

Convergence des productivités européennes Transition, rupture et racine unitaire

Stéphane ADJEMIAN *

RÉSUMÉ. – Ce papier réconcilie les conclusions des approches classique et en série temporelles de la convergence. En utilisant une statistique KPSS multivariée et en autorisant une rupture commune, il apparaît que la dynamique des productivités européennes se découpe en deux phases : après une phase de rattrapage les économies évoluent le long d'un sentier de croissance commun. On montre alors que, les écarts de productivités sont stationnaires au sens stochastique, sur l'ensemble de l'échantillon.

European Productivity Convergence, Transition Break and Unit Root

ABSTRACT. – This paper aims to resolve the conflict between the time series approach and the classical approach to test the convergence hypothesis. Using a multivariate KPSS statistic and allowing for one common break, it appears that European productivity dynamics may be split into two sub-periods. During the first, productivity gaps are reducing. During the second, productivity gaps are zero mean processes. Moreover, productivity discrepancies are shown to be stationary over the entire period.

* S. ADJEMIAN: EPEE, Université d'Evry Val d'Essonne, Département d'économie. Je remercie Jérôme GLACHANT et Stéphane GREGOIR, qui ont bien voulu encadrer le mémoire de DEA à l'origine de ce travail. Cette version a bénéficié du soutien financier du CNRS dans le cadre du contrat *Les enjeux économiques de l'innovation* et des remarques de deux rapporteurs anonymes. Les erreurs et omissions subsistantes sont miennes.

1 Introduction

Depuis les années 80, un nombre important d'articles discutent du phénomène de convergence des économies, ainsi que de la meilleure approche pour l'appréhender. L'émergence de cette littérature correspond à un regain d'intérêt pour l'économie de la croissance mais aussi à un changement de perspective important, on s'intéresse dorénavant plus spécifiquement aux comparaisons des performances nationales en terme de croissance (Cf. DURLAUF et QUAH [1998, pages 1 à 5]).

On peut tester l'hypothèse de convergence des économies en recourant à différentes technologies. Les initiateurs de cette problématique militent pour ce que l'on peut nommer, avec SALA-I-MARTIN [1996], l'*approche classique* de la convergence (les tests de β – convergence). On cherche alors à vérifier qu'une économie croît d'autant plus vite qu'elle est initialement défavorisée. Cette approche s'intéresse à la transition des économies vers leurs états stationnaires respectifs. Par ailleurs, au début des années 1990, avec QUAH [1990], une approche alternative s'est développée autour d'études en séries temporelles. Celle-ci associe la convergence des économies à la stabilité des écarts en terme de productivités ou de revenu par tête. On s'intéresse alors aux trajectoires d'états stationnaires, en se demandant si, à long terme, les économies évoluent le long de sentiers parallèles.

Les conclusions obtenues par ces deux approches sont différentes ; l'*approche classique* reproduit des résultats de convergence, au moins conditionnelle, pour des larges coupes d'économies (voir MANKIW, ROMER, WEIL [1992]), mais l'approche par les séries temporelles reporte rarement ces résultats pour des échantillons de tailles équivalentes, à l'exception notable d'EVANS et KARRAS [1996]. Cette incompatibilité ne doit pas étonner, elle s'explique par l'incapacité de l'approche par les séries temporelles à rendre compte de la phase de transition des économies (Cf. QUAH [1993] et BERNARD et DURLAUF [1996]).

Le but de cet article est de montrer *comment on peut réconcilier une approche par les séries temporelles et l'existence d'une longue phase de transition des économies*.

Je propose d'employer l'extension multivariée, proposée par CHOI et AHN [1999], du test de KWIATKOWSKI, PHILLIPS, SCHMIDT et SHIN [1992]. Il s'agit d'un test de stationnarité reposant sur une écriture composée du processus générateur de données. Cette spécification autorise un traitement cohérent des tendances déterministes, contrairement aux technologies de détection de racine unitaire généralement employée dans cette littérature. Pour dix économies européennes, de 1950 à 1990, il apparaît que les écarts de productivité sont stationnaires au sens stochastique et non stationnaires au sens déterministe.

Ce résultat doit-il nécessairement amener à conclure à la divergence des économies ? HOBIN et FRANSES [2000], qui emploient une généralisation équivalente du test de KWIATKOWSKI *et al.* [1992] pour réaliser un exercice de *clustering* sur 112 économies de 1960 à 1989, ne peuvent distinguer la non stationnarité stochastique de la non stationnarité déterministe des écarts de

productivités. En effet, ils testent la stationnarité contre la non stationnarité stochastique et déterministe des écarts de productivités. Pour ces deux auteurs la non stationnarité déterministe conduit nécessairement à conclure à la divergence des économies. Mais l'hétérogénéité des taux de croissance moyens estimés peut parfaitement rendre compte d'un rapprochement des économies, si les écarts de revenus par tête ne sont pas des processus invariants (dans le sens où l'on doit distinguer les taux de croissance moyens des taux de croissance de long terme des économies).

La problématique envisagée ici rejoint le problème de la réconciliation des approches par les propriétés des séries temporelles et les tests de β – convergence. CARLINO et MILLS [1993], qui étudient la convergence des régions américaines, proposent d'introduire une rupture dans la partie déterministe des écarts entre les régions, afin de montrer que les chocs sur les écarts ne persistent pas indéfiniment (Cf. la problématique de PERRON [1989]). Je propose un pas supplémentaire dans cette voie, en testant l'existence d'une date à partir de laquelle les économies évoluent le long de sentiers parallèles ou le long d'un sentier de croissance commun. La dynamique des économies se décompose alors en deux périodes, une première période de transition, se manifestant par une convergence ou une divergence des économies, puis une période finale où les économies évoluent sur des sentiers de croissance parallèles (ou confondus) et vérifient la définition de la convergence en termes de propriétés des séries temporelles. Sur données européennes, il apparaît que l'on ne peut rejeter l'hypothèse de convergence absolue des économies sur la période 1950-1990.

La section 2 donne la définition de la convergence et présente différents résultats obtenus dans la littérature concernée par l'approche en série temporelle de la convergence. La section 3 utilise la version multivariée du test KPSS pour tester la convergence des économies européennes. La section 4 propose un amendement à la définition de la convergence, afin de réconcilier les résultats de la section précédente avec les conclusions de l'*approche classique* de la convergence.

2 Tester la convergence

2.1 Définitions

Dans un modèle de croissance néo-classique, du fait de la décroissance des rendements du capital, chaque économie est attirée par une tendance dont la pente est déterminée par le progrès technique exogène. Cette impulsion technologique commune à toutes les économies, constitue une tendance dirigeant la dynamique mondiale. Par construction, une économie – imaginaire – moyenne capture cette tendance. En considérant des revenus par travailleur en déviation au revenu par travailleur moyen, on élimine la tendance commune et on obtient alors des processus stationnaires.

DÉFINITION : On accepte l'hypothèse de convergence des économies, contre la divergence, si les processus de productivité en déviation à la moyenne (en coupe) de l'échantillon sont stationnaires.

Cette définition de la convergence peut donc correspondre à la prédiction du modèle néo-classique, elle est vérifiée lorsque les économies évoluent le long de sentiers de croissance parallèles et lorsque les chocs affectant les écarts entre les économies ne persistent pas indéfiniment.

En suivant CARLINO et MILLS [1993], on distingue une convergence conditionnelle et une convergence absolue. Cette dernière apparaît lorsque les revenus par tête en déviation à la moyenne spatiale sont stationnaires de moyennes nulles, c'est-à-dire lorsque les économies ne peuvent s'éloigner les unes des autres durablement. La convergence conditionnelle correspond à une stationnarité en niveau des revenus par tête en déviation à la moyenne spatiale, c'est-à-dire lorsque les écarts entre les économies ne peuvent augmenter durablement¹.

Tester la convergence revient à tester la stationnarité stochastique (l'absence de racine unitaire) et la stationnarité déterministe (l'égalité des taux de croissance de long terme des productivités) des écarts de productivités.

2.2 Quelques résultats

Les travaux empiriques abordent cette définition de la convergence par différents biais. On distingue deux « traditions ». La première, avec les travaux de BERNARD et DURLAUF [1995], cherche à identifier des relations de cointégration entre les processus de revenus par tête. En effet, pour une collection de N économies, sous l'hypothèse de convergence, on doit avoir $N - 1$ relations de cointégration de la forme $(1, -1)$ entre, par exemple, chaque économie *follower* et l'économie *leader*. Ainsi, il existe une unique tendance commune dirigeant la dynamique mondiale des revenus par tête. BERNARD et DURLAUF [1995] proposent une application à 15 pays de l'OCDE. Les deux auteurs mettent en évidence l'existence de deux ou trois tendances stochastiques communes. L'hypothèse de convergence n'est donc pas vérifiée, les économies divergent de façons « ordonnées ».

La deuxième approche est plus contrainte. On teste directement la non stationnarité des écarts de revenus par tête, ou, de façon équivalente, la non stationnarité des processus de revenus par tête en déviation à la moyenne des revenus par tête dans la coupe des économies. Si on parvient à rejeter la non stationnarité des écarts de revenu par tête, alors on admet l'existence d'une relation de cointégration $(1, -1)$ entre les économies *follower* et l'économie *leader*. Cette approche repose généralement sur des tests ADF univariés ou des extensions en panel (voir QUAH [1990], EVANS et KARRAS [1996], BERNARD et JONES [1996]). À l'exception d'EVANS et KARRAS [1996], qui obtiennent un résultat de convergence conditionnelle des revenus par travailleur pour 54 économies de 1950 à 1990, et de BERNARD et JONES

1. Cette notion de convergence conditionnelle ne correspond pas, *a priori*, à la définition donnée par SALA-I-MARTIN [1996], puisque rien n'assure que l'hétérogénéité des niveaux de long terme soit liée aux fondamentaux respectifs des économies.

[1996], qui montrent la convergence des Productivités Totales des Facteurs pour 14 économies de l'OCDE de 1970 à 1987, les résultats sont défavorables à la convergence.

Pour expliquer ces conclusions généralement négatives, relativement aux résultats obtenus par l'*approche classique*, BERNARD et DURLAUF [1996] montrent que cette définition exige que les économies soient à leurs états stationnaires respectifs. Cette hypothèse forte ne semble pas acceptable pour l'ensemble des échantillons considérés dans cette littérature. Dès lors, l'incapacité à rejeter la divergence des économies paraît moins étonnante.

2.3 Limites

La limite mise en évidence par *Bernard* et *Durlauf* n'est pas l'unique problème soulevé par l'approche en série temporelle. On peut voir que les stratégies de tests reposant sur une écriture de type ADF des écarts de productivité ne sont pas adaptées pour tester l'hypothèse de convergence des économies. Par exemple, la procédure d'EVANS et KARRAS [1996] repose sur l'écriture suivante de l'évolution de la productivité de l'économie i en écart à la productivité moyenne dans la coupe des économies,

$$(1) \quad \Delta(y_{i,t} - \bar{y}_{\cdot,t}) = \delta_i + \rho_i(y_{i,t-1} - \bar{y}_{\cdot,t-1}) + \sum_{k=1}^q \varphi_{i,k} \Delta(y_{i,t-k} - \bar{y}_{\cdot,t-k}) + u_{i,t}$$

Si pour tout i on a $\rho_i = 0$, alors au moins une économie, parmi N , diverge. Si pour tout i on a $\rho_i < 0$, les économies convergent. Dans une deuxième étape, les deux auteurs discriminent, s'il y a lieu, entre une convergence absolue et une convergence conditionnelle, en testant l'égalité des δ_i pour $i = 1, \dots, N$.

Les deux auteurs proposent de tester $\rho_i = \rho = 0$ dans la dimension panel, en estimant l'équation (1), corrigée de l'hétéroscédasticité, et en supposant que les effets individuels sont différents de zéro (sous la nulle les taux de croissance de long terme sont distincts). Ils montrent que la statistique de student associée à l'estimateur du paramètre ρ est, sous la nulle, asymptotiquement distribuée comme une loi normale centrée réduite.

Le passage à la dimension panel préserve les défauts du test ADF, en particulier la présence de paramètres de nuisances dans la distribution de la statistique de test et le traitement confus des tendances déterministes.

Ce test de racine unitaire discrimine entre une non stationnarité stochastique et déterministe (on note $I(1) + T$) et une stationnarité en niveau ($I(0) + c$) pour tous les processus scalaires. La nulle et l'alternative ne peuvent rendre compte du cas où les processus d'intérêt sont stationnaires au sens stochastique et tendanciels au sens déterministe ($I(0) + T$). Dans le cas univarié (se reporter à STOCK [1994, pages 2783-2784]), si le processus générateur des données est stationnaire autour d'une tendance linéaire (*i.e.*, si $y_t = a + b.t + c.y_{t-1} + \varepsilon_t$ avec $\varepsilon_t \sim iid(0, \sigma^2)$, et $|c| < 1$), alors on montre qu'asymptotiquement l'estimateur de ρ (obtenu en estimant la version univariée de l'équation (1)) converge en probabilité vers zéro. Asymptotiquement,

la présence d'une tendance linéaire amène le test de *Dickey-Fuller* à accepter la présence d'une racine unitaire (à tort), car la statistique de test ne diverge pas. Mais à distance finie, la réponse de ce test de non stationnarité est moins claire. Tout dépend de l'importance de la tendance linéaire dans le processus générateur de données (*ie*, de l'importance du paramètre b par rapport à la *taille* de l'innovation, σ^2). Un simple exercice de simulations montre qu'à distance finie il existe un seuil $\tilde{b}(\sigma^2)$, pour les valeurs possibles de $|b|$, en deçà duquel le test de *Dickey-Fuller* conclut généralement à la stationnarité (à tort puisque le processus est non stationnaire au sens déterministe).

Cette technologie, mais surtout l'usage qui en est fait par les empiristes de la croissance, semble donc peu convenable pour tester la convergence des économies, car elle ne peut dire clairement pourquoi on rejette ou accepte la divergence des économies, et car elle peut amener à conclure à tort à la convergence des économies au sens de la section (2.1). Nous pourrions envisager d'introduire une tendance linéaire dans l'équation (1), mais il est plus simple de privilégier directement des technologies de test ne souffrant pas d'un traitement incohérent de la partie déterministe. L'objet de la prochaine section est de montrer comment on peut utiliser une extension multivariée du test de KWIATKOWSKI *et al.* [1992] pour tester l'hypothèse de convergence des économies.

3 Un test de convergence

3.1 Tester la stationnarité dans un cadre multivarié

La version multivariée du test de KWIATKOWSKI *et al.* [1992], repose sur le processus générateur de données suivant,

$$(2) \quad \begin{cases} Y_t = \xi \cdot t + R_t + \varepsilon_t \\ R_t = R_{t-1} + u_t, R_0 \text{ donné} \end{cases}$$

où Y_t , ε_t , u_t et la marche aléatoire R_t sont des processus vectoriels de dimension $N \times 1$, ξ est un vecteur de dimension $N \times 1$. Le processus $\{u_t\}_{t=1}^T$ est un bruit blanc multivarié, avec pour matrice de variance-covariance la matrice, de dimension $N \times N$, Θ . Le processus multivarié $\{\varepsilon_t\}_{t=1}^T$ est stationnaire au second ordre, avec pour matrice de variance-covariance Σ , définie positive. Pour rendre compte de processus pouvant être caractérisés par de la dépendance temporelle, on suppose que la ; perturbation $\{\varepsilon_t\}_{t=1}^T$ vérifie les conditions énoncées par PHILLIPS et DURLAUF [1986, corollaire 2.2 page 476]. Ces conditions sont liées au degré de dépendance temporelle et permettent d'envisager des cas suffisamment généraux (comme, par exemple, une modélisation VARMA de la perturbation). La matrice de variance covariance de long terme, notée Ω , résume alors les structures de dépendances temporelle et

spatiale. La condition initiale de la marche aléatoire joue le rôle de la constante dans le processus générateur de Y_t .

Sous l'hypothèse nulle, les N processus scalaires sont stationnaires au sens stochastique. Dans ce cas, la marche aléatoire multivariée, R_t , doit être dégénérée, au sens où $\Theta = 0$. Le processus générateur des données s'écrit alors de la façon suivante,

$$Y_t = R_0 + \xi \cdot t + \varepsilon_t$$

Sous l'hypothèse alternative, au moins un des processus scalaire est non stationnaire au sens stochastique. Dans ce cas la matrice Θ est non nulle dans au moins une direction.

Soit e_t le résidu estimé par les moindres carrés ordinaires du modèle sous l'hypothèse nulle, puis $S_t \equiv \sum_{\tau=1}^t \varepsilon_\tau$ la somme partielle vectorielle et \hat{S}_t la somme partielle estimée. La version multivariée de la statistique de test de KPSS est donnée par l'expression suivante,

$$\kappa = \text{trace} \left\{ T^{-2} \cdot \hat{\Omega}(l)^{-1} \cdot \sum_{t=1}^T \hat{S}_t \hat{S}_t' \right\}$$

où $\hat{\Omega}(l) \equiv T^{-1} \sum_{t=1}^T e_t e_t' + T^{-1} \sum_{s=1}^l k(s/l) \sum_{t=s+1}^T (e_t e_{t-s}' + e_{t-s} e_t')$ est un estimateur de la matrice de variance-covariance de long terme, où $k(x)$ définit la fenêtre pondérant la somme des matrices d'autocorrélation et où le paramètre de troncation ou de *bandwidth* (l) est une fonction de T telle que $l = O_p(T^\delta)$, avec $\delta \in]0, 2^{-1}[$.

On généralise simplement les résultats de KPSS pour obtenir le comportement asymptotique de κ sous l'hypothèse nulle, avec différentes formulations de la partie déterministe du processus². Lorsque $R_0 \neq 0$ et $\xi = 0$ la statistique κ adopte le comportement asymptotique suivant,

$$\kappa_0 \implies \text{trace} \left\{ \int_0^1 V(r) V(r)' dr \right\}$$

où $V(r) \equiv W(r) - rW(1)$ est un pont brownien de dimension $N \times 1$, « \implies » désignant la convergence en loi lorsque T tend vers l'infini. Si $R_0 \neq 0$ et $\xi \neq 0$, la distribution asymptotique de κ est définie par,

$$\kappa_1 \implies \text{trace} \left\{ \int_0^1 \tilde{V}(r) \tilde{V}(r)' dr \right\}$$

où $\tilde{V}(r) \equiv W(r) + (2r - 3r^2)W(1) + (6r^2 - 6r) \int_0^1 W(s) ds$ est un pont brownien du second ordre de dimension $N \times 1$.

2. On peut obtenir ces distributions limites en retirant la rupture dans la proposition de la section (4.2). Un rapporteur anonyme a justement remarqué que l'on n'obtient pas les mêmes distributions asymptotiques en utilisant le point (a,iv) du théorème 2 dans CHOI et AHN [1999]. Celui-ci donne la distribution limite de la statistique κ ($SBDH_B$ dans leurs notations) avec une tendance polynomiale d'ordre p . Le résultat de deux auteurs est faux, car ils devraient obtenir conformément à KPSS une fonctionnelle de pont brownien d'ordre $p + 1$. Une note généralisant la distribution limite sous la nulle et sous l'alternative est à disposition des lecteurs intéressés par ce point.

Sous l'hypothèse alternative de non stationnarité stochastique des p premiers processus scalaires, on a $\kappa = O_p(T^{1-\delta})$. Par exemple, si on utilise la fenêtre de *Bartlett* pour estimer la matrice de variance-covariance de long terme, alors on peut montrer, de la même manière que *KWIATKOWSKI et al.* [1992] dans le cas univarié, que la statistique de test κ_1 se comporte asymptotiquement de la façon suivante,

$IT^{-1}.\kappa_1 \implies \text{trace}$

$$\left\{ \left(\int_0^1 W_1^*(r)W_1^*(r)'dr \right)^{-1} \int_0^1 \left(\int_0^r W_1^*(s)ds \right) \left(\int_0^r W_1^*(s)ds \right)' dr \right\}$$

où $W_1^*(s) \equiv W_1(s) + (6s - 4) \int_0^1 W_1(r)dr + (6 - 12s) \int_0^1 r W_1(r)dr$ est un processus de *Wiener* centré et *détrendé* de dimension $p \times 1$. *CHOI et AHN* [1999] montrent que κ_0 diverge à la même vitesse lorsque les $p < N$ processus scalaires sont stationnaires autour de tendances linéaires.

Cette dernière propriété motive le choix de *HOBIN et FRANCES* [2000], dans l'utilisation de la statistique de test κ_0 pour tester la convergence des économies. Dans la mesure de l'erreur de seconde espèce, ce test ne conduit pas à accepter à tort l'hypothèse nulle de convergence au sens de la section (2.1), contrairement aux approches basées sur des représentations ADF. Mais en utilisant cette statistique de test, pour tester la stationnarité en niveau sous l'hypothèse nulle, ils ne peuvent identifier les raisons qui les poussent à rejeter la convergence. Le rejet de l'hypothèse nulle peut être causé par une non stationnarité stochastique et/ou une non stationnarité déterministe (présence de tendances linéaires dans le processus générateur de données). La procédure proposée ici permet de discriminer entre ces deux causes de non stationnarité.

3.2 Une procédure de test de la convergence

L'approche empruntée par *HOBIN et FRANCES* [2000] n'est pas entièrement satisfaisante. En effet, (i) elle ne délivre pas toute l'information disponible, (ii) elle mène à rejeter la convergence des économies alors que l'aspect tendanciel au sens déterministe des écarts de revenus peut résulter d'un rapprochement des économies. Leurs résultats, peu favorables à la convergence, peut s'expliquer par (ii). Pour pallier la propriété (i), on peut utiliser la statistique de test κ_1 . Celle-ci permet d'identifier l'origine de la divergence de la statistique de test κ_0 . On peut alors décrire une procédure de test de la convergence en trois étapes³ :

Étape (1) Tester la stationnarité stochastique des processus de productivité en déviation à la moyenne des productivités dans la coupe des économies, à l'aide de la statistique de test κ_1 . Si on rejette la stationnarité alors les économies divergent, car les chocs sur les écarts de productivité ont un effet permanent.

3. Celles-ci sont appliquées au même processus stochastique multivarié, $\{Z_t\}_{t=1}^T = \{(y_{i,t} - \bar{y}_{..t})_{i=1}^N\}_{t=1}^T$.

Étape (2) Si ces processus sont stationnaires au sens stochastique, alors on teste la stationnarité déterministe, c'est-à-dire la nullité du vecteur ξ dans l'équation (2). Si on ne rejette pas la stationnarité déterministe, alors les économies convergent au sens de la définition données dans la section (2.1). Dans le cas contraire, la définition donnée plus haut n'est pas vérifiée, mais cela peut être causé par un phénomène de rattrapage des économies.

Étape (3) Si on accepte la convergence au sens de la section (2.1), alors on peut discriminer entre une convergence absolue et une convergence conditionnelle en testant la nullité du vecteur R_0 dans l'équation (2).

Notons que la procédure envisagée ici ne repose pas explicitement, comme si nous avons utilisé la technologie ADF, sur un conditionnement négatif du taux de croissance des productivités relatives à leur niveau. Ne pas utiliser explicitement ce conditionnement négatif, n'est pas essentiel puisque la définition de la convergence au sens des séries temporelles (Cf. BERNARD et DURLAUF [1996]) n'emploie pas explicitement ce mécanisme. Cependant, cette représentation des données ne nous permet pas de calculer une vitesse de convergence des économies, facilitant ainsi la comparaison avec d'autres études.

Enfin, notons que cette procédure, contrairement aux approches en panel, permet de prendre en compte la structure de dépendance spatiale – *via* la matrice de variance-covariance – liant les économies.

3.3 Le cas européen

On considère les revenus réels par travailleurs pour dix pays européens⁴ de 1950 à 1990 en données annuelles (la variable RGDPW, dans la base de données de *Summers et Heston PWT 5.6*). Avant d'entamer l'exposé des résultats, il convient de remarquer la singularité du modèle complet. En considérant des variables en déviation à la moyenne de l'échantillon on introduit une singularité, puisque la somme (sur les individus) des productivités en déviation à la moyenne spatiale est nulle à chaque date. Pour tester la stationnarité il faut donc retirer une des économies (les résultats sont indépendants du choix de l'économie exclue).

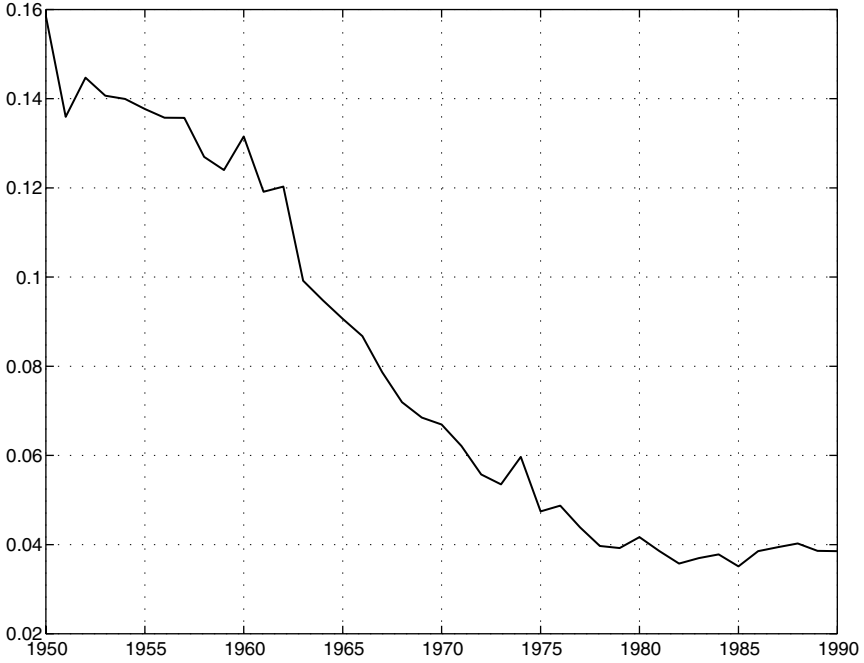
La première étape consiste à régresser, pour chaque économie, la productivité en déviation à la moyenne spatiale sur une constante et une tendance linéaire. En posant le paramètre de troncation $l = 4$ et en employant la fenêtre *Quadratic Spectral* (QS) pour estimer la matrice de variance-covariance de long terme, on obtient $\hat{\kappa}_1 = 0,7574$, à comparer avec la valeur critique au seuil de 5 % : 0,8295 (se reporter au tableau 1 pour les distributions à distance finie). On accepte donc la stationnarité stochastique des dix processus de revenus par tête en déviation à la moyenne de l'échantillon, contre la présence d'une racine unitaire dans un des processus scalaire, au seuil de 5 %⁵.

4. La Belgique, le Danemark, la France, l'Allemagne de l'ouest, la Grèce, l'Irlande, l'Italie, le Luxembourg, les Pays-Bas et la Grande Bretagne.

5. Pour $l = 2, 3$ et la même fenêtre, on accepte la stationnarité stochastique avec un niveau au moins supérieur à 2,5 %. On obtient respectivement $\hat{\kappa}_1 = 0,6994 / 0,7574$ pour les fenêtres de *Bartlett* et *QS*, lorsque le paramètre de *bandwidth* est déterminé par la méthode de NEWBY et WEST [1994] sans *prewhitening* des résidus (afin d'assurer asymptotiquement la divergence de la statistique de test sous l'alternative).

FIGURE 1

Évolution de la variance en coupe des productivités pour dix économies européennes de 1950 à 1990



Pour tester la stationnarité déterministe, c'est-à-dire l'égalité à zéro de l'estimateur de ξ , on applique un test de *Wald* d'égalité des scalaires dans ce vecteur. Puisque les résidus (éventuellement autocorrélés) sont tirés dans une loi inconnue, on utilise la technique du *Bootstrap* séquentiel⁶. On obtient $\hat{W} = 288,108$ comme statistique de *Wald*, et la valeur critique au seuil de 5 % est donnée par 39,973. On rejette donc la stationnarité déterministe contre la non stationnarité déterministe au seuil de 5 %.

Il faut s'interroger sur l'origine de ce résultat, afin de conclure sur la divergence ou la convergence des économies. Selon le test de stationnarité stochastique, les chocs sur les écarts entre les économies n'ont qu'un effet transitoire. Selon le test de *Wald* sur la nullité du vecteur ξ , les économies se déplacent sur des sentiers de croissance non parallèles. Il semble donc qu'une divergence déterministe (une hétérogénéité des taux de croissance moyens), soit à l'œuvre parmi les dix économies européennes. Cette conclusion, de

6. La technique du bootstrap consiste à tirer de façon équiprobable dans les résidus estimés (contrairement à la technique de Monte Carlo qui utilise des perturbations issues d'un générateur de nombres au hasard) afin de former des échantillons sous l'hypothèse nulle et de calculer la distribution de la statistique de test à distance finie. Cette technique n'est pas valide si les perturbations sont caractérisées par de la dépendance temporelle. Dans ce dernier cas, on généralise alors cette technique en tirant plusieurs blocs (dont les tailles sont définies de façon aléatoire) de résidus estimés consécutifs pour former chaque échantillon sous l'hypothèse nulle. Il s'agit du *bootstrap* séquentiel. Pour une description détaillée des généralisations du *bootstrap* aux processus avec dépendance temporelle, on peut se reporter à POLITIS, ROMANO et WOLF [1999].

FIGURE 2

Rattrapage des économies européennes (L'axe des ordonnées donne les coefficients associés à la tendance linéaire dans la régression des écarts de productivités sur un polynôme d'ordre un, l'axe des abscisses donne les conditions initiales en termes de productivités en déviation à la moyenne spatiale)

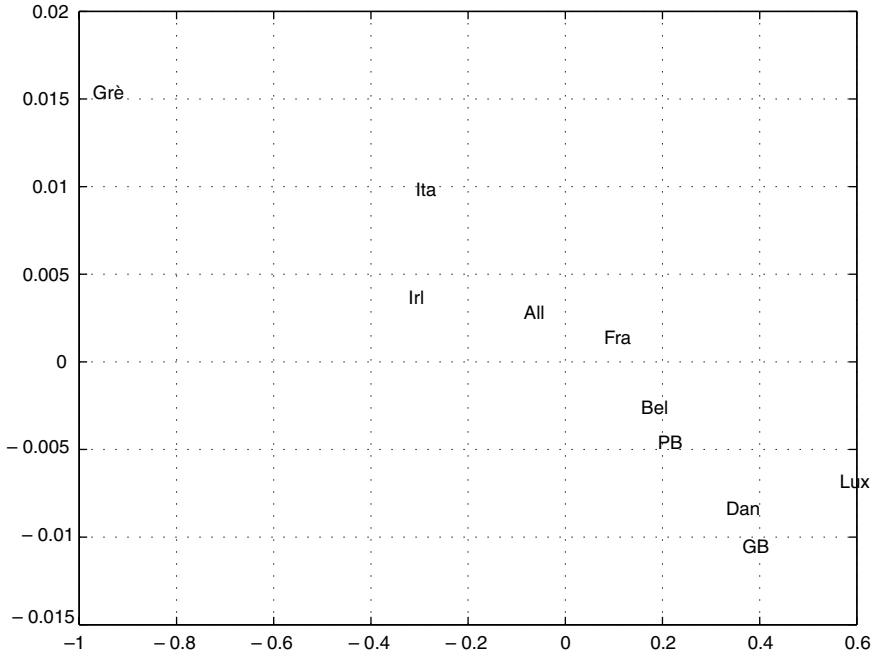


TABLEAU 1

Les quantiles des distributions pour les différents modèles et statistiques associées sont obtenues par Monte Carlo (20000 simulations), avec $T = 41$, $N = 9$ et $T_B = 1979$. L'issue de l'exercice de monte-carlo ne dépend pas des valeurs des paramètres μ_1 , μ_2 , γ_1 et γ_2 (voir, dans l'annexe A, l'expression analytique de la somme partielle \hat{S}_t)

Modèle sous la nulle	0,80	0,90	0,95	0,975	0,99
$I(0) + T + c$	0,6974	0,7644	0,8236	0,8785	0,9446
$M(0)$	0,7919	0,8780	0,9518	1,0221	1,1104
$M(1)$	0,4856	0,5259	0,5608	0,5924	0,6292
$M(2)$	0,4143	0,4520	0,4847	0,5136	0,5518

divergence des économies (au sens de la section (2.1)), est peu satisfaisante ; elle ne résiste pas à la simple observation des graphiques 1 et 2. Le graphique 2 montre clairement que pour chaque économie le coefficient estimé associé à la tendance linéaire $\hat{\xi}$ (représenté sur l'axe des ordonnées) est d'au-

tant plus important que la condition initiale (figurant sur l'axe des abscisses) est faible. Ce constat correspond à l'idée de rattrapage des économies européennes. L'évolution de la variance en coupe (graphique 1) confirme la réduction importante des écarts de productivités européennes. Ce phénomène ne peut être capturé, de façon satisfaisante par la définition de la convergence et la procédure de test envisagées dans cette section.

C'est l'utilisation de la statistique de test κ_1 plutôt que κ_0 (comme HOBIJN et FRANSES [2000]) qui permet cette conclusion nuancée. L'utilisation de la statistique de test de stationnarité stochastique κ_0 amènerait à conclure à la divergence (car les processus d'intérêt sont tendanciels au sens déterministe) mais ne permettrait pas d'associer ce résultat à l'existence d'un éventuel phénomène de rattrapage des économies.

4 Rattrapage et rupture de tendance

4.1 Une définition amendée

L'objet de cette section est de montrer comment une définition de la convergence basée sur les propriétés des séries temporelles peut rendre compte de l'existence du phénomène de rattrapage entre les économies. On peut imaginer le cas où dans une première phase les économies se déplacent le long de sentiers distincts, à des rythmes différents, puis qu'à partir d'une date T_B l'ensemble de ces économies évoluent sur des sentiers de croissance parallèles ou le long d'un sentier commun. Si dans la première phase les économies se rapprochent, alors elle correspond à une phase de rattrapage des économies. Dans ce cas il existe une date à partir de laquelle les économies vérifient la définition de la convergence donnée dans la section (2.1).

DÉFINITION : N économies convergent si (i) les écarts de productivité sont stationnaires au sens stochastique, (ii) il existe une date T_B à partir de laquelle ces écarts sont stationnaires au sens déterministe, (iii) antérieurement à cette date on constate un rapprochement des économies.

La convergence des économies est absolue si, à partir de la date T_B les écarts sont de moyennes nulles.

On rend compte de l'instabilité des interdépendances entre les économies en ne travaillant que sur la partie déterministe du processus multivarié décrivant l'évolution des productivités en déviation à la moyenne spatiale⁷. Cette

7. ST AUBYN [1996] propose d'utiliser le filtre de Kalman pour tester la convergence, au sens des séries temporelles, en tenant compte de l'existence d'une phase de transition. Il propose de tester la convergence vers zéro de la variance de l'innovation de la marche aléatoire dans l'équation (2) (il travaille dans un cadre univarié et ne prend donc pas en compte la totalité des interdépendances entre les N économies). Cette technologie de test de la convergence lui permet d'envisager l'absence d'invariance dans la partie stochastique du processus, alors que je considère l'absence d'invariance dans la partie déterministe.

méthodologie correspond à celle de CARLINO et MILLS [1993], mais ces derniers n'utilisent pas la même technologie de test, ils travaillent dans un cadre univarié et avec des statistiques de tests standards (ADF), qui ne sont pas les plus satisfaisantes pour tester la convergence des économies (Cf. *supra*). De plus, les motivations sont différentes. Ils introduisent des ruptures, en suivant la problématique de PERRON [1989], pour se donner les moyens de rejeter la présence de racine unitaire dans les écarts de revenus par tête des régions américaines, sans revenir sur la définition de la convergence. Ici, l'introduction d'une rupture commune, en modifiant la définition de la convergence des économies, permet de réconcilier une analyse en série temporelle avec l'existence d'une longue phase de transition et de montrer que l'aspect tendanciel au sens déterministe des écarts de productivités ne doit pas nécessairement amener à conclure à la divergence des économies.

Si on note Z_t le vecteur de dimension $N \times 1$ des productivités en déviation à la moyenne spatiale, alors on adopte la représentation suivante de l'évolution des productivités « centrées » sous la nulle de stationnarité stochastique,

$$(3) \quad Z_t = \gamma_1 + \gamma_2 \iota(t) + \mu_1 t + \mu_2 \iota(t)t + \varepsilon_t$$

avec $\iota(t) \equiv \mathbb{I}_{\{t > T_B\}}(t)$, où $T_B = [\lambda T]$ et $\lambda \in [0, 1]$, une fonction indicatrice. Le processus multivarié Z_t se décompose en la somme d'une tendance déterministe segmentée, d'une perturbation stationnaire. L'aspect segmenté de la partie déterministe rend compte de l'instabilité de la croissance. $\gamma_1, \gamma_2, \mu_1, \mu_2$ sont des vecteurs de dimension $N \times 1$. On distingue trois modèles,

$$(M0) \quad \gamma_1 \neq 0, \mu_1 \neq 0, \mu_2 = \gamma_2 = 0$$

$$(M1) \quad \gamma_1 \neq 0, \gamma_2 \neq 0, \mu_1 \neq 0, \mu_2 = 0$$

$$(M2) \quad \gamma_1 \neq \gamma_2, \mu_1 \neq \mu_2 \text{ et } \gamma_1, \gamma_2, \mu_1, \mu_2 \neq 0$$

Le modèle (M1) envisage le cas où à partir de la date T_B toutes les économies évoluent sur des sentiers de croissance parallèles. Le modèle (M0) envisage le cas où, à un aléa stationnaire près, les économies partagent le même sentier de croissance à partir d'une date T_B . Pour le modèle (M2) l'ensemble des économies n'évoluent pas sur des sentiers de croissance parallèles à partir de cette date.

Pour tester la convergence, absolue ou conditionnelle, tout en intégrant la nature instable de la dynamique des productivités, on peut utiliser la même statistique de test qu'auparavant (κ). La section suivante a pour objet d'établir ce point.

4.2 Stationnarité et rupture de tendance

AHN [1994] propose plusieurs tests de stationnarité, l'un d'eux correspond à la statistique du test KPSS multivarié, intégrant la possibilité d'existence de ruptures de tendances. Il commence par montrer que la statistique du test KPSS diverge lorsque l'on ne prend pas en compte l'existence de telles

ruptures, alors même que les processus scalaires sont stationnaires au sens stochastique. Puis il donne un théorème décrivant, de façon très générale, la distribution sous la nulle de la statistique de test. La proposition suivante envisage un cas particulier. Elle donne le comportement asymptotique de la statistique de test lorsque l'on considère les modèles (M0), (M1) ou (M2) sous la nulle. On suppose que la date de rupture $T_B \equiv \lfloor \lambda T \rfloor$ est connue.

PROPOSITION : (a) On suppose que la perturbation multivariée ε_t vérifie les conditions énoncées par PHILLIPS et DURLAUF [1986]. Sous l'hypothèse nulle (M0), la statistique κ adopte le comportement asymptotique suivant,

$$\kappa_{1,1/0,0}(\lambda) \implies \text{trace} \left\{ \int_0^\lambda \mathcal{H}_2(r,\lambda) \mathcal{H}_2(r,\lambda)' dr + \int_\lambda^1 \mathcal{H}_0(r,\lambda) \mathcal{H}_0(r,\lambda)' dr \right\}$$

Sous l'hypothèse nulle (M1) la statistique κ adopte le comportement asymptotique suivant,

$$\kappa_{1,1/1,0}(\lambda) \implies \text{trace} \left\{ \int_0^\lambda \mathcal{H}_2(r,\lambda) \mathcal{H}_2(r,\lambda)' dr + \int_\lambda^1 \mathcal{H}_1(r,\lambda) \mathcal{H}_1(r,\lambda)' dr \right\}$$

Sous l'hypothèse nulle (M2), la statistique κ adopte le comportement asymptotique suivant,

$$\kappa_{1,1/1,1}(\lambda) \implies \text{trace} \left\{ \int_0^\lambda \mathcal{H}_2(r,\lambda) \mathcal{H}_2(r,\lambda)' dr + \int_\lambda^1 \mathcal{H}_3(r,\lambda) \mathcal{H}_3(r,\lambda)' dr \right\}$$

$$\text{avec } \mathcal{H}_2(r,\lambda) \equiv W(r) - \frac{r}{\lambda} \left(4 - 3 \left(\frac{r}{\lambda} \right)^2 \right) \int_0^\lambda dW(r) - \frac{6r(r-\lambda)}{\lambda^3} \int_0^\lambda r dW(r),$$

$$\mathcal{H}_0(r,\lambda) \equiv W(r) - \int_0^\lambda dW(r), \mathcal{H}_1(r,\lambda) \equiv W(r) - \int_0^1 dW(r) + \frac{1-r}{1-\lambda} \int_\lambda^1 dW(r)$$

$$\text{et } \mathcal{H}_3(r,\lambda) \equiv W(r-\lambda) - \frac{(r-\lambda)\{4(1+\lambda+\lambda^2) - 3(\lambda+1)(r+\lambda)\}}{(1-\lambda)^3} \int_\lambda^1 dB(r) \\ - \frac{6(r-\lambda)(r-1)}{(1-\lambda)^3} \int_\lambda^1 r dB(r).$$

(b) Sous l'hypothèse alternative de non stationnarité stochastique, les statistiques de test $\kappa_{./}(\lambda)$ divergent, $\kappa_{./}(\lambda) = O_p(T^{1-\delta})$. (c) Si la date de rupture commune λ ne correspond pas à la vraie date de rupture (λ^*), alors $\kappa_{./}(\lambda) = O_p(T^{1-\delta})$. (d) Si le vraie modèle correspond à (M2) ou (M1), alors $\kappa_{1,1/0,0} = O_p(T^{1-\delta})$. (e) Si le vraie modèle correspond à (M2), alors $\kappa_{1,1/1,0} = O_p(T^{1-\delta})$.

| PREUVE : Se reporter à l'annexe.

On a $\mathcal{H}_3(r,0) \equiv \mathcal{H}_2(r,1) \equiv \tilde{V}(r)$, $\mathcal{H}_0(r,0) \equiv W(r)$ et $\mathcal{H}_1(r,0) \equiv V(r)$, on retrouve ainsi les distributions limites données dans la section (3.2), en l'absence de rupture dans la partie déterministe du processus. La date de rupture, caractérisée par $\lambda \in]0, 1[$, intervient comme un paramètre de nuisance. La

mise en œuvre du test demeure donc relativement simple. Dès lors que la vraie date de rupture commune est connue, il suffit d'estimer le modèle, (M1) ou (M0), sous l'hypothèse nulle de stationnarité stochastique, puis de calculer la statistique de test κ à partir des résidus estimés⁸. Puisque λ est un paramètre de nuisance, il faudra retabuler la loi de la statistique de test pour chaque valeur envisagée de la vraie date de rupture.

4.3 Retour au cas européen

La statistique de test $\kappa_{1,1/0,0}(\lambda)$, qui correspond au modèle (M0), permet de tester si à partir d'une date T_B , toutes les économies évoluent le long d'un sentier de croissance commun (convergence absolue). La première étape consiste à estimer par les MCO l'équation suivante pour $N - 1$ économies,

$$z_{i,t} = y_{i,t} - \bar{y}_{.,t} = \alpha(1 - \iota(t)) + \beta(1 - \iota(t))t + \varepsilon_{i,t}$$

Pour estimer cette équation j'ai considéré la date de rupture $T_B = 1979$ (les résultats de cette estimation sont donnés dans le tableau 2). Ce choix est motivé par l'observation de l'évolution de la variance en coupe des revenus par tête⁹ (graphique 1). En posant $l = 4$ et en employant la fenêtre QS, on trouve $\hat{\kappa}_{1,1/0,0} = 0.8674$. Au seuil de 5 %, on ne parvient pas à rejeter la stationnarité stochastique et l'existence d'une date (1979) à partir de laquelle les productivités évoluent le long d'un sentier de croissance commun contre la non stationnarité stochastique, la convergence conditionnelle, ou la divergence des taux de croissance¹⁰.

TABLEAU 2

Estimation des coefficients associés à la constante (α) et à la tendance linéaire (β) par les MCO sur la période (1950-1979)

Pays	$\hat{\alpha}$	$\hat{\beta}$
Belgique	0,21544	- 0,0031
Danemark	0,23437	- 0,0091
France	0,04772	0,0030
Allemagne	- 0,01561	0,0058
Grèce	- 1,00579	0,0200
Irlande	- 0,36259	0,0005
Italie	- 0,25835	0,0114
Luxembourg	0,50287	- 0,0110
Pays-Bas	0,30176	- 0,0016
Grande Bretagne	0,34016	- 0,0158

8. Se reporter à l'annexe B pour une discussion sur le niveau et la puissance du test à distance finie.

9. Évidemment, la « constance », à un aléa près de la variance en coupe après 1979, ne signifie pas que les économies évoluent sur un même sentier de croissance ou sur des sentiers de croissance parallèles.

10. Pour $l = 2,3$ on obtient respectivement $\hat{\kappa}_{1,1/0,0} = 1.0469$ et $\hat{\kappa}_{1,1/0,0} = 0.8329$. En utilisant la méthode de NEWEY et WEST [1994] pour déterminer le paramètre de troncation, on obtient avec les fenêtres QS et de Bartlett respectivement $\hat{\kappa}_{1,1/0,0} = 0.8674$ et $\hat{\kappa}_{1,1/0,0} = 0.8232$.

Il semble donc que les dynamiques des économies européennes puissent se découper en deux phases. Dans une première phase, les économies sont caractérisées par des dynamiques de transition. Après 1979, elles évoluent le long d'un sentier de croissance commun et vérifient donc à partir de cette date la définition de la convergence absolue au sens de la section (2.1). La phase de transition correspond bien à un phénomène de rapprochement des dix économies européennes considérées.

L'introduction de la rupture (modèle (M0)) éclaire les résultats obtenus dans la section (3.3). Les tendances linéaires alors obtenues rendent compte de l'instabilité dans la chronologie des dynamiques de croissance respectives des dix économies européennes, elles ne doivent pas nous amener à conclure à la divergence des productivités européennes.

5 Conclusion

Une conclusion émerge clairement ; la définition de la convergence basée sur les propriétés des séries temporelles peut rendre compte de l'existence d'une longue phase de transition, si l'on associe le phénomène de rattrapage à la partie déterministe des processus d'écart de revenu par tête. Pour dix économies européennes, on ne peut rejeter l'hypothèse de l'existence d'une date (1979) après laquelle les économies convergent au sens des séries temporelles, c'est-à-dire à partir de laquelle les processus de revenu par tête en déviation à la moyenne spatiale sont stationnaires de moyenne nulle.

Il convient cependant de s'interroger sur la représentation des données employée pour rendre compte de l'existence d'une longue phase de rattrapage des économies. Est-il légitime d'utiliser des tendances linéaires segmentées pour reconnaître ce phénomène alors que le modèle de croissance néo-classique prédit un ajustement « lisse » vers la trajectoire de long terme ? La conclusion pour le cas européen n'est-elle pas influencée par cette approximation ?

Ces questions incitent à rechercher une formulation plus souple de la partie déterministe, *ie*, sans recourir à des tendances linéaires par morceaux, pour rendre compte du processus de rattrapage des économies. Une telle formulation devra permettre de travailler sur une plus grande variété d'échantillons, puis de reprendre l'exercice de HOBIN et FRANCES [2000] en intégrant l'instabilité du processus de rattrapage des économies. ▼

• Références bibliographiques

- AHN B.C. (1994). – « Testing the Null of Stationarity in the Presence of Structural Breaks for Multiple Time Series », *Mimeo*, Department of Economics, The Ohio State University, Columbus.
- ANDREWS D.W.K. (1991). – « Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation », *Econometrica*, 59, p. 817-858.
- BERNARD A.B., DURLAUF S.N. (1995). – « Convergence in International Output », *Journal of Applied Econometrics*, 10, p. 97-108.
- BERNARD A.B., DURLAUF S.N. (1996). – « Interpreting Tests of the Convergence Hypothesis », *Journal of Econometrics*, 85, p. 161-173.
- BERNARD A.B., JONES C. (1996). – « Productivity across Industries and Countries: Time Series Theory and Evidence », *Review of Economics and Statistics*, 106, p. 135-146.
- CARLINO G.A., MILLS L.O. (1993). – « Are US Regional Incomes Converging? », *Journal of Monetary Economics*, 32, p. 335-346.
- CHOI I., AHN B.C. (1999). – « Testing the Null of Stationarity for Multiple Time Series », *Journal of Econometrics*, 88, p. 41-77.
- DURLAUF S.N., QUAH D.T. (1998). – « The New Empirics of Economic Growth », *Discussion Paper*, 384, CEP.
- EVANS P., KARRAS G. (1996). – « Convergence Revisited », *Journal of Monetary Economics*, 37, p. 249-265.
- HOBIIJN B., FRANSES P.H., OOMS M. (1998). – « Generalizations of the KPSS-Test for Stationarity », *Working Paper 9802/A*, Erasmus University of Rotterdam.
- HOBIIJN B., FRANSES P.H. (2000). – « Asymptotically Perfect and Relative Convergence of Productivity », *Journal of Applied Econometrics*, 15(1).
- KWIATKOWSKI D., PHILLIPS P.C.B., SCHMIDT P., SHIN Y. (1992). – « Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure are we that Economic Time Series have a Unit Root? », *Journal of Econometrics*, 54, p. 159-178.
- LEE J. (1996). – « On the Power of Stationarity Tests Using Optimal Bandwidth Estimates », *Economics Letters*, 51, p. 131-137.
- MANKIW N.G., ROMER D., WEIL D.N. (1992). – « A Contribution to the Empirics of Economic Growth », *Quarterly Journal of Economics*, 107, p. 407-437.
- NEWBY W.K., WEST K. (1994). – « Automatic Lag Selection in Covariance Matrix Estimation », *Review of Economic Studies*, 61, p. 631-653.
- PERRON P. (1989). – « The Great Crash, the Oil-Price Shock, and the Unit Root Hypothesis », *Econometrica*, 57, p. 1361-1401.
- PHILLIPS P.C.B. (1991). – « Spectral Regression for Cointegrated Time Series », in William Barnett, editor, *Nonparametric and Semiparametric Methods in Economics and Statistics*, chapter 16, p. 413-435, Cambridge University Press.
- PHILLIPS P.C.B., DURLAUF S.N. (1986). – « Multiple Time Series Regression with Integrated Processes », *Review of Economic Studies*, 53, p. 473-495.
- POLITIS D.N., ROMANO J.P., WOLF M. (1999). – *Subsampling*, Springer-Verlag, New York.
- QUAH D.T. (1990). – « International Patterns of Growth: I. Persistence in Cross-Countries Disparities », *Working Paper*, MIT.
- QUAH D.T. (1993). – « Galton's Fallacy and Tests of the Convergence Hypothesis », *Scandinavian Journal of Economics*, 95, p. 427-443.
- SALA-I-MARTIN X. (1996). – « The Classical Approach to Convergence Analysis », *Economic Journal*, 106, p. 1019-1036.
- ST AUBYN M. (1996). – « Convergence across Industrialized Countries (1890-1989): New Results using Time Series Methods », *Working Paper 2/96*, Universidade Técnica de Lisboa.
- STOCK J. (1994). – « Unit Roots, Structural Breaks and Trends », In Engle et McFadden, editors, *Handbook of Econometrics, volume IV*, p. 2739-2841, Elsevier Science.

ANNEXES

A. Preuve de la proposition

(a) On ne dérive la distribution limite de la statistique de test que pour le modèle (M1), pour les modèles (M0) et (M2) la preuve est identique. Sous l'hypothèse nulle le processus $\{Y_t\}_{t=1}^T$, s'écrit de la façon suivante :

$$(A.1) \quad \Leftrightarrow Y_t = Ax_t + \varepsilon_t$$

avec $A \equiv (\alpha_1 \ \alpha_2 \ \xi)$ une matrice de dimension $N \times 3$, $x_t \equiv (1 \ \iota(t) \ (1 - \iota(t))t)'$ un vecteur de dimension 3×1 . Le processus multivarié $\{\varepsilon_t\}_{t=1}^T$ est stationnaire au second ordre d'espérance nulle, avec pour matrice de variance-covariance Σ . On suppose que ce processus vérifie les conditions énoncées par PHILLIPS et DURLAUF [1986]. On note $\Omega \equiv \lim_{T \rightarrow \infty} E[T^{-1}S_T S_T'] = \Sigma + \sum_{i=1}^{\infty} (\Gamma(i) + \Gamma(i)') = E[\varepsilon_t \varepsilon_t'] + \sum_{i=1}^{\infty} \{E[\varepsilon_t \varepsilon_{t+i}'] + E[\varepsilon_{t+i} \varepsilon_t']\}$, la matrice de variance covariance de long terme. Si cette matrice est définie positive, alors PHILLIPS et DURLAUF [2000, page 476] établissent le principe d'invariance suivant,

$$T^{-\frac{1}{2}} \sum_{t=0}^{[rT]} \varepsilon_t \Longrightarrow B(r) = \Omega^{\frac{1}{2}} W(r)$$

où $W(r)$ est un processus de Wiener standard de dimension $N \times 1$, et « \Longrightarrow » indique la convergence en loi lorsque T tend vers l'infini.

La somme partielle estimée, \hat{S}_t , par les moindres carrés ordinaires, est donnée par l'expression suivante,

$$\hat{S}_t = \sum_{\tau=1}^t \varepsilon_{\tau} - \sum_{t=1}^T \varepsilon_t x_t' \left(\sum_{t=1}^T x_t x_t' \right)^{-1} \sum_{\tau=1}^t x_{\tau}$$

Soit $\delta_T = \text{diag}\{1, 1, T\}$. On a alors,

$$\delta_T \left(\sum_{t=1}^T x_t x_t' \right)^{-1} \delta_T = \frac{1}{[\lambda T]([\lambda T] - 1)} \begin{pmatrix} 2(2[\lambda T] + 1) & -2(2[\lambda T] + 1) & -6T \\ -2(2[\lambda T] + 1) & 2(2[\lambda T] + 1) + \frac{[\lambda T]([\lambda T] - 1)}{[(1 - \lambda)T]} & 6T \\ -6T & 6T & \frac{12T^2}{[\lambda T] + 1} \end{pmatrix}$$

$$T^{-\frac{1}{2}} \sum_{t=1}^T \varepsilon_t x_t' \delta_T^{-1} \implies \left(\int_0^1 dB(r) \quad \int_\lambda^1 dB(r) \quad \int_0^\lambda r dB(r) \right)$$

$$\delta_T^{-1} \sum_{t=1}^{\lfloor rT \rfloor} x_t = \begin{cases} \left(\lfloor rT \rfloor \quad 0 \quad \frac{\lfloor rT \rfloor (\lfloor rT \rfloor + 1)}{2T} \right)' & \text{si } r < \lambda \\ \left(\lfloor rT \rfloor \quad \lfloor (r - \lambda)T \rfloor \quad \frac{\lfloor \lambda T \rfloor (\lfloor \lambda T \rfloor + 1)}{2T} \right)' & \text{si } r > \lambda \end{cases}$$

La déduction du comportement asymptotique de la somme partielle estimée suit directement,

Cas 1 : $r < \lambda$

$$(A.2) \quad T^{-\frac{1}{2}} \hat{S}_{\lfloor rT \rfloor} \implies B(r) - \frac{r}{\lambda} \left(4 - 3 \frac{r}{\lambda} \right) \int_0^\lambda dB(r) - \frac{6r(r - \lambda)}{\lambda^3} \int_0^\lambda r dB(r) \equiv \mathcal{H}_2(r, \lambda)$$

On a $\mathcal{H}_2(r, 1) = \Omega^{\frac{1}{2}} \tilde{V}(r)$, où $\tilde{V}(r)$ est un pont brownien du second ordre de dimension $N \times 1$, on retrouve donc le comportement asymptotique de la somme partielle estimée en l'absence de rupture de tendance dans le cas de la stationnarité stochastique autour d'une tendance linéaire.

Cas 2 : $r > \lambda$

$$(A.3) \quad T^{-\frac{1}{2}} \hat{S}_{\lfloor rT \rfloor} \implies B(r) - \int_0^1 dB(r) + \frac{1 - r}{1 - \lambda} \int_\lambda^1 dB(r) \equiv \mathcal{H}_1(r, \lambda)$$

On a $\mathcal{H}_1(r, 0) = \Omega^{\frac{1}{2}} V(r)$, où $V(r)$ est un pont brownien de dimension $N \times 1$, on retrouve donc le comportement asymptotique de la somme partielle estimée en l'absence de rupture dans le cas de la stationnarité en niveau.

On s'intéresse à la distribution limite de l'expression suivante,

$$T^{-2} \sum_{t=1}^T \hat{S}_t \hat{S}_t' \# \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} \left(T^{-\frac{1}{2}} \hat{S}_{\lfloor rT \rfloor} \right) \left(T^{-\frac{1}{2}} \hat{S}_{\lfloor rT \rfloor} \right)' dr$$

$$\Leftrightarrow T^{-2} \sum_{t=1}^T \hat{S}_t \hat{S}_t' \# \int_0^\lambda \left(T^{-\frac{1}{2}} \hat{S}_{\lfloor rT \rfloor} \right) \left(T^{-\frac{1}{2}} \hat{S}_{\lfloor rT \rfloor} \right)' dr + \int_\lambda^1 \left(T^{-\frac{1}{2}} \hat{S}_{\lfloor rT \rfloor} \right) \left(T^{-\frac{1}{2}} \hat{S}_{\lfloor rT \rfloor} \right)' dr$$

Par substitution de (A.2) et (A.3) dans l'expression précédente, on obtient,

$$T^{-2} \sum_{t=1}^T \hat{S}_t \hat{S}_t' \implies \text{trace} \left\{ \Omega \left(\int_0^\lambda \mathcal{H}_2(r, \lambda) \mathcal{H}_2(r, \lambda)' dr + \int_\lambda^1 \mathcal{H}_1(r, \lambda) \mathcal{H}_1(r, \lambda)' dr \right) \right\}$$

d'où le résultat asymptotique annoncé dans la proposition.

(b) Pour montrer la divergence de la statistique de test sous l'alternative de non stationnarité stochastique on procède comme CHOI et AHN [1999].

(c) Voir AHN [1994].

(d) & (e) La statistique de test diverge, dans le cas de stationnarité stochastique, lorsque la partie déterministe du processus multivarié est mal spécifiée sur la deuxième période, c'est-à-dire lorsqu'en fait le processus est tendanciel au sens déterministe sur les deux sous périodes (si on considère $\kappa_{1,1/1,0}(\lambda)$ ou $\kappa_{1,1/0,0}(\lambda)$), ou lorsqu'en fait le processus est stationnaire en niveau sur la deuxième période (si on considère $\kappa_{1,1/0,0}(\lambda)$). Montrons que la statistique $\kappa_{1,1/1,0}(\lambda)$ diverge lorsque la partie déterministe du processus est mal définie sur la deuxième période (la preuve est identique pour la statistique $\kappa_{1,1/0,0}(\lambda)$). Dans ce cas, on estime modèle (A.1) alors que le processus générateur de données est de forme suivante,

$$(A.4) \quad \Leftrightarrow Y_t = \tilde{A} \tilde{x}_t + u_t$$

où u_t est un processus stationnaire vérifiant les mêmes conditions que ε_t , avec $\tilde{A} \equiv (\alpha_1 \ \alpha_2 \ \xi_1 \ \xi_2)$ une matrice de dimension $N \times 4$, $\tilde{x}_t \equiv (1 \ \iota(t) \ t \ \iota(t)t)'$ un vecteur de dimension 4×1 . Par substitution de (A.4) dans l'expression analytique du résidu estimé, on obtient,

$$.e_t = u_t - \sum_{t=1}^T u_t x_t' \left(\sum_{t=1}^T x_t x_t' \right)^{-1} x_t + \tilde{A} \tilde{x}_t - \sum_{t=1}^T \tilde{A} \tilde{x}_t x_t' \left(\sum_{t=1}^T x_t x_t' \right)^{-1} x_t$$

Soient $\hat{S}_t = \sum_{\tau=1}^t e_\tau$, $S_t^u = \sum_{\tau=1}^t u_\tau$, $S_t^x = \sum_{\tau=1}^t x_\tau$ et $S_t^{\tilde{x}} = \sum_{\tau=1}^t \tilde{x}_\tau$. Selon ces notations, on a alors,

$$(A.5) \quad \hat{S}_{[rT]} = S_{[rT]}^u - \sum_{t=1}^T u_t x_t' \left(\sum_{t=1}^T x_t x_t' \right)^{-1} S_{[rT]}^x \\ + \tilde{A} S_{[rT]}^{\tilde{x}} - \sum_{t=1}^T \tilde{x}_t x_t' \left(\sum_{t=1}^T x_t x_t' \right)^{-1} S_{[rT]}^x$$

Asymptotiquement, on a,

$$\begin{aligned} -S_{[rT]}^u &= O_p(T^{\frac{1}{2}}), \\ -\sum_{t=1}^T u_t x_t' \left(\sum_{t=1}^T x_t x_t' \right)^{-1} S_{[rT]}^x &= O_p(T^{\frac{1}{2}}), \\ -S_{[rT]}^{\tilde{x}} - \sum_{t=1}^T \tilde{x}_t x_t' \left(\sum_{t=1}^T x_t x_t' \right)^{-1} S_{[rT]}^x &= O(T^2). \end{aligned}$$

On a donc,

$$\sum_{t=1}^T \hat{S}_t \hat{S}_t' = O(T^5)$$

Pour déterminer la vitesse de divergence de la statistique de test il reste à étudier le comportement asymptotique de l'estimateur de la matrice de variance-covariance de long terme. En suivant les notations de PHILLIPS [1991],

$$\frac{1}{lT^2} \hat{\Omega}(l) = \frac{2\pi}{lT^2} \hat{f}_{ee}(0) = \frac{1}{l} \sum_{s=-l}^l k\left(\frac{s}{l}\right) \frac{1}{T^2} c(s)$$

où $\hat{f}_{ee}(0)$ est l'estimateur de la densité spectrale en zéro du résidu estimé e_t , $k(x)$ est une fenêtre de pondération des matrices d'auto-covariance (la fenêtre de Bartlett par exemple) et où,

$$c(s) = T^{-1} \sum_{t+s=1}^T e_t e_{t+s}' = T^{-1} \sum_{t+s=1}^T \Delta \hat{S}_t \Delta \hat{S}_{t+s}' = O(T^2)$$

On a alors $\hat{\Omega}(l) = O_p(lT^2)$ où $l = O_p(T^\delta)$, avec $\delta \in \left] 0, \frac{1}{2} \right[$. Finalement $\kappa_{1,1/1,0} = O_p(T^{1-\delta})$ lorsque le processus sous-jacent est tendanciel au sens déterministe dans les deux sous-périodes de l'échantillon.

B. Puissance et niveau empiriques

HOBIN, FRANSES et OOMS [1998] montrent que la méthode de sélection proposée par NEWBY et WEST [1994], utilisée sans *prewhitening* des résidus, conduit à une statistique de test convergente. Des simulations montrent alors que, même pour des petits échantillons, la statistique de test possède un niveau empirique et une puissance empirique satisfaisants. L'introduction de ruptures dans la partie déterministe n'enlève rien aux propriétés à distance finie de ce test. On considère quatre modèles pour décrire le processus générateur des données sous différentes hypothèses alternatives. Le tableau page suivante donne une description de ces modèles.

v_t est une marche aléatoire bivariée, ε_t est un processus stationnaire, $\iota(t) = \mathbb{I}_{t \leq T_B}(t)$ avec $T_B = [\lambda T]$ et $\lambda = 0,73$ (comme dans l'application sur les économies Européennes), ξ un vecteur 2×1 de paramètres. On considère

Modèle 1	$y_t = \xi t + v_t$	I(1)+T
Modèle 2	$y_t = \xi t + \varepsilon_t$	I(0)+T
Modèle 3	$y_t = \mu_1 t \iota(t) + \gamma_1 \iota(t) + (1 - \iota(t)) \gamma_2 + \varepsilon_t$	(M1)
Modèle 4	$y_t = \mu_1 t \iota(t) + \gamma_1 \iota(t) + \varepsilon_t$	(M0)

que l'innovation de la marche aléatoire est un bruit blanc de variance $\Sigma = I_2$, que le processus stationnaire ε_t est un processus autorégressif d'ordre un, avec pour matrice autorégressive $A = ((a_{i,j}))_{i,j=1,2}$ où $a_{1,1} = 0,5$, $a_{1,2} = 0$, $a_{2,1} = 0,2$ et $a_{2,2} = 0,4$, que l'innovation du processus stationnaire est donnée par $\Omega = ((\omega_{i,j}))_{i,j}$ où $\omega_{1,1} = 1$, $\omega_{1,2} = 0,5$, $\omega_{2,1} = 0,5$ et $\omega_{2,2} = 1,5$, que la condition initiale du processus bivarié y_t est nulle.

Les modèles 1, 2 et 3 permettent d'évaluer la puissance de la statistique de test $\kappa_{1,1/0,0}$ contre plusieurs alternatives, le modèle 4 permet d'évaluer le niveau empirique de cette même statistique. Le tableau suivant reporte la puissance du test lorsque l'on utilise la procédure de NEWBY et WEST [1994] et la fenêtre *Quadratic Spectral*¹¹ pour estimer la matrice de variance covariance de long terme¹².

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Puissance ou niveau	0,8388	0,6918	0,9659	0,0529

Ce test semble donc raisonnablement puissant contre les alternatives de non stationnarité stochastique et/ou linéaire, contre la convergence conditionnelle. Ainsi, en tenant compte de la singularité liée à la considération de processus en déviation à la moyenne spatiale, si trois économies sont caractérisées par de la convergence conditionnelle, alors on rejette 96 fois sur 100 la convergence absolue.

Cependant ces résultats, en terme de puissance et de niveau, sont très sensibles à la spécification du processus générateur des données sous l'alternative. Par l'expression (A.5) on sait que la valeur des paramètres associés à la tendance et à la constante sous l'alternative constituent des paramètres de nuisances à distance finie (contrairement à la valeur des paramètres sous la nulle). De plus, à distance finie la distribution de la statistique de test dépend de la *taille de la racine unitaire*, même si asymptotiquement la variance de l'innovation de la marche aléatoire multivariée ne constitue pas un paramètre de nuisance. LEE [1996] montre, dans le cas univarié, l'importance de ces paramètres de nuisances à distance finie sur la puissance du test KPSS, lorsque l'on utilise une procédure de sélection du paramètre de troncation basée sur les propriétés de dépendance caractérisant les données. Si ces procédures améliorent le niveau du test, elles détériorent dans de nombreux

11. Qui selon ANDREWS [1991] et NEWBY et WEST [1994] donne un estimateur de la variance de long terme plus précis à distance finie.

12. Les nombres reportés dans ce tableau sont obtenus sur la base de 5 000 simulations des modèles 1 à 4 pour un échantillon de taille $T = 50$. Les valeurs critiques sont retabulées par Monte-Carlo (20 000 simulations) pour le cas $N = 2$ et $T = 50$. Le niveau théorique des tests est de 5 %.

cas la puissance du test. C'est pourquoi il peut sembler finalement préférable de se donner le paramètre de troncation. Procéder ainsi assure une plus grande puissance du test, mais un biais positif du niveau. C'est l'approche finalement retenue par HOBIN et FRANCES [2000]. Ce choix est aussi légitimé par la nature multivariée de la statistique de test. En effet, sélectionner un paramètre de troncation trop grand par une procédure automatique réduit d'autant plus la puissance que la dimension individuelle du processus multivarié est importante (l'estimation des matrices d'autocorrélation, nécessaires à la construction de la matrice de variance-covariance de long terme est d'autant plus coûteuse en données que la dimension individuelle est importante). Dans le texte on donne les statistiques de test en considérant un paramètre de troncation comme donné ($l = 2, 3, 4$, comme HOBIN et FRANCES [2000]) ce qui induit un niveau empirique trop important, puis compare ces statistiques avec le cas où le paramètre de troncation est obtenu avec la méthode de NEWBY et WEST [1994], qui réduit le biais du niveau mais aussi la puissance du test.