

ALAIN DESDOIGTS<sup>1</sup>

# VERS UNE CONVERGENCE GLOBALE OU LOCALE ?

**RÉSUMÉ.** Ces dix dernières années, un investissement important a été réalisé pour comprendre pourquoi les pays les plus défavorisés et les pays les plus riches ne convergent pas en terme de niveaux de revenus par tête ou par travailleur. Dans l'après-guerre, la distribution des revenus mondiaux a évolué vers une polarisation où les écarts entre les pays les plus pauvres et les plus riches ont tendu à s'accroître. Cette dynamique observée de la distribution des revenus va à l'encontre de mécanismes telles que la mobilité internationale du capital et la diffusion du progrès technique, invoqués dans un schéma de croissance purement néo-classique qui privilégie l'hypothèse de rendements décroissants dans l'accumulation des facteurs de production.

Pour expliquer la persistance, voire l'accroissement, des inégalités entre les nations, certains auteurs invoquent des différences dans les paramètres qui définissent l'état stationnaire d'une économie en présence de rendements décroissants, tels que

la préférence pour le présent, la démographie, ou encore les choix en termes de politiques économiques. D'autres privilégient la thèse selon laquelle ces inégalités résultent, toute chose égale par ailleurs, de différences dans la mise en place d'une dynamique de croissance ; ils suggèrent l'existence de non-linéarités dans les technologies avec des étapes où les rendements seraient croissants, qui succéderaient à des phases de développement où les rendements seraient décroissants. Enfin, un troisième et dernier groupe porte son attention sur le rôle des croyances et des anticipations des individus qui influenceraient le processus de développement d'une nation, et cela, même pour des économies qui seraient parfaitement identiques par ailleurs.

Partant d'un large échantillon de nations, cet article considère un ensemble de variables économiques qui rendent compte à la fois des circonstances initiales dans lesquelles le processus de croissance s'installe, et des paramètres qui définissent

1. Alain Desdoigts est maître de Conférences à l'Université d'Evry-Val d'Essonne et chercheur associé au centre d'Etudes politiques et économiques d'Evry. E-mail : Alain.desdoigt@eco.univ-evry.fr.

L'auteur tient à remercier Isabelle Bensedoun, Patrick Bolton, Fabio Canova, Olivier Cortes, Jérôme Glachant, Wolfgang Härdle, Sigbert Klinkke, Lucrezia Reichlin, Thierry Verdier, Robert Waldmann, ainsi qu'un rapporteur anonyme ; il remercie également pour leurs commentaires les participants aux séminaires organisés par le CORE (Université de Louvain), le CREST, le DELTA, et ECARE (Université Libre de Bruxelles), ainsi que des participants aux European Economic Association Meeting (Toulouse, 1997), Latin American and Caribbean Economic Association Meeting (Bogota, 1997), et Latin American Econometric Society Meeting (Santiago du Chili, 1997). Les erreurs et insuffisances sont de sa seule responsabilité.

l'état stationnaire d'une économie dans un schéma purement néoclassique. L'examen des sources de l'hétérogénéité économique met en évidence des "clubs" homogènes de nations et l'importance relative des variables impliquées dans ces regroupements. Ces clubs d'économies reflètent des différences dans les conditions initiales (stocks de capital physique et surtout humain) à partir desquelles s'est engagé le processus de développement, ainsi que la capacité des économies à développer, absorber et partager les nouvelles technologies. La relation entre le taux de croissance et la structure de groupes mise en évidence montre l'existence de fortes non-linéarités et d'effets de seuil qui suggèrent l'existence d'équilibres multiples à long terme, ce qui remet en cause l'hypothèse d'une technologie commune à l'ensemble des économies invoquée dans un schéma purement néoclassique.

Les pays appartenant à ces clubs ont des

caractéristiques communes tels que les arrangements institutionnels (pays de l'OCDE), le capital ethnique et culturel (protestant *versus* catholique), et la localisation géographique. Ces caractéristiques, difficilement quantifiables, émergent de manière endogène sur la base des variables économiques introduites dans le modèle. Ce résultat suggère l'importance, pour rendre compte de la mobilité des nations, du rôle des anticipations formulées par des groupes d'individus dont les choix seraient délibérés, voire auto-réalisateurs, reflète d'attitudes, de valeurs, et de connaissances qui seraient partagées au sein d'une société mais varieraient d'une société à l'autre. Ces choix semblent être largement influencés par une idéologie, qu'elle soit religieuse, ou qu'elle relève d'une tradition ethnique ou locale, pour ne citer que ces deux exemples qui ressortent de l'analyse.

*Classification JEL* : C14, O40.

Ces dix dernières années ont connu un débat intense sur l'hypothèse de la convergence à long terme des économies vers un niveau de revenu identique. Autrement dit, existe-t-il une tendance, ou au moins la possibilité, pour les économies les plus pauvres de rattraper celles qui bénéficient d'un niveau de revenu par tête plus élevé ? Ce débat s'est concentré sur la confrontation entre les enseignements de la théorie et les observations de la réalité <sup>2</sup>.

Dans le modèle de croissance original de Solow, les productivités marginales sont supposées décroissantes pour chaque facteur de production (Solow, 1956). Les économies en retard sur un groupe de pays *leaders* les rattrapent car le rythme des gains de productivité ralentit au fur et à mesure que l'on se rapproche de l'état stationnaire. Bien sûr, les pays diffèrent par leur technologie, leurs préférences, leur structure de marché, etc. Etant donnée une telle hétérogénéité, le cadre d'analyse néoclassique de la croissance implique que les différences initiales en termes de niveau de revenu par tête ont tendance à s'éliminer si et seulement si, elles ne proviennent pas de différences dans les technologies, les préférences, ou les structures de marché.

Traditionnellement, ce résultat théorique est testé sur une économie représentative composée d'un échantillon de nations en coupe transversale. On régresse le

2. Cette hypothèse de convergence a été formulée dans le contexte d'un scepticisme marqué à l'égard du rattrapage des pays en voie de développement. Elle n'est en aucun cas une orthodoxie ancienne, au contraire, elle a elle-même marqué une rupture notable dans les théories de la croissance et l'économie du développement.

taux de croissance annuel moyen sur le niveau de revenu par tête initial tout en contrôlant les différences en termes d'état stationnaire. Toutes choses égales par ailleurs, les différences de niveau de revenu par tête ont, en moyenne, tendance à disparaître si le coefficient qui associe le taux de croissance et le niveau de revenu par tête initial, communément dénoté  $\beta$ , est significativement différent de zéro et négatif. On parle alors de  $\beta$ -convergence conditionnelle. Le concept de convergence conditionnelle se distingue de celui de convergence absolue. Dans le premier cas, chaque économie tend vers son propre état stationnaire et le phénomène de rattrapage ne s'observe que pour des économies identiques par ailleurs. Ainsi, on admet qu'il puisse y avoir des inégalités persistantes au sein d'un large échantillon d'économies. Dans le second, l'inégalité initiale se résorbe de toute façon, et les économies les plus pauvres initialement rattrapent, en moyenne, les plus riches. Certes, quel que soit le concept adopté, des chocs aléatoires peuvent perturber ce processus de rattrapage, mais leur effet reste transitoire. Ces deux concepts de  $\beta$ -convergence se distinguent également de la réduction de la dispersion des niveaux de revenu par tête au cours du temps au sein d'une large coupe d'économies ( $\sigma$ -convergence).

Deux critiques peuvent être formulées à l'égard de ces concepts. Premièrement, les régressions en coupe ne fournissent des informations que sur le comportement moyen dans la coupe transversale, occultant ainsi l'information sur des phénomènes à l'œuvre dans la réalité, et en particulier, sur la mobilité des nations. Deuxièmement, il importe de comprendre dans l'étude de la convergence, comment les économies se comportent les unes par rapport aux autres, ce qu'aucun des concepts décrits ci-dessus ne permet vraiment.

Pour y parvenir, on peut étudier les dynamiques inhérentes à la distribution des niveaux de revenus par tête et regarder si cette distribution tend à se concentrer vers une seule et même valeur, ou si, au contraire, elle se stratifie autour de certaines valeurs. Apparaissent ainsi d'éventuels clubs de convergence dont l'existence est occultée par les concepts de  $\beta$  et  $\sigma$ -convergence. Le test de la convergence conditionnelle permet d'identifier un ensemble de caractéristiques économiques servant de proxy à l'état stationnaire vers lequel tend une économie-type. Cependant, il ne permet pas de conclure sur la croissance relative des économies les plus pauvres, sur l'évolution de la distribution des revenus dans le temps, ou encore sur la mobilité de l'ensemble des nations dans le cadre de la mondialisation.

Si l'analyse de la dynamique sous-jacente à la distribution des revenus met en évidence une stratification du monde, elle ne permet, ni d'expliquer, ni de fonder l'existence des clubs de convergence qui en résultent, sur les caractéristiques propres à chaque économie. En particulier, cette stratification peut résulter de l'hétérogénéité micro-économique entre les nations de l'échantillon, auquel cas, elle est parfaitement compatible avec un cadre de croissance néoclassique. Par exemple, si deux économies diffèrent uniquement de par leurs préférences intertemporelles, elles seront caractérisées par des taux d'épargne et des niveaux de revenus par tête différents à l'état stationnaire. On observera alors une stratification de la distribution des revenus, sans pour autant pouvoir rejeter l'hypothèse d'une technologie commune à rendements décroissants pour l'ensemble des

économies<sup>3</sup>. Au contraire, si la stratification ne dépend que des conditions initiales, toutes choses égales par ailleurs, c'est-à-dire, uniquement des circonstances dans lesquelles la dynamique de croissance s'installe, nous pouvons conclure en faveur de l'existence d'équilibres multiples à long terme, c'est-à-dire en faveur de l'hypothèse d'une convergence locale.

La nouvelle théorie de la croissance s'attache à modéliser ce type de mécanismes. Par exemple, Romer introduit des économies d'échelles qui, si elles sont suffisamment importantes et externes à la firme, mais internes à l'industrie prise dans son ensemble, génèrent des rendements d'échelle sociaux croissants (Romer, 1986). La difficulté à accéder à la phase préliminaire de développement nécessaire pour des gains de productivité importants peut alors expliquer la croissance plus rapide et le cumul d'inégalités entre des économies relativement prospères et des économies relativement pauvres. Lucas suggère que ces externalités ne s'étendent pas forcément à toutes les formes d'activités : elles peuvent se concentrer au sein de zones géographiques telles que les villes ou les régions, voire même au sein de coalitions politiques ou institutionnelles (Lucas, 1988). Elles sont telles qu'un certain niveau de développement est nécessaire pour observer le rattrapage des économies en retard sur les innovations des nations les plus avancées technologiquement<sup>4</sup>.

Cet article identifie, entre l'hétérogénéité micro-économique et les conditions initiales, ce qui conduit des clubs de pays à être localisés dans différents champs d'attraction vers un même état stationnaire. Dans un espace multidimensionnel qui tient compte d'un large ensemble de conditions initiales et de variables économiques émergent des clubs d'économies qualifiés de "clubs de convergence conditionnels", illustration des différences dans les stocks initiaux de capital physique et humain, ainsi que dans la capacité des économies à absorber et à utiliser les nouvelles technologies. La formation de ces clubs reflète également les arrangements institutionnels, le capital ethnique et culturel, ainsi que la localisation géographique qui émergent de manière endogène, sans avoir fait l'objet d'une spécification *ex ante*. Il existe bien un déterminisme culturel qui influence directement la vie et les comportements économiques, et qui ne se limite pas à un épiphénomène dans l'environnement économique. Sur la période considérée, ces clubs sont compatibles avec la mobilité observée des nations. Enfin, la relation entre le taux de croissance et les clubs de nations mis en évidence, suggère l'existence de non linéarités et d'effets de seuils qui nous permettent d'invoquer l'existence d'équilibres multiples à long terme.

3. Il en sera de même, par exemple, si les économies sont caractérisées par des taux de croissance différents de la population ou de la population active.

4. Pour une présentation exhaustive des inférences possibles à partir des modèles théoriques et concernant l'hypothèse de convergence, se référer par exemple à Galor (1996) et Azariadis (1996).

# Les limites de l'approche néo-classique de la convergence

L'approche empirique classique de l'hypothèse de convergence dit qu'au sein d'un large échantillon de pays, il y a rattrapage si le taux de croissance des revenus par tête ou par travailleur est inversement corrélé avec le niveau de revenu initial. Autrement dit, dans une régression en coupe transversale, un coefficient significativement négatif du niveau de revenu indique une forme de décroissance des rendements de la fonction de production exprimée en termes soit de produit par tête, soit de produit par travailleur. A l'origine de cette thèse, les travaux d'Abramowitz (1986), de Baumol (1986), ainsi que de Baumol et Wolff (1988) qui ont répondu à la critique formulée par De Long (1988) concernant le biais lié à la sélection de l'échantillon. Dowrick et Nguyen suggèrent que la convergence pour les pays industrialisés est relativement robuste à la sélection de l'échantillon, mais qu'elle ne s'étend pas aux nations les plus pauvres, au moins en termes d'une régression simple entre taux de croissance annuel moyen et niveau de productivité initial (Dowrick & Nguyen, 1989). Ainsi, l'hypothèse de convergence absolue doit être rejetée pour un large échantillon de pays. Dans le même temps, la dispersion des revenus par tête a tendance à s'accroître. Ces faits stylisés vont clairement à l'encontre d'une convergence globale. Cependant, la seule persistance, voire l'accroissement des inégalités, ne permet pas de rejeter l'idée d'une technologie commune à l'ensemble des pays.

Mankiw, Romer et Weil montrent ce que la notion de convergence désigne dans le modèle de croissance néo-classique (Mankiw, Romer & Weil, 1992). Les économies en retard vont croître plus vite que les économies les plus avancées, si et seulement si, elles sont identiques par ailleurs. L'idée sous-jacente de la convergence conditionnelle est que les analystes s'attendent à ce que chaque économie converge vers son propre état stationnaire. Ce dernier est alors identique pour un ensemble d'économies homogènes, c'est-à-dire ne différant que par leurs conditions initiales. L'hypothèse nulle d'un coefficient  $\beta$  non significatif étant rejetée, le modèle de croissance néoclassique tend à apparaître un peu comme la "fin de l'histoire". Parallèlement, cela conduit à rejeter l'hypothèse alternative dans laquelle toutes les idées proposées dans les nouveaux modèles de croissance sont implicitement incorporées. Autrement dit, sur la base des résultats de Barro (1991), Barro et Sala-i-Martin (1992), ainsi que de l'étude citée ci-dessus<sup>5</sup>, on invalide et on exclut implicitement bon nombre des idées originales, voire prometteuses, que l'on peut trouver dans la nouvelle littérature théorique sur la croissance, ce qui semble pour le moins, plutôt restrictif.

En effet, il est plutôt surprenant que la littérature empirique se soit concentrée sur une hypothèse qui vise à rejeter toutes les alternatives proposées au modèle de croissance néoclassique, alors que la littérature théorique s'est, elle, concentrée sur la possibilité pour des groupes d'économies d'appartenir à différents champs d'attraction vers un même état stationnaire. Sans aucun doute, la force de la

5. Pour une revue de cette littérature et des faits stylisés qui lui sont associés, voir de la Fuente (1997).

convergence conditionnelle provient du fait que cette approche dérive explicitement du modèle de croissance néoclassique. Néanmoins, cette approche a reçu des objections à la fois méthodologiques, statistiques et conceptuelles. En particulier, il se peut qu'il existe des non linéarités dans les technologies dont dispose chaque économie. Cette éventuelle caractéristique est occultée par l'approche traditionnelle qui, implicitement, fait l'hypothèse d'une spécification linéaire commune à toutes les économies.

## La dynamique à l'œuvre dans l'évolution de la distribution des revenus

Quah propose une approche alternative aux traditionnelles régressions en coupe ; il se concentre sur les dynamiques inhérentes à l'évolution de la distribution des niveaux de revenus par tête ou par travailleur et sur la mobilité des nations (Quah, 1993, 1996) <sup>6</sup>. Il fait remarquer que ni la  $\beta$ -convergence, ni la  $\sigma$ -convergence ne permettent de mettre en évidence l'éventuelle existence de clubs de convergence. Une insuffisance de ces régressions en coupe tient au fait qu'elles ne fournissent des informations que sur le comportement moyen dans la coupe transversale. De plus, pour appréhender la convergence, il est plus intéressant de comprendre comment les économies se comportent les unes par rapport aux autres, que d'identifier un ensemble de variables qui vont servir de proxy à l'état stationnaire vers lequel tend une économie-type, pour tester ensuite si le taux de croissance tend à décroître lorsque cette économie tend vers cet état stationnaire.

La littérature empirique analyse si la distribution des revenus tend à se concentrer vers une seule et même valeur, ou, au contraire, à se stratifier autour de certains niveaux de revenus. Elle cherche si les données apportent des preuves de l'existence d'une trappe à pauvreté et/ou de clubs de convergence, comme l'a suggéré Baumol (1986).

L'évolution de la distribution des niveaux de revenus par tête, sur la période 1960-1985, montre la pertinence de cette approche (GRAPHIQUE 1) <sup>7</sup>. Ce graphique donne des estimations de la densité des logarithmes des niveaux de revenus par tête <sup>8</sup>. La transformation logarithmique se justifie parce que, dans le modèle de croissance néoclassique avec progrès technique exogène, la convergence est attendue en termes relatifs.

Alors qu'en 1960, la distribution est unimodale, une structure multimodale apparaît en 1985. Un grand nombre de pays qui, en début de la période, se concentraient autour de la moyenne, ont transité vers les queues de la distribution,

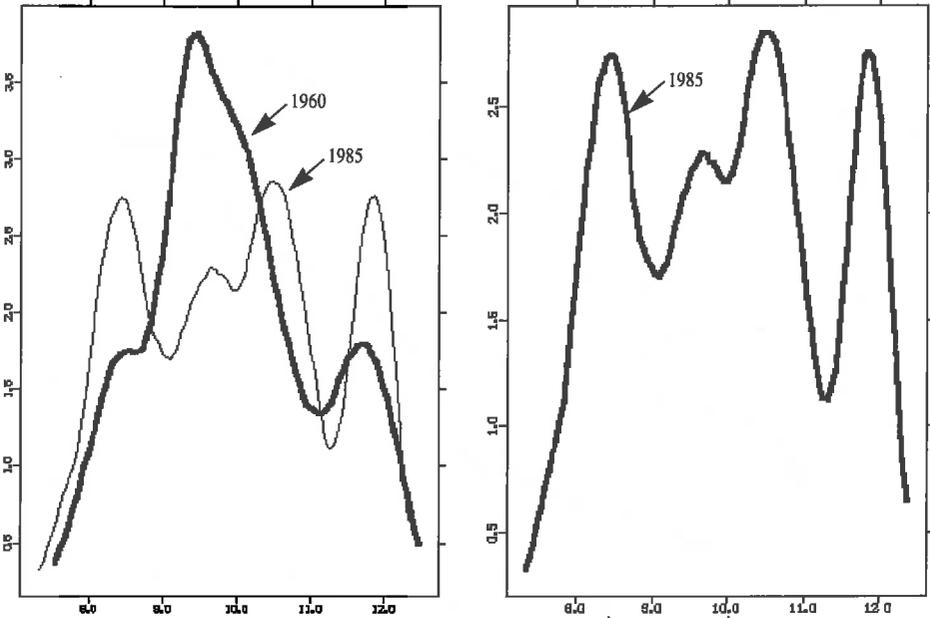
6. Voir aussi Desdoigts (1994) et Quah (1997).

7. Si la littérature empirique sur la croissance a connu un tel essor depuis maintenant dix ans, elle le doit aussi à la naissance d'une base de données (Penn World Table) qui permet de confronter des observations internationales à la fois en coupes transversales et en séries temporelles par une normalisation tenant compte des parités de pouvoir d'achat de chaque pays par rapport aux États-Unis. Pour plus d'information sur cette base de données, se référer à l'article de Summers et Heston (1991) ou se connecter à : <http://www.columbia.edu/~xc8/study/mydata.htm>.

8. Les détails concernant la construction de ces estimateurs de densité sont disponibles en ANNEXE.

GRAPHIQUE 1

Distribution des niveaux de revenus par tête (en logarithme)  
 pour un échantillon de 120 pays en 1960 et en 1985\*



\* Les observations sont normalisées par rapport à la moyenne.

Source : calculs de l'auteur à partir de la base de données de Summers et Heston (1991).

certains rattrapant les pays les plus développés, d'autres, au contraire, accroissant leur retard vis-à-vis des économies les plus riches initialement. Des économies ayant le même revenu initial empruntent des sentiers de croissance différents dont les trajectoires divergent. La conjecture de Quah selon laquelle, l'économie globale tend à se polariser, les fameux "twin peaks", apparaît plutôt séduisante, et, somme toute, cette information est occultée dans l'approche traditionnelle par les régressions en coupe transversale (Quah, 1996a).

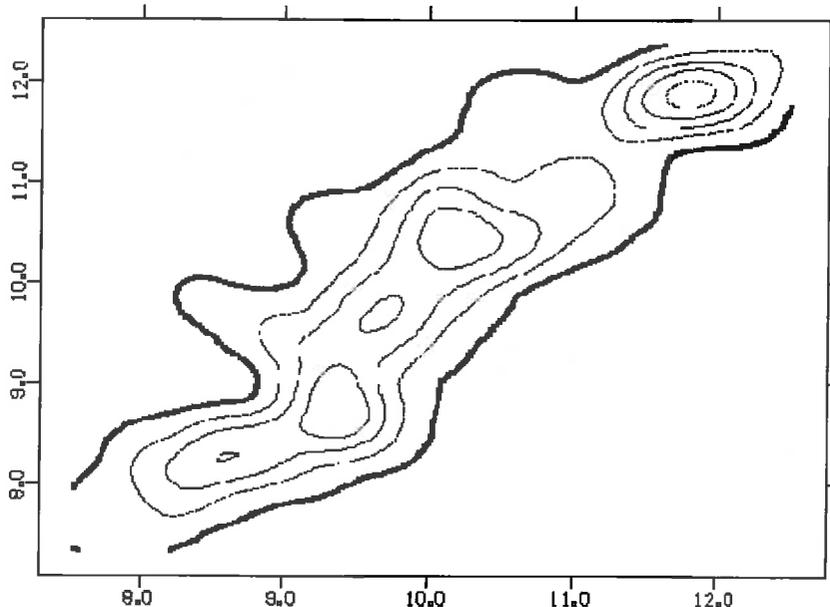
Les différents champs d'attraction vers un même état stationnaire peuvent aussi être illustrés par une variante des matrices de transition calculées par Quah (1993)<sup>9</sup>. Le GRAPHIQUE 2 décrit la probabilité pour un pays de passer d'une classe de revenu en 1960 à une autre classe de revenu en 1985. Cette probabilité est illustrée par des courbes de niveau (respectivement de la plaine vers le sommet : 10 %, 30 %, 50 %, 70 %, 90 %), correspondant à la densité jointe estimée des logarithmes des niveaux de revenus par tête pour les années 1960 et 1985<sup>10</sup>. Les

9. Une matrice de transition telle qu'elle a été proposée par Quah (1993) consiste à considérer les niveaux de revenus par tête à deux instants donnés du temps, et à les normaliser, par exemple, par rapport à la moyenne. Pour chacune des deux dates considérées, on rassemble ensuite ces observations en intervalles, par exemple,  $[0, 1/4], [1/4, 1/2], [1/2, 1], [1, 2], [2, +\infty]$ , où 1 représente le niveau moyen de revenu par tête. Pour chacune des classes de départ, on calcule ensuite les probabilités de transition d'une classe de revenu à une autre au cours de la période séparant les deux dates.

10. Les détails concernant la construction de cette densité jointe sont disponibles en ANNEXE.

GRAPHIQUE 2

Courbes de niveau (0,1 ; 0,3 ; 0,5 ; 0,7 ; 0,9) de la densité jointe des logarithmes des niveaux de revenus par tête en 1960 (en abscisse) et en 1985 (en ordonnée)\*



\* Les observations sont normalisées par rapport à la moyenne.

Source : calculs de l'auteur à partir de la base de données de Summers et Heston (1991).

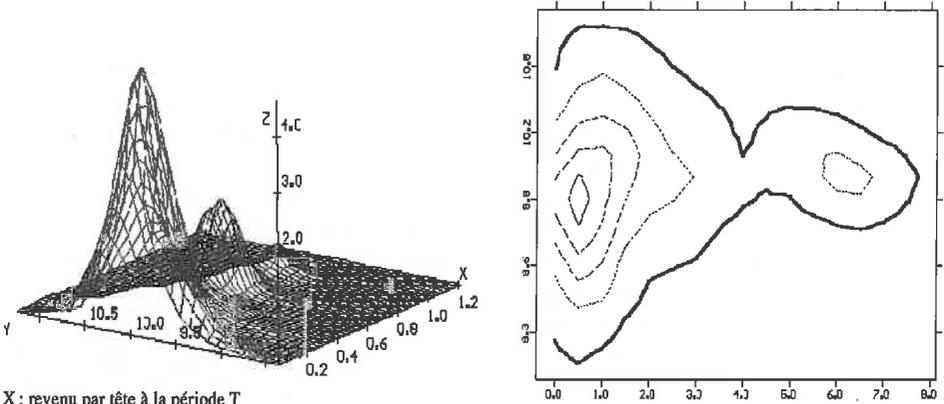
portions des courbes de niveau au-dessus (au-dessous) de la première bissectrice illustrent le phénomène de rattrapage (un approfondissement des inégalités). Les courbes plus éloignées de la première bissectrice correspondent à ce que l'on appelle couramment les "miracles" ou "désastres" économiques caractérisés par une faible probabilité de réalisation (inférieure à 10 %). Enfin, les courbes de niveau entourant les sommets représentent les principaux champs d'attraction.

Au vu de cette distribution, Quah conclut que persistance et immobilité caractérisent la dynamique inhérente à l'économie globale (Quah, 1993, 1996). La probabilité de transiter d'une classe de revenus à une autre est très faible, les observations sur le GRAPHIQUE 2 ont tendance à se concentrer autour de la première bissectrice. Cependant, cette apparente immobilité n'est en aucun cas un fait stylisé et trouve une explication très simple dans ce l'on nomme "une spirale d'inégalité". Prenons un exemple : partageons les revenus en intervalles de 5 % après les avoir normalisés par rapport au revenu des Etats-Unis <sup>11</sup>. Une économie qui débute la période avec un revenu équivalent à 5 % (respectivement 60 %) du revenu des Etats-Unis, dispose d'un revenu par tête d'environ 500\$ (5990\$). Pour

11. Les Etats-Unis sont le pays qui dispose du niveau de revenu par tête le plus élevé sur la période et sont communément admis comme étant le *leader*, c'est-à-dire celui qui est le plus proche de la frontière technologique.

GRAPHIQUE 3

Densité jointe (0,1 ; 0,3 ; 0,5 ; 0,7 ; 0,9) des niveaux de revenus par tête (en abscisse) et des taux de croissance des revenus par tête annuels moyens\* (en ordonnée) sur 5 ans



X : revenu par tête à la période T  
 Y : taux de croissance annuels moyens sur 5 ans  
 Z : densité estimée

\* Normalisés par rapport à celui des Etats-Unis.

Source : calculs de l'auteur à partir de la base de données de Summers et Heston (1991).

observer le rattrapage de cette économie dans une matrice de transition, il faut qu'elle ait un revenu au moins supérieur à 10 % (65 %) relativement à celui des Etats-Unis à la fin de la période, et donc supérieur à 1150\$ (7580\$). Autrement dit, elle devra être caractérisée par un taux de croissance annuel moyen d'environ 17 % (4,8 %). On observe ainsi une mobilité ascendante (respectivement descendante) relativement faible pour les pays appartenant à la classe de revenu la plus faible (respectivement la plus élevée). Dans ce cadre analytique, même si les pays aux niveaux de revenus les plus faibles ont tendance à rattraper les pays les plus riches, l'immobilité paraît les caractériser.

La densité jointe des niveaux de revenus par tête, normalisés par rapport à celui des Etats-Unis, et des taux de croissance annuels moyens fait mieux apparaître la mobilité à l'intérieur de notre système (GRAPHIQUE 3). D'une part, la possibilité de converger n'est pas un privilège des économies bénéficiant déjà d'un certain niveau de revenu initial. D'autre part, la possibilité de prendre ou d'accumuler du retard n'est pas l'apanage des pays les plus pauvres. Enfin et surtout, la présence des pays les plus pauvres dans les queues de la distribution marginale correspondant aux taux de croissance, suggère au contraire une forte mobilité des pays les plus pauvres. Malheureusement, ces taux de croissance restent trop faibles pour leur permettre d'échapper à la spirale d'inégalité décrite ci-dessus, ou encore à la trappe à pauvreté illustrée par le GRAPHIQUE 1.

Aussi éclairante que puisse être cette analyse qui met en évidence l'existence de clubs de convergence, elle ne dit cependant rien sur les caractéristiques du processus de convergence locale illustré par les GRAPHIQUES 1 et 2. De plus, le

GRAPHIQUE 3 n'est rien d'autre que le triangle de Baumol-Barro-Romer<sup>12</sup> qui a conduit à formuler le concept de convergence conditionnelle.

Comment, par exemple, concilier l'hypothèse de  $\beta$ -convergence avec la stratification de la distribution des revenus ? La réponse ne réside certainement pas dans la seule spirale d'inégalité décrite ci-dessus. En fait, une fois introduits dans une régression en coupe transversale les déterminants de l'état stationnaire, si ces derniers sont significativement différents de zéro, alors on ne peut rejeter l'idée selon laquelle les états stationnaires sont différents au sein de l'économie mondiale. On peut observer une stratification de la distribution des revenus, sans pour autant rejeter l'hypothèse d'une technologie à rendements décroissants commune à l'ensemble des économies<sup>13</sup>.

Une seconde hypothèse qui explique l'émergence de clubs de convergence consiste à relâcher la spécification commune à rendements décroissants et à envisager des non linéarités dans les technologies. Dès lors, des économies qui ne diffèrent, toutes choses égales par ailleurs, que par leurs conditions initiales, peuvent être localisées dans des champs d'attraction vers des états stationnaires différents. Cette seconde hypothèse est également compatible avec la stratification de la distribution des revenus. Cependant, elle admet l'existence d'équilibres multiples à long terme, et en particulier de clubs de convergence.

L'enjeu est donc le suivant : si tant est qu'une telle dichotomie soit à l'œuvre dans la réalité, les clubs de convergence sont-ils le fait d'une hétérogénéité micro-économique qui empêche les économies d'exploiter pleinement les avantages en termes de productivité décrits dans un schéma néoclassique ? Ou, résultent-ils uniquement des circonstances initiales dans lesquelles la dynamique de croissance s'installe, contraignant les économies à appartenir à des champs d'attraction vers des états stationnaires, différents, et suggérant la présence de non linéarités dans les technologies ?

## L'émergence de clubs de convergence "conditionnels"

Durlauf et Johnson mettent en évidence l'existence de clubs de convergence. Ils ajoutent aux régressions en coupe un degré de liberté en autorisant d'éventuelles non linéarités (Durlauf & Johnson, 1995). Deux variables de contrôle, le niveau de revenu par tête initial et le stock de capital humain initial servent de base à une classification en groupes des pays. Pour ce faire, ils balayent tout l'échantillon par des régressions en coupe traditionnelles le long des axes de ces deux variables. Si la somme des carrés des résidus est plus faible pour deux sous-échantillons que pour l'échantillon dans son ensemble, alors les pays sont classés dans ces deux sous-échantillons et sont soumis à une analyse identique à celle de

12. Sur un graphique décrivant, pour un large échantillon de pays, le taux de croissance annuel moyen associé au niveau de revenu par tête initial, on observe une configuration triangulaire du type de celle décrite par le GRAPHIQUE 3.

13. En dehors, bien sûr, de toute considération liée à la possibilité de chocs aléatoires qui viendraient perturber le cheminement d'une économie vers son état stationnaire, sans affecter nécessairement l'ensemble de l'économie mondiale et qui ne pourrait faire l'objet d'une spécification dans le cadre d'une régression en coupe transversale.

l'approche traditionnelle. Leurs résultats vont en faveur d'une convergence locale. L'hypothèse d'une spécification linéaire commune à tous les pays est rejetée.

Au vu des trois démarches présentées jusqu'ici, suivons l'approche traditionnelle mais procédons à rebours, comme le suggère Sala-i-Martin (1996)<sup>14</sup>. Autrement dit, au lieu de contrôler l'hétérogénéité économique à l'intérieur d'une seule et même économie représentative, examinons en premier lieu les sources d'hétérogénéité sur la base de variables qui reflètent à la fois, les conditions initiales, le rythme d'accumulation et l'allocation des ressources, et cela sans aucune information *a priori* sur cette hétérogénéité. Analysons ensuite dans quelle mesure les économies sont homogènes et identifions des clubs de convergence qualifiés de "conditionnels".

Cette analyse s'inspire de l'étude de Durlauf et Johnson dans le sens où elle cherche à mettre en évidence l'existence d'équilibres multiples, mais elle est conduite en considérant simultanément toutes les variables économiques les plus communément introduites dans les régressions en coupe transversale. Ainsi, elle évite l'inconvénient majeur de leur méthodologie économétrique, c'est-à-dire, la partition de l'échantillon en rectangles (puisque'il y a deux variables de contrôle) dont les côtés doivent nécessairement être parallèles aux axes des variables de contrôle sélectionnées. Ainsi, des anomalies, d'ailleurs notées par les auteurs, viennent se glisser dans leur étude. Par exemple, le Japon et la Corée du Sud appartiendraient à un club dans lequel on trouve également le Salvador et Madagascar, alors que l'Uruguay appartiendrait au club dans lequel figurent les Etats-Unis.

Comme il est impossible de visualiser un espace au delà de trois dimensions, il faut envisager une procédure qui réduise l'espace original de façon à mettre en évidence, si elle existe, une structure de groupes. Une variante de l'analyse factorielle est ici mise en œuvre. En bref, il s'agit de projeter cet espace multidimensionnel sur un espace de dimension réduite et d'analyser la distribution des points projetés. Ces derniers sont sélectionnés en maximisant un indice destiné à détecter des structures non spécifiées et non anticipées présentes dans l'espace multidimensionnel d'origine (Friedman & Tukey, 1974). Plus précisément, supposons que nous disposons de données pour  $n$  pays, chacun étant caractérisé par un ensemble de  $q$  variables qui prennent en compte à la fois les conditions initiales, le rythme d'accumulation et l'allocation des ressources. Les observations concernant le pays  $i$  peuvent être placées dans un vecteur de dimension  $q$ , noté  $X_i$ , et les données sont alors regroupées dans un ensemble  $(X_1, \dots, X_n)$ . Soit maintenant un vecteur  $\alpha$  de dimension  $q$ , une projection univariée des données se définit par :  $(z_1, \dots, z_n) = (\alpha' X_1, \dots, \alpha' X_n)$ , où  $z_i = \alpha' X_i$ . Avec  $\hat{f}_\alpha$ , un estimateur de la densité des  $z$ , alors nous pouvons définir un indice  $I(\hat{f}_\alpha)$ , qui pour un ensemble de données  $(X_1, \dots, X_n)$  fixé, peut être considéré comme une fonction de  $\alpha$ . Une caractéristique intéressante des données s'obtient par un choix approprié de  $I$ , et en maximisant cet indice sous la contrainte  $\alpha' \alpha = 1$ . C'est ici une variante de l'analyse classique par les composantes principales où l'on cherche à maximiser la part de la variance prise en compte par les différentes composantes principales :

14. Desdoigts (1996).

$I_{PC}(\hat{f}_\alpha) = \int (z - \int x \cdot \hat{f}_\alpha(z))^2 \hat{f}_\alpha(z) dz = \text{Var}(z)$ . La variance donne une mesure globale de la structure présente dans  $(X_1, \dots, X_n)$ . Ainsi, l'analyse par les composantes principales est parfaitement justifiée si la distribution jointe de  $(X_1, \dots, X_n)$  est normale, c'est-à-dire, si aucune information concernant la structure présente dans l'espace d'origine n'est occultée en se concentrant sur l'information fournie par ses premiers moments. En fait, l'indice à maximiser est de la forme :

$I(\hat{f}_\alpha) = \int (\hat{f}_\alpha(z) - \Phi(z))^2 \Phi(z) dz$ , où  $\Phi$  est la fonction de densité de la loi normale standard. Il a les mêmes propriétés que celui développé à l'origine par Friedman et Tukey (1974)<sup>15</sup>. Il permet d'identifier et de mettre en évidence une éventuelle structure locale présente dans l'espace d'origine. On parle souvent d'une structure qui serait "intéressante". S'il est plutôt difficile de s'entendre sur ce que l'on entend par "intéressant", la normalité, elle, est clairement "inintéressante". Littéralement, cet indice cherche à maximiser la distance de l'estimateur de densité des points projetés sélectionnés par rapport à la loi normale. Ainsi, si deux points projetés sont proches l'un de l'autre, on peut en conclure que les deux pays sont identiques respectivement à certaines de leurs caractéristiques. Pour discriminer entre l'influence de toutes ces variables, il faut se référer à la valeur prise par les coefficients du vecteur  $\alpha$ . La valeur d'un coefficient indique l'importance relative de la variable à laquelle il est associé dans la détermination des groupes, et permet donc de rendre compte en quoi les pays sont similaires ou, au contraire, différents.

Cette procédure peut être généralisée à des projections en deux dimensions.  $\alpha$  devient alors une matrice dont les colonnes doivent être orthogonales. Dans le cas d'une projection en deux dimensions,  $\hat{f}_\alpha$  est un estimateur de la densité jointe des projections qui dépendent de  $\alpha$  et  $\Phi$  est remplacé par une fonction de densité normale bivariée. On choisit alors  $\alpha \in R^q \times R^q$  qui maximise  $I$  sous la contrainte  $\alpha' \alpha = I_2$ , où  $I_2$  est la matrice identité. Dans cette étude, nous présentons des projections sur un plan afin de ne pas perdre de l'information qui serait localisée dans la seconde dimension.

Comme dans l'analyse de la distribution des revenus présentée ci-dessus, la structure qui émerge de la distribution correspondant aux projections sélectionnées, permet de dégager des informations sur l'existence ou l'absence de clubs de convergence, ainsi que sur les facteurs économiques qui conduisent à leur formation. La différence avec les modèles précédents de dynamique des distributions tient à l'aspect "conditionnel" des clubs qui ressortent de notre analyse. Cette caractéristique ne peut en aucun cas être capturée par l'approche traditionnelle de l'hypothèse de convergence.

## Les données utilisées

La croissance économique a de nombreuses facettes. Elle dépend de l'utilisation des ressources, du taux de croissance de la population, du savoir-faire technologique, des politiques économiques mises en œuvre, etc. Les données utilisées sont empruntées à l'annexe de De Long et Summers (1991). Le choix des

15. Voir ANNEXE.

variables résulte d'une volonté de comparaison des résultats avec ceux de l'approche empirique classique de l'hypothèse de convergence. Elles sont très proches de celles utilisées par Mankiw, Romer et Weil (*op. cit.*)<sup>16</sup> dont le modèle empirique dérive explicitement du modèle de croissance néoclassique augmenté du stock de capital humain. Cependant, l'analyse proposée ici considère l'investissement en capital physique à un niveau désagrégé et ajoute les dépenses gouvernementales. L'échantillon comprend 61 pays et l'espace original est de dimension 9. Les quatre premières variables sont l'investissement moyen au cours de la période 1960-85 exprimé en pourcentage du PIB.  $X_1$  représente l'investissement en équipement non électrique,  $X_2$ , en équipement électrique,  $X_3$ , en transport, et  $X_4$ , dans la construction. De Long et Summers mettent en évidence une relation très forte et particulièrement robuste entre l'investissement à forte teneur en technologie nouvelle (équipement électrique et non électrique) et la croissance, par rapport à une relation plus faible entre cette dernière et l'investissement en infrastructures (transport et construction) (De Long & Summers, 1991, 1992). L'investissement en équipements de haute technologie, en opposition à ceux en infrastructures, reflète la capacité d'une économie à absorber et à utiliser les nouvelles technologies, ainsi que l'état de l'art technologique qui établit les conditions dans lesquelles une économie initialement peu développée surmonte éventuellement son désavantage et rejoint un club de nations plus riches. Enfin, la distinction entre l'allocation et l'accumulation des ressources se justifie si des taux élevés d'investissements ne suffisent pas à assurer une croissance forte, lorsqu'ils sont alloués de manière inefficace.

Les cinq autres variables sont celles communément introduites dans l'approche traditionnelle :  $X_5$  est le taux de croissance moyen de la population active sur la période,  $X_6$  et  $X_7$  servent de proxy au stock initial de capital humain évalué au début de la période, en 1960, et représentent le pourcentage d'une classe d'âge aux niveaux de scolarisation, respectivement, primaire et secondaire,  $X_8$  mesure la part moyenne des dépenses gouvernementales dans le PIB sur la période et  $X_9$ , l'écart initial de productivité par rapport aux Etats-Unis en 1960.

## Le club des pays de l'OCDE : l'influence de l'éthique religieuse

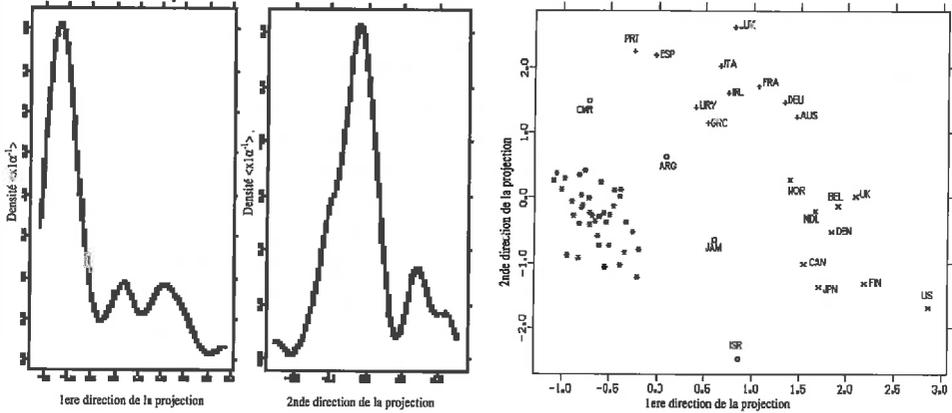
### L'émergence de l'OCDE comme club de nations

Le GRAPHIQUE 4 représente des estimateurs des densités de la projection sélectionnée (graphique de gauche) selon la procédure décrite ci-dessus et en ANNEXE, et une vue de cette même projection sur un plan (graphique de droite). Le graphique de gauche propose une vue intéressante de l'espace original mais réduit, ici, à deux dimensions. Une structure locale apparaît, qui permet d'identifier trois groupes de nations dans la première direction, et deux, dans la seconde direction. Le graphique de droite permet d'identifier ces groupes. Il s'agit, entre

16. Ce modèle sert également de point de départ à l'étude réalisée par Durlauf et Johnson (1995).

## GRAPHIQUE 4

Densités de la projection sélectionnée dans les deux directions (graphique de gauche) et vue de cette même projection sur un plan (graphique de droite) pour les 61 pays de l'échantillon



Source : calculs de l'auteur à partir de la base de données de De Long et Summers (1991).

TABLEAU 1

Projecteurs correspondant aux projections sélectionnées sur le graphique 4 pour 61 pays\*

	1 <sup>re</sup> direction	2 <sup>e</sup> direction	Variables
$X_1$	0,177	0,024	Équipement non électrique
$X_2$	- 0,013	- 0,078	Équipement électrique
$X_3$	0,055	0,229	Transports
$X_4$	- 0,085	- 0,138	Construction
$X_5$	- 0,148	- 0,995	Croissance de la population active
$X_6$	- 0,205	0,591	Taux de scolarisation primaire
$X_7$	0,684	- 1,253	Taux de scolarisation secondaire
$X_8$	- 0,017	- 0,099	Dépenses gouvernementales
$X_9$	- 0,318	- 0,209	Ecart initial

\* Pays de l'échantillon de la base de données de De Long et Summers (1991).

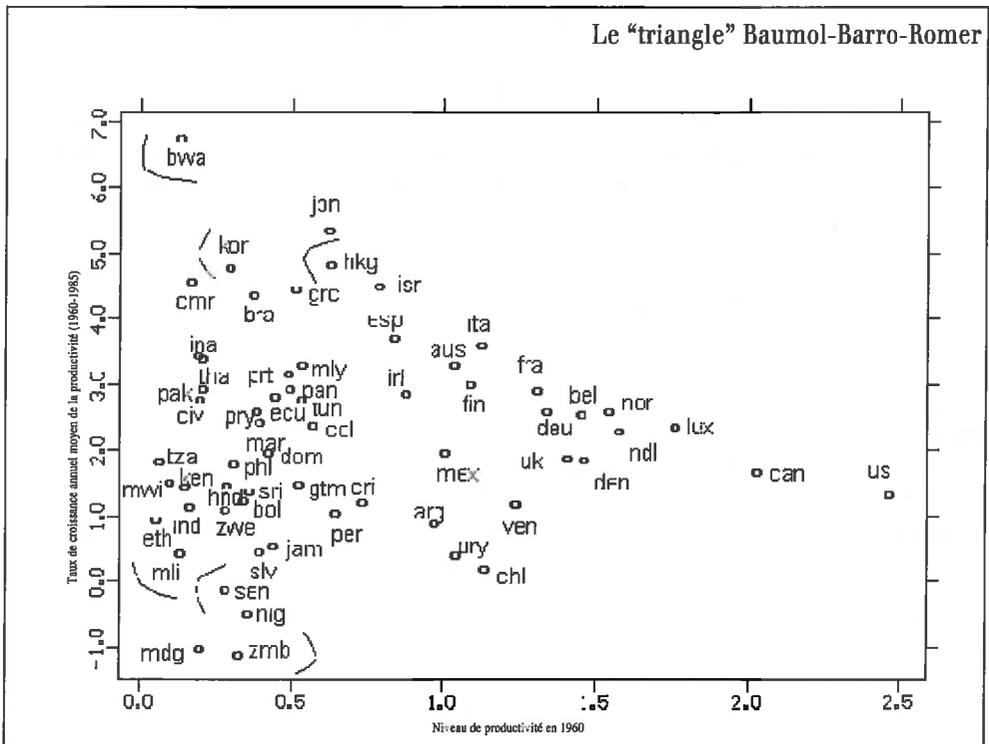
autres, des pays membres de l'OCDE d'une part, et des pays appartenant au reste du monde, d'autre part.

Dans l'analyse statistique mise en œuvre ici, aucune spécification *ex ante* ne permet de distinguer un pays appartenant à l'OCDE d'un pays qui n'y appartient pas. L'OCDE se forme de manière endogène sur la base des variables décrites précédemment qui reflètent les choix d'un pays en termes d'accumulation et d'allo-

cation de ses ressources ainsi que ses conditions initiales. Plus formellement, elle est le résultat d'une combinaison linéaire des variables introduites dans l'analyse statistique. La question est alors de comprendre quelles sont les caractéristiques économiques qui ont conduit à l'émergence de ce club. Référons-nous aux coefficients  $\alpha$  (TABLEAU 1). La valeur prise par chaque coefficient indique l'importance relative de la variable à laquelle il est associé dans la détermination des groupes qui émergent.

Les pays membres de l'OCDE émergent comme un club de nations distinct du reste du monde du fait des différences dans les conditions initiales, c'est-à-dire, stock de capital humain ( $X_6$ ,  $X_7$ ) en premier lieu, et stock de capital physique ( $X_9$ ) en second lieu. Ces variables rendent compte des circonstances dans lesquelles la dynamique de croissance s'est installée au cours de la période considérée. On est tenté de conclure que cette structure est largement influencée par l'écart initial entre les deux catégories de pays. Autrement dit, un effet de richesse et d'accumulation préliminaire à celle-ci domine dans l'émergence de la dimension institutionnelle. Cette idée doit toutefois être nuancée et considérée avec précaution. En effet, il est tout-à-fait possible d'identifier des nations telles que le Venezuela, le Mexique, et le Chili, qui, au début de la période, bénéficient d'un stock de capital physique plus important que, par exemple, l'Espagne, la Grèce, et le Portugal. Pour bien voir que l'émergence des pays membres de l'OCDE n'est pas uniquement liée à un effet de richesse, le triangle de Baumol-Barro-Romer

GRAPHIQUE 5



Source : calculs de l'auteur à partir de la base de données de De Long et Summers (1991).

qui correspond à l'échantillon, est reproduit sur le GRAPHIQUE 5 : la seule prise en compte de l'écart initial ne suffit pas à distinguer les pays membres de l'OCDE des autres pays. L'émergence de l'OCDE n'est pas le résultat de la seule accumulation préliminaire des richesses mais s'explique surtout par la détention d'un stock important de capital humain, relativement aux autres nations, ainsi que par leur capacité à absorber les nouvelles technologies et à développer leurs infrastructures ( $X_1$ ,  $X_3$ ).

### Le facteur culturel dans la différenciation au sein des pays de l'OCDE

Un autre enseignement de ces projections concerne l'existence de deux groupes distincts au sein des pays membres de l'OCDE. Le GRAPHIQUE 4 permet en effet d'identifier deux groupes de pays au sein de l'OCDE. A l'intérieur de ces deux groupes, on trouve soit des nations à dominante protestante (à l'exception du Japon et de la Belgique), soit des nations à dominante catholique (à l'exception de la Grèce). Ceci est en accord avec l'intuition de De Long (1988) qui dans le prolongement des travaux de Max Weber, insiste sur l'importance de la culture religieuse dominante dans le processus de développement. Les pays à dominante protestante sont caractérisés, en moyenne, par des taux d'investissements, des taux de croissance de la population active (qui varient du simple au double), et des stocks de capital humain plus élevés (TABLEAU 2). Notons que pour ces derniers, la différence tient essentiellement dans la variable "taux de scolarisation secondaire". Ces résultats illustrent la pertinence de l'argument développé par Max Weber dans *L'éthique protestante et l'esprit du capitalisme*.

Pour expliquer le succès économique de certaines nations, Hall et Jones mettent en évidence le rôle primordial joué par les institutions et les politiques économiques dans l'aménagement de l'environnement productif et commercial (Hall & Jones, 1997). Nos résultats suggèrent, à côté des facteurs institutionnels, l'impor-

TABLEAU 2

Moyennes et écart-types\* des variables de contrôle selon les groupes de pays\*\*

GROUPE	$^aX_1$	$^aX_2$	$^aX_3$	$^aX_4$	$^bX_5$	$^cX_6$	$^cX_7$	$^dX_8$	$^eX_9$
PROTESTANT	6,02 (0,016)	2,32 (0,009)	2,74 (0,006)	16,0 (0,04)	1,3 (0,006)	105,8 (0,08)	66,4 (0,1)	15,4 (0,037)	0,385 (0,2)
CATHOLIQUE	5,29 (0,018)	1,61 (0,008)	2,42 (0,008)	15,5 (0,04)	0,6 (0,003)	118,5 (0,13)	36,8 (0,1)	14,3 (0,025)	0,582 (0,15)

\* Entre parenthèses.

\*\* Les pays appartenant à ces groupes sont disponibles en annexe.

(a) Différentes composantes de l'investissement, en % du PIB (voir tableau 1).

(b) Taux de croissance de la population active en pourcentage.

(c) Taux de scolarisation dans le primaire et le secondaire, respectivement.

(d) Part des dépenses gouvernementales dans le PIB.

(e) Ecart relatif par rapport aux Etats-Unis en 1960 sur une échelle [0,1].

tance des facteurs culturels. Il existe bien un déterminisme culturel qui influence directement la vie et les comportements économiques, et qui ne se limite pas à un épiphénomène dans l'environnement économique. Nul doute, par exemple, que l'émergence de l'éthique protestante, – réussite économique, individualisme, et innovation –, a facilité la naissance et le développement du capitalisme, qui a, lui, contribué à la fois à la révolution industrielle et à la révolution démocratique (Inglehart, 1997).

Durlauf et Johnson voient dans l'importance de cette dimension culturelle l'explication à certaines des anomalies dans les groupes de pays qu'ils ont identifiés (Durlauf & Johnson, 1995). L'intérêt de la méthodologie retenue ici réside dans sa capacité à faire émerger de façon endogène ce type de caractéristiques. Celles-ci font référence à des conditions initiales telles que les stocks initiaux de capital humain et physique et la capacité à développer, absorber, et partager les nouvelles technologies, qui conduisent à la formation de clubs ou encore de coalitions d'économies. De Long et Summers ont également insisté sur l'importance du capital ethnique et culturel dans le développement d'une nation, importance qui, selon eux, doit être fortement liée à la localisation, et en particulier au continent d'accueil d'une nation (De Long & Summers, 1993).

Avant d'approfondir cette intuition, il est utile, pour qualifier les groupes de nations qui émergent de l'analyse en termes de "clubs de convergence conditionnels", d'étudier le lien entre la structure présente dans l'espace original et le taux de croissance de ces mêmes économies. La surface de la régression bivariée du taux de croissance annuel moyen par travailleur sur la projection sélectionnée est donnée par le GRAPHIQUE 6. Il s'agit d'une estimation non paramétrique qui met en évidence d'éventuelles non linéarités entre le taux de croissance et la structure présente dans l'espace original des variables économiques introduites dans l'analyse.<sup>17</sup>

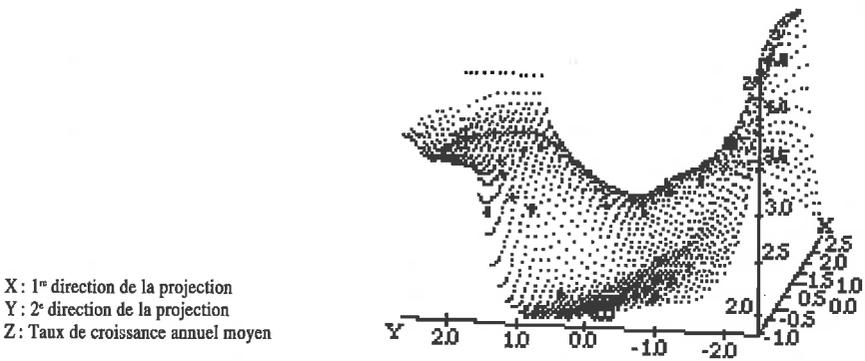
Le plan (X,Y) correspond au plan de la projection sélectionnée. Une structure non linéaire ressort clairement de cette régression. Le groupe de l'OCDE bénéficie en moyenne de taux de croissance plus élevés par rapport aux pays non membres de cette institution. Ils sont de plus localisés sur deux sommets distincts dont l'altitude varie selon que l'on appartient au groupe dit "protestant", ou au groupe dit "catholique". Dans la vallée, on trouve les pays du reste du monde, en quelque sorte enfermés dans une trappe à pauvreté. Les croix sur le graphique indiquent les observations qui ont conduit à l'estimation de cette structure. D'une part, il est intéressant de noter que pour sortir de cette vallée, les pays non membres de l'OCDE ont tendance à se diriger vers le sommet qui accueille les pays dits "protestants". D'autre part, il semble, au vu de la faible densité d'observations sur les flancs de cette montagne, qu'il existe un effet de seuil important pour opérer une transition de la vallée vers les sommets. Comme le suggèrent certains nouveaux modèles de croissance, par exemple la théorie du *big push*<sup>18</sup>, la difficulté à accéder à la phase préliminaire de développement nécessaire à des gains de productivité importants peut expliquer la croissance plus rapide et le cumul d'inégalités entre des économies relativement prospères et des économies relativement pauvres. De telles non linéarités dans les technologies où il existe

17. Voir l'ANNEXE pour une description détaillée de cette régression.

18. Murphy, Shleifer & Vishny (1989).

## GRAPHIQUE 6

Surface de la régression bivariée du taux de croissance annuel par travailleur (1960-1985) sur les directions de la projection délectionnée comme décrite dans le graphique 4



Source : calculs de l'auteur à partir de la base de données de De Long et Summers (1991).

des zones de rendements décroissants suivies par des zones de rendements croissants conduisent par exemple, à l'émergence d'une trappe à pauvreté telle qu'elle est illustrée par le GRAPHIQUE 6.

Ces résultats permettent de parler de "clubs de convergence conditionnels" et d'invoquer l'existence d'équilibres multiples à long terme. Il est utile de rappeler que la structure sur le plan (X,Y) reflète principalement des différences dans les conditions initiales et que les combinaisons linéaires sélectionnées qui conduisent à l'émergence de cette structure reflètent une dimension institutionnelle et culturelle qui a trait à la religion dominante d'une nation<sup>19</sup>. Plus généralement, par culture, on entend un système d'attitudes, de valeurs, et de connaissances largement partagées au sein d'une société qui se transmet de génération en génération. Elle peut cependant varier d'une société à l'autre (Inglehart, *op. cit.*). Si l'on admet que l'héritage culturel acquis au cours des premières années de sa vie est relativement résistant à tout changement, alors, on admet également l'existence à long terme de champs d'attraction vers des états stationnaires différents.

Certains faits stylisés sont masqués dans l'analyse classique de l'hypothèse de convergence, et notre méthodologie est une alternative statistique pertinente lorsqu'il s'agit de rendre compte de la stratification de la distribution des revenus. Pour le moins, elle montre l'importance d'une analyse précise des données avant d'entreprendre des régressions, sans se poser la question à laquelle il faut répondre : convergence globale ou locale ?

A ce stade de l'analyse, une étude interne au groupe des pays non membres de l'OCDE doit être mise en œuvre. Cette technique d'isolation est souvent utilisée dans l'analyse des clubs en général, et plus spécifiquement dans les techniques

19. Nous rappelons que les différences dans les taux de croissance de la population active (X5) contribuent largement à l'identification de nos deux groupes au sein de l'OCDE. Pour expliquer la mobilité des nations, les différentes approches, i.e., convergence conditionnelle et équilibres multiples, sont donc complémentaires et chacune permet de rendre compte de phénomènes à l'œuvre dans la réalité.

mises en œuvre dans notre étude, du fait de l'incapacité des projections en une ou deux dimensions à faire apparaître toute la structure présente dans l'espace original.

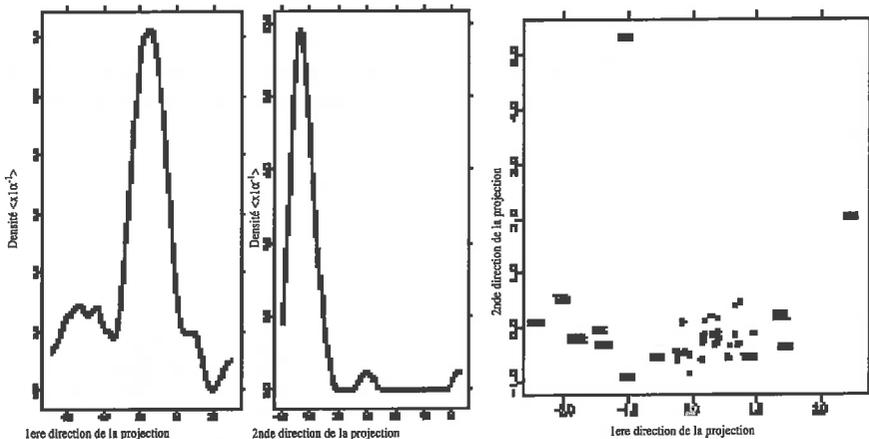
# Le cas des pays en développement

## Le miracle des pays du sud-est asiatique et la trappe à pauvreté africaine

Concentrons nous sur les pays non membres de l'OCDE. Le GRAPHIQUE 7 est construit de manière identique au GRAPHIQUE 4. Le continent africain et la zone des pays du sud-est asiatique<sup>20</sup> émergent de manière endogène sur la base de leurs caractéristiques économiques et cela, le long de l'axe correspondant à la première direction. La seconde direction met en évidence la présence de deux *outliers* que sont le Botswana et Hong Kong<sup>21</sup>. Les variables économiques ici à l'œuvre pour mettre en évidence cette structure géographique le long de l'axe horizontal reflètent essentiellement les dotations initiales en capital humain ( $X_7$ ) et avec une influence moindre, l'écart initial ( $X_9$ ), mais aussi, et surtout, les dépenses gouvernementales ( $X_8$ ) (TABLEAU 3). Notons également l'importance relative de la variable transport ( $X_3$ ) dans la structure émergeant dans la seconde direction de la projection sélectionnée.

GRAPHIQUE 7

Densités de la projection sélectionnée dans les deux directions (graphique de gauche) et vue de cette même projection sur un plan (graphique de droite) après avoir isolé les pays non membres de l'OCDE\*



\* L'Argentine, le Cameroun, Israël, la Jamaïque et l'Uruguay sont également exclus.

Source : calculs de l'auteur à partir de la base de données de De Long et Summers (1991).

20. Singapour et Taiwan ne sont pas disponibles dans l'échantillon original.

21. L'influence de la variable équipement électrique est due à la présence du Botswana qui dispose du ratio le plus élevé parmi tous ces pays, (exploitation des mines de diamants), en dépit d'un stock de capital humain parmi les plus bas dans l'échantillon à l'étude.

TABLEAU 3

Projecteurs correspondant aux projections sélectionnées sur le graphique 7 pour les pays non membres de l'OCDE\*

	1 <sup>re</sup> direction	2 <sup>e</sup> direction	Variables
$X_1$	0,010	- 0,019	Équipement non électrique
$X_2$	0,104	1,141	Équipement électrique
$X_3$	0,008	- 0,351	Transports
$X_4$	- 0,098	0,031	Construction
$X_5$	- 0,144	- 0,028	Croissance de la population active
$X_6$	0,046	- 0,079	Taux de scolarisation primaire
$X_7$	0,642	0,078	Taux de scolarisation secondaire
$X_8$	- 0,821	0,054	Dépenses gouvernementales
$X_9$	0,229	- 0,037	Ecart initial

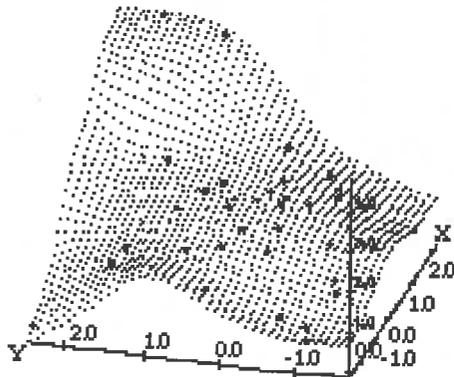
\* L'Argentine, le Cameroun, Israël, la Jamaïque et l'Uruguay sont également exclus.

Le GRAPHIQUE 8 donne la surface qui résulte de la régression du taux de croissance annuel moyen par travailleur sur les deux directions de la projection sélectionnée. Le Botswana et Hong Kong sont exclus de cette régression. Ce graphique peut être interprété comme une vue rapprochée de la vallée illustrée par le GRAPHIQUE 6. A nouveau, le plan (X,Y) représente la projection sélectionnée et les croix, les observations qui ont permis l'estimation de cette surface. La relation est clairement positive le long des deux directions de cette projection. Un effet de seuil apparaît également entre les pays du sud-est asiatique et les autres nations ; la pente de la surface de la régression ayant tendance à se raidir, cela suggère à nouveau la présence de non linéarités dans les technologies.

GRAPHIQUE 8

Surface de la régression bivariée du taux de croissance annuel moyen par travailleur (1960-1985) sur les directions de la projection sélectionnée comme décrite dans le graphique 7

X : 1<sup>re</sup> direction de la projection  
Y : 2<sup>e</sup> direction de la projection  
Z : Taux de croissance annuel moyen

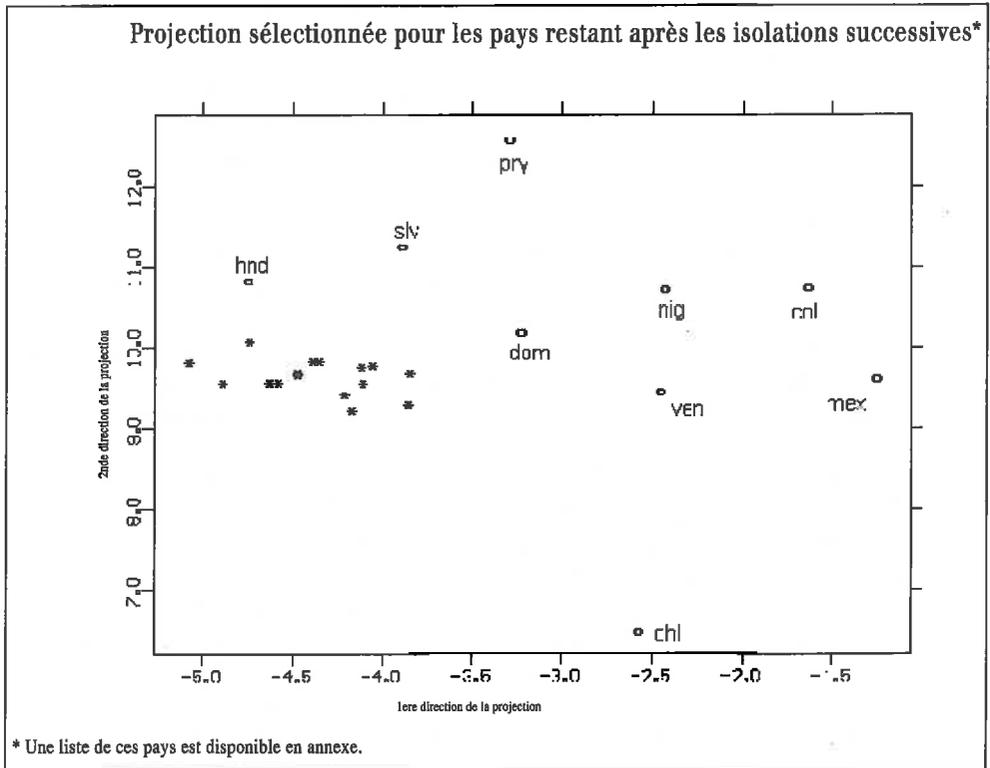


Source : calculs de l'auteur à partir de la base de données de De Long et Summers (1991).

## L'Amérique latine

L'Amérique latine est traitée dans une dernière étape dont sont exclus les pays non membres de l'OCDE, les pays du sud-est asiatique et la plupart des pays africains. Les projections correspondantes sélectionnées sont illustrées dans le GRAPHIQUE 9. Cette procédure permet ainsi, après avoir isolé les pays membres de l'OCDE, de reconstituer l'ensemble de l'Atlas, ce qui confère une dimension géographique particulièrement forte aux variables économiques introduites dans l'analyse. Ce résultat vient corroborer l'intuition de De Long et Summers (1993) selon laquelle la localisation géographique, c'est-à-dire le continent d'accueil d'une économie, a une forte connexion avec les facteurs ethniques et culturels, intuition qui rejoint le concept subjectif de *social capability* évoqué par Abramowitz (1986).

GRAPHIQUE 9



Barro (1991), De Long (1988), et Sala-i-Martin (1997) mettent également en évidence des corrélations partielles robustes entre le taux de croissance et des variables indicatrices de la localisation géographique et de la religion dominante au sein d'une nation. Dans une régression en coupe transversale, ces variables sont supposées indépendantes des variables économiques introduites par ailleurs. Autrement dit, être africain "et/ou" protestant, toutes choses égales par ailleurs, explique une part de la variance du taux de croissance dans la coupe transversale. Cette analyse met en scène ces mêmes variables qui émergent de manière

endogène et sont maintenant le résultat d'une combinaison linéaire des variables économiques introduites dans l'analyse. Dans le prolongement d'Azariadis (1996), toutes ces variables sont-elles déterminées de façon complètement exogène ou, au contraire, par un ensemble d'individus agissant délibérément, leurs comportements et leurs choix en termes d'accumulation et d'allocation des ressources reflétant un système d'attitudes, de valeurs, et de connaissances partagées au sein d'une société et variant d'une société à l'autre ?

Les caractéristiques qui dominent dans l'émergence de la structure de groupes mise en évidence, font référence à des conditions initiales qui permettent d'envisager l'existence de barrières qui empêchent les transferts et la diffusion du progrès technique et donc, l'exploitation des avantages en termes de productivité invoquée dans un schéma purement néoclassique. De plus, compte tenu des caractéristiques de ces barrières qui ont trait à l'environnement institutionnel, culturel et géographique d'une nation, elles peuvent conduire à la formation de clubs d'économies localisés dans différents champs d'attraction vers un même état stationnaire<sup>22</sup>. Notons toutefois, avec l'exemple de la Corée du Sud, que ces barrières ne sont pas infranchissables. Comme le montre le GRAPHIQUE 2, il existe, dans le long terme, une probabilité, certes faible, mais non nulle de transiter d'un équilibre vers un autre. En effet, les conditions initiales sont, elles-mêmes, impliquées dans un processus dynamique. A ce stade de l'analyse, il est important d'envisager un moyen d'étudier la dynamique propre à chacun des clubs qui sont certainement destinés à évoluer.

Notre analyse montre les limites d'une agrégation d'un ensemble d'économies hétérogènes dans une seule et même économie représentative. Elle indique toute la pertinence de certaines hypothèses des nouveaux modèles de croissance qui se concentrent sur l'hétérogénéité des individus pour montrer la possibilité d'une stratification en classes de revenus distinctes. Cette hétérogénéité a-t-elle simplement trait aux déterminants de l'état stationnaire dans un schéma néoclassique ? Il ne semble pas. Les conditions initiales, tels que les stocks de capital physique et humain et l'état de l'art de la technologie, qui caractérisent une économie, semblent être des déterminants fondamentaux dans la stratification de la distribution des revenus. Elles conditionnent l'environnement économique de telle sorte qu'à long terme des équilibres multiples ne peuvent pas être exclus.

De plus, les combinaisons linéaires sélectionnées de nos variables conduisent à l'émergence de caractéristiques difficilement quantifiables tels que les arrangements institutionnels, le capital ethnique et socioculturel, et la localisation géographique. Ainsi, à côté de l'influence prépondérante des conditions initiales, c'est-à-dire de l'histoire à laquelle se rattache une nation, vient se greffer une troisième hypothèse qui rend compte de la mobilité des nations : celle, selon laquelle, les choix en termes d'accumulation et d'allocation des ressources, refléteraient des anticipations ou des choix délibérés d'un groupe "dominant" d'agents<sup>23</sup>. Ces choix obéissent à une idéologie, qu'elle soit religieuse, ou qu'elle relève d'une tradition ethnique ou locale, pour ne citer que ces deux exemples qui ressortent clairement de notre analyse. Les croyances ne sont pas les simples

22. Voir, également, Quah (1997) dont les résultats suggèrent l'importance de la localisation géographique ainsi que les échanges, en particulier leur structure, pour expliquer la mobilité des nations sur la période considérée.

23. Voir par exemple, Krugman (1991) et Benhabib & Gali (1995).

conséquences des changements intervenant dans la vie économique et sociale, elles peuvent également modeler les conditions socio-économiques du développement, c'est-à-dire, influencer à la fois la vie économique et politique.

## CONCLUSION

Le renouveau des théories de la croissance impulsé par les articles de Romer (*op. cit.*) et Lucas (*op. cit.*) a ouvert la voie à un vaste débat sur l'hypothèse de convergence. Les économies les plus pauvres tendent-elles à rattraper les plus riches ?

Alors que les modèles théoriques modélisent des mécanismes et des implications logiques conduisant à des équilibres multiples, la littérature empirique s'est concentrée sur une hypothèse de convergence explicitement dérivée du modèle de croissance néoclassique. Ces modèles empiriques testent la significativité du coefficient, dans une régression en coupe transversale d'économies, qui associe le taux de croissance annuel moyen et le niveau initial de produit par tête ou par travailleur, tout en contrôlant les différences en termes d'état stationnaire grâce à un ensemble de variables économiques, politiques, et/ou institutionnelles. Comme ce coefficient est significatif, le modèle de croissance néoclassique est apparu un peu comme la "fin de l'histoire". Cependant, cette méthodologie exclut implicitement bon nombre des idées prometteuses contenues dans les nouveaux modèles de croissance.

Ces régressions ne fournissent des informations que sur le comportement moyen d'une économie construite de toute pièce en agrégeant un ensemble de nations hétérogènes afin d'obtenir une seule et même économie représentative. Or, il faudrait comprendre comment les économies se comportent les unes par rapport aux autres, et pas seulement identifier les variables qui approximent l'état stationnaire vers lequel tend une économie-type. Une littérature alternative s'est concentrée sur l'analyse de la dynamique sous-jacente à la distribution des revenus, qui met en évidence un monde qui a tendance à se stratifier, voire à se polariser, de telle sorte que des champs d'attraction vers des états stationnaires différents se dégagent.

Cependant, cette méthodologie reste silencieuse sur les variables économiques qui conduisent à l'émergence de clubs de convergence. Or, il est essentiel de dégager les caractéristiques structurelles qui conduisent des groupes de pays à se situer sur des trajectoires les conduisant vers des états stationnaires différents. C'est pourquoi, la méthodologie qui a été exposée vise à mettre en évidence l'existence de clubs de convergence sur la base des caractéristiques économiques propres à chaque nation. Autrement dit, des clubs de pays pour lesquels l'hypothèse d'un état stationnaire identique semble être réaliste, émergent-ils naturellement sur la base de leurs caractéristiques économiques ?

Sur la période considérée, des clubs d'économies émergent, qui reflètent principalement les conditions initiales (stocks de capital physique et humain), et la capacité des économies à développer, absorber et partager les nouvelles technologies. Ils sont compatibles avec la mobilité des nations sur la même période. De plus, la mise en évidence de non linéarités dans les technologies et d'effets de seuils suggère l'existence de "clubs de convergence conditionnels", où la difficulté à accéder à la phase préliminaire de développement explique la croissance

plus rapide et le cumul d'inégalités entre des économies relativement prospères et d'autres relativement pauvres. Ces résultats renvoient à l'importance des circonstances dans lesquelles le processus de croissance s'installe pour expliquer les trajectoires prises par les différents pays et remet en cause l'idée d'un processus unique de convergence invoquée dans un schéma purement néoclassique.

Les pays appartenant à ces clubs ont des caractéristiques communes, tels que les arrangements institutionnels, le capital ethnique et culturel, et la localisation géographique. Dans des régressions en coupe transversale, ces caractéristiques sont corrélées de manière robuste avec le taux de croissance moyen. L'intérêt de cette méthodologie réside dans sa capacité à faire émerger ces variables, de manière endogène, c'est-à-dire, par des combinaisons linéaires de variables purement économiques qui reflètent à la fois la phase préliminaire de développement, le rythme d'accumulation et l'allocation des ressources. Ces résultats suggèrent également l'importance, pour rendre compte de la mobilité des nations, du rôle des anticipations de groupes d'individus dont les choix seraient délibérés, voire auto-réalisateurs, et reflétant un système d'attitudes, de valeurs, et de connaissances qui serait partagé au sein d'une société, mais qui peut varier d'une société à l'autre. Ces choix sont largement influencés par une idéologie, qu'elle soit religieuse, ou qu'elle relève d'une tradition ethnique ou locale.

A. D.

#### ANNEXE

La dynamique de la distribution des niveaux de revenus est analysée dans cette étude avec les techniques de lissage par les noyaux<sup>24</sup>. L'utilisation de ces techniques se justifie pour deux raisons. On pourrait en effet se demander dans un premier temps, pourquoi nous n'avons pas eu recours à l'estimateur de densité le plus fréquemment utilisé : l'histogramme ? En fait, son estimation présente un inconvénient majeur : la structure qui émerge de l'estimation d'une densité par un histogramme dépend de façon cruciale du choix de l'origine. Or, l'exercice consiste précisément à mettre en évidence la vraie structure sous-jacente à la distribution des revenus. L'existence de cette structure dans la distribution, qui illustre la présence de clubs de convergence ou de champs d'attraction vers des états stationnaires différents, ne peut donc pas être le résultat d'un choix arbitraire. Pourtant, dans un second temps, nous désirons conserver la flexibilité des estimateurs non paramétriques tel que l'histogramme, plutôt que d'imposer une forme fonctionnelle paramétrique spécifique.

Supposons que nous disposons d'une variable aléatoire d'observations identiquement et indépendamment distribuées  $\{X_i\}_{i=1}^n$ . Rosenblatt (1956) propose alors de centrer autour

de chaque observation  $X_i$ , une fonction  $K$  appelée noyau et définie par :  $K\left(\frac{x-X_i}{h}\right)$ , où  $h$  est un paramètre qui détermine le montant de lissage inhérent à la procédure d'estimation. L'estimateur de densité par les noyaux est alors :

24. Une introduction complète et très pédagogique de ces techniques est proposée dans les manuels de Silverman (1986) et Härdle (1990).

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right)$$

La fonction  $K$  est symétrique, continue et satisfait la condition,  $\int K(x)dx = 1$ , ce qui garantit  $\int \hat{f}_h(x)dx = 1$ . Par conséquent, pour une fonction  $K$  particulière et un paramètre de lissage donné, il n'existe qu'un seul estimateur de densité possible.

La généralisation à des observations multivariées  $\{x_i\}_{i=1}^n$  de dimension  $d$ , où  $x_i = (X_{i1}, \dots, X_{id})^T$ , est immédiate. La fonction  $K$  doit être remplacée par un noyau multivarié  $K^d$  défini :

$$K^d\left(\frac{x - X_i}{h}\right) = K^d\left(\frac{x_1 - X_{i1}}{h_1}, \dots, \frac{x_d - X_{id}}{h_d}\right) = \prod_{j=1}^d K\left(\frac{x_j - X_{ij}}{h_j}\right)$$

L'estimateur de densité multivarié devient alors :

$$\hat{f}_h(x) = \frac{1}{nh_1 \dots h_d} \sum_{i=1}^n K^d\left(\frac{x - X_i}{h}\right) \text{ avec } h = (h_1, \dots, h_d).$$

Maintenant que nous nous sommes assurés que nos estimateurs sont indépendants du choix arbitraire de l'origine, il nous reste à choisir le paramètre de lissage  $h$  ou  $d$  paramètres selon la dimension de notre espace d'origine. La qualité de l'information issue de l'analyse de la distribution des revenus dans le texte dépend de la validité de nos estimateurs qui eux-mêmes dépendent du paramètre de lissage. Il est bien entendu que le choix de ce paramètre ne peut en aucun cas être arbitraire, ce qui invaliderait les résultats présentés dans le texte. Aussi la mise en œuvre de ces techniques nécessite de construire un estimateur qui soit le plus proche possible de la vraie valeur prise par la fonction de densité. On peut par exemple imaginer un critère de convergence en moyenne quadratique qui se définit par :

$$MSE(x) = E_f((\hat{f}_h(x) - f(x))^2),$$

et qui devient par les propriétés de l'espérance et de la variance, égal à :

$$MSE(x) = Var(\hat{f}_h(x)) + Biases^2(\hat{f}_h(x))$$

On peut maintenant montrer (cf., par exemple, Silverman (1986)) que l'estimateur  $\hat{f}_h(x)$  est asymptotiquement sans biais lorsque le paramètre de lissage tend vers zéro et que sa variance est asymptotiquement proportionnelle à  $1/nh$ . Ainsi, la vitesse de convergence qui dépend du paramètre  $h$  peut s'interpréter de manière simple et intuitive. Si l'on choisit  $h$  élevé, on privilégie la stabilité de l'estimateur, ce qui conduit à accroître le montant de lissage inhérent à la procédure d'estimation, c'est-à-dire à la disparition de l'éventuelle structure présente dans la distribution. Au contraire, si l'on choisit  $h$  petit, alors on privilégie le biais de l'estimateur, mais ce dernier devient alors trop volatile. Il faut donc faire un arbitrage entre la stabilité de l'estimateur et son biais.

Une méthode couramment utilisée pour choisir le paramètre de lissage, est la validation croisée par les moindres carrés introduite indépendamment par Rudemo (1982) et Bowman (1984). Dans ce cas,  $\hat{f}_h(x)$  est un estimateur convergent de  $f(x)$  lorsque la somme des carrés résiduels est minimum. Elle est définie de la manière suivante :

$$SCR(h) = \int (\hat{f}_h(x) - f(x))^2 dx$$

$$SCR(h) - \int (f(x))^2 dx = \int (\hat{f}_h(x))^2 dx - 2 \int \hat{f}_h(x) f(x) dx$$

Puisque  $\hat{f}_h(x)$  est une variable aléatoire en tant que fonction de nos variables aléatoires originales, on peut écrire :

$$E_f(\hat{f}_h(x)) = \int \hat{f}_h(x) f(x) dx$$

$$- \quad E_f(\hat{f}_h(x)) = \frac{1}{n} \sum_{i=1}^n \hat{f}_{h,i}(x_i) \quad \text{avec} \quad \hat{f}_{h,i}(x) = \frac{1}{(n-1)h} \sum_{i \neq j} K\left(\frac{(x - X_j)}{h}\right).$$

Stone (1984) montre qu'un paramètre de lissage  $\hat{h}$  qui minimise asymptotiquement  $SCR(h)$  est donné par :

$$\hat{h} = \text{Arg min } CV(h), \quad \text{avec} \quad CV(h) = \int (\hat{f}_h(x))^2 dx - 2n^{-1} \sum_{i=1}^n \hat{f}_{h,i}(x_i).$$

Finalement, notons que le théorème de Stone (1984) s'applique également dans le cas multivarié. La méthode de validation croisée par les moindres carrés est donc tout à fait appropriée à la sélection de paramètres de lissage de données multidimensionnelles. Elle est alors développée en complète analogie avec le cas univarié. Il est important de noter ici est que les estimateurs de densité présentés dans le texte sont issus d'une estimation qui utilise des paramètres de lissage qui minimisent notre critère ( $SCR(h)$ ). Ils sont donc asymptotiquement optimaux et ne relèvent pas d'un choix arbitraire qui résulterait d'un *a priori* sur l'existence de clubs de convergence et/ou de champs d'attraction vers des états stationnaires différents <sup>25</sup>.

On fait appel à cette même méthode de lissage par les noyaux pour estimer les surfaces de régression bivariée du taux de croissance annuel moyen par travailleur sur les projections sélectionnées. Supposons que nous disposons de  $n$  observations indépendantes  $\{(X_i, Y_i)\}_{i=1}^n$ , la régression de  $y$  sur  $x$  prend communément la forme suivante :

$$Y_i = m(X_i) + \varepsilon_i,$$

avec  $\varepsilon$ , un terme d'erreur identiquement et indépendamment distribué.

Si la forme de la densité jointe est connue, alors  $m(x)$  peut être déterminée de manière analytique. Cependant, si cette forme est inconnue,  $m(x)$  est également inconnue. Traditionnellement, la forme de  $m(x)$  est supposée linéaire et la densité jointe de  $y$  et  $x$  est supposée suivre une loi Normale. Les modèles de régression non paramétriques sont, eux, obtenus sans faire d'hypothèse quant à la forme fonctionnelle de  $m(x)$ . Pour approximer  $m(x)$ , on peut, par exemple, avoir recours à la méthode de lissage par les noyaux. Lorsque l'on estime une fonction de densité par la méthode des noyaux, on considère les points dans un voisinage proche de  $x$  et on pondère la fréquence des observations par le paramètre de lissage  $h$ . Ici, la procédure consiste à pondérer la variable dépendante autour d'un point  $x$ . Un choix naturel est la moyenne de la variable dépendante autour de ce point. Puisque les observations de la variable dépendante qui sont relativement éloignées de  $x$  prendront des valeurs moyennes différentes, cette moyenne locale doit être construite de telle sorte qu'elle est définie dans un voisinage proche de  $x$ . Cette procédure peut être interprétée comme l'idée de base du lissage.

25. Pour les aspects pratiques de la mise en œuvre de ces techniques appliquées au problème de la convergence, voir, par exemple, Desdoigts (1996).

Afin de pondérer les observations  $Y_i$  en fonction de la distance qui sépare  $X_i$  de  $x$ , l'estimateur suivant est utilisé :

$$\hat{m}_h(x) = n^{-1} \sum_{i=1}^n W_{h,i}(x) Y_i,$$

avec  $\{W_{h,i}(x)\}_{i=1}^n$  qui représente une séquence de pondérations qui dépend du paramètre de lissage  $h$  et de l'échantillon  $\{X_i\}_{i=1}^n$  de la variable explicative et qui se définit de la manière suivante :

$$\{W_{h,i}\}_{i=1}^n = \frac{K_h(x - X_i)}{\hat{f}_h(x)},$$

où  $\hat{f}_h(x) = n^{-1} \sum_{i=1}^n K_h(x - X_i),$

l'estimateur de densité par les noyaux et  $K_h(u) = h^{-1} K(u/h).$

Cela conduit à l'estimateur de Nadaraya-Watson :

$$\hat{m}_h(x) = n^{-1} \frac{\sum_{i=1}^n K_h(x - X_i) Y_i}{\sum_{i=1}^n K_h(x - X_i)}$$

Si, par exemple, le noyau est Gaussien, alors, plus  $x$  sera proche de  $X_i$ , le plus important sera le poids affecté à  $Y_i$  sera important. La généralisation à des observations bivariées  $\{X_{i1}, X_{i2}\}_{i=1}^n$  comme c'est le cas dans notre application est immédiate. La fonction  $K$  doit être remplacée par un noyau bivarié  $K^2$  en complète analogie avec le cas de l'estimation de la densité multivariée décrit ci-dessus.

Venons-en maintenant à la technique utilisée dans la seconde partie de cet article où il s'agit d'expliquer la formation des clubs de convergence issus de l'analyse de la dynamique sous-jacente à la distribution des niveaux de revenus et des taux de croissance. Il faut pour cela prendre en compte simultanément l'ensemble des caractéristiques économiques propres à chaque pays. L'espace multidimensionnel qui en résulte doit faire l'objet de transformations avant de pouvoir être analysé. C'est précisément l'objet de la technique statistique décrite ici. Une référence incontournable sur le sujet est la contribution offerte par Jones et Sibson, 1987.

Soit une matrice  $X$  de dimension  $q \times n$ , où  $n$  est le nombre d'individus et  $q$  la dimension de l'espace considéré qui reflète les caractéristiques économiques de nos individus. Si maintenant, nous projetons ces données dans un espace à une dimension, nous obtenons un vecteur de projection ( $Z$ ) tel que :

$$Z = \alpha^T X = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_q X_q,$$

avec  $\alpha$  un vecteur de dimension  $q \times 1$ .

Maintenant, définissons un indice  $I(\alpha)$ , qui dépend de la direction prise par la projection, c'est-à-dire du projecteur  $\alpha$ , et maximisons cet indice sous la contrainte  $\alpha^T \alpha = 1$ . Si l'on utilise un estimateur de densité tel qu'il a été défini plus haut, la forme empirique de

notre indice peut prendre la forme suivante (Friedman et Tukey, 1974) :

$$\hat{I}_{FT,h}(\alpha) = \frac{1}{n^2 h} \sum_{i=1}^n \sum_{j=1}^n K \left( \frac{\alpha^T (x_i - x_j)}{h} \right).$$

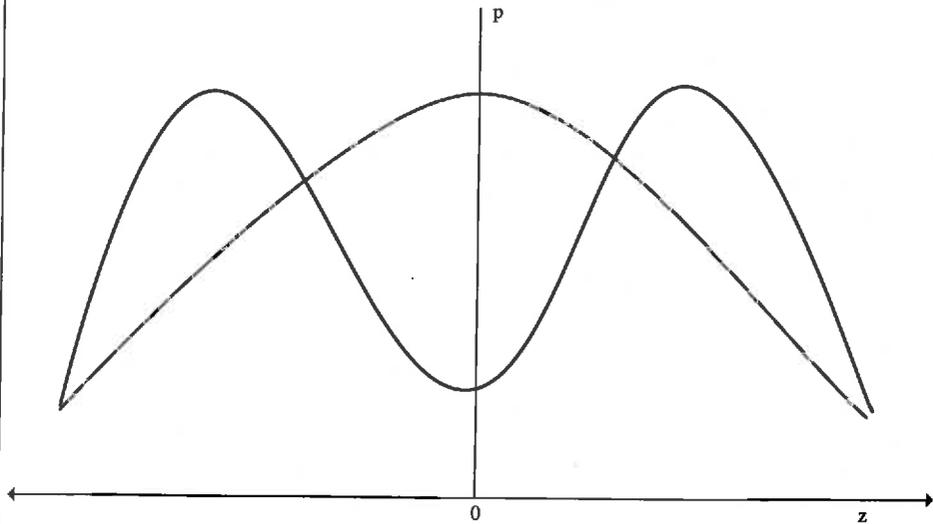
$\hat{I}_{FT,h}$  est en fait un estimateur de :

$$I_{FT} = \int_{\mathbb{R}} p_Z^2(z) dz, \text{ où } p_Z \text{ est la fonction de densité de } Z.$$

Littéralement, cet indice est la somme au carré des densités aux points projetés sélectionnés. Si les données ont été transformées de sorte qu'elles sont de moyenne nulle et de variance unitaire, la projection sélectionnée hérite de ces propriétés. Ainsi, l'indice  $I(\alpha)$  sera maximum lorsqu'il déviara d'une forme parabolique révélant alors la structure présente dans l'espace multidimensionnel d'origine, si elle existe. Le graphique ci-dessous donne une intuition particulièrement claire de l'enjeu de cette technique par rapport à une simple analyse de la distribution des revenus. La densité en gras conduira clairement à un indice supérieur à celui correspondant à une densité qui tend à prendre une forme parabolique.

GRAPHIQUE 10

Un exemple d'une projection "intéressante" de l'espace original (en gras) en opposition à une projection qui ne révèle aucune structure : c'est la parabole



Notons que ce graphique où seulement deux groupes d'individus émergent est une représentation purement imaginaire de ce que l'on entend par une réduction utile de l'espace original qui décrit l'ensemble des fondamentaux économiques propres à chaque pays. Les modes ne reflètent plus l'évolution de la distribution des revenus, mais bien les éventuels champs d'attraction vers des états stationnaires différents, où ces derniers sont définis par la prise en compte simultanée des choix en termes d'allocation et d'accumulation des ressources des nations présentes dans l'échantillon.

Dans cet article, nous utilisons en fait l'indice construit par Cook & al. (1993). Cet indice appelé "Natural Hermite" mesure les déviations de la distribution d'une projection



pondération qui privilégie les différences au centre de la distribution, ce qui permet d'éliminer l'influence que peuvent avoir sur l'indice des observations trop éloignées du corps principal de l'espace d'origine.

L'estimation de cet indice nécessite le recours à des séries polynomiales. Ces séries sont d'ordre infini, et il faut donc spécifier un ordre, dont le rôle est identique à celui joué par le paramètre de lissage dans l'indice de Friedman et Tukey. La structure qui émerge est stable pour des ordres variant de 5 à 9. L'ordre utilisé ici, est égal à 7.

Enfin, le processus d'exploration d'une éventuelle structure locale peut être amélioré par une transformation initiale des données de telle sorte qu'elles soient de moyenne nulle et qu'elles aient une matrice de variance-covariance égale à la matrice Identité. En conséquence, lorsque l'on optimise  $I_{NH}$ ,  $z$  est en fait égal à  $\beta^T Y$ , avec  $Y = D^{-1/2} U^T (X - E(X)) \in R^q$ , et où  $U$  et  $D$  résultent d'une décomposition de la matrice de variance-covariance telle que,

$$\sum = E((X - E(X))(X - E(X))^T) = UDU^T.$$

Une fois que l'indice est maximisé,  $\beta$  peut être traduit dans l'espace des variables originales *via* la transformation suivante :  $\alpha = UD^{-1/2}\beta$ . Si toutes les variables ont été transformées avant l'analyse de sorte qu'elles soient de moyenne nulle et de variance unitaire, le coefficient associé à une variable représente son importance relative dans la détermination de la structure qui émerge, et par conséquent des similarités et des différences entre les différentes économies. Ce sont ces coefficients qui sont présentés dans le texte.

A. D.

#### RÉFÉRENCES

- Abramowitz M. (1986), "Catching-up, Forging Ahead, and Falling Behind", *Journal of Economic History*, XLVI (2), pp. 385-406.
- Azariadis C. (1996), "The Economics of Poverty Traps. Part One : Complete Markets", *Journal of Economic Growth*, (4), pp. 449-486.
- Barro R.J. (1991), "Economic Growth in a Cross-Section of Countries", *Quarterly Journal of Economics*, 106 (2), pp. 407-443.
- Barro R.J. & X. Sala-i-Martin (1992), "Convergence", *Journal of Political Economy*, 100 (2), pp. 223-251.
- Baumol W.J. (1986), "Productivity Growth, Convergence, and Welfare : What the Long-Run Data Show", *The American Economic Review*, 76 (5), pp.1072-1085.
- Baumol W.J. & E.N. Wolff (1988), "Productivity Growth, Convergence, and Welfare : Reply", *The American Economic Review*, 78 (5), pp.1155-1159.
- Bowman A.W. (1984), "An Alternative Method of Cross-Validation for the Smoothing of Density Estimates", *Biometrika*, 71, pp.353-360.
- Benhabib J. & J. Gali (1995), "On Growth and Indeterminacy : Some Theory and Evidence", *Carnegie-Rochester Series on Public Policy*.
- Cook D., A. Buja & J. Cabrera (1993), "Projection Pursuit Indexes Based on Orthonormal Function Expansions", *Journal of Computational and Graphical Statistics*, 2 (3), pp.225-250.
- De la Fuente A. (1997), "The Empirics of Growth and Convergence : A Selective Review", *Journal of Economic Dynamics and Control*, 21, pp.23-73.

- De Long B.J. (1988), "Productivity Growth, Convergence, and Welfare : a Comment", *American Economic Review*, 78 (5), pp.1138-1154.
- De Long B.J. & L.H. Summers (1991), "Equipment Investment and Economic Growth", *Quarterly Journal of Economics*, 106(2), pp.445-502.
- De Long B.J. & L.H. Summers (1992), "Equipment Investment and Economic Growth : How Strong is the Nexus ?" *Brookings Papers on Economic Activity*, 2, pp.157-211.
- De Long B.J. & L.H. Summers (1993), "Equipment Investment in Developing Countries", *Journal of Monetary Economics*, 32, pp.395-415.
- Desdoigts A. (1994), *Changes in the World Income Distribution : A Non Parametric Approach to Challenge the Neoclassical Convergence Argument*, PhD Dissertation, European University Institute, Florence.
- Desdoigts A. (1996), "Smoothing Techniques Applied to a Key Economic Issue : The Convergence Hypothesis", *Computational Statistics*, 11, pp. 481-494.
- Desdoigts A. (1996), *Patterns of Economic Development and the Formation of Clubs*, Miméo, Université d'Evry-Val d'Essonne.
- Dowrick S. & D.T. Nguyen (1989), "OECD Comparative Economic Growth 1950-85 : Catch-Up and Convergence", *American Economic Review*, 79 (5), pp. 1010-1030.
- Durlauf S.N. & P.A. Johnson (1995), "Multiple Regimes and Cross-Country Growth Behavior", *Journal of Applied Econometrics*, 10, pp. 365-384.
- Friedman J.H. & J.W. Tukey (1974), "A Projection Pursuit Algorithm for Exploratory Data Analysis", *IEEE Transactions on Computers*, C 23, pp. 881-890.
- Galor O. (1996), "Convergence ? Inferences from Theoretical Models", *Economic Journal*, 106 (437), pp. 1056-1069.
- Härdle W. (1990), *Smoothing Techniques : With Implementation in S*, Springer Verlag, New York Inc.
- Hall R.E. & C.I. Jones, (1997), "Levels of Economic Activity Across Countries", *American Economic Review* (Papers and Proceedings), pp. 173-177.
- Inglehart R. (1997), *Modernization and Postmodernization : Cultural, Economic, and Political Change in 43 societies*, Princeton University Press.
- Jones M.C. & R. Sibson (1987), "What is Projection Pursuit ?" (with discussion) *Journal of the Royal Statistical Society*, Series A, 150 (1), pp. 1-36.
- Krugman P. (1991), "History versus Expectations", *Quarterly Journal of Economics*, 106, pp. 651-667.
- Lucas Jr R.E. (1988), "On the Mechanics of Economic Development", *Journal of Monetary Economics*, 22, pp. 3-42.
- Mankiw N.G., D. Romer & D. Weil (1992) "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics*, 107 (2), pp. 407-437.
- Murphy K.M., A. Shleifer & R.W. Vishny (1989) "Industrialization and the Big Push", *Journal of Political Economy*, 97 (5) : 1003-26.
- Quah D. (1993), "Galton's Fallacy and Tests of the Convergence Hypothesis", *Scandinavian Journal of Economics*, 95 (4), pp. 427-443.
- Quah D. (1996), "Twin Peaks : Growth and Convergence in Models of Distribution Dynamics", *Economic Journal*, 106 (437), pp. 1045-1055.
- Quah D. (1997), "Empirics for Growth and Distribution : Stratification, Polarization, and Convergence Clubs", *CEPR Discussion Paper No. 1586*, Londres : CEPR.
- Romer P. (1986), "Increasing Returns and Long Run Growth", *Journal of Political Economy*, 99 (juin), pp. 500-521.

- Rosenblatt M. (1956), "Remarks on Some Non Parametric Estimates of a Density Function", *The Annals of Mathematical Statistics*, 27, pp. 832-837.
- Rudemo M. (1982), "Empirical Choice of Histograms and Kernel Density Estimators", *Scandinavian Journal of Statistics*, 9, pp. 65-78.
- Sala-i-Martin X. (1996), "The Classical Approach to Convergence Analysis", *Economic Journal*, 106 (437), pp. 1019-1036.
- Sala-i-Martin X. (1997), "I Just Ran Two Millions Regressions", *American Economic Review (Papers and Proceedings)*, pp. 178-183.
- Silverman B.W. (1986), "Density Estimation for Statistics and Data Analysis", *Chapman and Hall*, Londres.
- Solow R.M. (1956), "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, 70, pp. 65-94.
- Stone C.J. (1984), "An Asymptotically Optimal Window Selection Rule for Kernel Density Estimates", *Annals of Statistics*, 12, pp.1285-1297.
- Summers R. & A. Heston (1991), "The Penn World Table (Mark 5) : An Expanded Set of International Comparisons, 1950-1988", *Quarterly Journal of Economics*, 106 (2), pp. 327-368.
- Weber M. (1964), *L'éthique protestante et l'esprit du capitalisme*, Edition Plon.