & & \dots \\ \Delta y_{1,T} & & \Delta y_{N,T} \end{pmatrix}$$

La procédure de test peut alors être décomposée en deux étapes. Dans une première étape, on estime  $f_t = \Delta F_t$  et  $\lambda_i$  dans le modèle (49) par une analyse en composantes principales. L'estimateur  $\hat{f}$  de la matrice  $f$  correspond alors au produit du scalaire  $\sqrt{T-1}$  par une matrice dont les colonnes sont définies par les  $r$  vecteurs propres associés aux  $r$  plus grandes valeurs propres de la matrice  $XX'$ . L'estimateur  $\hat{\Lambda}$  est défini par  $\hat{\Lambda} = X'\hat{f}/(T-1)$ . On note alors  $\hat{z}_{i,t} = \Delta y_{i,t} - \hat{\lambda}'_i \hat{f}_t$ .

Dans une deuxième étape, connaissant le vecteur des variations de la composante commune  $\hat{f}_t = \Delta \hat{F}_t$  ainsi que  $\hat{z}_{i,t}$ , on détermine les variables cumulées définies par :

$$\hat{F}_{m,t} = \sum_{s=2}^t \Delta \hat{F}_{m,s} = \sum_{s=2}^t \hat{f}_{m,s} \quad \hat{e}_{i,t} = \sum_{s=2}^t \hat{z}_{i,s} \quad (50)$$

avec  $t = 1, \dots, T$ ,  $m = 1, \dots, r$  et  $i = 1, \dots, N$ . On teste alors l'hypothèse nulle de racine unitaire dans la composante idiosyncratique  $e_{i,t}$  et dans les facteurs communs  $F_t$  à l'aide des variables estimées  $\hat{e}_{i,t}$  et  $\hat{F}_{m,t}$ .

Afin de tester la non stationnarité de la composante idiosyncratique, Bai et Ng ont proposé d'empiler les  $t$ -statistiques des tests ADF calculées sur la base des composantes estimées  $\hat{e}_{i,t}$  dans le cadre d'un modèle ne contenant pas de terme déterministe :

$$\Delta \hat{e}_{i,t} = \delta_{i,0} \hat{e}_{i,t-1} + \delta_{i,1} \Delta \hat{e}_{i,t-1} + \dots + \delta_{i,p} \Delta \hat{e}_{i,t-p} + \mu_{i,t} \quad (51)$$

Soit  $ADF_{\hat{e}}^c(i)$  la  $t$ -statistique du test ADF de la composante idiosyncratique du  $i^{\text{ème}}$  individu. La distribution asymptotique de la statistique  $ADF_{\hat{e}}^c(i)$  est identique à celle de la statistique usuelle de Dickey-Fuller du modèle sans constante. Un test de racine unitaire peut alors être effectué sur chacune des composantes idiosyncratiques du panel. La principale différence par rapport aux tests de racine unitaire en séries temporelles réside dans le fait que les facteurs communs ont été éliminés des données.

<sup>19</sup>Pour une estimation de  $r$ , voir Bai et Ng (2002) ou Moon et Perron (2004).

Naturellement, ces tests menés sur séries temporelles sont peu puissants pour des échantillons  $T$  de petite taille. C'est pourquoi Bai et Ng proposent d'utiliser soit une statistique moyenne à la IPS, soit une statistique à la Maddala et Wu construite comme suit :

$$Z^c = \frac{1}{\sqrt{4N}} \left[ -2 \sum_{i=1}^N \log p^c(i) - 2N \right] \quad (52)$$

où  $p^c(i)$  désigne la p-value associée à la statistique de Dickey-Fuller  $ADF_e^c(i)$ . Bai et Ng sont à ce moment là obligés de supposer l'indépendance entre individus des composantes individuelles inobservables  $e_{i,t}$  (théorème 3, page 8) pour dériver la loi de cette statistique. Cela peut paraître paradoxal puisque le test de Bai et Ng est justement censé prendre en compte ces dépendances inter-individuelles. Mais il faut bien comprendre ici que Bai et Ng ne supposent l'indépendance que des composantes individuelles  $e_{i,t}$  de la variable  $y_{i,t}$  définie par exclusion des composantes communes. On est bien loin dans ce cas de l'hypothèse d'indépendance inter-individuelle retenue par IPS ou Maddala et Wu qui portait sur la série totale  $y_{i,t}$ . C'est donc une hypothèse qui, sur le plan économique, est parfaitement justifiable dès lors qu'il n'y a pas d'erreur de spécification notamment sur le nombre  $r$  de composantes communes<sup>20</sup>.

Sous l'hypothèse d'indépendance des  $e_{i,t}$ , les statistiques de test basées sur les composantes estimées  $\hat{e}_{i,t}$  sont elles aussi indépendantes. Dès lors, les p-values  $p^c(i)$  sont distribuées de façon indépendante selon des lois uniformes sur  $[0, 1]$ . En conclusion, sous l'hypothèse que l'ensemble des composantes individuelles  $e_{i,t}$  pour  $i = 1, \dots, N$  sont  $I(1)$ , la statistique de test  $Z^c$  suit une loi  $N(0, 1)$  quelle que soit la taille  $N$  du panel. On peut ici adopter la standardisation de Choi (2001) pour des panels de taille  $N$  importante (cf. section précédente).

Afin de tester la non stationnarité des facteurs communs, Bai et Ng (2004) distinguent deux cas<sup>21</sup>. Lorsqu'il y a un seul facteur commun parmi les  $N$  variables ( $r = 1$ ), ils utilisent un test ADF standard dans le cas d'un modèle avec constante :

$$\Delta \hat{F}_{1,t} = c + \gamma_{i,0} \hat{F}_{1,t-1} + \gamma_{i,1} \Delta \hat{F}_{1,t-1} + \dots + \gamma_{i,p} \Delta \hat{F}_{1,t-p} + v_{i,t} \quad (53)$$

La  $t$ -statistique correspondante, notée  $ADF_{\hat{F}}^c$ , a la même distribution asymptotique que la statistique usuelle de Dickey-Fuller dans le cas d'un modèle avec constante. S'il existe plus d'un facteur commun ( $r > 1$ ), Bai et Ng testent le nombre  $r_1$  de tendances stochastiques communes dans ces facteurs communs. Bien évidemment, si  $r_1 = 0$ , alors il existe  $N$  vecteurs de cointégration pour  $N$  facteurs communs, et tous les facteurs sont  $I(0)$ . Tester individuellement la présence d'une racine unitaire dans chacun des facteurs tend généralement à surestimer le nombre de tendances communes. Bai et Ng proposent alors deux statistiques basées sur les  $r$  facteurs estimés  $\hat{F}_{m,t}$  pour  $m = 1, \dots, r$ . Ces statistiques sont similaires à celles proposées par Stock et Watson (1988). L'objet est de tester si la partie réelle de la plus petite valeur propre de la matrice des coefficients autorégressifs est égale à l'unité. Les deux statistiques nécessitent des tests successifs de séquences d'hypothèses, à l'instar des tests du nombre de vecteurs de cointégration de Johansen (1988). L'hypothèse nulle est définie par  $H_0 : r_1 = m$ . Si l'hypothèse nulle est rejetée, on pose  $m = m - 1$  et on réapplique le test. Dans le cas inverse, le nombre estimé de tendances communes, noté  $\hat{r}_1$ , est égal à  $m$ . On teste l'égalité entre le nombre de tendances communes et le nombre de facteurs communs, *i.e.*  $r_1 = r$ . La première statistique de test,

<sup>20</sup> En pratique, le nombre de composantes communes est rarement connu et doit être estimé. Comme l'ont montré Bai et Ng (2002), pour des panels d'au moins 20 individus, le nombre de composantes communes peut être estimé avec une précision parfaite. En revanche, cela n'est plus le cas lorsque le nombre d'individus est moindre. Dans ce dernier cas, la méthode d'estimation de Bai et Ng (2002) tend à fortement surestimer le nombre de facteurs communs (voir Moon et Perron, 2004). Cette surestimation du nombre de facteurs communs, lorsque le nombre d'individus est faible, engendre en conséquence d'importantes distorsions de taille.

<sup>21</sup> Dans leur premier document de travail (Bai et Ng, 2001), la procédure était identique quel que soit le nombre de facteurs communs et reposait uniquement sur des tests ADF.

notée  $MQ_f$ , suppose que les composantes non stationnaires suivent un processus autorégressif vectoriel d'ordre fini. La deuxième statistique, notée  $MQ_c$ , autorise la présence de dynamiques plus générales pour le processus de racine unitaire.

Une démarche similaire peut être adoptée dans le cas d'un modèle incluant des tendances déterministes propres à chaque individu ( $\beta_i \neq 0$  dans le modèle 46). La seule différence réside alors dans l'étape 1, où la variable  $\Delta y_{i,t}$  doit être remplacée par  $\Delta y_{i,t} - \overline{\Delta y_i}$ , avec  $\overline{\Delta y_i} = (T-1)^{-1} \sum_{t=2}^T \Delta y_{i,t}$ . De la même façon, dans le modèle (49), l'accroissement de la composante commune doit être centré, ce qui revient à remplacer  $f_t = \Delta F_t$  par  $\Delta F_t - \overline{\Delta F_t}$ , de même pour la composante idiosyncratique  $\Delta e_{i,t} - \overline{\Delta e_{i,t}}$ . Dans ce cas, la statistique de test associée à la composante commune de l'étape 2, notée  $ADF_F^r(m)$ , suit sous  $H_0$  la distribution de Dickey-Fuller pour un modèle avec constante et tendance. La statistique de test associée à la composante individuelle de l'individu  $i$ , notée  $ADF_e^r(i)$ , admet une distribution asymptotique reliée à celle d'un pont brownien identique à celle du test de Schmidt et Lee (1991).

Ainsi que le notent Banerjee et Zanghieri (2003), la mise en oeuvre des tests de Bai et Ng (2001) telle qu'elle est décrite ici montre bien le rôle que peut avoir la présence de co-mouvements entre les individus. En effet, pour les panels pour lesquels il existe une forte dépendance entre les individus, les tests de Bai et Ng — en prenant en compte les facteurs communs aux diverses séries — acceptent l'hypothèse nulle de racine unitaire dans les facteurs, ce qui les amènent à conclure à la non stationnarité des séries. Pour finir, notons que les simulations menées par Bai et Ng (2001) montrent que leur test donne des résultats satisfaisants en termes de taille et de puissance, même pour des panels de taille modérée ( $N = 20$ ).

## 7.2 Les tests de Phillips et Sul (2003a) et Moon et Perron (2004)

Les tests de Phillips et Sul (2003a) et Moon et Perron (2004), contrairement à Bai et Ng (2001), testent directement la présence d'une racine unitaire dans la série observable  $y_{i,t}$  et non pas de façon séparée dans les composantes individuelle et commune. Au delà de cette différence fondamentale, il existe certaines similitudes entre les deux approches qui reposent sur l'utilisation d'un modèle factoriel. Nous ne présenterons en détail que le test de Moon et Perron (2004) qui admet la spécification la plus générale des composantes communes.

Moon et Perron (2004) considèrent un modèle autorégressif standard avec effets individuels fixes dans lequel les résidus satisfont un modèle factoriel. La structure du modèle est donc différente de celle de Bai et Ng (2001). Si l'on conserve les mêmes notations que précédemment, leur modèle s'écrit :

$$y_{i,t} = \alpha_i + y_{i,t}^0 \quad (54)$$

$$y_{i,t}^0 = \phi_i y_{i,t-1}^0 + \mu_{i,t} \quad (55)$$

$$\mu_{i,t} = \lambda_i' F_t + e_{i,t} \quad (56)$$

On suppose dans un premier temps que la dimension  $r$  du vecteur  $F_t$  est *a priori* connue et que les chocs idiosyncratiques  $e_{i,t} = \sum_{j=0}^{\infty} d_{i,j} v_{i,t-j}$ , avec  $v_{i,t}$  *i.i.d.*  $(0, 1)$  sont non corrélés dans la dimension individuelle. Dès lors, la corrélation inter-individuelle des variables  $y_{i,t}$  est déterminée par le vecteur  $\lambda_i$  puisque  $E(\mu_{i,t} \mu_{j,t}) = \lambda_i' E(F_t F_t') \lambda_j$ . On teste l'hypothèse nulle de racine unitaire pour tous les individus du panel  $H_0 : \phi_i = 1, \forall i = 1, \dots, N$  contre  $H_1 : \phi_i < 1$  pour au moins un individu  $i$ .

L'intuition de la démarche de Moon et Perron (2004) est alors la suivante : il s'agit de transformer le modèle de sorte à éliminer les composantes communes de la série  $y_{i,t}$ , puis de tester la racine unitaire sur les séries en écarts aux facteurs communs. Ainsi, on supprime les dépendances inter-individuelles et l'on peut se ramener à des distributions asymptotiques normales. On retrouve alors des distributions normales comme chez IPS ou Levin et Lin, mais la

différence fondamentale est que les statistiques de tests sont construites ici à partir de données transformées, prises en écarts aux composantes communes et donc indépendantes dans la dimension individuelle.

On suppose que l'on dispose initialement de  $T + 1$  observations de la variable d'intérêt  $y_{i,t}$ . Soit la matrice  $Z$  des observations individuelles de  $y_{i,t}$  et  $Z_{-1}$  la matrice des observations retardées. On note enfin  $\Lambda$  la matrice  $(N, r)$  des coefficients  $\lambda_i$  (cf. section précédente) et  $F$  la matrice des composantes communes :

$$\underset{(T,N)}{Z} = \begin{pmatrix} y_{1,2} & \dots & y_{N,2} \\ \dots & & \\ y_{1,T+1} & & y_{N,T+1} \end{pmatrix} \quad \underset{(T,N)}{Z_{-1}} = \begin{pmatrix} y_{1,1} & \dots & y_{N,1} \\ \dots & & \\ y_{1,T} & & y_{N,T} \end{pmatrix} \quad \underset{(T,r)}{F} = \begin{pmatrix} F'_1 \\ \dots \\ F'_T \end{pmatrix}$$

Comme l'avons précédemment mentionné, l'idée de Moon et Perron (2004) consiste à tester la racine unitaire sur les écarts aux composantes communes. Supposons que les paramètres de  $\Lambda$  soient connus, les composantes d'écart s'écrivent tout simplement comme la projection de  $Z$  sur l'orthogonal du sous espace vectoriel engendré par les colonnes de  $\Lambda$ . En effet, considérons pour simplifier le modèle (55) sous l'hypothèse nulle de racine unitaire ( $\rho_i = 1$ ) et en l'absence d'effets individuels ( $\alpha_i = 0$ ) écrit sous forme vectorielle :

$$Z = Z_{-1} + F \Lambda' + e$$

où  $e$  désigne la matrice  $(T, N)$  des composantes idiosyncratiques. Soit  $Q_\Lambda = I_N - \Lambda(\Lambda'\Lambda)^{-1}\Lambda'$  la matrice de projection sur l'orthogonal du sous espace vectoriel engendré par les colonnes de  $\Lambda$ . En pré-multipliant à droite par  $Q_\Lambda$ , on exprime alors le modèle à partir des données en écarts aux composantes communes  $ZQ_\Lambda$  :

$$ZQ_\Lambda = Z_{-1}Q_\Lambda + eQ_\Lambda \quad (57)$$

La projection  $ZQ_\Lambda$  correspond aux données exprimées en écarts aux facteurs communs, tandis que les résidus  $eQ_\Lambda$  ne présentent par construction aucune corrélation inter-individuelle. C'est alors à partir de ces données transformées que l'on va effectuer le test de racine unitaire. Moon et Perron construisent une statistique de test à partir de l'estimateur *pooled*<sup>22</sup> ( $\phi_i = \phi_j$ ) de la racine autorégressive. Plus précisément, ils considèrent un estimateur *pooled* corrigé en raison de l'éventuelle autocorrélation (intra-individuelle) des résidus  $eQ_\Lambda$  :

$$\hat{\phi}_{pool}^+ = \frac{\text{trace}(Z_{-1}Q_\Lambda Z') - NT\lambda_e}{\text{trace}(Z_{-1}Q_\Lambda Z'_{-1})} \quad (58)$$

avec  $\lambda_e = N^{-1} \sum_{i=1}^N \lambda_e^i$  où le terme  $\lambda_e^i$  désigne la somme des autocovariances positives des composantes idiosyncratiques :

$$\lambda_e^i = \sum_{l=1}^{\infty} \sum_{j=0}^{\infty} d_{i,j} d_{i,j+l}$$

A partir de cet estimateur  $\hat{\phi}_{pool}^+$ , Moon et Perron proposent deux statistiques de test de l'hypothèse nulle de racine unitaire, notées respectivement  $t_a$  et  $t_b$ . Ces deux statistiques convergent lorsque  $T$  et  $N$  tendent vers l'infini et que le rapport  $N/T$  tend vers 0.

$$t_a = \frac{T\sqrt{N} \left( \hat{\phi}_{pool}^+ - 1 \right)}{\sqrt{2\gamma_e^4/w_e^4}} \xrightarrow{T,N \rightarrow \infty} N(0,1) \quad (59)$$

<sup>22</sup>Ils justifient ce choix par le fait que cet estimateur simplifie l'étude des lois asymptotiques et leur permet en outre d'étudier leur test sous une alternative locale de quasi-racine unitaire.

$$t_b = T\sqrt{N} \left( \hat{\phi}_{pooled}^+ - 1 \right) \sqrt{\frac{1}{NT^2} \text{trace} (Z_{-1} Q Z'_{-1})} \frac{w_e^2}{\gamma_e^4} \xrightarrow[T, N \rightarrow \infty]{l} N(0, 1) \quad (60)$$

Les quantités  $w_e^2$  et  $\gamma_e^4$  correspondent respectivement aux moyennes sur  $N$  des variances individuelles de long terme  $w_{e,i}^2$ , et des variances individuelles de long terme au carré  $\gamma_{e,i}^4$ , de la composante idiosyncratique  $e_{i,t}$  avec  $w_{e,i}^2 = \left( \sum_{j=0}^{\infty} d_{i,j} \right)^2$ . Si la réalisation de la statistique  $t_a$  (ou  $t_b$ ) est inférieure au seuil de la loi normale, on rejette l'hypothèse nulle de racine unitaire pour tous les individus du panel.

Notons au passage que l'on retrouve la vitesse de convergence en  $T\sqrt{N}$  de l'estimateur *pooled* (corrigé ou non) de la racine autorégressive que l'on avait obtenu dans le cas du modèle de Levin et Lin (équation 11). Ceci est parfaitement normal, puisque une fois que les dépendances inter-individuelles ont été purgées dans le modèle de Phillips et Moon, on obtient sur données transformées un modèle avec racine autorégressive commune (estimateur *pooled*) identique à celui de Levin et Lin sous l'hypothèse d'indépendance inter-individuelle.

Naturellement, puisque la composante idiosyncratique  $e_{i,t}$  est inobservable, les quantités  $w_e^2$ ,  $\lambda_e$  et  $\gamma_e^4$  sont inconnues. Les définitions (59) et (60) ne sont pas directement applicables. Ainsi, afin d'appliquer ces tests, il nous faut deux ingrédients supplémentaires (liés) :

1. Un estimateur des composantes commune et idiosyncratique : plus précisément un estimateur  $\hat{Q}_\Lambda$  de la matrice de projection sur l'orthogonal du sous espace vectoriel engendré par les colonnes de  $\Lambda$ , matrice des paramètres  $\lambda_i$  associés aux composantes communes.
2. Des estimateurs de la variance de long terme ( $\hat{w}_e^2$  et  $\hat{\gamma}_e^4$ ) et de la somme des autocovariances positives ( $\hat{\lambda}_e$ ), construits à partir des estimations des composantes individuelles  $\hat{e}_{i,t}$ .

### Estimation de la matrice de projection

Tout comme Bai et Ng, Moon et Perron proposent d'estimer  $\Lambda$  par une analyse en composantes principales (cf. section précédente) des résidus. On commence par estimer les résidus  $\mu_{i,t}$  à partir du modèle (55). Pour cela, on utilise un estimateur *pooled*  $\hat{\phi}_{pooled}$  sur les données brutes centrées sur leurs moyennes individuelles pour tenir compte des effets individuels :

$$\hat{\mu} = \tilde{Z} - \hat{\phi}_{pooled} \tilde{Z}_{-1} \quad (61)$$

où  $\tilde{Z} = Q_x Z$  et  $Q_x = I_T - T^{-1} l_T l_T'$  où  $l_T = (1, \dots, 1)'$ , l'estimateur *pooled* étant défini par  $\hat{\phi}_{pooled} = \text{trace}(\tilde{Z}'_{-1} \tilde{Z}) / \text{trace}(\tilde{Z}'_{-1} \tilde{Z}_{-1})$ . A partir des résidus  $\hat{\mu}$ , on applique une analyse en composantes principales. Un estimateur  $\tilde{\Lambda}$  de la matrice  $(N, r)$  des coefficients des composantes communes correspond alors au produit du scalaire  $\sqrt{N}$  par une matrice dont les colonnes sont définies par les  $r$  vecteurs propres associés aux  $r$  plus grandes valeurs propres de la matrice  $\hat{\mu}' \hat{\mu}$ . Moon et Perron utilisent un estimateur  $\hat{\Lambda}$  re-paramétré :

$$\hat{\Lambda} = \tilde{\Lambda} \left( \frac{1}{N} \tilde{\Lambda}' \tilde{\Lambda} \right)^{\frac{1}{2}} \quad (62)$$

A partir de cet estimateur  $\hat{\Lambda}$ , on définit un estimateur de la matrice de projection  $\hat{Q}_\Lambda$  qui nous permettra d'obtenir par la suite une estimation des composantes idiosyncratiques :

$$\hat{Q}_\Lambda = I_N - \hat{\Lambda} \left( \hat{\Lambda}' \hat{\Lambda} \right)^{-1} \hat{\Lambda}' \quad (63)$$

Une estimation des données en écarts aux composantes communes est ainsi définie par  $\tilde{Z} \hat{Q}_\Lambda$ . On peut donc en déduire un estimateur *pooled* corrigé  $\hat{\phi}_{pooled}^+$  de la racine autorégressive uniquement sur ces composantes d'écarts selon la formule (58) en remplaçant  $Z Q_\Lambda$  par  $\tilde{Z} \hat{Q}_\Lambda$ .

### Estimation des variances de long terme

Un estimateur de la composante idiosyncratique est alors défini par  $\hat{e} = \hat{\mu}\hat{Q}_\Lambda$ . Soit  $\hat{e}_{i,t}$  la composante individuelle des résidus de l'individu  $i$  à la date  $t$ . Pour chaque individu  $i = 1, \dots, N$ , on définit l'autocovariance empirique des résidus  $\{\hat{e}_{i,t}\}_{t=1}^T$  :

$$\hat{\Gamma}_i(j) = \frac{1}{T} \sum_{t=1}^{T-j} \hat{e}_{i,t} \hat{e}_{i,t+j}$$

A partir de  $\hat{\Gamma}_i(j)$  on construit un estimateur à noyau de la variance de long terme et de la somme des autocovariances positives comme suit :

$$\hat{w}_{e,i}^2 = \sum_{j=-T+1}^{T-1} w(q_i, j) \hat{\Gamma}_i(j) \quad (64)$$

$$\hat{\lambda}_{e,i} = \sum_{j=1}^{T-1} w(q_i, j) \hat{\Gamma}_i(j) \quad (65)$$

où  $w(q_i, j)$  désigne une fonction noyau et  $q_i$  un paramètre de troncature. Ne reste plus alors qu'à définir les estimateurs des moyennes des variances de long terme individuelles :

$$\hat{w}_e^2 = \frac{1}{N} \sum_{i=1}^N \hat{w}_{e,i}^2 \quad \hat{\lambda}_e = \frac{1}{N} \sum_{i=1}^N \hat{\lambda}_{e,i} \quad \hat{\gamma}_e^4 = \frac{1}{N} \sum_{i=1}^N (\hat{w}_{e,i}^2)^2 \quad (66)$$

Les fonctions noyau et le paramètre de troncature  $q$  doivent satisfaire un ensemble de trois hypothèses pour que les statistiques  $\tilde{t}_a$  et  $\tilde{t}_b$  définies par les équations (59) et (60) en substituant  $w_e^2$ ,  $\lambda_e$  et  $\gamma_e^4$  par leurs estimateurs, convergent vers des lois normales. Moon et Perron utilisent notamment une fonction de type noyau QS (voir Salanié, 1999) :

$$w(q_i, j) = \frac{25}{12\pi^2 x^2} \left[ \frac{\sin(6\pi x/5)}{6\pi x/5} - \cos\left(\frac{6\pi x}{5}\right) \right] \quad \text{pour } x = \frac{j}{q_i}$$

où le paramètre de troncature optimal  $q_i$  est défini par la relation :

$$q_i = 1.3221 \left[ \frac{4\hat{\rho}_{i,1}^2 T_i}{(1 - \hat{\rho}_{i,1})^4} \right]^{1/5}$$

où  $\hat{\rho}_{i,1}^2$  désigne l'estimateur de l'autocorrélation d'ordre 1 de la composante individuelle  $\hat{e}_{i,t}$  de l'individu  $i$ .

Phillips et Sul (2003a) considèrent un modèle plus restrictif que celui de Moon et Perron (2004) puisqu'il ne contient qu'un seul facteur indépendamment distribué dans le temps. Le vecteur de facteurs communs  $F_t$  est réduit à une variable  $\theta_t$  *N.i.d.* (0, 1). Au delà de cette distinction, la principale différence réside dans le fait que Phillips et Sul adoptent une méthode de moments pour éliminer les facteurs communs au lieu d'une analyse en composantes principales. L'idée est de tester la racine unitaire sur des données orthogonalisées : celles-ci étant par construction indépendantes dans la dimension individuelle, il est possible alors d'appliquer les tests de racine unitaire de première génération. Ainsi, à partir des estimateurs des racines autorégressives individuelles obtenues sur données orthogonalisées, Phillips et Sul construisent différentes statistiques de tests de racine unitaire : soit des statistiques moyennes similaires à

celles d'IPS, soit des statistiques définies comme des combinaisons de p-values associées à des tests individuels de racine unitaire.

Moon et Perron (2004) ont effectué des simulations de Monte Carlo afin d'étudier les propriétés de leurs tests en termes de taille et de puissance. Leurs résultats peuvent être résumés comme suit. En premier lieu, en l'absence de tendance déterministe, le test  $t_b$  donne de meilleurs résultats en termes de taille que le test  $t_a$ . Par ailleurs, plus les composantes communes sont importantes par rapport à la composante idiosyncratique, plus il est difficile de contrôler la taille des tests pour de faibles valeurs de  $N$ . Les deux tests donnent de très bons résultats en termes de puissance, y compris face à l'hypothèse alternative d'une valeur moyenne égale à 0,99 pour le paramètre autorégressif (ce dernier variant en outre selon les individus). En second lieu, lorsqu'une tendance déterministe est présente, les résultats en termes de taille restent valables, mais la puissance des tests diminue de façon drastique. Au final, Moon et Perron (2004) concluent que leur test nécessite un minimum de 20 individus afin de disposer d'une estimation précise du nombre de facteurs et de donner des résultats fiables (voir note de bas de page numéro 20). Dans ces conditions, les tests ont de bonnes propriétés en termes de taille et de puissance.

### 7.3 Le test de Choi (2002)

Tout comme Moon et Perron (2004), Choi (2002) teste la racine unitaire à partir d'une transformation de la série observée  $y_{i,t}$  permettant d'éliminer les corrélations inter-individuelles et les éventuelles composantes de tendance déterministes. Mais si dans le principe l'approche de Choi (2002) est identique à celle de Moon et Perron, elle en diffère sur deux aspects essentiels. Tout d'abord, Choi considère un modèle à erreurs composées :

$$y_{i,t} = \alpha_i + \theta_t + v_{i,t} \quad (67)$$

$$v_{i,t} = \sum_{j=1}^{p_i} d_{i,j} v_{i,t-j} + \varepsilon_{i,t} \quad (68)$$

où  $\varepsilon_{i,t}$  est *i.i.d.*  $(0, \sigma_{\varepsilon,i}^2)$  et indépendamment distribué entre individus. L'effet temporel  $\theta_t$  est représenté par un processus stationnaire. Dans ce modèle, contrairement aux approches de Bai et Ng (2001) et Moon et Perron (2004), il n'existe qu'un seul facteur commun ( $r = 1$  dans nos notations) représenté par l'effet temporel  $\theta_t$ . Mais, plus fondamentalement, le modèle de Choi impose que les variables individuelles  $y_{i,t}$  répondent de façon homogène à l'unique facteur commun. C'est là une différence fondamentale avec Phillips et Sul (2003a) qui eux aussi considèrent un seul facteur commun, mais qui envisagent une spécification hétérogène de la sensibilité à ce facteur, du type  $\lambda_i \theta_t$ . Choi justifie son choix par le fait que le passage en logarithme du modèle (67) permet d'introduire une telle sensibilité. Mais surtout Choi avance l'argument qu'il est possible dans son modèle de tester l'hypothèse de stationnarité du processus  $\theta_t$ , ce qui n'est pas le cas avec une sensibilité hétérogène. C'est un point important puisque, dans une optique macro-économique, ces effets temporels sont souvent censés capter les inflexions de la conjoncture internationale et rien ne garantit que ces effets soient stationnaires.

Dans le modèle (67), on teste l'hypothèse nulle d'une racine unitaire dans la composante idiosyncratique  $v_{i,t}$  pour tous les individus du panel, ce qui s'écrit  $H_0 : \sum_{j=1}^{p_i} d_{i,j} = 1, \forall i = 1, \dots, N$  contre l'hypothèse alternative selon laquelle il existe des individus  $i$  tels que  $\sum_{j=1}^{p_i} d_{i,j} < 1$ .

La seconde différence avec l'approche de Moon et Perron (2004), liée pour partie à la spécification du modèle, réside dans la manière d'orthogonaliser les séries individuelles  $y_{i,t}$  qui seront utilisées par la suite pour tester la racine unitaire uniquement sur la composante individuelle. Pour supprimer les dépendances inter-individuelles, Choi isole  $v_{i,t}$  en éliminant la

constante (effet individuel)  $\alpha_i$ , mais aussi et surtout le terme d'erreur commun  $\theta_t$  (effet temporel). Pour éliminer ces composantes déterministes, Choi procède en deux étapes : élimination de la constante par l'approche d'Elliott, Rothenberg et Stock (1996), ERS par la suite, puis élimination de l'effet temporel par centrage sur la moyenne individuelle. En effet, si la composante  $v_{i,t}$  est stationnaire, l'application des MCO permet d'obtenir un estimateur efficace du terme constant. Toutefois, si  $v_{i,t}$  est  $I(1)$  ou présente une quasi racine unitaire, l'approche d'ERS consistant à estimer le terme constant sur données quasi-différenciées par les MCG, permet *in fine* d'obtenir de meilleures propriétés à distance finie pour les tests de racine unitaire. Pour cette raison, Choi (2002) utilise l'approche d'ERS, ce qui constitue, à notre connaissance, la première extension de cette approche en panel. Reprenons à présent dans le détail ces deux étapes de l'orthogonalisation des séries individuelles.

**Étape 1 :** Si l'on suppose que la plus grande racine du processus  $v_{i,t}$  est  $1 + c/T$  (processus de quasi racine unitaire), pour chaque individu  $i = 1, \dots, N$ , on construit deux séries  $\tilde{y}_{i,t}$  et  $\tilde{c}_{i,t}$  quasi différenciées telles que pour  $t \geq 2$  :

$$\tilde{y}_{i,t} = y_{i,t} - \left(1 + \frac{c}{T}\right) y_{i,t-1} \quad \tilde{c}_{i,t} = 1 - \left(1 + \frac{c}{T}\right) \quad (69)$$

Choi considère pour tous les individus du panel la valeur de la constante  $c = -7$  fournie par ERS dans le cas d'un modèle sans dérive. On régresse ensuite par les MCG  $\tilde{y}_{i,t}$  sur la variable déterministe  $\tilde{c}_{i,t}$ . Soit  $\hat{\alpha}_i$  l'estimateur des MCG obtenu pour chaque individu  $i$  sur données quasi-différenciées. Pour  $T$  suffisamment grand, on doit alors observer :

$$y_{i,t} - \hat{\alpha}_i \simeq \theta_t - \theta_1 + v_{i,t} - v_{i,1}$$

et ceci que le processus  $v_{i,t}$  soit  $I(1)$  ou quasi intégré. Ne reste plus alors qu'à éliminer la composante commune  $\theta_t$  source de corrélation entre individus : c'est précisément l'objet de la deuxième étape.

**Étape 2 :** Choi propose de centrer la variable  $y_{i,t} - \hat{\alpha}_i$  sur sa moyenne individuelle et de définir ainsi une nouvelle variable  $z_{i,t}$  comme suit :

$$z_{i,t} = (y_{i,t} - \hat{\alpha}_i) - \frac{1}{N} \sum_{i=1}^N (y_{i,t} - \hat{\alpha}_i) \quad (70)$$

En effet, compte tenu des résultats précédents, pour chaque individu  $i = 1, \dots, N$ , on peut montrer que :

$$z_{i,t} \simeq (v_{i,t} - \bar{v}_t) - (v_{i,1} - \bar{v}_1)$$

où  $\bar{v}_t = (1/N) \sum_{i=1}^N v_{i,t}$ . Ainsi, les composantes déterministes  $\alpha_i$  et  $\theta_t$  sont éliminées de la définition de  $z_{i,t}$ . Les processus  $z_{i,t}$  peuvent être considérés comme indépendants dans la dimension individuelle puisque les moyennes  $\bar{v}_1$  et  $\bar{v}_t$  convergent en probabilité vers 0 dès lors que  $N$  tend vers l'infini. Ne reste plus alors qu'à effectuer un test de racine unitaire sur les séries transformées  $z_{i,t}$ .

Dans le cas d'un modèle avec dérivées individuelles, la démarche est identique à deux différences près : pour obtenir les estimateurs  $\hat{\alpha}_i$  et  $\hat{\gamma}_i$  (coefficient associé à la tendance individuelle) on régresse par MCG la série  $\tilde{y}_{i,t}$  sur  $\tilde{c}_{i,t}$  et  $\tilde{d}_{i,t} = 1 - c/T$ . On choisit alors la valeur  $c = -13.5$  fournie par ERS dans le cas d'un modèle avec dérive. On construit  $z_{i,t}$  de la manière suivante :

$$z_{i,t} = (y_{i,t} - \hat{\alpha}_i - \hat{\gamma}_i t) - \frac{1}{N} \sum_{i=1}^N (y_{i,t} - \hat{\alpha}_i - \hat{\gamma}_i t)$$

A partir des séries  $\{z_{i,t}\}_{t=2}^T$ , on effectue des tests individuels de racine unitaire sans constante ni tendance quel que soit le modèle considéré puisque toutes les composantes déterministes ont été retirées :

$$\Delta z_{i,t} = \rho_i z_{i,t-1} + \sum_{j=1}^{p_i-1} \beta_{i,j} \Delta z_{i,t-j} + u_{i,t} \quad (71)$$

Choi (2002) montre que dans un modèle sans constante, la t-statistique  $t_{\rho_i}^{ERS}$  de Dickey-Fuller suit sous  $H_{0,i} : \rho_i = 0$  la distribution asymptotique de Dickey-Fuller lorsque  $T$  et  $N$  tendent vers l'infini. En présence d'une dérive, cette statistique admet une distribution tabulée par ERS (voir Salanié, 1999). Encore une fois, ces tests menés sur séries temporelles bien que plus puissants que les tests ADF standard, restent tout de même peu puissants au regard des tests en panel. C'est pourquoi, Choi propose des statistiques de test en panel fondées sur les statistiques  $t_{\rho_i}^{ERS}$  individuelles qui sont, et c'est ce qui est fondamental, par construction indépendantes les unes des autres. Naturellement Choi (2002) s'inspire de ses travaux antérieurs (Choi, 2001) et propose finalement trois statistiques de test fondées sur des combinaisons de niveaux de significativité de tests individuels (cf. section précédente) :

$$P_m = -\frac{1}{\sqrt{N}} \sum_{i=1}^N [\ln(p_i) + 1] \quad (72)$$

$$Z = -\frac{1}{\sqrt{N}} \sum_{i=1}^N \Phi^{-1}(p_i) \quad (73)$$

$$L^* = \frac{1}{\sqrt{\pi^2 N/3}} \sum_{i=1}^N \ln\left(\frac{p_i}{1-p_i}\right) \quad (74)$$

où  $p_i$  désigne le niveau de significativité associé à la statistique  $t_{\rho_i}^{ERS}$  et  $\Phi(\cdot)$  la fonction de répartition de la loi normale centrée et réduite. Sous l'hypothèse nulle de racine unitaire pour tous les individus du panel, Choi (2002) montre que ces trois statistiques convergent lorsque  $T$  et  $N$  tendent conjointement vers l'infini vers une loi normale centrée réduite. Les règles de décision sont les suivantes. Pour la statistique de Fisher transformée  $P_m$ , si la réalisation est supérieure au seuil de la loi normale (1.64 à 5% de risque de première espèce), on rejette  $H_0$ . Pour les deux statistiques  $Z$  et  $L^*$ , si la réalisation est inférieure au seuil de la loi normale (-1.64 à 5% de risque de première espèce), on rejette  $H_0$ .

Tout comme pour les travaux de Choi (2001) ou Maddala et Wu (1999), la principale difficulté de cette approche réside dans la nécessité de simuler par Bootstrap les p-values  $p_i$  utilisées dans la construction des statistiques  $P_m$ ,  $Z$  et  $L^*$  et ce d'autant plus que l'on utilise les distributions d'ERS. Choi reprend la méthodologie de simulation de MacKinnon (1994). Il convient de noter que les p-values ainsi obtenues peuvent être relativement sensibles à la loi utilisée pour générer le processus  $u_t$ . Pour des tailles  $T$  importantes, le fait de supposer à tort la normalité des résidus  $u_t$  ne pose pas de problème. En revanche, lorsque  $T$  est faible les p-values obtenues sous l'hypothèse de normalité pourraient être sensiblement différentes des vraies valeurs si la distribution est différente de la loi normale. Plus spécifiquement, les simulations menées par Choi (2002) visant à comparer les trois tests proposés  $P_m$ ,  $Z$  et  $L^*$  et à étudier l'impact de la non normalité des erreurs montrent que :

- Les trois tests donnent de bons résultats en termes de taille, d'autant plus que  $T$  augmente.
- La puissance des trois tests augmente lorsque  $N$  augmente, ce qui justifie l'utilisation de données de panel. Lorsqu'une tendance déterministe est incluse dans le modèle, la puissance des trois tests tend à diminuer.

- Globalement, les tests  $P_m$  et  $Z$  conduisent à de meilleurs résultats, en termes de taille et de puissance, que le test  $L^*$ , ce qui conduit Choi (2002) à recommander les deux premiers tests pour les applications empiriques.
- Enfin, lorsque les résidus  $u_t$  ne suivent pas une loi normale, mais une loi uniforme, il apparaît que les conclusions précédentes restent valides. En conséquence, le fait que les résidus ne suivent pas une loi normale n'a pas d'impact sur les résultats en termes de taille et de puissance des tests.

## 7.4 Le test de Pesaran (2003)

Pesaran (2003) propose un test unique (contrairement à Bai et Ng, 2001) permettant de tenir compte des éventuelles dépendances entre individus. Mais, à la différence des deux approches précédentes (Moon et Perron, 2003 ou Choi, 2002), Pesaran ne teste pas la racine unitaire sur des variables transformées prises en écart aux composantes déterministes. Il choisit au contraire de conserver les séries brutes  $y_{i,t}$  en augmentant le modèle  $DF$  ou  $ADF$  par l'introduction des moyennes individuelles de  $y_{i,t-1}$  et des différences premières  $\Delta y_{i,t}$  : on obtient alors un modèle augmenté de type  $CADF$  (*Cross Sectionally Augmented Dickey-Fuller*). De ce point de vue, revenant à une approche en termes de statistiques à la Dickey-Fuller, le test de Pesaran (2003) se démarque des tests précédemment présentés dans la mesure où les distributions asymptotiques sont non standards.

Si l'on veut faire une analogie avec le problème de l'estimation de la relation de cointégration en séries temporelles en présence de biais d'endogénéité de second ordre, on pourrait opposer, d'une part, les approches à la Moon et Perron comparables aux *Fully Modified*<sup>23</sup> (MCO sur variables transformées) et, d'autre part, l'approche de Pesaran (2003) sensiblement proche de la logique des *Moindres Carrés Dynamiques*, *DOLS*, de Stock et Watson (1993) ; le principe des *DOLS* consistant en effet à conserver les données brutes mais à augmenter le modèle en introduisant des termes différenciés retardés et avancés. Les t-statistiques individuelles obtenues dans le modèle  $CADF$  ont alors une distribution asymptotique indépendante de tout paramètre de nuisance. Toutefois, ces t-statistiques individuelles ne sont indépendantes dans la dimension individuelle que conditionnellement à un mouvement Brownien  $W_f$ . On peut dès lors construire à partir de ces statistiques individuelles soit une statistique moyenne de type IPS (appelée CIPS, pour *Cross-Sectionally Augmented IPS*), soit une statistique de type Maddala et Wu (1999). Mais l'application du théorème central limite à la statistique moyenne ne permet pas d'aboutir à un résultat de normalité asymptotique comme dans IPS (2003) du fait que la propriété d'indépendance des statistiques  $CADF$  individuelles n'est ici que conditionnelle. Il est alors nécessaire d'utiliser des seuils simulés de la loi asymptotique ou de la loi à distance  $T$  finie pour la statistique moyenne. C'est donc une approche qui est relativement simple à mettre en oeuvre. En outre, comme le signale Pesaran (2003), cette approche pourra être étendue très facilement à des tests individuels autres que les tests ADF (de type ERS ou de type Max-ADF de Leybourne, 1995).

Etant donné la popularité d'IPS (1997), Pesaran considère exactement le même modèle et la même structure de test à la différence près qu'il introduit un facteur commun  $\theta_t$  avec une sensibilité hétérogène à la Phillips et Sul (2003a) :

$$\Delta y_{i,t} = \alpha_i + \rho_i y_{i,t-1} + u_{i,t}$$

$$u_{i,t} = \gamma_i \theta_t + \varepsilon_{i,t}$$

où l'effet individuel  $\alpha_i$  est défini par  $\alpha_i = -\rho_i \gamma_i$  avec  $\gamma_i \in \mathbb{R}$ . Le facteur commun  $\theta_t$  *i.i.d.* (0,1) est inobservable. Les hypothèses du test de Pesaran sont identiques à celles d'IPS. Pesaran montre qu'en l'absence d'autocorrélation des  $\varepsilon_{i,t}$ , l'introduction dans le modèle de la moyenne

<sup>23</sup> Voir par exemple Hurlin et N'Diaye (1998).

$\bar{y}_t = (1/N) \sum_{i=1}^N y_{i,t}$  et de sa valeur retardée  $\bar{y}_{t-1}$  est suffisante pour filtrer asymptotiquement les effets de la composante commune inobservable  $\theta_t$  dès lors que  $N$  tend vers l'infini. Ainsi, Pesaran considère un modèle *DF* augmenté dans la dimension inter-individuelle ou modèle CADF :

$$\text{Modèle CADF : } \Delta y_{i,t} = \alpha_i + \rho_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_i \Delta \bar{y}_t + v_{i,t} \quad (75)$$

Pour chaque individu  $i = 1, \dots, N$  on estime ce modèle et l'on construit de façon standard la *t*-statistique associée à l'hypothèse nulle de racine unitaire pour l'individu  $i$ , notée  $t_i(N, T)$ . Pesaran montre que la distribution exacte de  $t_i(N, T)$  sous l'hypothèse nulle de racine unitaire dépend de paramètres de nuisance, mais que l'influence de ces paramètres disparaît dès lors que  $N$  tend vers l'infini et cela que  $T$  soit fixe ou tende vers l'infini. Lorsque  $T$  est fixe, on doit toutefois s'assurer qu'aucun effet ne puisse transiter *via* le niveau de la moyenne individuelle à la date initiale  $\bar{y}_0$  : ceci peut être obtenu en appliquant le test non pas directement à  $y_{i,t}$ , mais à la différence  $y_{i,t} - \bar{y}_0$ . Lorsque  $T$  et  $N$  tendent vers l'infini de façon séquentielle ou jointe, la statistique CADF individuelle  $t_i(N, T)$  converge vers la même distribution asymptotique. Cette distribution est une généralisation de celle de Dickey-Fuller pour le modèle avec constante, dans le sens où en l'absence d'effet temporel on retrouve cette dernière.

A partir des statistiques CADF individuelles  $t_i(N, T)$ , Pesaran propose notamment une statistique moyenne de type IPS appelée CIPS, pour *Cross-Sectionally Augmented IPS*.

$$CIPS(N, T) = \frac{1}{N} \sum_{i=1}^N t_i(N, T) \quad (76)$$

Il s'agit d'une généralisation de la statistique  $t_{\text{bar}NT}$  (équation 27) d'IPS (2003). Pesaran montre que l'on peut réécrire  $CIPS(N, T)$  sous la forme suivante :

$$CIPS(N, T) = \frac{1}{N} \sum_{i=1}^N t_i(N, T) = \frac{1}{N} \sum_{i=1}^N CADF_{i,f} + O_p(1)$$

où  $CADF_{i,f}$  désigne la distribution asymptotique lorsque  $N$  tend vers l'infini des statistiques  $t_i(N, T)$ . L'indice  $f$  indique que ces distributions sont asymptotiquement dépendantes, ce qui exclut l'application d'un théorème central limite dans ce cas. La distribution de  $\overline{CADF} = N^{-1} \sum_{i=1}^N CADF_{i,f}$  lorsque  $N$  tend vers l'infini est non standard et les seuils de cette variable sont tabulés par Pesaran pour différentes tailles  $T$  et  $N$  ainsi que pour différentes valeurs du risque de première espèce. Dans le cas d'un modèle avec constante, pour un risque de première espèce de 5%, les seuils sont reportés dans le tableau 5. On approxime alors la distribution de la statistique moyenne  $CIPS(N, T)$  par celle de  $\overline{CADF}$ . Ainsi, on rejette l'hypothèse nulle de racine unitaire si la réalisation de  $CIPS(N, T)$  est inférieure au seuil de la distribution de  $\overline{CADF}$ .

Pesaran propose en outre une statistique de type CIPS tronquée, notée  $CIPS^*(N, T)$ , construite comme la moyenne de statistiques individuelles tronquées, telle que :

$$CIPS^*(N, T) = \frac{1}{N} \sum_{i=1}^N t_i^*(N, T) \quad (77)$$

où la statistique individuelle  $t_i^*(N, T)$  tronquée est définie par :

$$t_i^*(N, T) = \begin{cases} -K_1 & \text{si } t_i(N, T) \leq -K_1 \\ t_i(N, T) & \text{si } -K_1 < t_i(N, T) < K_2 \\ K_2 & \text{si } t_i(N, T) \geq K_2 \end{cases}$$

Les paramètres de troncature  $K_1$  et  $K_2$  dépendent du modèle retenu. Il s'agit de constantes positives suffisamment importantes pour que  $\Pr[-K_1 < t_i(N, T) < K_2]$  soit suffisamment importante (supérieure à 0,9999). Pour un modèle sans constante, ni tendance  $K_1 = 6.12$  et

TAB. 5 – Seuils Critiques à 5% de la Distribution de  $\overline{CADF}$  et  $\overline{CADF}^*$

$T/N$	10	15	20	30	50	70	100
10	-2.52 (-2.47)	-2.40 (-2.35)	-2.33 (-2.29)	-2.25 (-2.22)	-2.19 (-2.16)	-2.14 (-2.13)	-2.10 (-2.11)
15	-2.37	-2.28	-2.22	-2.17	-2.11	-2.07	-2.04
20	-2.34	-2.26	-2.21	-2.15	-2.11	-2.07	-2.04
30	-2.33	-2.25	-2.20	-2.15	-2.11	-2.07	-2.05
50	-2.33	-2.25	-2.20	-2.16	-2.11	-2.08	-2.06
70	-2.33	-2.25	-2.20	-2.15	-2.12	-2.08	-2.07
100	-2.32	-2.25	-2.20	-2.16	-2.12	-2.08	-2.07

Source : Pesaran (2003), Table 3b. Entre parenthèses figurent les seuils de la statistique tronquée.

$K_2 = 4.16$ , pour un modèle avec constante :  $K_1 = 6.19$  et  $K_2 = 2.61$ , enfin dans un modèle avec dérive  $K_1 = 6.42$  et  $K_2 = 1.70$ <sup>24</sup>. La distribution de  $CIPS^*(N, T)$  est approximée par celle de  $\overline{CADF}^* = N^{-1} \sum_{i=1}^N CADF_{i,f}^*$  où  $CADF_{i,f}^*$  désigne la distribution asymptotique tronquée de la statistique  $t_i(N, T)$ , compte tenu des mêmes seuils  $K_1$  et  $K_2$ . Les seuils critiques de  $\overline{CADF}^*$  et de  $\overline{CADF}$  sont pratiquement confondus dès lors que  $T > 15$ . En dessous de cette taille, les seuils de  $\overline{CADF}^*$  à 5% sont reportés entre parenthèses dans le tableau 5. Pour des échantillons de petite taille ( $T < 10$ ), la taille de la statistique non tronquée est généralement trop importante, c'est pourquoi on doit lui préférer la statistique tronquée.

Enfin, en présence d'autocorrélation des résidus, la même démarche doit être appliquée dans un modèle "doublement" augmenté : augmenté dans la dimension inter-individuelle et augmenté par les termes habituels des spécifications ADF standard. Le modèle CADF devient alors :

$$\text{Modèle CADF : } \Delta y_{i,t} = \alpha_i + \rho_i y_{i,t-1} + c_i \bar{y}_{t-1} + \sum_{j=0}^p d_{i,j} \Delta \bar{y}_{t-j} + \sum_{j=0}^p \delta_{i,j} \Delta y_{i,t-j} + e_{i,t} \quad (78)$$

Les distributions des statistiques moyennes  $CIPS^*(N, T)$  et  $CIPS(N, T)$  sont alors identiques à celles du cas précédent.

Afin d'appréhender les propriétés en termes de taille et de puissance des statistiques proposées, Pesaran (2003) a effectué des simulations de Monte Carlo. Les principaux résultats obtenus peuvent être synthétisés comme suit :

- En l'absence de corrélation sérielle des erreurs, les statistiques  $CIPS$  et  $CIPS^*$  ne présentent pas de distorsion de taille et conduisent à des résultats similaires en termes de puissance. Ainsi, lorsque les erreurs ne sont pas autocorrélées, il n'existe pas de différence majeure entre les deux statistiques.
- En présence de corrélation sérielle des erreurs, il ressort d'importantes distorsions de taille lorsque les régressions  $CADF$  ne tiennent pas compte de la dépendance entre les séries. En revanche, l'utilisation du modèle doublement augmenté conduit à une taille de l'ordre de 5% et les tests basés sur ce modèle ont donc de bonnes propriétés en termes de taille. Pour des valeurs de  $T$  relativement faibles (inférieures à 20), la statistique  $CIPS$  présente toutefois des distorsions de taille plus importantes que la statistique tronquée qui ne souffre pas d'un tel problème, même pour  $T = 10$ . La puissance de la statistique  $CIPS^*$  croît rapidement dès lors que  $T \geq 20$  et est plus importante en cas de corrélation sérielle négative qu'en cas de corrélation sérielle positive.

<sup>24</sup>Ces valeurs sont obtenues par Pesaran en retenant une approximation normale pour  $t_i(N, T)$ , conduisant à :  $K_1 = -E(CADF_{i,f}) - \Phi^{-1}(\varepsilon/2) \sqrt{Var(CADF_{i,f})}$  et  $K_2 = E(CADF_{i,f}) + \Phi^{-1}(1 - \varepsilon/2) \sqrt{Var(CADF_{i,f})}$ ,  $\varepsilon$  désignant une constante positive proche de zéro.

- Enfin, en présence de corrélation sérielle des erreurs et de tendances linéaires, la taille du test *CIPS\** reste tout à fait correcte (y compris pour  $T = 10$ ), contrairement à celle de la statistique non tronquée.

Ces résultats mettent donc globalement en avant la supériorité de la statistique *CIPS\** par rapport à la statistique *CIPS*, dès lors que les termes d'erreurs présentent de l'autocorrélation.

## 7.5 Le test de Chang (2002)

Comme nous l'avons mentionné précédemment, il existe une approche alternative dans les tests de seconde génération qui n'est pas fondée sur un modèle factoriel. Les tests les plus représentatifs sont ici ceux de Chang (2002, 2004). Nous nous contenterons de présenter le test de 2002, fondé sur un estimateur des variables instrumentales non linéaires qui permet de résoudre le problème des paramètres de nuisance lié aux corrélations inter-individuelles. Considérons le modèle suivant :

$$\Delta y_{i,t} = \alpha_i + \rho_i y_{i,t-1} + \sum_{j=1}^{p_i} \beta_{i,j} \Delta y_{i,t-j} + \varepsilon_{i,t} \quad (79)$$

où les innovations  $\varepsilon_{i,t}$  sont *i.i.d.* ( $0, \sigma_{\varepsilon_i}^2$ ) et peuvent être corrélées dans la dimension individuelle. Pour éliminer ces dépendances, Chang propose d'utiliser pour chaque individu  $i$ , une variable instrumentale générée par une fonction non linéaire  $F(y_{i,t-1})$  des valeurs passées  $y_{i,t-1}$ . Cette fonction, appelée fonction génératrice d'instrument (*Instrument Generating Function*), est une fonction régulière, intégrable telle  $\int_{-\infty}^{\infty} xF(x) dx \neq 0$ . Cette hypothèse peut s'interpréter comme le fait que l'instrument non linéaire  $F(\cdot)$  doit être corrélé avec le régresseur  $y_{i,t-1}$ .

On note  $x_{it}$  le vecteur des différences premières centrées passées  $(\Delta y_{i,t-1}, \dots, \Delta y_{i,t-p_i})'$  et  $X_i = (x_{i,p_i+1}, \dots, x_{iT})'$  la matrice  $(T, p_i)$  correspondante. Soit  $y_{i,i} = (y_{i,p_i}, \dots, y_{i,T-1})'$  le vecteur des valeurs passées et  $\varepsilon_i = (\varepsilon_{i,p_i+1}, \dots, \varepsilon_{i,T})'$  le vecteur des résidus. Sous l'hypothèse nulle de racine unitaire, l'estimateur des variables instrumentales non linéaires du paramètre  $\rho$  est alors défini par :

$$\hat{\rho}_i = \left[ F(y_{i,i})' y_{i,i} - F(y_{i,i})' X_i (X_i' X_i)^{-1} X_i' y_{i,i} \right]^{-1} \left[ F(y_{i,i})' \varepsilon_i - F(y_{i,i})' X_i (X_i' X_i)^{-1} X_i' \varepsilon_i \right] \quad (80)$$

La variance de cet estimateur est égale à :

$$\hat{\sigma}_{\hat{\rho}_i}^2 = \hat{\sigma}_{\varepsilon_i}^2 \left[ F(y_{i,i})' y_{i,i} - F(y_{i,i})' X_i (X_i' X_i)^{-1} X_i' y_{i,i} \right]^{-2} \left[ F(y_{i,i})' F(y_{i,i}) - F(y_{i,i})' X_i (X_i' X_i)^{-1} X_i' F(y_{i,i}) \right] \quad (81)$$

où  $\hat{\sigma}_{\varepsilon_i}^2 = (1/T) \sum_{t=1}^T \hat{\varepsilon}_{it}^2$ . Chang montre que la  $t$ -statistique utilisée pour tester la racine unitaire, notée  $Z_i$ , et construite à partir de  $\hat{\rho}_i$ , converge asymptotiquement vers une loi normale centrée réduite :

$$Z_i = \frac{\hat{\rho}_i}{\hat{\sigma}_{\hat{\rho}_i}} \xrightarrow[T \rightarrow \infty]{l} N(0, 1) \quad \text{pour } i = 1, \dots, N$$

Ce résultat asymptotique inhabituel est entièrement dû à la non linéarité des variables instrumentales. Chang fournit différents exemples de fonctions pouvant générer de tels instruments. Citons trois exemples :  $IGF_1(x) = x \exp(-c_i |x|)$  avec  $c_i = 3T^{-1/2} s^{-1} (\Delta y_{it})$  où  $s^2 (\Delta y_{it})$  désigne la variance empirique de  $\Delta y_{it}$ ,  $IGF_2(x) = \mathbb{I}(|x| < K)$  et  $IGF_3(x) = \mathbb{I}(|x| < K) * x$  où  $K$  désigne un paramètre de troncature. De plus, Chang montre que les  $t$ -statistiques  $Z_i$

sont asymptotiquement indépendantes dans la dimension individuelle, ce qui bien évidemment autorise l'utilisation de statistiques moyennes :

$$S_N = \frac{1}{\sqrt{N}} \sum_{i=1}^N Z_i \quad (82)$$

Cette statistique est elle aussi asymptotiquement normalement distribuée. L'approche de Chang (2002) est *a priori* extrêmement séduisante puisqu'elle ne fait aucune hypothèse sur la nature des dépendances inter-individuelles, contrairement aux modèles factoriels utilisés dans les autres tests. Il est toutefois important de signaler qu'Im et Pesaran (2003), en utilisant un modèle à facteurs communs, mettent en évidence de très fortes distorsions de taille pour ce test, y compris lorsque la dimension  $N$  est relativement faible par rapport à la dimension  $T$ .

## 8 Application

Considérons à présent une application de ces différents tests portant sur le PNB réel par tête de 25 pays de l'OCDE. Le panel est cylindré<sup>25</sup> sur la période 1965-2000 puisque certains tests étudiés ne peuvent pas être implémentés lorsque les dimensions temporelles varient avec les individus.

On constate tout d'abord que les trois tests de première génération conduisent à un diagnostic mitigé quant à la présence d'une racine unitaire dans le PNB réel par tête (tableau 6). Le test de Levin, Lin et Chu (2002) conduit au rejet de la non stationnarité quelle que soit l'hypothèse formulée sur la composante déterministe (modèle avec effets individuels ou modèle avec effets individuels et tendances déterministes). Dans cet exemple, la statistique de Levin et Lin est fondée sur une estimation de la variance de long terme des résidus utilisant une fonction noyau de type Bartlett et un paramètre de troncature commun pour tous les pays, déterminé par la valeur  $q = 3.21T^{1/3}$  (Levin, Lin et Chu, 2002)<sup>26</sup>. Mais ce résultat est robuste à la modification de la fonction noyau ainsi qu'au choix de paramètres individuels par la méthode de Newey et West (1994). Les retards individuels  $p_i$  ont été déterminés par la méthode du général au spécifique d'Hall (1994) et les résultats s'avèrent robustes au choix d'une autre méthode de sélection des retards. Le test de Levin et Lin conduit donc à un résultat assez contre-intuitif selon lequel le PIB par tête serait stationnaire pour les pays de l'OCDE de notre panel.

Les résultats sont légèrement différents lorsque l'on prend en compte la dimension hétérogène de la racine autorégressive avec le test d'Im, Pesaran et Shin (2003). A partir des mêmes retards individuels que dans le test de Levin et Lin, la statistique de test  $W_{tbar}$  conduit à accepter l'hypothèse nulle de racine unitaire dans un modèle avec effets individuels. Toutefois, ce résultat n'est pas robuste à l'inclusion de tendances déterministes. En outre, les tests de Maddala et Wu (1999) et Choi (2001) — basés eux aussi dans cet exemple sur les statistiques de tests ADF individuelles — conduisent au rejet de la non stationnarité, y compris dans un modèle sans tendance. Rappelons toutefois que, pour ces tests, le rejet de l'hypothèse nulle n'implique pas

<sup>25</sup>Les données sont tirées des *World Development Indicators*, World Bank (Code : NY.GDP.PCAP.KD) et correspondent au PNB par tête exprimé en dollars constants 1995, base 100 en 1995. Six pays de l'OCDE ont été exclus de l'échantillon : la Turquie, l'Allemagne, la République Tchèque, la Pologne et la République Slovaque. Les programmes de ces tests ont été réalisés sous Matlab.

<sup>26</sup>L'estimateur  $\hat{\Omega}_i^2$  de la variance de long terme  $\Omega_i^2$  dans le modèle sans constante est défini par :  $\hat{\Omega}_i^2 = \frac{1}{T-1} \sum_{t=2}^T \Delta y_{i,t}^2 + 2 \sum_{j=1}^{q_i} w(q_i, j) \times \left( \frac{1}{T-1} \sum_{t=2+j}^T \Delta y_{i,t} \Delta y_{i,t-j} \right)$   $i = 1, \dots, N$  où  $w(q_i, j)$  désigne une fonction noyau et  $q_i$  un paramètre de troncature. Levin et Lin utilisent une fonction noyau de type Bartlett et, dans leurs simulations de Monte Carlo, ils considèrent une taille de fenêtre identique pour tous les individus  $q_i = q$ , telle que  $q$  soit défini comme l'entier le plus proche de la quantité  $q = 3.21 \tilde{T}^{1/3}$  avec  $\tilde{T} = (T - \bar{p} - 1)$  où  $\bar{p} = (1/N) \sum_{i=1}^N p_i$  désigne la moyenne des retards sur les  $N$  individus. Dans le cas du modèle avec constante, l'estimateur est défini en remplaçant les variations  $\Delta y_{i,t}$  par les variations centrées  $\Delta y_{i,t} - \overline{\Delta y}_i$ . Dans le cas du modèle 3, il convient de remplacer  $\Delta y_{i,t}$  par les variations prises en écart à une tendance déterministe  $\Delta y_{i,t} - a_i - b_i t$  où les paramètres  $a_i$  et  $b_i$  sont estimés par MCO sur données individuelles.

la stationnarité des PNB par tête des 25 pays de l'échantillon, mais signifie qu'il existe au moins un pays pour lequel il n'y a pas de racine unitaire dans la dynamique du PNB. On obtient ainsi globalement des résultats très mitigés, y compris avec le test d'Im, Pesaran et Shin (2003).

TAB. 6 – Résultats des Tests de Première Génération

	Statistique	Modèle sans tendance	Modèle avec tendance
Levin, Lin et Chu (2002)	$t_{\rho}^*$	-6.739 (0.00)	-4.753 (0.00)
Im, Pesaran et Shin (2003)	$Z_{tbar}$	0.594 (0.72)	-6.020 (0.00)
	$W_{tbar}$	0.545 (0.70)	-2.519 (0.00)
Maddala et Wu (1999)	$P_{MW}$	68.01 (0.05)	85.29 (0.00)
Choi (2001)	$Z_{MW}$	1.80 (0.04)	3.529 (0.00)

Notes : Les p-values associées aux différentes statistiques figurent entre parenthèses

Une des raisons qui pourrait expliquer ces résultats tient à la présence d'éventuelles dépendances entre les PNB réels par tête des pays de la zone OCDE. La procédure de test<sup>27</sup> du nombre de facteurs communs proposée par Bai et Ng (2002) nous conduit ainsi à retenir l'existence d'un facteur commun dans le PNB ( $r = 1$ ). A partir des estimations  $\hat{e}_{it}$  des composantes idiosyncratiques obtenues par l'approche de Bai et Ng (2004), les tests sur ces composantes conduisent au rejet de l'hypothèse nulle de racine unitaire. Ainsi, la réalisation de la statistique à la Choi (2001), notée  $Z^c$ , construite à partir des composantes idiosyncratiques estimées, est égale à 3.461. Toutefois, ce test ne conduit pas au rejet global de la non stationnarité des PNB réels par tête puisque la composante de facteur commun présente elle au contraire une propriété de racine unitaire. En effet, la réalisation de la statistique de test ADF construite à partir du facteur commun  $\hat{F}$  ( $ADF_{\hat{F}}^c = -1.212$ ) conduit à accepter l'hypothèse nulle de racine unitaire. Ainsi, dans la perspective de Bai et Ng (2004), le PNB réel par tête admet une racine unitaire, mais celle-ci provient uniquement de la composante de facteur commun qui peut s'assimiler dans notre exemple à un facteur de croissance mondiale. Les composantes d'écart à ce facteur commun sont quant à elle stationnaires.

Ces résultats sont globalement confirmés par les autres tests de racine unitaire de seconde génération (tableau 7). Le test de Moon et Perron, défini uniquement sur la composante d'écart au facteur commun, conduit à rejeter l'hypothèse nulle de racine unitaire. Mais, contrairement au cas du test de Bai et Ng, on rejette alors la racine unitaire pour le PNB réel par tête dans cet échantillon. Ce résultat n'est toutefois pas robuste à la prise en compte de tendances déterministes dans le modèle. Dans ce cas, les statistiques modifiées  $t_a^{\#}$  et  $t_b^{\#}$  (Moon et Perron, 2004) conduisent à accepter l'hypothèse nulle. Pour tous les autres tests, l'hypothèse nulle de racine unitaire n'est pas rejetée et ce de façon robuste au choix des statistiques et de la composante déterministe du modèle. Les trois statistiques de Choi (2002), la statistique CIPS de Pesaran (2003) calculée pour trois ordres de retard différents (CIPS<sub>p</sub> avec  $p = 1, 2$  et  $3^{28}$ ) ainsi que les statistiques de Chang calculées pour trois fonctions non linéaires instrumentales, permettent de conclure à la présence d'une racine unitaire dans la dynamique du PNB réel par tête des pays de l'OCDE. Il est ainsi très intéressant de constater à quel point la prise en

<sup>27</sup>Dans cette étude, nous retenons le critère d'information  $IC_2$ .

<sup>28</sup>L'utilisation des statistiques tronquées  $CIPS^*$  donne des résultats similaires.

compte des éventuelles dépendances inter-individuelles peut modifier la nature du diagnostic quant à la non stationnarité des séries macro-économiques.

TAB. 7 – Résultats des Tests de Deuxième Génération

	Modèle sans tendance			Modèle avec tendance		
Moon et Perron (2004)	$\hat{r}$	$t_a$	$t_b$	$\hat{r}$	$t_a^\#$	$t_b^\#$
	1	-12.26 (0.00)	-5.910 (0.00)	1	-0.310 (0.37)	-0.293 (0.38)
Choi (2002)	$P_m$	$Z$	$L^*$	$P_m$	$Z$	$L^*$
	-3.508 (0.99)	4.323 (1.00)	4.137 (1.00)	1.558 (0.06)	-0.678 (0.24)	-0.523 (0.30)
Pesaran (2003)	$CIPS_1$	$CIPS_2$	$CIPS_3$	$CIPS_1$	$CIPS_2$	$CIPS_3$
	-1.848 (0.37)	-1.823 (0.41)	-1.911 (0.29)	-2.465 (0.26)	-2.375 (0.40)	-2.196 (0.69)
Chang (2002)	$IGF_1$	$IGF_2$	$IGF_3$	$IGF_1$	$IGF_2$	$IGF_3$
	12.88 (1.00)	16.86 (1.00)	14.52 (1.00)	6.411 (1.00)	1.389 (0.91)	6.911 (1.00)

Notes : Les p-values associées aux différentes statistiques figurent entre parenthèses

## 9 Conclusion et extensions : vers une troisième génération de tests

Dans ce travail, nous nous sommes attachés à dresser une revue de la littérature des principaux tests de racine unitaire en panel. Nous avons mis en évidence une double évolution de ces tests depuis les travaux fondateurs de Levin et Lin (1992) : une évolution vers des modélisations hétérogènes avec les travaux d’Im, Pesaran et Shin (1997) et de Maddala et Wu (1999) et, plus récemment, une évolution vers une prise en compte des dépendances inter-individuelles. Cette dernière introduit une dichotomie entre deux générations de tests. La seconde génération est aujourd’hui en pleine construction étant donné la diversité des formes possibles de corrélations inter-individuelles.

A côté des deux générations de tests présentés ici, il convient de mentionner qu’une troisième catégorie de tests se développe depuis le début de la décennie actuelle : les tests de racine unitaire en panel incluant la possibilité de ruptures structurelles. Depuis les travaux pionniers de Perron (1989), il est en effet bien connu que le fait de ne pas tenir compte des ruptures structurelles — lorsqu’elles existent — peut engendrer un biais en faveur du non rejet de l’hypothèse de racine unitaire conduisant à une perte de puissance significative de ces tests. Un tel problème, initialement mis en évidence dans le cas des tests sur séries temporelles, se pose également dans le cas des tests en panel. En effet, les tests de racine unitaire en panel, reposant sur une combinaison linéaire des statistiques de tests sur séries individuelles, souffrent ainsi du même problème de biais et de perte de puissance si un changement structurel est présent.

L’un des travaux fondateurs dans le domaine est le test développé par Im, Lee et Tieslau (2002), reprenant les développements de Im et Lee (2001). Ce test, constituant une extension du test de Schmidt et Phillips (1992) et Amsler et Lee (1995) basé sur le principe du multiplicateur de Lagrange, est relativement flexible dans la mesure où les ruptures structurelles n’ont pas nécessairement lieu à la même date pour tous les individus et le nombre de ruptures peut être différent pour chaque individu du panel. Ce test repose toutefois sur l’hypothèse restrictive d’absence de corrélation inter-individuelle des termes d’erreur, ce qui a conduit au développe-

ment d'autres types de tests, en particulier par Carrion *et al.* (2001, 2002). Le test proposé par Carrion *et al.* (2001) constitue une généralisation du test de Harris et Tzavalis (1999) et vise à tenir compte d'un changement structurel intervenant dans chacune des séries individuelles à la même date. Prolongeant ces travaux, Carrion *et al.* (2002) développent un test de l'hypothèse nulle de stationnarité tenant compte de l'existence possible de ruptures multiples et constituant une extension du test de Hadri (2000). Ce dernier test est assez général au sens où il considère *(i)* des ruptures structurelles multiples, *(ii)* des ruptures structurelles qui peuvent être situées à des dates différentes pour chacune des séries et *(iii)* un nombre différent de ruptures structurelles pour chaque individu du panel.

## Références

- [1] AMEMIYA, T. (1985), *Advanced Econometrics*, Harvard University Press.
- [2] AMSLER, C. ET LEE, J. (1995), “An LM Test for a Unit Root in the Presence of a Structural Change”, *Econometric Theory*, 11, 359-368.
- [3] ANDREWS, D.W.K. (1991), “Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation”, *Econometrica*, 59, 817-858.
- [4] BAI, J. ET NG, S. (2001), “A PANIC Attack on Unit Roots and Cointegration”, Boston College, Department of Economics, Unpublished Manuscript.
- [5] BAI, J. ET NG, S. (2002), “Determining the Number of Factors in Approximate Factor Models”, *Econometrica*, 70(1), 191-221.
- [6] BAI, J. ET NG, S. (2004), “A PANIC Attack on Unit Roots and Cointegration”, *Econometrica*, 72(4), 1127-1178.
- [7] BALTAGI, B.H. ET KAO, C. (2000), “Nonstationary Panels, Cointegration in Panels and Dynamic Panels : a Survey”, *Advances in Econometrics*, vol 15, Elsevier Science, 7-51.
- [8] BANERJEE, A. (1999), “Panel Data Unit Root and Cointegration : an Overview”, *Oxford Bulletin of Economics and Statistics*, Special Issue, 607-629.
- [9] 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.
- [10] BANERJEE, A., MARCELLINO, M. ET OSBAT, C. (2003), “Testing for PPP : Should we use Panel Methods?”, *Mimeo*.
- [11] BANERJEE, A. ET ZANGHERI, P. (2003), “A New Look at the Feldstein-Horioka Puzzle Using an Integrated Panel”, *Working Paper* CEPII, 2003-22.
- [12] CARRION-I-SILVESTRE, J.L., DEL BARRIO-CASTRO, T. ET LOPEZ-BAZO, E. (2001), “Level Shifts in a Panel Data Based Unit Root Test. An Application to the Rate of Unemployment”, *Technical Report*, Département d'économétrie, Université de Barcelone.
- [13] CARRION-I-SILVESTRE, J.L., DEL BARRIO-CASTRO, T. ET LOPEZ-BAZO, E. (2002), “Breaking the Panels. An Application to the GDP per Capita”, *Document de travail*, Département d'économétrie, Université de Barcelone.
- [14] CHANG, Y. (2002), “Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency”, *Journal of Econometrics*, 110, 261-292.
- [15] CHANG, Y. (2004), “Bootstrap Unit Root Tests in Panels with Cross-Sectional Dependency”, *Journal of Econometrics*, 120, 263-293.
- [16] CHOI, I. (2001), “Unit Root Tests for Panel Data”, *Journal of International Money and Finance*, 20, 249-272.
- [17] CHOI, I. (2002), “Combination Unit Root Tests for Cross-Sectionally Correlated Panels”, *Mimeo*, Hong Kong University of Science and Technology.
- [18] DICKEY, D.A. ET FULLER, W.A. (1979), “Distribution of the Estimators for Autoregressive Time Series with a Unit Root”, *Journal of the American Statistical Association*, 74, 417-31.
- [19] DICKEY, D.A. ET FULLER, W.A. (1981), “Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root”, *Econometrica*, 49, 1057-1072.
- [20] ELLIOTT, G., ROTHENBERG, T. ET STOCK, J. (1996), “Efficient Tests for an Autoregressive Unit Root”, *Econometrica*, 64, 813-836
- [21] EVANS, P. ET KARRAS, G. (1996), “Convergence Revisited”, *Journal of Monetary Economics*, 37, 249-265.

- [22] FISHER, R.A. (1932), *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh.
- [23] GAULIER, G., HURLIN, C. ET JEAN-PIERRE, P. (1999), "Testing Convergence : A Panel Data Approach", *Annales d'Economie et de Statistique*, 55-56, 411-428.
- [24] HADRI, K. (2000), "Testing for Unit Roots in Heterogeneous Panel Data", *Econometrics Journal*, 3, 148-161.
- [25] HALL, A. (1994), "Testing for a Unit Root in Time Series with Pretest Data-Based Model Selection", *Journal of Business and Economic Statistics*, 12, 461-70.
- [26] HAMILTON, J.D. (1994), *Times Series Analysis*, Princeton.
- [27] HARRIS, R.D.F. ET TZAVALIS, E. (1999), "Inference for Unit Roots in Dynamic Panels where the Time Dimension is Fixed", *Journal of Econometrics*, 91, 201-226.
- [28] HARVEY, A. ET BATES, D. (2003), "Multivariate Unit Root Tests and Testing for Convergence", Faculty of Economics and Politics, University of Cambridge, *Unpublished Manuscript*.
- [29] HENIN, P-Y., JOLIVALDT, P. ET NGUYEN, A. (2001), "Testing for Unit Root on Heterogeneous Panels : A Sequential Approach", *Couverture Orange CEPREMAP*, 2001-08
- [30] HURLIN, C. (2001), *Econométrie des Données de Panel. Modèles Linéaires Simples*, Ecole Doctorale Edocif, Université Paris IX Dauphine.
- [31] HURLIN, C. ET N'DIAYE (1998), "La Méthode d'Estimation des Fully Modified ou Moindres Carrés Modifiés : Avantages et Limites.", *Cahiers Eco & Maths 98.26*, Université de Paris I.
- [32] HSIAO, C. (1986), "Analysis of Panel Data", *Econometric society Monographs NO 11*. Cambridge University Press.
- [33] IM, K.S. ET LEE, J. (2001), "Panel LM Unit Root Tests with Level Shifts", *Manuscript*, University of Central Florida.
- [34] IM, K.S., LEE, J. ET TIESLAU, M. (2002), "Panel LM Unit Root Tests with Level Shifts", *Manuscript*, University of Central Florida.
- [35] IM, K.S. ET PESARAN, M.H. (2003), "On the Panel Unit Root Tests Using Nonlinear Instrumental Variables", *Manuscript*, University of Southern California.
- [36] IM, K.S., PESARAN, M.H. ET SHIN, Y. (1997), "Testing for Unit Roots in Heterogeneous Panels", *DAE, Working Paper 9526*, University of Cambridge.
- [37] IM, K.S., PESARAN, M.H. ET SHIN, Y. (2002), "Testing for Unit Roots in Heterogeneous Panels", revised version of the *DAE, Working Paper 9526*, University of Cambridge.
- [38] IM, K.S., PESARAN, M.H. ET SHIN, Y. (2003), "Testing for Unit Roots in Heterogeneous Panels", *Journal of Econometrics*, 115, 1, 53-74.
- [39] JOHANSEN, S. (1988), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 12, 231-254.
- [40] KWIATKOWSKI, D., PHILLIPS, P.C.B., SCHMIDT, P. ET SHIN, Y. (1992), "Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root", *Journal of Econometrics*, 54, 91-115.
- [41] LEVIN, A. ET LIN, C.F. (1992), "Unit Root Test in Panel Data : Asymptotic and Finite Sample Properties", *University of California at San Diego*, Discussion Paper 92-93.
- [42] LEVIN, A. ET LIN, C.F. (1993), "Unit Root Test in Panel Data : New Results", *University of California at San Diego*, Discussion Paper 93-56.
- [43] LEVIN, A., LIN, C.F., ET CHU, C.S.J. (2002), "Unit Root Test in Panel Data : Asymptotic and Finite Sample Properties", *Journal of Econometrics*, 108, 1-24.

- [44] LEYBOURNE, S.J. (1995), “Testing for Unit Roots Using Forward and Reverse Dickey-Fuller Regressions”, *Oxford Bulletin of Economics and Statistics*, 57, 559-571.
- [45] LOPEZ, J.H. (1997), “The Power of the ADF Test”, *Economics Letters*, 57, 5-10.
- [46] MACKINNON, J.G. (1994), “Approximate Asymptotic Distribution Functions for Unit-Root and Cointegration”, *Journal of Business and Economic Statistics*, 12, 167-176.
- [47] MADDALA, G.S. ET WU, S. (1999), “A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test”, *Oxford Bulletin of Economics and Statistics*, special issue, 631-652.
- [48] MAGNUS, J.R. (1986), “The Exact Moments of a Ratio of Quadratic Forms in Normal Variables”, *Annales d'Economie et de Statistique*, 4, 96-109.
- [49] MOON, H. R. ET PERRON, B. (2004), “Testing for a Unit Root in Panels with Dynamic Factors”, *Journal of Econometrics*, 122, 81-126.
- [50] MOON, H. R., PERRON, B. ET PHILLIPS, P.C.B (2003), “Incidental Trends and the Power of Panel Unit Root Tests”, *Mimeo*, Université de Montréal.
- [51] NABEYA, S., (1999), “Asymptotic Moments of some Unit Root Test Statistics in the Null Case”, *Econometric Theory*, 15, 139-149.
- [52] NEWEY, W. ET WEST, K. (1994), “Automatic Lag Selection and Covariance Matrix Estimation”, *Review of Economic Studies*, 61, 631-653.
- [53] NG, S. ET PERRON, P. (1995), “Unit Root Tests in ARMA Models with Data Dependent Methods for the Selection of the Truncation Lag”, *Journal of the American Statistical Association*, 90, 268-281.
- [54] O'CONNELL, P. (1998), “The Overvaluation of Purchasing Power Parity”, *Journal of International Economics*, 44, 1-19.
- [55] PERRON, P. (1989), “The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis”, *Econometrica*, 57, 1361-1401.
- [56] PESARAN, H.M. (2003), “A Simple Panel Unit Root Test in the Presence of Cross Section Dependence”, *Mimeo*, University of Southern California.
- [57] PESARAN, H.M. ET SMITH, R (1995), “Estimating Long-Run Relationships from Dynamic Heterogeneous Panels”, *Journal of Econometrics*, 68, 79-113.
- [58] PHILLIPS, P.C.B. ET MOON, H. (1999), “Linear Regression Limit Theory for Nonstationary Panel Data”, *Econometrica*, 67, 1057-1111.
- [59] PHILLIPS, P.C.B. ET PERRON, P. (1988), “Testing for a Unit Root in Time Series Regression”, *Biometrika*, 75, 335-346.
- [60] PHILLIPS, P.C.B. ET SUL, D. (2003a), “Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence”, *Econometrics Journal*, 6(1), 217-259.
- [61] PHILLIPS, P.C.B. ET SUL, D. (2003b), “The Elusive Empirical Shadow of Growth Convergence”, *Cowles Foundation Discussion Paper 1398*, Yale University.
- [62] PIERCE, R.G. ET SNELL, A.J. (1995), “Temporal Aggregation and the Power of Tests for a Unit Root”, *Journal of Econometrics*, 65, 333-345.
- [63] PLOBERGER, W. ET PHILLIPS, P.C.B. (2002), “Optimal Testing for Unit Roots in Panel Data”, *Mimeo*.
- [64] QUAH, D. (1994), “Exploiting Cross-Section Variations for Unit Root Inference in Dynamic Data”, *Economics Letters*, 44, 9-19.
- [65] SALANIE, B. (1999), “Guide Pratique des Séries Non Stationnaires”, *Economie et Prévision*, 137, 119-141.
- [66] SCHMIDT, P. ET LEE, J. (1991), “A Modification of Schmidt-Phillips Unit Root Test”, *Economics Letters*, 36, 285-289.

- [67] SCHMIDT, P. ET PHILLIPS, P.C.B. (1992), "LM Tests for a Unit Root in the Presence of Deterministic Trends", *Oxford Bulletin of Economics and Statistics*, 54, 257-587.
- [68] SEVESTRE, P. (2002), *Econométrie des Données de Panel*, Dunod.
- [69] STOCK, J.H AND WATSON, M.W. (1988), "Testing for Common Trends", *Journal of the American Statistical Association*, 83, 1097-1107.
- [70] STOCK, J.H ET WATSON, M.W. (1993), "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems", *Econometrica*, 61, 783-820.
- [71] STRAUSS, J. ET YIGIT, T. (2003), "Shortfalls of Panel Unit Root Testing", *Economics Letters*, 81, 309-313.
- [72] TAYLOR, M.P. ET SARNO, L. (1998), "The Behavior of Real Exchange Rates During the Post-Bretton Woods Period", *Journal of International Economics*, 46, 281-312.