

Mesures et tests de convergence : une revue de la littérature

Catherine Fuss

Département des études
Université Libre de Bruxelles

De nombreuses questions économiques font référence au concept de convergence, qu'il s'agisse, par exemple de l'intégration économique, ou du thème du développement. Différentes définitions de convergence peuvent être utilisées, distinguant les séries convergentes des séries ayant convergé, les séries convergeant à une constante près de celles tendant à l'égalité. Différentes procédures de test sont à la disposition de l'analyste. Les analyses en panel et en coupe testent la convergence globale de l'échantillon vers le même niveau. L'analyse de l'évolution de la distribution offre une grande souplesse et permet de détecter des « clubs de convergence ». Les méthodes de séries chronologiques sont adaptées à l'analyse de la convergence d'un petit nombre de séries. Différentes procédures permettent de prendre en compte un grand nombre de schémas de convergence, de tenir compte de la dynamique des séries et de la dynamique du processus de convergence. En somme, les différents tests offrent à l'analyste des méthodes complémentaires dont l'utilisation dépend du type de données.

La notion de convergence est fréquemment utilisée dans l'analyse économique. L'intégration européenne, par exemple, fait référence à la convergence d'un ensemble d'indicateurs macroéconomiques. Le traité de Maastricht imposait des conditions précises de convergence des taux d'inflation, taux d'intérêt, déficit public et dette publique, entre pays. Les thèmes du développement et de la croissance introduisent également des questions de convergence de revenu, de niveau de vie ou de productivité entre pays. L'analyse microéconomique se base parfois sur ce type de notion. Par exemple, la collusion entre firmes situées sur différents marchés peut se traduire par une convergence des prix entre ces marchés. La question de convergence se pose donc dans un grand nombre de domaines, et pour une multitude de variables économiques.

Cependant, si la convergence est un concept intuitivement simple, sa définition précise l'est beaucoup moins et son analyse empirique s'est effectuée par différentes méthodes, selon, entre autres, le type de données considérées. Chaque type de test apporte différentes informations.

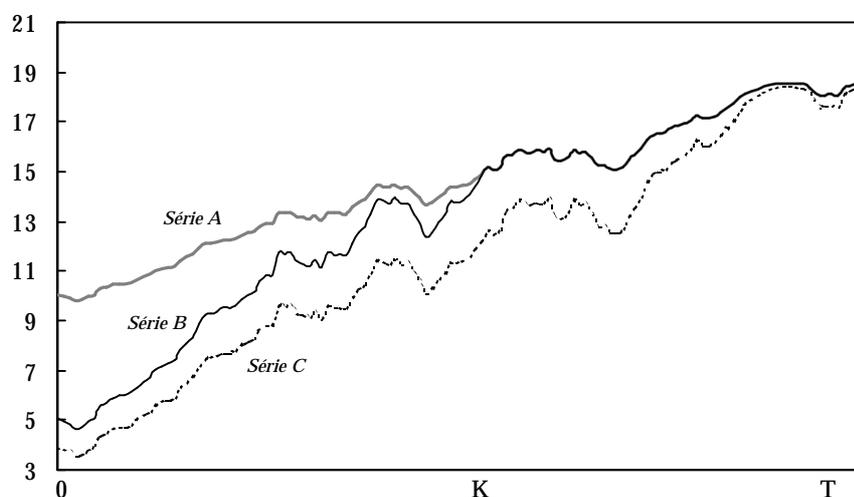
Cet article propose une revue des tests de convergence couramment utilisés. Après une discussion sur la définition de la convergence, l'article expose les différents tests de convergence, et discute leurs avantages et inconvénients.

Qu'est-ce que la convergence ?

À première vue, le concept de convergence est relativement aisé à définir. On parle de convergence lorsque la différence entre les séries ou leur dispersion se réduit dans le temps. Cependant, il existe différents schémas de convergence. On peut chercher à les caractériser de façon précise en opérant une distinction entre séries qui ont convergé et séries qui convergent, et entre convergence vers le même niveau et convergence à une constante près.

Les graphiques ci-dessous illustrent ces différents concepts. Sur le graphique 1, il est clair que *les séries A, B et C convergent*, la différence entre les séries et leur dispersion décroît dans le temps. On peut parler de rattrapage : les séries B et C partant d'un niveau inférieur à A « rattrapent » le niveau de A. La dispersion entre les séries tend vers zéro, il y a convergence vers le même niveau. Si on affine l'analyse en considérant deux sous-échantillons, de 0 à K et de K à T, on note que, sur la période [K, T], la différence et la dispersion entre les séries A et B est nulle, le processus de convergence est achevé, *les séries A et B ont convergé*.

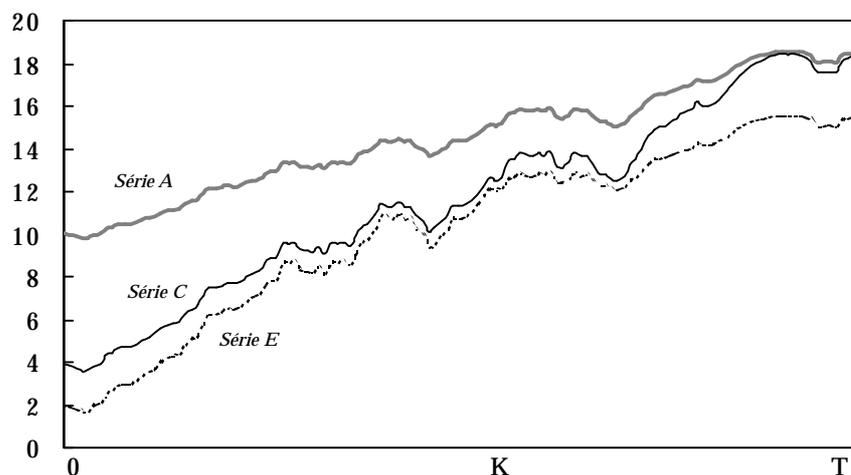
1. Exemples de séries convergentes



Sources : Calculs de l'auteur.

Sur le graphique 2, l'écart entre A et E se réduit dans le temps; il est constant sur la période $[K, T]$. Un critère strict de convergence, exigeant que la différence entre les séries tende vers zéro, conclurait à la divergence. Un critère plus souple, exigeant que la dispersion diminue, conclurait que les séries convergent. *Les séries convergent à une constante près; sur la période $[K, T]$, elles ont convergé à une constante près.*

2. Exemples de séries convergentes



Sources : Calculs de l'auteur.

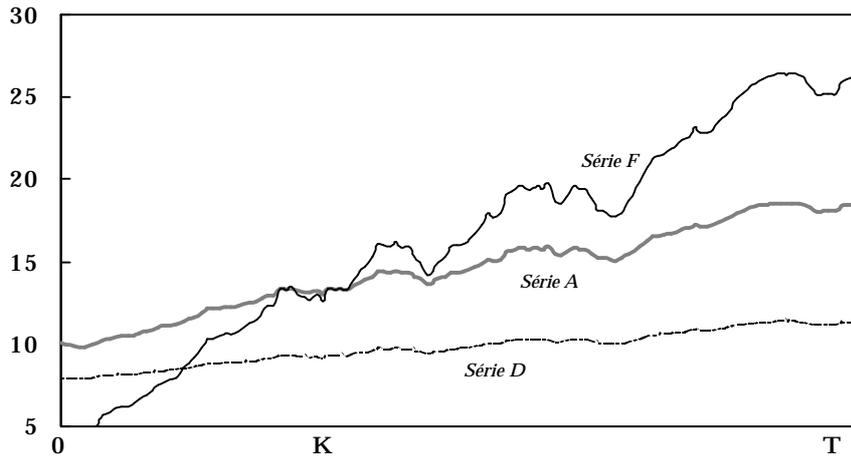
On voit déjà à partir de ces quelques exemples graphiques qu'on peut opérer une distinction entre séries convergentes, séries ayant convergé, séries convergentes à une constante près, séries ayant convergé à une constante près. La convergence à une constante près peut se justifier dans un grand nombre de cas pratiques. Par exemple, le traité de Maastricht imposait que les indicateurs économiques se situent dans une fourchette définie, c'est-à-dire qu'ils devaient converger à une constante près, pourvu que la constante n'excède pas la norme de Maastricht. Dans un autre domaine, Stigler et Sherwin (1985) définissent un marché économique par la stabilité du rapport des prix sur les différents marchés, autrement dit par la convergence des prix à une constante près. La constante peut être interprétée en termes de coûts de transports et autres coûts d'arbitrage entre les marchés. En finance, l'arbitrage entre deux actifs implique que les taux de rendement entre ces actifs convergent à une constante près, où la constante peut représenter le différentiel de prime de risque.

Tout comme il y a différentes façons de converger, il y a différentes façons de diverger. Sur le graphique 3, les séries F, D et A divergent. F « dépasse » la série A. La différence et la dispersion entre F et A

décroissent puis augmentent à partir de K. D reste en dessous de A et ne « rattrape » jamais la série A, la différence et la dispersion entre A et D se creusent au cours du temps.

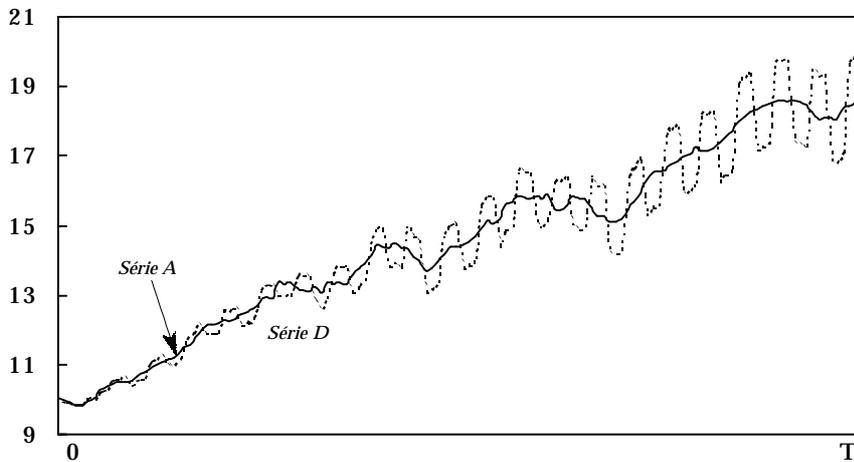
Le cas de la série G du graphique 4 est un peu particulier. En moyenne G est égale à A. Cependant, on ne peut pas parler de convergence car la dispersion entre A et G se creuse au cours du temps, la série G peut s'écarter de plus en plus de la série A. G est non seulement plus volatile que A, mais de plus sa volatilité croît dans le temps. La relation entre A et G est instable.

3. Exemples de séries divergentes



Sources : Calculs de l'auteur.

4. Exemple de séries divergentes



Sources : Calculs de l'auteur.

Caractérisation statistique

On peut identifier globalement deux façons de définir la convergence : en fonction de l'évolution moyenne de leur différence, ou en fonction de l'évolution de leur dispersion. Les deux notions ne sont pas équivalentes. Analyser la différence en moyenne amène à négliger les écarts occasionnels des séries à la moyenne. Par contre, la dispersion d'un ensemble de séries prend en compte les fluctuations des séries, c'est-à-dire leur volatilité.

De façon plus formelle, une adaptation de la classification de Hall, Robertson et Wickens (1993) donne les définitions suivantes ¹ :

(i) deux séries y_t et y_t^* convergent point à point si leur différence tend vers une constante :

$$\lim_{t \rightarrow \infty} (y_t - y_t^*) = a$$

(ii) deux séries y_t et y_t^* convergent en espérance si l'espérance de leur différence tend vers une constante :

$$\lim_{t \rightarrow \infty} E(y_t - y_t^*) = a$$

(iii) deux séries y_t et y_t^* convergent en probabilité si l'espérance de leur différence tend vers une constante :

$$\lim_{t \rightarrow \infty} E(y_t - y_t^*) = a,$$

et si la variance de leur différence tend vers zéro :

$$\lim_{t \rightarrow \infty} \text{var}(y_t - y_t^*) = 0$$

Le cas (i) correspond au cas des séries A et B du graphique 1, ces séries tendent à être égales, et sont égales à partir de la date K. Les séries A et E convergent point à point à une constante près. Cette définition est trop stricte pour l'analyse de séries économiques qui peuvent être affectées par des chocs économiques différents aux effets temporaires. La convergence en espérance laisse plus de souplesse dans l'analyse et autorise les séries à dévier l'une de l'autre pourvu qu'en moyenne elles soient égales. Autrement dit, des chocs aléatoires peuvent influencer distinctement chacune des séries, tant que ces chocs sont d'espérance nulle.

Ce critère peut s'avérer trop souple. Il conduirait, par exemple, à accepter l'hypothèse que les séries A et G du graphique 4 convergent puisqu'elles sont égales en moyenne. Par contre, l'application du critère de convergence en probabilité conduirait à rejeter la convergence entre les séries A et G puisqu'il exige que la variance de la différence entre les séries tende vers zéro.

1. En réalité, Hall et al (1993) considèrent le cas plus général $y_t^* - b \cdot y_t$, plutôt que la différence $y_t^* - y_t$.

Le critère de convergence en probabilité est extrêmement contraignant puisqu'il exige que la dispersion entre les séries tende vers zéro, c'est-à-dire que les séries tendent à converger point à point. Pour des séries économiques qui subissent l'effet de chocs aléatoires transitoires différents, ce critère est trop exigeant. Il est plus vraisemblable que la dispersion tende vers une constante, plutôt que vers zéro.

En conséquence je suggère un critère situé à mi-chemin entre la convergence en espérance et la convergence en probabilité :

deux séries y_t et y_t^* convergent si l'espérance de leur différence tend vers une constante :

$$\lim_{t \rightarrow \infty} E(y_t - y_t^*) = a,$$

et si la variance de leur différence tend vers une constante, S^2 :

$$\lim_{t \rightarrow \infty} \text{var}(y_t - y_t^*) = S^2.$$

Dans ce cas,

(*) les séries ont convergé (à une constante près) si l'espérance de leur différence et la variance de leur différence sont constantes :

$$E(y_t - y_t^*) = a \text{ et } \text{var}(y_t - y_t^*) = S^2.$$

(**) les séries convergent si l'espérance de leur différence tend vers une constante et si la variance de leur relation est constante :

$$\lim_{t \rightarrow \infty} E(y_t - y_t^*) = a \text{ et } \text{var}(y_t - b \cdot y_t^* - a) = S^2$$

ou si l'espérance de leur différence tend vers une constante et si la variance de leur relation tend vers une constante :

$$\lim_{t \rightarrow \infty} E(y_t - y_t^*) = a \text{ et } \lim_{t \rightarrow \infty} \text{var}(y_t - b \cdot y_t^* - a) = S^2$$

L'importance de la condition sur la variance apparaît clairement lorsqu'on applique ces critères à des séries non stationnaires.

1. Séries non stationnaires et cointégration

Une série est non stationnaire* lorsque sa moyenne, ou sa variance et ses autocovariances dépendent du temps. Des séries non stationnaires sont usuellement représentées par des processus à racine unitaire. Les séries sont dites intégrées d'ordre un si leur différence première est stationnaire. Un exemple de processus intégré d'ordre un est la marche aléatoire avec dérive :

$$(a) \quad x_t = a + x_{t-1} + h_t$$

où h_t est un bruit blanc d'espérance nulle et de variance S^2 . Un tel processus est de variance croissante dans le temps : $\text{var}(x_t) = S^2 \cdot t$, l'auto-

* Pour simplifier, on se concentre sur la stationnarité faible d'ordre deux.

corrélation d'ordre i , r_{i^p} dépend également du temps :

$$r_{i^p} = \sqrt{1 - \frac{i}{t}}$$

Par induction récursive, (a) peut se réécrire comme :

$$(b) \quad x_t = a.t + x_0 + \sum_{i=1}^t h_i$$

h_i a un effet sur toutes les valeurs de x_t postérieures à i , le choc h_i a des effets permanents sur la série x_t

$\sum_{i=1}^t h_i$ représente la tendance stochastique de x_t , $a.t$ représente la tendance déterministe.

Deux séries intégrées d'ordre un, y_t^* et y_t sont cointégrées s'il existe des paramètres a et b tels que le résidu de l'équation suivante, e_t , soit stationnaire :

$$(c) \quad y_t^* = a + b.y_t + e_t$$

Ceci est le cas lorsque les chocs permanents affectant les séries y_t^* et y_t ne sont pas indépendants. Lorsque les séries sont cointégrées, elles ne peuvent pas dévier de la relation (c) au-delà de marges de fluctuations constantes, puisque e_t est de variance constante. Au contraire, lorsqu'elles ne sont pas cointégrées, e_t est de variance croissante et les séries peuvent s'écartier de plus en plus de la relation (c).

Considérons, pour simplifier, deux marches aléatoires, y_t^* et y_t :

$$(1) \quad y_t^* = a^* + y_{t-1}^* + e_t^* \quad \text{avec } E(e_t^*) = 0 \text{ et } \text{var}(e_t^*) = S^{*2}$$

$$(2) \quad y_t = a + y_{t-1} + e_t \quad \text{avec } E(e_t) = 0 \text{ et } \text{var}(e_t) = S^2$$

Par induction récursive, (1) et (2) peuvent se réécrire comme :

$$(3) \quad y_t^* = a^*.t + y_0^* + \sum_{k=1}^t e_k^*$$

$$(4) \quad y_t = a.t + y_0 + \sum_{k=1}^t e_k$$

La différence entre les séries vaut :

$$(5) \quad y_t^* - y_t = (a^* - a).t + (y_0^* - y_0) + \left(\sum_{k=1}^t e_k^* - \sum_{k=1}^t e_k \right)$$

L'espérance de cette différence est donc :

$$(6) \quad E(y_t^* - y_t) = (a^* - a).t + (y_0^* - y_0)$$

La variance de la différence vaut :

$$(7) \quad \text{var}(y_t^* - y_t) = \text{var} \left(\sum_{k=1}^t e_k^* - \sum_{k=1}^t e_k \right)$$

Si y_0^* est supérieur à y_0 les séries convergent en espérance dès que $a^* < a$. Dans ce cas, la différence initiale entre les séries est positive et l'espérance de la différence est une fonction décroissante du temps. Les séries ont convergé lorsque $a^* = a$, dans ce cas, l'espérance de la différence

est constante. Si les séries ne sont pas cointégrées, c'est-à-dire si les chocs permanents qui affectent les séries sont indépendants, la variance de la différence est croissante dans le temps, elle vaut $(\mathbf{s}^{*2}.t + \mathbf{s}^2.t)$. L'écart entre les séries peut croître dans le temps, même s'il est constant en moyenne. Par contre lorsque les séries sont cointégrées, l'écart $y_t^* - b.y_t$ est de variance constante, il fluctue donc dans des marges constantes. Le critère (***) revient à exiger que les séries soient cointégrées ou tendent à être cointégrées.

Différents tests ont été proposés, des tests pour coupes transversales, d'autres pour panels, pour séries chronologiques ou encore pour l'analyse de l'évolution de la distribution. Les différences portent sur le type de données auxquels les tests s'appliquent aussi bien qu'aux informations qui en découlent. Les sections suivantes exposent les différents tests, leurs avantages et leurs inconvénients.

Tests de convergence en coupes transversales

Les tests de convergence en coupes transversales ont été développés dans la littérature sur la croissance, en particulier, par Barro et Sala-i-Martin, pour tester le rattrapage des revenus par tête des pays en développement vers ceux des pays développés. Cette question découle du modèle de croissance de Solow qui indique qu'à coefficients techniques et préférences identiques, les pays convergent vers le même état stationnaire (voir par exemple, Barro, 1991, Barro et Sala-i-Martin, 1991, Baumol, 1986, Cohen, 1992, Mankiw, Romer et Weil, 1992, Sala-i-Martin, 1995). Deux types de tests ont été proposés, le test de *b*-convergence qui étudie le comportement de retour à la moyenne d'un ensemble de variables, et les analyses de *s*-convergence qui étudient l'évolution de la dispersion de la coupe transversale. Ces tests étudient la convergence en moyenne transversale.

b-convergence

Partant de l'idée que les pays pauvres rattraperont les pays riches s'ils réalisent un taux de croissance supérieur, la procédure de test consiste à régresser le taux de croissance de la série sur le niveau initial de la série :

$$(8) \quad (\ln(y_{jt}) - \ln(y_{j0}))/t = \mathbf{a} - \mathbf{b}.\ln(y_{j0}) + \mathbf{e}_{jt}$$

On conclut à la convergence si le coefficient du niveau initial, $-\mathbf{b}$, est négatif. Ce test est aussi connu sous le nom de test de *b*-convergence².

2. Le test a été étendu au cas de convergence conditionnelle (voir parmi de multiples exemples, Barro, 1991, Barro et Sala-i-Martin, 1991, Mankiw et al., 1992). Plusieurs variables

Deux observations temporelles de la coupe transversale, en 0 et en t , suffisent pour effectuer le test. Les revenus s'écartent d'autant plus vite des conditions initiales — c'est-à-dire que l'écart entre y_t et y_0 est d'autant plus grand — et par conséquent les pays convergent d'autant plus rapidement, que b est grand. La vitesse de convergence, f , peut être calculée à partir de l'équation : $b = (1 - e^{-f \cdot t}) / t$.

Les inconvénients de ce test ont été longuement discutés dans la littérature. Les critiques les plus sévères ont été adressées par Quah dans une série d'articles (1993a, b, 1994, 1996) ; elles portent sur l'interprétation des résultats et remettent sérieusement en cause l'utilisation du test de b -convergence. Premièrement, Quah critique le manque de clarté dans l'interprétation du résultat de rattrapage. En effet, si les pays pauvres croissent plus vite que les pays riches, cette situation peut conduire à un dépassement des pays riches par les pays pauvres, aussi bien qu'à la convergence à une constante près ou qu'à la convergence vers le même niveau. Deuxièmement, la b -convergence ne donne aucune information sur l'évolution de la dispersion de la coupe transversale, ni sur d'autres éléments de la distribution (le nombre de modes par exemple). Troisièmement, l'estimateur de b est biaisé vers zéro lorsque les séries sont non stationnaires, ce qui est très probablement le cas avec des séries de revenu ou de revenu par tête. Enfin, la critique la plus sévère est que le test ne peut fournir aucune information pertinente sur le processus de convergence car ce test souffre d'une erreur de Galton³. Quah (1993a) démontre comment un coefficient de revenu initial négatif peut être associé avec une distribution de dispersion constante voire croissante.

2. Démonstration de l'erreur de Galton (Quah, 1993a)

Cas où la dispersion est constante

Réécrivons l'équation (8), avec $Y_t = \ln(y_{jt})$ et $Y_0 = \ln(y_{j0})$, comme :

$$Y_t = c + d \cdot Y_0 + e_t$$

L'estimateur des moindres carrés de d est

$$\hat{d} = \frac{\text{cov}(Y_t, Y_0)}{\text{var}(Y_0)}$$

sont ajoutées comme régresseurs, afin de prendre en compte l'hétérogénéité des états stationnaires entre pays. Dans ce cas, chaque pays converge mais vers son propre état stationnaire ; puisque les états stationnaires sont différents, les pays divergent.

3. Le paradoxe de Galton est très ancien. Il peut être résumé comme suit : on observe qu'en moyenne les enfants de parents de grande taille sont plus petits que leurs parents. Ceci indique un comportement de retour à la moyenne, et laisse supposer que la taille des personnes d'une population va converger vers le même niveau. Pourtant, la dispersion des tailles ne se réduit pas au cours du temps. Par conséquent, la tendance à retourner à la moyenne ne garantit pas que la dispersion de la distribution se réduise.

Si la coupe transversale est stationnaire, la variance de la coupe est constante au cours du temps, $\text{var}(Y_\rho) = \text{var}(Y_t)$, et l'estimateur de \hat{d} devient :

$$\hat{d} = \frac{\text{cov}(Y_t, Y_0)}{\sqrt{\text{var}(Y_0)} \cdot \sqrt{\text{var}(Y_t)}}$$

L'inégalité de Cauchy-Schwartz est $\|x \cdot y\| \leq \|x\| \cdot \|y\|$, pour $\|z\|$ la norme de z . En choisissant $\sqrt{E[(z_t - \bar{z}) \cdot (z_t - \bar{z})]}$ comme norme on a :

$$\text{cov}(Y_t, Y_0) \leq \sqrt{\text{var}(Y_0)} \cdot \sqrt{\text{var}(Y_t)}$$

Par conséquent, $\hat{d} \leq 1$, et dans l'équation (8), $-b \leq 0$

Cas où la dispersion est croissante

Lorsque la dispersion est croissante, l'estimateur de b peut toujours être négatif. Si on suppose, par exemple, que la dispersion de la distribution croît au taux I supérieur à un : $I \cdot \text{var}(Y_\rho) = \text{var}(Y_t)$, l'estimateur de \hat{d} devient

$$\hat{d} = \frac{\text{cov}(Y_t, Y_0)}{\sqrt{\text{var}(Y_0)} \cdot \sqrt{\text{var}(Y_t)}} \cdot \sqrt{I}$$

Par l'inégalité de Cauchy-Schwartz, on a donc $\hat{d} \leq \sqrt{I}$, et $-b \leq \sqrt{I} - 1$

La pertinence de l'erreur de Galton dépend donc de l'écart entre $\text{cov}(Y_t, Y_0)$ et $\sqrt{\text{var}(Y_0)} \cdot \sqrt{\text{var}(Y_t)}$ et de la valeur de I .

Les critiques émises par Bernard et Durlauf (1991) portent sur le fait que le test de b -convergence ne tient pas compte du caractère non stationnaire des séries de revenu. Considérer le taux de croissance annuel moyen du revenu, revient à faire l'hypothèse implicite que le taux de croissance est constant, c'est-à-dire que la tendance est déterministe; cette modélisation n'est pas appropriée si la tendance est un processus stochastique⁴. Tester la convergence de cette manière revient à supposer que la dispersion entre les séries résulte uniquement d'un choc initial, et que les séries ne sont pas affectées par des chocs aléatoires à chaque période.

L'équation (8) peut être dérivée du modèle de croissance de Solow. En introduisant le capital physique et le taux de croissance de la population, elle permet alors d'estimer les coefficients de la fonction de production macroéconomique. Mankiw et *al.* (1992) montrent que le taux de convergence généralement retenu (2 % par an) est incompatible avec les données de part du capital dans l'output. Ceci résulte de l'omission

4. La littérature sur les processus à racine unitaire montre bien les différences conceptuelles, de modélisation et d'interprétation économique entre une tendance déterministe et une tendance stochastique (voir par exemple, à ce sujet, Hamilton, chapitre 15). Le caractère non stationnaire des séries de revenu a été mis en évidence à plusieurs reprises dans la littérature, et peut s'interpréter comme résultant du fait que le progrès technique, à la source de l'évolution du revenu, est un processus aléatoire plutôt que déterministe.

de certaines variables, en particulier le capital humain. La prise en compte de ces variables implique que les pays convergent chacun vers son état stationnaire mais vers des états stationnaires différents, c'est-à-dire qu'ils divergent. Canova et Marcet (1995) et Desdoigts (1997), sur base de méthodes différentes montrent également que les pays convergent, mais vers des états stationnaires différents.

Enfin, ce genre de test est soumis aux critiques usuelles des coupes transversales, les coefficients sont identiques pour tous les pays, ce qui peut s'avérer discutable lorsqu'on compare une centaine de pays, développés et en développement. Ce type de test fournit une information moyenne sur la coupe transversale, négligeant les comportements individuels. Ceci peut conduire à un résultat de divergence dès qu'un sous-ensemble de séries diverge, malgré une convergence des autres séries de l'échantillon.

Certaines de ces critiques peuvent être rencontrées par l'utilisation d'un autre mode d'analyse pour coupes transversales, les analyses de *S*-convergence.

S-convergence

L'analyse de la *S*-convergence consiste à étudier l'évolution de la dispersion de l'ensemble de la coupe transversale. On conclut à la convergence de l'ensemble de l'échantillon lorsque la dispersion diminue dans le temps. Barro et Sala-i-Martin (1991), par exemple, regardent l'évolution de l'écart-type en coupe ou du coefficient de variation⁵.

Quah (1993a) a montré que la *b*-convergence n'implique pas nécessairement la *S*-convergence. Hénin et Le Pen (1995) montrent que ceci peut résulter de l'occurrence de chocs aléatoires qui maintiennent constante ou croissante la dispersion de la distribution. La *b*-convergence traduit un comportement de retour à la moyenne, les séries retournent naturellement vers leur moyenne suite à un choc aléatoire. Lorsque les séries ont le même état stationnaire, la même moyenne, cette propriété a tendance à ramener toutes les séries vers le même niveau, donc à réduire la dispersion de l'ensemble des séries. Cependant l'occurrence de chocs aléatoires distincts pour chaque série à chaque période fait dévier de cette tendance et donc limite la réduction de la dispersion.

Hénin et Le Pen (1995) et Hart (1995) proposent une autre façon d'analyser la *S*-convergence, plutôt que de se baser sur l'évolution de coefficients de dispersion, un test de réduction de la dispersion peut être effectué. La procédure se base sur l'équation suivante :

5. Pour des variables à tendance, le coefficient de variation (rapport de l'écart-type à la moyenne) est une mesure plus appropriée de la dispersion puisqu'elle ne dépend pas de l'unité de mesure. L'écart-type sous-évaluera la *S*-convergence dans le cas de séries croissantes.

$$(9) \ln(y_{jt}) - \ln(y_{j0}) = a - b \ln(y_{j0}) + e_{jt}$$

où comme précédemment, on s'attend à ce que le coefficient du niveau passé soit négatif. Il y a **S**-convergence si $(1-b)^2/R^2 < 1$, avec R^2 , le coefficient de corrélation multiple entre $\ln(y_{jt})$ et $\ln(y_{j0})$ de l'équation (9). Ceci peut être testé à partir d'un simple test de Student; par exemple, pour $1-b$ positif cette condition s'écrit $b > 1-R$. Il faut donc que l'effet de convergence qui résulte de la propriété de retour vers la moyenne soit suffisamment important pour contrebalancer l'effet de divergence induit par l'occurrence de chocs aléatoires.

3. Lien entre **b**- et **S**-convergence (Hénin et Le Pen, 1995)

Considérons l'équation (9) :

$$(9) \ln(y_{jt}) - \ln(y_{j0}) = a - b \ln(y_{j0}) + e_{jt}$$

Il y a **b**-convergence si $-b$ est négatif. Il y a **S**-convergence si la variance de y_{jt} décroît dans le temps.

L'équation (9) peut se réécrire comme :

$$\ln(y_{jt}) = a + (1-b) \ln(y_{j0}) + e_{jt}$$

En prenant la variance de chacun des membres on trouve

$$\text{var}(\ln y_{jt}) = (1-b)^2 \text{var}(\ln y_{j0}) + \text{var}(e_{jt})$$

$$\hat{U} \quad \text{var}(\ln y_{jt}) / \text{var}(\ln y_{j0}) = (1-b)^2 + \text{var}(e_{jt}) / \text{var}(\ln y_{j0})$$

$$\hat{U} \quad \frac{\text{var}(\ln y_{jt}) - \text{var}(e_{jt})}{\text{var}(\ln y_{j0})} = (1-b)^2$$

$$\hat{U} \quad \frac{\text{var}(\ln y_{jt})}{\text{var}(\ln y_{j0})} - \frac{\text{var}(\ln y_{jt}) - \text{var}(e_{jt})}{\text{var}(\ln y_{jt})} = (1-b)^2$$

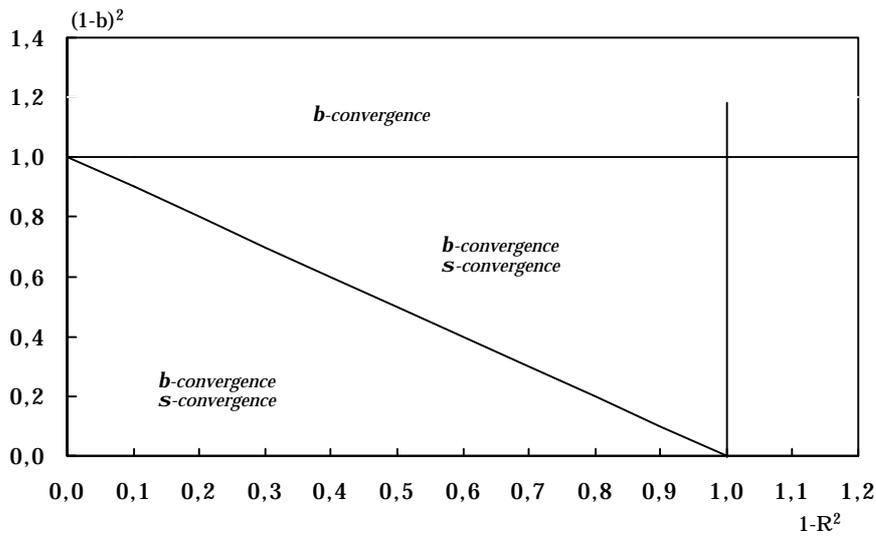
$$\hat{U} \quad \frac{\text{var}(\ln y_{jt})}{\text{var}(\ln y_{j0})} R^2 = (1-b)^2$$

avec R^2 le coefficient de corrélation multiple de l'équation (9). On a finalement :

$$\frac{\text{var}(\ln y_{jt})}{\text{var}(\ln y_{j0})} = \frac{(1-b)^2}{R^2}$$

Il y a **S**-convergence lorsque $\text{var}(\ln y_{jt}) / \text{var}(\ln y_{j0}) < 1$, c'est-à-dire lorsque $(1-b)^2 < R^2$

Hénin et le Pen (1995) résument cette condition par le graphique ci-dessous. Pour $(1-b)^2$ inférieur à 1, il y a **b**-convergence. Il y a de plus **S**-convergence en dessous de la droite d'équation $(1-b)^2 = 1 - (1-R^2)$.

5. Liens entre *b*- et *s*-convergence

Il reste néanmoins que la *s*-divergence tout comme la *b*-convergence sont des concepts relativement réducteurs de l'évolution de la distribution de la coupe transversale, la distribution étant résumée par deux indicateurs, la moyenne et la variance. D'autres méthodes fournissent des informations plus riches mais nécessitent des échantillons de dimension temporelle supérieure.

Un exemple d'application

Afin d'illustrer les méthodes ci-dessus les tests de *b*- et *s*-convergence sont appliqués aux séries de revenu par habitant des 144 régions de douze pays de l'Union européenne, pour un échantillon allant de 1980 à 1992⁶. L'estimation de l'équation (8) donne

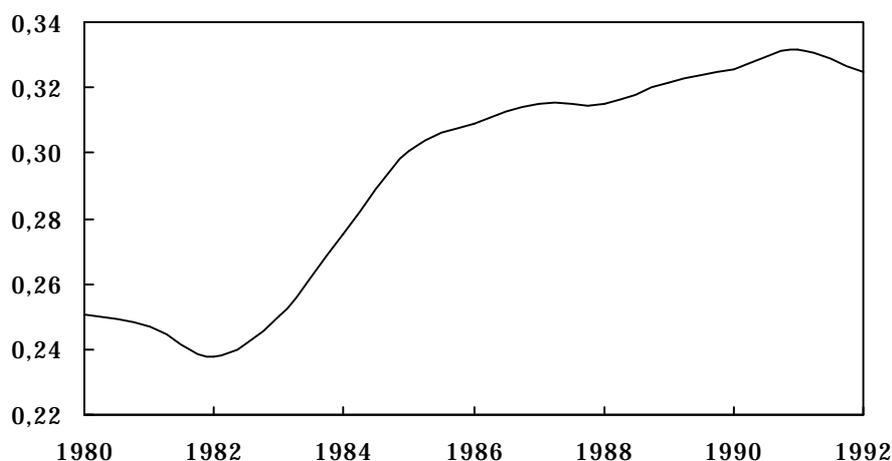
$$\begin{aligned} (\ln(y_{jt}) - \ln(y_{j0})) / t &= 163 - 0.0076 \ln(y_{j0}) + e_{jt} \\ \text{(erreur standard)} & \quad (105) \quad (0.0111) \end{aligned}$$

b est non significativement différent de zéro; on ne peut pas rejeter l'hypothèse de non convergence des revenus par habitant régionaux.

L'évolution du coefficient de variation des revenus par habitant régionaux est donnée au graphique 6. Il apparaît nettement que la dispersion des revenus par habitant régionaux, loin d'être stable, s'est creusée sur l'ensemble de la période, en particulier dans la première moitié des années quatre-vingt; il y a *s*-divergence.

6. Les données proviennent de la banque REGIO d'EUROSTAT. Le niveau de désagrégation est NUTS2.

**6. Coefficient de variation des revenus par habitant
de 144 régions européennes**



Sources : Regio-Eurostat et calculs de l'auteur.

La procédure de Hart permet de tester statistiquement cette hypothèse de *S*-divergence. L'estimation de l'équation (9) donne :

$$\ln(y_{jt}) - \ln(y_{j0}) = 2124 - 0,01 \cdot \ln(y_{j0}) + e_{j0}$$

(erreur standard) (1369) (0,14)

Le coefficient du revenu initial est comme précédemment non significativement différent de zéro. La racine carrée du coefficient de corrélation multiple entre $\ln(y_{jt})$ et $\ln(y_{j0})$, R , vaut 0.0058. La t-statistique pour l'hypothèse nulle $H_0 : b=1-R$ vaut -6,83. On rejette donc l'hypothèse de *S*-convergence, $b > 1-R$, puisque la t-statistique est significativement négative. Dans le cas des régions européennes, les trois critères fournissent des indications cohérentes et rejettent l'hypothèse de convergence.

Tests de convergence en panel

Les données de panel ont une dimension temporelle plus riche. Cette caractéristique peut être exploitée pour améliorer la précision des estimateurs, ou tenir compte de la dynamique des séries.

Islam (1995) discute l'utilité des méthodes de panel pour étudier la convergence. Les méthodes de panel améliorent la précision des estimateurs, permettent d'appréhender les problèmes de variables omises aux travers d'effets individuels. Dans le cadre de la convergence des revenus par habitant, ces effets individuels sont interprétés comme le

reflet de l'état de la technologie. Ils peuvent être représentés par des variables indicatrices par pays, comme chez Canova et Marcet (1995), ou être corrélés entre pays. Canova et Marcet (1995) estiment l'équation suivante :

$$(10) \quad y_{i,t} - y_{i,t-1} = c_i - \mathbf{b} \cdot y_{i,t-1} + \mathbf{e}_{it}$$

Les effets fixes, c_p , permettent d'introduire des états stationnaires différents. Les auteurs montrent qu'il y a \mathbf{b} -convergence. Mais ce résultat implique que chaque série converge vers un état stationnaire différent, qui peut être estimé par $\hat{c}_i / \hat{\mathbf{b}}$

Beine, Docquier et Hecq (1998) profitent de la dimension temporelle de l'échantillon pour tenir explicitement compte du caractère non stationnaire des séries. Ils réalisent des tests de racine unitaire en panel. Les tests de racine unitaire testent la stabilité de la série, c'est-à-dire l'hypothèse que les écarts à la moyenne sont stationnaires, autrement dit que les séries ont tendance à retourner à leur moyenne. Les auteurs se basent sur le modèle de Gaulier, Hurlin et Jean-Pierre (1997)⁷ :

$$(11) \quad \Delta(y_{i,t} - \bar{y}_t) = c_i - \mathbf{b}_i(y_{i,t-1} - \bar{y}_{t-1}) + \sum_{j=1}^{p_i} \Delta(y_{i,t-j} - \bar{y}_{t-j}) + \mathbf{e}_{it}$$

où y_t représente la moyenne transversale des séries. Le test revient à analyser l'hypothèse d'un comportement de retour à la moyenne, tout comme dans les tests de \mathbf{b} -convergence. L'avantage du test de racine unitaire est qu'il utilise des statistiques de test appropriées au caractère non stationnaire des séries. De plus, le modèle (11) permet de tester trois types de convergence : si $\mathbf{b}_i = \mathbf{b} \forall i$ avec $\mathbf{b} > 0$, et $c_i = c_j$ pour i différent de j , il y a convergence globale de l'ensemble de l'échantillon vers le même état stationnaire à la même vitesse; il y a convergence conditionnelle, et donc convergence vers des états stationnaires différents si $c_i \neq c_j$ pour i différent de j ; enfin, les vitesses de convergence diffèrent d'une série à l'autre si $\mathbf{b}_i \neq \mathbf{b}_j$ pour i différent de j .

Lorsque les séries sont stationnaires, la moyenne et la variance de ces écarts sont constantes. Dès lors, ni l'espérance de la différence entre les séries, ni leur dispersion ne varient dans le temps. Par conséquent cette méthode revient à tester que les séries ont convergé à une constante près, et non pas qu'elles convergent.

L'inconvénient de ce type de test repose, comme pour les coupes transversales, sur l'application de coefficients identiques pour toutes les séries, à moins d'autoriser des effets fixes et des vitesses de convergence propres à chaque série. Le modèle propose un test de convergence globale, vers la moyenne transversale. Par conséquent, dès qu'une série ou

7. Levin et Lin (1992) proposent un modèle à vitesses de convergence et structures dynamiques identiques entre les séries : $\mathbf{b}_i = \mathbf{b}$, $p_i = p \forall i$. Evans et Karras (1996) autorisent différentes vitesses de convergence tout en conservant une même structure dynamique pour toutes les séries : $p_i = p \forall i$.

un ensemble de séries diverge, le test peut indiquer une divergence globale. Une analyse plus fine de la distribution peut être réalisée dans l'analyse de l'évolution des distributions ou en séries chronologiques. Ceci est l'objet des sections suivantes.

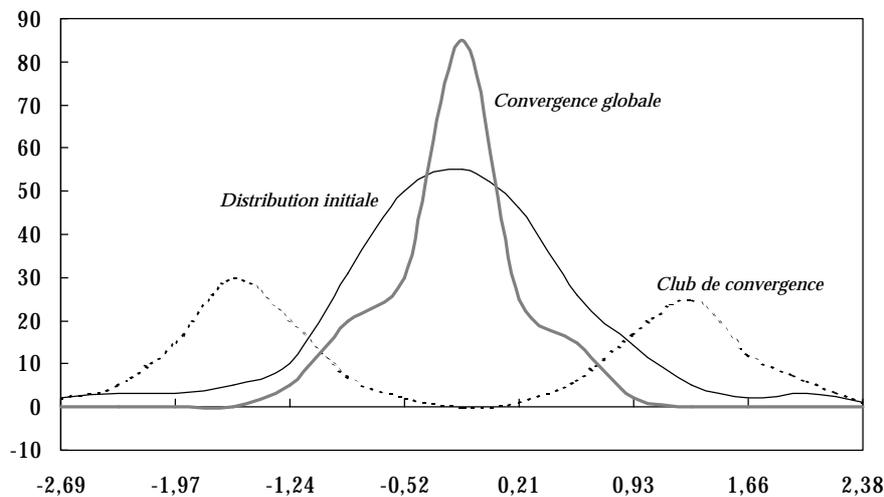
Tests de convergence en distribution

Quah a développé une méthode permettant d'analyser l'évolution de la distribution dans le temps. Desdoigts (1995) et Quah (1993a, b, 1994, 1996) ont utilisé cette méthode pour étudier la convergence des revenus par habitant.

Le graphique 7 expose deux modes d'évolution de la distribution au cours du temps, correspondant à deux phénomènes de convergence. De la distribution initiale, les économies peuvent évoluer vers une distribution resserrée, la dispersion diminue, tous les éléments de la distribution convergent vers le même niveau. La distribution peut également évoluer vers une distribution multimodale, indiquant plusieurs « clubs de convergence », plusieurs groupes de séries convergeant vers des niveaux différents. La dispersion de l'ensemble de la distribution peut rester constante ou même s'accroître, mais la dispersion à l'intérieur de chacun des groupes, se réduit.

L'analyse de l'évolution de la distribution d'un ensemble de séries peut se faire selon deux stratégies complémentaires. Comparer la distri-

7. Convergence et évolution de la distribution



Sources : Calculs de l'auteur.

bution en début et en fin de période, ou inférer la distribution limite, dite distribution ergodique, c'est-à-dire la distribution vers laquelle tend la distribution observée. Il y a convergence vers le même niveau si la distribution tend vers une distribution dégénérée en un point.

La distribution peut être estimée de façon discrète (Quah, 1993a,b) ou de façon non paramétrique (Desdoigts, 1995, Quah, 1994, 1996). L'évolution de la distribution est donnée par une matrice de transition, à partir de laquelle la distribution ergodique peut être estimée. Bianchi (1995) propose un test de multimodalité.

La méthode s'applique particulièrement aux panels à dimension temporelle importante. Le panel doit avoir une dimension transversale suffisante pour une estimation précise de la distribution, et une dimension temporelle suffisante pour estimer avec précision la dynamique de la distribution.

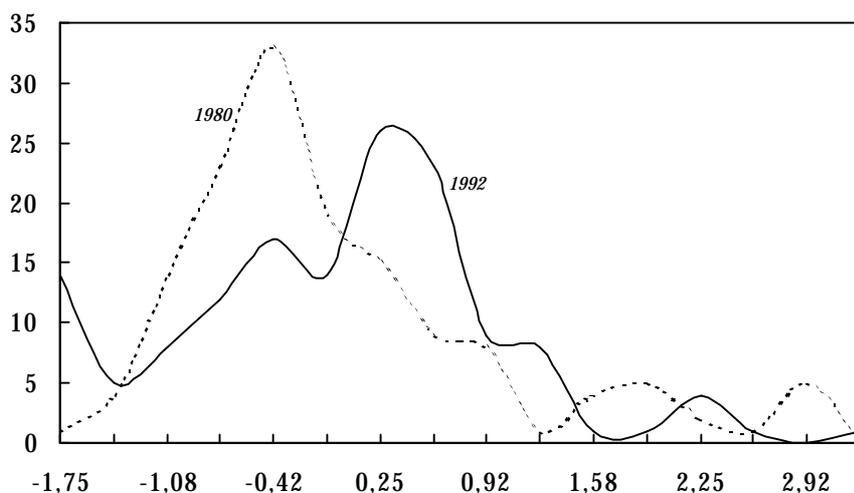
Les avantages de ce type d'analyse sont multiples. La méthode n'impose pas de structure *a priori* sur les relations entre les séries et permet de considérer tous les paramètres de la distribution. L'analyse permet de dégager des comportements spécifiques à l'intérieur de la distribution; en particulier, elle permet de détecter la présence de « clubs de convergence ». Elle permet de réaliser une analyse de long terme au travers de la distribution ergodique.

Cependant, si la méthode n'impose pas de structure *a priori* sur les relations entre les variables, elle impose une structure dynamique précise sur l'évolution de la distribution. La méthode est plus appropriée à la détection des clubs de convergence qu'à l'analyse de la convergence au sein des clubs, celle-ci peut se faire plus aisément par une analyse en séries chronologiques.

Un exemple : PIB réels par habitant régionaux en Europe

A titre d'illustration, le graphique 8 ci-dessous reporte les histogrammes des revenus par habitant régionaux des 144 régions de douze pays de l'Union européenne en 1980 et 1992. Les séries sont normalisées, c'est-à-dire de moyenne nulle et de variance unitaire, pour faciliter la comparaison des formes des deux distributions. L'évolution de la dispersion est donnée au graphique 6. Il n'y a pas eu sur la période, de tendance claire à la convergence. On peut éventuellement voir l'apparition d'une distribution bimodale.

8. Histogrammes des revenus réels par habitant de 144 régions européennes



Sources : Régio-Eurostat et calculs de l'auteur.

Tests de convergence en séries chronologiques

En ce qui concerne l'analyse de la convergence à long terme, les tests reposent usuellement sur des analyses de cointégration ou sur des modèles à paramètres variables. Une distinction est faite selon qu'on considère que le processus de convergence est stable ou qu'il évolue dans le temps à mesure que les séries convergent.

La convergence vue comme un processus de long terme

Dans le cadre de la problématique de l'intégration européenne, la convergence a fréquemment été étudiée comme un processus de long terme (voir, par exemple, Artis et Nachane, 1991, pour la convergence des salaires et des prix) et plus rarement dans le cadre de la croissance (voir Bernard et Durlauf, 1991, 1995, pour la convergence des revenus). Les tests se sont développés autour de la notion de cointégration.

La cointégration se rapporte aux propriétés de la relation entre deux (ou plusieurs) variables non stationnaires⁸. Deux séries non stationnaires, y_t^* et y_t sont cointégrées s'il existe une relation entre ces variables qui soit stationnaire, c'est-à-dire si il existe des coefficients a et b tels que le résidu, e_t , de l'équation suivante soit stationnaire :

8. Voir l'encadré 1. Pour simplifier on considère des variables stationnaires (faible) après une transformation en différence première.

$$(12) \quad y_t^* = a + b.y_t + e_t$$

Une telle relation est appelée une relation de cointégration.

Lorsque les séries ne sont pas cointégrées, elles peuvent s'écarter de plus en plus de la relation (12) puisque le résidu, e_t , est de variance croissante. Lorsque les séries sont cointégrées, leur évolution jointe peut être prévue avec une précision constante, les séries ne peuvent pas s'écarter de la relation de long terme au delà de marges constantes. La relation de cointégration est une relation d'équilibre de long terme : équilibre dans le sens où les chocs ne peuvent avoir que des effets temporaires sur la relation d'équilibre ; suite à un choc, les séries retournent vers cette relation ; long terme car elle mesure l'évolution jointe des séries affectées par le même choc permanent, indépendamment des perturbations temporaires ou de court terme.

La première condition de la convergence est que les séries soient cointégrées. La cointégration est une condition nécessaire mais non suffisante à la convergence. La convergence impose, de plus, des contraintes sur les paramètres de long-terme, a et b . Bernard et Durlauf (1991, 1995) imposent que b soit égal à un. Ceci revient à tester que les séries ont convergé, qu'elles ont atteint un équilibre où elles sont égales (à une constante, a , près), et non pas qu'elles sont en train de converger.

Deux stratégies peuvent être utilisées pour y remédier : utiliser une analyse à paramètres variables ou analyser plus précisément les coefficients a et b . Fuss (1998a) propose la classification suivante :

Convergence et divergence lorsque y_t et y_t^ sont cointégrés*

$a=0$ et $b=1$	les séries ont convergé
$a=0$ et $b=1$	les séries ont convergé à une constante, a , près
$a>0$ et $b<1$	les séries convergent (pour y_t et y_t^* croissantes, y_t rattrape y_t^*)
$a<0$ et $b>1$	les séries convergent (pour y_t et y_t^* croissantes, y_t^* rattrape y_t)
$a>0$ et $b>1$	les séries divergent (y_t reste en dessous de y_t^*)
$a<0$ et $b<1$	les séries divergent (y_t^* reste en dessous y_t)

Analyser la convergence au travers de la cointégration revient à analyser la convergence à long terme. Le critère de convergence ainsi défini se situe à mi-chemin entre le concept de convergence en espérance et le concept de convergence en probabilité, définis plus haut. La relation de convergence ne doit pas être satisfaite en tout point ; les séries peuvent dévier temporairement de la relation d'équilibre suite à des chocs économiques transitoires. L'accent mis sur la relation de long terme peut avoir des interprétations économiques pertinentes, par exemple dans le cadre de la convergence des revenus par tête où la relation de cointégration peut être interprétée comme le reflet de l'état stationnaire. Enfin,

une telle analyse a l'avantage de prendre en considération un grand nombre de schémas de convergence (et de divergence). Par contre les paramètres sont fixes dans le temps.

Si on réalisait une analyse de cointégration sur les séries présentées dans les graphiques de la section introductive, avec la série A dans le membre de gauche de l'équation (12), on trouverait les résultats suivants : les séries A et C convergent sur l'ensemble de la période, $a > 0$ et $b < 1$. Les séries A et D divergent sur l'ensemble de la période, $a > 0$ et $b > 1$; A et F divergent sur la période [K, T], $a = 0$ et $b < 1$. Les séries A et G ne sont pas cointégrées. Les paramètres des relations entre les séries A et B et A et E ont varié sur l'ensemble de la période. Sur la période [0, K], les séries convergent, $a > 0$ et $b < 1$. Sur la période [K, T], les séries ont convergé, $b = 1$. Un modèle à paramètres variables est nécessaire pour tenir compte de ce type d'évolution.

La convergence vue comme un processus évolutif

Le processus de convergence peut être un processus plus complexe qu'un processus linéaire tel celui décrit dans la section précédente. Au fur et à mesure que les séries sont en train de converger, un ou plusieurs paramètres peuvent varier. Cette section passe en revue quatre façons de tenir compte de ces évolutions.

Gundlach (1993) utilise un modèle de cointégration linéaire construit de façon à ce que l'élasticité de y_t relativement à y_t^* tende vers l'unité lorsque y_t tend vers y_t^* :

$$(13) \quad y_t/y_t^* = a + b.y_t^* + e_t$$

Dès lors, y_t converge vers y_t^* (à une constante, a , près) si $b > 0$, et les séries ont convergé (à une constante, a , près) si $b = 0$. L'avantage de ce test est de conserver les propriétés de cointégration tout en restant dans un cadre d'analyse simple. L'inconvénient est qu'un seul paramètre est autorisé à varier au cours du processus de convergence.

Canarella, Polard et Lai (1990) ont utilisé un modèle de cointégration à paramètres variables, pour tester la parité de pouvoir d'achat, c'est-à-dire l'hypothèse que les taux de change bilatéraux ont convergé avec les prix relatifs. Après avoir rejeté l'hypothèse de cointégration dans un modèle linéaire usuel (tel que décrit dans l'équation (12)), ils estiment un modèle à paramètres variables, par filtre de Kalman :

$$(14) \quad \begin{aligned} y_t^* &= a_t + b_t y_t + e_t \\ a_t &= g_0 + g_1 a_{t-1} + u_t \\ b_t &= h_0 + h_1 b_{t-1} + h_t \end{aligned}$$

où e_t , u_t et h_t sont des variables stationnaires, d'espérance nulle et de variance constante.

Ils trouvent que b_t est constant et non significativement différent de l'unité. De plus, a_t varie dans le temps avec $g_1=1$. Enfin, les trois séries y_t , y_t^* et a_t sont cointégrées. Sous de telles hypothèses, le modèle (14) devient :

$$(15) \quad \begin{aligned} y_t^* &= a_t + y_t + e_t \\ a_t &= g_0 + a_{t-1} + u_t \end{aligned}$$

avec e_t stationnaire. Ils concluent que ceci confirme l'hypothèse de convergence, et que a_t est le choc permanent qui induit des déviations de la relation de convergence.

Cette interprétation me semble erronée. Premièrement, comme l'a souligné Honohan (1993) trouver un processus non stationnaire a_t tel que y_t^* , y_t et a_t soient cointégrés n'apporte pas de nouvelle information par rapport au fait que y_t et y_t^* ne sont pas cointégrés⁹.

Deuxièmement, le fait que la différence entre y_t et y_t^* soit fonction d'une série non stationnaire implique que la différence entre les séries peut dévier de plus en plus de zéro, car la variance de a_t (et donc la variance de la différence) croît dans le temps et tend vers l'infini¹⁰, comme dans le cas des séries A et G au graphique 4. Pour le dire autrement, le fait que la différence entre les séries dépende d'un choc permanent signifie que l'occurrence d'un tel choc fait dévier les séries de l'égalité de façon permanente. En conclusion, cette approche n'apporte pas d'information nouvelle sur le processus de convergence et conduit, me semble-t-il, à accepter de façon erronée l'hypothèse de convergence.

Hall, Robertson et Wickens (1993) utilisent un modèle où la différence entre les séries est fonction d'un processus évoluant dans le temps selon le modèle suivant, estimé par filtres de Kalman :

9. Considérons deux processus à racine unitaire, y_t^* et y_t intégrés d'ordre un :

$$y_t^* = y_{t-1}^* + u_t^*$$

$$y_t = y_{t-1} + u_t$$

avec u_t^* , u_t stationnaires. Si ces deux processus ne sont pas cointégrés, dans l'équation ci-dessous, le résidu e_t est un processus à racine unitaire intégré d'ordre un :

$$y_t^* = a + b.y_t + e_t$$

autrement dit :

$$e_t = e_{t-1} + x_t$$

avec x_t stationnaire. Dans ce cas, on peut aisément construire un processus a_t non stationnaire, cointégré avec y_t^* et y_t tel que les trois soient cointégrés. Par exemple, il suffit de construire a_t comme :

$$a_t = y_{t-1}^* - b.y_{t-1} + e_{t-1}$$

qui est bien un processus intégré d'ordre un. Et dans l'équation :

$$y_t^* = a_t + b.y_t + x_t$$

on aura x_t stationnaire. Autrement dit y_t^* , y_t et a_t sont cointégrés.

10. Pour $y_t^*.y_t = a_t + x_t$ avec $a_t = a_{t-1} + x_t$ et $S_x = \text{var}(x_t)$, on a $\text{var}(a_t) = \text{var}(y_t^*.y_t) = S_x.t$ (voir l'encadré 1).

$$\begin{aligned}
 (16) \quad & y_t^* - y_t = a_t + e_t && \text{avec } e_t \sim N(0, S^2) \\
 & a_t = a_{t-1} + u_t && \text{avec } u_t \sim N(0, W_t) \\
 & W_t = f \cdot W_{t-1}
 \end{aligned}$$

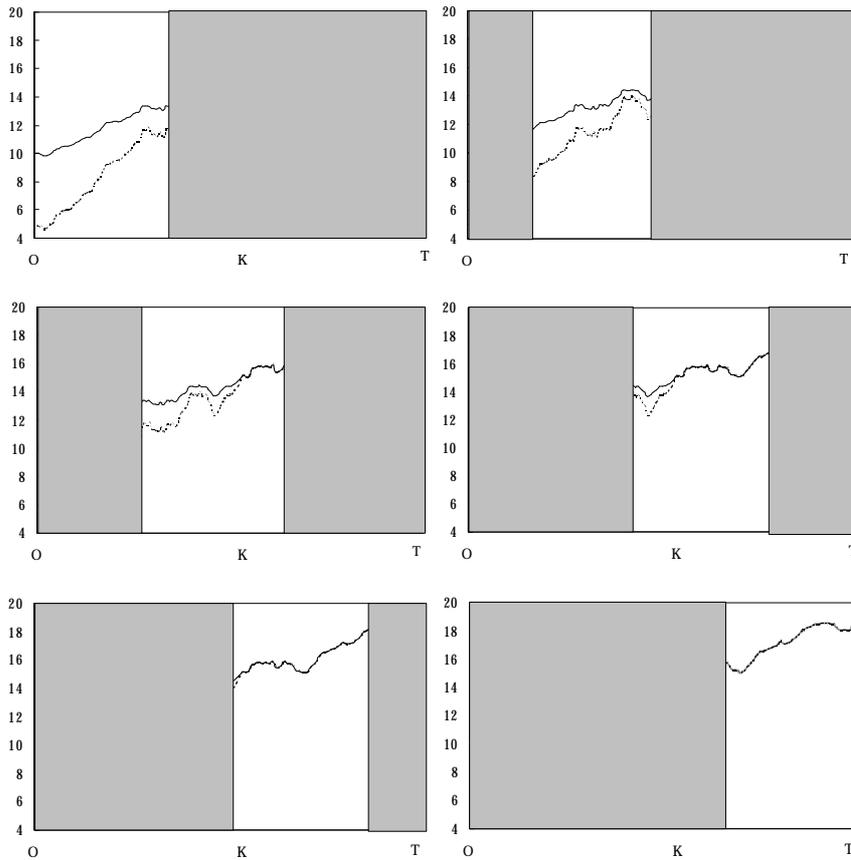
Il y a convergence si f est inférieur à un. Dans ce cas, en effet, la variance du bruit, u_t , tend vers zéro. Par conséquent, a_t tend vers une constante et la différence entre les séries est stationnaire autour d'une constante. Les séries peuvent s'écarter temporairement l'une de l'autre dans des bornes de fluctuations limitées (le concept de convergence se situe entre la convergence en espérance et la convergence en probabilité).

L'avantage d'une telle méthode est sa souplesse. Les paramètres varient librement en chaque point du temps. La relation entre les séries tend vers une relation de cointégration où les séries ont convergé (à une constante près). L'inconvénient du modèle est qu'il ne prend pas en compte tous les schémas de convergence. Par exemple, le modèle s'adapte mal au cas des séries A et C reportées au graphique 1. Ces séries sont cointégrées et leur différence décroît dans le temps de façon monotone, ce qui correspondrait dans le modèle (16) à un cas où a_t tend en décroissant vers zéro (ou une constante a) ; ceci n'est possible que si u_t est de moyenne négative qui contredit l'hypothèse du modèle que u_t est d'espérance nulle. La seule façon qu'ont les paramètres de tendre vers une constante dans ce modèle est par une diminution de leur variance, c'est-à-dire que les paramètres fluctuent de moins en moins autour de leur valeur limite, pour tendre vers cette valeur limite. Donc l'application du modèle aux séries A et C ne peut donner un coefficient a_t qui décroisse vers zéro, que lorsqu'une hypothèse du modèle ($u_t \sim N(0, W_t)$) est violée. Dans le cas contraire, le modèle conclut à la divergence.

Enfin, une dernière méthode repose sur l'analyse de cointégration récursive (Fuss, 1998a, b). Les tests récursifs sont fréquemment utilisés pour détecter la présence de paramètres variant dans le temps ou de ruptures dans l'évolution des séries. Dans le cas des tests de convergence, ils permettent de prendre en compte le fait que les paramètres peuvent varier au cours du temps, au cours du processus de convergence. La méthode consiste à appliquer les tests de convergence sur différents sous-échantillons, ou fenêtres, de même taille. Les graphiques 9 ci-dessous illustrent l'intérêt de la méthode dans le cas des séries A et B du graphique 1. La méthode permet de détecter le fait que les séries A et B sont en train de converger en début de période mais ont achevé le processus de convergence en fin de période.

Dans le cas des tests de convergence basés sur la cointégration, la méthode prend en compte différents types d'évolution du processus de convergence. Premièrement, la relation entre les séries peut évoluer vers une relation de cointégration lorsque la variance de la relation entre les séries évolue vers une constante. Deuxièmement, les paramètres de cette relation peuvent évoluer dans le temps.

9. Analyse réursive avec des fenêtres de taille constante



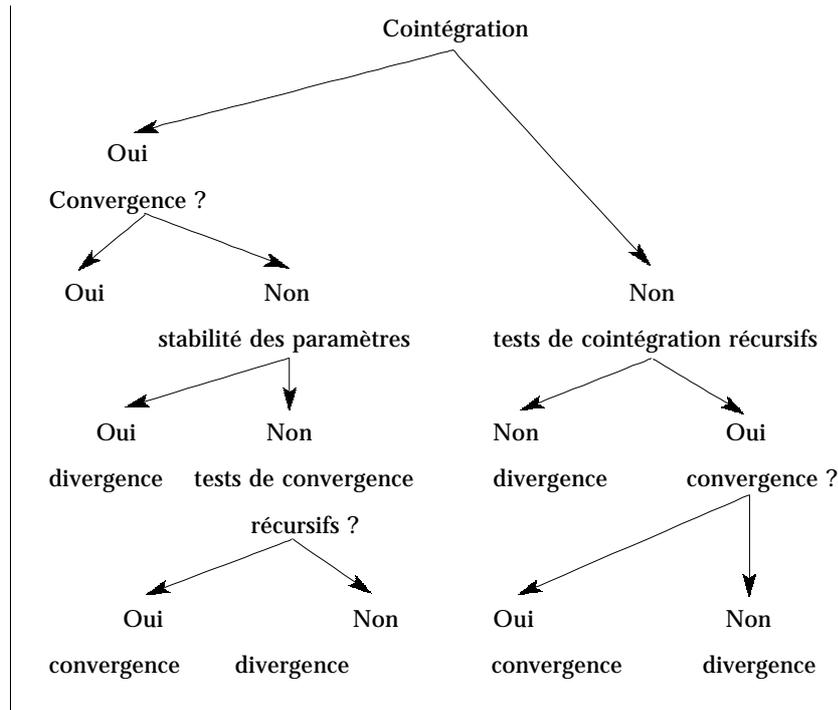
Sources : Calculs de l'auteur.

Fuss (1998a, b) procède en deux étapes. La première étape consiste à définir la période sur laquelle les séries sont cointégrées au moyen de tests de cointégration sur l'ensemble de la période et de tests récurifs. Ensuite, deux procédures complémentaires peuvent être utilisées. La première consiste à estimer les paramètres de long terme sur l'ensemble de la période où il y a cointégration, puis à réaliser des tests de stabilité des paramètres de façon réursive sur chaque sous-échantillon, c'est-à-dire, pour \bar{a} et \bar{b} les paramètres estimés sur l'ensemble de la période où il y a cointégration, à tester sur chaque sous-échantillon, l'hypothèse jointe $H_0 : a = \bar{a}$ et $b = \bar{b}$. Pour les échantillons où cette hypothèse est rejetée une analyse des paramètres de long terme peut être réalisée telle que décrite à la table 1. La seconde procédure se rapproche de la méthode de Bernard et Durlauf. Elle consiste à tester l'hypothèse $H_0 : b = 1$ sur chaque sous-échantillon où les séries sont cointégrées. Cette seconde procédure est moins riche mais plus facile à mettre en œuvre.

Elle permet de détecter les échantillons sur lesquels les séries ont convergé.

L'analyse de convergence en cointégration peut être réalisée selon le schéma ci-dessous. Après avoir analysé la convergence sur l'ensemble de l'échantillon, des tests récursifs sont utilisés pour vérifier que l'absence de convergence ne résulte pas d'une évolution de la relation entre les séries vers une relation de cointégration qui représente la convergence.

10. Stratégie de test de convergence en cointégration



Outre les avantages d'une analyse de convergence en cointégration — prise en compte explicite du caractère non stationnaire des séries, analyse du long terme, prise en considération de différents schémas de convergence — une telle procédure autorise tous les paramètres, coefficients et variance à varier dans le temps. Ceci permet de prendre en compte d'autres schémas de convergence impliquant des non linéarités dans un ou plusieurs paramètres. En termes de précision des estimateurs, si l'estimation récursive risque d'être moins précise car elle se fait sur une période plus réduite, elle peut être plus précise si elle est réalisée sur des périodes plus homogènes.

Les inconvénients résident dans le fait que la mise en œuvre de cette procédure est relativement lourde. La multiplication des tests (tests de racine unitaire, tests de cointégration, tests sur les paramètres) réduit la puissance globale de la procédure, ce qui est d'autant plus vrai pour l'analyse récursive. L'analyse bivariée peut donner des résultats incohérents. Cependant l'analyse peut être étendue au cas multivarié pour un petit nombre de séries¹¹. Une limite de l'analyse récursive est que si elle autorise bien les paramètres à varier entre deux sous-échantillons, ceux-ci sont constants à l'intérieur d'un sous-échantillon.

Exemple : l'analyse des taux d'intérêt à trois mois

Fuss (1998b) étudie la convergence des taux d'intérêt à trois mois par une analyse récursive en cointégration¹². L'analyse récursive permet de mettre en évidence le fait que le taux autrichien converge vers le taux allemand depuis 1972, jusqu'en 1993. Le taux norvégien converge vers les taux allemand et français depuis 1978.

De plus, l'analyse récursive permet de détecter l'effet de la crise sur les marchés de changes européens de septembre 1992 à août 1993. En effet, à partir de cette période, la première condition de convergence, l'existence d'une relation de cointégration, n'est plus satisfaite. Cette crise spéculative a donc déstabilisé les relations entre taux européens.

Les limites de cette analyse apparaissent particulièrement bien dans le cadre de l'étude de cette crise récente. Etant donné que les coefficients sont constants à l'intérieur d'un sous-échantillon, et que les sous-échantillons doivent être de taille suffisante pour mesurer le long terme, la procédure ne permet pas d'analyser la situation postérieure à la crise. Fuss (1998b) considère des sous-échantillons de quinze années, et ne peut donc pas tester la convergence dans la période d'après crise. Au contraire, Campa et Chang (1996), à l'aide d'une méthode différente et de données journalières sur les cours des options, montrent que les taux de change franc-mark, mark-livre sterling et mark-lire ont retrouvé leur crédibilité après la crise. Sur la base d'une méthode différente, sans analyser la convergence, Caporale, Hassapis et Pitis (1995) montrent, à l'aide de données mensuelles, que l'élargissement des bandes de fluctuations en août 1993 a éliminé les anticipations de réévaluations.

11. Le nombre de paramètres à estimer pour un système de variables est une fonction exponentielle du nombre de variables.

12. Artis et Zhang (1998) utilisent une analyse de cointégration récursive sans tester la convergence. Ils montrent que les liens entre taux d'intérêt européens et américain se sont relâchés, dans le sens où l'existence d'une relation de cointégration disparaît au cours des dernières années, tandis que le lien avec le taux d'intérêt allemand se resserre.

Conclusion

Cet article passe en revue différents tests de convergence utilisés dans la littérature empirique. Chaque test a, bien entendu, ses avantages et inconvénients. Plus fondamentalement, ils diffèrent par la question à laquelle chacun d'eux permet de répondre.

Le praticien est avant tout dépendant du type de données dont il dispose. Il aura à sa disposition des outils bien différents selon qu'il a à traiter un panel de dimension temporelle limitée ou étendue, ou un ensemble réduit de séries chronologiques. Certaines procédures peuvent être utilisées de façon complémentaire.

Les tests en coupes transversales et en panels fournissent une information moyenne, étudient le comportement moyen de l'ensemble des séries. Des conclusions de divergence peuvent apparaître dès qu'un sous-ensemble de séries divergent. De tels tests masquent les comportements individuels. En ce qui concerne ces tests, une analyse de *b*-convergence est très clairement insuffisante, puisque, comme l'a démontré Quah (1993a), une telle analyse est erronée. L'analyse de *s*-convergence seule ou conjointement à l'analyse de *b*-convergence est plus recommandée.

Les tests qui prennent en compte l'ensemble de la distribution sont nettement plus riches. Ils permettent de détecter des comportements différenciés à l'intérieur de la distribution, des clubs de convergence.

Les tests en séries chronologiques reposent généralement sur une analyse bivariée. Ceci permet de tenir compte des comportements individuels des séries. Ils peuvent prendre en compte de nombreux schémas de convergence, entre chaque paire d'éléments de l'ensemble considéré. L'application de tests de convergence dans une analyse de cointégration permet de tenir compte d'un grand nombre de schémas de convergence, mais leur mise en œuvre peut s'avérer relativement lourde. L'utilisation de modèles à paramètres variables laisse une grande souplesse dans l'évolution des paramètres, mais ne peut pas prendre en compte tous les schémas de convergence. Les analyses en cointégration récursive sont intermédiaires entre les analyses en cointégration et celles à paramètres variables. Elles sont limitées par le fait que les coefficients restent constants à l'intérieur des périodes d'estimation. Elles peuvent servir à tester l'existence de paramètres variables.

Références bibliographiques

- ARTIS M. J. e D.t NACHANE, 1991 : « Wages and prices in Europe : a test of the German leadership thesis », *Weltwirtschaftliches Archiv*, 1, 59-77.
- ARTIS M.J. et W. ZHANG, 1998 : « The linkage of interest rates within the EMS », *Weltwirtschaftliches Archiv*, 134, 117-132.
- BARRO R.J., 1991 : « Economic growth in a cross section of countries », *The Quarterly Journal of Economics*, CVI, 407-444.
- BARRO R.J. et X. SALA-I-MARTIN, 1991 : « Convergence accross states and regions », *Brooking Papers on Economic Activity*, 1, 107-182.
- BAUMOL W.J., 1986 : “Productivity growth, convergence and welfare : what the long-run data show”, *The American Economic Review*, 76, 1072- 1085.
- BEINE M., F. DOCQUIER et A. HECQ, 1998 : « Convergence des groupes en Europe : une analyse sur données régionales », Ministère de la Région Wallonne, Service des Etudes et de la Statistique, discussion paper n° 9801.
- BERNARD A.B. et S.N. DURLAUF, 1991 : « Convergence of international output movements », *NBER Working Paper* n° 3717.
- BERNARD A.B. et S.N. DURLAUF, 1995 : « Convergence of international output movements », *Journal of Applied Econometrics*, 10, 97-108.
- BIANCHI M., 1995 : « Testing for convergence : a bootsrap test for multimodality », *Working Paper, Bank of England*.
- CAMPA J.-M. et P.H. CHANG, 1996 : « Arbitrage-based tests of target zone credibility : evidence from ERM cross-rate options », *The American Economic Review*, 86, 726-740.
- CANARELLA G., S.K. POLLARD et K.S. LAI, 1990 : « Cointegration between exchange rates and relative prices : Another view », *European Economic Review*, 34, 1303-1322.
- CANOVA F. et A. MARCET, 1995 : « The poor stay poor : non convergence across countries and regions », *CEPR Workshop on Empirical Macroeconomics*, Brussels.
- CAPORALE G.M., C. HASSAPIS et N. PITIS, 1995 : « Excess returns in the EMS : do “weak” currencies still exist after the widening of fluctuation bands? », *Weltwirtschaftliches Archiv*, 131, 326-338.
- COHEN D., 1992 : « Test on the convergence hypothesis : a critical note », *CEPR Discussion Paper* n° 691.

- DESDOIGTS A., 1995 : « Changes in the world income distribution : a non-parametric approach to challenge the neo-classical convergence argument », *Humboldt-Universität zu Berlin, Quantifikation und Simulation Ökonomischer Prozesse*, Discussion Paper n° 15.
- DESDOIGTS A., 1997 : « Patterns of economic development and the formation of clubs », Université d'Evry-Val d'Essone, mimeo.
- EVANS P. et G. KARRAS, 1996 : « Convergence revisited », *Journal of Monetary Economics*, 37, 249-265.
- FUSS C., 1998a : « Convergence among industrialised countries : a time series investigation », *Cahiers Economiques de Bruxelles*, 158, 177-202.
- FUSS C., 1998b : « Contributions to the empirical analysis of convergence in the European Union », *thèse de doctorat*, Université Libre de Bruxelles.
- GAULIER G., C. HURLIN et P. JEAN-PIERRE, 1997 : « Conditional convergence : panel data approach », présenté à la 7^e conférence sur les données de panel, Paris, juin 1997.
- GUNDLACH E., 1993 : « Empirical evidence for alternative growth models : time series results », *Weltwirtschaftliches Archiv*, 129, 103-119.
- HALL S., D. ROBERTSON et M. WICKENS, 1993 : « How to measure convergence with an application to the EC economies », 1993 ESEM Meeting, Upsala.
- HART P., 1995 : « Galtonian regression across countries and the convergence of productivity », *Oxford Bulletin of Economic and Statistics*, 57, 287-293.
- HÉNIN P-Y. et Y. LE PEN, 1995 : « Les épisodes de la convergence européenne », *Revue économique*, 46, 667-677.
- HONOHAN P., 1993 : « Cointegration and time-varying parameters : A comment », *European Economic Review*, 37, 1279-1283.
- ISLAM N., 1995 : « Growth empirics : a panel data approach », *Quarterly Journal of Economics*, 110, 1127-1170.
- LEVIN A. et C.F. LIN, 1992 : « Unit root tests in panel data : asymptotic and finite sample properties », *Working Paper* n° 9223, University of San Diego.
- MANKIW N.G., D. ROMER et D.N. WEIL, 1992 : « A contribution to the empirics of economic growth », *Quarterly Journal of Economics*, CVII, 407-438.
- QUAH D., 1993a : « Galton's fallacy and tests of the convergence hypothesis », *CEPR Discussion Paper*, n° 820.

-
- QUAH D., 1993b : « Empirical cross section dynamics in economic growth », *European Economic Review*, 37, 426-434.
- QUAH D., 1994 : « One business cycle and one trend from (many,) many disaggregates », *European Economic Review*, 38, 605-613.
- QUAH D., 1996 : « Empirics for economic growth and convergence », *European Economic Review*, 40, 1353-1376.
- SALA-I-MARTIN X., 1995 : « The classical approach to convergence analysis », *CEPR Discussion Paper* n° 1254.
- STIGLER G.J. et R.A. SHERWIN, 1985 : « The extent of the market », *Journal of Law and Economics*, 28, 555-585.

