

**Institut National de la Planification  
et de la Statistique d'Alger, Algérie  
(I.N.P.S.)**

**Université JEAN MONNET de  
SAINT-ETIENNE, France  
(U.J.M.)**

**Thèse de co-tutelle**

***Convergence régionale de l'emploi et dépendances  
spatiales :  
Le cas de l'Algérie  
Approche par l'économétrie Spatiale***

**Thèse pour le Doctorat en Economie**

**Présentée : Yacine BELARBI**

**Directeurs de thèse :**

Monsieur **Mohamed Yacine FERFERA**  
Professeur de Sciences Économiques à l'INPS

Monsieur **Khaled BOUABDALLAH**  
Professeur de Sciences Économiques à l'UJM

**Membres du Jury**

Monsieur **Youcef BENABDALLAH**  
Professeur de Sciences Économiques à l'INPS

**Rapporteur**

Mme **Julie LE GALLO**  
Professeur de Sciences Économiques Université de Franche-Comté CRESE.

**Rapporteur**

Monsieur **Cem ERTUR**  
Professeur des Universités, Université d'ORLEANS

**Membre**

Monsieur **Abdallah ZOUACHE**  
Maître de conférences (HDR), UJM

**Membre**

***Juin 2009***

## Table des matières

<b>INTRODUCTION GENERALE .....</b>	<b>3</b>
<b>CHAPITRE 1 : CROISSANCE, GEOGRAPHIE ECONOMIQUE ET CONVERGENCE :</b>	
<b>MODELISATION ET MESURE. ....</b>	<b>9</b>
<b>INTRODUCTION : .....</b>	<b>9</b>
<b>1- CONVERGENCE DANS LES MODELES DE CROISSANCE : .....</b>	<b>11</b>
1-1 LE MODELE NEO-CLASSIQUE : .....	11
1-1-1 <i>Le processus de convergence</i> : .....	14
1-1-2 <i>Vitesse de convergence</i> : .....	16
1-1-3 L'EFFET DU CAPITAL HUMAIN SUR LA CONVERGENCE : .....	18
1-2 LA CONVERGENCE DANS LES MODELES DE CROISSANCE ENDOGENE : .....	22
1-2-1 <i>Le modèle AK</i> : .....	23
1-2-2 <i>La théorie basée sur l'innovation</i> : .....	25
1-2-3 <i>Les défis empiriques de la convergence</i> : .....	28
1-3 L'ECONOMETRIE DE LA CONVERGENCE : .....	29
1-3-1 <i>la convergence et les conditions initiales</i> : .....	29
1-3-2 <i>La <math>\beta</math> – convergence</i> : .....	31
1-3-3 <i>La <math>\beta</math> – convergence et la divergence économique</i> .....	32
1-4 LA CROISSANCE ET LA GEOGRAPHIE ECONOMIQUE : .....	34
1-4-1 <i>les théories de l'espace et de la localisation</i> : .....	35
1-4-2 <i>L'économie géographique et l'explication de la formation des agglomérations</i> : .....	35
1-4-3 <i>Effet de la taille du marché et la structure centre périphérique</i> : .....	39
A) <i>L'effet de la taille du marché</i> : .....	40
B) <i>La structure centre – périphérie</i> .....	44
C) <i>Croissance et agglomération</i> : .....	50
C-1) <i>le modèle de Martin et Ottaviano (99)</i> .....	50
C-2) <i>le modèle d'Englmann et Walz (1995)</i> .....	52
<b>CONCLUSION.....</b>	<b>53</b>
<b>CHAPITRE II : ECONOMETRIE SPATIALE.....</b>	<b>56</b>
<b>INTRODUCTION : .....</b>	<b>56</b>
2.1 <i>Le concept de processus spatial aléatoire (champ aléatoire)</i> .....	57
2.2 <i>La régression linéaire avec des données spatiales</i> : .....	63
2.3 METHODE D'ESTIMATION BAYESIENNE : .....	71
2.3.1 <i>Les principes de l'économétrie Bayésienne</i> : .....	72
2.3.2 <i>Le choix de distributions a priori</i> .....	74
2.3.3 <i>Estimateurs Bayésiens et fonction de coût</i> : .....	75
2.3.4 <i>Méthodes de calcul des estimateurs</i> .....	77
2.3.5 <i>Méthode de Monte Carlo par chaîne de Markov</i> .....	77

2.3.6 Algorithme de Metropolis – Hastings .....	83
2.3.7 Echantillonnage de Gibbs : .....	85
CONCLUSION : .....	87
<b>CHAPITRE III.....</b>	<b>89</b>
<b>ETAT DE L’ESPACE ET DE L’ECONOMIE EN ALGERIE .....</b>	<b>89</b>
INTRODUCTION : .....	89
1- TENDANCE MACROECONOMIQUE .....	89
2- HISTORIQUE DE LA CREATION DES WILAYATES : .....	93
3- DESCRIPTION SOCIODEMOGRAPHIQUE DE L’ESPACE : .....	98
<i>A - Photographie sociodémographique des wilayates : .....</i>	<i>98</i>
<i>B- Distribution spatiale des ressources humaines .....</i>	<i>108</i>
<i>C - Infrastructures de logement et de transport .....</i>	<i>110</i>
<i>F- Infrastructures de formation .....</i>	<i>112</i>
<b>CONCLUSION : .....</b>	<b>115</b>
<b>CHAPITRE IV : ANALYSE ECONOMETRIQUE DE LA CONVERGENCE SPATIALE ENTRE LES WILAYATES .....</b>	<b>116</b>
INTRODUCTION .....	116
1. LE CADRE THEORIQUE.....	117
2. CROISSANCE REGIONALE DE L’EMPLOI ET DEPENDANCES SPATIALES.....	121
3. L’EFFET DE LA CROISSANCE DE L’EMPLOI PAR TETE DES REGIONS VOISINES .....	124
<i>3.1. Dépendances spatiales et convergence régionale.....</i>	<i>127</i>
<i>3.2. Dépendances spatiales et chocs communs régionaux.....</i>	<i>131</i>
CONCLUSION .....	133
<b>CHAPITRE V : L’HETEROGENEITE SPATIALE DANS LA CROISSANCE DE L’EMPLOI REGIONALE EN ALGERIE. ....</b>	<b>135</b>
INTRODUCTION .....	135
1- PRESENTATION DU MODELE BSALE : .....	137
2- ESTIMATION BAYESIENNE SPATIALE AUTOREGRESSIVE LOCALE (BAYESIAN SPATIAL AUTOREGRESSIVE LOCAL ESTIMATION BSALE) : .....	137
3-ESTIMATION DU MODELE.....	139
4- L’ECHANTILLONNEUR DE MCMC : .....	140
5- DONNES : .....	141
6- RESULTATS ET INTERPRETATIONS : .....	142
6- CONCLUSION : .....	144
<b>CONCLUSION GENERALE .....</b>	<b>146</b>
ANNEXE 1 : MODELE SPATIAL BAYESIEN HETEROSCEDASTIQUE : .....	150
ANNEXE 2 : RESULTATS D’ESTIMATION PAR L’APPLICATION DU MODELE BSALE.....	154
<b>BIBLIOGRAPHIE.....</b>	<b>157</b>

## INTRODUCTION GENERALE

Depuis le lancement des réformes économiques structurelles, au début des années 80, l'économie algérienne observe un passage d'une économie centralisée vers une économie du marché. L'adoption des mécanismes du marché a été fortement soutenue par les accords, engagés dans le cadre du plan d'ajustement structurel, avec les institutions financières internationales. La libéralisation de l'activité économique nationale devrait conduire au développement et à la consolidation de l'investissement privé. Un tel processus ne peut se produire sans interactions avec le développement des zones d'activités composant l'espace économique algérien.

Les premières observations sur les zones d'activités (régions) dévoilent des dynamiques différentes dans l'intensité de l'investissement, les taux de création d'entreprises et les types d'activités développées.

L'objectif recherché, à travers la formulation de cette problématique, est d'explorer les éléments théoriques et empiriques qui nous aideront à analyser l'origine des différences dans les dynamiques observées. En d'autres termes, le processus de convergence entre les régions économiques en Algérie. Les questions fondamentales qui orienteront ce travail sont les suivantes : *Pourquoi les dynamiques de la croissance régionale des activités sont différentes ? Y a-t-il une tendance vers la persistance des disparités dans la répartition du développement ? Quelles sont les variables qui conditionnent un tel processus de convergence ?*

La littérature classique en économie de développement résume les changements par la lecture des tendances d'évolution dans les disparités entre les régions. Dans la courbe de Kuznet et tel que montré par Williamson (1965), la différence régionale dans le revenu augmente, se stabilise, puis diminue dans les périodes naturelles de croissance. Les modèles de croissance donnent uniquement l'explication de la croissance des disparités régionales. En effet, le modèle de Solow – Swan prévoit que sous la condition d'un taux fixe d'épargne, un faible ratio du capital par tête est associé à une croissance proportionnellement rapide du stock de capital vers le sentier de croissance équilibrée. Ce résultat implique que les économies pauvres tendent à avoir un taux de croissance par tête plus rapide que celui des économies riches. Un tel phénomène est dit : convergence. Plusieurs applications empiriques, ( USA 1880-1990, Japan 1930-1990, 90 régions des pays de l'Europe) montrent une forte convergence régionale absolue. Néanmoins, la théorie de la croissance n'explique pas la

persistance des disparités (ou divergences) dans les pays en voie de développement (exemple : les applications de Treyz 1995, Jian et al 1996).

La théorie de l'agglomération fournit une alternative de réponse pour l'explication du phénomène de la persistance des disparités dans les pays en voie de développement. Selon cette théorie, les firmes, et les activités de production en général, tendent à se localiser en proximité. Les économies d'agglomération sont engendrées par les externalités technologiques et pécuniaires, communication et diversité dans la production et des inputs. En effet, le modèle d'économie géographique de Krugman (1991) montre que la combinaison des rendements croissants, la baisse du coût de transport et les externalités technologiques peuvent engendrer un processus d'agglomération.

Selon Maurel (1996), l'ensemble des facteurs analysés par la théorie de l'agglomération peuvent être du côté de la demande comme : l'importance de la taille d'une ville. Celle-ci, engendre une diversité dans la demande et du coup incite les entreprises à augmenter la diversité des outputs. Du côté de l'offre, la diversité des inputs, la qualification de la main d'œuvre et les politiques économiques locales sont importantes pour l'émergence d'une agglomération.

L'analyse de Henderson (1986 et 1988) sur les activités économiques Brésiliennes a clairement montrée l'effet des économies d'agglomérations sur les gains de productivité. Mills et Becker (1986) et Becker et al (1992), montrent que ces gains augmentent à chaque fois que la taille des villes augmente. Fujita et Hu (2001), montrent que la libéralisation et la globalisation économique peuvent offrir beaucoup de chance à la convergence et l'agglomération. L'ouverture sur l'économie mondiale permet d'exploiter l'avantage géographique des zones périphériques en leur permettant d'intégrer le marché mondial. Cet avantage engendre, par contre, deux effets :

- Si les activités de production initiales étaient concentrées dans les régions centrales, l'avantage de localisation des zones périphériques peut réduire les forces d'agglomération vers les régions du centre.
- Si les activités de production initiales étaient concentrées dans les régions périphériques, la globalisation renforce les forces d'agglomération vers ces régions et crée des disparités régionales entre le centre et la périphérie.

Les causes des inégalités entre le développement des régions peuvent avoir pour origine : des politiques préférentielles de développement et qui conduisent les régions développées à opérer dans un environnement économique soutenu par les institutions publiques, ce qui permet aux régions préférées de devenir plus attractives à l'investissement. Des positions géographiques permettant un accès avec des coûts faibles de transport. Des tailles importantes des agglomérations (urbanisation) et des structures industrielles diversifiées. Une prédominance d'un capital humain de qualité.

Les études sur l'expérience Américaine montrent que la convergence se produit si avec des institutions juridiques et économiques sont similaires, des niveaux d'éducation et de santé publique comparable, et avec une libre migration du capital et de la main d'œuvre.

Fleicher et Chen (1997), Mody et Wang (1997) et Démurger (2001), en se basant sur les différences régionales dans l'investissement en infrastructure, ils ont trouvé que ces différences ont un impact statistiquement significatif sur la croissance des régions. Hell et Jones (96), Chang (94) et Ravallion et Jalan (96 et 98), ont testé l'importance des variables géographiques, institutionnelles et politiques dans l'explication des disparités régionales dans le revenu. Selon ces auteurs les différences dans le climat, les infrastructures, l'accès aux services publics et méconnaissance de la réalité physique du territoire et des technologies adéquates peuvent affecter les rendements marginaux du capital. Aussi, l'imperfection dans la mobilité de la main d'œuvre peut créer les conditions de la persistance des inégalités.

Selon Krugman (1998) deux facteurs peuvent être déterminants pour l'importance économique d'une ville : la taille de la population urbaine et la structure politique du système. Un système centralisé a tendance à favoriser l'apparition d'une méga-ville.

Les prédictions de convergence déduites du modèle Slow - Swan ont été remises en cause par la théorie de la croissance endogène. Les modèles d'économie géographique montrent que l'intégration économique s'accompagne le plus souvent de polarisation croissante des territoires. Dans le modèle CP (Krugman 1990), une demande de consommation finale plus importante dans les régions centrales, auto renforce la migration des travailleurs, et du coup la disparité régionale. Deux critiques sont adressées à ce modèle. Cependant, deux insuffisances ont été mises en évidence : La mobilité du travail n'explique pas la formation de grandes régions industrielles dans des économies caractérisées par une faible mobilité de la main d'œuvre ; la demande de biens intermédiaire, plus importante que la demande de consommation finale, fait que la localisation des producteurs de biens intermédiaires conditionne celle des producteurs de biens finaux et vice versa. la concentration des entreprises

du secteur finale, attirées par le faible coût des biens intermédiaire dans les régions centrales, peut enclencher un processus d'agglomération cumulatif.

La mobilité du travail qualifié proposée par Krugman ne permet pas, à elle seule, l'apparition des disparités entre les régions économiques. Le modèle de Krugman et Venables (1995) en jouant sur la mobilité des entreprises, au lieu de celle de la main d'œuvre qualifiée, montrent que quand les entreprises se localisent dans une région sans une augmentation correspondante de l'offre de travail, les rémunérations augmentent ce qui induit deux forces dans le phénomène d'agglomération : 1) La demande finale augmente au centre puisque les consommateurs disposent de revenus plus élevés, cette force d'agglomération est liée à la demande finale. Chez Krugman, cette force est provoquée par l'augmentation de la taille de la population. 2) L'enrichissement du coût du travail induit une force de dispersion. Les ETP peuvent trouver des avantages de se localiser en périphérie, car le salaire est moins cher et ce même si la demande est moindre quand le coût de transfert est faible.

Avec une telle approche, le renforcement des inégalités régionales, continue d'apparaître jusqu'à un certain niveau d'intégration, au-delà de qui un processus inverse s'enclenche pour conduire à l'affaiblissement des inégalités régionales par la réindustrialisation de la périphérie.

De nombreux travaux montrent qu'une concentration spatiale croissante irait de paire avec les premières étapes du processus de développement économique, et plus ce processus avance plus on assiste à un redéploiement spatial des activités. Ceci rejoint le résultat montré par la courbe de Kuznet. Cette forme en cloche, mise en évidence par le modèle de Venables et Krugman, reliant intégration économique et inégalité régionale, rend la convergence un phénomène conditionnel qui s'inscrit dans un processus temporel de développement. Ce résultat semble mieux adapté au cas des pays de l'UE, caractérisé par de faibles migrations internationales et une importance dans la consommation des biens intermédiaires dans la production. La convergence telle que citée par Abramovitz (1986) implique une tendance de long terme vers l'égalisation du revenu par tête ou des niveaux de production.

Le modèle de Krugman et Venables (1995) peut expliquer le lien entre localisation géographique et disparité économique entre les régions (pays) de l'U.E, mais peut – il expliquer les disparités observées dans les régions des pays en voie de développement, et par voie de conséquence prévoir le processus de convergence.

Globalement, notre réflexion va aborder les caractéristiques qui permettent de déceler les disparités géographiques entre les régions d'activité en Algérie, les outils qui autorisent une meilleure prise en compte des disparités géographiques dans l'analyse de la convergence des économies locales, et le poids de ces disparités sur le processus de convergence. La prise en compte des disparités spatiale dans la croissance des régions nous impose la mise au point de modèles appropriés pour tester la convergence. Le choix d'un modèle permettra d'illustrer les différents débats théoriques et les applications empiriques sur la question, et l'importance de la prise en compte de la dimension spatiale dans le cadre de ces modèles de convergence.

Ce travail est composé de quatre chapitres. Dans le premier chapitre de ce travail nous allons situer le cadre théorique dans lequel ont s'inscrit, et présenter les outils théoriques utilisés. Dans cette partie, Il s'agit de présenter un aperçu de la littérature sur la croissance économique et le rôle de l'apport de la nouvelle économie géographique dans l'explication des disparités géographiques de la croissance et de la convergence économique. Les différentes synthèses de la littérature présentées au cours de ce chapitre vont nous permettre de retracer les débats les plus récents sur la question de la convergence économique et de donner plus d'éclaircissement sur notre démarche.

Dans le deuxième chapitre nous aborderons les techniques de l'économétrie spatiale utilisée dans les applications développée dans cette thèse. Le développement de l'économétrie spatiale, relativement récent dans la littérature statistique, permet de répondre à la violation des hypothèses classiques, à cause des similarités entre observations causées par la proximité spatiale. L'économétrie spatiale permet d'expliquer le phénomène de diffusion spatial rencontré dans plusieurs domaines tels que : épidémiologie, géographie, agriculture, traitement de l'image, sciences régionales, astronomie, et dans d'autres domaines (Haining, 2003). Ainsi, nous développerons les concepts de base liés à l'économétrie spatiale, puis nous introduirons les techniques de l'approche Bayésienne utilisée dans les modélisations faites dans cette thèse.

Dans le troisième chapitre nous discuterons de l'économie Algérienne ainsi que sur les caractéristiques socioéconomique de l'espace algérien. Nous présenterons les caractéristiques des différentes périodes traversées par l'économie Algérienne. Nous analyserons les déterminants macroéconomiques de la croissance de l'économie, que celles des différents secteurs économique. Un aperçu sur la distribution spatiale des facteurs humains et socioéconomiques, sera présenté avec une photographie sociodémographique des wilayates, et

les potentialités de chacune d'elles en termes de ressources humaines et d'infrastructure susceptibles d'être à la base de l'attractivité et de la compétitivité de chaque espace.

Le quatrième et le cinquième chapitre seront consacrés à deux applications sur l'étude des déterminants de la croissance régionale de l'emploi en Algérie. Dans la première application nous examinons les forces économiques qui peuvent expliquer la croissance de l'activité économique des régions Algériennes lors des huit dernières années. Et dont la démarche trouve ses origines dans les approches développées dans les travaux de Glaeser, Kallal, Scheinkman et Shleifer (1992), Henderson, Kuncoro et Tuner (1995), et d'une manière indirecte dans le travail de Rappaport (1999). Glaeser et al. On s'interroge sur la relation entre la croissance économique de l'emploi manufacturier par tête dans 48 régions d'Algérie et la localisation géographique de ces régions (en terme de voisinage régional immédiat), l'emploi par tête d'habitant initial, la structure de l'industrie locale, l'intensité de la concurrence locale, la densité de la population, et le niveau d'éducation de la population des adultes. Dans la deuxième application, nous étudierons le phénomène de l'hétérogénéité spatiale et qui se traduit par une instabilité spatiale des paramètres de la régression, les modèles ne sont pas stationnaires dans l'espace et les paramètres varient systématiquement avec la localisation. Nous mobiliserons pour ce cas des modélisations qui permettent de prendre en compte les caractéristiques particulières de chaque localisation de l'échantillon. Les unités spatiales ne sont ni de formes régulières, ni homogènes : les régions peuvent avoir des formes et des airs différents, des niveaux de développement technologique variables, des populations plus ou moins importantes, etc. Plusieurs travaux économétriques sur la convergence ont soulevés les problèmes économétriques induits par l'hétérogénéité. Pour remédier à ce problème, nous introduisant une analyse basée sur la notion de la convergence locale avec une estimation bayésienne du modèle spatial autorégressif local (Bayesian Spatial Autoregressive Local Estimation).

Enfin, pour terminer nous conclurons ce travail par une reprise synthétique des différents résultats obtenus.

# **Chapitre 1 : Croissance, géographie économique et convergence : Modélisation et mesure.**

---

## **Introduction :**

L'objet de ce chapitre est de situer le cadre théorique de notre travail, et de présenter les fondements théoriques utilisés dans ce travail. Dans cette partie, Il s'agit de présenter un aperçu de la littérature sur la croissance économique et le rôle de l'apport de la nouvelle économie géographique dans l'explication des disparités géographiques de la croissance et de la convergence économique.

La théorie à l'origine des analyses de la croissance régionale peut être appréciée comme une réaction ou une adaptation des hypothèses fondamentales de la théorie néo-classique de la croissance, telle qu'elle a été décrite par Solow (1956). Les améliorations apportées, telle que l'hypothèse des rendements ou du sentier de la croissance équilibrée, renvoient à une volonté de rendre la théorie plus proche des considérations de la réalité. Dans l'espoir de réconcilier la théorie avec les réalités empiriques, l'évidence théorique d'une convergence à un sentier de croissance équilibrée unique, notamment pour les pays du monde ou les région de l'Union Européenne, a été reformulée par les travaux de Barro et Sala-i-Martin (1992, 1995) et Barro (1991, 1997). Ces deniers, ainsi que Mankiw et al. (1992) et Revelt (1992), considèrent plusieurs sentiers au lieu d'un sentier unique. On note aussi que Lee et al. (1997) montrent un biais dans l'estimateur du taux de convergence et avancent l'argument selon lequel les tests de signification utilisant la statistique de Student (t - statistics) ne sont pas valides.

L'hypothèse des rendements d'échelle constants est fondamentale dans la théorie néo-classique. Cependant, l'introduction des effets substantiels des économies d'externalités internes et externes (agglomération) sur la production (Fingleton et McCombes, 1998) rend les rendements croissants préférables dans les analyses régionales (théoriques et empiriques). En effet, les politiques d'intégration régionales et la diffusion technologique peuvent booster la croissance de la productivité dans les régions à faible productivité, ce qui entraîne une convergence du niveau régionale de la productivité, et compense l'effet des rendements

croissants qui peuvent conduire à un comportement divergent des économies. Dans le contexte américain, caractérisé par une forte mobilité régionale du capital et du travail, plusieurs applications empiriques (McCombie et Ridder, 1984 ; Bernat, 1996) soutiennent l'hypothèse de rendements d'échelles croissants. Il est commun, même avec le style de régression à la Barro modifiée, de trouver la présence des rendements d'échelles croissants due au capital humain, ce qui va dans le même sens de la théorie de la croissance endogène (Romer, 1986 ; Lucas, 1988) et des théories de la diffusion technologique (Nelson et Phelps, 1966).

La nouvelle économie géographique (Krugman, 1991a,b ; Krugman et Venables, 1995 ; Ottaviano et Puga, 1998 ; Puga et Venables, 1997, 1999)<sup>1</sup> donne une place marquante aux rendements croissants. Cette remise en cause des modèles néoclassiques de commerce international et le développement des théories de la formation de l'espace urbain (Ogawa et Fujita, 1980, 1989 ; Fujita, 1989) ou les modèles d'équilibres inter-urbains (Henderson, 1987 ; Abdel-Rahman, 1994, 1996) donnent beaucoup d'importance au rôle des rendements croissants dans l'explication de la localisation spatiale des agents économiques et la formation et le développement des agglomérations.

Dans les modèles interrégionaux trois résultats théoriques sont importants à mettre en évidence. Ils sont fondés sur l'hypothèse de la concurrence monopolistique, des rendements de production croissants, et l'importance des innovations et des externalités de connaissance. Les agents économiques sont incités à se regrouper dans l'espace, et les disparités régionales sont expliquées par la distribution spatiale des facteurs (travail qualifié et capital). En effet, la persistance de la concentration géographique de régions riches ou de régions pauvres conforte la compréhension de l'interaction entre l'évolution des disparités économiques et les disparités spatiales.

Il existe une masse importante de travaux empiriques relatifs aux hypothèses et prédictions des modèles théoriques, avec un grand nombre de contributions empiriques sur la croissance économique régionale (Neven and Gouyette 1995 ; Sala-i-Martin 1996 ; Carlino et Mills 1993, 1996 ; Bernard et Jones 1996 ; Chatterji et Dewhurst 1996 ; Armsrong 1956). L'ensemble de ces travaux est caractérisé par le fait que les régions sont considérées comme des économies isolées, en d'autres termes, leurs spécifications empiriques excluent les

---

<sup>1</sup> Propose des modèles d'équilibre interrégionaux pour expliquer la formation des agglomérations sans faire appel aux avantages comparatifs.

interactions entre les régions. Cependant, plusieurs arguments théoriques et empiriques suggèrent que les régions ne sont ni homogènes, ni indépendantes (Rey Montouri 1999).

Nous abordons ainsi l'analyse des disparités géographiques selon la démarche suivante : en premier lieu, nous étudions les concepts de convergence associés aux différentes théories de la croissance néoclassique. En deuxième temps, nous montrons comment la prise en compte des résultats de la nouvelle économie géographique peut donner plus de consistance pour la compréhension des disparités géographiques dans la convergence économique.

Les différentes synthèses de la littérature présentées au cours de chapitre permettent de retracer le débat et de donner plus d'éclaircissement sur notre démarche.

## **1- Convergence dans les modèles de croissance :**

Dans cette section, nous abordons les implications du modèle de croissance de la théorie néo-classique et ceux de la théorie endogène sur la convergence des économies. Le modèle néoclassique de Solow-Swan, basé sur des rendements d'échelle constants et une productivité marginale du capital décroissante, prévoit que des économies semblables en termes de technologies et de préférences, convergent vers le même taux de croissance à long terme. Ces prédictions ont été remises en cause par la théorie de la croissance endogène (Romer 1986, Lucas 1988, Aghion et Howitt 1992). Cette dernière montre que la formation du taux de croissance de long terme (exogène dans le modèle néo-classique) dépend des caractéristiques des économies et que les disparités du produit par tête auront tendance à persister dans le long terme.

### **1-1 Le modèle néo-classique :**

Pour le modèle néo – classique de Solow-Swan, une économie converge vers un état de croissance régulier de long terme. Si on considère le modèle (Solow 1956, Swan 1956) d'une économie fermée à deux facteurs de production (capital et travail), dont la consommation des ménages est de type Keynésien (la propension à épargner est une part constante du revenu) et les entreprises produisent un bien qui sert à l'investissement et à la consommation. La concurrence sur le marché des biens et des facteurs est considérée parfaite.

Supposons une fonction de type Cobb – Douglas avec des facteurs substituables à rendements constants :

$$Y_t = f(K_t, A_t L_t) = K_t^\alpha (A_t L_t)^{1-\alpha} \quad 0 < \alpha < 1 \quad (1)$$

$Y_t$  représente le niveau de production à la date  $t$ ,  $K_t$  le stock de capital physique,  $L_t$  la quantité du facteur travail et  $A_t$  le niveau technologique, augmente l'efficacité du travail, considéré neutre au sens de Harrod.  $A_t L_t$  est le travail effectif, composé du travail et de la productivité induite par le progrès technique. Cette fonction de production satisfait la condition de concavité (les productivités marginales des facteurs sont positives et décroissantes), et la production est à rendements d'échelle constants. La forme intensive de cette fonction de production est donnée par :

$$y_t = f(k_t) = k_t^\alpha \quad (2)$$

où  $y_t = Y_t/A_t L_t$  et  $k_t = K_t/A_t L_t$  respectivement production et capital par unité de travail effectif. Les conditions d'Inada (1963), stipulant que la productivité marginale des facteurs tend vers l'infini lorsque ces derniers tendent vers 0 et inversement, sont respectées.

On suppose que le travail et le progrès technique croissent aux taux exogènes respectifs de  $n$  et  $g$  :

$$L_t = L_0 e^{nt} \quad (3)$$

$$A_t = A_0 e^{gt} \quad (4)$$

$A_t L_t$  croît au taux de  $n + g$ .

Le comportement dynamique de l'économie est décrit par les équations suivantes :

- La croissance du stock du capital physique correspondant à un point dans le temps est donné l'investissement moins la dépréciation du capital

$$\dot{K} = \frac{dK_t}{dt} = I - \delta K = s \cdot f(K, AL) - \delta K \quad (5)$$

- Si on divise les deux côtés de l'équation (5) par  $A_t L_t$  on obtient la forme suivante :

$$\begin{aligned} \dot{K}/AL &= s \cdot f(k) - \delta k \\ \dot{k}_t &= \frac{d(K_t/A_t L_t)}{dt} = \dot{K}/AL - (n + g)k \\ \dot{k}_t &= s \cdot f(k) - (n + g + \delta)k \end{aligned} \quad (6)$$

$\delta$  est le taux de dépréciation du capital et  $s$  la part de la production investie. L'équation (6) donne la forme de la variation du stock de capital par unité de travail effectif. Les hypothèses (concavité, rendement marginal décroissant du capital) assurent l'existence, la stabilité et l'unicité de la solution de long terme (état régulier) donnée par  $\dot{k}_t = 0$ . La résolution de l'équation  $s \cdot f(k^*) = (n + g + \delta)k^*$  (avec  $f(k^*) = k^*$ ) détermine le niveau de capital par unités de travail effectif à l'état régulier :

$$k^* = \left( \frac{s}{n + g + \delta} \right)^{\frac{1}{1-\alpha}} \quad (7)$$

soit

$$y^* = \left( \frac{s}{n + g + \delta} \right)^{\frac{\alpha}{1-\alpha}} \quad (8)$$

L'équation (8) montre que la production par tête de l'état régulier dépend positivement du taux d'épargne et négativement du taux de croissance de la population. La prédiction centrale du modèle de Solow concerne l'impact de l'épargne et de la croissance de la population sur la production réelle. La linéarisation de l'équation (8) est donnée par la forme suivante :

$$\ln(y^*) = \frac{\alpha}{1-\alpha} \ln(s) - \frac{\alpha}{1-\alpha} \ln(n + g + \delta) \quad (9)$$

Le niveau de la production par tête est donné par :

$$\ln\left(\frac{Y}{L}\right) = \ln(A_0) + gt + \frac{\alpha}{1-\alpha} \ln(s) - \frac{\alpha}{1-\alpha} \ln(n + g + \delta) \quad (10)$$

Le taux de croissance de la production par tête, donnée par  $\frac{d \ln\left(\frac{Y}{L}\right)}{dt} = g$ , dépend du progrès technique. Par hypothèse, le modèle de Solow conduit à l'égalisation des taux de croissance de long terme des différentes économies puisque le progrès technique est considéré comme un bien non rival et accessible à tous. En l'absence du progrès technique ( $g = 0$ ), et à cause de l'hypothèse de la décroissance de la productivité marginale du capital, le capital et la production sont constants à long terme.

Du point de vue de la dynamique de transition vers l'état régulier (convergence), le modèle possède des implications optimistes. En effet, les économies les plus pauvres

épargnent plus et présentent des rendements de l'investissement plus importants, ce qui les rapproche graduellement vers les économies les plus riches. Ces dernières présentent des stocks de capital par tête importants et des rendements plus faibles, par conséquent une accumulation du capital par tête plus lente que dans les pays pauvres. Ces conditions conduisent à une convergence pour le même niveau de production et de capital par tête, et donc une convergence vers le même taux de croissance par tête de l'état régulier. Ce dernier est égal au taux de croissance du progrès technique. L'hypothèse de la convergence fait que les différences dans les taux de croissance observés pour les économies du monde ne sont que transitoires. Cependant, cette conclusion n'est valide que si les économies possèdent des paramètres structurels identiques. Dans le cas contraire, chaque économie converge vers son propre état régulier.

### **1-1-1 Le processus de convergence :**

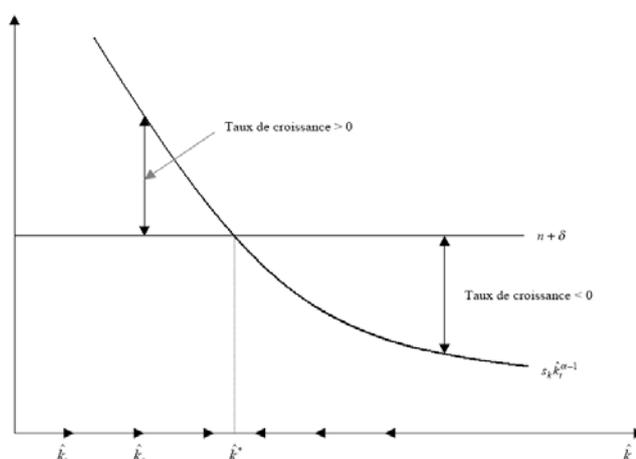
Considérons l'équation de dynamique du modèle de croissance néoclassique. Le taux de croissance du capital par unité de travail effectif  $\hat{k}_t$ , s'obtient comme suit :

$$\frac{\dot{k}_t}{k_t} = s.k_t^{\alpha-1} - (n + g + \delta) = g_k \quad (11)$$

Selon cette équation, le stock de capital par unité de travail effective d'une économie croît tant que  $s.k_t^{\alpha-1}$  est supérieur à  $(n + g + \delta)$ .

Dans le cas de deux économies possédant la même fonction de production et les mêmes caractéristiques structurelles  $n, \delta$  et  $s$ , avec des niveaux initiaux de stock de capital par unité de travail effectif différents. Les deux économies vont converger vers le même point d'équilibre  $k^*$ .

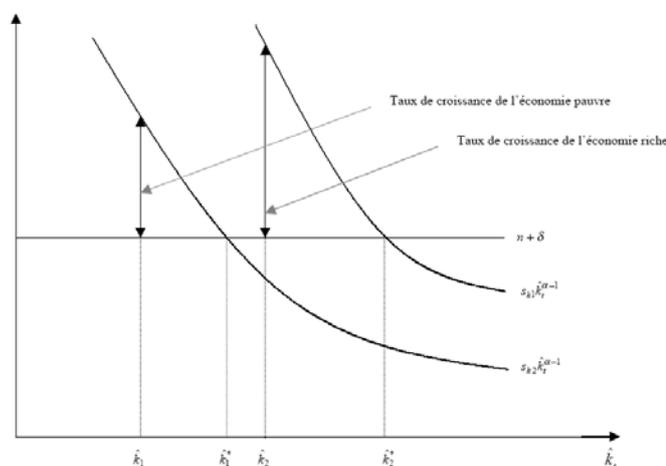
**Figure 1 : Dynamique de croissance dans le modèle Solow-Swan : Economies identiques**



La dotation en capital des économies pauvres croît plus vite que celle des économies riches, ce processus de rattrapage décrit la convergence en dotation de capital par tête. Cette propriété peut facilement être étendue à la production par tête  $y$ . Cette description de la convergence absolue implique que quelque soit le niveau initial du capital ou de la production par tête, les économies pauvres convergent avec des vitesses aussi importantes que la distance qui les sépare de l'état d'équilibre est grande.

Si les deux économies sont conditionnées par des paramètres structurels différents, elles ne convergent pas vers le même état régulier. Des taux d'épargne différents ( $s_{k_1} < s_{k_2}$ ), impliqueront des quantités différentes de capital par unités de travail effectif à l'état régulier  $k_1^* < k_2^*$ .

**Figure 2: Dynamique de croissance dans le modèle Solow-Swan : Economies différentes**



La convergence vers des états réguliers différents, mais avec un même taux de croissance à l'état régulier, est dite convergence conditionnelle.

### 1-1-2 Vitesse de convergence :

On pratique on ne s'intéresse pas uniquement à l'effet éventuel d'une variation donnée (comme celle du taux d'épargne) mais aussi à la rapidité de ces effets. Pour cela, nous recourons à des approximations au voisinage de l'équilibre à long terme. Un développement de Taylor de l'ordre de 1 de l'expression (11) au voisinage de  $\ln(\hat{y}_t) = \ln(\hat{y}^*)$ , on obtient :

$$\frac{\dot{k}_t}{k_t} \simeq s \cdot (\alpha - 1) k^{*(\alpha-1)} (\ln(k_t) - \ln(k^*)) \quad (11)$$

Si on remplace  $s$  par son expression en fonction de  $k^*$  en utilisant (7), on obtient l'expression de l'économie autour de son état régulier :

$$\begin{aligned} \frac{\dot{k}_t}{k_t} &= \frac{d \ln(k_t)}{dt} \simeq -\gamma (\ln(k_t) - \ln(k^*)) \quad (12) \\ \lambda &= (1 - \alpha)(n + g + \delta) \end{aligned}$$

$\gamma$  mesure la vitesse de convergence vers son état régulier. Cette même vitesse est obtenue pour la production par unité de travail effectif puisque  $y_t = k_t^\alpha$  :

$$\frac{d \ln(y_t)}{dt} = \alpha \cdot \frac{d \ln(k_t)}{dt} \simeq -\lambda (\ln(y_t) - \ln(y^*)) \quad (13)$$

La résolution de l'équation différentielle (13) donne l'évolution de la production par unité de travail effectif :

$$\ln(y_t) = (1 - e^{-\lambda t}) \ln(y^*) + e^{-\lambda t} \ln(y_0) \quad (14)$$

$y_0$  représente le niveau de production par unité de travail effectif à la date initiale. Si on remplace  $y^*$  par son expression (8), on obtient :

$$\begin{aligned} \ln(y_t) - \ln(y_0) &= (1 - e^{-\lambda t}) \frac{\alpha}{1 - \alpha} \ln(s) \\ &\quad - (1 - e^{-\lambda t}) \frac{\alpha}{1 - \alpha} \ln(n + g + \delta) - (1 - e^{-\lambda t}) \ln(y_0) \quad (15) \end{aligned}$$

L'évolution de la production par tête est donnée par :

$$\ln\left(\frac{Y_t}{L_t}\right) - \ln\left(\frac{Y_0}{L_0}\right) = (1 - e^{-\lambda t}) \ln(A_0) + gt + (1 - e^{-\lambda t}) \frac{\alpha}{1 - \alpha} \ln(s) - (1 - e^{-\lambda t}) \frac{\alpha}{1 - \alpha} \ln(n + g + \delta) - (1 - e^{-\lambda t}) \ln\left(\frac{Y_0}{L_0}\right) \quad (17)$$

Cette forme estimable (Barro et Sala-i-Martin, 1991, 1995) de l'équation (17) est la base du test de  $\beta$ -convergence, où  $\beta$  est le coefficient de régression reliant le niveau initial du produit par tête au taux de croissance. Le taux de croissance dépend négativement du niveau initial du produit par tête.

La présentation du modèle néo-classique nous permet soulever ses caractéristiques de base de l'état régulier telles que résumées par Mankiw, 1995 :

1. la convergence d'une économie vers son état régulier de long terme est indépendante des conditions initiales ;
2. Le niveau de la production par tête dépend du taux d'épargne et du taux de la croissance de la population. Plus le taux d'épargne est important, plus le niveau de la production par tête de l'état régulier est important. A l'inverse, plus le taux de croissance de la population est important, plus le niveau de la production par tête à l'état régulier est faible;
3. Le taux de croissance de la production par tête de long terme dépend uniquement du taux du progrès technique ;
4. Le stock de capital croît au même taux que la production ;
5. La productivité marginale du capital est constante et la productivité marginale du travail croît au taux du progrès technique.

Selon Mankiw et al. (92), l'examen des données d'un grand échantillon de pays montre que le modèle de Solow prédit bien le sens de l'effet de l'épargne et de la croissance de la population sur la croissance, mais qu'il ne le fait pas correctement pour la magnitude. Deux raisons sont avancées en faveur du rôle potentiel du stock du capital humain pour corriger la magnitude de l'effet de l'épargne et de la croissance de la population. Premièrement, l'accumulation du capital physique et la croissance de la population ont un grand impact sur le revenu quand l'accumulation du capital humain est prise en compte. Deuxièmement : l'accumulation du capital humain peut être corrélée avec les taux d'épargnes et les taux de croissance de la population. Dans ce cas, l'omission de l'accumulation du

capital humain biaise les coefficients estimés de l'épargne et de la croissance de la population. Mankiw et al. (1992) proposent un modèle de Solow augmenté par l'incorporation, en plus du capital physique, le capital humain dans la fonction de production.

### 1-1-3 L'effet du capital humain sur la convergence :

Les économistes ont longtemps insisté sur l'importance du capital humain dans le processus de la croissance. En effet, ignorer le capital humain peut conduire à des conclusions incorrectes. Kendrick (1976) estime que le capital humain représente plus de la moitié du total du stock du capital des Etats Unis en 1969.

L'introduction du stock du capital humain dans la fonction de production de Solow, telle que présentée par Mankiw et al. (92), prend la forme suivante :

$$Y_t = K_t^\alpha H_t^\phi (A_t L_t)^{1-\alpha-\phi} \quad 0 < \alpha + \phi < 1 \quad (18)$$

$H_t$  : stock du capital au temps  $t$

Le reste des variables est défini comme précédemment. Soit  $s_k$  la fraction du revenu investie dans le capital physique et  $s_h$  la fraction investie dans le capital humain. L'évolution de l'économie est déterminée par :

$$\dot{k}_t = s_k y_t - (n + g + \delta) k_t \quad (19)$$

$$\dot{h}_t = s_h y_t - (n + g + \delta) h_t \quad (20)$$

$h_t = \frac{H_t}{A_t L_t}$  représente le stock du capital humain par unité de travail effectif. On

suppose que les deux types de capitaux se déprécient au même taux  $\delta$ .

A l'état régulier, le niveau du capital physique et humain par unité de travail effectif est déduit par les équations suivantes :

$$k^* = \left[ \frac{s_k^{1-\phi} s_h^\phi}{n + g + \delta} \right]^{1/(1-\alpha-\phi)} \quad (21)$$

$$h^* = \left[ \frac{s_k^\alpha s_h^{1-\alpha}}{n + g + \delta} \right]^{1/(1-\alpha-\phi)} \quad (22)$$

La production par tête à l'état régulier est donnée par le remplacement des expressions (21) et (22) dans la forme intensive :  $y_t = k_t^\alpha h_t^\phi$ . Ce qui donne :

$$\ln\left(\frac{Y}{L}\right) = \ln(A_0) + gt - \left(\frac{\alpha + \phi}{1 - \alpha - \phi}\right) \ln(n + g + \delta) + \left(\frac{\alpha}{1 - \alpha - \phi}\right) \ln(s_k) + \left(\frac{\phi}{1 - \alpha - \phi}\right) \ln(s_h) \quad (23)$$

Cette équation montre comment la production par tête dépend de la croissance de la population et de l'accumulation du capital humain et physique. L'élasticité de la production par rapport à l'investissement devient égale à  $\frac{\alpha}{1 - \alpha - \phi}$  dans le modèle augmenté, au lieu

de  $\frac{\alpha}{1 - \alpha}$ . Un taux d'épargne important induit une importante production, et un niveau

important de capital humain à l'état régulier, même si la part allouée au capital humain est sans changement. En d'autres termes, la présence de l'accumulation du capital humain augmente l'effet de l'investissement dans le capital physique sur le niveau de la production à l'état régulier. Cependant, les estimations basées sur le modèle de Solow augmenté avec capital humain trouvent un très faible taux de convergence vers l'état régulier. Ce qui est consistant avec les évidences sur les flux de capitaux internationaux (Barro, Mankiw, et Sala-i-Martin, 1995 ; Manzocchi et Martin, 1996).

La dérivée de l'expression (23) par rapport au temps donne un taux de croissance de l'économie égal au taux de croissance du progrès technique.

Une méthode alternative permet d'exprimer la production dans ce modèle a été donnée par Mankiw et al. (92). Si on combine (23) avec l'équation de l'état régulier du capital humain donnée par l'équation (22), on obtient une équation de la production par tête fonction du taux d'investissement dans le capital physique, le taux de croissance de la population, et le niveau du capital humain :

$$\ln\left(\frac{Y_t}{L_t}\right) = \ln(A_0) + gt + \frac{\alpha}{1 - \alpha} \ln(s_k) - \frac{\alpha}{1 - \alpha} \ln(n + g + \delta) + \frac{\phi}{1 - \alpha} \ln(h^*) \quad (24)$$

Cette équation est similaire à l'équation (10) du modèle simple de Solow. Dans ce modèle le niveau du capital humain est un élément composant le terme de l'erreur dans le modèle non augmenté de Solow. Puisque le taux d'épargnes et celui de la croissance de la population influencent  $h^*$  (le niveau du capital humain), on doit s'attendre à ce que le capital humain soit positivement corrélé avec le taux d'épargne et négativement corrélé avec la croissance de la population. Cependant, omettre le terme du capital humain du modèle de croissance biaise les coefficients de l'épargne et de la croissance de la population.

Une approximation de Taylor d'ordre 1 autour des points de l'état régulier donne la vitesse de convergence comme suit : (Mankiw et al.1992)

$$\frac{d \ln(y_t)}{dt} = \lambda [\ln(y^*) - \ln(y_t)] \quad (25)$$

$$\lambda = (n + g + \delta)(1 - \alpha - \phi)$$

Le modèle suggère une régression naturelle pour étudier le taux de convergence. La solution de l'équation différentielle (25) implique que :

$$\ln(y_t) = (1 - e^{-\lambda t}) \ln(y^*) + e^{-\lambda t} \ln(y_0) \quad (26)$$

où  $y_0$  est la production par unité de travail effectif à une date initiale.

A l'état régulier, le taux de croissance de l'économie est alors gouverné par l'équation suivante :

$$\ln(y_t) - \ln(y_0) = (1 - e^{-\lambda t}) \frac{\alpha}{1 - \alpha - \phi} \ln(s_k) + (1 - e^{-\lambda t}) \frac{\phi}{1 - \alpha - \phi} \ln(s_h) \quad (27)$$

$$- (1 - e^{-\lambda t}) \frac{\alpha + \phi}{1 - \alpha - \phi} \ln(n + g + \delta) - (1 - e^{-\lambda t}) \ln(y_0)$$

En terme de production par tête on obtient :

$$\ln\left(\frac{Y_t}{L_t}\right) - \ln\left(\frac{Y_0}{L_0}\right) = (1 - e^{-\lambda t}) \ln(A_0) + g t + (1 - e^{-\lambda t}) \frac{\alpha}{1 - \alpha - \phi} \ln(s_k) \quad (28)$$

$$+ (1 - e^{-\lambda t}) \frac{\phi}{1 - \alpha - \phi} \ln(s_h)$$

$$- (1 - e^{-\lambda t}) \frac{\alpha + \phi}{1 - \alpha - \phi} \ln(n + g + \delta) - (1 - e^{-\lambda t}) \ln(y_0)$$

Dans le modèle de Solow la croissance de la production est une fonction des paramètres déterminants de l'état régulier et du niveau initial de la production. Le niveau des taux d'investissements en capital humain et en capital physique est susceptible de renforcer ou de contrecarrer le mécanisme de rattrapage.

L'équation (28) est la base des travaux empiriques sur la croissance. Si l'on pose que  $t^{-1} \cdot \left( \ln\left(\frac{Y_t}{L_t}\right) - \ln\left(\frac{Y_0}{L_0}\right) \right) = \gamma$ ,  $\beta = -t^{-1} (1 - e^{-\lambda t})$  et que l'on ajoute un terme de l'erreur à

l'équation (28), celle-ci devient :

$$\begin{aligned} \gamma = & g + \beta \ln(y_0) + \beta \frac{\alpha + \phi}{1 - \alpha - \phi} \ln(n + g + \delta) \quad (29) \\ & - \beta \frac{\alpha}{1 - \alpha - \phi} \ln(s_k) - \beta \frac{\phi}{1 - \alpha - \phi} \ln(s_h) - \beta \ln(A_0) + v_i \end{aligned}$$

Mankiw et al.(1992) considère le terme  $A_0$  comme les influences spécifiques aux pays sur la croissance, comme : les ressources propres, climat et les institutions. Ils supposent que ces différences varient d'une manière aléatoires dans le sens où :

$$\log(A_0) = \log(A) + e \quad (30)$$

où  $e$  représente le choc spécifique au pays. L'intégration de (30) dans (29), on posant  $\varepsilon = v - \beta e$  donne :

$$\begin{aligned} \gamma = & g - \beta \ln(A) + \beta \ln(y_0) + \beta \frac{\alpha + \phi}{1 - \alpha - \phi} \ln(n + g + \delta) \quad (31) \\ & - \beta \frac{\alpha}{1 - \alpha - \phi} \ln(s_k) - \beta \frac{\phi}{1 - \alpha - \phi} \ln(s_h) + \varepsilon \end{aligned}$$

Plusieurs études empiriques élargissent la forme de Mankiw et al.(1992) par l'introduction de variables de contrôle  $Z$  dans la régression (31). Ces variables permettent de capter l'hétérogénéité dans les termes,  $g$  (taux de croissance de l'état régulier) et  $A_0$  (la technologie initiale), supposés constants entre les pays. Formellement,  $g - \beta \log(A_0)$  dans (29) est remplacé par  $g - \beta \log(A) + \pi Z - \beta e$  ce qui donne la régression suivante :

$$\begin{aligned} \gamma = & g - \beta \ln(A) + \beta \ln(y_0) + \beta \frac{\alpha + \phi}{1 - \alpha - \phi} \ln(n + g + \delta) \quad (32) \\ & - \beta \frac{\alpha}{1 - \alpha - \phi} \ln(s_k) - \beta \frac{\phi}{1 - \alpha - \phi} \ln(s_h) + \pi Z + \varepsilon \end{aligned}$$

Temple (1999) justifie l'introduction de  $Z$ , comme une proxy pour rendre compte des différences dans les TPF de long terme, par l'invraisemblance de l'hypothèse de l'égalité des TFP dans un échantillon de pays relatant un intervalle de temps de plus de 20 ou 30 années. En effet, il semble plausible de penser que les variables de contrôle  $Z$  rendent compte des différences dans les taux de croissance de long terme  $g$ , plutôt que du niveau technologique initial  $A_0$ .

La représentation générique de la régression (32) prend la forme suivante : (Durlauf, Johnson et Temple, 2004)

$$\gamma = \beta \log(y_0) + \psi X + \pi Z + \varepsilon \quad (33)$$

La variable  $X$  contient la constante,  $\log(n + g + \delta)$ ,  $\ln(s_k)$  et  $\ln(s_h)$ . Les variables couvertes par  $\log(y_0)$  et  $X$  représentent celles des déterminants de la croissance proposée dans le cadre du modèle de croissance de Solow, les variables  $Z$  représentent les déterminants non prévu par la théorie originale de Solow. L'équation (33) représente la base de l'économétrie de la croissance. Cette régression est souvent connue sous le nom « régression de Barro ». Dans les applications empiriques modernes, cette équation a été généralisée dans plusieurs directions, elle a été appliquée aux séries temporelles, données de panel et plus récemment à l'économétrie spatiale.

## **1-2 La convergence dans les modèles de croissance endogène :**

La théorie néoclassique a été pour longtemps le cadre favorisé pour l'étude de la croissance. A partir de la moitié des années 80, un groupe de théoriciens de la croissance chapeauté par Paul Romer (1986) remettent en cause les explications des effets exogènes sur la croissance de long terme. Cette insatisfaction, théorique et empirique, était à la base de la construction d'une nouvelle classe de modèles dans lesquels les clés déterminantes de la croissance sont endogènes aux modèles. Les premiers modèles de la croissance endogène (Romer, 1986 ; et Lucas 1988) ont pour objet d'expliquer la croissance de long terme et les disparités de croissance observées au niveau international. Ces modèles sont caractérisés par l'hypothèse de rendements non décroissants des facteurs reproductibles de la production. Mankiw, Romer et Weil (1992) constatent que, sous cette hypothèse, les pays qui épargnent plus croissent indéfiniment plus vite, et que ces pays ne convergent pas nécessairement, même s'ils ont les mêmes préférences et technologies. La faiblesse empirique du modèle de Solow pour expliquer les différences entre les pays a été bien argumentée dans le paragraphe suivant de Barro (1989) : « Dans les modèles de croissance néoclassiques avec des rendements décroissants, tels que Solow (1956), Cass (1965) et Koopmans (1965), la croissance par tête d'un pays tend à être inversement liée au niveau initial du revenu par tête. Par conséquent, dans l'absence de chocs, les pays riches et pauvres peuvent converger en termes de niveaux de revenu par tête. Cependant, cette hypothèse de convergence semble inconsistante avec les évidences empiriques, qui indiquent que les taux de croissances par tête ne sont pas corrélés avec le niveau initial de la production ».

L'objectif principal des modèles de croissance endogène est de rompre avec les hypothèses d'un progrès technique exogène et des rendements décroissants dans le but d'expliquer l'origine de la croissance à long terme et la différence des taux de croissance entre pays. La première vague de la théorie de la croissance endogène était le modèle AK. La nouveauté dans ce modèle consiste à considérer le progrès technologique comme une autre forme de l'accumulation du capital. Ce modèle a été suivi par une seconde vague, avec Aghion et Howitt (1998) qui ont élaboré la théorie Schumpétérienne ou la théorie basée sur l'innovation. Cette dernière fait la distinction entre la connaissance technologique et le capital, et analyse le processus technologique d'innovation comme une activité séparée de l'épargne.

### 1-2-1 Le modèle AK :

La propriété clé dans les modèles de croissance endogène est l'absence des rendements décroissants du capital. Selon Howitt (2008), une ancienne version du modèle *AK* a été proposée par Frankel (1962)<sup>2</sup>, qui justifie que la fonction de production agrégée peut présenter une productivité marginale du capital constante ou même croissante, ceci est dû au fait qu'une partie de la croissance du stock du capital des firmes peut être le capital intellectuel qui crée le progrès technologique. Ce dernier peut arrêter la tendance à la baisse de la productivité marginale du capital. La version simple de la fonction de production (Robelo, 1991) est donnée par :

$$Y_t = AK_t \quad (34)$$

$A$  est une constante positive reflétant le niveau technologique. L'idée de l'absence globale des rendements décroissants devient plus plausible si on considère  $K$  dans le sens général où il comprend le capital humain. La production par tête  $y = Ak$ , et la productivité moyenne et marginale est constante au niveau  $A > 0$ . Il s'agit d'un modèle sans progrès technique  $g = 0$  car le but est de montrer que la croissance par tête de long terme peut se produire même en l'absence d'un progrès technique exogène.

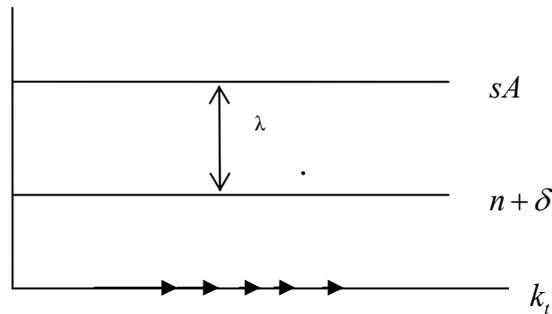
$$\gamma_k = \frac{\dot{k}_t}{k_t} = \frac{\dot{y}_t}{y_t} = sA - (n + \delta) \quad (35)$$

---

<sup>2</sup> Barro et Xavier Sala-i-Martin considèrent que c'est Von Neumann (1937) le premier à utiliser une fonction de production de type AK

Pour la représentation graphique, la différence avec la figure 1 du modèle de Solow est que la courbe  $s \frac{f(k_t)}{k_t}$  est remplacée par une ligne horizontale au niveau  $sA$

**Figure 3 : Le modèle AK**



$\gamma_k$  est la distance verticale entre les deux lignes  $sA$  et  $n + \delta$  et tel que  $sA > (n + \delta)$  de sorte que

$\gamma_k = \frac{\dot{k}_t}{k_t} > 0 \forall k$ . Puisque les deux lignes sont parallèles,  $\gamma_k$  est constant et indépendant de  $k$ . En d'autres

termes,  $k$  croît toujours au taux de l'état régulier  $\gamma_k^* = sA - (n + \delta)$ .

Puisque  $y = Ak$ ,  $\gamma_k$  est égal à  $\gamma_k^*$  quelque soit  $t$ . La croissance par tête décrite par l'équation (30) dépend du comportement des paramètres structurels du modèle, tels que le taux d'épargne et le taux de la croissance de la population. Contrairement au modèle néo-classique l'augmentation de  $s$  entraîne une augmentation du taux de croissance par tête de long terme  $\gamma_k^*$ . D'une manière similaire, le changement dans le niveau technologique  $A$ , ou le taux de dépréciation  $\delta$  et le taux de la croissance de la population  $n$ , entraînent des changements sur le taux de croissance par tête de long terme.

Ce modèle de croissance peut être considéré comme un cas limite du modèle néo-classique de Mankiw (1995). En effet, la formulation AK ne prévoit pas la convergence conditionnelle ou inconditionnelle, qui est  $\frac{d \ln(y_t)}{dt} = 0 \forall y$ . Cette interprétation est rendue possible dès lors que l'on considère la variable  $K$  comme du capital élargi. Le travail est assimilé au capital humain et agrégé dans le capital physique.

Si on considère un groupe d'économie structurellement similaire par rapport aux paramètres  $s$ ,  $A$ ,  $n$  et  $\delta$ . Ces économies diffèrent seulement dans leurs conditions initiales, telles que le stock initial du capital par tête et le revenu par tête initial. A cause de l'hypothèse d'absence de rendement décroissant, ces économies vont croître aux même taux par tête. En

effet, si on observe le modèle  $AK$  sous l'optique du modèle de Cobb-Douglas avec une part unitaire du capital :  $\alpha = 1$ . L'analyse de la convergence telle que vue précédemment montre que la vitesse de convergence donnée par l'équation (12)  $\gamma = (1 - \alpha)(g + n + \delta)$  implique que  $\gamma = 0$  pour  $\alpha = 1$ .

Romer (1986) produit une analyse similaire avec une fonction de production plus générale, sous l'hypothèse que l'épargne est générée par la maximisation d'une fonction d'utilité intertemporelle au lieu d'un taux fixe d'épargne. Lucas (1988) produit une analyse basée sur le capital humain au lieu du capital physique ; suivant Uzawa (1965), il suppose que le capital humain et la connaissance technologique sont la même chose.

### 1-2-2 La théorie basée sur l'innovation :

Cette théorie connue sous le nom anglais de « innovation-based growth theory » reconnaît que le capital intellectuel, source du progrès technique, est différent du capital humain et physique. Ces derniers sont accumulés à travers l'épargne et la scolarité, mais la croissance du capital intellectuel se fait par l'innovation.

Une version de la théorie basée sur l'innovation a été initiée par Romer (1990), elle suppose que la productivité agrégée est une fonction croissante du degré de variété d'un produit. Dans cette théorie, l'innovation induit la croissance de la productivité par la création de nouvelles variétés de produits. Il utilise une fonction de production à la Dixit-Stiglitz et dans la quelle la production est la résultante du travail et d'un continuum de produits intermédiaires.

$$Y = L^{1-\alpha} \int_0^A x^\alpha(i) \cdot di, \quad 0 < \alpha < 1 \quad (36)$$

$L$  représente l'offre agrégée du travail (supposée constante),  $x(i)$  est le flux d'input du produit intermédiaire  $i$ , et  $A$  mesure les différents produits intermédiaires utilisés. Une augmentation dans la variété des produits, mesurée par  $A$ , augmente la productivité. Cet effet, qui capture une forme du progrès technique, reflète le gain obtenu par l'élargissement de la gamme des produits intermédiaires  $A \cdot x(i)$  sur un intervalle  $A$ . Les rendements décroissants se produisent quand l'expansion des bien intermédiaires  $A \cdot x(i)$  se produit à travers la croissance des quantités des bien  $x(i)$  sachant leur nombre  $A$ . Cependant, la décroissance des rendement ne ce produit pas si la croissance dans  $A \cdot x(i)$  prend la forme d'une croissance

dans le nombre de variétés  $A$  sachant  $x(i)$ . En effet, un changement technologique sous la forme d'une croissance continue dans  $A$  évite la tendance des rendements décroissants. Cette propriété de la fonction de production fournit la base de la croissance endogène.

Une autre version de la théorie de la croissance basée sur l'innovation est la théorie « Schumpétérienne » développée par Aghion et Howitt (1992) et Gossman et Helpman (1991). La théorie Schumpétérienne s'intéresse à la qualité des innovations qui rend les anciens produits obsolètes, à travers le processus que Schumpeter appelle la « destruction créative ».

Dans la théorie Schumpétérienne la production agrégée est produite par un continuum de produits intermédiaire, suivant la forme suivante :

$$Y = L^{1-\alpha} \int_0^1 A^{1-\alpha}(i) x^\alpha(i) \cdot di, \quad (37)$$

Avec une mesure fixe de la variété du produit, normalisée à l'unité. Chaque produit intermédiaire possède son propre paramètre de productivité  $A(i)$ . Chaque secteur est monopolisé et produit son propre bien intermédiaire avec un coût marginal constant et égal à l'unité. La courbe de la demande adressée, par le secteur final, au secteur  $i$  est donnée par le produit marginal  $\alpha \left( \frac{A(i)L}{x(i)} \right)^{1-\alpha}$ . L'égalisation entre le revenu marginal (produit de  $\alpha$  avec le produit marginal) et le coût marginal d'une unité, donne le profit maximum du revenu intermédiaire du monopole :

$$\begin{aligned} x(i) &= \xi LA(i) & (38) \\ \xi &= \alpha^{2/1-\alpha} \end{aligned}$$

La substitution de (33) dans la fonction de production (32), donne la fonction de production agrégée :

$$Y = \theta AL \quad (39)$$

où  $\theta = \xi^\alpha$  et  $A$  le paramètre de la productivité moyenne tel que  $A \equiv \int_0^1 A(i) \cdot di$ .

L'innovation dans la théorie Schumpétérienne crée une version améliorée des anciens produits. Une innovation dans un secteur  $i$  consiste en une version où le paramètre de

productivité  $A(i)$  excède celui de la version précédente par un facteur fixe  $\lambda > 1$ . Supposons que la probabilité d'une innovation dans un secteur  $i$ , sur un petit intervalle de temps d'étendue  $dt$ , est égale à  $\mu$ . Le taux de croissance de  $A(i)$  est donné par :

$$\frac{dA(i)}{A(i)} \cdot \frac{1}{dt} = \left\{ \begin{array}{l} (\lambda - 1) \cdot dt \text{ avec une probabilité } \mu \cdot \frac{1}{dt} \\ 0 \quad \quad \quad \text{avec une probabilité } 1 - \mu \cdot dt \end{array} \right\} \quad (40)$$

Le taux espéré de la croissance de  $A(i)$  est donné par :

$$E(g) = \mu(\lambda - 1) \quad (36)$$

La probabilité  $\mu$  d'une innovation dans un secteur donné est proportionnelle au flux actuel de dépenses en recherche et développement (RD) ajustés par la productivité.

$$\mu = \frac{\varphi R}{A} \quad (41)$$

$R$  est la part de la production finale dépensée en RD, et la division sur  $A$  permet de prendre en compte la force de l'augmentation de la complexité. En effet, l'avancée technologique devient plus complexe, et la société engage une augmentation des dépenses en RD juste pour maintenir le taux précédent de la croissance de l'innovation.

Il découle de (39) que le taux de croissance  $g$  de la production agrégée est le même que le taux de croissance du paramètre de la productivité moyenne  $A$ . La loi des grands nombres garantit que  $g$  est égal au taux de croissance espéré (41) de chaque paramètre individuel de productivité. De là, on peut écrire que :

$$g = E(g) = (\lambda - 1)\varphi \frac{R}{A} \quad (42)$$

De cette expression et celle de (34), il s'en suit que le taux de croissance dépend de la fraction de la production dépensée en recherche et développement :

$$g = (\lambda - 1)\varphi \theta L n \quad (43)$$

$$n = \frac{R}{Y}$$

La théorie basée sur l'innovation implique que le chemin d'une croissance rapide n'est pas d'épargner une part importante de la production, mais de consacrer une grande part de la production à la recherche développement. La théorie est explicite sur comment l'activité de

RD est influencée par les différentes politiques, quels sont les gagnants et les perdants du progrès technique, comment les gains et les pertes dépendent des arrangements sociaux, et comment ces arrangements affectent la société et son habilité à s'adapter avec le changement technologique, source ultime de la croissance.

### **1-2-3 Les défis empiriques de la convergence :**

Le premier problème posé avec les modèles de croissance endogène de première génération vient de l'observation que plusieurs pays semblent converger vers le même taux de croissance du PIB par tête de long terme (Barro et Sala-i-Martin, 92 ; Evans, 96). La prédiction stipule que le taux de croissance de long terme de chaque pays dépend de l'effet de la politique et des institutions de diffusion de l'innovation. Pour Mankiw et al. (92), la convergence observée dans les taux de croissance correspond à la prédiction de Solow-Swan, sous l'hypothèse que les pays sont capable de s'inscrire dans la même frontière technologique et présentent le même taux de progression technique.

Pour prendre en compte ces limites, la théorie endogène intègre l'importance des transfères technologiques entre pays, d'une manière similaire aux transferts technologiques entre industries dans une économie fermée (Howitt, 2000). Cette modification est justifiée par le fait que la croissance de la productivité totale des facteurs (TFP) dans un pays dépend des dépenses en RD du pays, mais aussi des celle des pays partenaires commerciaux (Coe et Helpmann, 95 ; Eaton et Kortum, 96). L'incorporation des transferts technologiques dans la théorie, met en évidence le rôle des performances en RD dans la convergence vers un même taux de croissance. Les pays qui présentent de faibles taux de croissance que les leaders mondiaux de la technologie, peuvent améliorer leur vitesse de croissance, simplement par l'implémentation des nouvelles technologies découvertes dans les autres pays, c'est ce que (Gerschenkron, 1952) appel « l'avantage du retard » (advantage of backwardness).

Howitt (2000) montre que la théorie modifiée est capable d'expliquer les caractéristiques de la distribution du PIB par tête au même titre que le modèle de Solow-Swan, et aussi d'autres caractéristiques additionnelles. A titre d'exemple, la théorie prédit que les pays différents dans les niveaux du PIB par tête non seulement à cause de la différence des taux d'épargne et de la croissance de la population, mais aussi à cause d'une différence dans les intensités en RD (Lichtenberg, 1993) . Elle prédit aussi que les pays peuvent être différents dans les niveaux des TFP et du capital par tête, et que ces déterminants du PIB par tête peuvent être positivement corrélés entre les pays (Klenow et Rodríguez-Clare, 1997 ; Hall et

Jones, 1999). Cependant, dans les pays à faible motivation pour améliorer la RD, les firmes n'investissent pas dans le changement technologique et ne sont pas capables de bénéficier des transferts technologiques. Ces pays ne convergent pas et présentent des taux de croissances faibles, même dans le long terme. Ce résultat s'accorde avec celui des clubs de convergence, développé par plusieurs chercheurs empiristes (Quah, 1996).

### **1-3 l'économétrie de la convergence :**

Beaucoup de travaux empiriques se sont intéressés à l'hypothèse de la convergence. Cet intérêt est devenu plus répandu après les travaux d'Abramovitz (1986) et de Baumol (1986).

Dans cette section nous traitons l'hypothèse de la convergence selon une spécification qui relie les conditions initiales à la production de long terme et nous explorons les techniques employées dans l'étude de la dépendance de long terme.

#### **1-3-1 la convergence et les conditions initiales :**

L'effet des conditions initiales sur la production de long terme représente la première question empirique explorée par les économistes de la croissance. L'effet de conditions initiales est la base de l'hypothèse de convergence. Le but des différents travaux réalisés dans cette optique, est de répondre aux questions concernant les différences dans les revenus par tête (entre pays ou régions). Les différences de revenus par tête sont-elles permanentes ou temporaires ? S'ils sont permanentes, cette permanence reflète-t-elle une certaine hétérogénéité structurelle dans la croissance de long terme ? Et est-ce que cette hétérogénéité permet de vérifier la convergence conditionnelle ? Dans le cas où les différences sont temporaires, est-ce que la convergence absolue (inconditionnelle) est vérifiable ? Si les conditions initiales déterminent la production de long terme, et les pays avec les mêmes conditions initiales présentent les mêmes taux de croissance de long terme, alors on peut parler de club de convergence (Galor, 1996).

Dans la formalisation de l'idée des conditions initiales nous posons  $\log(y_{it})$  le niveau de la production par tête dans le pays (ou région)  $i$  au temps  $t$ . On associe avec  $\log(y_{it})$  les conditions initiales  $\psi_{i0}$ . Ces dernières n'ont pas d'effet sur le long terme si :

$$\lim_{t \rightarrow \infty} P(\log(y_{it}) | \psi_{i0}) \text{ ne dépend pas de } \psi_{i0} \quad (44)$$

Se basant sur ce cadre d'analyse, les travaux empiriques sur la convergence s'intéressent à la dépendance entre la production par tête de long terme et le stock initial du capital humain et physique. La question de la convergence a pour origine la question sur les effets des conditions initiales sur la persistance des différences dans la production par tête entre les pays (régions). Deux pays  $i$  et  $j$  montrent une convergence si :

$$\lim_{t \rightarrow \infty} \left\| P(\log(y_{it}) | \psi_{i0}) - P(\log(y_{jt}) | \psi_{j0}) \right\| = 0 \quad (45)$$

$\|\cdot\|$  est une métrique pour calculer la distance entre les deux mesures de probabilités. Les économistes de la croissance s'intéressent généralement au niveau moyen de la production. L'équation (45) implique que les deux pays  $i$  et  $j$  montrent une convergence dans les niveaux de production moyenne dans le sens où :

$$\lim_{t \rightarrow \infty} E(\log(y_{it}) - \log(y_{jt}) | \psi_{i0}, \psi_{j0}) = 0 \quad (46)$$

Dans la mesure où l'on s'intéresse à un taux commun de croissance de l'état régulier. L'équation (46) peut être modifiée pour que la différence espérée entre  $\log(y_{it})$  et  $\log(y_{jt})$  soit bornée (Pesaran 2004). Bernard et Durlauf (96) suggèrent une forme de la convergence partielle qui montre comment les différences de revenus contemporaines tendent à diminuer. Si  $\log(y_{i0}) > \log(y_{j0})$ , leur définition revient à chercher si :

$$E(\log(y_{it}) - \log(y_{jt}) | \psi_{i0}, \psi_{j0}) < \log(y_{i0}) - \log(y_{j0}) \quad (47)$$

Plusieurs modifications à cette définition ont été proposées, on cite celle de Hall, Robertson et Wickens (1997), basée sur une tendance vers zéro des différences entre les variances des niveaux de production. Plus récemment, Pesaran (2004) a proposé une définition de la convergence basée sur la vraisemblance d'une déviation importante de long terme. Cette idée est basée sur une différence importante entre les niveaux de production, avec l'affectation d'une probabilité qui donne une certaine flexibilité à la réalisation de cet événement.

La faiblesse de ces définitions réside dans le fait qu'ils ne permettent pas de faire la distinction entre l'effet de long terme des conditions initiales et de l'hétérogénéité structurelle de long terme. Cependant, dans les travaux empiriques, il est important de faire la distinction entre les conditions initiales  $\psi_{i0}$  et les caractéristiques structurelles  $\omega_{i0}$ . A la différence des effets des caractéristiques structurelles, les effets des conditions initiales sur l'état régulier impliquent l'existence de clubs de convergence. Pour prendre en compte cette différence

d'effet, Durlauf, Johnson et Temple (2004) proposent une modification de la forme (45), comme suit :

$$\lim_{t \rightarrow \infty} \left\| P(\log(y_{it}) | \psi_{i0}, \omega_{i0}) - P(\log(y_{jt}) | \psi_{j0}, \omega_{j0}) \right\| = 0 \text{ si } \omega_{i0} = \omega_{j0} \quad (48)$$

Cette forme implique que les pays  $i$  et  $j$  convergent. La notion de convergence en terme d'espérance prend la forme :

$$\lim_{t \rightarrow \infty} E(\log(y_{it}) - \log(y_{jt}) | \psi_{i0}, \omega_{i0}, \psi_{j0}, \omega_{j0}) = 0 \text{ si } \omega_{i0} = \omega_{j0} \quad (49)$$

Dans la pratique la distinction entre les conditions initiales et l'hétérogénéité structurelle équivaut à traiter le stock du capital humain et le capital physique en premier lieu et les autres variables en dernier. Ainsi, les variables de Solow  $X$  et les variables de contrôle  $Z$  dans l'équation de régression (33) sont souvent interprétées comme capturant l'hétérogénéité structurelle. Cette pratique peut être critiquée si ces dernières variables sont elles mêmes déterminées d'une manière endogène par les conditions initiales.

### 1-3-2 La $\beta$ – convergence :

Les analyses statistiques de la convergence s'intéressent aux propriétés de  $\beta$  dans la régression (33). La convergence est vérifiée si  $\beta$  est estimée inférieure à zéro. Si deux pays ont les mêmes déterminants de l'état régulier et convergent vers le même sentier de croissance, le pays qui a un faible niveau initial de revenu par tête a un faible ratio de capital par tête et donc une forte productivité marginale du capital. Un taux d'investissement donné entraîne une croissance plus rapide du pays pauvre. Cette évidence est souvent utilisée comme une remise en cause des modèles de croissance endogène de première génération (type Romer et Lucas). Ces derniers prévoient, après introduction des taux d'épargne et de croissance de la population, que les pays dont les revenus initiaux sont importants, présentent les taux de croissance les plus importants.

Actuellement, il existe plusieurs études sur la  $\beta$  – convergence. Ces études diffèrent selon les groupes de pays étudiés, les périodes et le choix des variables de contrôle. L'hypothèse de la convergence absolue a souvent été vérifiée pour des groupes d'économies relativement homogènes tels que les Etats d'Amérique du Nord (US), les pays de l'OCDE ou pour certaines régions de l'Europe. Quand l'échantillon de pays est très hétérogène, il n'existe généralement pas de corrélation entre le revenu initial et la croissance.

En dépit d'une large confirmation de la convergence dans la littérature, la revendication d'une convergence conditionnelle globale reste très controversée (Durlauf, Johnson et Temple, 2004). La complexité dans la convergence conditionnelle réside dans le choix des variantes  $Z$  pour spécifier le revenu de l'état régulier. Cette complexité se justifie par le fait que la théorie n'est pas toujours un bon guide dans le choix des éléments de  $Z$ , et par la dépendance du revenu de l'état régulier de ce même choix. Par conséquent, l'évidence de l'hypothèse de la convergence conditionnelle semble plus robuste, mais dépend fortement du choix des variables de contrôle.

### 1-3-3 La $\beta$ – convergence et la divergence économique

La  $\beta$  – convergence n'implique pas automatiquement la convergence telle que définie par l'équation (44). L'estimation d'un  $\beta < 0$  n'est pas inconsistante avec des alternatives économiques non convergentes. L'exemple du modèle de croissance et des seuils d'externalités développé par Azariadis et Drazen (1990), montre une discontinuité dans la fonction de production agrégée sur un échantillon d'économies agrégées. Cette discontinuité signifie que l'état régulier d'une économie donnée, dépend de la position du niveau du stock du capital initial par rapport à ce seuil. Ce modèle met en évidence l'existence de deux états réguliers distincts, et que la convergence statistique peut coexister avec la présence de plusieurs état réguliers.

Ce dernier résultat a été illustré dans le travail effectué, sur une régression simplifiée de la croissance, par Bernard et Durlauf (1996). A cet effet, si on suppose que pour chaque pays les variables de Solow  $X_i$  et de contrôle  $Z_i$  sont identiques, et qu'il n'y a pas de changement technique, ni de croissance de la population. La régression induite par l'hypothèse de Azariadis-Drazen sur la fonction de production agrégée est de la forme suivante :

$$\gamma_i = k + \beta \left( \log y_{i,0} - \log y_{l(i)}^* \right) + \varepsilon_i \quad (50)$$

Où  $i$  est l'indice du pays,  $l(i)$  indice de l'état régulier du pays  $i$  et  $y_{l(i)}^*$  la production par tête à l'état régulier  $l(i)$ . Tous les pays associés à un même état régulier ont la même valeur  $\log y_{l(i)}^*$ .

Le modèle basé sur le seuil d'externalité, tel que défini dans (50), ne produit pas de convergence économique, tant qu'il existe au moins deux états réguliers. Cependant, les

données générées par un échantillon de pays présentant plusieurs états réguliers peuvent montrer une convergence statistique. En effet, à cause de ce fait stylisé, l'équation de régression de la croissance quand tous les pays ont le même état régulier peut se réécrire comme :

$$\gamma_i = k + \beta(\log y_{i,0}) + \varepsilon_i \quad (51)$$

Si les données sont générées par une forme semblable à (50), la régression standard (51) est mal spécifiée. Une estimation de (51) sur des données générées par un processus qui suit (50) (c'est-à-dire quand l'équation (50) est la bonne spécification) donne une estimation du paramètre de convergence de la forme :

$$\beta_{ols} = \beta \left( 1 - \frac{\text{cov}(\log y_{i(i)}^*, \log(y_{i0}))}{\text{var}(\log y_{i0})} \right) \quad (52)$$

La forme (52) montre que le signe de  $\beta$  dépend de  $1 - \frac{\text{cov}(\log y_{i(i)}^*, \log(y_{i0}))}{\text{var}(\log y_{i0})}$  qui est une fonction de la covariance entre le revenu initial et celui de l'état régulier d'un pays dans l'échantillon. Suivant cette formulation, il est possible d'avoir un  $\beta_{ols}$  négatif même si l'échantillon comprend des pays associés avec des états réguliers différents. En effet, on peut s'attendre à avoir un  $\beta_{ols} < 0$  si les pays à faible revenu tendent à être initialement au dessous de leur état régulier alors que les pays à fort revenu commencent avec des revenus au dessus de leur état régulier.

L'exemple traité ci-dessus montre combien la convergence statistique peut être consistante avec une divergence économique. A cet effet, on peut noter deux observations importantes : en premier lieu, certains modèles statistiques conformes avec l'existence de régimes multiples, estiment mieux les données sur pays que le modèle linéaire de Solow (Durlauf et Johnson, 1995). En deuxième lieu, d'autres études ont mis en évidence la non constance du paramètre  $\beta$  et qui semble dépendre d'une manière non linéaire des conditions initiales (Liu et Stengos, 1999). Ces résultats confirment la compatibilité de l'existence de sentiers de croissance avec des différences permanentes de revenus entre économies dotées de mêmes taux de croissance de la population, de l'épargne et d'un accès identique à la technologie.

Un autre problème d'identification soulevé par Cohen (1996) et Goetz et Hu (1996) concernant l'effet de l'endogénéité sur la relation entre la  $\beta$ -convergence et la convergence économique. Dans le modèle de Solow,  $\beta$  n'illustre pas d'une manière incontestable comment les conditions initiales peuvent affecter les différences futures de revenus, si les taux de croissance de la population et de l'épargne sont eux-mêmes fonction du revenu. Un  $\beta \geq 0$  peut être compatible avec une convergence économique partielle si les taux d'accumulation du capital humain et physique dépendent du niveau du revenu, et un  $\beta < 0$  peut être compatible avec une divergence économique si les taux d'accumulation du capital humain et physique pour les pays riches et les pays pauvres divergent dans le temps. Cohen (1996) montre que l'équation conventionnelle d'accumulation du capital humain, où l'accumulation est proportionnelle à la production par tête, ne rend pas compte du processus d'accumulation. En effet, les pays pauvres avec de faibles stocks initiaux de capital humain n'accumulent pas aussi rapidement que les pays riches. Dans la même logique d'idées, Goetz et Hu (1996) testent la réaction de l'accumulation du capital humain au niveau du revenu. L'effet indirect de la croissance du revenu (à travers la croissance du capital humain) peut compenser l'effet direct, et dans ce cas on peut observer une divergence dans le temps.

#### **1-4 La croissance et la géographie économique :**

L'espace a toujours occupé une place périphérique dans la pensée économique. Samuelson (1952) a fait remarquer que « *les questions spatiales ont été tellement négligées en théorie économique que cela seul suffit pour en justifier l'étude* ». Les théories de l'économie géographique montrent que la dimension spatiale n'est pas neutre, que les interactions entre économies sont dépendantes des formes d'accessibilités aux marchés et des avantages de proximités des lieux de concentration des activités économiques. Les choix de localisation, découlant de ces principes, expliquent l'inégalité observée dans la répartition des activités économiques, de la croissance et des richesses entre les pays et les régions. La nouvelle économie géographique a été initiée par trois auteurs à savoir Fujita (1988), Krugman (1991) et Venables (1996). Les trois auteurs utilisent des modèles d'équilibre général avec concurrence monopolistique. La majorité de ces modèles supposent l'existence deux secteurs, moderne et traditionnel.

Nous commençons par la présentation des différentes théories de l'espace et de la localisation, des fondements théoriques de la nouvelle économie géographique, ainsi qu'une présentation simplifiée du modèle de Krugman (1991), développée par Forshild et Ottaviano

(1999). En fin, comment les modèles de la synthèse géographie - croissance permettent l'étude des liens entre la croissance économique et la répartition spatiale des activités économiques.

#### **1-4-1 les théories de l'espace et de la localisation :**

Les théories qui rendent compte des différentes forces qui poussent vers plus au moins d'agglomération, peuvent être renvoyées aux théories qui traitent de l'espace et de l'agglomération telles que la théorie traditionnelle du commerce, et celle de l'économie géographique.

La Théorie traditionnelle du commerce sous rendements croissants et concurrence parfaite telle que la théorie des avantages comparatifs, stipule que chaque localisation se spécialise dans la production de biens pour lesquelles elle possède des avantages comparatifs. Dans les modèles Ricardiens, les avantages comparatifs sont le résultat de différences exogènes dans les technologies, ou de différences exogènes dans les dotations suivant la théorie d'Heckscher-Ohlin. Ces théories sont peu éclairantes quand à l'explication de la persistance des disparités régionales (Thisse, 1997 ;Walz, 1999, Martin,2000). Les avantages comparatifs devaient conduire à une spécialisation qui réduira les différences des prix des facteurs. L'absence de la dimension économique de l'espace est une lacune (Thisse, 1997) qui rend les prédictions de la théorie traditionnelle incapable d'expliquer la distribution géographique des activités économiques.

La nouvelle économie géographique est un ensemble de modèles supposant des rendements d'échelle croissants pour des technologies identiques entre les localisations. La concurrence est imparfaite et les dotations sont identiques, mais les facteurs peuvent être mobiles entre les localisations tels que les revenus et les prix des facteurs peuvent être entraînés de manière endogène par le lieu de localisation. Partant des imperfections de la théorie traditionnelle du commerce, Krugman (1991) développe un cadre théorique pour expliquer l'agglomération spatiale des activités et les divergences de spécialisation régionales.

#### **1-4-2 L'économie géographique et l'explication de la formation des agglomérations :**

L'explication de la formation des agglomérations dans un nombre réduit de localisations passe par la supposition de l'existence d'un environnement économique marqué par des imperfections de marché et des externalités. L'apparition d'une

agglomération est le produit de l'équilibre entre les forces d'agglomération et les forces de dispersion. Les deux concepts de base pour expliquer les mécanismes de formation de ces agglomérations sont : la concurrence imparfaite et les externalités. La présence de rendements d'échelle croissants explique l'agglomération des activités dans un espace homogène. L'hypothèse des rendements rend l'hypothèse de concurrence parfaite intenable. L'analyse du processus d'agglomération suppose, alors, à la fois rendements croissants et concurrence imparfaite au sein d'une structure d'équilibre général, et ce afin que les interactions entre marchés des produits et du travail soient présentes. Mais il faut encore qu'il y ait commerce à l'équilibre si l'on veut que les régions ou les nations ne soient pas en autarcie.

***1- la concurrence monopolistique*** : Le concept de concurrence monopolistique remonte à Chamberlin (1933), il peut être décrit au moyen des quatre hypothèses suivantes :

- les biens sont différenciés et les entreprises vendent des produits de même nature mais imparfaitement substituables. Chaque firme a un pouvoir de monopole sur sa variété ;

- chaque firme produit une seule variété à rendements croissants, dont elle choisit le prix ;

- le nombre d'entreprises appartenant à l'industrie considérée est suffisamment élevé pour que chacune d'elles soit négligeable par rapport à l'ensemble des autres ;

- il y a libre entrée et sortie dans la branche de sorte que les profits sont nuls.

La différence principale de ces hypothèses avec celles de la concurrence parfaite réside dans le fait que chaque firme vend un produit spécifique dont elle choisit le prix, ce qui lui assure un marché propre. Toutefois, l'existence d'autres variétés implique que la taille de ce marché dépend du comportement des autres firmes, restreignant ainsi la liberté du producteur dans son choix de prix, sans toutefois la supprimer. Autrement dit, si la firme n'est pas en situation de concurrence parfaite, elle n'est pas davantage en situation de monopole absolu. Enfin, pour les firmes opérant à rendements croissants, les ressources disponibles dans l'économie imposent une limite au nombre de variétés susceptibles d'être produites. Ce nombre dépend évidemment des éventuelles barrières à l'entrée que les firmes sont susceptibles d'ériger. Dans le cas de la concurrence monopolistique, on suppose que le coût fixe, associé au lancement d'une nouvelle variété, est la seule barrière effective, laquelle ne peut être manipulée par les firmes.

En dépit d'un succès certain dans les années 1930, les idées de Chamberlin ont longtemps sommeillé jusqu'à ce que Spence (1976) et, surtout, Dixit et Stiglitz (1977), proposent une modélisation susceptible d'être utilisée dans différents domaines de la théorie économique. Il est intéressant de remarquer que ces deux articles comptent parmi les premiers qui vont marquer le renouveau de l'économie industrielle. Le modèle de Spence, bien que proche de celui de Dixit et Stiglitz, procède de l'analyse d'équilibre partiel à la différence du second qui se situe d'emblée dans un contexte d'équilibre général. Dans le modèle de Dixit et Stiglitz, l'économie est caractérisée par des biens différenciés et chaque firme produit une variété unique de sorte que celle-ci en détient le monopole. Le nombre de firme set donc de biens est important de sorte que les interactions stratégiques sont absentes. Il n'existe pas de barrières à l'entrée et les nouvelles firmes peuvent rentrer dans le marché jusqu'à ce que le profit s'annule. L'ensemble des biens est valorisé de manière symétrique, ce qui requiert une fonction d'utilité de type CES (Constant Elasticity Substitution).

Fujita, Krugman et Venables résument les avantages du modèle de Dixit-Stiglitz dans trois points : en premier lieu la possibilité qu'il permet pour analyser les effets de localisation de la concurrence imparfaite et des rendements croissants sans avoir recours au caractère complexe des interactions stratégiques, traités dans le modèle de duopole à la Hotelling (1929). En second lieu, il offre un cadre d'analyse qui permet d'expliquer conjointement la localisation des consommateurs et des producteurs, par la substitution du cadre de l'équilibre partiel par à celui de l'équilibre général. En dernier lieu, la taille importante des firmes dans un marché de type Dixit-stiglitz, rend l'approche microéconomique des choix de localisation sous rendements croissants compatible avec une analyse agrégée des conséquences de ces choix (Matsuyama, 1995).

## ***2- Les externalités :***

Suivant la distinction de Scitovski (1954) et selon la manière dont ils se manifestent, il existe deux types d'externalités : les externalités pécuniaires et les externalités technologiques. Les externalités pécuniaires se propagent à travers les mécanismes du marché c'est-à-dire par les prix (Fujita et Thisse, 1997). Les externalités technologiques affectent, directement la fonction d'utilité des consommateurs ou la fonction de production des entreprises, par l'effet de proximité et d'échange entre agents. Ils touchent les activités d'innovation et comme le souligne Arrow (1962), posent le problème de la sous optimalité des incitations à innover émanant du marché. Contrairement aux externalités technologiques, les externalités pécuniaires décrivent un bénéfice induit par des interactions économiques

associés aux mécanismes d'offre et de demande et donc transitant par le système des prix. Les externalités correspondent souvent à des formes de non-convexité de la fonction de production ou de consommation.

À la différence des économies d'échelle (interne à la firme), les économies de localisation (ou de spécialisation à la MAR (Marshall, Arrow et Romer), selon la classification de **Glaser** et al. 1992) sont externes à la firme et affectent toutes les firmes appartenant au même secteur Hoover (1936). En effet, la concentration géographique de firmes travaillant dans la même branche est souvent évoquée pour expliquer la spécialisation productive de certains espaces (Marshall, 1920 ; Krugman 1991, Hederson, 1988 ; Enright, 1990 ; Kim, 1995 ; Ellison et Glaser, 1997, 1999). Les économies d'urbanisation (où à la Jacobs) sont externes à la branche mais internes au système économique : elles apparaissent lors de l'agglomération d'industries différentes et sont ainsi représentatives des espaces urbains. Les externalités peuvent agir en tant que forces d'attraction, si elles sont positives, ou des forces de répulsion, si elles sont négatives.

Deux raisons sont avancées pour expliquer comment les externalités influencent le choix de localisation. Une première, d'ordre général, fait intervenir les bénéfices tirés de la proximité. Ces derniers sont d'autant plus élevés que les firmes et les individus sont proches les uns des autres (Fujita et thisse, 1997). Une seconde raison est liée à la transmission des connaissances technologiques. En effet, la transmission est plus intense entre les firmes quand de nombreux face-à-face sont répétés (Dosi, 1988). La proximité géographique facilite ces face-à-face. Krugman (1991) décrit les mécanismes de causalité circulaire enclenchés par les externalités externes, par la tendance de la production industrielle à se concentrer dans les espaces où existent des marchés de taille importante, et par l'importance de ces derniers dans les endroits où la production est concentrée.

Les externalités locales donnent un rôle important à la géographie dans la croissance de la production et les activités d'innovation, mais ce n'est pas la seule voie (Baldwin et Martin, 2004). Martin et Ottaviano (2001), dans leur étude du cas des innovations optent pour l'effet des liens verticaux plutôt que l'effet des externalités de localisation sur les innovations. Le secteur d'innovation utilise en input les biens industriels, la localisation des industries affecte le coût de l'innovation à travers les coûts commerciaux. Yamamoto (2002) présente un modèle similaire où des interactions circulaires entre la croissance et l'agglomération viennent des liens verticaux entre le secteur des biens intermédiaires et celui de l'innovation. Tel que Grossman et Helpman (1991), Quah (2002) propose un type de modèle sans échange

commercial et où l'externalité de connaissance est imparfaite dans le temps et dans l'espace. Dans ce cas, les externalités de connaissance sont localisées et les activités à rendement croissant se concentrent dans l'espace.

La nouvelle économie géographique renvoie les causes des disparités spatiales aux externalités pécuniaires et aux rendements croissants. Ces deux notions amènent à modéliser les relations économiques entre agents avec introduction des coûts de transfert (coûts de transport et autres coûts liés à l'échange). Ces coûts sont fondamentaux dans l'explication des différenciations spatiales et de l'intégration régionale. Les objectifs initiaux de la nouvelle économie géographique renvoient aux conditions rendants soutenable la concentration spatiale des activités économiques et l'apparition des agglomérations, ainsi que la définition des conditions d'instabilité d'une répartition spatiale équilibrée des activités, et qui rend l'équilibre spatial symétrique instable aux perturbations (Fujita, Krugman et Venables, 1999).

#### **1-4-3 Effet de la taille du marché et la structure centre périphérique :**

La théorie moderne de l'agglomération est très dominée par le principe que la place du marché est le centre de l'économie parce que cette place représente son marché principal. Ce principe est appelé l'effet de la taille du marché (HME : Home Market Effect). Suivant Helpman et Krugman (1985), quand on introduit les coûts de transport, l'effet de la taille du marché se produit quand des industries en concurrence imparfaite tendent à concentrer leurs productions dans les grands marchés et à exporter aux petits marchés. La taille du marché joue le rôle de force gravitationnel qui attire les secteurs en concurrence imparfaite vers les grands marchés. Cette notion est au cœur de la plupart des modèles d'agglomération. En utilisant deux modèles différents de la concurrence monopolistique, on peut voir qu'une partie plus que proportionnelle du secteur moderne se localise dans les grandes régions. Le capital est utilisé dans chaque région et peut se déplacer d'une région à une autre pour les taux de rendement nominal les plus importants.

L'effet de la taille ne peut pas expliquer pourquoi des petits chocs temporaires peuvent avoir des effets permanents sur le paysage économique. Les modèles centre et périphérie (CP), partant d'un faible niveau des coûts de transport et de l'hypothèse que les travailleurs dépensent leurs revenus dans leur région de travail, montrent que cette immobilité des revenus exacerbe l'effet transitoire de la taille du marché causé par des chocs temporaires. En d'autres termes, la mobilité du travail renforce d'une manière plus que proportionnelle l'avantage initial de la taille du marché. Au contraire, quand les coûts de transport sont élevés, la taille du

marché n'a pas d'effet sur la taille des régions. Comme conséquence à de faibles coûts de transport et à la mobilité des facteurs de productions, le modèle CP est le premier équilibre général à produire un espace économique irrégulier à partir d'un espace physique régulier. Comparé à l'effet de la taille du marché, dans le modèle CP le capital humain se localise là où il y a les taux réels de rendements les plus importants.

Venables (1996) dans l'étude des modèles de liens verticaux, montre que l'agglomération peut se produire même en absence de la mobilité du travail. La taille du marché régional est déterminée par les liens input - output. En d'autres termes, à chaque fois qu'une entreprise s'installe sur un marché, elle augmente la taille de celui-ci en amont et en aval.

### **A) L'effet de la taille du marché :**

La première étape pour comprendre la nouvelle économie géographique consiste à considérer l'effet de la taille du marché (HME). Dans le cas de deux économies, l'HME implique que la localisation avec une large demande locale attire une partie plus que proportionnelle de firmes appartenant aux industries en concurrence monopolistiques. Ainsi, la demande induit la spécialisation spatiale dans les flux commerciaux et génère la prédiction théorique qui stipule que les grandes régions peuvent être des exportateurs nets de biens produits sous des rendements croissants et une concurrence imparfaite. L'espace est décrit, comme dans la théorie commerciale, par deux régions. La maximisation du profit des firmes conduit à la minimisation des coûts de transport induit par le transfère des outputs. Dans ce sens et toutes choses égales par ailleurs, les grands marchés sont les places dominantes pour les entreprises. Cependant, pour relaxer la pression de la concurrence des prix et par ce que les prix de transport sont positifs, certaines entreprises préfèrent se localiser dans les petits marchés et loin de leurs concurrents.

Considérons deux régions (A et B), et deux secteurs économiques, le secteur moderne X et le secteur traditionnel Z. Il y a deux facteurs de production, capital (K) et le travail (L).

Les deux régions peuvent être différentes dans leurs dépenses, chaque économie est dotée avec K capitalistes et L travailleurs où chacun offre une unité de son facteur correspondant (inelastically). A la différence du travail, le capital est parfaitement mobile entre les régions. Les travailleurs sont distribués tel qu'une fraction  $\theta \in [0,1]$  réside dans la région A:  $L_A = \theta L$ , cette région est supposée la plus importante en terme de travail ( $\theta > \frac{1}{2}$ ).

Pour exclure les avantages comparatifs à la Heckscher-Ohlin, les capitalistes sont distribués avec la même fraction  $\theta$  dans la région  $A$ :  $K_A = \theta K$ . La dotation relative des facteurs est alors la même entre les deux régions.

Le secteur  $Z$  produit un bien homogène sous rendements d'échelle constants et une concurrence parfaite. En particulier, une unité d'output requiert une unité de travail. La maximisation du profit implique que  $p_i^Z = w_i$  ( $w_i$  est égal au salaire). Le prix de ce bien est le même quelque soit la région, il est égal au numéraire :  $p_A^Z = p_B^Z = 1$  et  $w_A = w_B = 1$ .

Le secteur  $X$  produit, sous des rendements croissants, un continuum de variétés horizontalement différenciées. La production de  $q(s)$  de la variété  $s$  nécessite une quantité  $f$  de capital une quantité variable  $mq(s)$  de travail. Le coût total pour une firme produisant la variété  $s$  est de :

$$TC_i(s) = r_i f + w_i m q_i(s) \quad (1)$$

$r_i$  est le taux de rendement du capital dans la région  $i$  chaque variété est produite par une seule firme. Les consommateurs sont supposés avoir une préférence pour la variété, chaque firme obtient une part plus importante du marché en produisant une variété différenciée qu'en répliquant une variété existante. Les régions sont homogènes, et les coûts intra-régionaux sont nuls. Les marchés régionaux sont segmentés, chaque firme fixe son prix en fonction de la spécificité du marché où sa variété est vendue (Head et Mayer, 2000).

L'équilibre du marché est le produit de l'interaction entre les forces de dispersion et les forces d'agglomération. Les forces centripètes se trouvent dans l'accès au marché : la localisation dans les grands marchés économise sur les coûts de transport, et compense les effets de la concurrence spatiale. Les forces centrifuges se trouvent dans l'encombrement du marché : la concurrence féroce se produit quand les firmes sont localisées côté à côté, ce qui correspond à l'effet du prix en concurrence spatiale.

Soit  $\lambda \in [0,1]$  une fraction du capital investi en  $A$ , tel que  $(\theta - \lambda)K > 0$  ou  $< 0$  mesure l'importance du flux du capital investi en  $A$  (ou hors de  $A$ ). Soit  $r_i(\lambda)$  le taux de rendement du capital dans la région  $i = A, B$  quand la distribution spatiale est  $(\lambda, 1 - \lambda)$ . Un équilibre spatial se produit à  $\lambda \in [0,1]$  quand :  $\Delta r(\lambda) \equiv r_A(\lambda) - r_B(\lambda) = 0$  où à  $\lambda = 0$  quand  $\Delta r(0) \leq 0$ , où à

$\lambda=1$  quand  $\Delta r(1) \geq 0$ . un tel équilibre existe toujours quand  $r_i(\lambda)$  est une fonction continue de  $\lambda$  (Grinsburgh et al., 1986).

La préférence des consommateurs de la région  $i$  est présentée par la fonction d'utilité suivante : (Ottaviano et Thisse, 2003)

$$U_i = Q_i^u Z_i^{1-u} \quad (2)$$

où

$$Q_i = \left( \int_0^N q_i(s)^{\frac{\sigma-1}{\sigma}} ds \right)^{\frac{\sigma}{\sigma-1}} \quad (3)$$

$Q_i$  est la consommation du bien X,  $Z_i$  est la consommation du numéraire,  $q_i(s)$  la consommation de la variété  $s$  du bien X, et N le nombre (masse) des variétés. Parce que chaque variété est négligeable,  $\sigma > 1$  est à la fois élasticité de la demande de chaque variété et l'élasticité de substitution entre deux variétés<sup>3</sup>.

La consommation des variétés étrangères impose un coût de transfert de type iceberg. Si un bien doit être transporté de la région A vers la région B et si  $q$  unités de ce bien doivent arriver à destination, il faut en expédier  $\tau q$  unités sachant que  $\tau \geq 1$  (le cas où  $\tau = 1$  équivaut à admettre que les coûts de transport sont nuls).

Le revenu régional  $Y_i$  consiste en la somme des rendements du capital  $R_i$  et des salaires  $w_i L$ . Le taux de rendement dans la région  $i$  est donné par :

$$r_A(\lambda) = \frac{u}{\sigma K} \left[ \frac{Y_A}{\lambda + \phi(1-\lambda)} + \frac{Y_B}{\phi\lambda + (1-\lambda)} \right] \quad (4)$$

$\phi = \tau^{1-\sigma} \in (0, 1]$  mesure le degré d'ouverture économique, il augmente lorsque  $\tau$  diminue.

---

<sup>3</sup> L'élasticité de substitution varie donc entre 1 et  $\infty$ . Les variétés sont totalement indépendantes lorsque  $\sigma = 1$  et des substituts parfaits lorsque  $\sigma \rightarrow \infty$ ; En revanche, pour toute valeur intermédiaire de  $\sigma$ , elles sont des substituts imparfaits.

Pour  $\lambda \in (0,1)$  la distribution d'équilibre des firmes qui résolvent (ou résout)  $r_A(\lambda) = r_B(\lambda) = r$  implique :

$$Y_A = \theta(rK + L) \text{ et } Y_B = (1-\theta)(rK + L) \quad (5)$$

( $w_i = 1 \equiv$  numéraire puisque le travail est immobile). La combinaison des équations du revenu et du taux de rendement donne la distribution d'équilibre des firmes suivante :

$$\lambda^* = \frac{1}{2} + \frac{1+\phi}{1-\phi} \left( \theta - \frac{1}{2} \right) \succ \frac{1}{2} \quad (6)$$

Cette expression révèle la présence de l'effet de la taille du marché puisque :  $\frac{d\lambda^*}{d\theta} = \frac{1+\phi}{1-\phi} \succ 1$ . Ainsi, la plus grande région reçoit une part plus que proportionnelle de firmes. Pour une interprétation plus facile, l'équation de  $\lambda^*$  peut se réécrire de la manière suivante :

$$(1+\phi) \left( \theta - \frac{1}{2} \right) - (1-\phi) \left( \lambda^* - \frac{1}{2} \right) = 0 \quad (7)$$

Le premier terme de gauche dépend de la distribution spatiale des consommateurs  $\theta$ . Puisque le coefficient de  $\left( \theta - \frac{1}{2} \right)$  est positif, ce terme mesure les avantages de l'accès au marché des grandes régions en présence de barrières commerciales. Le second terme dépend de la distribution interrégionale des firmes  $\lambda$ . Puisque le coefficient de  $\left( \lambda^* - \frac{1}{2} \right)$  est négatif, ce terme mesure l'inconvénient de l'encombrement du marché de la région qui reçoit un grand nombre de firmes. De faibles barrières commerciales ( $\tau$  proche de 1) et une faible élasticité de substitution ( $\sigma$  proche de 1) procure une large valeur à  $\phi$  (proche de 1), ce qui renforce l'avantage d'accès et affaiblit les désavantages de l'encombrement (car  $(1-\phi)$  tend vers 0). En particulier,  $\frac{d^2\lambda^*}{d\theta d\phi} \left( = \frac{1}{(1-\phi)^2} \right) \succ 1$ , révèle l'accroissement de l'HME avec l'ouverture commerciale.

**Proposition 1 :** *considérons une économie avec deux régions et deux secteurs. Si les deux régions diffèrent dans leurs dépenses sur le bien moderne, alors l'équilibre du marché implique une partie plus que proportionnelle du secteur moderne dans la région qui dépense le plus.*

Par conséquence, des chocs faibles et permanents sur la demande relative induisent de larges et permanentes différences dans la spécialisation régionale.

### **B) La structure centre – périphérie**

Cette structure est due à Krugman (1991), elle admet que même des petits chocs temporaires induisent de larges et permanentes différences entre les régions. Le modèle proposé met la lumière sur les aspects spatiaux du processus d'industrialisation.

Dans ce modèle,  $K$  ne représente plus le capital physique, mais plutôt la capital humain contenu dans le facteur travail. A la différence de ce qui précède, le propriétaire du capital offre ses services seulement dans la région où il réside. Par conséquence,  $K$  et  $L$  représentent respectivement le travail qualifié et non qualifié. Le travail qualifié est mobile entre les régions, et le travail non qualifié est sédentaire. De plus, le travail non qualifié est identiquement distribué entre les deux régions, parce que ces dernières sont considérées à priori symétriques.

L'équilibre est le résultat de l'interaction entre l'effet d'encombrement et celui d'accès au marché. Cependant, à la différence de la structure HME, la mobilité du travail affecte les côtés offre et demande dans la région de destination (non pas le côté offre uniquement). En effet, la localisation des firmes est renforcée par celle du travail qualifié et vice versa. Le travail qualifié, pour un meilleur bien être, migre dans les régions dotées d'un nombre important de firmes. La croissance du nombre de consommateurs augmente la demande sur les biens différenciés, et ainsi conduit à la création de nouvelles firmes.

Soit  $N = K/f$  le nombre total des firmes, et  $f$  le capital fixe. Le nombre de firmes dans chaque région est proportionnel à la quantité de capital investi localement, il est donné par l'équation (8) suivante :

$$n_A = \frac{\lambda K}{f} \quad n_B = \frac{(1-\lambda)K}{f} \quad (8)$$

L'équation (8) implique que  $\lambda$  mesure la fraction de firmes et du travail qualifié dans la région A (on suppose que chaque travailleur dispose d'une unité de capital (Combes et al, 2005)). Soit  $v_i(\lambda, \phi)$  l'utilité indirecte du travail qualifié dans la région  $i$ , l'équilibre spatial au point  $\lambda \in (0,1)$  se produit quand :  $\Delta v(\lambda) \equiv v_A(\lambda) - v_B(\lambda) = 0$  ou au point  $\lambda = 0$  quand  $\Delta v(0) \leq 0$ ,

ou au point  $\lambda = 1$  quand  $\Delta v(1) \geq 0$ . Un tel équilibre existe toujours quand  $v_r(\lambda)$  est une fonction continue de  $\lambda$ , mais cet équilibre n'est pas unique (et n'est pas toujours stable).

La stabilité d'un équilibre est étudiée à l'aide de l'équation suivante :

$$\frac{d\lambda}{dt} = \begin{cases} \Delta v(\lambda) & \text{si } 0 < \lambda < 1 \\ \min\{0, \Delta v(\lambda)\} & \text{si } \lambda = 1 \\ \max\{0, \Delta v(\lambda)\} & \text{si } \lambda = 0 \end{cases} \quad (9)$$

Si  $\Delta v(\lambda)$  est positive et  $\lambda \in (0,1)$  les travailleurs migrent de B à A ; si elle est négative, ils migrent dans le sens opposé. Chaque équilibre spatial est un état régulier du processus (9).

Un équilibre spatial est localement stable si, pour chaque déviation marginale de la distribution de la population de la situation d'équilibre, le processus (9) ramène cette distribution à la position d'équilibre de départ.

L'utilité indirecte est donnée par l'expression suivante :

$$\Delta v(\lambda, \phi) = u''(1-u)^{1-u} \left\{ \frac{w_A(\lambda, \phi)}{P_A(\lambda, \phi)^u} - \frac{w_B(\lambda, \phi)}{P_B(\lambda, \phi)^u} \right\} \quad (10)$$

où  $w_i(\lambda, \phi)$  représente le salaire prévalant dans la région  $i$ . L'indice de prix dans la région A est donné par :

$$P_A(\lambda, \phi) = \frac{m\sigma}{\sigma-1} [\lambda + \phi(1-\lambda)]^{\frac{1}{1-\sigma}} \left( \frac{K}{f} \right)^{\frac{1}{1-\sigma}} \quad (11)$$

La présence de  $P_A$  et de  $P_B$  dans l'équation (10) ajoute de nouveaux items à la liste des effets de localisation. En particulier (11) montre que, pour un salaire donné, la région avec un nombre important de travailleurs qualifiés, et donc de firmes manufacturières, présente un faible niveau général des prix et donc un pouvoir d'achat important. Un faible indice de prix et un grand nombre de firmes domestiques implique moins de variété manufacturière importée et donc moins de coûts (de transport) à assumer par les consommateurs (effet du coût de la vie). Ces effets additionnels, associés à l'effet de la taille du marché, militent plus pour une agglomération des industries contre l'effet d'encombrement du marché.

La détermination des salaires des travailleurs qualifiés d'équilibre  $w_A^*$  et  $w_B^*$  conduit au calcul du rapport suivant :

$$\frac{w_A^*(\lambda, \phi)}{w_B^*(\lambda, \phi)} = \frac{2\phi\lambda + [1 - u/\sigma + (1 + u/\sigma)\phi^2](1 - \lambda)}{2\phi(1 - \lambda) + [1 - u/\sigma + (1 + u/\sigma)\phi^2]\lambda} \quad (12)$$

La différentiation de (12) par rapport à  $\lambda$  montre que la région avec le plus grand nombre de travailleurs offre des salaires larges (faibles) chaque fois que  $\phi$  est supérieur (inférieur) au seuil :

$$\phi_r = \frac{1 - u/\sigma}{1 + u/\sigma} \quad (13)$$

avec  $\phi_r \in (0, 1)$ . Ce ci est le résultat de l'échange entre des forces opposées. D'un côté, pour des coûts de transport donnés, un grand nombre de travailleurs qualifiés dans une région donnée entraîne un grand nombre de firmes industrielles concurrentes, ce qui baisse l'indice du prix local par effet de l'augmentation de l'offre, et induit une baisse de la demande locale adressée à chaque firme à cause de la fragmentation de la demande (effet d'encombrement du marché). Une faible demande conduit à de faibles profits, et donc à de faibles rémunérations du travail qualifié. D'un autre côté, recevoir plus d'entreprises implique aussi des profits additionnels et plus de revenus pour le travail qualifié, dont une fraction (de ces revenus) est dépensée sur les industries locales. Par conséquent, les dépenses locales sont importantes et qui, pour un indice de prix donné, accroissent la demande adressée aux firmes (effet de la taille du marché). Le premier (second) domine le second (le premier) à chaque fois que  $\phi < (>) \phi_r$ . Ce résultat est vrai quelque soit  $\lambda$ .

L'expression (13) montre que l'effet d'encombrement du marché est fort lorsque le coût du transport  $\tau$  est important, les firmes vendent surtout dans leur marché domestique à cause de l'importance des coûts de transport. Cet effet est aussi important quand l'élasticité du prix-demande ( $\sigma$ ) adressée aux firmes est large parce que la demande est très sensible aux variations de l'indice des prix. Enfin, comme le montre l'équation (13), l'effet de la taille du marché est important lorsque la part des revenus dépensés sur le produit industriel ( $u$ ) est importante. Le salaire du travail qualifié est important dans les régions où le nombre de la main-d'œuvre qualifiée est important pour des petites valeurs de  $\tau$  et  $\sigma$ , et une valeur importante de  $u$ .

On peut calculer  $\Delta v(\lambda, \phi)$  comme une fonction de  $\Delta V(\lambda, \phi)$  où :

$$\Delta V(\lambda, \phi) \equiv \frac{2\phi\lambda + [1 - u/\sigma + (1 + u/\sigma)\phi^2](1 - \lambda)}{[\lambda + \phi(1 - \lambda)]^{\frac{u}{1-\sigma}}} - \frac{2\phi(1 - \lambda) + [1 - u/\sigma + (1 + u/\sigma)\phi^2]\lambda}{[(1 - \lambda) + \phi\lambda]^{\frac{u}{1-\sigma}}} \quad (14)$$

$\Delta V(\lambda, \phi)$  est tout ce qui importe dans la détermination de l'équilibre. Tous les équilibres internes sont des solutions à  $\Delta V(\lambda, \phi) = 0$ , bien que les configurations totalement agglomérées  $\lambda = 0$  et  $\lambda = 1$  sont des équilibres si et seulement si  $\Delta V(0, \phi) < 0$  et  $\Delta V(1, \phi) > 0$  respectivement.

$$\Delta V(0, \phi) = -\Delta V(1, \phi) = \frac{[1 - u/\sigma + (1 + u/\sigma)\phi^2]}{[\phi]^{\frac{u}{1-\sigma}}} - 2\phi$$

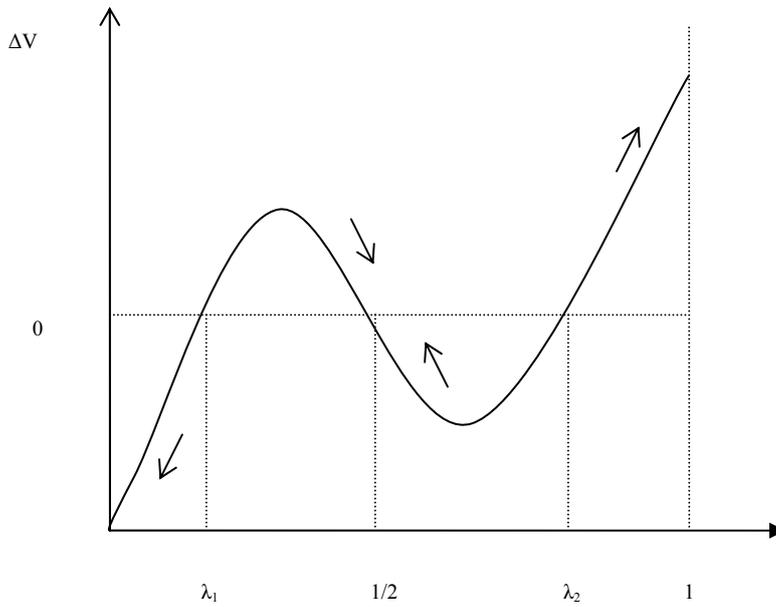
Une agglomération totale dans chaque région (structure CP) est un équilibre spatial stable chaque fois que les coûts de transport sont assez faible pour que  $\phi$  soit au dessus de  $\phi_s$ , appelé le point de soutient et défini par<sup>4</sup> :

$$\Delta V(0, \phi) = -\Delta V(1, \phi) = 0 \Rightarrow 1 - u/\sigma + (1 + u/\sigma)\phi_s^2 - 2\phi_s^{1+\frac{u}{1-\sigma}} = 0 \quad (15)$$

---

<sup>4</sup>  $\phi = \tau^{1-\sigma}$  sachant que  $1-\sigma < 0$  quand  $\tau \downarrow \Rightarrow \phi \uparrow$

**Figure 4 : Dynamique de la structure Centre – Périphérique**



Pour  $0 < \lambda < 1$  on peut montrer que  $\Delta V(\lambda, \phi) = 0$ , a au plus trois solutions. Il est facilement vérifiable que la structure symétrique, obtenue par  $\lambda = 1/2$  et qui entraîne une même distribution géographique des firmes et du travail qualifié, existe quelque soit la valeur des autres paramètres. Cette solution est stable chaque fois que  $\Delta V_{\lambda}(1/2, \phi) < 0$  (Pente de  $\Delta V(\lambda, \phi)$  au point  $\lambda = 1/2$ ). Ceci est le cas si et seulement si les coûts de transport sont suffisamment large pour que  $\phi$  soit au dessous de la valeur seuil  $\phi_b$  appelé point de brisure, défini comme :

$$\phi_b = \frac{1 - 1/\sigma - u/\sigma}{1 - 1/\sigma + u/\sigma} \phi_r \quad (16)$$

Le point de brisure décroît avec  $u$  et croît avec  $\sigma$ . Cependant, si  $\phi_b < 0$  et où  $u \geq \sigma - 1$  (condition du trou noir), la configuration symétrique n'est jamais stable et l'effet d'encombrement du marché est toujours dominé par celui de la taille du marché et celui du coût de la vie. Nous supposons pour la suite que  $u < \sigma - 1$ .

Autour de l'équilibre stable donné par  $\lambda = 1/2$ , il existe au moins deux autres équilibres symétriques instables. Ce dernier résultat provient de l'étude  $\Delta v(\lambda)$ , qui est symétrique autour de  $\lambda = 1/2$  et change de concavité au plus deux fois.

Les propriétés locales suivantes peuvent être élaborées au voisinage de  $\lambda = 1/2$  :

$$\Delta v_{\lambda_1}(1/2, \phi_b) = 0 \quad \Delta v_{\lambda_1 \lambda_1}(1/2, \phi_b) < 0 \quad (17)$$

$$\Delta v_{\lambda_2}(1/2, \phi_b) = 0 \quad \Delta v_{\lambda_2 \lambda_2}(1/2, \phi_b) > 0 \quad (18)$$

parce que  $\lambda = 1/2$  est toujours un équilibre stable,  $\Delta V(\lambda, \phi)$  pivote autour  $(\lambda, \Delta V) = (1/2, 0)$  quand  $\phi$  change. L'équation (17) indique que l'état régulier  $\lambda = 1/2$  passe de la situation stable à la situation instable quand  $\phi$  devient au dessus  $\phi_b$ . Comme l'équation (18) l'indique aussi, le changement de stabilité fait émerger deux équilibres additionnels symétriques. Ces propriétés montrent que l'équation (9) subit des bifurcations locales au point  $\phi = \phi_b$ . L'extension globale  $\Delta v_{\lambda_2 \lambda_2}(1/2, \phi_b) > 0$  implique que  $\phi_b > \phi_s$ .

Le fait que les coûts de transport au point de brisure sont inférieurs au coûts du point de soutient implique que le modèle montre une attraction locale. Une fois l'équilibre CP est atteint, les coûts de transport peuvent augmenter au dessus du point de brisure avant que l'agglomération ne cesse d'être un équilibre.

**Proposition 2 :** *considérons une économie avec deux régions et deux secteurs. Le secteur traditionnel emploie une main-d'œuvre géographiquement immobile distribuée d'une manière égale entre les deux régions. Le secteur moderne emploie une main-d'œuvre qualifiée géographiquement mobile. Pour de faibles coûts de transport le seul équilibre stable est celui où le secteur moderne s'agglomère en une seule région.*

En d'autres termes, pour des coûts de transport suffisamment bas, des petits chocs transitoires sur une situation initialement symétrique peuvent entraîner un déséquilibre dans la répartition spatiale au profit d'une seule région. La différence par rapport à la première proposition est due au fait que les travailleurs sont immobiles. La mobilité rend la taille du marché endogène, conduisant à l'émergence d'une causalité cumulative. Quand les coûts de transport sont faibles, l'attraction des consommateurs vers les grands marché rend ces derniers plus larges : la réallocation des consommateurs augmente les avantages de l'accès au marché et réduit les avantages d'encombrement. Ceci conduit à l'agglomération des toutes les firmes dans une seule région. Au contraire, si les coûts de transport sont importants, les inconvénients de l'encombrement du marché dominant les avantages d'accès, ce qui encourage la dispersion des firmes.

D'autres modèles d'agglomération ont été mis en oeuvre, ils font introduire les liens verticaux entre secteurs et firmes Krugman et Venables (1995) et Ottaviano (2002). Le travail

est supposé immobile entre les régions et mobile entre les secteurs. Pour Ottaviano et Thisse (2004), la présence des liens input – output entre les firmes, et sous réserve d'un faible coût de transport, rend l'agglomération de toutes les firmes en une seule région le seul équilibre stable.

### **C) Croissance et agglomération :**

Les déterminants spatiaux de la croissance ont été portés par les travaux de : Englmann et Walz (95), Walz (96), Baldwin et Forslid (2000), Martin et Ottaviano (99) et Baldwin, Martin et Ottaviano (2001). Ces modèles réalisent une synthèse entre la théorie de la croissance endogène et l'économie géographique. Parallèlement, au coût de transport et aux rendements croissants, la diffusion des externalités technologiques dans un espace géographique restreint rend la polarisation spatiale des activités plus bénéfique pour la croissance que la dispersion.

La présentation de ces modèles va nous montrer comment la localisation des externalités de la connaissance est étroitement liée celle des firmes et à la croissance. Les modèles de Martin et Ottaviano (99) et d'Englmann et Walz (95) font l'hypothèse que l'économie est marquée par des externalités de connaissance strictement locales puis globales. Dans le premier cas, ils définissent une autarcie technologique totale des régions et dans le second cas des environnements technologiques ouverts.

#### **C-1) le modèle de Martin et Ottaviano (99)**

Ce modèle suppose deux régions et deux facteurs de production (capital et travail), les travailleurs sont immobiles et l'agglomération est le résultat de la mobilité des firmes. La croissance est modélisée d'une manière similaire à Grossman et Helpman (1991). L'innovation concerne les biens de consommation finale et la croissance est soutenue par l'introduction d'une nouvelle variété d'un bien sur le marché, et les préférences sont données par la fonction CES. Chaque firme produit une seule variété. Il existe deux secteurs : un qui produit des biens homogènes en concurrence parfaite avec des coûts de transport nuls entre les régions, et un second qui produit des variétés différenciées sous concurrence monopolistique. Toute nouvelle variété introduite sur le marché est considérée comme innovation, elle nécessite un brevet à durée de vie illimitée. Chaque firme exerce une activité de recherche développement (RD) et une activité de production, ces deux activités peuvent être localisées dans deux régions différentes sans coût additionnel. L'échange de variétés entre régions exige un coût de type Iceberg.

La force de travail est considérée homogène et à la différence du modèle de Krugman (91), les salaires versés sont exogènes et normalisés à l'unité. Les marchés locaux de travail n'exercent aucun effet sur la localisation des firmes. Le capital est parfaitement mobile, et l'introduction d'une nouvelle firme est financée par l'investissement privé des travailleurs des deux régions. Le revenu total est la somme du revenu de l'investissement et celui du travail. La différence entre les régions réside dans la dotation initiale en capital, ce qui crée une différence dans les revenus des deux régions. De ce fait, il existe un écart permanent entre les tailles des deux marchés régionaux.

Les firmes produisant les biens différenciés sont plus portées à se localiser dans la région favorisée par le plus grand marché et à exporter à l'autre région, toutefois cette attirance vers les grands marchés est régulée par le coût de transport. L'effet de la taille du marché fait en sorte que l'équilibre est réalisé par l'agglomération des firmes dans la région offrant le revenu du capital le plus important.

Les activités d'innovation se localisent en fonction des externalités technologiques qu'elles peuvent capter, ces externalités sont strictement locales et sont générées par des activités productives. Les coûts d'innovation sont fonction décroissante du nombre de firmes. Les nouvelles firmes innovantes se localisent dans les régions où les coûts de RD sont les plus faibles. L'agglomération des activités de RD dans la région favorisée induit un équilibre CP. Toutes les activités RD se concentrent dans la région où le volume des externalités technologiques est le plus élevé.

Martin et Ottaviano (99) montrent que plus le niveau d'agglomération est important, plus le volume des externalités locale est important et le coût de l'innovation faible, ce qui augmente l'attractivité de la région à d'autres entrants et donc son taux de croissance. L'agglomération agit positivement sur la croissance, la baisse du coût de transport étant un déterminant positif indirect.

Pour les externalités globale, Martin et Ottaviano font l'hypothèse d'une diffusion globale des connaissances. Contrairement à une situation d'externalités locales, l'innovation va bénéficier d'un même volume d'externalités quelle que soit sa localisation. Tous les équilibres de localisation des activités d'innovation sont possibles y compris celui d'une répartition égale entre les régions. Le coût de l'innovation n'est pas sensible au niveau d'agglomération des activités industrielles. Puisque les connaissances se diffusent à l'ensemble de l'économie, la productivité des activités d'innovation dans une région est indépendante du nombre de firmes localisées dans cette région. La croissance globale de

l'économie ne dépend pas de la distribution spatiale des activités industrielles et du niveau de l'agglomération.

### **C-2) le modèle d'Englmann et Walz (1995)**

Le processus de croissance est emprunté aux modèles de Romer (1990) et Grossman et Helpman (91). Les biens intermédiaires sont à l'origine de deux types d'externalités : l'activité d'innovation préalable à leurs production, le deuxième est d'ordre pécuniaire. Les bien intermédiaires ne font pas l'objet d'échange inter-régions et les firmes de production des bien de localisation se localisent là où l'offre des biens intermédiaires est importante. Aussi, les firmes innovantes produisant les biens intermédiaires favorisent la localisation dans les marchés où la demande sur leurs outputs est importante.

Le facteur travail est composé de la main-d'œuvre non qualifiée et immobile, et de la main-d'œuvre qualifiée. Cette dernière migre d'une région à une autre à la recherche d'un salaire nominal plus élevé, déterminé de manière endogène sur le marché du travail.

Le secteur produisant les variétés différenciées n'utilise que le travail qualifié. La production de chaque variété est précédée par une activité d'innovation financée par des fonds privés et bénéficiant d'externalités strictement locales. Contrairement à Martin et Ottaviano (99), les activités de production et de RD d'une même firme ne peuvent pas se localiser dans deux régions différentes.

Le secteur produisant les biens de consommation ne sera présent dans les deux régions que si les activités d'innovation sont réparties initialement de manière égales. Toute perturbation donnant l'avantage à une région en terme de biens intermédiaires provoquera une divergence permanente en terme de taux régionaux d'innovation. Une région mieux dotée en terme de production de biens intermédiaires offre de meilleur gains de productivité et donc attire plus de firmes produisant les variétés de consommation finales. En retour, les firmes produisant les biens de consommation intermédiaire choisissent de se localiser dans les régions offrant des marchés importants en terme de production de biens de consommation finale. Par un processus de causalité cumulatif, l'économie converge vers une configuration de CP où toutes les activités innovantes et de production de biens intermédiaires s'agglomèrent dans une seule région.

Les externalités inter-régionales rendent les coûts engendrés par l'innovation insensibles à la distribution spatiale des activités. Une région créant plus de biens intermédiaires, et donc d'externalités de connaissance ne possède que l'avantage comparatif

dû à la production de biens intermédiaires puisque l'avantage de celui de la RD disparaît. Les deux auteurs montrent qu'un équilibre CP peut prévaloir sous certaines conditions. Des dotations initiales en biens intermédiaires différentes créent une convergence vers un état asymétrique où la production de ces biens se concentre en une seule région.

Selon Baldwin, Forshild, Martin, Ottaviano, Røbet-Nicoud (2001) les externalités technologiques ne semblent en effet ni globales ni entièrement locales. En présence d'externalité locales, l'agglomération, soutenu par la baisse des coûts de transport, peut être bénéfique à la croissance et engendrer des trajectoires de développement inégales entre les régions (Martin et Ottaviano (99), Englmann et Walz (95)). L'imparfaite mobilité du capital, associée à la baisse des coûts de transport, est un facteur déstabilisant d'une forte agglomération (Baldwin, Martin et Ottaviano (2001)). La croissance est un facteur d'agglomération, seule une grande diffusion interrégionale des connaissances empêcherait l'apparition d'une relation causale entre les deux phénomènes et pousserait vers une croissance plus équilibrée entre les régions (Baldwin, Martin et Ottaviano (2001), Baldwin et Forshild (2000))

## **Conclusion**

Dans ce chapitre, nous avons relaté les principes fondateurs et les différents travaux de recherche sur la question de la convergence des économies. Le premier fondement de la convergence se base sur des développements issus du modèle néoclassique de Solow. Le concept de convergence est différencié selon que les états réguliers des économies étudiées sont identiques ou non. Si les économies ne convergent pas vers le même état régulier, on parle de  $\beta$ -convergence conditionnelle, sinon la convergence est dite absolue. Les modèles de croissance endogène, se basant sur le rôle des diffusions technologiques dans le processus de rattrapage entre les économies permettent de donner une explication théorique au concept de club de convergence développé par plusieurs chercheurs empiristes.

La définition de la convergence n'échappe pas à certaines faiblesses et limites. En effet, elle ne permet pas de faire la distinction entre l'effet de long terme des conditions initiales et de l'hétérogénéité structurelle de long terme. À la différence des effets des caractéristiques structurelles, les effets de conditions initiales sur l'état régulier impliquent l'existence de club de convergence. La convergence absolue est souvent conditionnée par l'homogénéité de l'échantillon étudié, généralement la corrélation entre le revenu initial et la croissance n'est pas vérifiée dans le cas de pays hétérogènes. L'estimation  $\beta < 0$  n'est pas

consistante avec des alternatives économiques non convergentes. Le modèle de Azariadis et Drazen (90) montre que la convergence statistique peut coexister avec la présence de plusieurs états réguliers, et l'exemple traité par Bernat et Duralauf (96) montre que la convergence statistique peut être consistante avec une divergence économique. Ces deux résultats confirment la compatibilité de l'existence de sentier de croissance avec des différences permanentes de revenus entre économies dotées de même taux de croissance de la population, de l'épargne et d'un même accès à la technologie.

Une autre limite des l'approches néoclassiques de la croissance réside dans la négligence du rôle de l'espace dans la répartition et la localisation des activités économiques. Les théories de l'économie géographique apportent des réponses aux causes de la concentration spatiale des activités économiques et expliquent les interactions qui existent entre l'agglomération et la croissance. Dans la section croissance et géographie économique nous avons montré le rôle des forces qui agissent pour l'agglomération des activités industrielles et de RD, et ceux qui agissent pour la dispersion. Ces modèles mettent en avant l'effet de la taille du marché (Home Market Effect) et celle de la configuration centre - périphérie. Le niveau des coûts de transport entre les régions est un facteur important dans l'explication du schéma d'équilibre dominant dans la répartition des activités économiques, et des dynamiques de croissance. En effet, les modèles de synthèse, entre la théorie de la croissance endogène et l'économie géographique, montrent que la diffusion des externalités technologiques dans un espace géographique restreint rend la polarisation spatiale des activités plus bénéfique pour la croissance que le dispersion. La croissance économique peut en retour renforcer les processus de polarisation. Enfin, les effets d'externalités technologiques entre les régions renforcent la polarisation si elles sont locales et conduisent à une répartition plus équilibrée entre les régions si elles sont globales.

Ces résultats théoriques montrent l'importance de la prise en compte des facteurs déterminants la répartition spatiale des activités dans le processus de croissance des économies. Ces effets doivent être considérés dans l'analyse du processus de convergence. L'hétérogénéité spatiale doit être introduite dans l'analyse et la mesure de la convergence, puisque l'existence d'une configuration centre périphérique est plus adaptée avec les clubs de convergence. Les données sur les économies ne sont pas aléatoirement distribuées dans l'espace et les phénomènes d'auto corrélation spatiale rendent, s'ils ne sont pas considérés, inconsistantes les estimations de la convergence. Les méthodes de l'économétrie spatiale doivent prendre en compte les effets dus à l'hétérogénéité spatiale dans la distribution de la

croissance. Nous consacrons le chapitre suivant au développement des concepts de l'économétrie spatiale utilisés pour l'étude de la convergence régionale de l'emploi en Algérie.

## **Chapitre II : Économétrie Spatiale**

### **Introduction :**

Le développement de l'économétrie spatiale est relativement récent dans la littérature statistique. La violation des hypothèses classiques, à cause des similarités entre observations causées par la proximité spatiale, a induit des développements pour expliquer le phénomène de diffusion spatial rencontré dans plusieurs domaines tels que : épidémiologie, géographie, agriculture, traitement de l'image, sciences régionales, astronomie, et dans d'autres domaines (Haining, 2003).

Les techniques de la statistique spatiale sont à la base de l'économétrie spatiale. Les premières contributions remontent aux travaux de Whittle (1954, 1962, 1963), Bartlett (1963, 1975) et les contributions de Besag (1974). Les différents résultats obtenus ont conduit au premières codifications dans les années soixante-dix avec d'importantes publications, tel que le célèbre ouvrage de Cliff et Ord (1973).

Le terme économétrie spatiale a été inventé par Jean Paelinck dans la fin des années soixante-dix. Dans l'avant propos du premier livre consacré entièrement à ce sujet, Paelinck et Klaassen (1979) suggèrent que cette nouvelle branche de l'économétrie est un mélange de la théorie économique, de la formalisation mathématique et de la statistique mathématique. Cette nouvelle discipline s'identifie dans cinq caractéristiques (Paelinck et Klaassen 1979) : (1) le rôle de la dépendance spatiale dans les modèles spatiaux, (2) l'asymétrie des relations spatiales, (3) l'importance des facteurs explicatifs localisés dans d'autres espaces, (4) la différenciation entre les interactions ex-post et ex-ant, et (5) la modélisation explicite de l'espace.

A part quelques exemples de la possibilité d'utiliser les méthodes de la statistique spatiale dans l'économie (tels que ceux de Clive Granger, 1969, 1974), l'ouvrage publié par Luc Anselin en 1988, constitue une étape importante dans l'histoire du développement de cette discipline. Se basant sur des travaux précédents, il présente pour la première fois un traitement clair de sujets, tels que : la dépendance spatiale et l'hétérogénéité spatiale, fondamentales dans le traitement des données économiques spatiales. Il définit le sujet comme une collection de techniques qui traitent les particularités causées par l'espace dans l'analyse statistique des modèles des sciences régionales.

Actuellement un grand intérêt est porté à l'utilisation des méthodes de l'économétrie spatiale, cet intérêt témoigne du nombre important de papiers faisant référence aux problèmes spatiaux dans les journaux de l'économétrie et de l'économétrie appliquée. Beaucoup d'applications commencent à s'étendre à d'autres techniques économétriques : comme celles de donnée de panel, les modèles de convergence régionale, les modèles spatio-temporels, les variables qualitatives, les externalités spatiales, les modèles bayésiens et les techniques non paramétriques, etc...

Nous commençons, dans ce chapitre, par le développement des concepts de base liés à l'économétrie spatiale, puis nous introduirons les techniques de l'approche Bayésienne utilisée dans les modélisations faites dans cette thèse.

## **2.1 Le concept de processus spatial aléatoire (champ aléatoire)**

**Définition 1 :** soit  $(\Omega, B, P(\cdot))$  un espace probabiliste, avec  $\Omega$  espace des événements,  $B$  ensemble de Borel associé et  $P(\cdot)$  mesure de probabilité. De plus, soit  $S$  un ensemble non vide dans  $\mathfrak{R}^2$ , et soit la fonction  $X(\cdot, \cdot)$  une application de  $\Omega \times S \rightarrow \mathfrak{R}$ . La séquence ordonnée de variables aléatoires  $\{X(\cdot, s), s \in S\} = \{X(s), s \in S\}$  est appelée un processus spatial aléatoire.

Il y a deux caractéristiques fondamentales pour le processus spatial aléatoire  $\{X(s), s \in S\}$ . La première concerne la nature de l'indice  $s \in S$ , la topologie des observations. La seconde concerne la structure de dépendance spatiale affichée par l'ensemble des variables constituant le processus spatial.

### **2.1.1 La nature de l'indice $S$**

#### **2.1.1.1 Généralités :**

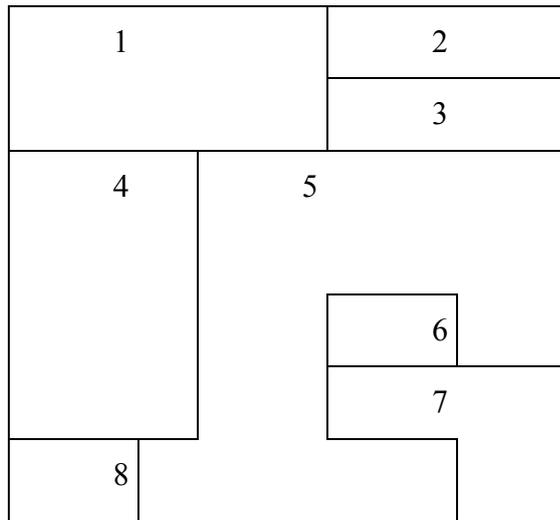
Dans notre cas, les indices appartenant à l'ensemble  $S$  représentent une série de valeurs ordonnées relatant un ensemble fini de polygones (ou régions en économie géographique) (Voir figure 1).

Le vecteur  $s$  représente des scalaires  $s \in N^+$  relatant des nombres arbitraires assignés à chaque polygone.

Les informations statistiques dans l'analyse de données économiques se réfèrent à des points dans l'espace, elles permettent de positionner les agents économiques, ou d'agréger les observations dans une unité territoriale telle que la région ou la commune.

On définit un champ aléatoire de paramètres continus par :  $\{X(s), s = (r, s), r, s \in \mathfrak{R}\}$  et un champ aléatoire de paramètres discrets par :  $\{X(s), s = s, s \in N^+\}$

**Figure 5 : Ensemble de polygones**



**2.1.1.2 La topologie du champ aléatoire :**

Dans le cas d'un champ aléatoire de paramètres continus, la topologie de l'espace de référence est spécifiée par le concept de métrique ou distance. Dans le cas d'une distance euclidienne dans un espace de taille n, on a :

$$d_{ij} = d(s_i, s_j) = \sqrt{(s_i - s_j)'(s_i - s_j)}$$

Avec  $d_{ij}$  la distance entre les points  $s_i$  et  $s_j$ . On considère souvent la distance comme représentative de celle entre les centroïdes des polygones. Cependant, cette solution n'est pas satisfaisante dans le cas de formes irrégulières des polygones. Comme alternative, on peut utiliser le concept de distance de Hausdorff entre les polygones. La distance de Hausdorff entre les polygones A et B est définie comme :  $H(A, B) = \max_{a \in A} \left\{ \min_{b \in B} [d(a, b)] \right\}$  où a et b représentent respectivement des points de A et B et  $d(a, b)$  une distance quelconque. Cette

métrique décrite comme la distance maximale d'un polygone au point le plus proche dans un autre polygone.

Dans le contexte économique, le choix de la distance appropriée dépend de la représentativité des agents économiques par rapport à un lien géographique. Des définitions alternatives considèrent le temps nécessaire pour joindre les points  $s_i$  et  $s_j$  ou le coût de transport entre ces points. Plusieurs autres définitions ont été suggérées à l'utilisation ; nous citons à titre d'exemple : les distances sociales (Doreian, :1980), les distances économiques (Case et al., 93 ; Conley et Topa, 2002), et les flux empiriques observés entre les régions (Murdoch et al.,97).

### 2.1.1.3 le concept de voisinage :

Plusieurs définitions sont possibles :

**Définition 2 : Seuil critique de voisinage** Deux régions  $s_i$  et  $s_j$  sont voisines, si  $0 \leq d_{ij} < d^*$ ,  $d^*$  représente le seuil critique.

**Définition 3 : Voisin le plus proche** Deux régions  $s_i$  et  $s_j$  sont voisines, si  $d_{ij} = \text{Min}(d_{ik}) \quad \forall i, k$

**Définition 4 : Voisinage basé sur la contiguïté** Dans le cas de processus à paramètres discrets, une définition simple de voisinage peut être basée sur l'adjacence entre deux polygones. Dans ce cas, deux régions  $s_i$  et  $s_j$  sont voisines, si elles partagent la même frontière.

Soit  $N(i)$  l'ensemble des voisins d'une région  $s_i$ , et soit  $\eta_i$  la cardinalité de cet ensemble. La matrice de contiguïté  $W$  d'éléments génériques  $w_{ij}$  est donnée par :

$$w_{ij} = \begin{cases} 1 & \text{si } j \in N(i) \\ 0 & \text{si non} \end{cases} \quad (1)$$

Notons que  $\eta_i = \sum_j w_{ij}$ . La matrice de contiguïté exprime les proximités existantes entre tous les couples de régions composant le processus spatial aléatoire (le champ aléatoire). Au lieu d'une matrice de poids binaire, on présente la matrice de poids généralisé de la manière suivante :

$$w_{ij} = g[d_{ij}] \quad (2)$$

$g$  est la fonction inverse de la distance, un exemple est donné par une pondération de type gravitationnelle :

$$w_{ij} = d_{ij}^{-\alpha} \quad \alpha > 0 \quad (3)$$

Souvent, il est utile de considérer une matrice de poids avec des lignes standardisées, dans le sens où la somme des éléments de chaque ligne est égale à 1.

$$w_{ij}^* = \frac{w_{ij}}{\sum_j w_{ij}} \quad (4)$$

$$\sum_j w_{ij}^* = 1$$

Cette dernière définition est très importante pour l'introduction du concept de décalage (retard) spatial.

### 2.1.2 Structure de dépendance dans un champ aléatoire

**Définition 5** Soit un champ aléatoire  $\{X(s), s \in S\} = \{X(s_1), X(s_2), \dots, X(s_n)\}$  et soit la topologie induite par la définition du voisinage  $N(i)$ , la valeur spatialement décalée de la variable aléatoire  $X(s_i)$  est définie comme :

$$L[X(s_i)] = \frac{1}{\eta_i} \sum_{s_j \in N(i)} X(s_j) \quad (5)$$

Où d'une manière équivalente :

$$L[X(s_i)] = \frac{1}{\eta_i} \sum_{j=1}^n w_{ij} X(s_j) = \sum_{j=1}^n w_{ij}^* X(s_j) \quad (6)$$

L'écriture matricielle donne :

$$L[X(s_i)] = W^* X(s) \quad (7)$$

$$X(s) = (X(s_1), X(s_2), \dots, X(s_n))'$$

**Définition 6** Pour chaque séquence de variables aléatoires  $X(s_1), X(s_2), \dots, X(s_n), n \geq 1$ , appartenant au champ  $\{X(s), s \in S\}$ , la fonction de densité marginale de  $X(s_1)$  peut être définie comme suit :

$$f_{X(s_1)}(x(s_1)) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_{X(s_1), X(s_2), \dots, X(s_n)}(x(s_1), x(s_2), \dots, x(s_n)) dx(s_2) \dots dx(s_n) \quad (8)$$

**Définition 7** Pour chaque variable aléatoire appartenant au champ  $\{X(s), s \in S\}$ , la moyenne, la variance et le moment d'ordre supérieur sont définis comme suit :

$$\begin{aligned} E(X(s_i)) &= \mu(s_i) \quad \forall i \\ V(X(s_i)) &= \sigma^2(s_i) \quad \forall i \\ E(X(s_i)^r) &= \mu_r(s_i) \quad \forall i, \forall r \geq 1, s \in S \end{aligned}$$

**Définition 8** soit le processus aléatoire spatial  $\{X(s), s \in S\}$ , la fonction d'auto-covariance spatiale du champ est définie par :

$$\gamma(s_i, s_j) = E[(X(s_i) - \mu(s_i))(X(s_j) - \mu(s_j))] \quad s_i, s_j \in \mathfrak{R}^2 \quad (9)$$

La forme standardisée donne la fonction d'auto-corrélation spatiale, elle est donnée par :

$$\rho(s_i, s_j) = \frac{\gamma(s_i, s_j)}{\sqrt{\gamma(s_i) \gamma(s_j)}} \quad s_i, s_j \in \mathfrak{R}^2 \quad (10)$$

Un des processus les plus utilisé en économétrie, est le processus Gaussien (Normal). La distribution de plusieurs processus empiriques, peut être approximée par ce dernier.

**Définition 9** un processus aléatoire spatial  $\{X(s), s \in S\}$  est dite Gaussien si, pour chaque sous ensemble fini dans  $\mathfrak{R}^2$ ,  $(s_1, s_2, \dots, s_n)$ ,  $(X(s_1), X(s_2), \dots, X(s_n)) = X(s)$  est tel que :

$$X(s) \sim N(\mu, V)$$

où

$$f_{X(s_1), X(s_2), \dots, X(s_n)}(x(s_1), x(s_2), \dots, x(s_n)) = \frac{|V|^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2}(X(s) - \mu(s))' V^{-1}(X(s) - \mu(s))\right\} \quad (11)$$

$\gamma(s_i, s_j) \in V, i=1, \dots, n; j=1, \dots, n$  ;  $V$  matrice de variance covariance de taille  $n \times n$  et  $\mu \equiv (\mu(s_1), \mu(s_2), \dots, \mu(s_n))'$  un vecteur de taille  $n \times 1$  des valeurs des moyennes mathématiques.

En analyse économétrique, nous analysons conjointement plusieurs phénomènes spatialement distribués. A titre d'exemple, le taux de croissance et le PIB dans toutes les localisations de l'espace. Pour prendre en compte cet aspect on introduit le concept de *vecteur de champs aléatoires de dimension k*.

**Définition 10** un vecteur de champs aléatoires  $\{X(s), s \in S\}$  est défini comme le champ  $X(s) = (X_1(s), X_2(s), \dots, X_k(s))'$  où chaque composante  $X_i(s)$  représente un champ aléatoire  $\{X_i(s), s \in S\}$ .

Plusieurs des définitions introduites par le champ aléatoire peuvent être généralisées pour le vecteur de champs aléatoires

**Définition 11.** Soit le vecteur de champs aléatoires  $\{X(s), s \in S\}$ , les quantités :

$$\gamma_{lk}(s_i, s_j) = E \left[ (X_l(s_i) - \mu_l(s_i)) (X_k(s_j) - \mu_k(s_j)) \right] \quad \forall l, k; \forall s_i, s_j \in \mathfrak{R}^2 \quad (12)$$

et

$$\rho_{lk}(s_i, s_j) = \frac{\gamma_{lk}(s_i, s_j)}{\sqrt{\gamma_{ll}(s_i) \gamma_{kk}(s_j)}} \quad \forall l, k; \forall s_i, s_j \in \mathfrak{R}^2 \quad (13)$$

sont appelées respectivement les fonctions spatiales de variance covariance et de corrélation.

**Définition 12 :** un vecteur de champs aléatoires  $\{X(s), s \in S\}$  est dite Gaussien si :

$$\begin{bmatrix} X(s_1) \\ X(s_2) \\ \dots \\ X(s_n) \end{bmatrix} \sim N \left( \begin{bmatrix} \mu(s_1) \\ \mu(s_2) \\ \dots \\ \mu(s_n) \end{bmatrix}, \begin{bmatrix} V(s_1) & C(s_1, s_2) & \dots & C(s_1, s_n) \\ & V(s_2) & & \\ & & \dots & \\ C(s_n, s_1) & & & V(s_n) \end{bmatrix} \right) \quad (14)$$

avec  $\mu(s_i)$   $k \times 1$  vecteur de moyennes,  $V(s_i)$   $k \times k$  matrice de variance dans la région  $i$  et  $C(s_i, s_j)$   $k \times k$  matrice de covariance entre les régions  $i$  et  $j$ .

$$V(s_i) = \begin{bmatrix} \gamma_{11}(s_i, s_i) & \gamma_{12}(s_i, s_i) & \dots & \gamma_{1k}(s_i, s_i) \\ \gamma_{k1}(s_i, s_i) & \gamma_{k2}(s_i, s_i) & \dots & \gamma_{kk}(s_i, s_i) \end{bmatrix} \quad (15)$$

et

$$C(s_i, s_j) = \begin{bmatrix} \gamma_{11}(s_i, s_j) & \gamma_{12}(s_i, s_j) & \dots & \gamma_{1k}(s_i, s_j) \\ \gamma_{k1}(s_i, s_j) & \gamma_{k2}(s_i, s_j) & \dots & \gamma_{kk}(s_i, s_j) \end{bmatrix}$$

## 2.2 La régression linéaire avec des données spatiales :

Soit  $Y(s_i)$  la variable indépendante du modèle à la localisation  $s_i$ ,  $X(s_i)$  vecteur des variables explicatives de dimension  $k$ .  $Z(s_i) = \left[ Y(s_i), X(s_i)' \right]'$  une collection de variables aléatoires appartenant au champ aléatoire  $\{Z(s), s \in S\}$  défini sur l'espace probabiliste  $(\Omega, B, P(\cdot))$  et qui génère un ensemble de données observées dans  $n$  localisations de coordonnées  $(s_1, s_2, \dots, s_n)$  dans un espace continu ou discret. Nous cherchons à construire un modèle qui explique le comportement de variables économiques  $Y(s_i)$  dans la localisation  $s_i$  en fonction des variables aléatoires  $X(s_i)$  constituant le processus aléatoire.

L'hypothèse de base derrière le modèle empirique est que les données sont tirées aléatoirement à partir d'une distribution conditionnelle  $Y$  sachant  $X$ . En d'autres termes,  $y \equiv (y_1, y_2, \dots, y_n)'$  est un échantillon indépendant généré de la densité  $f_{y_i|x_i}(y_i | X_i = x_i, \theta_i), i = 1, \dots, n$ . Cette hypothèse est violée puisque les variables dépendantes sont supposées corrélées dans l'espace. Les estimateurs des moindres carrés (OLS) de  $\beta$  et  $\sigma^2$  sont inefficients et inconsistants même s'ils restent sans biais. Les variances empiriques sont biaisées et, dans la majorité des cas, sous estimées. Comme conséquence, le coefficient de détermination multiple ( $R^2$ ) aussi bien que les statistiques  $t$  et  $F$  tendent à être gonflées. L'étude de la violation de la situation idéale de dépendance dans l'espace, nous conduit à l'étude de la corrélation entre les unités spatiales ou la corrélation spatiale.

### 2.2.1 Procédure pour tester la dépendance spatiale :

Moran (1950) et Cliff et Ord (1972) proposent un test qui peut être appliqué à l'étude de la dépendance spatiale dans les erreurs, du moins pour une analyse exploratoire. Le test se base sur la forme :

$$I = h \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} \hat{e}_i \hat{e}_j}{\sum_{i=1}^n \hat{e}_i^2} \quad (16)$$

n le nombre des observations spatiales,  $\hat{e}_i = y_i - \hat{\beta}'x_i$  les résidus de la régression linéaire,  $w_{ij} \in W$  matrice de contiguïté, et  $h = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}}$  facteur de normalisation. L'écriture matricielle

$$\text{donne : } I = h(\hat{e}'\hat{e})^{-1}(\hat{e}'W\hat{e}) = (\hat{e}'\hat{e})^{-1}(\hat{e}'W^*\hat{e}) \quad (17)$$

Cliff et ord (72) ont étudiés la distribution asymptotique de la statistique I, calculée à partir des résidus d'une régression OLS. Si les résidus sont distribués selon une loi normale, la statistique I est asymptotiquement normale avec les espérances mathématiques suivantes :

$$E(I) = \frac{k-1}{n-k-1} \text{tr}(MW^*)$$

et

$$V(I) = \frac{\text{tr}(MW^*MW^{*'}) + \text{tr}(MW^*)^2 + [\text{tr}(MW^*)]^2}{(n-k-1)(n-k+1)} - E(I)^2 \quad (18)$$

$\text{tr}(\cdot)$  dénote la trace d'une matrice,  $M = I - X(X'X)^{-1}X'$  matrice idempotente et  $k$  le nombre de variables indépendantes dans le modèle.

### 2.2.2 Re - spécification des hypothèses du modèle linéaire :

1. la normalité de la distribution conditionnelle :

$$f_{y_i|x_i, y_j} (y_i | X_i = x_i, Y_j = y_j; j \in N(i), \theta_i) \sim N$$

2. normalité des espérances :

$$E\left(Y_i \mid X_i = x_i, Y_j = y_j; j \in N(i); \theta_i\right) = \alpha' \sum_{j=1}^n w_{ij} x_j + \beta' x_i + \rho \sum_{j=1}^n w_{ij} y_j$$

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)' , \sum_{j=1}^n w_{ij} x_j = \left( \sum_{j=1}^n w_{ij} y_{1j}, \sum_{j=1}^n w_{ij} y_{2j}, \dots, \sum_{j=1}^n w_{ij} y_{kj} \right)'$$

$\alpha, \beta$  et  $\rho$  sont des paramètres à estimer. En particulier  $\alpha$  et  $\rho$  sont les paramètres qui régulent la quantité de la dépendance spatiale dans les variables indépendantes et dans la variable dépendante décalée (endogène).

3. homoshédasticité :

$$V\left(Y_i \mid X_i = x_i, Y_j = y_j; j \in N(i), \theta\right) = \sigma^2 \quad \forall x_i, y_i$$

4. les paramètres sont constants dans l'espace:

$$\theta_i(\alpha', \beta', \rho, \sigma^2) = \theta \quad \forall i$$

### 2.2.3 Auto corrélation globale (régression linéaire avec décalage spatial dans la variable dépendante)

Le premier cas de dépendance spatiale qu'on va considéré est le modèle spatial autorégressif. Ce dernier nous permet capter la dépendance spatiale entre les données en ajoutant la variable  $y$  spatialement décalée, comme une variable indépendante, dans le modèle. Ce modèle est référé souvent comme un modèle spatial décalé (Anselin et Berra, 1998), ou modèle spatial autorégressif (SAR) (LeSage, 1999).

Le modèle peut s'écrire de la manière suivante :

$$y_i = \beta' x_i + \rho \sum_{j=1}^n w_{ij} y_j + u_i \quad (19)$$

$\rho$  le paramètre autorégressif,  $f_u(u|X) \sim N(0, \sigma^2 I_n)$  et  $I_n$   $n \times n$  matrice identité.

A la différence de la spécification en séries chronologiques, la présence du terme spatialement décalé parmi les variables explicatives introduit une corrélation entre le terme de l'erreur et la variable décalée (Anselin et Berra, 1998). La méthode des moindres carrée ne fournit pas des estimateurs consistants.

La forme matricielle de l'équation (16) est donnée par la forme :

$$y = X\beta + \rho W y + u \quad (20)$$

$X$  de dimension  $n \times k$

On peut fournir une justification probabiliste de l'équation (20), on utilisant la théorie de champ aléatoire et ainsi, interpréter cette dernière comme une équation différentielle stochastique conduisant à un champ aléatoire simultanément autorégressif (Arbia, 2005).

Le modèle peut être réécrit comme suit :

$$y = \rho W y + \varepsilon \quad (21)$$

$$\varepsilon = X\beta + u, \quad u \sim N(0, \sigma^2 I_n) \quad \text{et} \quad \varepsilon \sim N(X\beta, \sigma^2 I_n)$$

L'équation (21) devient :

$$y = (I - \rho W)^{-1} \varepsilon \quad (22)$$

Les propriétés du processus aléatoire spatial  $y$  sont générés par ce qui suit :

- L'espérance mathématique est donnée par :

$$E(y) = E\left((I - \rho W)^{-1} \varepsilon\right) = (I - \rho W)^{-1} X\beta \quad (23)$$

- La matrice de variance covariance est donnée par (Arbia, 2005) :

$$E(yy') = V(\sigma^2, \rho) = \sigma^2 \left[ (I - \rho W)' (I - \rho W) \right]^{-1} \quad (24)$$

De (23) et (24), on peut avoir la fonction de vraisemblance gaussienne de l'échantillon des observations :

$$L(\sigma^2, \rho, \beta; y) = c(y) |V(\sigma^2, \rho)|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} (y - (I - \rho W)^{-1} X\beta)' V^{-1} (y - (I - \rho W)^{-1} X\beta) \right\} \quad (25)$$

Le log de la vraisemblance est donnée par :

$$l(\sigma^2, \rho, \beta; y) = c(y) - \frac{n}{2} \ln(\sigma^2) + \ln |I - \rho W| - \frac{1}{2\sigma^2} \left[ ((I - \rho W) y - X\beta)' ((I - \rho W) y - X\beta) \right] \quad (26)$$

Sachant que :

$$|V(\sigma^2, \rho)| = \left| \sigma^2 \left[ (I - \rho W)' (I - \rho W) \right]^{-1} \right| = \sigma^{2n} \left| \left[ (I - \rho W)' (I - \rho W) \right]^{-1} \right|$$

$$\left| \left[ (I - \rho W)' (I - \rho W) \right]^{-1} \right| = |(I - \rho W)|^{-2}$$

La structure de cette matrice de variance est telle que chaque localisation (ou observation) est corrélée avec les autres localisations de l'espace étudié. La dépendance est supposée autant plus forte que les localisations sont proches entre elles. L'auto corrélation globale peut être décomposée de la manière suivante :

$$(I - \rho W)^{-1} = I + \rho W + \rho^2 W^2 + \dots$$

### **2.2.4 Auto corrélation locale (régression linéaire avec décalage spatial dans les erreurs) :**

Dans ce type de modèles on suppose la présence d'une dépendance spatiale dans les erreurs. Ce qui remet en cause l'hypothèse classique de bruit blanc. Cette formulation est connue en économétrie sous le nom de «Spatial Error Model ou SEM» (Anselin et Berra, 1998 et Anselin et al., 2004).

Le modèle linéaire est redéfini de la manière suivante :

$$y_i = \beta' x_i + e_i \quad (27)$$

Avec une expression autorégressive simultanée pour les erreurs de la forme :

$$e_i = \rho \sum_{\substack{j=1 \\ j \neq i}}^n w_{ij} e_j + u_i \quad (28)$$

$$u_i \sim N(0, \sigma^2)$$

Une réécriture matricielle compacte donne les formes suivantes :

$$y = X\beta + e \quad (29)$$

$X$  une matrice  $n \times k$

$$e = \lambda W e + u \quad (30)$$

Ce processus est caractérisé par une matrice de variance covariance de la forme (Arbia, 2005) :

$$V(\lambda, \sigma^2) = \left[ (I - \lambda W)' \Sigma^{-1} (I - \lambda W) \right]^{-1} \quad (31)$$

$\Sigma$  est une matrice diagonale de termes génériques :  $\sigma_i^2 = \text{Var}(u_i)$ . Dans le cas homoshédastique  $\sigma_i = \sigma \forall i$ , la matrice  $V$  devient :

$$V(\lambda, \sigma^2) = \sigma^2 \left[ (I - \lambda W)' (I - \lambda W) \right]^{-1} = \sigma^2 (I - \lambda W)^{-1} (I - \lambda W')^{-1} \quad (32)$$

La fonction de vraisemblance est donnée par :

$$L(\sigma^2, \lambda, \beta; e) = c(e) |V(\sigma^2, \lambda)|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} e' V(\sigma^2, \lambda)^{-1} e \right\} \quad (33)$$

La substitution de  $e = y - X\beta$  donne :

$$L(\sigma^2, \lambda, \beta; y, X) = c(y, X) |V(\sigma^2, \lambda)|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (y - X\beta)' V(\sigma^2, \lambda)^{-1} (y - X\beta) \right\} \quad (34)$$

La fonction log de vraisemblance est donnée par :

$$\begin{aligned} l(\sigma^2, \lambda, \beta; y, X) = \\ c(y, X) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2} \ln \left| (I - \rho W)^{-1} (I - \rho W')^{-1} \right| \\ - \frac{1}{2\sigma^2} (y - X\beta)' \left| (I - \rho W)^{-1} (I - \rho W')^{-1} \right|^{-1} (y - X\beta) \end{aligned} \quad (35)$$

### 2.2.5 Méthodes d'estimation :

Comme il déjà été annoncé ci-dessus, l'estimateur des moindres carrés ordinaires n'est pas convergent lorsqu'il y a une variable endogène décalée, et il n'est pas efficace en présence d'auto corrélations spatiales des erreurs. L'introduction d'autres méthodes d'estimation devient nécessaire, tels que la méthode des maximums de vraisemblance, celle des variables instrumentales (IV), ou des moments généralisés (Anselin, 1988).

#### 2.2.5.1 l'inconsistance de la méthode des MCO :

Pour étudier simultanément les propriétés des deux modèles précédents, nous introduisant le modèle spatial général de Anselin (1988) (ou SARAR (1,1)), avec à la fois une variable endogène décalée et une auto-corrélation spatiale dans les erreurs.

$$y = \rho W_1 y + X\beta + e \quad (36)$$

$$e = \lambda W_2 e + u$$

$$u \sim N(0, \sigma^2 I_n)$$

Avec  $V(\rho, \lambda, \sigma^2) = \sigma^2 (I - \rho W_1)^{-1} (I - \lambda W_2)^{-1} (I - \rho W_1')^{-1} (I - \lambda W_2')^{-1}$  matrice de variance covariance.

**Proposition 1 :** *la variable endogène décalée est corrélée avec le terme des erreurs. L'estimateur des MCO n'est pas convergent.*

**Démonstration :**

La convergence des estimateurs du MCO est conditionnée par le fait que la covariance, entre la variable décalée  $W_1 y$  et l'erreur  $e$ , doit être nulle. Le modèle (31) peut être réécrit comme suit :

$$y = (I - \rho W_1)^{-1} X\beta + (I - \rho W_1)^{-1} e \quad (37)$$

Alors :

$$E(W_1 y e') = E(W_1 (I - \rho W_1)^{-1} X\beta e' + W_1 (I - \rho W_1)^{-1} e e') \quad (38)$$

Or comme  $E(W_1 (I - \rho W_1)^{-1} X\beta e') = 0$ , la covariance entre  $W_1 y$  et l'erreur  $e$  est égale à :

$$\begin{aligned} E(W_1 y e') &= W_1 (I - \rho W_1)^{-1} E(e e') \quad (39) \\ &= W_1 (I - \rho W_1)^{-1} (I - \lambda W_2)^{-1} (I - \lambda W_2')^{-1} \sigma^2 \neq 0 \end{aligned}$$

Pour le modèle SAR ( $\lambda = 0$ ), où pour le modèle SEM ( $\rho = 0$ ), l'estimateur des MCO n'est pas convergent à cause de l'endogénéité. L'application des méthodes des moindres carrés quasi généralisés, ou celles des variables instrumentales sont limités pour ce type d'estimation (Le Gallo, 2000).

**2.2.5.2 la méthode du maximum de vraisemblance:**

Partant du modèle général (31), celui-ci peut se réécrire de la manière suivante :

$$(I - \rho W_1) y = X\beta + (I - \lambda W_2)^{-1} u \quad (40)$$

et

$$u = (I - \lambda W_2) [(I - \rho W_1)y - X\beta] \quad (41)$$

La fonction de vraisemblance des variables gaussiennes  $u$  est donnée par :

$$L(\sigma^2, u) = c(u) |V|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} u' V^{-1} u\right) \quad (42)$$

$V$  est la matrice de variance covariance de  $u$ .  
 $V = \sigma^2 I$ ,  $|V| = \sigma^{2n} |I| = \sigma^{2n}$  et  $V^{-1} = \sigma^{-2} I$ . Alors :

$$L(\sigma^2, u) = c(u) (\sigma^{2n})^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2} u'u\right) \quad (43)$$

Le Jacobien du changement de variable de  $u$  à  $y$  dans la transformation (40) donne :

$$J = \left| \frac{\partial u}{\partial y} \right| = |I - \lambda W_2| |I - \rho W_1| \quad (44)$$

L'expression de l'équation (43) en fonction de  $y$  et sa multiplication par le Jacobien donne la vraisemblance de  $y$  suivante :

$$L(\rho, \lambda, \sigma^2, \beta; y) = c(y) (\sigma^{2n})^{\frac{1}{2}} |I - \lambda W_2| |I - \rho W_1| \exp\left[-\frac{1}{2\sigma^2} \left\{ (I - \lambda W_2) [(I - \rho W_1)y - X\beta] \right\}' \left\{ (I - \lambda W_2) [(I - \rho W_1)y - X\beta] \right\} \right] \quad (45)$$

Le log de vraisemblance est donné par l'expression suivante :

$$l(\rho, \lambda, \sigma^2, \beta; y) = c(y) - \frac{n}{2} \ln \sigma^2 + \ln |I - \lambda W_2| + \ln |I - \rho W_1| - \frac{1}{2\sigma^2} \left\{ (I - \lambda W_2) [(I - \rho W_1)y - X\beta] \right\}' \left\{ (I - \lambda W_2) [(I - \rho W_1)y - X\beta] \right\} \quad (46)$$

L'équation (46) ne peut pas être maximisée analytiquement car les paramètres présentent un haut niveau de non linéarité. Par conséquent, nous procédons par le remplacement de ces paramètres (considérés, initialement, comme des paramètres de nuisance) avec des estimations consistantes. Cette méthode de procéder est décrite dans LeSage (1999) et consiste à estimer le vecteur des paramètres  $\beta$  de la forme réduite (47) ci-dessous, avec la méthode des moindres carrés pondéré.

$$(I - \lambda W_2)y = X\beta + \varepsilon \quad (47)$$

On utilise  $\Omega = (I - \rho W_1)' (I - \rho W_1)$  comme la matrice de poids.

$$\hat{\beta} = (X' \Omega X)^{-1} X' \Omega (I - \lambda W_2) y \quad (48)$$

On définit les résidus estimés du modèle comme :

$$\hat{\varepsilon} = (I - \lambda W_2) y - X \hat{\beta} \quad (49)$$

L'estimateur consistant de la variance des bruits blancs est donnée par :

$$\hat{\sigma}^2 = \frac{\hat{\varepsilon}' \hat{\varepsilon}}{n} \quad (50)$$

La substitution des équations (48) et (50) dans l'équation de log de vraisemblance donne la forme suivante :

$$l(\rho, \lambda, y) = c(y) + \ln |I - \lambda W_2| + \ln |I - \rho W_1| - \frac{1}{2\hat{\sigma}^2} \left\{ (I - \lambda W_2) [(I - \rho W_1) y - X \hat{\beta}] \right\}' \left\{ (I - \lambda W_2) [(I - \rho W_1) y - X \hat{\beta}] \right\} \quad (51)$$

L'équation (51) peut être maximisée numériquement pour obtenir des estimations de la pseudo - vraisemblance des paramètres  $\lambda$  et  $\rho$ . La substitution des ces dernières estimations dans les équations (48), (49) et (50) donne les estimations des paramètres  $\sigma^2$  et  $\beta$ .

La technique du maximum de vraisemblance semble largement utilisée dans l'économétrie fréquentielle (El Ouardighi, 2002 ; Baumont et al. 2004). L'introduction des méthodes bayésiennes permet de répondre aux questions concernant le temps de calcul en présence d'un grand nombre d'observations, ainsi que la prise en compte des observations extrêmes, très courantes dans les données spatiales. Dans la section suivante nous présentons les techniques d'estimation bayésiennes, et leurs apports à l'économétrie spatiale.

### 2.3 Méthode d'estimation Bayésienne :

La méthode du maximum de vraisemblance, dominante jusqu'à la fin des années 90 (Anselin et Berra, 98), peut être coûteuse en terme de temps de calcul du Hessien quand les échantillons sont de tailles importantes. De surcroît, dans les modèles d'économétrie spatiale, la normalité des erreurs est souvent remise en cause (LeSage, 1997). L'étude des modèles spatiaux, notamment ceux de l'économie géographique, nécessite l'introduction de techniques d'estimations aptes à prendre en compte des phénomènes complexes tels que : les variations des paramètres dans l'espace, l'asymétrie, l'incertitude dans les mesures de distance ou de

proximité, etc. Les méthodes bayésiennes nous permettent de tenir compte à la fois des problèmes d'hétérogénéités, des observations aberrantes et de mieux considérer les fondements économiques dans la mesure des interactions spatiales.

Dans ce chapitre nous exposerons les méthodes de simulations MCMC (Monte Carlo Markov Chain) qui permettent d'estimer les modèles Bayésiens.

### 2.3.1 Les principes de l'économétrie Bayésienne :

L'économétrie Bayésienne est devenue de plus en plus attractive avec l'apparition de l'ouvrage fondateur de Zellner (1971). Depuis, de nombreux travaux spécialisés ont été publiés tels que celui de Bauwens et al. (1995) sur les séries chronologiques, Poirier (1995) sur les aspects théoriques de l'économétrie Bayésienne, Koop (2003) présente un cadre général de l'économétrie Bayésienne et LeSage (1997) introduit ces méthodes à l'économétrie spatiale.

Les modèles économétriques peuvent être estimés selon une approche fréquentiste (statistique) ou probabiliste (Bayésienne) (Robert, 2006). Comparé à la modélisation probabiliste, l'analyse statistique est une méthode qui permet, à partir de l'observation d'un phénomène aléatoire contrôlé par le paramètre  $\theta$ , de déduire de ces observations une inférence (un résumé, une caractérisation) sur  $\theta$ . A l'inverse, la modélisation probabiliste caractérise le comportement des observations futures conditionnellement à  $\theta$ .

Cette inversion de probabilité est décrite par le théorème de Bayes : Si A et B sont des événements tels que  $P(E) \neq 0$ ,  $P(A|E)$  et  $P(E|A)$  sont reliés par :

$$P(A|B) = \frac{P(E|A)P(A)}{P(E|A)P(A) + P(E|A^c)P(A^c)} = \frac{P(E|A)P(A)}{P(E)}$$

En particulier :

$$\text{Si } P(A) = P(B) \text{ alors } \frac{P(A|E)}{P(B|E)} = \frac{P(E|A)}{P(E|B)} \quad (52)$$

L'équation (52) exprime le fait fondamental que, pour deux causes équiprobables, le rapport de probabilités pour un effet donné est égal au rapport de probabilités de ces deux causes. Ce théorème décrit aussi la vraisemblance de A de  $P(A)$  vers  $P(A|E)$ , une fois E observé. La version continue de ces résultats est donnée par Bayes (1763), à savoir, que pour

deux variables aléatoires  $x$  et  $y$ , de distributions conditionnelle  $f(x|y)$  et marginale  $g(y)$ , la distribution conditionnelle de  $y$  sachant  $x$  est :

$$g(y|x) = \frac{f(x|y)g(y)}{\int f(x|y)g(y)dy}$$

Bien que ce résultat d'inversion soit naturel d'un point de vu probabiliste, Bayes et Laplace sont allés plus loin pour considérer que l'incertitude sur le paramètre  $\theta$  d'un modèle peut être décrite par une distribution de probabilité  $\pi$  sur  $\Theta$ , appelée distribution à priori. L'inférence est fondée sur la distribution de  $\theta$  conditionnelle à  $x$ ,  $\pi(\theta|x)$ , appelée distribution à posteriori est donnée par :

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta)d\theta} \quad (53)$$

$\pi(\theta|x)$  est proportionnelle à la distribution de  $x$  conditionnelle à  $\theta$  (distribution à posteriori), qui est aussi la vraisemblance multipliée par la distribution à priori de  $\theta$ . L'apport principal de l'approche Bayésienne est de considérer en plus une distribution aléatoire des paramètres.

Ainsi, un modèle statistique bayésien est composé d'un modèle statistique paramétrique,  $f(x|\theta)$ , et d'une distribution à priori pour les paramètres,  $\pi(\theta)$ . Le théorème de Bayes actualise l'information sur  $\theta$  en extrayant l'information contenue dans l'observation  $x$ , il permet de mettre « causes et effets » sur le même niveau conceptuel, puisque les deux sont aléatoires.

Le développement de l'outil informatique et des techniques de simulation ont eu de grands effets sur le développement de l'application des méthodes bayésiennes. En effet, lorsque le calcul analytique des distributions à posteriori  $\pi(\theta|y)$  est difficile à faire, on peut faire recours à des méthodes de simulation pour générer des échantillons de variables aléatoires suivant la distribution  $\pi(\theta|y)$ . Les méthodes de simulations font appel à des algorithmes stochastiques faisant intervenir des procédures itératives d'échantillonnage. L'échantillonnage des composant du paramètre  $\theta$  permet de créer une suite dépendante de valeurs de ces composantes, possédant des propriétés Markoviennes. Les algorithmes les plus utilisés appartiennent à la famille des MCMC, il s'agit de l'échantillonneur de Gibbs (Tanner

et Wong, 1987), de l'algorithme de Metropolis (Metropolis et al., 1953), et de son extension par Hastings (Hastings, 1970).

Lorsque les lois utilisées appartiennent à la famille exponentielle nous pouvons avoir recours aux lois dites conjuguées : si  $\pi(y|\theta)$  appartiennent à la famille exponentielle alors il existe une loi conjuguée naturelle pour  $\pi(\theta)$  et  $\pi(\theta|y)$ . Des exemples avec démonstration sont illustrés dans les ouvrages de Berger (1985), Robert (1992), et Bernardo et Smith (1994).

Sans conteste, le point crucial et le plus critiqué dans l'analyse bayésienne est le choix de la distribution à priori. Ce dernier a un impact considérable sur l'inférence de la loi à posteriori. Il existe plusieurs manières, mais souvent arbitraires, de choisir une loi à priori. En particulier, l'utilisation systématique de lois usuelles (normale, gamma, bêta, etc) et des lois conjuguées n'est pas toujours justifiée. La détermination subjective de la loi a priori qui en résulte ignore une partie de l'information a priori.

Les critiques contre l'approche bayésienne ont une certaine validité puisqu'elles attirent l'attention sur le fait qu'il n'y a pas une façon unique de choisir une loi a priori, et que ce choix a un impact sur l'inférence résultante. Le travail de Jeffreys (1946), sur les a priori non informatives, propose une méthode de construction de la loi a priori directement déduite de la distribution des observations. Plus récemment, les avancées en terme de robustesse et d'analyse de sensibilité ont fourni une base solide pour l'analyse bayésienne dans la situation d'information a priori incomplète. L'introduction de la modélisation hiérarchique permet de classer la sélection d'un a priori avec une diminution notable de son impact sur l'inférence résultante.

### **2.3.2 Le choix de distributions a priori**

L'approche bayésienne se focalise sur la considération de la distribution a posteriori, proportionnelle au produit de la vraisemblance et de la densité a priori des paramètres. Cependant, le choix de cette dernière est sujet à l'expérience de l'utilisateur et n'est pas forcément unique. Lorsque l'information a priori sur les paramètres est vague, on peut considérer des lois a priori conjuguées comme point de départ pour élaborer des distributions a priori fondées sur une information a priori limitée.

Dans ce qui suivra dans nos applications, nous utiliserons des distributions uniformes et conjuguées, ce qui permet d'affecter les mêmes probabilités à toutes les valeurs possibles

de la variable. Ce choix doit être considéré avec précaution, surtout si les distributions à priori sont impropres, et justifié à partir de la théorie économique.

**Définition 13 :** *une famille  $\wp$  de distribution de probabilité sur  $\Theta$  est dite conjuguée (ou fermée sur échantillonnage) par une fonction de vraisemblance  $f(x|\theta)$  si, pour tout  $\pi \in \wp$ , la distribution a posteriori  $\pi(\theta|x) \in \wp$*

L'approche a priori conjuguée, introduite par Raiffa et Schlaifer (1961) peut être justifiée par un raisonnement d'invariance. En effet, quand l'observation  $x \sim f(x|\theta)$  modifie  $\pi(\theta)$  en  $\pi(\theta|x)$ , l'information transmise par  $x$  sur  $\theta$  est limitée, et elle ne devrait pas modifier toute la structure de  $\pi(\theta)$ , mais simplement celle de ces paramètres. Finetti (1974) considère que l'information a priori pouvait être interprétée comme des observations passées virtuelles, ce qui mène forcément à des lois a priori conjuguées pour des familles exponentielles.

Une vraisemblance de type exponentielle est donnée par :

$$f(x|\theta) = h(x) \exp(\theta x - \phi(\theta)) \quad (54)$$

L'équation (54) est une loi générique d'une famille exponentielle. Cette loi admet une famille conjuguée. La famille conjuguée pour  $f(x|\theta)$  est donnée par :

$$\pi(\theta|\mu, \lambda) = K(\mu, \lambda) \exp(\theta\mu - \lambda\phi(\theta)) \quad (55)$$

$K(\mu, \lambda)$  représente la constante de normalisation de la densité. La loi à posteriori correspondante est donnée par  $\pi(\theta|\mu+x, \lambda+1)$ .

Cette approche est justifiée dans la pratique. De nombreuses distributions usuelles sont de type exponentiel. Notons le choix artificiel de ce type de choix pour prendre en compte l'information a priori. Cependant, la flexibilité qu'offre cette approche permet d'obtenir une variance importante pour des densités a priori.

### 2.3.3 Estimateurs Bayésiens et fonction de coût :

Dans le sens classique, un estimateur est une fonction des observations (une statistique). Il dépend des paramètres de la fonction de distribution.

Soit  $L(\theta, \delta) \geq 0$  une fonction qui mesure le coût du choix de l'estimateur  $\delta = \delta(x)$  du paramètre  $\theta$ . Le coût est minimum quand  $\delta = \theta$ . La fonction du coût  $L$  représente le coût (l'erreur) dû à une mauvaise évaluation de  $\theta$ . L'approche Bayésienne de la théorie de la décision est fondée sur *le coût moyenne à posteriori* donné par la forme suivante :

$$E[L(\theta, \delta)|x] = \int_{\Theta} L(\theta, \delta) \pi(\theta|x) d\theta$$

$\theta \in \Theta$  l'espace des paramètres. Ce coût moyenne l'erreur (c'est-à-dire le coût) selon la distribution à priori de  $\theta$  conditionnellement à la valeur observée  $x$ . Le coût moyenne à posteriori est une fonction de  $x$ .

**Définition 14 :** un estimateur de Bayes associé à une distribution a priori  $\pi$  et une fonction de coût  $L$  est un estimateur qui minimise la fonction de coût moyenne a posteriori donné par :

$$\delta^*(x) = \arg \min_{\theta \in \Theta} E[L(\theta, \delta)|x]$$

Les fonctions de coût usuelles utilisées en analyse bayésienne sont :

1. le coût quadratique : introduit par Legendre (1805) et Gauss (1810)

$$L(\theta, \delta) = (\theta - \delta)^2 \quad (56)$$

l'estimateur de Bayes  $\delta$  associé à la loi a priori  $\pi$  et au coût quadratique (56) est la *moyenne a posteriori*  $\pi^* = E(\theta|x)$

2. L'erreur de coût absolu : est une solution alternative au coût quadratique en dimension un.

$$L(\theta, \delta) = |\theta - \delta|$$

où plus généralement une fonction linéaire par morceaux

$$L_{k_1, k_2}(\theta, \delta) = \begin{cases} k_2(\theta - \delta) & \text{si } \theta > \delta \\ k_1(\delta - \theta) & \text{si non.} \end{cases} \quad (57)$$

l'estimateur de Bayes  $\delta$  associé à la loi a priori  $\pi$  et à la fonction linéaire par morceaux (57) est le fractile  $k_1/(k_1 + k_2)$  de  $\pi(\theta|x)$ . Dans le cas du coût absolu ( $k_1 = k_2$ ), l'estimateur de Bayes est la médiane a posteriori.

3. Le coût 0 – 1 : pour ce coût la pénalité associée à un estimateur  $\delta$  est 0 si la réponse est correcte et 1 sinon.

$$L(\theta, \delta) = \begin{cases} (1 - \delta) & \text{si } \theta \in \Theta \\ \delta & \text{si non.} \end{cases} \quad (58)$$

l'estimateur de Bayes  $\delta$  associé à la loi a priori  $\pi$  et au coût (58) est donnée par :

$$\delta^*(x) = \begin{cases} 1 & \text{si } P(\theta \in \Theta|x) \succ P(\theta \notin \Theta|x) \\ 0 & \text{si non} \end{cases}$$

$\delta^*(x)$  vaut 1 si et seulement si  $P(\theta \in \Theta|x) \succ 1/2$

### 2.3.4 Méthodes de calcul des estimateurs

La simplicité de l'approche bayésienne réside dans le fait que pour une fonction de coût  $L$  et une loi a priori  $\pi$ , l'estimateur bayésien associé à une observation  $x$  est la décision  $\delta$  minimisant le coût a posteriori :

$$L(\pi, \delta|x) = \int_{\Theta} L(\theta, \delta) \pi(\theta|x) d\theta \quad (59)$$

Dans la pratique, la minimisation de (59) soulève deux problèmes : le calcul de la loi a posteriori  $\pi(\theta|x)$  peut être impossible même si  $\pi(\theta|x)$  est connue car la minimisation de (59) peut nécessiter un temps de calcul long surtout si  $\Theta$  et l'ensemble des décisions sont de grandes dimensions. Une réponse à ces difficultés est apportée par les méthodes d'approximations utilisées pour simuler les lois a posteriori qui n'admettent pas d'expression analytique. Cette approche permet alors de pallier les problèmes d'intégration et d'optimisation rencontrés dans la construction des estimateurs bayésien. Robert et Casella (2004) proposent une revue approfondie de ces techniques. Andrieu et al. (2004) mettent en avant les avantages procurés par l'utilisation de ces méthodes.

### 2.3.5 Méthode de Monte Carlo par chaîne de Markov

Dans un problème statistique, l'approximation de l'intégrale

$$\int_{\Theta} g(\theta) f(x|\theta) \pi(\theta) d\theta \quad (60)$$

doit tirer avantage du fait que  $\pi(\theta)$  soit une densité de probabilité (loi a priori propre), ou que  $f(x|\theta)\pi(\theta)$  soit proportionnelle à une densité. En effet, s'il est possible de produire des variables aléatoires  $\theta_1, \dots, \theta_m$  de loi  $\pi(\theta)$ , la moyenne  $\frac{1}{m} \sum_{i=1}^m g(\theta_i) f(x|\theta)$  converge presque sûrement vers (60) lorsque  $m$  tends vers l'infini. La mise en oeuvre de cette méthode consiste à générer un échantillon iid d'une loi uniforme  $U([0,1])$  puis transformer en variable de la loi d'intérêt suivant la méthode générale suivante :

Si  $U \sim U_{[0,1]}$ , alors la variable aléatoire  $F^{-1}(u)$  a pour distribution  $F$ .  $F$  une fonction sur  $R$ . l'inverse généralisée de  $F$ ,  $F^{-1}$ , est définie par :

$$F^{-1}(u) = \inf \{x; F(x) \geq u\}$$

pour générer une variable aléatoire  $X \sim F$ , il suffit de générer  $U \sim U_{[0,1]}$  et puis faire la transformation  $x = F^{-1}(u)$

La méthode MCMC permet d'approcher la génération de variables aléatoires d'une loi a posteriori  $\pi(\theta|x)$  quand celle-ci ne peut pas être simulée directement. La méthode des MCMC tire son nom de l'idée que, pour produire des approximations acceptables d'intégrales et d'autres fonctions dépendantes d'une loi d'intérêt, il suffit de générer une chaîne de Markov  $(\theta^m)_m$  de loi limite la loi d'intérêt. Nous présenterons ici les propriétés garantissant la convergence des chaînes de Markov vers leur distribution d'équilibre : réversibilité, apériodicité, irréductibilité, récurrence pour que le comportement de simulation de la chaîne de Markov soit robuste aux conditions initiales.

**Définition 15 :** (Chaîne de Markov) un processus aléatoire  $X(\mathfrak{N}, \mathcal{B}(\mathfrak{N}))$  est une chaîne de Markov, si pour tout ensemble  $\mathfrak{N}$  mesurable  $\in \mathcal{B}(\mathfrak{N})$ ,

$$P(X_n \in \mathfrak{N} | X_0, \dots, X_{n-1}) = P(X_n \in \mathfrak{N} | X_{n-1}) \quad (62)$$

la chaîne est définie suivant son noyau de transition.

**Définition 16 :** (Noyau de transition) une probabilité de transition sur  $(\mathfrak{N}, \mathcal{B}(\mathfrak{N}))$  est une application  $K$  de  $\mathfrak{N} \times \mathcal{B}(\mathfrak{N})$  dans  $[0,1]$  telle que :

- $\forall x \in \mathfrak{N}, K(x, \cdot)$  est une probabilité mesurable

-  $\forall A \in \mathcal{B}(\mathbb{S}), K(\cdot, A)$  est mesurable

Notons que  $K(x, dy)$  la probabilité de transition sur  $\mathcal{B}(\mathbb{S})$  étant donné  $x$  sur  $\mathbb{S}$ . Dans le cas où  $\mathbb{S}$  est un ensemble discret (fini ou dénombrable), le noyau de transition est une matrice  $\mathcal{K}$  définie par les éléments :

$$P_{xy} = P(X_n = y | X_{n-1} = x), \quad x, y \in \mathbb{S} \quad (63)$$

puisque  $\forall x \in \mathbb{S}, K(x, \cdot)$  est une probabilité, on doit avoir

$$P_{xy} \geq 0 \text{ et } K(x, A) = \sum_{y \in \mathbb{S}} P_{xy} = 1 \quad (64)$$

La matrice  $\mathcal{K}$  est une matrice de transition de Markov (ou matrice de transition). Soit  $K$  une probabilité de transition, une séquence de variable aléatoire  $X_0, X_1, \dots, X_n, \dots$  est une chaîne de Markov notée  $(X_n)_{n \geq 0}$  est homogène de probabilité initiale  $u_0$  et de transition  $K$ , si  $\forall A \in \mathcal{B}(\mathbb{S}), K(X_n, A) = P(X_{k+1} \in A | x_0, x_1, \dots, x_k)$  satisfait  $u_0(A) = \int_{\mathbb{S}} u_0(dx) K(x, A)$ .

Cette probabilité de transition nous permet de générer itérativement une séquence  $(x_n)_{n \geq 0}$ , correspondant à la réalisation d'une chaîne de Markov  $(X_n)_{n \geq 0}$  telle que, pour  $n$  très grand, la distribution de  $x_n$  est très proche de  $u_0$  quelque soit le choix de l'état initial  $x_0$ .

L'algorithme de simulation est de la forme suivante :

**Algorithme 1 : Simulation par chaîne de Markov**

1. *initialiser*  $x_0 \sim u_0, l = 1$
2. *itération*  $l : x_l \sim K(x_{l-1}, x_l)$
3.  $l \leftarrow l + 1$  et aller en (2)

L'échantillon  $x_l$  est simulé en fonction de l'échantillon  $x_{l-1}$ , la séquence ainsi produite est une chaîne de Markov. Le noyau  $K$  de la chaîne de Markov utilisé dans l'algorithme doit être construit de façon à ce que les échantillons  $x_l$  soient asymptotiquement distribués suivant la distribution invariante d'intérêt.

La loi  $u_0$  de  $X_0$  étant la distribution initiale, la probabilité conditionnelle de  $X_n | X_0$  s'écrit :

$$P(X_n \in A | X_0 = x_0) = K^n(x_0, A) \quad (65)$$

avec  $K^n$  la  $n^{\text{ième}}$  application du noyau  $K$ , telle que :

$$K^n(x_0, A) = \int_{\mathfrak{S}} K^{n-1}(y, A) K(x_0, y) dy \quad (66)$$

Il s'agira d'établir les propriétés de convergence des algorithmes de simulation des chaînes de Markov vers la distribution d'équilibre :

$$\lim_{x \rightarrow \infty} K^n(x, A) = u(A) \quad (67)$$

pour tout les ensembles mesurables  $A$  et pour tout  $x$ .

### 2.3.5.1- Irréductibilité

Est une mesure de la sensibilité de la chaîne de Markov aux conditions initiales. Elle garantie la convergence des algorithmes MCMC. Dans le cas discret, une chaîne est irréductible si tous les états communiquent entre eux :

$$P_x(\tau_y < \infty) > 0, \quad \forall x, y \in \mathfrak{S} \quad (68)$$

$\tau_y$  temps d'arrêt en  $y$ . la probabilité de passage d'un état  $x$  vers état  $y$  en un nombre fini d'itérations est alors non nulle.  $P_x(\cdot)$  est la probabilité conditionnelle de la chaîne en partant de  $x$ .

**Définition 17 (Irréductibilité pour une chaîne continue):** Une chaîne est  $\varphi$ -irréductible pour une mesure  $\varphi$  et  $n \geq 1$ , si pour tous points  $x \in \mathfrak{S}$  et pour chaque  $A \in \mathcal{B}(\mathfrak{S})$  avec  $\varphi(A) > 0$  :  $K^n(x, A) > 0$

La condition de  $\varphi$ -irréductible nous assure que pour des ensembles suffisamment grands mesurés par  $\varphi$ , il est possible des les atteindre à partir de n'importe quel point en un temps fini. L'irréductibilité garantie que chaque ensemble  $A$  a une chance d'être visiter par la chaîne de Markov, mais n'assure pas une fréquence suffisante de passage des trajectoires  $(x_n)_{n \geq 0}$  afin de pouvoir en estimer la probabilité. Pour garantir une approximation acceptable du modèle simulé la propriété suivante est importante.

### 2.3.5.2 Récurrence et transition

Dans un espace d'états fini, le nombre moyen de passage dans un état  $\omega$  est défini par  $\eta_\omega = \sum_{i=1}^{\infty} H_\omega(X_i)$ . L'état est dit récurrent si  $IE_\omega(\eta_\omega) = \infty$  et transitoire si  $IE_\omega(\eta_\omega) < \infty$ . La stabilité de la chaîne  $(X_n)_{n \geq 0}$  est renforcée si le nombre de passage est infini pour toute la trajectoire. Dans ce cas la chaîne est dite récurrente au sens de Harris.

**Définition 18 : (récurrence au sens de Harris)** un ensemble  $A$  est dit récurrent au sens de Harris si :

$$P_x(\eta_A = \infty) = 1 \quad \forall x \in A \quad (69)$$

la chaîne  $(X_n)_{n \geq 0}$  est  $\varphi$ -Harris récurrente si elle est :  $\varphi$ -irréductible et pour tout ensemble  $A$  (avec  $\varphi(A) > 0$ ),  $A$  est récurrente au sens de Harris. Il faut que :

$$P_x(\eta_A = \infty) = 1 \Rightarrow IE(\eta_A) = \infty \quad (70)$$

### 2.3.5.3 Mesures invariantes :

Une chaîne  $(X_n)$  est stable si la distribution marginale de  $(X_n)$  est indépendante de  $n$ . cette situation nécessite l'existence d'une distribution de probabilité  $\pi$  tel que :  $X_{n+1} \sim \pi$  si  $X_n \sim \pi$

**Définition 19 :** Une mesure  $\pi$  est invariante pour un noyau  $K(.,.)$  si :

$$\pi(A) = \int_{\mathfrak{S}} K(x, A) \pi(dx), \quad \forall A \in \mathcal{B}(\mathfrak{S}) \quad (71)$$

Une chaîne est positive récurrente si  $\pi$  est une mesure de probabilité. Si  $\pi$  est une mesure de probabilité,  $\pi$  est aussi appelée une distribution stationnaire puisque :  $X_0 \sim \pi \Rightarrow X_n \sim \pi \quad \forall n$ . Dans ce cas la stationnarité de la distribution est unique. La propriété de la mesure de probabilité invariante est importante non seulement parce qu'elle permet de définir un processus stationnaire, mais aussi parce que c'est une mesure qui définit le comportement de long terme (ergodique) de la chaîne de Markov. Tierney (1994) montre que si une chaîne a une distribution invariante  $\pi$  et si de plus, elle est irréductible et apériodique, alors  $\pi$  est unique et elle est la distribution d'équilibre de la chaîne

### 2.3.5.6 Ergodicité et convergence:

La notion d'ergodicité permet de garantir la convergence de la chaîne vers sa distribution d'équilibre indépendamment des conditions initiales. La distribution invariante  $\pi$  est le candidat naturel de la distribution limite. Dans le cas discret, un état  $\omega$  est ergodique si :

$$\lim_{x \rightarrow \infty} |K^n(\omega, \omega) - \pi(\omega)| = 0 \quad (72)$$

La convergence est établie en utilisant la norme de la variation total et en cherchant à ce que  $\left\| \int K^n(x, \cdot) u dx - \pi \right\|_{VT} = \sup \left| \int K^n(x, A) u dx - \pi(A) \right|$  soit la plus petite possible.

**Théorème 1 :** *Si la chaîne  $(X_n)$  est récurrente positive au sens de Harris et apériodique, alors  $\lim_{x \rightarrow \infty} \left\| \int K^n(x, \cdot) u dx - \pi \right\|_{VT} = 0$  quelque soit la distribution initiale  $u$*

une chaîne récurrente positive au sens de Harris et apériodique est équivalente à ergodique (Meyn et Tweedie, 1993). La convergence dans la variation totale implique que :  $\lim_{x \rightarrow \infty} |IE_u[h(X_n)] - IE^\pi[h(X)]| = 0$  pour chaque fonction bornée  $h$ .

**Théorème 2 Ergodicité** *Si une chaîne de Markov  $(X_n)$  est récurrente au sens de Harris, alors pour toute fonction  $h$  avec  $E(h) < \infty$ ,*

$$\lim_{x \rightarrow \infty} \frac{1}{n} \sum_i h(X_i) = \int h(x) d\pi(x) \quad (73)$$

La démonstration de ce théorème est réalisée par le recours au théorème central limite. Ce dernier permet d'avoir une bonne approximation de  $\int g(x) \pi(x) dx$  par la moyenne

$$\bar{g}_n = \frac{1}{n} \sum_i g(X_i) \text{ si :}$$

$$\sqrt{n} (\bar{g}_n - IE_\pi[g(X)]) \xrightarrow{d} N(0, \gamma_g^2) \quad (74)$$

Le théorème central limite permet de contrôler la convergence des algorithmes de Monte Carlo par chaîne de Markov au sens où, lorsque  $\gamma_g^2 > 0$ , on peut vérifier si les moyennes  $\bar{g}_n$  convergent vers la quantité  $IE_\pi(g)$ .

### 2.3.6 Algorithme de Metropolis – Hastings

Il n'est pas nécessaire d'utiliser un échantillon iid distribué exactement selon la fonction  $\pi$  pour faire une approximation de l'intégrale :

$$I = \int h(x)\pi(x)dx \quad (75)$$

L'utilisation d'une chaîne de Markov ergodique de loi stationnaire  $\pi$  nous permet d'obtenir une chaîne  $X_1, \dots, X_n \sim$  une approximation de  $\pi$  pour approcher la quantité (75). La mise en oeuvre de ce principe nécessite la construction d'un mécanisme de génération pour produire de telles chaînes de Markov. L'algorithme de Metropolis et al.(1953) et Hastings (1970) satisfait cette contrainte.

Pour une valeur arbitraire initiale  $x^{(0)}$ , une chaîne ergodique  $(X^{(t)})$  est générée en utilisant un noyau de transition (en deux étapes) avec une distribution stationnaire  $\pi$ . Pour des valeurs assez grandes de  $T_0$ ,  $X^{(T_0)}$  peut être considérée comme étant distribuée suivant  $\pi$ .

L'algorithme de Metropolis - Hastings utilise la densité objective (loi cible ou loi objective)  $\pi$  avec une densité conditionnelle  $q(y|x)$  appelée la distribution instrumentale (ou loi de proposition).

#### Algorithme 2 (Metropolis - Hastings) :

1. initialiser avec une valeur arbitraire  $x^{(0)}$ ,  $t = 0$

2. itération  $t$  : générer  $Y_t \sim q(y|x^{(t)})$

accepter  $Y_t$  avec la probabilité  $\rho(x, y) = \min \left\{ \frac{\pi(y) q(x|y)}{\pi(x) q(y|x)}, 1 \right\}$

simuler  $u \sim U_{[0,1]}$

si  $u \leq \rho(x, y)$ , alors prendre  $x^{(t+1)} = Y_t$ , sinon poser  $x^{(t+1)} = x^{(t)}$

$t \leftarrow t + 1$ , allez en (2)

Le schéma de simulation est d'autant plus efficace que la distribution  $q(y|x^{(t)})$  est simulable rapidement. Les propriétés de convergence sont présentées par :

1. La chaîne de Markov de Metropolis – Hastings est réversible, de densité  $\pi$  invariante/stationnaire puisqu'elle satisfait la condition suivante :

$$\pi(y)K(y,x) = \pi(x)K(x,y)$$

2. comme  $\pi$  est une mesure de probabilité, la chaîne est récurrente positive.

$$3. \quad \text{si } P \left[ \frac{\pi(Y_t)q(X^{(t)}|Y_t)}{\pi(X^{(t)})q(Y_t|X^{(t)})} \geq 1 \right] < 1 \quad (76),$$

l'événement  $\{X^{(t+1)} = X^{(t)}\}$  est possible, alors la chaîne est apériodique.

$$4. \quad \text{si } q(y|x) > 0 \quad \forall (x,y) \quad (77)$$

la chaîne est irréductible

5. Pour Metropolis – Hastings,  $\pi$ -irréductibilité implique récurrence au sens de Harris

6. alors, pour Metropolis – Hastings qui satisfait (76) et (77)

$$(i) - \text{pour } h, \text{ avec } \mathbb{E}_f |h(X)| < \infty, \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T h(X^{(t)}) = \int h(x) d\pi(x)$$

$$(ii) - \text{et } \lim_{n \rightarrow \infty} \left\| \int K^n(x, \cdot) u(dx) - \pi \right\|_{VT} = 0 \text{ pour chaque distribution initiale } u, \text{ où}$$

$K^n(x, \cdot)$  représente le noyau pour  $n$  transitions.

Une propriété stupéfiante de cet algorithme est d'autoriser un nombre infini de lois de propositions en produisant toutes chaînes de Markov convergeant vers la loi d'intérêt. (Robert, 2006). Notons que le schéma de simulation M-H ne fonctionne pas bien si les zones de forte probabilités de  $q$  sont localisées surtout dans les queues de distribution de  $\pi$ . Ainsi, pour une densité  $\pi$  fixée, le choix de la distribution candidate  $q$  avec support étroit freine la convergence de l'algorithme, et le choix d'une candidate avec support très étendu crée un gaspillage de simulations en dehors du support  $\pi$ .

L'algorithme M-H peut présenter des difficultés à explorer convenablement le support de la distribution d'intérêt. Cette situation s'aggrave dans le cas de distributions multidimensionnelles, le nombre d'échantillons nécessaire pour obtenir une couverture suffisante du support de la distribution d'intérêt devient très important. L'échantillonnage de Gibbs est plus adapté à la simulation de distributions.

### 2.3.7 Echantillonnage de Gibbs :

D'un point de vue général, l'échantillonnage de Gibbs tire profit des structures hiérarchiques d'un modèle. Lorsqu'un modèle s'écrit de la forme :

$$\pi(\theta|x) = \int \pi_1(\theta|x, \lambda) \pi_2(\lambda|x) d\lambda \quad (78)$$

Lorsque les deux lois  $\pi_1(\theta|x, \lambda)$  et  $\pi_2(\lambda|x)$  sont connus et peuvent être simulées, la génération de  $\theta$  de  $\pi(\theta|x)$  est équivalente à la génération de  $\lambda$  de  $\pi_2(\lambda|x)$ , puis de  $\theta$  de  $\pi_1(\theta|x, \lambda)$ .

Une première technique d'échantillonnage de Gibbs (appelée augmentation de données) a été introduite par Tanner et Wong (1987) pour tirer profit des lois conditionnelles selon l'algorithme suivant :

#### Algorithme 3: Echantillonnage de Gibbs bivarié

*Initialisation : Commencer par une valeur arbitraire  $\lambda^{(0)}$ .*

*Itération  $t$  : pour  $\lambda^{(t-1)}$  donné, générer*

*a.  $\theta^{(t)}$  selon  $\pi_1(\theta|x, \lambda^{(t-1)})$*

*b.  $\lambda^{(t)}$  selon  $\pi_2(\lambda|x, \theta^{(t)})$ .*

**Lemme 1 :** Si  $\pi_1(\theta|x, \lambda) > 0$  sur  $\Theta$  ( $\pi_2(\lambda|x, \theta) > 0$  sur  $\Lambda$  respectivement), les deux suites  $(\theta^{(m)})$  et  $(\lambda^{(m)})$  sont des chaînes de Markov ergodiques de lois univariantes  $\pi(\theta|x)$  et  $\pi(\lambda|x)$  respectivement.

La généralisation de l'algorithme (3) de Gibbs bivarié consiste à considérer plusieurs groupes de paramètres,  $\theta, \lambda_1, \dots, \lambda_p$ , tels que :

$$\pi(\theta|x) = \int \dots \int \pi(\theta, \lambda_1, \dots, \lambda_p|x) d\lambda_1 \dots d\lambda_p \quad (79)$$

Cette généralisation correspond à l'introduction de niveaux additionnels dans le modèle hiérarchique (78), pour des raisons de simulation, ou de modélisation de l'hyperparamètre  $\lambda$  ou du paramètre  $\theta$  en des composantes de plus petites dimensions. L'échantillonnage de Gibbs fournit des simulations de la loi jointe  $\pi(\theta, \lambda_1, \dots, \lambda_p|x)$ , lorsque

certaines des lois additionnelles associées à  $\pi$  sont disponibles . Lorsque  $\pi(\theta|x)$  se décompose elle-même en lois conditionnelles, il n' y a pas besoin d'introduire des paramètres additionnelles  $\lambda_i (1 \leq i \leq p)$  . Dans ce cadre, la mise en oeuvre de l'échantillonnage de Gibbs peut être faite de plusieurs façons. Si la décomposition de  $(\theta, \lambda)$  en  $(\theta, \lambda_1, \dots, \lambda_p)$  correspond à la décomposition du modèle selon ses niveaux hiérarchiques :

$$\pi(\theta|x) = \int \dots \int \pi_1(\theta|\lambda_1, x) \pi_2(\lambda_1|\lambda_2) \dots \pi_{p+1}(\lambda_p) d\lambda_1 \dots d\lambda_p \quad (80)$$

Il semble logique de simuler selon des lois conditionnelles (Robert, 2006)

$$\begin{aligned} \pi(\theta, \lambda_1, \dots, \lambda_p) &= \pi_1(\theta|\lambda_1, x) \\ \pi(\lambda_i | x, \theta, (\lambda_j)_{j \neq i}) &= \pi(\lambda_i | \lambda_{i-1}, \lambda_{i+1}) \quad (1 < i < p) \quad (81) \\ \pi(\lambda_1 | x, \theta, (\lambda_j)_{j \neq 1}) &= \pi(\lambda_1 | \theta, \lambda_2) \\ \pi(\lambda_p | x, \theta, (\lambda_j)_{j \neq p}) &= \pi(\lambda_p | \lambda_{p-1}) \end{aligned}$$

Quelles que soient les dimensions de  $\theta$  et  $\lambda_j$  .

Un algorithme alternatif proposé par Gelfand et Smith (1990) est l'échantillonneur de Gibbs direction par direction, qui ne prend pas en compte les divisions hiérarchiques et ne considère que les paramètres unidimensionnels, afin de générer conditionnellement aux autres paramètres.

La formulation générale de l'algorithme d'échantillonnage de Gibbs pour une loi jointe  $\pi(\theta_1 \dots \theta_p)$ , de lois conditionnelles complètes  $\pi_1, \dots, \pi_p$  est donnée par :

#### **Algorithme (4) Echantillonnage de Gibbs**

Pour  $(\theta_1^{(t)} \dots \theta_p^{(t)})$  donnés, simuler

1.  $\theta_1^{(t+1)} \sim \pi_1(\theta_1 | \theta_2^{(t)}, \dots, \theta_p^{(t)})$ ,
2.  $\theta_2^{(t+1)} \sim \pi_2(\theta_2 | \theta_1^{(t+1)}, \theta_3^{(t)}, \dots, \theta_p^{(t)})$ ,
- .
- .
- .

$$p. \theta_p^{(t+1)} \sim \pi_p \left( \theta_p \mid \theta_1^{(t+1)}, \dots, \theta_{p-1}^{(t+1)} \right).$$

La loi jointe de  $\pi$  est stationnaire à chaque étape de cet algorithme, car les  $\pi_j$  sont des lois conditionnelles complètes de  $\pi$ . L'échantillonnage de Gibbs n'exige pas que les  $\theta_i$  soient unidimensionnelles et le choix de la décomposition peut être entièrement fondé sur des raisons de simulation. En pratique, chaque fois que la simulation d'une loi conditionnelle donnée  $\pi(\theta_i \mid \theta_{j \neq i})$  est difficile, cette étape de simulation peut être remplacée par une seule étape de M-H de loi cible  $\pi(\theta_i \mid \theta_{j \neq i})$  ce qui ne modifie pas la loi stationnaire de la chaîne et est entièrement valable de point de vue MCMC (Robert, 2006).

### **Conclusion :**

L'objectif de ce chapitre est de montrer comment on peut modéliser et introduire l'autocorrélation et l'hétérogénéité spatiales dans les modèles économétriques. En effet, la méthode des moindres carrés ordinaires produit des estimateurs inefficients. L'approche de l'économétrie spatiale nous permet d'introduire l'effet spatial soit sous forme de variable endogène décalée et/ou sous forme d'autocorrélation spatiale dans les erreurs. Plusieurs méthodes d'estimation sont possibles. Les plus célèbres sont celle de la méthode du maximum de vraisemblance, et l'approche bayésienne introduite par LeSage (99). Des tests statistiques permettent de vérifier et de rendre compte de l'existence d'une corrélation spatiale dans les données.

L'hétérogénéité spatiale peut se présenter par une instabilité des paramètres dans l'espace. Dans un tel cas, il faut spécifier des régimes spatiaux et où les paramètres dépendent de la localisation spatiale des données.

La méthode du maximum de vraisemblance, dominante jusqu'à la fin des années 90 (Anselin et Berra, 98), peut être coûteuse en terme de temps de calcul du Hessien avec des échantillons de tailles importantes. De surcroît, dans les modèles d'économétrie spatiale, la normalité des erreurs est souvent remise en cause (LeSage, 1997). L'étude des modèles spatiaux, notamment ceux de l'économie géographique, nécessite l'introduction de techniques d'estimations aptes à prendre en compte des phénomènes complexes tels que : les variations des paramètres dans l'espace, l'asymétrie, l'incertitude dans les mesures de distance ou de proximité, etc. Les méthodes bayésienne nous permettent de tenir compte à la fois des

problèmes d'hétérogénéités, des observations aberrantes et de mieux considérer les fondements économiques dans la mesure des interactions spatiales.

La simulation d'une variable aléatoire de loi  $\pi$  est un outil numérique très utile pour résoudre des problèmes qui n'ont pas de solution analytique évidente. La simulation par les méthodes de Monte Carlo par Chaîne de Markov, repose sur l'utilisation de la dynamique d'une chaîne de Markov ergodique de loi invariante. Nous avons présenté les deux principales méthodes de simulation MCMC, l'algorithme de Metropolis et Hastings et la méthode d'échantillonnage de Gibbs que nous allons utiliser dans les modélisations des problèmes étudiés dans cette thèse.

Les techniques de l'économétrie spatiale permettent d'améliorer les estimations produites par les MCO, de détecter la présence des effets de débordement géographiques sur les comportements économiques et d'évaluer l'impact de ces effets à travers l'estimation des paramètres spatiaux. Les études empiriques que nous allons conduire dans les deux derniers chapitres ont pour objet d'introduire les effets de dépendance et d'hétérogénéité spatiale dans la mesure de la convergence des taux de croissance de l'emploi régional en Algérie.

## **Chapitre III**

### **Etat de l'espace et de l'économie en Algérie**

#### **Introduction :**

Dans ce chapitre nous présenterons un aperçu sur l'économie algérienne, ainsi que les caractéristiques socioéconomique de l'espace algérien. Nous précéderons par la présentation des caractéristiques des différentes périodes traversées par l'économie Algérienne et les performances de la croissance hors hydrocarbures. Nous analyserons les déterminants macroéconomiques de la croissance de l'économie, et dont le principal secteur en terme de part des exportation est celui des hydrocarbures. Nous donnerons un aperçu sur la part des différents secteurs dans la croissance globale, ainsi leurs performances en terme de productivité des facteurs.

Une autre partie sera consacrée à l'analyse spatiale et à la distribution des facteurs humains et socioéconomiques. Nous donnerons une photographie sociodémographique des wilayates, et les potentialité de chacune d'elles en termes de ressources humaines et d'infrastructure susceptibles d'être à la base de l'attractivité et de la compétitivité de chaque espace.

#### **1- Tendances macroéconomiques**

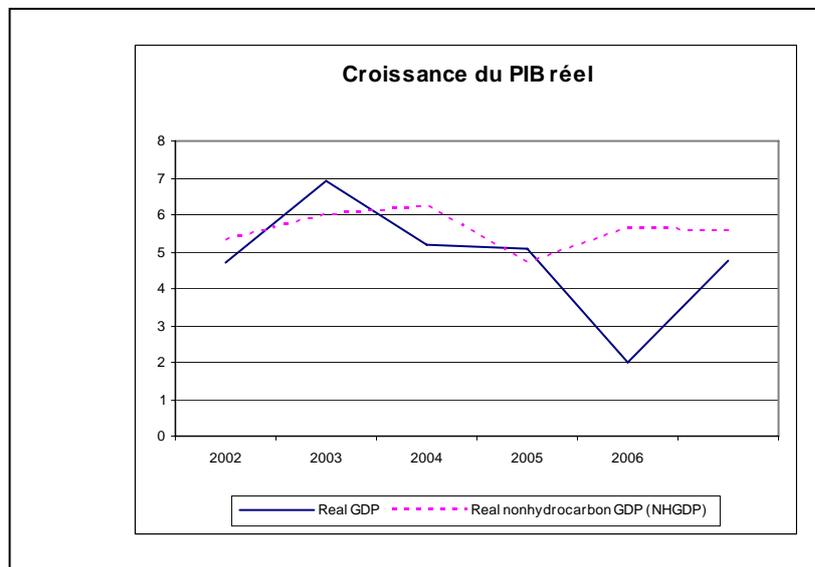
Le secteur hors hydrocarbures représente 2/3 du PIB et 98 % de l'emploi global, mais il ne représente qu'une faible part des exportations. La croissance du PIB réel enregistre une légère croissance avec une moyenne annuelle de 6% pour la période 2002 – 2006. Toute fois, au vu de fluctuations observées, la tendance générale de la croissance de cette période ne semble pas stable. Aussi, cette performance est modeste si on la compare aux taux annuels de croissance des pays MENA producteur de pétrole.

Les performances de la croissance dans le secteur hors hydrocarbure peut être divisée en trois périodes (IMF country report n° 07/61, 2007) :

- Période d'expansion 1980 – 85 : dans cette période on observe une croissance annuelle moyenne du secteur hors hydrocarbure de l'ordre de 5.1 %. Cette période était caractérisée par une stabilité politique, des prix élevés des hydrocarbures et une croissance rapide dans les dépenses publiques.

- Période de récession 1986 – 94 : dans cette période, caractérisée par une instabilité politique, une chute du prix du pétrole, la production hors hydrocarbures a chuté d'environ 1 % par année.
- Période de reprise de la croissance à partir de 1994. caractérisée par l'adoption de réformes structurelles conduites par le support du FMI et la banque mondiale. L'augmentation du prix du pétrole a permis aux autorités de lancer en 2002 le plan de soutien à la relance économique et à partir de 2004 le programme de consolidation de la croissance économique. A partir de 2002, la croissance réelle du PIB hors hydrocarbure a atteint un taux annuel moyen de croissance de 5 ½ % et la période 2002 – 05 et 5.56 % pour la période 2002 – 06.

**Figure 6 : Croissance di PIB réel**



Du côté de l'offre, la croissance est particulièrement forte dans l'agriculture, construction et travaux publics, et dans les services non gouvernementaux. Du côté demande, la croissance du PIB hors hydrocarbure était principalement menée par l'absorption domestique avec un taux de croissance des exportations hors hydrocarbures inférieur à 2 %.

Boileau Loko et Kangni Kpodar (2006) montrent que pour la période 1981 – 2005 l'accumulation des facteurs était la source principale de la croissance du PIB en Algérie. Cependant, la productivité totale des facteurs semble négative pour la même période. Les facteurs qui peuvent expliquer la faiblesse de la productivité sont englobés dans : la lenteur dans la transition vers l'économie de marché, faiblesse des institutions, la faiblesse de la taille

du secteur privé, climat non favorable des investissements, la distorsion du marché de travail, et la immigration de la main d'œuvre qualifié

**Tableau 1** : Sources de la croissance en Algérie

	1981 - 2005	1981- 85	1986- 94	1995- 2001	2002- 05
Croissance réelle du PIB en %	2.8	5.2	0.0	3.1	5.5
	Taux de croissance des facteurs en %				
Capital	2.8	6.8	2.1	1.2	2.3
Travail	4.1	4.3	3.5	3.6	6.5
Capital humain	9.6	11.3	9.7	7.8	10.6
	Contributions (en points de pourcentage)				
Capital	1.4	3.4	1.1	0.6	1.2
Travail	1.0	1.1	0.9	0.9	1.6
Capital humain	2.4	2.8	2.4	2.0	2.7
TFP	-2.0	-2.1	-4.3	-0.3	0.1

Source : Boileau Loko et Kangni Kpodar (2006) Calculé en supposant une fonction de production de type Cobb-Douglas

Suivant les données disponibles le taux de chômage (y compris celui des jeûnes) a remarquablement baissé depuis 2000. Celui-ci est passé de 30 % à 15.3 % en 2005, malgré qu'il reste supérieur à la moyenne de celui des pays du MENA. L'effort de la relance économique et de celui de la croissance a contribué à la baisse du taux de chômage. Le degré de la baisse du taux de chômage de la période 2004 – 05 est fortement déterminé par l'emploi à domicile, de nature temporaire, et qui inclus aussi l'emploi informel et le service national. En plus l'emploi permanent décroît par rapport à l'emploi temporaire. En 2004, plus de 90 % de l'emploi dans le secteur public était permanent contre seulement 11 % dans le secteur privé. Bien que l'emploi dans le secteur public est en décroissance, le moteur de la création de l'emploi reste l'Etat. A titre d'exemple dans le secteur de la construction et des travaux publics, la croissance de l'emploi de 3 % en 2001 à 8 % en 2004, avait pour raison principale la croissance des dépenses d'investissements de l'Etat induites par de la croissance des recettes des hydrocarbures. En outre, selon un rapport de L'OCDE (2004), l'impact du programme de relance économique sur la croissance et la création de l'emploi dans le long terme est mitigé et non claire.

Selon les résultats du rapport IMF, n° 07/61 (Algeria : Selected Issues) : l'élasticité de l'emploi et la croissance en Algérie montre, suivant des estimations OLS, une croissance d'un pourcent dans le PIB hors hydrocarbures peut induire une croissance de 0.9 pourcent dans l'emploi total (l'emploi domestique étant exclu). Les services non gouvernementaux et bâtiment et travaux publics sont les secteurs les plus intenses en terme du facteur travail. Dans le secteur industriel, l'emploi n'est pas corrélé avec la production, probablement à cause de la baisse considérable de la taille que ce secteur a connu.

Une analyse macro économique faite par l'utilisation d'une OLS sur des données sectorielles agrégées donne les résultats suivants :

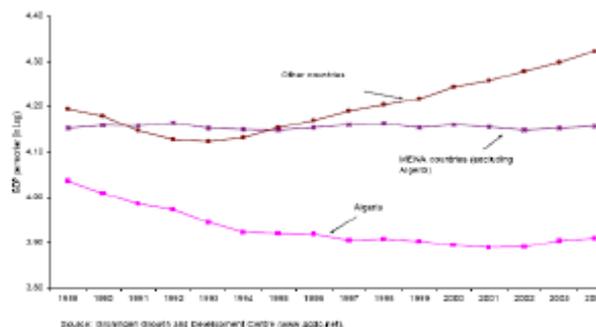
Tableau 2 : Elasticité emploi – production 1989 - 2006

Dependant variable: Employment	Total employment	Total employment (excluding work at home)	Employment by sector				
			Agriculture	Industry	Construction and public works	Government services	Nongovernment services
Nonhydrocarbon GDP	1.34*** (0.15)	0.90*** (0.09)					
GDP by sector							
Agriculture			0.87*** (0.07)				
Industry				0.31 (0.22)			
Construction and public works					0.94*** (0.04)		
Government services						0.67*** (0.09)	
Nongovernment Services							1.79*** (0.14)
Constant	-1.72 (1.14)	1.52** (0.71)	1.98*** (0.39)	4.49*** (1.26)	1.17*** (0.28)	3.12*** (0.55)	-4.45*** (0.95)
Observations	17	17	17	17	17	17	16
R-squared	0.85	0.86	0.92	0.12	0.97	0.79	0.92

La faiblesse de la productivité de travail explique l'importance relative de l'élasticité emploi – croissance. Cette dernière relation a aidé pour réduire substantiellement le taux de chômage dans les dernières années. Or une réduction durable du chômage ne peut pas se faire dans un environnement de faible productivité. L'amélioration de cette dernière conduit à la réduction du coût unitaire du travail, ce qui doit conduire à l'intensification de la création de l'emploi.

La création de l'emploi semble importante dans les secteurs de croissance faible (ou négative) de la productivité du travail. En effet, pendant la décennie 1990, les secteurs les plus dynamiques en terme de création de l'emploi étaient le secteur administration public et le secteur des services, et où la croissance de la productivité était nulle ou négative. Durant les dernières années, l'emploi a augmenté rapidement dans le secteur agricole et dans celui des services, et dont la croissance de la productivité était négative.

**Figure 7 : Productivité du travail dans les pays du MENA et des pays en transition 1989 - 2004**



La productivité de travail, mesurée par le rapport de la production hors hydrocarbures, a baissé de 2.3 % entre 1997 et 2004, par contre le salaire réel a augmenté de 14% durant la même période. La productivité dans le secteur agricole reste volatile et fortement dépendante des conditions pluviométriques. La croissance de la productivité dans le secteur industriel et dans celui des services gouvernementaux reflète une stagnation ou une baisse de l'emploi dans ces secteurs.

## 2- Historique de la création des wilayates :

L'espace Algérien, entre 1515 et 1848 était organisé dans la partie Nord du pays autour de trois beyliks (entités administratives) : le Ponant, Titteri et le Levant. Ces trois entités administratives sont placées sous l'autorité d'un bey dépendant directement du Dey d'Alger nommé lui-même par le pouvoir turc. Le beylik du Levant, est constitué de l'Algérie orientale, avec pour centre politique Constantine, et comme point d'ouverture sur la Méditerranée, le port de Bône. Le beylik du Titteri, en position centrale, a pour capitale Médéa et pour port Dellys. Enfin le dernier beylik, du Ponant, localisé dans l'Algérie occidentale, s'organise autour de Mascara, et d'Arzew comme base portuaire. Alger et la plaine de la Mitidja, constituent un quatrième territoire, appelé Dar-es-Soltan, sous l'autorité d'un Agha relevant directement du Dey. Le pouvoir turc, marin par excellence, installa sa capitale politique à Alger sur le littoral, ce qui constitue une rupture avec les précédents royaumes qui s'appuyaient sur des capitales localisées préférentiellement dans le Tell, et les Hautes Plaines.

En 1848, la France a partagé l'Algérie en trois départements : (départements d'Oran, d'Alger et de Constantine). Le sud de ces départements formait 6 territoires qui furent regroupés en Territoires du Sud en 1902 puis réorganisés en 4 territoires en 1905.

### Tableau 3 Territoires Algérien en 1905

Nom	Chef-lieu	Dates d'existence
Alger	Alger	1848 — 1957
Oran	Oran	1848 — 1957
Constantine	Constantine	1848 — 1957
Bône	Bône	1955 — 1957
Territoires du Sud	—	1902 — 1957

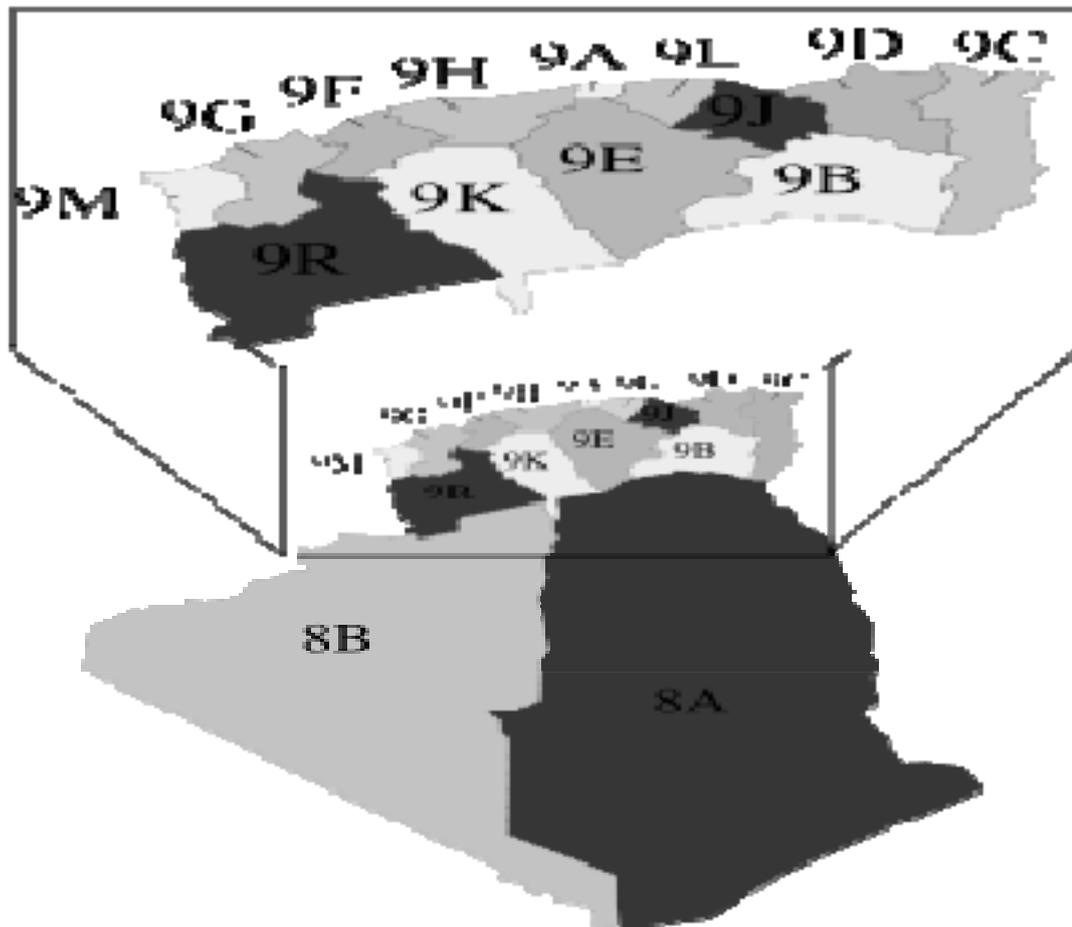
En 1957, l'organisation territoriale de l'Algérie fut modifiée : 14 départements furent créés en remplacement des 5 précédents.

En 1955, le Département de Bône fut créé à partir du département de Constantine. En 1958, trois nouveaux départements (Aumale, Bougie et Saïda) s'y ajoutèrent. En 1959, les départements d'Aumale et de Bougie furent supprimés. Les autres départements persistèrent jusqu'à l'indépendance de l'Algérie en 1962. Les divisions territoriales qu'ils représentaient furent en revanche conservées jusqu'en 1974, prenant le nom de wilayas à partir de 1968.

**Tableau 4 : Départements Algériens en 1957**

Numéro	Nom	Chef-lieu	Dates d'existence
8A	Oasis	Ouargla	1957-1962
8B	Saoura	Béchar	1957-1962
9A	Alger	Alger	1957-1962
9B	Batna	Batna	1957-1962
9C	Bône	Bône	1957-1962
9D	Constantine	Constantine	1957-1962
9E	Médéa	Médéa	1957-1962
9F	Mostaganem	Mostaganem	1957-1962
9G	Oran	Oran	1957-1962
9H	Orléansville	Orléansville	1957-1962
9J	Sétif	Sétif	1957-1962
9K	Tiaret	Tiaret	1957-1962

9L	Tizi-Ouzou	Tizi-Ouzou	1957-1962
9M	Tlemcen	Tlemcen	1957-1962
9P	Bougie	Bougie	1958-1959
9N	Aumale	Aumale	1958-1959
9R	Saïda	Saïda	1958-1962

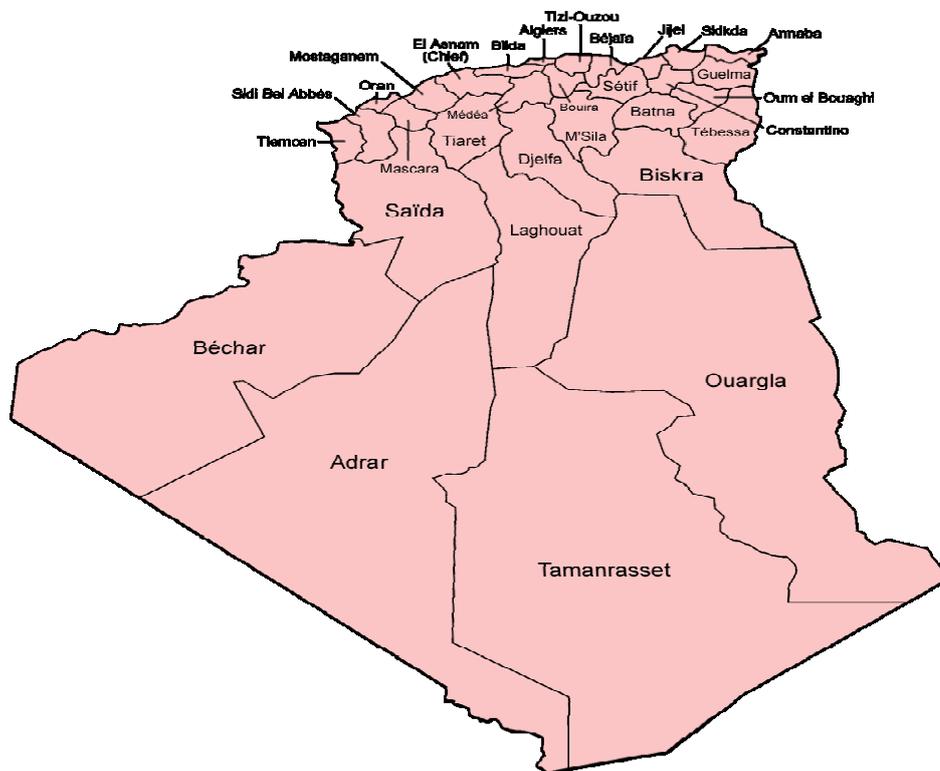


Durant la guerre de libération, le territoire algérien est découpé en six wilayat : l'Oranais, l'Algérois, la Kabylie, le Constantinois, l'Aurès, et le Sud. Alger retrouve un statut proche de celui qui lui était conféré par l'administration turque, le Dar es Soltan, et constitue une zone autonome.

Au lendemain de l'indépendance, l'intervention sur le maillage territorial s'inscrivait dans une conjoncture où la maîtrise du territoire l'emportait sur la logique spatiale : « sitôt l'indépendance acquise, l'une des préoccupations majeures de l'Algérie fut d'asseoir son autorité et de renforcer le contrôle de son territoire, (Troin, 1985, p. 312). Le nombre de départements, rebaptisés *wilayas*, fut maintenu mais les communes regroupées.

Les 15 wilayas issues de la colonisation française passent à 31, du fait de l'accroissement démographique. Un nouveau découpage intervient en 1974, avec pour vocation de créer un maillage plus cohérent et efficace, ainsi que de favoriser le développement économique du pays. L'Etat algérien voulait ainsi gommer davantage les disparités créées par l'ancienne puissance coloniale et engendrer un développement solidaire entre wilayat prospères et wilayat économiquement en difficulté. Le nombre de wilaya passe de quinze à trente et une et les dénominations des wilayat se simplifient, chaque wilaya prend dès lors le nom de son chef lieu.

**Carte 1 : Carte de l'Algérie entre 1974 et 1984**

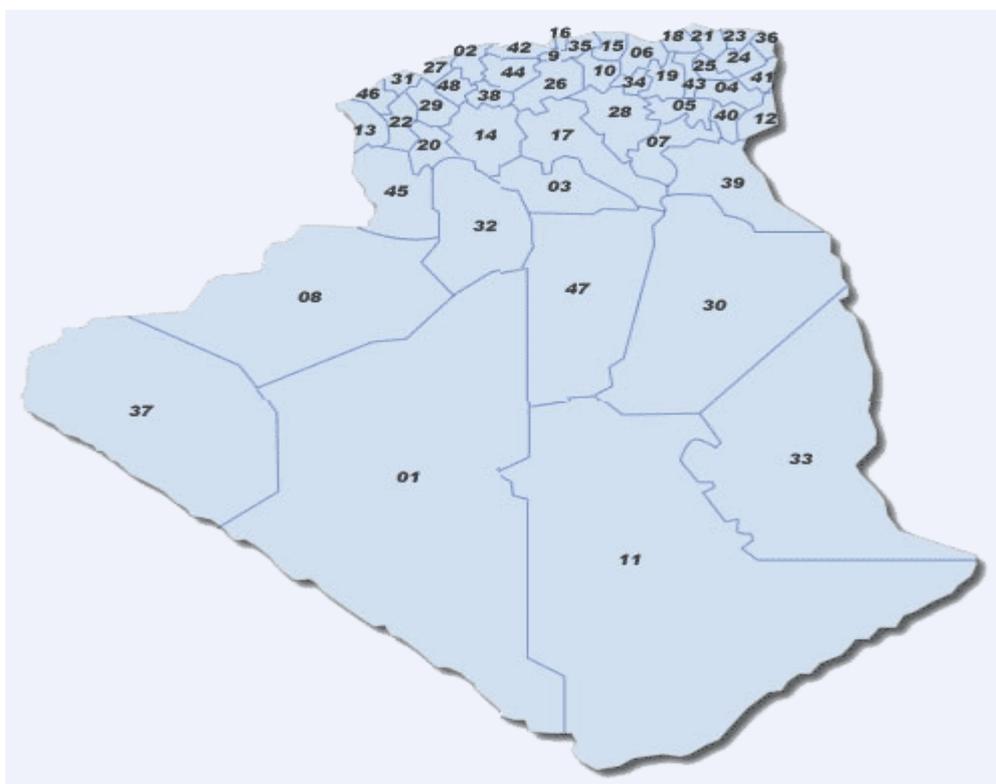


Ensuite on chercha à réduire les disparités régionales héritées du système macrorégional de la période coloniale. La logique de la nouvelle partition territoriale visait une homogénéisation de la trame administrative : promouvoir des wilayas plus comparables par la taille et appliquer une seule grille d'équipements. Il en est résulté une augmentation sensible du nombre de circonscriptions et une diminution de la taille des mailles, y compris de la maille communale.

La création de nouvelles unités administratives en 1984 a contribué à réduire l'écart statistique et donc à renforcer la cohérence entre les différentes circonscriptions. En effet, par le biais de ce découpage, il s'agissait d'impliquer le plus d'espace et le plus de population

possible dans le processus de développement en multipliant le nombre d'unités administratives. Bien que la densité de population ait augmenté, la moyenne de la population par wilaya a diminué suite à une densification du maillage territorial.

**Carte 2 : Carte de l'Algérie depuis 1984 à nos jours**



code	wilaya	code	wilaya	Code	wilaya	code	wilaya	code	wilaya
1	ADRAR	11	TAMANRASSET	21	SKIKDA	31	ORAN	41	SOUK AHRAS
2	CHLEF	12	TEBESSA	22	SIDI BEL ABBES	32	EL BAYADH	42	TIPAZA
3	LAGHOUAT	13	TLEMCEN	23	ANNABA	33	ILLIZI	43	MILA
4	OU M EL BOUAGHI	14	TIARET	24	GUELM	34	B.B.ARRERIDJ	44	AIN DEFLA
5	BATNA	15	TIZI OUZOU	25	CONSTANTINE	35	BOUMERDES	45	NAAMA
6	BEJAIA	16	ALGER	26	MEDEA	36	EL TARF	46	AIN TEMOUCHENT
7	BISKRA	17	DJELFA	27	MOSTAGANEM	37	TINDOUF	47	GHARDAIA
8	BECHAR	18	JIJEL	28	M SILA	38	TISSEMSILT	48	RELIZANE
9	BLIDA	19	SETIF	29	MASCARA	39	EL OUED		
10	BOUIRA	20	SAIDA	30	OUARGLA	40	KHENCHELA		

certaines limites de wilayat ne sont basées que sur des particularités physiques, ou encore sur des spécificités économiques comme c'est le cas pour la wilaya de Ouargla. Il y a donc parfois incohérence entre découpage administratif et espaces fonctionnels. Le découpage de

1984, appliqué à partir du 1er janvier 1985, tente de remédier à ces incohérences en multipliant le nombre de wilayat. Celles-ci sont désormais quarante-huit, le nombre de daïra augmente de 160 à 229, ainsi que le nombre de communes qui passe alors à 1541. Le territoire algérien, devient plus finement maillé, ce qui permet une action en matière d'aménagement plus ciblée, et davantage en cohérence avec les préoccupations et attentes locales. Ces deux derniers découpages, étaient marqués par une volonté d'un développement harmonieux du territoire, volonté qui s'est exprimée notamment par le biais des politiques de promotion administrative et industrielle.

### **3- Description sociodémographique de l'espace :**

Dans cette partie nous allons analyser les dynamiques territoriales à partir d'un certain nombre de variables et d'indicateurs sociodémographiques. C'est une analyse descriptive basée sur des données statistiques de wilayas et qui donnera une vue sur les potentiels démographiques, sociales et économique des territoires, et leurs caractéristiques structurelles

#### **A - Photographie sociodémographique des wilayates :**

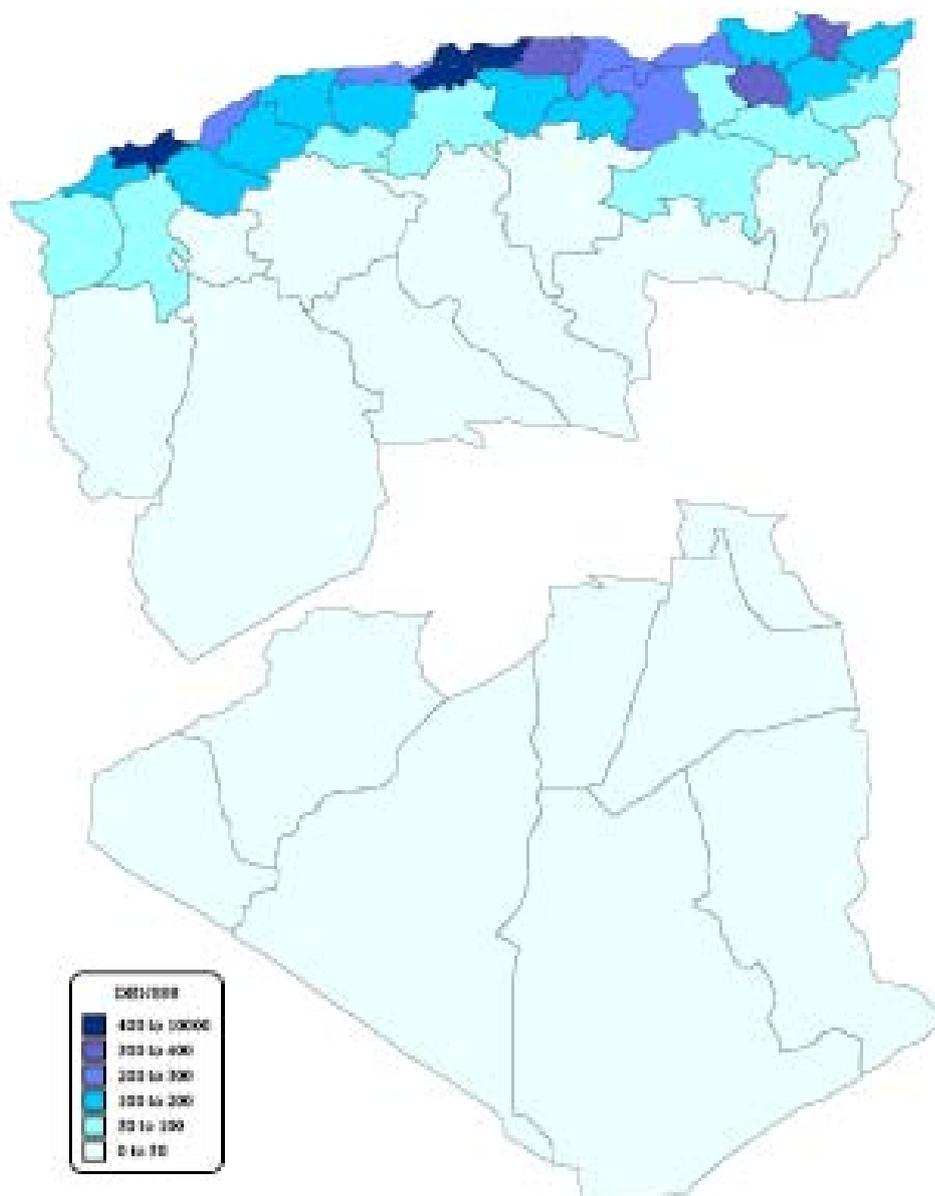
La pression démographique est une variable essentielle dans la description et la détection des agglomérations. Plusieurs variables peuvent être utilisées, comme : la variation de la population, la densité ou la variation de la densité. On peut aussi utiliser des variables telles que :

- **taux d'emploi** , qui est le rapport entre la population active et la population en age de travailler. Cet indicateur décrit la capacité d'une wilaya à mobilise sa ressource en main d'œuvre ;
- **taux d'activité** , qui est le rapport entre la population active et la population totale
- **taux d'actifs potentiels** , qui est le rapport entre la population en age de travailler (population active potentielle) et la population totale. Cet indicateur exprime la capacité d'une wilaya à produire de la richesse.

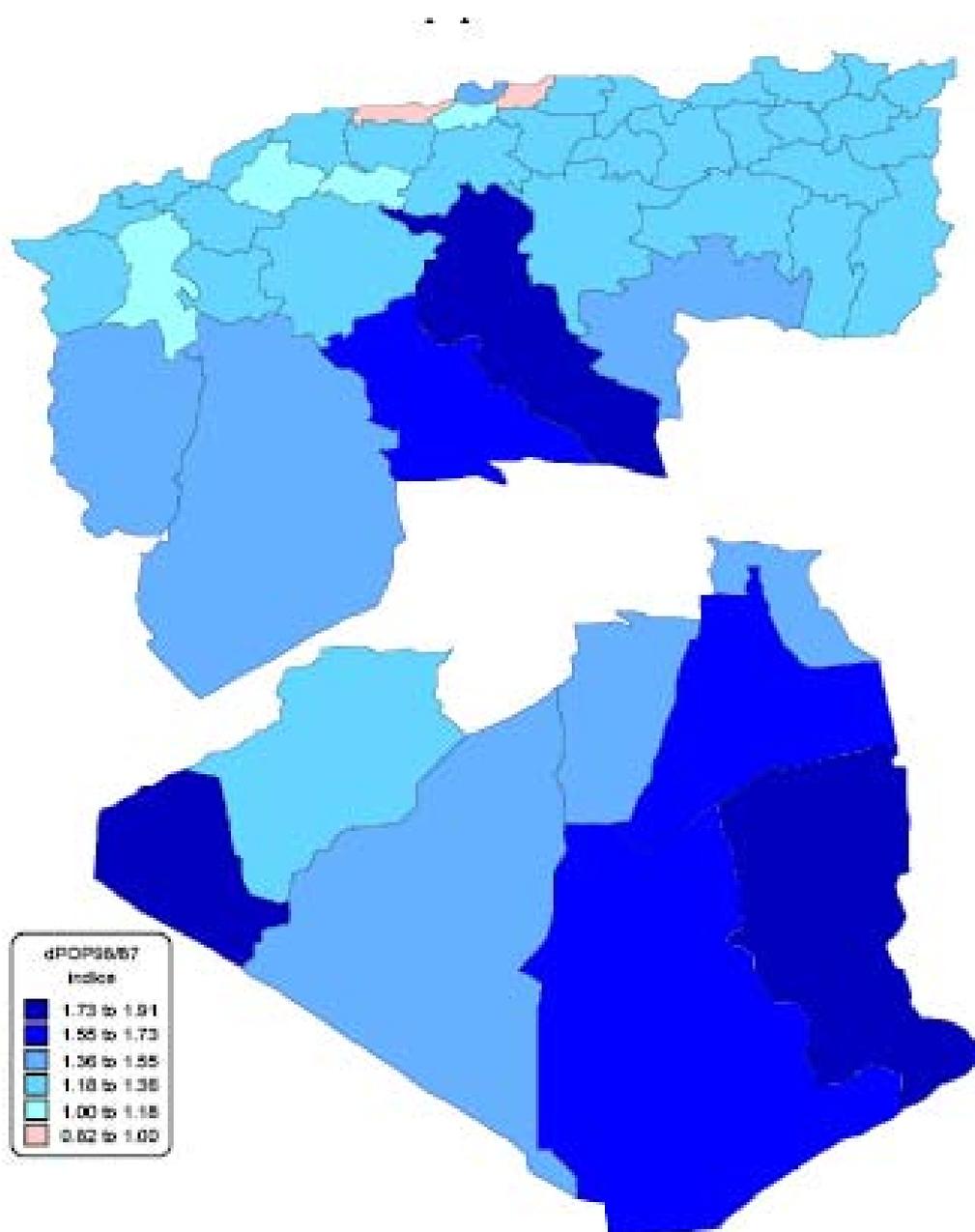
L'urbanisation et l'agglomération de la population sur l'espace algérien est indicateur important pour décrire la concentration de la demande potentiel. Dans les annuaires statistiques de l'ONS (Office National des Statistiques) la population des wilayas est classée selon trois modalités : chef lieu, agglomération secondaire et zones à habitat éparse. Une agglomération est un ensemble d'habitations espacées de moins de 200 mètres ; au-delà il

s'agit d'habitats épars. Une agglomération avec moins de 5000 habitants, est considérée rurale. Plus de 5000 habitants est une agglomération urbaine, si au moins 50 % de la population sont employés hors de l'agriculture.

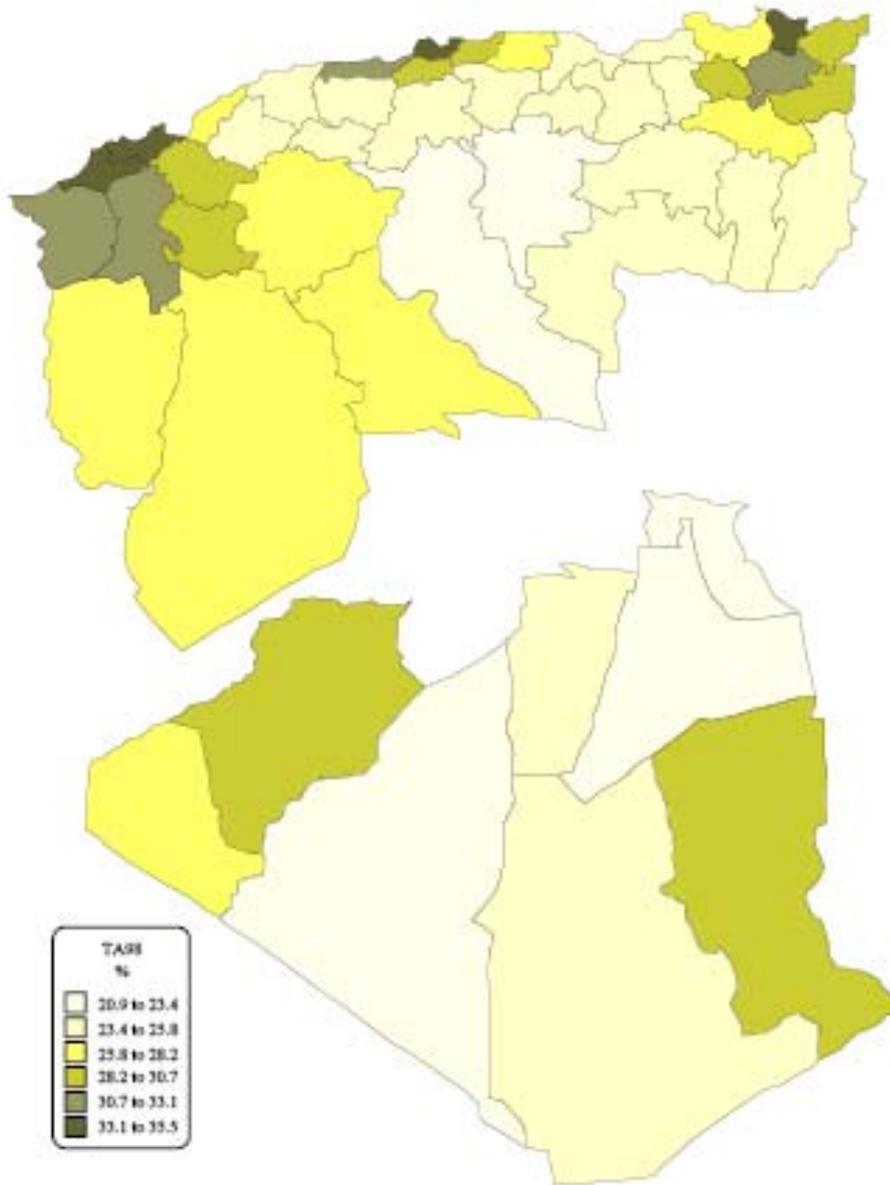
**Carte 3 : Densité de la population en 1998**



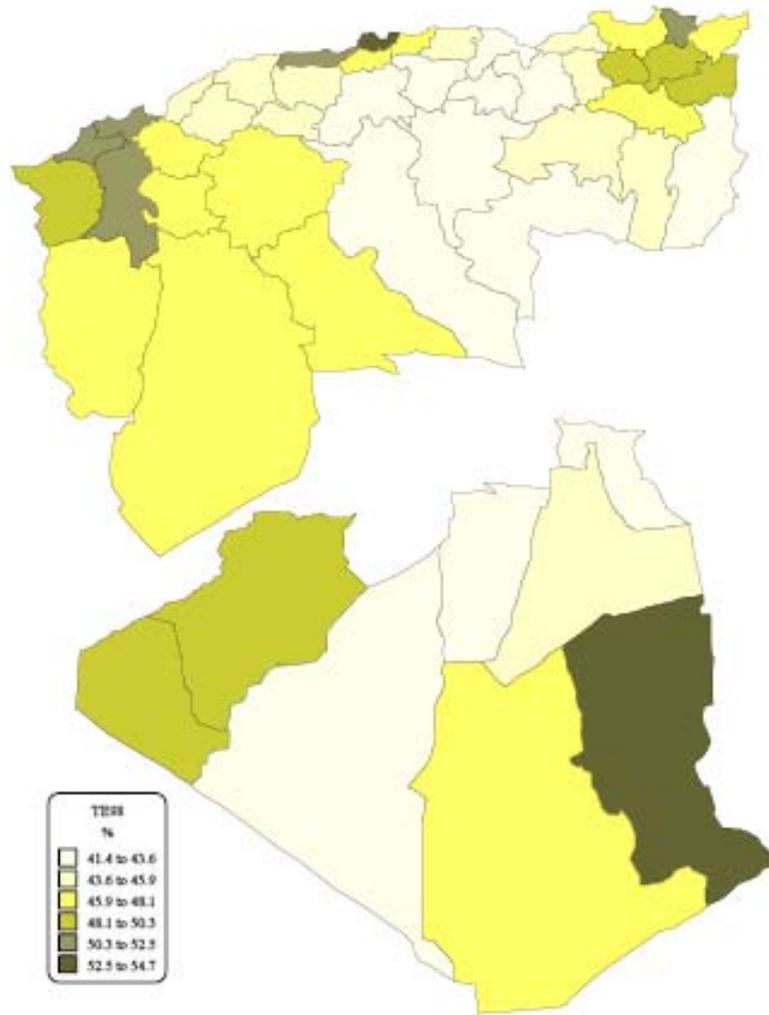
Carte 4 : Variation de la population entre 1987 et 1998



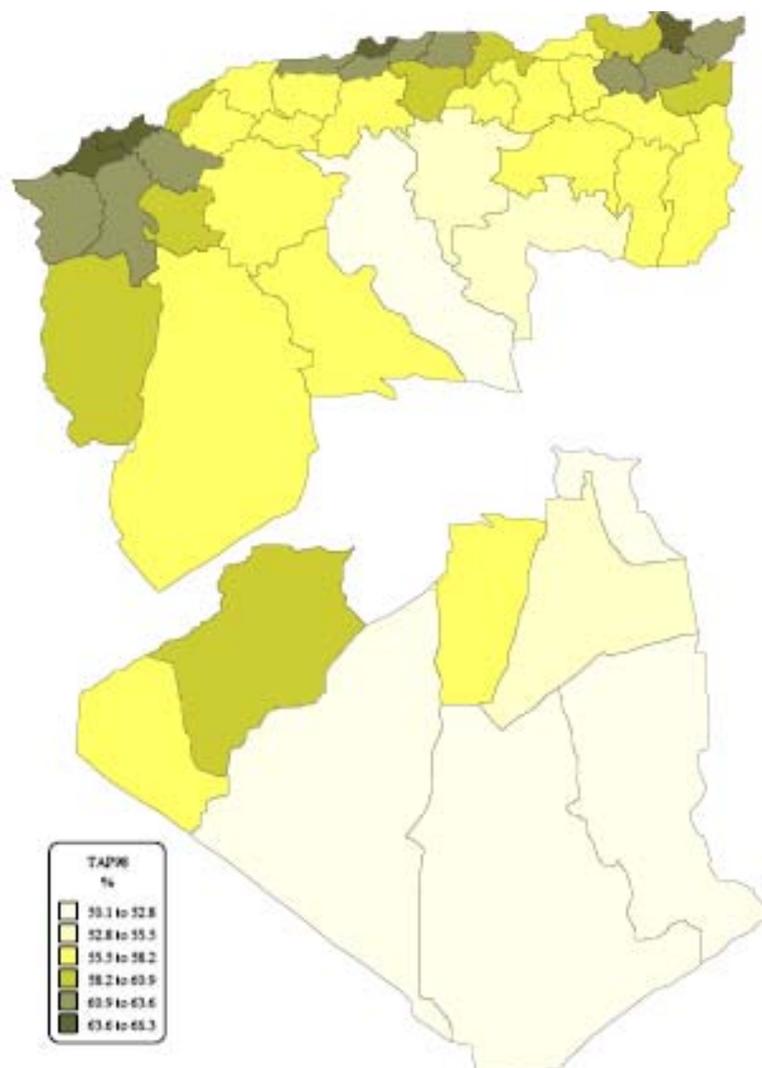
**Carte 5 : Taux d'activité de la population**



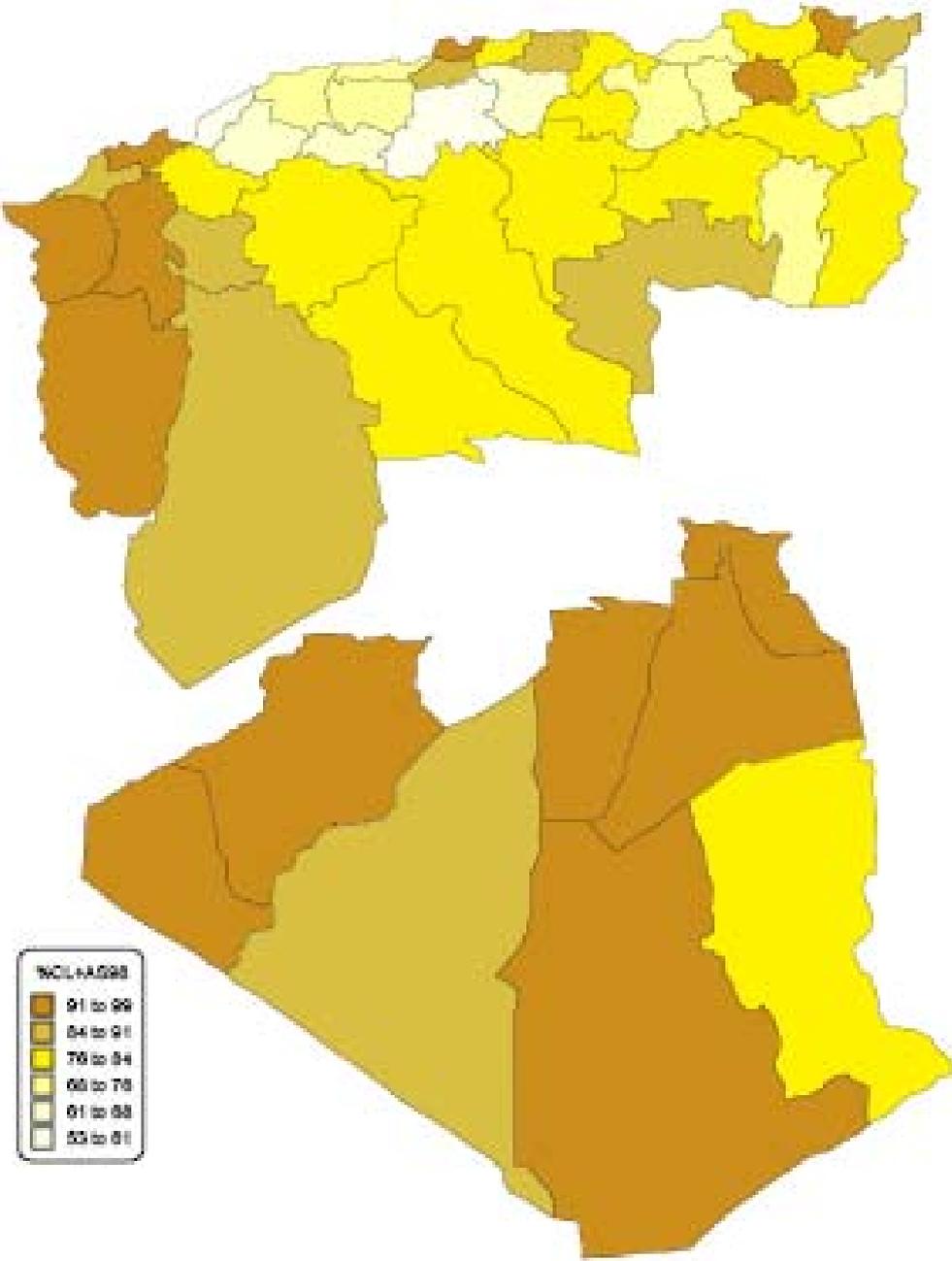
**Carte 6 : Taux d'emploi en 1998**



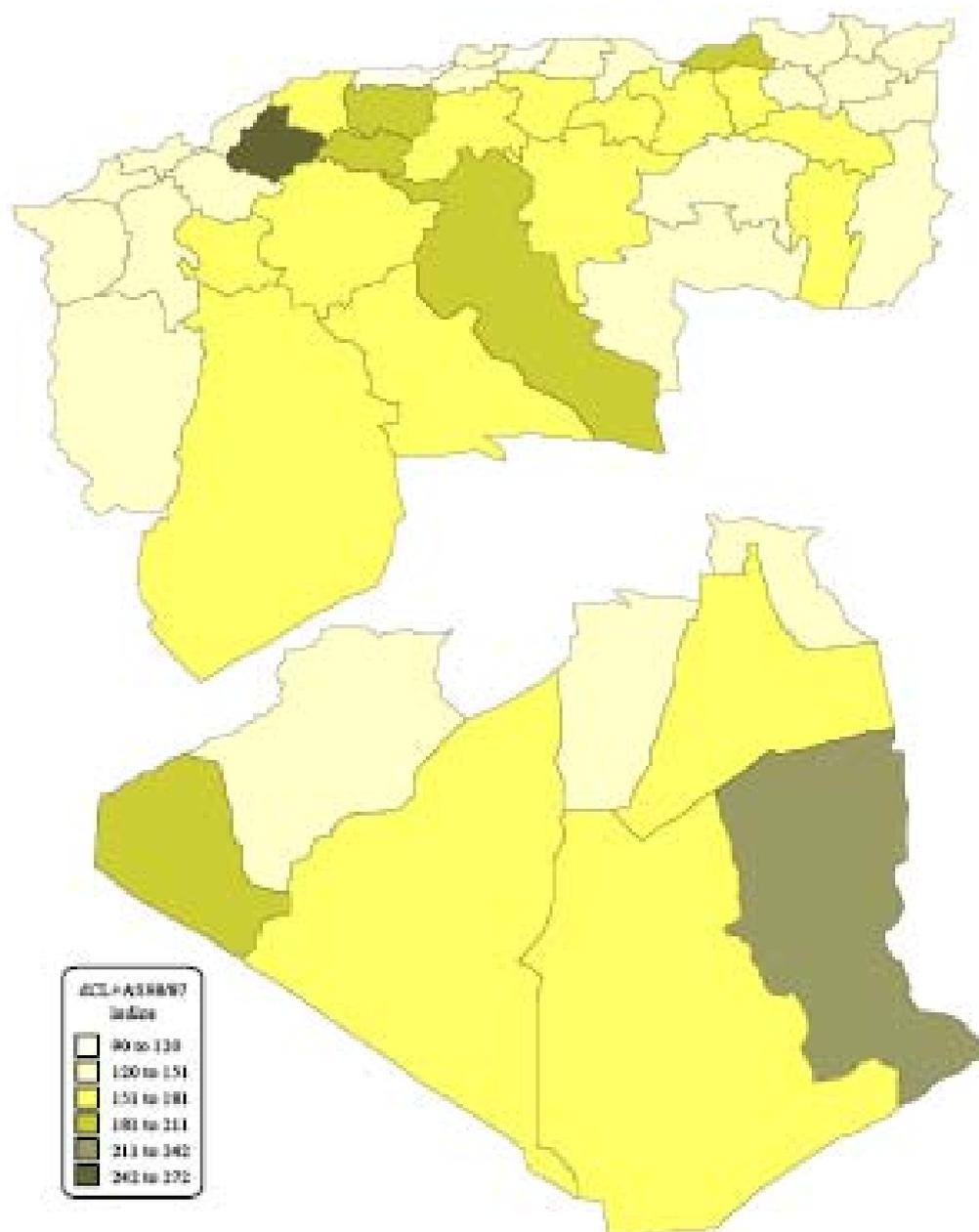
Carte 7 : Taux d'actifs potentiel en 1998



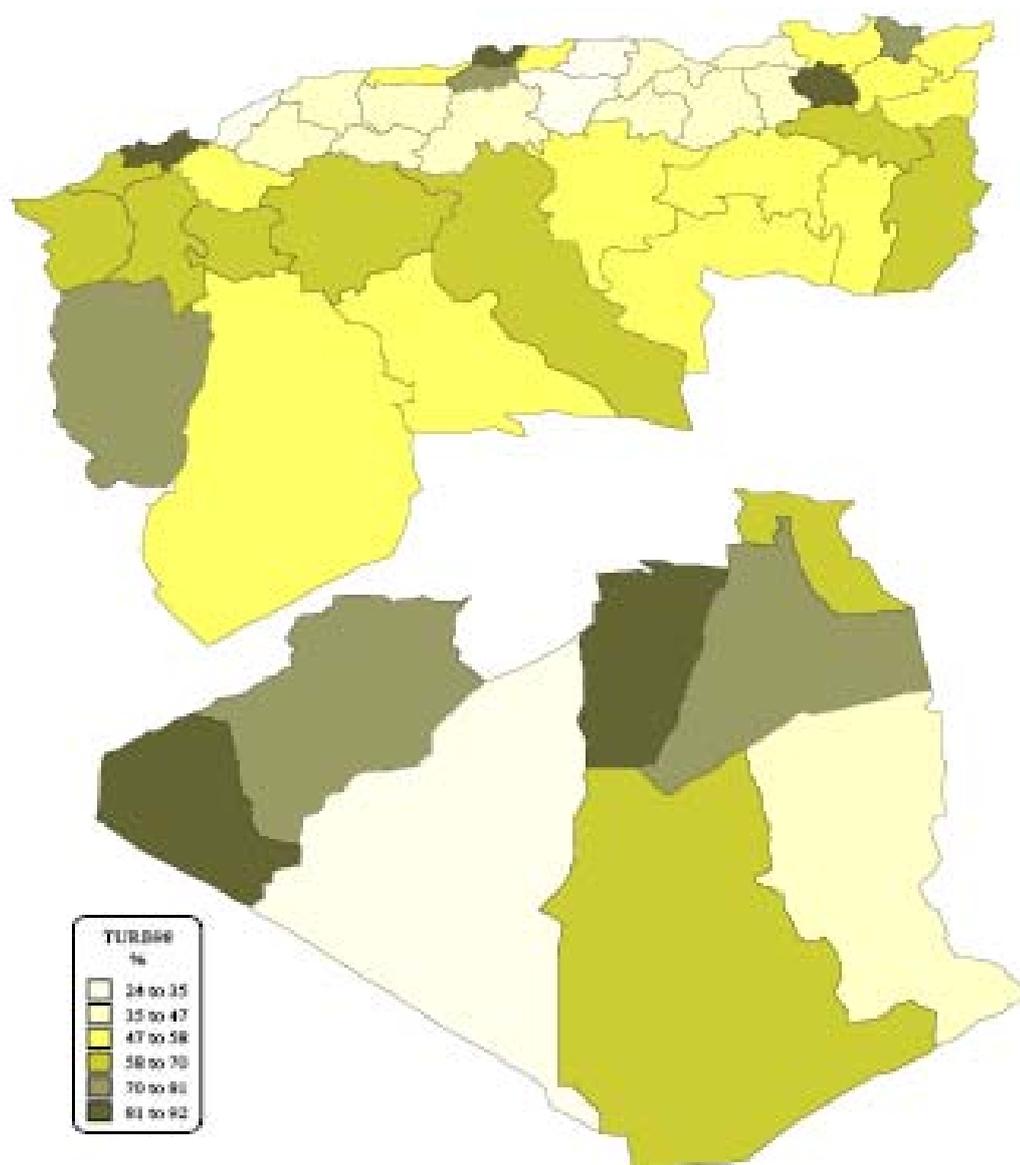
**Carte 8: Part de la population dans les agglomérations secondaires et chefs-lieux  
en 1998 (%)**



**Carte 9 : Variation de la population dans les agglomérations secondaires et chefs-lieux de 1987 à 1998 (%)**



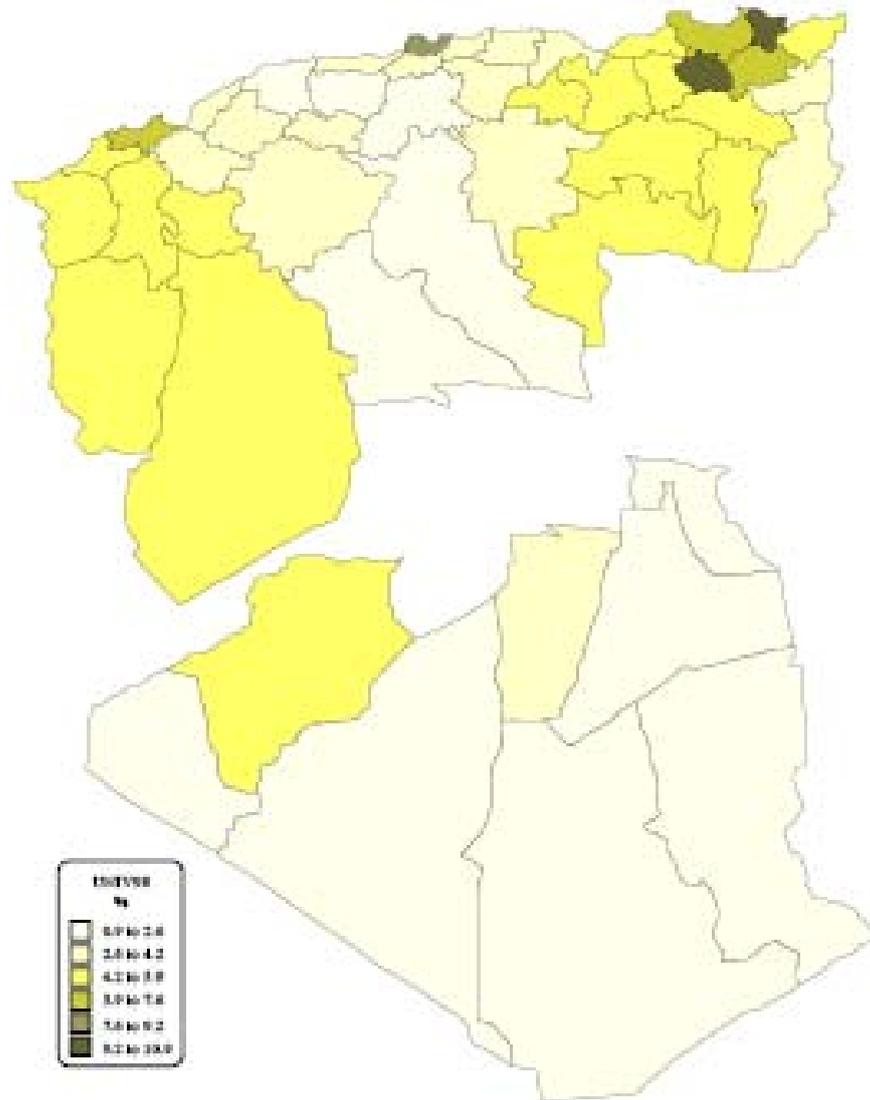
**Carte 10 : Taux d'urbanisation en 1998**



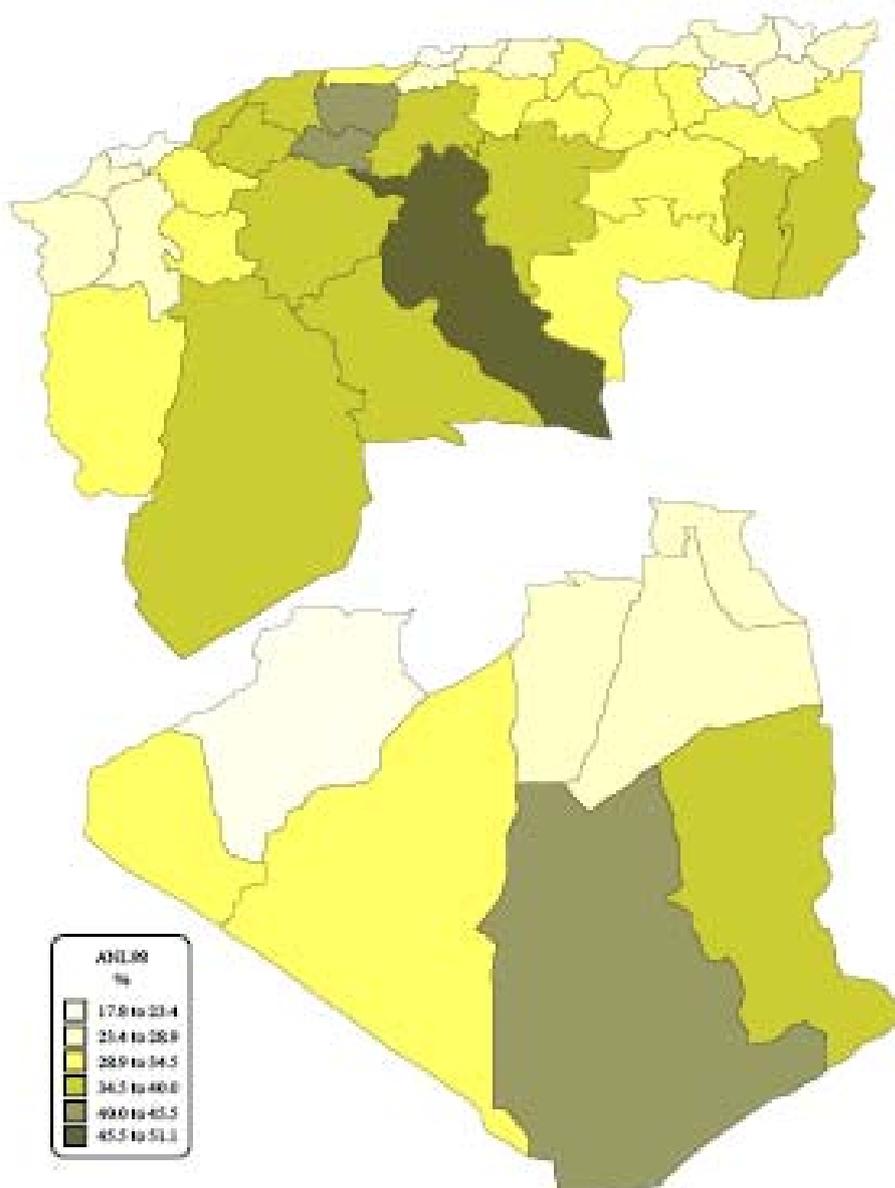
## B- Distribution spatiale des ressources humaines

La répartition spatiale du capital humain est un indicateur important dans la distribution spatiale de la croissance économique. Ce facteur donne des indications sur la mobilisation de la croissance régionale de long terme. Dans cette description, nous utilisons comme indicateur de la répartition du capital humain le **% de diplômés du supérieur** dans la population active, et le **taux d'analphabétisme**, qui est un bon indicateur de la **pauvreté**.

**Carte 11 : Part de la population ayant un diplôme universitaire en 1998**



**Carte 12 : Pourcentage de la population analphabète en 1998**

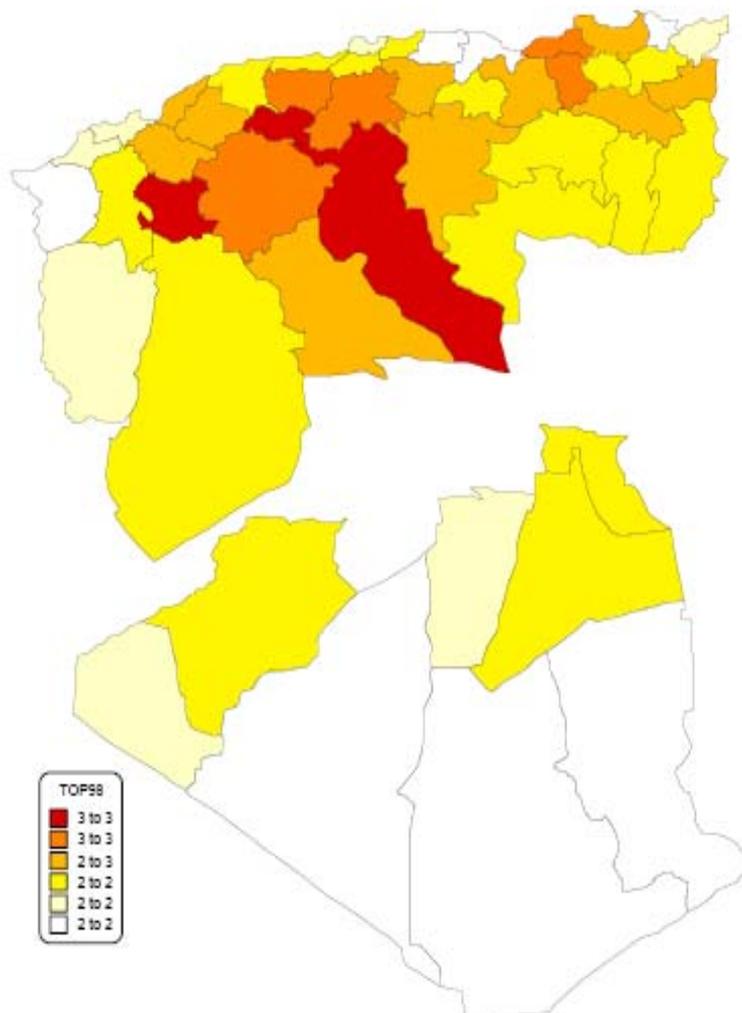


## C - Infrastructures de logement et de transport

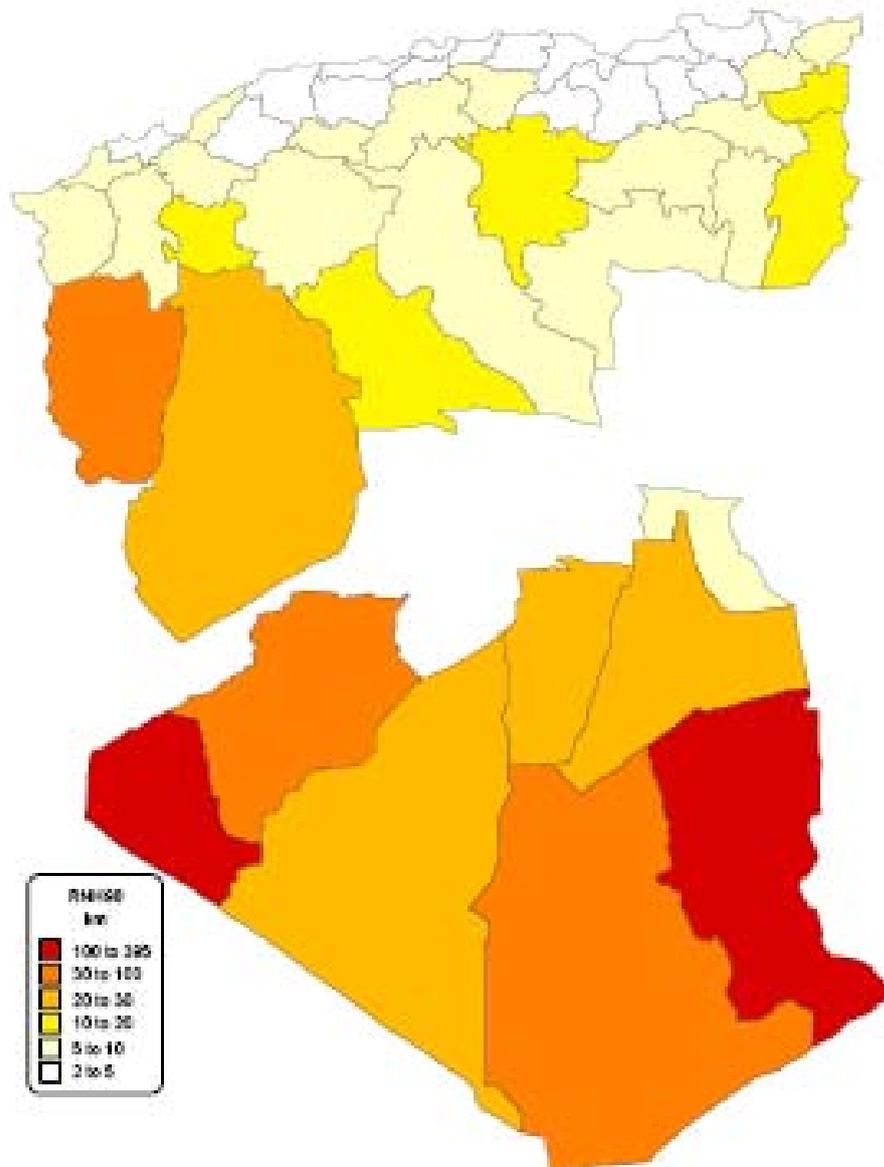
Le logement est un des facteurs de l'attractivité ou de la répulsivité des territoires. Il est tellement déficitaire en Algérie – et le gouvernement en a fait une priorité – qu'on peut s'interroger véritablement sur l'information apportée par cette variable. Pourtant le **taux d'occupation par pièce (TOP)** varie presque du simple au double, ce qui en fait un critère significatif de différenciation des wilayas.

Les infrastructures de transport sont également un des principaux critères de localisation des activités économiques. Elles représentent des externalités des territoires, et sont très importantes pour un pays étendu comme l'Algérie. Plusieurs variables peuvent être construites, nous avons choisi le **km de routes nationales revêtues**, rapportées au nombre d'habitants.

**Carte 13 : Taux d'occupation par pièce en 1998**



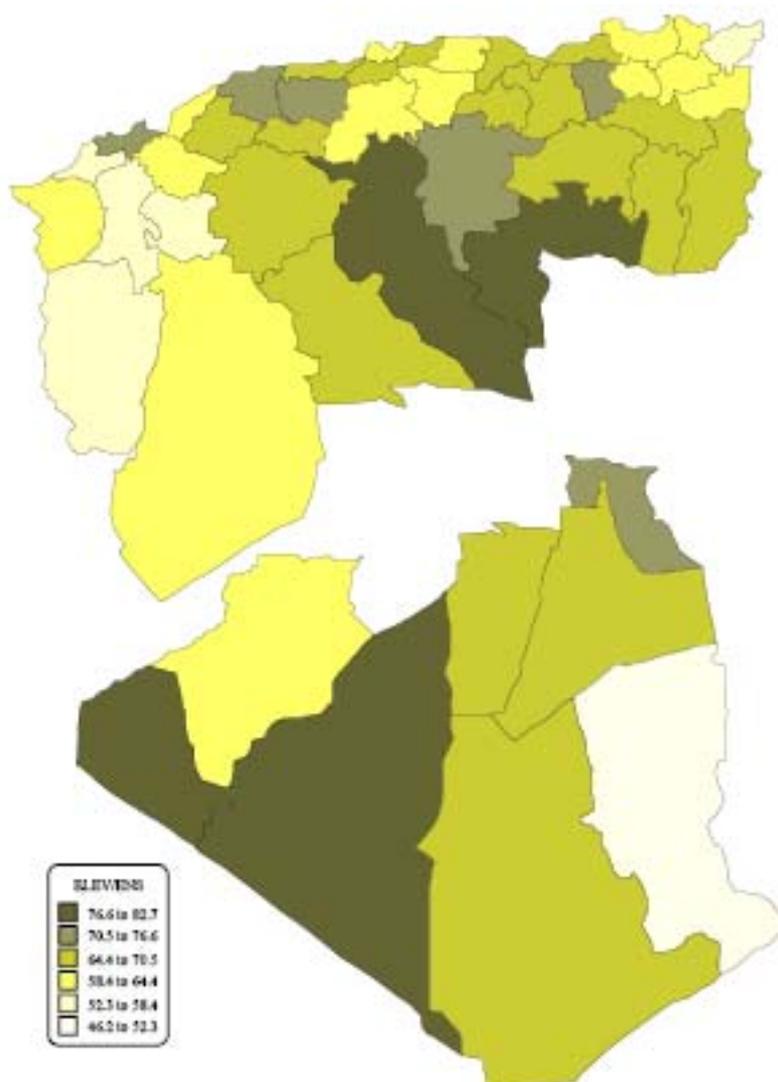
Carte 14 : KM de route nationale par habitant en 1998



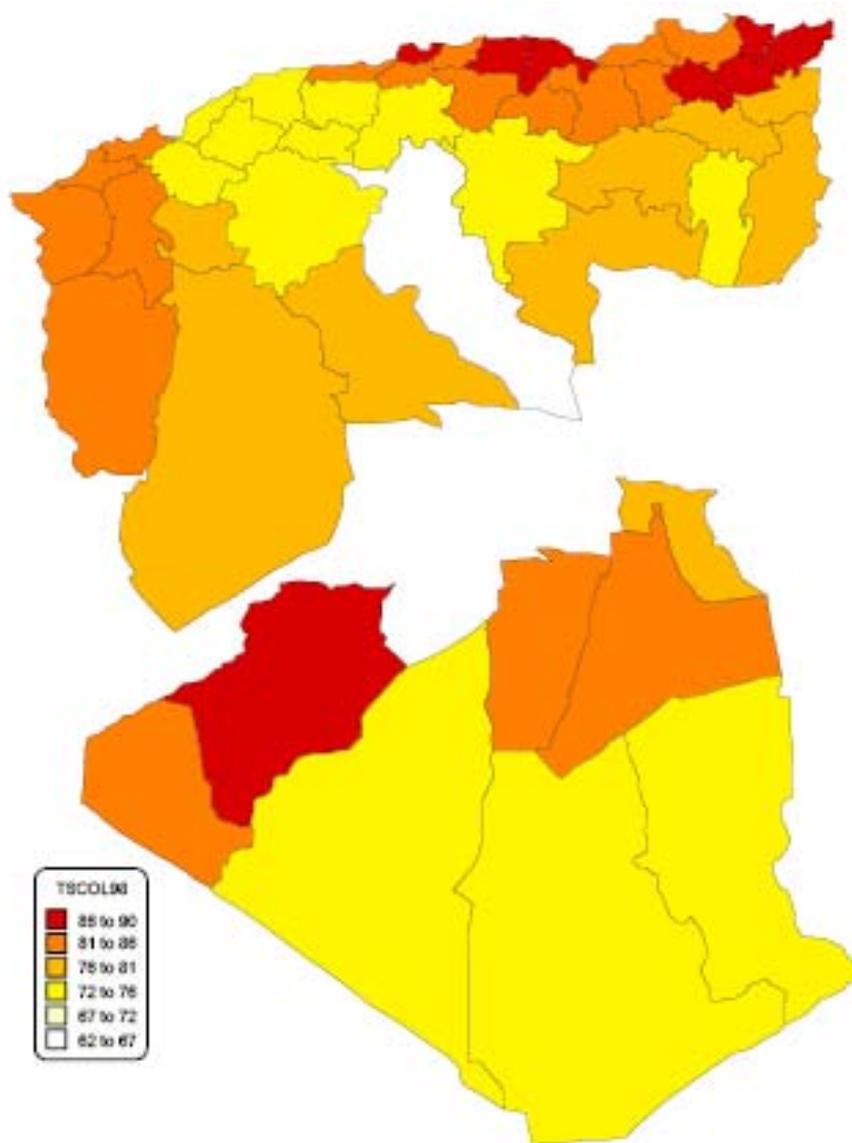
## F- Infrastructures de formation

Les infrastructures sociales, d'enseignement, et de la santé contribuent à la compétitivité et à l'attractivité des territoires. Pour les infrastructures de formation, on se concentre sur l'enseignement de base présent dans toutes les wilayates, et pour lequel nous avons regroupé les données de l'enseignement primaire, moyen et secondaire. Comme indicateurs, nous utilisons le nombre d'élèves par enseignants, le taux de scolarisation. Pour la santé nous utilisons le nombre de médecins pour 1000 habitants

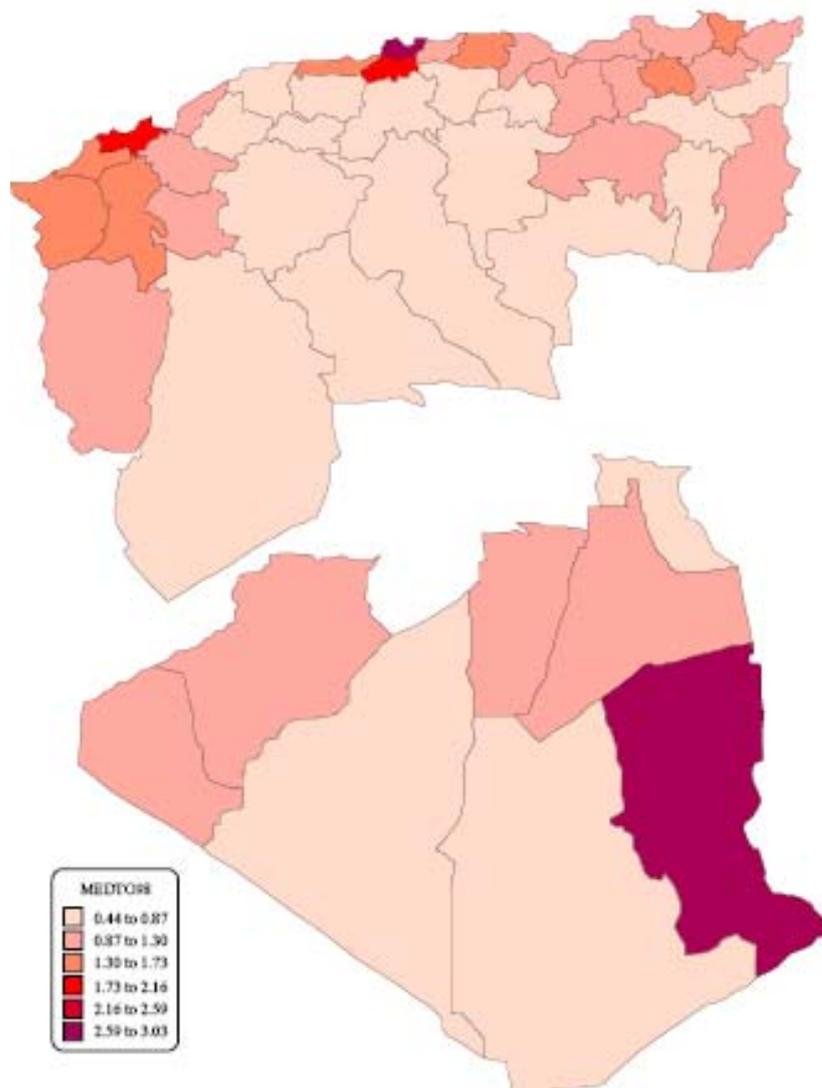
**Carte 15 Nombre d'élève par enseignant**



Carte 16 : Taux de scolarisation en 1998



**Carte 17 : Nombre de médecins pou 1000 habitants en 1998**



## **Conclusion :**

L'analyse macro économique a montré que la croissance de l'économie algérienne est dépendante de facteurs fortement volatiles. Les prix des hydrocarbures et la pluviométrie ne peuvent pas à eux seuls assurer une croissance structurellement stable.

Les dynamiques économiques locales en Algérie sont la résultante de la répartition de l'effort public pour stimuler l'investissement à travers les infrastructures et la différenciation des avantages, et d'un autre côté la transcription actuel marquée par de fortes disparités économiques. Les potentialités de développement sont fortement différentes entre les régions (wilayates). L'importance de l'espace Algérien et sa forte hétérogénéité géographique exige le passage d'une logique de gouvernance centralisée et souvent loin des de la réalités des besoins locales vers une logique décentralisée basée sur une optique de compétitivité des territoires. La mise en valeur des capacités de développement locales doit se faire en fonction des moyens possibles de chaque espace, et le soutien de l'Etat doit prendre en compte le fait qu'on ne peut pas réaliser une équilibre régionale en offrant les mêmes possibilités quelque soit la région considérée.

La description de l'espace Algérien montre une forte différenciation spatiale. Les potentialités de développement et l'émergence de clubs économiques convergent ne peuvent pas être les mêmes quelque soit le lieu de l'espace étudié.

## **Chapitre IV : Analyse économétrique de la convergence spatiale entre les wilayates**

### **Introduction**

Dans ce chapitre nous examinons les forces économiques qui peuvent expliquer la croissance de l'activité économique des régions Algériennes lors des huit dernières années. Notre démarche trouve ses origines dans les approches développées dans les travaux de Glaeser, Kallal, Scheinkman et Shleifer (1992), Henderson, Kuncoro et Tuner (1995), et d'une manière indirecte dans le travail de Rappaport (1999). Glaeser et al. ont mis en évidence l'impact de la structure industrielle locale et des types d'externalités technologiques induites sur la croissance de l'emploi. Henderson et al., se basant sur une spécification semblable à celle de Barro et Sala-i-Martin (1991), conditionnée par les caractéristiques historiques et actuelles des marchés locaux, trouvent une forte vitesse de convergence dans l'emploi industriel local. Rappaport (1999), se basant sur l'idée de Tiebot (1956), à savoir que les régions à forte densité de la population présentent de forts taux de productivité et une meilleure qualité de la vie, suggère que la variable densité de la population offre une meilleure métrique naturelle pour capturer les variations régionales de la productivité et de la qualité de la vie. Rappaport et Sachs (1999) expliquent la corrélation entre la densité de la population et les proximités côtières par l'effet combiné de la productivité et de la qualité de la vie. Ces dernières contribuent positivement à l'utilité (qualité de la vie et productivité via la croissance des salaires).

Des études, incluant celle de Barro et Sala-i-Martin (1991), ont trouvé une meilleure convergence entre les régions qu'entre les Etats. Selon Glaeser et al. (1995), l'avantage de l'étude de la convergence régionale renvoie aux deux faits suivants.

*Premièrement*, les régions sont des économies complètement ouvertes, il n'y a pas de barrière à la mobilité du capital, du travail et des idées. Les régions sont des unités économiques plus spécialisées (et moins arbitraires) que les Etats, et peuvent donner plus de sens à l'étude du mouvement des ressources et de la convergence entre les régions qu'entre les Etats.

*Deuxièmement*, les théories de la croissance insistent beaucoup sur l'importance de la diffusion des idées. Plusieurs versions de ces théories se focalisent sur les externalités technologiques induites par le capital physique (Romer, 1986), le capital humain (Lucas,

1988), les industries de haute intensité technologique (Porter, 1991), et plusieurs autres types d'externalités. Glaeser et al. (1992) trouvent de fortes évidences de l'impact des externalités technologiques entre les industries sur la croissance des régions d'un même pays.

Dans ce chapitre on applique la question de la convergence entre les régions à une économie en voie de développement. Depuis le lancement des réformes économiques structurelles, au début des années 80, l'économie algérienne observe un passage d'une économie centralisée vers une économie du marché. L'adoption des mécanismes du marché a été fortement soutenue par les accords, engagés dans le cadre du plan d'ajustement structurel, avec les institutions financières internationales. La libéralisation de l'activité économique nationale a conduit au développement et à la consolidation de l'investissement privé. Un tel processus ne peut se produire sans interactions avec le développement des zones d'activités composant l'espace économique algérien. Les premières observations sur les zones d'activités (régions-willayat<sup>5</sup>) dévoilent des dynamiques différentes dans l'intensité de création de l'emploi, les taux de création des entreprises et les types d'activités développées. Les dynamiques économiques des régions en Algérie sont très hétérogènes. Pour certaines régions la croissance de l'activité économique et de l'emploi est très significative, tandis qu'elle est moins évidente pour d'autres.

Aussi, on s'interroge sur la relation entre la croissance économique de l'emploi manufacturier par tête dans 48 régions d'Algérie et la localisation géographique de ces régions (en terme de voisinage régional immédiat), l'emploi par tête d'habitant initial, la structure de l'industrie locale, l'intensité de la concurrence locale, la densité de la population, et le niveau d'éducation de la population des adultes. La section 2 pose le cadre théorique de l'article. La section 3 dérive le modèle économétrique. La section 4 analyse les résultats.

## **1. Le cadre théorique**

Pour réaliser notre application empirique, nous nous inspirons du cadre développé par Glaeser et al. (1992, 1995) et Henderson et al. (1995). Nous supposons une structure simple de la fonction de production. Les régions sont considérées comme des économies séparées avec un même bassin du capital et de l'emploi (les marchés du travail et du capital sont supposés parfaitement intégrés). Les différences dans les taux de croissance régionale ne proviennent pas des différences dans les taux d'épargne ou dans des différences dans les

---

<sup>5</sup> Département administratif

dotations exogènes en facteurs. A cause de l'hypothèse de la libre mobilité du travail et du capital entre les régions, celles-ci diffèrent seulement dans le niveau de productivité et dans la qualité de la vie.

La fonction de production dans une région ( $i$ ) et une période ( $t$ ) est donnée par :

$$Y_{it} = A_{it} f(l_{it}) = A_{it} l_{it}^{1-\alpha} \quad (1)$$

$$0 < \alpha < 1$$

$$i = \overline{1,48}$$

$$t = \{1998, 2005\}$$

$A_{it}$  représente le niveau de productivité dans la région  $i$  au temps  $t$ . Selon Glaser et al. (1995),  $A$  est interprété d'une manière à permettre la possibilité que les forces sociales, technologiques et politiques agissantes au niveau régional déterminent la productivité globale. Les variables explicatives intégrées dans le modèle sont généralement vues comme affectant ( $A_{it}$ ) la base technologique et le niveau d'efficacité, ou le taux effectif de la croissance technologique (V. Henderson, 2000).  $\alpha$  est un paramètre de production national (n'est pas spécifique à une région  $i$  donnée, il est commun à toutes les régions).

L'absence de données fiables sur le PIB régional ne nous permet pas d'estimer directement la fonction de production. A cet effet, nous utilisons l'approche développée par Glaeser et al. (1992) et Henderson et al. (1995) qui étudient la croissance urbaine par le biais de la croissance de l'emploi. Cette approche repose sur l'idée que les économies d'agglomération améliorent la productivité et entraînent une croissance plus rapide des régions productives. D'après S.S Rosenthal et W.C. Strange (2004), cette approche nous permet d'évaluer indirectement les dynamiques régionales par l'étude de la croissance de l'emploi par région.

L'avantage de l'étude de la croissance régionale par la variable croissance de l'emploi total réside dans le fait que les données sur cette variable sont souvent disponibles, et qu'elle s'accorde mieux avec une spécification linéaire. Cependant, le choix a priori du niveau d'emploi à atteindre par les employeurs est conditionné par le niveau et le type de capital déjà existant dans la région. Le choix d'une période d'analyse suffisante pour une variation significative de l'emploi diminue l'effet des facteurs fixes des régions et renforce l'influence des dynamiques induites par les changements dans la structure initiale de l'investissement et du capital. La simultanéité entre l'effet de la croissance de l'emploi et l'effet de la structure régionale de l'emploi peut causer un problème d'endogénéité dans la spécification du modèle.

En effet, la croissance de l'emploi dans une région donnée est sensible à la composition régionale de l'emploi (effet de l'agglomération), et en même temps, affecte le niveau et la composition de l'emploi. Pour aborder ce problème, l'approche, proposée par Glaeser et al. (1992) et Henderson et al. (1995), consiste à prendre, comme régresseurs, les variables régionales décalées par rapport à la période initiale.

Selon Glaeser (2000), l'analyse de Rappaport (1999) sur les dynamiques d'ajustement résultants de chocs régionaux spécifiques montrent que les résultats de l'utilisation de la variable emploi dans la croissance régionale vont dans la même direction que ceux obtenus en utilisant les variables salaire ou revenu. Glaeser et al. (1995) et Rappaport (1999) montrent que la croissance du revenu et celle de la population tendent à évoluer dans la même direction. De surcroît, la variable population est usuellement très proche de celle de l'emploi (Glaeser, 2000).

Admettre un seul input, dans cette fonction de production, signifie qu'on ne peut pas capturer les innovations technologiques induites par une accumulation supplémentaire du capital physique. Dans chaque région, le niveau technologique, les prix et les salaires sont supposés connus. Si on suppose la condition d'équilibre, la maximisation de la fonction du profit est donnée par :

$$A_{it} f(l_{it}) - w_{it} l_{it} \quad (2)$$

ce qui implique :

$$A_{it} f'(l_{it}) = w_{it} \quad (3)$$

(3) peut être réécrite de la manière suivante :

$$\log \frac{A_{i,t+1}}{A_{it}} = \log \frac{w_{i,t+1}}{w_{it}} - \log \frac{f'(l_{i,t+1})}{f'(l_{it})} \quad (4)$$

Le niveau technologique local ( $A$ ) est supposé intégrer les différentes externalités technologiques présentes dans la région. La variable  $A$ , variable clé dans notre spécification, capture le degré de la productivité locale, les connaissances techniques de production, les avantages et facilités de production locale, et d'autres aspects particuliers à la région  $i$ , tels que les modes d'organisation et la structure industrielle locale.

On peut envisager  $A$  comme une fonction de toutes les caractéristiques spécifiques aux régions.

$$\log \frac{A_{i,t+1}}{A_{it}} = g_{it} (\text{structure industrielle,} \quad (5)$$

conditions initiales et autres variables)

Dans l'équation (5), la structure industrielle mesure le niveau de spécialisation et de la concentration régionale de l'industrie.

Rappaport (1999) montre que le différentiel dans la croissance du salaire, induit par la croissance de la productivité, est compensé en partie par l'augmentation du prix du foncier, due à la croissance de la densité de la population. Sous ces conditions, Glaeser (2000) explique que la croissance du salaire peut être utilisé comme une mesure de la croissance de la productivité locale. Cependant, pour Rappaport (1999), la difficulté de contrôler l'hétérogénéité locale et interrégionale dans la force de travail et le stock du foncier rend l'utilisation du salaire et du prix du foncier difficile. A cet effet, il suggère la variable densité de la population comme une meilleure métrique pour capturer les variations de la productivité et de la qualité de la vie. Suivant cette démarche, nous proposons, pour nos estimations, d'instrumentaliser la variable salaire par celle de la densité de la population<sup>6</sup> par mètre carré. Nous posons alors :

$$\log \left( \frac{w_{i,t+1}}{w_{it}} \right) = h_{it} (\text{variation dans la densité de la population}) \quad (6)$$

On associe aux conditions initiales les variables emploi initial et mobilité de la main d'œuvre. Selon Blanchard et Katz (1992) la différence dans le taux de croissance moyen de l'emploi entre les régions est du à l'immigration, plutôt qu'à la différence dans le taux de croissance naturel de la population. Ils trouvent pour le cas des USA pour les périodes 50 – 87 et 70 – 87 des coefficients de corrélation de l'ordre de 0.84 et 0.91, respectivement.

La combinaison de (4), (5), et (6) donne la forme fonctionnelle suivante :

---

<sup>6</sup> Voir aussi le travail de P.P. Combes et al. (2004) et G. Mion et P. Naticchioni (2005). L'élasticité des salaires par rapport à la densité est largement expliquée par les différences entre niveaux de compétence, et la répartition géographique des compétences est liée à la taille des villes en tant que productrice du savoir.

$$\log \frac{l_{i,t+1}}{l_{it}} = \text{cnst} + \beta * \text{densité de la population} + g_i (\text{structure industrielle, conditions initiales et autres variables}) + \varepsilon_i \quad (7)$$

La prise en compte de la structure industrielle dominante met en évidence les types d'externalités existante dans une région, et le mode d'accumulation des informations sur la productivité et la technologie. Selon la terminologie utilisée par Glaeser et al. (1992), il existe deux types d'externalités technologiques : celles de localisation, dites de type Marshall (1890)-Arrow(1962)-Romer(1986) (MAR) et celles d'urbanisation, dites de type Jacobs (1969). Contrairement aux externalités de localisation, favorables à une accumulation intra industries des connaissances technologiques, les externalités d'urbanisation sont plus propice à une accumulation des connaissances technologiques entre industries. Une région largement spécialisée est plus entraînée par des économies de type MAR qu'une région moins spécialisée. La variable structure industrielle dominante dans une région capte l'impact des externalités technologiques locales sur la vitesse de la croissance de l'emploi régional.

## 2. Croissance régionale de l'emploi et dépendances spatiales

Plusieurs travaux empiriques en économie de la croissance ont explicitement introduit les effets spatiaux dans leurs spécifications. La présence de dépendance spatiale est expliquée par l'existence des externalités entre les régions, ces dernières ne sont pas considérées homogènes et indépendantes. Rey et Montouri (1999) montrent, qu'en dehors du fait des mécanismes théoriques de diffusion technologique, la mobilité des facteurs et le transfert des paiements, utilisés pour justifier la convergence régionale, ont des composantes géographiques explicites. Ignorer le facteur de localisation spatiale peut conduire à des estimateurs biaisés et à des interprétations insuffisantes. Pour corriger ce problème, les suggestions d' Rey et Montouri (1999), Rey et Montouri (99) et López et al.(99) sont largement concordantes avec les hypothèses et les prédictions de la théorie de la croissance endogène et les modèles de la nouvelle économie géographique. Cette dernière montre que l'interaction entre agents économiques conduit à l'agglomération spatiale de l'activité dans un nombre limité de régions (Fujita et al. 1999). Les effets d'externalités sont supposés liés à la taille du marché, l'accès aux services, à l'intensité des relations économiques entre régions, à la diffusion technologique et à la similarité dans les règles institutionnelles et politiques dans les différentes régions. Les interactions entre des entreprises hétérogènes, localisées dans différentes régions, entraînent des régions hétérogènes et interdépendantes.

Bernat (96), Rey et Montouri (99) sont considérés parmi les premiers à avoir explicitement introduit l'effet spatial dans les spécifications économétriques de la croissance. Bernat (96) a testé une version simple des lois de Kaldor sur un ensemble d'Etats Américains. Rey et Montouri (99) ont testé la  $\beta$ -convergence absolue sous dépendance et hétérogénéité spatiale. Ces travaux ont été suivis par une série d'études (Fingleton et McCombie, 1998, López et al., 1999, Fingleton, 1999, et autres). Les spécifications proposées introduisent la dépendance spatiale, soit sous forme d'autocorrélation spatiale des erreurs, ou soit sous forme d'un modèle autorégressif spatial. La sélection d'un de ces modèles est basée principalement sur des critères statistiques proposés par Anselin et Rey (91), et Florax et Folmer (92). En dehors du fait que les externalités et les interactions entre les régions sont la source majeure des dépendances spatiales, ces travaux ont permis d'introduire les effets spatiaux d'une manière ad hoc dans les spécifications empiriques. Fingleton et López (2005)<sup>7</sup> concluent que les évidences empiriques sur une spécification préférée sont mitigées et semblent dépendre des régions et des périodes traitées dans la spécification empirique.

Le choix d'une spécification spatiale correcte (substantive ou nuisance) a des implications différentes sur l'interprétation du processus de la croissance économique. Dans le cas d'un modèle avec autocorrélation spatiale des erreurs, Bernat (1996) explique que la croissance d'une région est affectée par la croissance des régions voisines uniquement à travers l'étendue de l'écart (positif ou négatif) de la croissance des régions voisines par rapport à la croissance moyenne. Dans le modèle autorégressif spatial, la croissance d'une région est directement affectée par la croissance dans les régions voisines et cet effet est indépendant de celui des variables exogènes. Rey et Montouri (1999) expliquent que l'écart par rapport au sentier de croissance équilibrée, pour le cas d'un modèle avec autocorrélations spatiales dans les erreurs, n'est pas uniquement fonction des chocs spécifiques à la région, mais plutôt d'un ensemble complexe de chocs qui débordent. Cependant, dans une spécification avec une variable endogène spatialement décalée, le taux de croissance dans une région peut être lié à celui des régions voisines après conditionnement par le niveau initial du revenu.

---

<sup>7</sup> Fingleton et López (2005) citent les travaux de Armstrong (1995), Bernat (1996), Rey et Montouri (1999), Pon-Novell et Viladecans (1999), Vayá y Moreno (2002), Niebuhr (2001), Kosfeld et al. (2002), Baumont et al. (2003), Arbia et al. (2003), Ying (2003) et Fingleton (2001 et 2004).

Dans cette section, nous organisons notre réflexion en trois temps. Premièrement, nous nous demandons si la croissance de l'emploi par tête dans une région est liée à la croissance de celle des régions voisines. En particulier, cette relation révèle-t-elle une dépendance spatiale ? L'autocorrélation spatiale signifie que l'observation sur une variable dans une région est dépendante de l'observation dans les autres régions de voisinage. Deux sources sont à l'origine de cette autocorrélation. La première provient de la nature de la distribution des activités dans l'espace (Odland, 1988 ; Haining, 1990). En économie géographique, cette distribution est conditionnée par la proximité, la mobilité du travail et du capital. La deuxième peut provenir d'une mauvaise spécification du modèle, comme des variables omises spatialement autocorrélées (Le Gallo 2002). Deuxièmement, nous examinons l'impact des dépendances spatiales sur la convergence des régions en Algérie. Troisièmement, nous recherchons les sources de la dépendance spatiale entre les régions.

Souvent, dans le traitement de données spatiales, nous rencontrons ce qui est actuellement connu par « effet d'enclave ». Ceci signifie qu'une région particulière peut présenter un comportement différent de la majorité des observations spatiales. Cette effet conduit à une distribution asymétrique des erreurs, avec une longue queue (skewed distribution) et plus vraisemblablement suivant une loi de Student. En effet, ce phénomène s'observe par une non constance de la variance des erreurs, aussi bien que la présence d'observations extrêmes (outliers) à travers l'espace.

Dans la littérature empirique, la présence de données aberrantes (ou outliers) affectant les estimations des modèles de convergence ont été remarqués par De Long et Summers (1991) et Temple (1998, 1999). L'hypothèse d'hétéroscédasticité semble plus appropriée que l'hypothèse traditionnelle de Gauss Markov, à savoir que les variances des erreurs sont constantes dans l'espace. Geweke (1993), se basant sur l'approche de Lange, Little et Taylor (1989)<sup>8</sup>, propose un modèle linéaire bayésien hétéroscédastique, il montre que cette approche de modélisation des erreurs est semblable à celle d'un modèle qui suppose une distribution de Student des erreurs. Cette approche a été étendue pour le cas des modèles spatiaux par LeSage (1997, 2000).

Pour prendre en compte l'effet d'enclave sur nos estimations, nous utiliserons une approche bayésienne hétéroscédastique dans notre modélisation. Les modèles bayésien

---

<sup>8</sup> Proposent une méthode de régression non bayésienne basée sur l'hypothèse que les erreurs sont indépendantes et identiquement distribuées selon une loi de Student.

hétéroscédastiques supposent que les variances des erreurs ne sont pas constantes. Ces dernières prennent la forme  $\varepsilon_i \rightarrow N(0, \sigma^2 V)$  où  $V$  est une matrice diagonale contenant les paramètres  $(\nu_1 \ \nu_2 \ \dots \ \nu_n)$ , à estimer par la méthode Markov Chain Monte Carlo (MCMC), et représentant la magnitude des différences dans les variances. Les termes  $(\nu_1 \ \nu_2 \ \dots \ \nu_n)$  permettent de pondérer les observations qui présentent de larges variances. La distribution à priori des termes  $\nu_i$  prennent la forme d'une distribution indépendante de  $\chi(r)/r$ . LeSage (1997) propose, pour l'hyperparamètre  $r$ , de choisir des valeurs à priori variant entre 2 et 7 pour tous les modèles. Si les données ne contiennent pas de valeurs extrêmes ou des variances non constantes, ces valeurs de  $r$  produisent des  $\nu_i$  estimés relativement constantes ou proches de l'unité.

### 3. L'effet de la croissance de l'emploi par tête des régions voisines

Notre première spécification, met en évidence l'influence de la dépendance spatiale (externalité) sans l'introduction de l'effet conditionnel des variables de contrôle. Un modèle autorégressif de premier ordre est donné par :

$$e_i = \rho \sum_{j=1}^n w_{ij} e_j + \varepsilon_i \quad i = 1 \dots n \quad (1)$$

où

$$e_i = \ln \left( \frac{I_{i,05} / N_{i,05}}{I_{i,98} / N_{i,98}} \right)$$

La forme matricielle est donnée par :

$$(1 - \rho W)E = \varepsilon$$

$$\varepsilon_i \rightarrow N(0, \sigma^2 V)$$

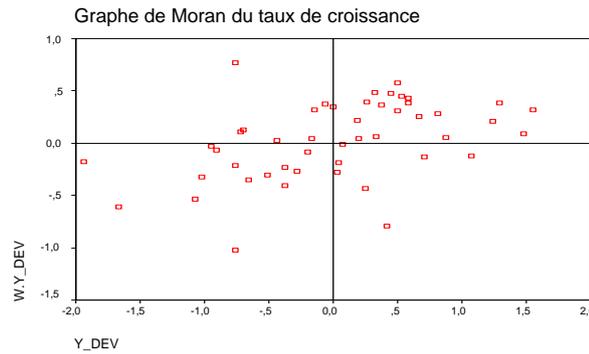
$e_i$  est le logarithme népérien du taux de croissance de l'emploi par tête,  $W$  est la matrice de contiguïté de taille  $(n \times n)$ . Dans cette spécification, la croissance de l'emploi dans une région  $i$  dépend de la moyenne pondérée des taux de croissance des régions voisines.

**Tableau 5 :***Test des externalités de croissance**Variable dépendante : taux de croissance de l'emploi par tête**W matrice de contiguïté*

Paramètre	Modèle homoscédastique	Modèle hétéroscédastique
	r = 100	r = 4
$\rho$	<b>0.4547</b>	<b>0.4379</b>
<i>sign</i>	<b>0.0125</b>	<b>0.0115</b>
$R^2$	<b>0.1825</b>	<b>0.6490</b>

Le tableau (1) présente les estimations pour le cas d'une distribution normale des erreurs avec une variance constante, et les résultats pour le cas de l'introduction de l'effet d'hétéroscédasticité sur les estimations ( $r = 4$ ). Les deux approches confirment la présence d'une dépendance spatiale dans le taux de croissance de l'emploi par tête d'habitant. Il existe un effet d'externalité d'un ordre de 0.43 (modèle hétéroscédastique), ce qui montre l'ampleur de l'effet du taux moyen pondéré de croissance des régions voisine sur la région en considération. Une croissance moyenne de l'emploi par tête de 1 % , pondérée par l'effet de proximité, dans le reste des régions est associées approximativement avec une croissance de l'ordre de 0,43 % dans la région en considération. Notons que cette spécification ne fournit aucune information sur les causes des externalités observées. Elle peut refléter l'effet de chocs dus à un ensemble complexe de facteurs spécifiques aux régions : comme la mobilité de la main d'œuvre, ou tout simplement, le fait que les régions contiguës ont tendance à avoir le même niveau d'éducation des adultes, ou des structures industrielles similaires ou complémentaires. Les résultats du tableau (1) montrent que la croissance dans une région donnée est dépendante de la croissance des régions avoisinantes. En d'autres termes, les régions, avec des niveaux de croissance de l'emploi par tête semblable, ont tendance à se regrouper dans l'espace.

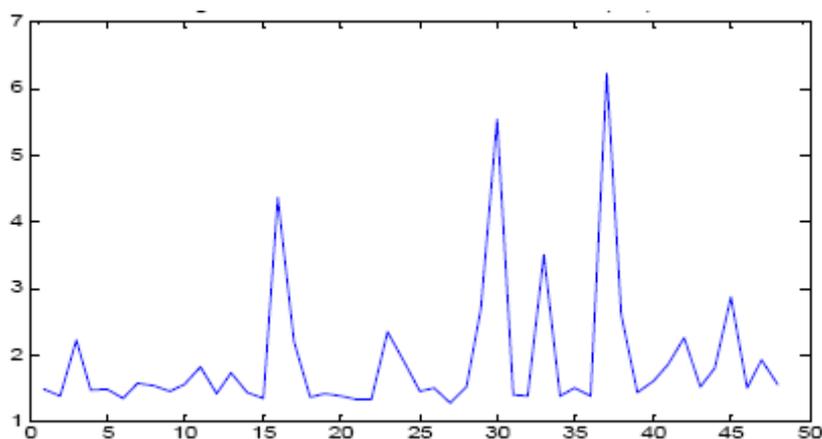
**Figure 8 Graphe de Moran du taux de croissance**



Le graphe de Moran ci-dessus montre que les régions avec des taux de croissance de l'emploi par tête similaires ont tendance à se localiser en proximité. La répartition du taux de croissance dans l'espace semble fortement influencée par l'effet de la proximité.

Les différences entre les deux approches dans les estimations du tableau (1) s'expliquent par la mise en place d'une certaine robustesse dans les estimations. Cette robustesse est due à la prise en compte, dans le modèle hétéroscédastique, de l'effet d'enclave dans les observations.

Pour le cas du modèle (1), la figure (1) met en évidence la non constance des variances où la présence d'observations aberrantes. Celles-ci sont représentées par de fortes valeurs des  $v_i$  estimées avec une valeur de  $r = 4$ .



**Figure 9 : Moyenne des estimations a posteriori de  $v_i$**

### 3.1. Dépendances spatiales et convergence régionale

Nous examinons ici, si le comportement régional de la croissance trouvé reflète la distribution régionale de certains déterminants de la croissance de l'emploi. Pour ce faire, nous estimons une forme économétrique plus générale que celle de la version (1).

$$E = const + \rho WE - (1 - e^{-\beta T}) \ln\left(\frac{I_{98}}{N_{98}}\right) + X\delta + \varepsilon \quad (2)$$

où

$$E = (1 - \rho W)^{-1} \left( const - (1 - e^{-\beta T}) \ln\left(\frac{I_{98}}{N_{98}}\right) + X\delta + \varepsilon \right)$$

Où  $E$  est un vecteur ( $n \times 1$ ),  $W$  matrice de contiguïté ( $n \times n$ ),  $X$  matrice des variables exogènes de taille ( $n \times k$ ),  $\beta$  paramètre de convergence et  $\delta$  vecteur des paramètres de taille ( $k \times 1$ ) et  $\varepsilon \rightarrow N(0, \sigma^2 V)$ . Cette spécification postule que la croissance de l'emploi dans une région  $i$  dépend de la croissance moyenne dans les régions voisines, du niveau initial de l'emploi par tête et d'un ensemble de variables exogènes contenu dans la matrice  $X$ . Cette dernière, reflète les facteurs déterminant les différences dans le comportement de convergence des régions.  $\beta$  mesure la vitesse de convergence entre les régions. Quand  $\beta > 0$  et  $\delta$  un vecteur dont les éléments ne sont pas significatifs, on conclut en faveur d'une  $\beta$  - convergence absolue, par contre dans le cas où  $\beta > 0$  et avec un vecteur dont les coefficients sont significatifs, on parle d'une  $\beta$  - convergence conditionnelle.

L'équation (2) permet de rendre compte de l'effet des externalités dans la croissance des régions quand l'effet des variables spécifiques aux régions est pris en compte. Il est à noter que normalement les variables conditionnelles ne doivent pas être très différentes entre des régions proches (contiguës), leur introduction peut être considérée comme un test de robustesse des résultats obtenus sur les externalités avec l'équation (1). La variable croissance décalée dans (1) peut capturer l'effet des variables régionales omises, spatialement corrélées. Pour réduire au minimum l'effet d'endogénéité, les données sur les variables conditionnelles se réfèrent tous à la période initiale, soit 1998.

Dans (3) ci-dessous, la matrice  $Z$  introduit les facteurs à l'origine des différences dans les taux de croissance entre les régions, elle peut entraîner une autocorrélation spatiale dans les erreurs. En d'autres termes, les externalités entre les régions peuvent s'exprimer à travers les dépendances spatiales dans les erreurs des équations de croissance. L'équation (2) peut ainsi être réécrite de la manière suivante :

$$E = (I - \rho W)^{-1}(Zb + \varepsilon) = (I - \rho W)^{-1} Zb + (I - \rho W)^{-1} \varepsilon \quad (3)$$

Suivant la typologie introduite par Anselin (2003), cette structure peut être associée avec la présence d'externalités globales dans le processus de croissance. La croissance dans chaque région est affectée par l'emploi initial, les variables conditionnelles, et l'ampleur de sa position, déterminée par la matrice  $W$ , dans le système régional et dont l'intensité décroît avec la distance. Ceci est représenté dans (3) par le produit de la matrice  $Z$  et l'inverse de la transformation spatiale  $(I - \rho W)$ . En plus, la croissance dans chaque région est affectée par des chocs aléatoires internes et par des chocs provenant du reste du système des régions, représenté par  $(I - \rho W)^{-1} \varepsilon$ , et dont l'effet décroît avec la distance. Le modèle (3) impose une contrainte importante dans la structure des externalités spatiales, de sorte que le mécanisme de diffusion spatial est typiquement le même dans  $Z$  et dans  $\varepsilon$ .

**Tableau 6 :**

*Test des externalités de la croissance conditionnée par les variables exogènes*

*Variable dépendante : taux de croissance de l'emploi par tête  $W$  matrice de contiguïté pondérée*

Variables	(1)		(2)	
	Modèle homoscédastique $r = 100$	Modèle hétéroscédastique $e$ $r = 4$	Modèle homoscédastique $r = 100$	Modèle hétéroscédastique $r = 4$
Constante	-4.0039***	-4.2487***	-7.8826***	-8.1092***
Log emploi initial par tête	0.3526***	0.3715***	0.4124**	0.3618**
Indice de concentration industrie			0.1602	0.3194**
Indice de concentration BTP et hydrocarbures			0.0693	0.0240
Variation de la Densité de la population			0.0924**	0.1352***
Mobilité de la main d'œuvre			-1.4668**	-2.3127***
AEP			0,9759	1.9141**
Education			4.8238	3.5448
$\rho$	0.3573**	0.3307**	0.4169**	0.4361**
<i>Sign</i>	0.0330	0.0460	0.0250	0.0104
<i>R</i>	0.1723	0.1664	0.3491	0.2187
Statistique I de Moran	2.3410**		2.8253***	
Sign de I - Moran	0.019		0.0047	

\*, \*\* et \*\*\* significatif respectivement à 10 %, 5 % et 1 %

Le tableau (3) synthétise les résultats obtenus pour l'estimation de l'équation (2) avec externalités entre les régions. La première colonne montre que la croissance de l'emploi par tête n'est pas, tel que postulé par l'hypothèse de la convergence, inversement corrélée à l'emploi par tête initial. Ce résultat montre une absence du mécanisme de convergence vers le sentier de croissance équilibrée entraîné, selon la théorie de croissance néoclassique, par la décroissance des rendements dans l'accumulation du capital. Les variables conditionnelles dans la spécification (colonnes 2) contrôlent les facteurs déterminant les différences dans les sentiers régionaux de croissance. La significativité de  $\beta$  et des paramètres associés au reste des variables exogènes met en évidence la nature du phénomène de  $\beta$ -convergence conditionnelle.

L'introduction des variables conditionnelles n'a pas d'effet important sur l'effet de l'emploi initial sur la croissance de l'emploi régional par tête d'habitant. La différence des résultats obtenus par les deux méthodes s'explique par le fait que la méthode d'estimation hétéroscédastique prend en compte l'hétérogénéité spatiale dans les observations, les estimateurs sont plus robustes à la différence des variances et les calculs des estimateurs sont basés sur des observations pondérées par rapport aux variances.

Les résultats dans les colonnes (2) sont étendus pour inclure les variables conditionnelles. Dans les deux cas présentés ci-dessus, les résultats sont obtenus par application de la matrice de contiguïté. Le coefficient,  $\rho$ , mesurant l'étendue des externalités, est estimé significativement différent de zéro dans tous les cas. Par rapport au modèle autorégressif de premier ordre (1), l'introduction des variables conditionnelles ne réduit pas la magnitude de l'étendue des externalités. L'introduction des variables exogènes améliore la significativité de l'effet d'externalité, ceci montre l'existence d'un effet de spillover plus confirmé après introduction de ces variables. La baisse du coefficient  $\rho$  suggère, pour le cas des modèles absolus de la colonne (1), que l'absence du processus de convergence globale entre les régions, réduit l'intensité de dépendance dans la variable croissance de l'emploi par tête.

Notre estimation indique qu'en moyenne la croissance de l'emploi dans les régions étudiées bénéficie au moins de 0,43 % de celle des régions voisines. Il est important de signaler que l'introduction des variables conditionnelles n'affecte pas la robustesse de l'estimation du coefficient le plus important dans notre analyse,  $\rho$ .

Le degré de la concentration de l'activité industrielle passée, mesuré par le rapport de la part de l'emploi industriel dans la région par rapport à la part de l'emploi dans l'industrie dans l'emploi national, affecte positivement la croissance de l'emploi. Le coefficient de cette variable est significatif à un niveau inférieur à 5 %. Au contraire, la présence d'une activité passée dominée par le secteur des hydrocarbures et des BTPH n'a pas d'effet significatif sur la croissance de l'emploi régional par tête. Selon ces estimations, il nous semble que l'existence d'un tissu industriel passé produit des effets d'externalités positives sur la croissance de l'activité. Le secteur des hydrocarbures, sources de valeur importante pour l'économie Algérienne, ne semble pas avoir d'impact significatif sur la croissance de l'activité régionale. La concentration passée de l'industrie crée un environnement attractif pour les investisseurs actuels. Globalement, les externalités de localisation ou de type MAR semble jouer en faveur de la croissance de l'emploi par tête dans les régions étudiées. Une augmentation de 1 % dans le niveau de la concentration industrielle passée d'une région augmente, selon la méthode d'estimation utilisée, le taux de croissance de l'emploi par tête de 0,31 %.

La variation de la densité de la population a un coefficient positif (0,135) et statistiquement significatif. Des recherches récentes reconnaissent l'importance du lien entre la croissance et la densité de la population et des entreprises. Une forte densité spatiale favorise la transmission et la production des idées, ceci crée un environnement stimulant l'innovation et la croissance. Jaffe, Trajtenberg, et Henderson (1993) trouvent que la distance par rapport à la source des idées influence l'intensité de la communication des nouvelles idées, et une forte concentration spatiale de la population et des firmes facilite la transmission de ces idées qui sont à la base de la création de nouveaux biens et des nouvelles méthodes pour la production des biens existants. Rappaport (1999) et Glaeser (2003a) considère la densité de la population comme meilleur indicateur de la croissance de la productivité et de la qualité de la vie. Selon Rappaport (1999), le défaut de la densité de la population est qu'elle ne permet pas de distinguer entre productivité et qualité de la vie. Dans un certain sens, la distinction est de moindre importance pour l'analyse régionale, puisque les deux contribuent positivement à l'utilité (qualité de vie directement et productivité indirectement via les hauts salaires). Cette distinction est importante dans les études entre pays. Comme la qualité de la vie est un bien normal, sa demande augmente avec le revenu, les individus dans les pays sous

développés placent peu de valeur (relativement aux pays développés) dans les attributs de la qualité de la vie<sup>9</sup>.

L'effet de la mobilité interrégionale et de l'éducation n'est pas significatif. La variable AEP (part des ménages connecté à un réseau d'eau potable) mesure l'effort de soutien public apporter à une région donnée. Cette variable est significative et a un impact positif sur la croissance régionale de l'emploi.

### 3.2. Dépendances spatiales et chocs communs régionaux

L'objet de ce test est de jeter plus de lumière sur les sources des liens qui existent entre les régions. Les autocorrélations spatiales dans les taux de croissance de l'emploi par tête peuvent être le résultat d'autocorrélations spatiales entre les déterminants usuels du taux de croissance de l'emploi, tel que l'immigration interrégionale et autres. Ici, nous testons la possibilité que les corrélations observées soient le résultat de chocs communs aux régions géographiques. La spécification à estimer est de la forme :

$$\begin{aligned} E &= Zb + \varepsilon & (4) \\ \varepsilon &= \lambda W\varepsilon + \mu \\ \mu &\rightarrow N(0, \sigma^2 V) \end{aligned}$$

Cette forme implique que la croissance dans une région  $i$  ne dépend pas intrinsèquement du taux de croissance de la région  $j$ , mais plutôt que les régions sont soumises à des chocs communs dont l'intensité diminue avec la distance ou tout simplement s'annule avec l'absence de contiguïté.

L'équation (4) peut être réécrite de la manière suivante :

$$\begin{aligned} E &= Zb + (I - \lambda W)^{-1} \mu & (5) \\ \mu &\rightarrow N(0, \sigma^2 I_n) \end{aligned}$$

---

<sup>9</sup> Selon Blanchard et Katz, l'emploi croît et décroît à un taux qui peut être déterminé par deux situations : dans les régions à forte attractivité de l'emploi, régions où l'offre de l'emploi est positif, l'afflux de la main d'œuvre réduit les salaires, qui à leur tour stimulent l'afflux de nouvelles entreprises et ainsi soutient la croissance de l'emploi. Les régions à forte attractivité des entreprises, régions où la demande du travail est positive, l'afflux des firmes augment les salaires, qui à leur tour stimulent l'afflux de la main d'œuvre, et soutiennent ainsi la croissance

Cette forme montre comment un choc aléatoire affecte la croissance de l'emploi par tête dans la région, et à travers la transformation spatiale, les autres régions. Dans la forme (5), les externalités globales sont associées exclusivement aux chocs aléatoires.

**Tableau 7 :**

*Test du choc commun conditionné par les variables exogènes*

*Variable dépendante : taux de croissance de l'emploi par tête*

*W matrice de contiguïté pondérée*

Variables	(1)		(2)	
	Modèle homoscédastique	Modèle hétéroscédastique	Modèle homoscédastique	Modèle hétéroscédastique
	r = 100	r = 4	r = 100	r = 4
Constante	-4.3096***	-4.4080**	-7.0065***	-7.3830***
log emploi initial par tête	0.3742***	0.3849***	0.3110*	0.3089*
Indice de concentration industrie			0.2047*	0.2772**
Indice de concentration BTP et hydrocarbures			0.0928	0.0537
Variation de la Densité de la population			0.1136**	0.1259**
Mobilité de la main d'œuvre			-1.7320**	-1.8760**
AEP			0.8131	1.2385
Education			5.1993	5.1866
$\lambda$	0.5276***	0.4436**	0.6425***	0.5235***
Sign	0.0060	0.0130	0.0020	0.0092
R	0,2761	0.2649	0.4565	0.4243
Statistique I de Moran	2,3410**		2.8253***	
Sign de I - Moran	0.0192		0.0047	

\*, \*\* et \*\*\* significatif inférieur respectivement à 10 %, 5 % et 1 %

Les résultats obtenus dans les tableaux (3) et (4) fournissent de fortes évidences quand à l'existence de l'effet spatial dans le modèle de convergence conditionnel (colonnes2) ou absolue (colonnes 1). Nos résultats suggèrent une forte autocorrélation dans la croissance de l'emploi par tête. De surcroît, la présence d'une dépendance spatiale significative dans les erreurs implique la propagation sur le reste des régions de chocs aléatoires subissant une région spécifique. L'introduction d'un choc positif de l'erreur d'une région spécifique induit manifestement un large effet relatif (en terme de croissance de l'emploi) sur la région avec

une propagation de ce choc aux régions voisines. La magnitude du choc ralentit avec la distance.

Les paramètres associés avec l'erreur spatiale et la variable dépendante décalée sont toujours significatifs. Ceci confirme l'existence d'une forte dépendance spatiale dans la croissance de l'emploi par tête entre les régions étudiées. Les tests de Moran et les tests du multiplicateur de Lagrange sont fortement significatifs. Or, selon Anselin et Rey (1991), ces tests sont très puissants pour détecter les deux formes de dépendance spatiale.

## **Conclusion**

Dans ce papier nous avons analysé l'importance de la localisation sur la croissance en Algérie. Nous avons trouvé un effet substantiel du débordement sur la croissance en utilisant une variété de spécifications différentes. La spécification basée uniquement sur le taux de croissance de l'emploi par tête (non contrôlée par les variables exogènes) montre une forte dépendance spatiale entre les régions étudiées. Cette spécification nous semble très importante parce que, comme on l'a vu, l'introduction des variables conditionnelles à la spécification standard ne réduit pas l'effet du débordement entre les régions. L'analyse du graphique de Moran montre que l'autocorrélation spatiale est graphiquement très apparente.

L'estimation par l'application de l'approche de l'échantillonnage par MCMC nous a offert un moyen simple de reconsidérer l'hypothèse d'homogénéité dans les erreurs. L'introduction a priori de valeurs informatives pour les paramètres  $\rho$ ,  $\lambda$  et  $\beta$ , mais avec une hypothèse a priori d'hétéroscédasticité des erreurs, permet d'améliorer la robustesse des résultats dans le cas de variances non constantes ou d'observations aberrantes. La méthode bayésienne hétéroscédastique n'exige pas de spécifier une forme pour la variation des variances dans l'espace, elle peut détecter automatiquement la non constance des variances et les observations aberrantes. L'introduction a priori d'informations sur les paramètres, peut être très utile pour maîtriser la multifonctionnalité et les autres problèmes qui peuvent dégrader la précision des estimateurs. Dans certains cas, l'estimation par l'échantillonnage de Gibbs peut donner de bons résultats, quand les problèmes de conditionnement ou d'identification empêchent les méthodes d'optimisation de maximiser la fonction de vraisemblance.

Les résultats obtenus montre l'absence d'un processus de convergence entre les régions Algérienne. En d'autres termes, les régions caractérisées par un passé dans la croissance continuent d'évoluer, et les régions pauvres restent pauvres. A cet effet, la

significativité du coefficient de dépendance spatial nous laissent soupçonner l'existence de club de convergence. La dynamique de croissance entre les régions n'est pas répartie d'une manière équitable. Les secteurs des hydrocarbures et celui des BTPH, n'ont pas d'effet d'externalité sur l'économie locale. La prédominance passée de telles activités, n'entraînent pas une dynamique de croissance locale. Le test I de Moran indique une forte signification de la dépendance spatiale.

## **CHAPITRE V : L'hétérogénéité spatiale dans la croissance de l'emploi régionale en Algérie.**

### **Introduction**

Le phénomène de l'hétérogénéité spatiale se traduit par une instabilité spatiale des paramètres de la régression, les modèles ne sont pas stationnaires dans l'espace. Les paramètres varient systématiquement avec la localisation, ce qui se réfère à une absence de stabilité des comportements ou des relations économiques dans l'espace. Les formes fonctionnelles et les paramètres varient selon leurs localisations et ne sont pas homogènes. Cependant, il est nécessaire de mobiliser des modélisations qui permettent de prendre en compte les caractéristiques particulières de chaque localisation de l'échantillon. L'hétérogénéité spatiale peut aussi se traduire par l'hétéroscédasticité, problème très fréquent dans les estimations en coupe transversal, à cause des variables manquantes ou de toute autre forme de mauvaise spécification. Dans notre situation, les unités spatiales ne sont ni de formes régulières, ni homogènes : les régions peuvent avoir des formes et des airs différents, des niveaux de développement technologique variables, des populations plus ou moins importantes, etc.

Dans les analyses économétriques de la convergence plusieurs travaux ont soulevés les problèmes économétriques induits par l'hétérogénéité. Pour remédier à ce problème, nous introduisant une analyse basée sur la notion de la convergence locale avec une estimation bayésienne du modèle spatial autorégressif local (Bayesian Spatial Autoregressive Local Estimation). Cette approche nous évite le besoin de spécifier l'instabilité des paramètres estimés dans l'espace. En effet, les observations sont partitionnées en traitant chaque région en fonction d'un ensemble de régions avoisinantes. Cette technique rend compte de l'instabilité des paramètres dans l'espace, en prenant en compte toutes les informations dans le reste des observations. La méthode adoptée présuppose que les différences dans les caractéristiques régionales peuvent créer une uniformité dans les structures économiques. Cette hétérogénéité conduit à des estimateurs locaux similaires pour les observations avoisinantes dans l'espace, et donc à des grappes dans la magnitude des paramètres estimés. Nous estimons les paramètres pour chaque région/observation dans l'échantillon, et dont l'examen nous permet de voir leur variation dans l'espace. Cette procédure produit des estimations linéaires locales qui varient systématiquement avec la taille du sous échantillon autour d'une région pour produit des estimations globales quand la totalité des observations de l'échantillon sont incluses. Ainsi, on

peut évaluer la différence entre les estimations linéaires locales, et tenir compte de l'hétérogénéité spatiale des paramètres.

Cette méthode, qui permet d'avoir des estimations à des points variés de l'espace, a été introduite par McMillen (96, 97) sous le nom de régression locale pondérée (LWR), et avec Brunson, Fortheringham et Charlton (1998) sous le nom de régression géographique pondérée (GWR). Ces approches rendent directement compte du problème de l'hétérogénéité spatiale par l'estimation de modèles séparés en utilisant les données avoisinantes de chaque observation. Cette manière de faire, nous permet des estimations de paramètres variables dans l'espace sans spécification de formes pour les variations spatiales.

Pour LeSage (2001) ce type de modèles pose deux problèmes. Premièrement, les modèles linéaires locaux ne tiennent pas compte de la sensibilité à la taille du sous échantillon d'une observation dans les estimations locales. L'incapacité de tenir compte de cette sensibilité, réduit l'utilité des estimations linéaires locales. Deuxièmement, les observations aberrantes dues aux effets d'enclaves spatiales, les changements de tendance ou les variables aberrantes peuvent exercer un impact important sur les estimations linéaires locales et qui pèsent fortement sur les petits échantillons. Comme résultat à ces inconvénients, l'utilisation des méthodes LWR conduit souvent à des estimations très variables dans l'espace. Un troisième problème réside dans le fait que la dépendance spatiale ne disparaît pas avec une utilisation d'un nombre optimal des observations de voisinage. La méthode proposée par Pace et LeSage (2002) (Spatial autoregressive local estimation SALE) et Ertur, et Le Gallo et LeSage (2003) (Bayesian spatial autoregressive local estimation BSALE) pour le cas bayésien, permet de produire des estimateurs non biaisés, en prenant en compte la dépendance spatiale dans les échantillons linéaires locaux.

Le modèle SALE estime un modèle autorégressif spatiale pour chaque observation en utilisant une séquence d'observations de son voisinage spatial. Cette approche ressemble à celle de Brunson et al. (98) qui intègre LWR et le modèle à erreurs autorégressives. Une caractéristique importante de l'estimation spatiale autorégressive locale est que la séquence des estimations obtenue des sous échantillons de tailles croissantes convergent aux estimations limites obtenues par le modèle autorégressif globale. Les estimations du modèle linéaire local convergent aux estimations du modèle des moindres carrés ordinaires. Les modèles SALE est une version des modèles LWR (quand le paramètre spatial autorégressif est ignoré)

## 1- Présentation du modèle BSALE :

Pour adapter dépendance spatiale et hétérogénéité, nous produirons des estimations pour  $n$  modèles, où  $n$  représente le nombre d'observation en coupe instantanée. La forme homoscédastique du modèle d'estimation spatial autorégressive locale est donnée par la forme suivante :

$$U(i)y = \rho_i U(i)Wy + U(i)X\beta_i + U(i)\varepsilon \quad (1)$$

$U(i)$  matrice de pondération diagonale de taille  $n \times n$  de l'observation  $i$ . Cette matrice permet d'affecter un poids à chacune des  $m$  observations voisines à l'observation  $i$ , et zéro pour le reste des observations. Des approches alternatives ont été proposé pour la construction de la matrice des poids  $U$  pour chaque observation  $i$  de la diagonale de la matrice  $U(i)$ . McMillen suggère une fonction de pondération tri cube de la forme suivante :

$$diagU(i) = \left( \left( 1 - \frac{d_i^j}{d_i^m} \right)^3 \right)^3 I_{(d_i^j < d_i^m)}$$

$d_i^j$  est la distance entre l'observation  $j$  et l'observation  $i$ ,  $d_i^m$  représente la distance entre le  $m^{ième}$  voisin le plus proche de l'observation  $i$ ,  $I_{(\cdot)}$  est la fonction indicatrice.  $U(i)y$  est un vecteur de taille  $m \times 1$  des taux de croissance de l'emploi par tête des  $m$  régions les plus proches (en distance euclidienne) de l'observation  $i$ . D'une manière similaire, le produit  $U(i)X$  extrait l'information des variables explicatives des  $m$  régions les plus proches de  $i$ . Le paramètre  $\rho_i$  mesure l'influence de la variable  $U(i)Wy$  sur  $U(i)y$ . On note que quand  $m \rightarrow n$ ,  $U(i) \rightarrow I_n$  et les estimations s'approchent de ceux construites à partir de toutes les  $n$  observations dans le cas du modèle SAR<sup>10</sup>

## 2- Estimation bayésienne spatiale autorégressive locale (Bayesian Spatial Autoregressive Local Estimation BSALE) :

Pour prendre en compte la non constance des variances, nous adoptons une approche Bayésienne. Le modèle est spécifié tel que montré ci dessous :

---

<sup>10</sup> SAR : Modèle spatiale autorégressif de forme :  $y = \rho Wy + X\beta + \varepsilon$

$$\begin{aligned}
U(i)y &= \rho_i U(i)Wy + U(i)X\beta_i + U(i)\varepsilon & (2) \\
U(i)\varepsilon_i &\rightarrow N(0, \sigma_i^2 U(i)V), \\
V &= \text{diag}(v_1 \quad v_2 \quad \dots \quad v_n)
\end{aligned}$$

$W$  est la matrice de contiguïté pondérée d'ordre  $n \times n$  et dont les sommes lignes sont normalisés à l'unité. Ce modèle introduit un ensemble de paramètre à estimer  $(v_1 \quad v_2 \quad \dots \quad v_n)$  pour pondérer les variances. On note que seulement  $m$  paramètres  $v_i$  prennent des valeurs non nulles. Cette approche, introduite par Geweke (1993) sur les modèles de moindres carrés, et généralisée par LeSage (1997) pour les modèles autorégressifs spatiales, rend les estimations robustes à la non constance des variances et à la présence de données aberrantes. Les équations suivantes fournissent les distributions à priori formelles à utiliser dans l'estimation du modèle :

$$\begin{aligned}
\pi(\beta) &\rightarrow N(c, T) & (3) \\
\pi\left(\frac{r}{v_i}\right) &\rightarrow IID \chi^2(r) \\
\pi\left(\frac{1}{\sigma^2}\right) &\rightarrow \Gamma(d, \nu) \\
\pi(\rho) &\rightarrow U[-1, 1]
\end{aligned}$$

Les distributions à priori assignées aux paramètres  $\beta$ ,  $\sigma$  et  $\rho$  dans le modèle prennent la forme de distributions à priori diffuses. Ceci peut s'accomplir, pour le paramètre  $\beta$ , on affectant à priori des valeurs nulles pour le vecteur  $c$  ( $c = 0$ ) et une valeur à priori de  $I_k(1e+10)$  pour la matrice variance covariance  $T$ . Un a priori diffuse pour  $\sigma$  est associé avec des valeurs  $d = 0$  et  $\nu = 0$ . Les informations à priori concernant les scalaires  $v_i$  de pondération des variances prennent la forme d'un ensemble indépendant identiquement distribué suivant une distribution  $\chi^2(r)/r$  où  $r$  est un paramètre de la distribution de  $\chi^2$ . Cette manière de faire va nous permettre d'estimer les  $m$  paramètres non nuls  $v_i$  en ajoutant uniquement un seul paramètre  $r$  au modèle. On utilise la même valeur de l'hyper paramètre  $r$  pour toutes les observation durant l'estimation. La taille optimal  $m$  des observations plausibles peut être obtenue dans l'intervalle :  $\frac{1}{4}n < m < \frac{3}{4}n$ , tel que le nombre d'observations utilisées est supérieur à  $\frac{1}{4}$  de l'ensemble des observations et inférieur à  $\frac{3}{4}$  des observation de l'échantillon. La valeur de  $m$  peut changer en fonction de la taille de l'échantillon utilisée. L'estimation des paramètres est conditionnée par le choix de la valeur de  $m$ . La sensibilité des paramètres estimés peut être examiner par l'analyse

de l'impact de la valeurs prise par  $m$  (la taille des sous échantillons) sur la variation des valeurs des paramètres estimés.

### 3-Estimation du modèle

Les paramètres  $\beta, V, \sigma$  et le paramètre de dépendance spatial  $\rho$  peuvent être estimés par des tirages séquentiels à partir des distributions conditionnelles de ces paramètres. Ce processus est dite échantillonnage conditionnel successif (alternating conditional sampling) ou méthode d'échantillonnage par Chaîne de Markov Monte Carlo (Markov Chain Monte Carlo sampling MCMC).

Pour implémenter cette méthode d'estimation, nous devons déterminer les distributions conditionnelles de chaque paramètre dans notre modèle BSALE. LeSage (1997) propose la distribution conditionnelle de  $\beta$ , sachant  $\rho, \sigma, U(i), V$  et  $m$ , comme suit:

$$\begin{aligned} P(\beta | \rho, \sigma, V, m) &\rightarrow N(\bar{g}, \sigma^2 G) \\ \bar{g} &= (X'U(i)V^{-1}X + \sigma^2 T^{-1})^{-1} (X'U(i)V^{-1}Ay + \sigma^2 T^{-1}c) \\ G &= (X'U(i)V^{-1}X + \sigma^2 T^{-1})^{-1} \\ A &= (I_n - \rho W) \end{aligned} \quad (9)$$

La distribution de  $\beta$  est une normale multivariée, et la distribution conditionnelle de  $\sigma$  prend la forme proposée par Gelman, Carlin, Stern et Rubin (1995) suivante :

$$\begin{aligned} P(\sigma | \beta, \rho, V, m) &\propto (\sigma^2)^{-\left(\frac{m+d}{2}+1\right)} \exp\left(-e'U(i)V^{-1}e + \frac{2v}{2\sigma^2}\right) \\ e &= (I_n - \rho W)y - X\beta \end{aligned} \quad (10)$$

Cette dernière distribution est proportionnelle à une distribution gamma inverse avec les paramètres  $\frac{m}{2} + d$  et  $e'U(i)V^{-1}e + 2v$ . Geweke(1993) montre que la distribution conditionnelle de  $V$  sachant les autres paramètres est proportionnelle à une distribution de type  $\chi^2(r+1)$ . La distribution conditionnelle de chaque terme  $v_i$  peut s'exprimer suivant la forme suivante :

$$P\left(\frac{e_i^2 + r}{v_i} | \beta, \rho, \sigma^2, v_{-i}, m\right) \rightarrow \chi^2(r+1) \quad (11)$$

où  $e_i = U(i)e$  et  $v_{-i} = (v_1 \dots v_{i-1}, v_{i+1} \dots v_n)$ .

La distribution conditionnelle de  $\rho$  est donnée par la forme suivante :

$$P(\rho|\beta, \sigma, V, m) \propto |A|(s^2)^{-(m-k)/2}$$

$$s^2 = (AU(i)y - Xb)' U(i)V^{-1}(AU(i)y - Xb) / (m-k)$$

Il n'existe pas d'algorithme de tirage aléatoire pour la forme de cette distribution ci dessus. Cependant, on peut utiliser une méthode numérique univariée d'intégration de la distribution à posteriori de  $\rho$ . On appliquant une transformation logarithmique à la distribution conditionnelle de  $\rho$  et l'estimateur de Monte Carlo du log du déterminant sur un grille de  $j=1 \dots n$  de valeurs du paramètre de  $\rho$ , allant de -1 à 1 (Barry and Pace (1999)), on peut exprimer  $s^2$  comme un vecteur sur une grille des valeurs affectées à  $\rho$ . Ce qui donne la forme suivante :

$$\begin{pmatrix} \ln(p_1) \\ \ln(p_2) \\ \cdot \\ \cdot \\ \cdot \\ \ln(p_q) \end{pmatrix} \propto \begin{pmatrix} \ln|I_n - \rho_1 W| \\ \ln|I_n - \rho_2 W| \\ \cdot \\ \cdot \\ \cdot \\ \ln|I_n - \rho_q W| \end{pmatrix} - \left(\frac{m-k}{2}\right) \begin{pmatrix} \ln(\phi(\rho_1)) \\ \ln(\phi(\rho_2)) \\ \cdot \\ \cdot \\ \cdot \\ \ln(\phi(\rho_q)) \end{pmatrix}$$

où  $\phi(\rho_j) = e_0' U(i) e_0 - 2\rho_j e_d' U(i) e_0 + \rho_j^2 e_d' U(i) e_d, j=1 \dots q$ , et on utilise  $\tilde{y}$  et  $\tilde{X}$  pour noter les variables  $\sqrt{V^{-1}}y$  et  $\sqrt{V^{-1}}X$ , avec :

$$e_0 = \tilde{y} - \tilde{X}\beta_0$$

$$e_d = W\tilde{y} - \tilde{X}\beta_d$$

$$\beta_0 = (X'V^{-1}X)^{-1} X'V^{-1}y$$

$$\beta_d = (X'V^{-1}X)^{-1} X'V^{-1}Wy$$

Cette méthode d'approche nous conduit à un problème d'intégration numérique simple qui peut être solutionné par la règle de Simpson.

#### 4- L'échantillonneur de MCMC :

Le schéma d'estimation par la méthode MCMC implique une affectation aux paramètres de valeurs initiales arbitraires qu'on note  $\beta^0, \sigma^0, V^0$  et  $\rho^0$  pour une valeur fixe  $m$  du nombre des régions du voisinage de l'observation  $i$ . On tire séquentiellement des distributions

conditionnelles des paramètres dans notre modèle pour l'échantillon des régions de voisinage de taille  $m$  et l'observation  $i$ .

1.  $p(\beta|\sigma^0, V^0, \rho^0, m)$  est une distribution normale multivariée dont la moyenne et la variance, définies par l'équation (9).
2.  $p(\sigma|\beta^1, V^0, \rho^0, m)$  est une distribution de Khi-Deux avec  $m+2d$  degré de liberté tel que montré dans l'équation (10). Notons que l'évaluation de cette densité se fait en prenant en compte la valeur mise à jour  $\beta^1$  de  $\beta$ . La valeur mise à jour  $\sigma^1$  de  $\sigma$  est employée pour mettre à jour le reste des paramètres tirés par une évaluation séquentielle des distributions suivantes.
3.  $p(v_i|\beta^1, \sigma^1, v_{-i}^0, \rho^0, m)$  est obtenue à partir de la distribution de Khi-Deux tel que montré en (11). Pour une meilleure vitesse de calcul, ce tirage s'effectue en vecteur.
4. L'évaluation de la distribution conditionnelle  $p(\rho|\beta^1, \sigma^1, V^1, m)$  se fait par une approche d'intégration numérique tel que montré ci dessus.

Une fois l'étape (4) réalisée, on revient à l'étape (1) on employant les valeurs mises à jour à la place des paramètres initiaux  $\rho^0, \beta^0, V^0$  et  $\sigma^0$ . Dans chaque passage de la séquence on collecte les paramètres tirés pour construire la distribution jointe à posteriori des paramètres du modèle associé avec le sous échantillon des observation de voisinage  $m$  et l'observation  $i$ .

## 5- Données :

Pour réaliser notre application empirique, nous utiliserons le cadre théorique de Glaeser et al. (1992, 1995) et Henderson et al. (1995) déjà développé dans le chapitre IV. Nous supposons une structure simple de la fonction de production. Les régions sont considérées comme des économies séparées avec un même bassin du capital et de l'emploi (les marchés du travail et du capital sont supposés parfaitement intégrés). Les différences dans les taux de croissance régionale ne proviennent pas des différences dans les taux d'épargne ou dans des différences dans les dotations exogènes en facteurs. A cause de l'hypothèse de la libre mobilité du travail et du capital entre les régions, celles-ci diffèrent seulement dans le niveau de productivité et dans la qualité de la vie.

La fonction de production dans une région ( $i$ ) et une période ( $t$ ) est donnée par :

$$\begin{aligned}
 Y_{it} &= A_{it} f(l_{it}) = A_{it} l_{it}^{1-\alpha} & (1) \\
 0 &< \alpha < 1 \\
 i &= \overline{1,48} \\
 t &= \{1998, 2005\}
 \end{aligned}$$

$A_{it}$  représente le niveau de productivité dans la région  $i$  au temps  $t$ . Selon Glaser et al. (1995),  $A$  est interprété d'une manière à permettre la possibilité que les forces sociales, technologiques et politiques agissantes au niveau régional déterminent la productivité globale. Les variables explicatives intégrées dans le modèle sont généralement vues comme affectant ( $A_{it}$ ) la base technologique et le niveau d'efficacité, ou le taux effectif de la croissance technologique (V. Henderson, 2000).  $\alpha$  est un paramètre de production national (n'est pas spécifique à une région  $i$  donnée, il est commun à toutes les régions).

L'absence de données fiables sur le PIB régional en Algérie ne nous permet pas d'estimer directement la fonction de production. A cet effet, nous utilisons l'approche développée par Glaeser et al. (1992) et Henderson et al. (1995) qui étudient la croissance urbaine par le biais de la croissance de l'emploi. Cette approche repose sur l'idée que les économies d'agglomération améliorent la productivité et entraînent une croissance plus rapide des régions productives. D'après S.S Rosenthal et W.C. Strange (2004), cette approche nous permet d'évaluer indirectement les dynamiques régionales par l'étude de la croissance de l'emploi par région.

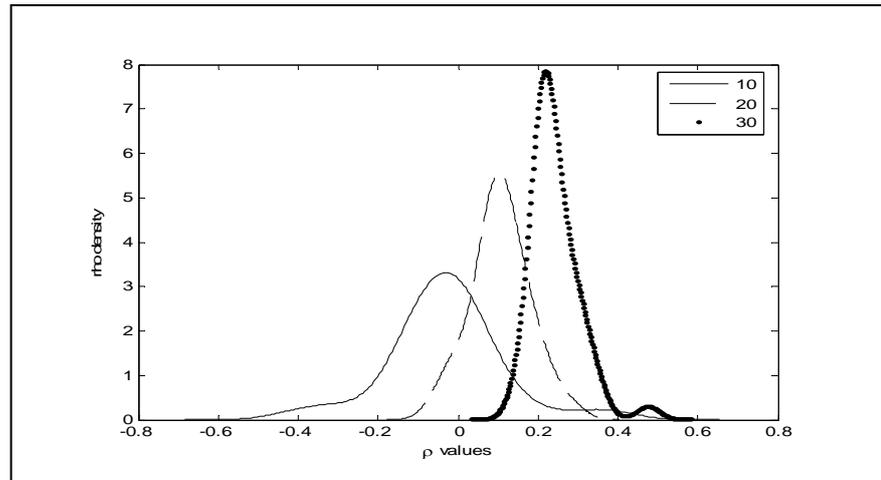
L'avantage de l'étude de la croissance régionale par la variable croissance de l'emploi total a été explicité dans le chapitre IV.

Nous utiliserons les données sur l'emploi par wilaya de travail pour les deux périodes 1998 et 2005. La wilaya représente un niveau de découpage administratif assez intégré pour permettre des analyses sur les dynamiques régionales.

## **6- Résultats et interprétations :**

Les modèles linéaires non paramétriques locaux ont pour objet de réduire l'effet du paramètre de dépendance spatiale  $\rho$ , on se basant sur des échantillons de petites tailles, et où la dépendance spatiale peut être faible ou inexistante. Nous présentons les estimations de la densité Kernel (proportionnelle) de la distribution de 48 estimations du paramètre  $\rho$  du modèle SALE basé sur des sous échantillons des régions de voisinage de taille  $m=10, 20$  et  $30$ .

**Figure 10 : Distribution de rho en fonction de la taille m**



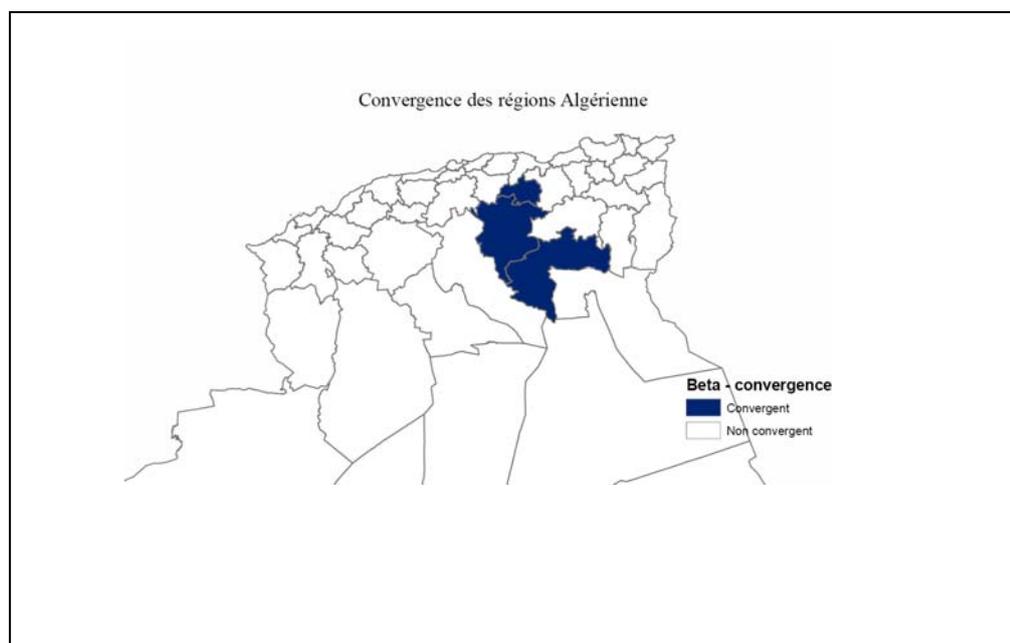
Les trois graphes ci-dessus (fig.1) montre la variation qui existe entre les paramètres de la dépendance spatiale dans notre échantillon de wilayat. En effet, plus la taille des sous échantillons augmente, plus la distribution du paramètre de la dépendance spatiale prend, en moyenne, des valeurs positives et différentes de zéro. Cependant, les valeurs de la dépendance spatiale sont différentes entre les régions, elles sont plus importantes (positives ou négatives) pour certaines régions. De larges valeurs indiquent la présence d'une dépendance spatiale entre les éléments du vecteur de la variable dépendante  $y$ . Dans les modèles spatiaux autorégressifs, des valeurs non nulles de  $\rho$  produisent des effets de biais et de simultanéité sur les estimateurs de  $\beta$ , similaire aux effets de la simultanéité (Anselin, 1988).

La distribution des estimateurs de  $\rho$  pour les sous échantillons de  $q = 20$  et  $q = 30$  montre des valeurs importantes et positives, la dépendance spatiale augmente avec celle de la taille du sous échantillon utilisé. La moyenne des 48 estimateurs de  $\rho$  passe 0,105 à 0,242 quand la taille du sous échantillon utilisé passe de 20 régions à 30 régions.

L'introduction du décalage spatial de la variable dépendante dans le modèle agit comme un proxy permettant de rendre compte des effets spatiaux non pris en compte par les variables explicatives. En outre, ce décalage permet de stabiliser les estimations et d'identifier l'hétérogénéité des structures spatiales.

Les résultats d'estimation des paramètres de convergence, basés sur une matrice de contiguïté d'ordre 3 et une matrice de poids de calibration d'ordre 10 ( $m=10$ ), sont présentés dans la carte ci-dessous :

**Carte 18 : Hétérogénéité dans la convergence des régions Algérienne**



Les différences dans la convergence entre les régions sont apparentes. La non convergence trouvée dans le modèle globale (cf. chapitre) n'est pas une situation générale pour toutes les régions considérée dans notre échantillon, celle-ci est plus affinée par l'étude des différences qui peuvent exister dans un espace qui semble présenter des structures hétérogènes. La convergence, indiquée par les coefficients de la variable log - emploi initial par tête de signes négatifs et significatifs, est vraisemblable pour les trois régions du Centre – Est du pays : Bordj Bou-arreridj, M'sila et Biskra. Ces trois régions tendent à former un club de convergence de comportement différent de celui du reste des régions Algériennes. L'utilisation d'un SAR globale ne montre pas ce type de distinctions. L'estimation du paramètre de convergence conduit à une inférence de la convergence conditionnelle globale, à la différence au modèle SAR hétéroscédastique. Le concept de la convergence locale développée, dans notre cas, par le modèle BSALE fournit des informations supplémentaires sur la présence de la convergence dans la croissance régionale en Algérie. Les estimations obtenues par le modèle BSALE suggèrent que la convergence régionale en Algérie est localisée dans un espace spécifique, ce qui n'est pas le cas pour le reste des régions du pays.

## **6- Conclusion :**

Les résultats obtenus montrent que les problèmes posés, par les changements de régimes quand on passe d'une région spatiale à une autre, dans les régressions conventionnelles peuvent

être adapté par l'estimation autorégressive locale. L'application de la méthodologie du modèle BSALE (Bayesian spatial autoregressive local estimation) montre clairement l'effet de l'hétérogénéité spatiale sur l'estimation globale du paramètre de la convergence régionale de l'emploi par tête en Algérie. En effet, la non convergence globale, obtenu par l'application du SAR globale, omit sa présence pour certaines régions. La tendance globale n'est pas une évidence quelque soit la région de l'espace global étudié.

La méthode BSALE est une approche bayésienne robuste de la version SALE (spatial autoregressive local estimation) exposée par Pace et LeSage (2002). Ce type de méthode d'estimation locale, basées sur des sous échantillons d'individus, produit des estimateurs qui convergent vers ceux obtenus par les modèles bayésien autoregressives quant la taille des sous échantillons d'individus tendent vers la taille globale de la population utilisée. En effet, l'intérêt de cette méthode réside dans le rôle de des estimateurs locaux dans la mise en évidence du caractère de la convergence locale et des clubs de convergence.

## CONCLUSION GENERALE

Ce travail nous a permis de relater les principes fondateurs et les différents travaux de recherche sur la question de la convergence des économies. Nous avons vu que le premier fondement de la convergence se base sur des développements issus du modèle néoclassique de Solow. Le concept de convergence est différencié selon que les états réguliers des économies étudiées sont identiques ou non. Si les économies ne convergent pas vers le même état régulier, on parle de  $\beta$ -convergence conditionnelle, sinon la convergence est dite absolue. Les modèles de croissance endogène, se basant sur le rôle des diffusions technologique dans le processus de rattrapage entre les économies permettent de donner une explication théorique au concept de club de convergence développé par plusieurs chercheurs empiristes.

La définition de la convergence n'échappe pas à certaines faiblesses et limites. En effet, elle ne permet pas de faire la distinction entre l'effet de long terme des conditions initiales et de l'hétérogénéité structurelle de long terme. A la différence des effets des caractéristiques structurelles, les effets de conditions initiales sur l'état régulier impliquent l'existence de club de convergence. La convergence absolue est souvent conditionnée par l'homogénéité de l'échantillon étudié, généralement la corrélation entre le revenu initial et la croissance n'est pas vérifiée dans le cas de pays hétérogènes. L'estimation  $\beta < 0$  n'est pas consistante avec des alternatives économiques non convergentes. En effet, la convergence statistique peut coexister avec la présence de plusieurs états réguliers ce qui confirme la compatibilité de l'existence de sentier de croissance avec des différences permanentes de revenus entre économies dotées de même taux de croissance de la population, de l'épargne et d'un même accès à la technologie.

Une autre limite des approches néoclassiques de la croissance réside dans la négligence du rôle de l'espace dans la répartition et la localisation des activités économiques. Les théories de l'économie géographique apportent des réponses aux causes de la concentration spatiale des activités économiques et expliquent les interactions qui existent entre l'agglomération et la croissance. Dans la section croissance et géographie économique nous avons montré le rôle des forces qui agissent pour l'agglomération des activités industrielles et de RD, et ceux qui agissent pour la dispersion. Ces modèles mettent en avant l'effet de la taille du marché (Home Market Effect) et celle de la configuration centre - périphérie. Le niveau des coûts de transport entre les régions est un facteur important dans

l'explication du schéma d'équilibre dominant dans la répartition des activités économiques, et des dynamiques de croissance. En effet, les modèles de synthèse, entre la théorie de la croissance endogène et l'économie géographique, montrent que la diffusion des externalités technologiques dans un espace géographique restreint rend la polarisation spatiale des activités plus bénéfique pour la croissance que la dispersion. La croissance économique peut en retour renforcer les processus de polarisation. Enfin, les effets d'externalités technologiques entre les régions renforcent la polarisation si elles sont locales et conduisent à une répartition plus équilibrée entre les régions si elles sont globales.

Ces résultats théoriques montrent l'importance de la prise en compte des facteurs déterminants la répartition spatiale des activités dans le processus de croissance des économies. Ces effets doivent être considérés dans l'analyse du processus de convergence. L'hétérogénéité spatiale doit être introduite dans l'analyse et la mesure de la convergence, puisque l'existence d'une configuration centre périphérique est plus adaptée avec les clubs de convergence. Les données sur les économies ne sont pas aléatoirement distribuées dans l'espace et les phénomènes d'auto corrélation spatiale rendent, s'ils ne sont pas considérés, inconsistantes les estimations de la convergence. Les méthodes de l'économétrie spatiale doivent prendre en compte les effets dus à l'hétérogénéité spatiale dans la distribution de la croissance. Nous consacrons le chapitre suivant au développement des concepts de l'économétrie spatiale utilisés pour l'étude de la convergence régionale de l'emploi en Algérie.

A la différence des techniques économétriques non spatiales, l'économétrie spatiale nous permet d'introduire l'effet spatial soit sous forme de variable endogène décalée et/ou sous forme d'autocorrélation spatiale dans les erreurs. Plusieurs méthodes d'estimation sont possibles. Les plus célèbres sont celle de la méthode du maximum de vraisemblance, et l'approche bayésienne. Des tests statistiques permettent de vérifier et de rendre compte de l'existence d'une corrélation spatiale dans les données.

L'hétérogénéité spatiale peut se présenter par une instabilité des paramètres dans l'espace. Dans un tel cas, il faut spécifier des régimes spatiaux et où les paramètres dépendent de la localisation spatiale des données.

La méthode du maximum de vraisemblance peut être coûteuse en terme de temps de calcul du Hessien avec des échantillons de tailles importantes. De surcroît, dans les modèles d'économétrie spatiale, la normalité des erreurs est souvent remise en cause. L'étude des modèles spatiaux, notamment ceux de l'économie géographique, nécessite l'introduction de

techniques d'estimations aptes à prendre en compte des phénomènes complexes tels que : les variations des paramètres dans l'espace, l'asymétrie, l'incertitude dans les mesures de distance ou de proximité, etc. Les méthodes bayésiennes nous permettent de tenir compte à la fois des problèmes d'hétérogénéités, des observations aberrantes et de mieux considérer les fondements économiques dans la mesure des interactions spatiales.

La simulation d'une variable aléatoire de loi  $\pi$  est un outil numérique très utile pour résoudre des problèmes qui n'ont pas de solution analytique évidente. La simulation par les méthodes de Monte Carlo par Chaîne de Markov, repose sur l'utilisation de la dynamique d'une chaîne de Markov ergodique de loi invariante.

Les techniques de l'économétrie spatiale permettent d'améliorer les estimations produites par les MCO, de détecter la présence des effets de débordement géographiques sur les comportements économiques et d'évaluer l'impact de ces effets à travers l'estimation des paramètres spatiaux.

L'analyse macro économique a montré que la croissance de l'économie algérienne est dépendante de facteurs fortement volatiles. Les prix des hydrocarbures et la pluviométrie ne peuvent pas à eux seules assurer une croissance structurellement stable.

Les dynamiques économiques locales en Algérie sont la résultante de la répartition de l'effort public pour stimuler l'investissement à travers les infrastructures et la différenciation des avantages, et d'un autre côté la transcription actuelle marquée par de fortes disparités économiques. Les potentialités de développement sont fortement différentes entre les régions (wilayates). L'importance de l'espace Algérien et sa forte hétérogénéité géographique exige le passage d'une logique de gouvernance centralisée et souvent loin des réalités des besoins locaux vers une logique décentralisée basée sur une optique de compétitivité des territoires. La mise en valeur des capacités de développement locales doit se faire en fonction des moyens possibles de chaque espace, et le soutien de l'Etat doit prendre en compte le fait qu'on ne peut pas réaliser un équilibre régional en offrant les mêmes possibilités quelque soit la région considérée.

La description de l'espace Algérien montre une forte différenciation spatiale. Les potentialités de développement et l'émergence de clubs économiques convergent ne peuvent pas être les mêmes quelque soit le lieu de l'espace étudié.

Notre analyse de l'espace économique Algérien nous a permis de trouver un effet substantiel du débordement sur la croissance en utilisant une variété de spécifications différentes. La spécification basée uniquement sur le taux de croissance de l'emploi par tête (non contrôlée par les variables exogènes) montre une forte dépendance spatiale entre les

régions étudiées. Cette spécification nous semble très importante parce que, comme on l'a vu, l'introduction des variables conditionnelles à la spécification standard ne réduit pas l'effet du débordement entre les régions. L'analyse du graphique de Moran montre que l'autocorrelation spatiale est graphiquement très apparente.

Les résultats obtenus montre l'absence d'un processus de convergence entre les régions Algérienne. En d'autres termes, les régions caractérisées par un passé dans la croissance continuent d'évoluer, et les régions pauvres restent pauvres. A cet effet, la significativité du coefficient de dépendance spatial nous laissent soupçonner l'existence de club de convergence. La dynamique de croissance entre les régions n'est pas répartie d'une manière équitable. Les secteurs des hydrocarbures et celui des BTPH, n'ont pas d'effet d'externalité sur l'économie locale. La prédominance passée de telles activités, n'entraînent pas une dynamique de croissance locale. Le test I de Moran indique une forte signification de la dépendance spatiale.

Les résultats obtenus montrent que les problèmes posés, par les changements de régimes quand on passe d'une région spatiale à une autre, dans les régressions conventionnelles peuvent être adapté par l'estimation autorégressive locale. L'application de la méthodologie du modèle BSALE (Bayesian spatial autoregressive local estimation) montre clairement l'effet de l'hétérogénéité spatiale sur l'estimation globale du paramètre de la convergence régionale de l'emploi par tête en Algérie. En effet, la non convergence globale, obtenu par l'application du SAR globale, omit sa présence pour certaines régions. La tendance globale n'est pas une évidence quelque soit la région de l'espace global étudié.

## ANNEXE 1 : Modèle spatial bayésien hétéroscédastique :

Nous considérons uniquement la solution du cas d'un modèle Bayésienne autorégressif (SAR). Pour faciliter le calcul de la distribution à priori du paramètre de dépendance spatiale  $\rho$ , nous considérons, en premier temps, le cas avec une distribution normale des erreurs et une variance constante. La fonction du maximum de vraisemblance s'écrit de la forme suivante :

$$L(\beta, \sigma, \rho, y, X) = (2\pi\sigma^2)^{-n/2} |A| \exp\left\{-\frac{1}{2\sigma^2} (Ay - X\beta)' (Ay - X\beta)\right\} \quad (A1)$$

La densité à priori est de la forme  $\pi(\beta, \sigma, \rho) = \pi(\beta) \times \pi(\sigma) \times \pi(\rho)$ . La fonction de vraisemblance peut être combinée avec les densité à priori de  $\beta$  et  $\sigma$ , en prenant la forme proposée par Jefferys (1939) soit,  $\pi(\beta, \sigma) \propto \sigma^{-1}$  ( $\pi(\beta) \propto \text{cnst}$  et  $\pi(\sigma) \propto \sigma^{-1}$ ), et une densité à priori générale pour le paramètre  $\rho$ , qu'on note  $\pi(\rho)$ . Le produit de la vraisemblance et les densités à priori donne la forme kernel de la distribution à posteriori de tous nos paramètres suivante :

$$p(\beta, \sigma, \rho | y, X) \propto \sigma^{-(n+1)} |A| \exp\left\{-\frac{1}{2\sigma^2} (Ay - X\beta)' (Ay - X\beta)\right\} p(\rho) \quad (A2)$$

On utilisant la propriété inverse de la distribution gamma, on peut intégrer la forme (A2) par rapport à  $\sigma$ , on obtient :

$$\begin{aligned} p(\beta, \rho | y, X) &\propto |A| \left\{ (Ay - X\beta)' (Ay - X\beta) \right\}^{n/2} p(\rho) \\ &= |A| \left\{ (n-k)s^2(\rho) + (\beta - b(\rho))' X'X(\beta - b(\rho)) \right\}^{n/2} p(\rho) \end{aligned} \quad (A3)$$

où

$$\begin{aligned} s^2(\rho) &= (Ay - Xb(\rho))' (Ay - Xb(\rho)) / (n-k) \\ b(\rho) &= (X'X)^{-1} X'Ay \\ A(\rho) &= A = (I_n - \rho W) \end{aligned}$$

L'intégration de la forme (A3) par rapport à  $\beta$  nous donne la distribution marginale à posteriori de  $\rho$  suivante :

$$p(\rho|y, X) \propto |A(\rho)| (s^2(\rho))^{-(n-k)/2} p(\rho)$$

L'expression formelle d'un modèle spatial bayésien autorégressif, auquel on ajoute une distribution a priori de type normal - gamma conjuguée pour les paramètres  $\beta$  et  $\sigma$ , une distribution a priori uniforme pour le paramètre  $\rho$  en plus d'une distribution à priori de khi - deux pour les paramètres de la matrice  $V$ , se présente sous la forme ci-dessous :

$$\begin{aligned} y &= \rho W y + X \beta + \varepsilon \\ \varepsilon &\rightarrow N(0, \sigma^2 V) & V &= \text{diag}(v_1 \dots v_n) \\ \pi(\beta) &\rightarrow N(c, T) \\ \pi\left(\frac{r}{v_i}\right) &\rightarrow \text{iid } \chi^2(r) & r &\text{ hyperparamètre} & (A4) \\ \pi\left(\frac{1}{\sigma^2}\right) &\rightarrow \Gamma(d, \nu) \\ \pi(\rho) &\rightarrow U[0,1] \end{aligned}$$

Fixer à priori  $c$  à 0 et  $T$  avec une grande valeur résulte en un a priori diffuse pour  $\beta$ . Des valeurs de  $d = 0$  et  $\nu = 0$ , implique un a priori diffuse pour  $\sigma$ .

Gewek (1993), en modélisant les outliers et l'hétéroscédasticité, montre que cette stratégie de choix de priori est équivalente à l'hypothèse que les erreurs dans un modèle de régression linéaire avec une distribution iid de Student avec  $r$  degrés de liberté. Ce qui est consistant avec le fait que quand  $r \rightarrow \infty$  la distribution de Student approche la distribution normale.

L'introduction des termes  $v_i$  est motivée par le fait que  $E\left(\frac{1}{v_i}\right) = 1$  et  $V\left(\frac{1}{v_i}\right) = \frac{2}{r}$ . Ce qui implique que quand  $r$  prend de grandes valeurs, les termes  $v_i$  s'approche de l'unité, ce qui résulte un  $V = I_n$ . Autrement dit, de faibles valeurs de  $r$  conduit à une distribution asymétrique avec une longue queue (skewed distribution) ou les larges valeurs de  $v_i$  s'éloignent de la valeur centrale (moyenne unitaire à priori). Le rôle des valeurs de  $v_i$  consiste à accommoder les observations extrêmes et celles avec de larges variances par la pondération des observations.

La densité à posteriori de l'ensemble des paramètres, telle que dans (A2), n'est pas commode pour l'analyse. Cependant, on peut dériver la distribution à priori de chaque

paramètre dans notre modèle et utiliser la méthode d'échantillonnage (Markov Chain Monte Carlo)<sup>11</sup> pour l'estimation.

La méthode de Monté Carlo Chaîne de Markov (MCMC) ce base sur l'idée que, au lieu de calculer la densité à posteriori telle que celle de (A2), il suffit d'avoir un grand échantillon aléatoire tiré des distributions à posteriori de chaque paramètre :  $p(\theta|y)$ ,  $\theta$  paramètre et  $y$  base de données Si l'échantillon tiré de  $p(\theta|y)$  est assez large, on peut avoir une forme approximative de la densité de probabilité en utilisant les estimateurs de la densité de kernel, sans avoir à connaître la forme analytique exacte de la density.

### Distributions conditionnelles :

Suivant la méthodologie bayésienne usuelle, on peut combiner la fonction de vraisemblance avec les distribution à priori de  $\beta, \sigma, V$  et  $\rho$  pour obtenir la distribution à posteriori de notre modèle. Cette distribution synthétise l'ensemble des informations, à priori et de la base de données, que nous avons sur les paramètres inconnus  $\beta, \sigma, V$  et  $\rho$ .

Pour implémenter cette méthode d'estimation, nous devons déterminer la distribution marginale à posteriori de chaque paramètre.

- La distribution conditionnelle de  $\beta$  sachant connu les paramètres  $\sigma, V$  et  $\rho$  est donnée par la forme suivante :

$$\begin{aligned}
 p(\beta|\rho, \sigma, V) &\rightarrow N(\bar{\beta}, \sigma^2 B) \\
 \bar{\beta} &= (X'V^{-1}X + \sigma^2 T^{-1})^{-1} (X'V^{-1}(I_n - \rho W)y + \sigma^2 T^{-1}c) \quad (A5) \\
 B &= (X'V^{-1}X + \sigma^2 T^{-1})^{-1}
 \end{aligned}$$

Notons que cette forme est très semblable à la version des moindres carrée généralisée de la formule d'estimation de Theil Goldberger (1961) et à la densité conditionnelle obtenue par Zellner (1971). La distribution conditionnelle de  $\beta$  suit une distribution normale multivariée, à partir de laquelle il est facile d'échantillonner le vecteur  $\beta$ .

- La distribution conditionnelle de  $\sigma$  sachant les autres paramètres, prend la forme suivante :

---

<sup>11</sup> Cette méthode se réfère aussi comme méthode d'échantillonnage de Gibbs

$$p(\sigma^2 | \beta, \rho, V) \propto (\sigma^2)^{-\left(\frac{n}{2} + d + 1\right)} \exp\left[-e'V^{-1}e + \frac{2\nu}{\sigma^2}\right] \quad (A6)$$

$$e = (I_n - \rho W)y - X\beta$$

Cette forme est proportionnelle à la distribution inverse de gamma avec comme paramètres  $(n/2) + d$  et  $e'V^{-1}e + 2\nu$  (soit une  $\chi^2_{(n+2d)}$ ). De plus, il est facile d'échantillonner le paramètre  $\sigma$  à partir de cette distribution.

- Geweke (1993) montre que la distribution conditionnelle de  $V$  sachant les autres paramètres est proportionnelle à une densité de  $\chi^2$  avec  $r+1$  degrés de liberté. Spécialement on peut exprimer la distribution conditionnelle de chaque terme  $v_i$  comme :

$$p\left(\frac{\sigma^{-2}e_i^2 + r}{v_i} \mid \beta, \rho, \sigma^2, v_{-i}\right) \rightarrow \chi^2_{(r+1)} \quad (A7)$$

$v_{-i} = (v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$  pour chaque  $i$ . Cette forme représente celle d'une distribution connue de laquelle il est facile d'échantillonner.

- La distribution conditionnelle à posteriori de  $\rho$  prend la forme

$$\begin{aligned} p(\rho | \beta, \sigma, v) &\propto |A(\rho)| (s^2(\rho))^{-(n-k)/2} p(\rho) \\ s^2(\rho) &= (Ay - Xb(\rho))' V^{-1} (Ay - Xb(\rho)) / (n - k) \\ b(\rho) &= (X'X)^{-1} X'Ay \\ A(\rho) &= A = (I_n - \rho W) \end{aligned} \quad (A8)$$

Pour échantillonner de la forme (A8), LeSage (1997, 2000) propose la méthode dite « Metropolis - within - Gibbs »

## ANNEXE 2 : Résultats d'estimation par l'application du modèle BSALE

Bayesian spatial autoregressive local estimation  
 Dependent variable = depvar  
 Nobs, Nvars = 48, 8  
  
 R-squared = 0.9683  
 Rbar-squared = 0.9628  
 ndraws,nomit = 1100, 100  
 r-value = 4.0000  
 q-nearest = 10

Posterior Estimates											
Obs	Const.		indus		Hydo/btp		Empl_98		Densité		
	Coef.	p-level	Coef.	p-level	Coef.	p-level	Coef.	p-level	Coef.	p-level	
1	-15.4575	0.6229	-0.5613	0.6523	0.0596	0.9688	0.4023	0.7312	0.2015	0.7628	
2	11.0365	0.5974	<b>-1.5016</b>	<b>0.0666</b>	<b>0.8604</b>	<b>0.0762</b>	0.4442	0.3869	-0.5384	0.1560	
3	-9.8467	0.8035	0.5496	0.7533	0.1085	0.8943	0.2847	0.7554	0.1032	0.8436	
4	-0.6874	0.9826	0.7341	0.5181	-0.0498	0.9397	-0.8498	0.8015	0.3687	0.4247	
5	5.1518	0.7483	0.1874	0.8330	-1.2959	0.2833	-1.5784	0.3341	0.0120	0.9685	
6	7.2576	0.8487	-0.2842	0.8706	-0.4239	0.6458	-1.1152	0.4788	0.1508	0.7523	
7	3.4877	0.7541	0.8527	0.7684	-0.5353	0.6887	<b>-0.9907</b>	<b>0.0447</b>	<b>0.1811</b>	<b>0.0472</b>	
8	-5.6280	0.8232	-0.8972	0.6462	-2.0400	0.6195	0.6507	0.4908	-0.1313	0.8467	
9	-46.5385	0.0012	-0.3641	0.2400	0.2087	0.8343	<b>2.4341</b>	<b>0.0041</b>	0.5851	0.1788	
10	34.2713	0.5139	0.4743	0.5415	-0.4693	0.5202	-1.6800	0.5462	-0.1153	0.7955	
11	-17.8360	0.2231	-0.5134	0.1312	0.1448	0.6273	0.6831	0.4705	0.3660	0.2567	
12	-14.5455	0.8776	0.5523	0.7684	-1.0569	0.1407	0.6311	0.9475	<b>0.3738</b>	<b>0.0575</b>	
13	-50.5524	0.3352	-3.0347	0.2394	-3.3614	0.3475	3.5395	0.1187	1.6164	0.4902	
14	6.7065	0.0010	<b>-1.3780</b>	<b>0.0000</b>	<b>0.7323</b>	<b>0.0000</b>	<b>0.4771</b>	<b>0.0000</b>	<b>-0.4609</b>	<b>0.0000</b>	
15	21.3750	0.5986	0.4935	0.6429	-0.5112	0.5864	-0.9940	0.5760	-0.0126	0.9779	
16	-47.1497	0.0000	<b>-0.3711</b>	<b>0.0169</b>	0.2107	0.5176	<b>2.4689</b>	<b>0.0000</b>	<b>0.5861</b>	<b>0.0000</b>	
17	4.5387	0.8794	0.8868	0.6864	-0.1659	0.8903	-0.5427	0.8908	-0.0088	0.9806	
18	-9.7294	0.6104	0.1456	0.8401	0.0724	0.8682	-0.5984	0.6824	0.5752	0.4710	
19	-5.2201	0.7598	0.5213	0.4789	-0.2066	0.7748	-0.6898	0.2594	0.5081	0.1341	
20	3.4910	0.9638	-2.8145	0.7242	-2.2367	0.8804	1.7742	0.8283	-0.7827	0.7001	
21	-25.3107	0.6623	0.2361	0.8375	-0.3865	0.7621	1.8699	0.7836	0.9999	0.6908	
22	-18.0397	0.7115	-3.1435	0.3182	-4.2713	0.3666	3.0637	0.2419	-0.0419	0.9855	
23	-32.6171	0.4496	0.2214	0.7759	-0.5738	0.5583	2.8028	0.5750	1.2006	0.5247	
24	-32.1123	0.0236	<b>0.5478</b>	<b>0.0498</b>	<b>-0.6885</b>	<b>0.0234</b>	<b>2.8314</b>	<b>0.0871</b>	<b>1.8351</b>	<b>0.0060</b>	
25	2.2229	0.9568	0.4601	0.6446	0.1961	0.7551	-2.1308	0.6697	1.0021	0.5531	
26	-42.1381	0.0884	-0.5325	0.5555	0.6694	0.7433	2.1052	0.1255	0.3923	0.3531	
27	14.4592	0.0841	<b>2.3334</b>	<b>0.0019</b>	0.0839	0.9329	<b>1.5683</b>	<b>0.0062</b>	<b>-1.4452</b>	<b>0.0000</b>	
28	3.0429	0.7885	0.0089	0.9893	-0.6285	0.2975	<b>-1.1246</b>	<b>0.0554</b>	0.1437	0.1042	
29	11.3173	0.6415	-2.4898	0.1449	-0.0070	0.9979	1.7748	0.3107	-1.3548	0.1495	
30	-8.5638	0.0613	<b>-0.8923</b>	<b>0.0458</b>	0.2214	0.2469	0.0776	0.7933	-0.0824	0.6173	
31	50.5014	0.8457	-2.9897	0.8051	-5.3867	0.8175	2.1929	0.7806	-3.8450	0.8071	
32	3.7062	0.5952	0.3092	0.6445	0.1339	0.7529	0.2736	0.8229	<b>-0.1897</b>	<b>0.0142</b>	
33	-4.9785	0.9510	-0.7626	0.6088	0.1577	0.9638	-0.2486	0.9784	-0.1171	0.9647	
34	11.9577	0.3688	-0.4640	0.6831	<b>-0.7886</b>	<b>0.0709</b>	<b>-1.3029</b>	<b>0.0410</b>	0.0407	0.7788	
35	-45.0628	0.5586	-0.1300	0.8315	-0.3319	0.4374	2.4503	0.5386	0.6936	0.3578	
36	-13.6267	0.8412	0.5531	0.6876	-0.2479	0.8929	0.4644	0.9566	0.4499	0.4576	
37	-6.4953	0.8320	-1.2689	0.3357	-3.3065	0.3412	0.5069	0.4927	-0.1801	0.8099	
38	7.9114	0.6962	0.7642	0.3838	1.0143	0.1557	0.9235	0.3090	-0.6698	0.1221	
39	-12.8630	0.8203	-0.4741	0.9186	0.1644	0.9046	0.5299	0.9124	-0.1799	0.8300	
40	11.6987	0.5698	0.9119	0.2825	-0.4131	0.5945	-2.2059	0.2958	0.3552	0.1417	
41	-8.7496	0.7425	0.5767	0.4351	-0.1380	0.7839	-0.0207	0.9942	0.3711	0.1728	
42	-28.5725	0.4978	-0.4104	0.4640	0.4436	0.6791	1.5073	0.4286	0.0158	0.9907	
43	-15.2064	0.7237	-0.2350	0.8902	0.2781	0.7985	-1.7133	0.6882	1.0900	0.6459	
44	-22.4569	0.2010	-0.4179	0.3729	0.5441	0.3919	1.1531	0.2047	-0.1646	0.7321	
45	-4.3859	0.6374	1.6493	0.3271	-0.5387	0.4930	1.0135	0.3606	0.0252	0.9415	
46	36.8540	0.5470	-2.5941	0.2886	-4.3424	0.3141	2.1623	0.2596	-2.9801	0.3707	
47	17.8846	0.8481	-0.0759	0.9516	0.4187	0.2116	-1.1250	0.8466	-0.2517	0.6995	
48	16.270	0.5324	-2.2741	0.4604	-0.0376	0.9893	1.5278	0.3030	<b>-1.4617</b>	<b>0.0469</b>	

Suite

Posterior Estimates

obs	Mobilité.		AEP.		Educat.		Rho.	
	Coef.	p-level	Coef.	p-level	Coef.	p-level	Coef.	p-level
1	7.9521	0.7503	2.6953	0.7874	18.5869	0.8346	0.3627	0.0642
2	1.3508	0.7612	-2.6955	0.1249	-25.5589	0.6084	0.3066	0.1030
3	-2.7214	0.8022	4.6493	0.4699	7.4387	0.9353	0.3523	0.0754
4	-11.3457	0.5270	3.0717	0.6281	15.1035	0.7297	0.3483	0.0731
5	-0.0919	0.9795	12.5234	0.1120	13.2903	0.5414	0.3528	0.0719
6	1.0151	0.8636	6.8003	0.3925	0.3175	0.9968	0.3451	0.0737
7	-3.1809	0.4939	5.5603	0.3826	6.4222	0.8607	0.3528	0.0621
8	-5.5499	0.7000	0.7965	0.9355	-5.2869	0.9201	0.3455	0.0780
9	-0.5529	0.3742	-6.0333	0.1064	53.0674	0.0002	0.2825	0.1192
10	-1.0570	0.5652	6.7699	0.4988	-56.8075	0.4306	0.3298	0.0862
11	1.8169	0.7878	4.3950	0.1747	3.2451	0.8733	0.3511	0.0619
12	-5.9837	0.6317	0.6229	0.9702	9.8478	0.6781	0.3375	0.0785
13	24.8056	0.4117	-2.5098	0.5886	-19.9752	0.7657	0.3553	0.0692
14	2.0998	0.0249	-2.3826	0.0011	-15.4879	0.0058	0.2300	0.1418
15	-1.3105	0.6679	4.5005	0.5724	-40.1742	0.5487	0.3460	0.0766
16	0.5806	0.0017	-6.0664	0.0000	53.7787	0.0006	0.3343	0.0796
17	-2.0082	0.6881	3.2759	0.7300	0.0936	0.9982	0.3319	0.0810
18	0.5621	0.9061	5.7593	0.3619	23.4482	0.6044	0.3632	0.0706
19	-0.0180	0.9931	5.1461	0.0525	14.0409	0.6372	0.3611	0.0635
20	27.8700	0.7792	-0.2734	0.9778	-57.3303	0.7166	0.3377	0.0840
21	-7.9169	0.6982	5.1700	0.5331	-33.7366	0.8369	0.3561	0.0678
22	35.1615	0.3332	-1.3661	0.8184	-58.5774	0.4288	0.3621	0.0609
23	-11.0799	0.4797	4.6608	0.4708	-51.1501	0.6836	0.3594	0.0606
24	-20.6626	0.0008	5.2937	0.0088	-76.7620	0.0710	0.3458	0.0680
25	1.9244	0.8503	7.5484	0.3895	25.0246	0.6240	0.3366	0.0841
26	-0.7443	0.7190	-4.1731	0.4998	54.1075	0.1247	0.3470	0.0808
27	17.0250	0.0262	0.9689	0.5034	-65.8964	0.0002	0.3015	0.1047
28	-2.6305	0.4954	3.9606	0.4649	18.5450	0.2317	0.3674	0.0599
29	17.9139	0.3625	1.2283	0.6591	-67.3578	0.0796	0.3113	0.0987
30	-1.0409	0.7608	8.0502	0.5792	9.6373	0.6545	0.3366	0.0749
31	51.2510	0.7982	0.9866	0.9637	-134.1037	0.7911	0.3535	0.0739
32	-0.7432	0.9209	5.8641	0.6707	-32.5559	0.6083	0.3457	0.0779
33	2.6160	0.9806	1.3569	0.9610	29.0032	0.9179	0.3482	0.0691
34	2.6193	0.5648	5.4938	0.3121	-1.9297	0.9363	0.3365	0.0809
35	-0.4935	0.5386	-6.5234	0.6139	44.6651	0.6580	0.3237	0.0941
36	-8.9754	0.6001	2.1822	0.8958	7.2633	0.9324	0.3387	0.0800
37	-9.9959	0.4539	1.5752	0.8732	1.1588	0.9855	0.3279	0.0854
38	1.5191	0.3988	-4.4057	0.3252	-27.5126	0.4655	0.3167	0.0964
39	-2.6757	0.9047	9.6953	0.8163	8.0302	0.9161	0.3562	0.0683
40	-11.8717	0.2896	5.0780	0.3293	21.1877	0.2770	0.3478	0.0694
41	-9.0033	0.2217	2.8875	0.5645	10.4880	0.6328	0.3505	0.0713
42	-0.8334	0.6140	-1.5796	0.8714	39.3736	0.3719	0.3382	0.0796
43	3.5382	0.7845	10.8758	0.5625	57.3452	0.6856	0.3661	0.0692
44	-1.2030	0.3609	0.5748	0.9062	34.4838	0.1018	0.3163	0.0967
45	-10.2415	0.4333	14.1552	0.2624	-62.3182	0.2815	0.3510	0.0691
46	41.7106	0.2611	0.2158	0.9667	-112.9763	0.2346	0.3434	0.0705
47	6.0733	0.8199	-10.6884	0.8216	21.8803	0.7462	0.3558	0.0674
48	17.9959	0.4778	0.6418	0.9360	-69.2132	0.3154	0.3354	0.0870

Bayesian spatial autoregressive local estimation

Dependent variable = depvar

Nobs, Nvars = 48, 8

R-squared = 0.8640

Rbar-squared = 0.8402

ndraws,nomit = 1100, 100

r-value = 4.0000

q-nearest = 15

Bayesian spatial autoregressive local estimation

Dependent variable = depvar

Nobs, Nvars = 48, 8

R-squared = 0.6451

Rbar-squared = 0.5830

ndraws,nomit = 1100, 100

r-value = 4.0000

q-nearest = 20

## Bibliographie

1. Abdel-Rahman H, Fujita M (1990) "Product variety, Marshallian externalities and city size" *Journal of Regional Science* 30: 165–183
2. Abdel-Rahman H.M. (1996) "When do cities specialize in production?", *Regional Science and Urban Economics*, 26, 1-22.
3. Abdel-Rahman, Hesham M (1995) "Toward a general-equilibrium theory of a core-periphery system of cities"*Regional Science & Urban Economics*; Vol 25 Issue 4, p529, 18p
4. Abdel-Rahman, Hesham M(1998)"INCOME DISPARITY, TIME ALLOCATION, AND SOCIAL WELFARE IN A SYSTEM OF CITIES".*Journal of Regional Science*,Vol. 38 Issue 1, p137-154, 18p
5. Abdel-Rahman, Hesham M. (1994) "Economies of scope in intermediate goods and a system of cities" *Regional Science & Urban Economics*;Vol. 24 Issue 4, p497, 28p
6. Abdel-Rahman, Hesham M.(2002)"Does the Structure of an Urban System Affect Income Disparities?"*Journal of Regional Science*, Vol. 42 Issue 2, p389, 21p
7. Abraham F., Von Rompuy P. (1995) "Regional convergence in the European Monetary Union", *Papers in Regional Science*, 74, 125-142.
8. Abramovitz (1986) "Catching-up, forging ahead, and falling behind" *Journal of Economic History*, pp. 385–406
9. Abreu M, de Groot HLF, Florax RJGM (2005) "Space and growth: a survey of empirical evidence and methods". *Région et Développement* 21: 12–43
10. Aghion P. and Howitt P. (1998) "Endogenous Growth Theory" The MIT Press
11. Aghion, P and Howitt,P. (1992)"A Model of Growth through Creative Destruction." *Econometrica*, 60 (2), pp. 323-51.
12. Anselin L (2003) "Spatial externalities, spatial multipliers and spatial econometrics" *International Regional Science Review* 26: 153–166
13. Anselin L, Bera AK (1998) "Spatial dependence in linear regression models with an introduction to spatial econometrics." In: Ullah A, Giles DEA (eds) *Handbook of applied economic statistics*. Marcel Dekker, New York, pp. 237-289
14. Anselin L, Bera AK, Florax R, Yoon MJ (1996) "Simple diagnostic tests for spatial dependence". *Regional Science and Urban Economics* 26: 77-104
15. Anselin, L (2002) "Spatial externalities, Spatial multipliers and spatial econometrics" W.P Regional economics applications laboratory.
16. Anselin, Luc (1988). *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer Academic Publishers).
17. Anselin, Luc and Serge Rey. 1991. "Properties of Tests for Spatial Dependence in Linear Regression Models," *Geographical Analysis*, 23, 112–131
18. Arbia G, Basile R, Salvatore M (2003) *Measuring spatial effects in parametric and nonparametric modelling of regional growth and convergence*. Mimeo, University G. D'Annunzio, Pescara
19. Arbia, G. (2005) "Spatial econometrics: with application to regional convergence", Springer-Verlag, Berlin, Forthcoming

20. Armstrong H. (1995a) Convergence among the regions of the European union, Papers in Regional Science, 74, 143-152.
21. Armstrong H. (1995b) An appraisal of the evidence from cross-sectional analysis of the regional growth process within the European Union, dans : Armstrong H., Vickerman R. (eds.), Convergence and Divergence among European Union, Pion, Londres.
22. Azariadis C., Drazen A. (1990) Threshold externalities in economic development, Quarterly Journal of Economics, 105, 501-526.
23. Azzoni, C. (1999). "Economic Growth and Regional Income Inequalities in Brasil", Annals of Regional Science, forthcoming.
24. Azzoni, C. (1999). Economic Growth and Regional Income Inequalities in Brasil, Annals of Regional Science, forthcoming.
25. Azzoni.C et Filho..N.M, Menezes.T et Neto.R.S (2003) "Geography and regional income convergence among Brazilian states" WP Department d'Economie Université de Saopaulo
26. Baldwin R.E., Forslid R. (2000) The core-periphery model and endogenous growth : stabilizing and destabilizing integration, Economica, 67, 307-324.
27. Baldwin R.E., Martin P., Ottaviano G.I.P. (2001) Global income divergence, trade, and industrialization : the geography of growth take-offs, Journal of Economic Growth, 6, 5-37.
28. Barro R.J. (1991) Economic growth in a cross-section of countries, Quarterly Journal of Economics, 106, 407-444.
29. Barro R.J. (1997) Determinants of Economic Growth. A Cross-Country Empirical Study, MIT Press, Cambridge.
30. Barro R.J., Sala-I-Martin X. (2003) "Economic growth" The MIT Press
31. Barro R.J., Sala-I-Martin X. (1992) Convergence, Journal of Political Economy, 100, 223-251
32. Barro R.J., Sala-I-Martin X. (1995) Economic Growth, McGraw-Hill, New York.
33. Barro R.J., Sala-I-Martin X. (1997) Technological diffusion, convergence and growth, Journal of Economic Growth, 2, 1-26.
34. Barro, R.J (2003)"Detrminants of economic growth in a panel of countries"Annals of Economics and Finance, 4, 231-274
35. Barro,R.J (2000) "Inequality and Growth in a Panel of Countries"Journal of Economic Growth Volume 5, Number 1, 5-32
36. Becker G.S., Murphy K.M., Tamura R. (1990) "Human capital, fertility, and economic growth", Journal of Political Economy, 98, S12-37.
37. Becker,Edwin, C.M, Mills,S and Williamson,J.G (1986) "Modeling Indian Migration and City Growth, 1960-2000"Economic Development and Cultural Change, Vol. 35, No. 1, 1-33
38. Bernard A.B., Durlauf S.N. (1995) Convergence in international output, Journal of Applied Econometrics, 10, 97-108.
39. Bernard, A.B., Durlauf, S.N. (1996) "Interpreting test of conergence hypothesis" Journal of econometrics, 71 : 161-173.
40. Bernardo, JM and Smith, AFM (1994) "Bayesian Theory"New York, 2000

41. Bernat G.A. (1996) Does manufacturing matter ? A spatial econometric view of Kaldor's laws, *Journal of Regional Science*, 36, 463-477
42. Besag, J.E (1974) "Spatial interaction and the statistical analysis of lattice systems" *Journal of the Royal Statistical Society, Series B Volume*, 36. Pages, 196-236
43. Black, D. and Henderson, J.V. (1998) "Urban evolution in the USA". Working Paper No. 98-21, Brown University.
44. Blanchard, O. and L. Katz, (1992) " Regional evolutions" *Brookings Papers on Economic Activity* 1, 1-76
45. Brunson C, Fotheringham AS, and Charlton, ME (1998) Spatial nonstationarity and autoregressive models. *Environment and Planning A* 30: 957-973
46. Brunson C, Fotheringham AS, and Charlton, ME (1999) Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Science* 39: 497-524
47. Brunson C, Fotheringham AS, Charlton ME (1996) Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical Analysis* 28: 281-298
48. Brunson C, Fotheringham AS, Charlton ME (2000) Geographically weighted regression as a statistical model. (Unpublished working paper, Department of Geography, University of Newcastle)
49. Brunson, Chris, A. Stewart Fotheringham, and Martin Charlton 1998 "Geographically Weighted Regression: Modelling Spatial Non-Stationarity," *The Statistician*, 47, 431-443
50. Brunson, Chris, A. Stewart Fotheringham, and Martin Charlton. 1996. "Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity," *Geographical Analysis*, 28, 281-298
51. Carlino G, Mills L (1993) Are U.S. regional incomes converging?: A time series analysis, *Journal of Monetary Economics* 32: 335-346
52. Carlino G, Mills L (1996) Convergence and the US States: a time series analysis. *Journal of Regional Science* 36: 597-616
53. Chang, R. (1994), *Income Inequality and Economic Growth: Evidence and Recent Theories*, *Economic Review*, July/August, pp 1 - 9
54. Chang, R. (1994), *Income Inequality and Economic Growth: Evidence and Recent Theories*, *Economic Review*, July/August, pp 1 - 9
55. Chemingui, M.A and Ayadi, N (2003) "Understanding the Poor Human Capital Contribution to Economic Growth in Algeria" *IMF Working Paper*
56. Ciccone A. and Hall R.E. (1996) "Productivity and the density of economic activity" *American Economic Review* 86, 54-70.
57. Ciccone A. (2002) "Agglomeration effects in Europe" *European Economic Review* 46, 213-227.
58. Ciccone, A and Hall, R.E (1996) Productivity and the density of economic activity, *American Economic Review*, 86, 54-70
59. Cliff AD, Ord JK (1981) *Spatial processes*. Pion, London
60. Coe, DT et Helpmann, E (1995) "International R&D spillovers " *European Economic Review* 39; 859-887

61. Cohen, D., (1996), "Tests of the Convergence Hypothesis: Some Further Results," *Journal of Economic Growth*, 1, 3, 351-361.
62. Combes P.P (2000a) "Economic structure and local Growth : France, 1984-1993" *Journal of Urban Economics* 47, 329-355.
63. Combes P.P, Duraton G, and Gobillon L(2004) "Spatial wage disparities: Sorting matters" CEPR Discussion Paper N° 4240.
64. Combes P.P, Magnac T and Robin J. M. (2004) "The dynamics of local employment in France" *Journal of Urban Economics* 56, 217-243.
65. Combes P.P, Overman H.G (2004) "The spatial distribution of economic activities in the European Union" *Handbook of urban and regional economics vol. 4 Cities and Geography*.
66. Combes, P.P, Lfourcade, M (2005) "Transport costs: measures, determinants, and regional policy implications for France" *Journal of Economic Geography*; Vol. 5 Issue 3, p319-349
67. Combes, P.P. (2000b) "Economic structure and location growth: France, 1984-1993" *Journal of urban economics* 47, 329-355.
68. Conley, T.G., Ligon, E. (2002) "Economic distance and cross-country spillovers" *Journal of economic growth*, 7 : 157-187.
69. Conley, TG and Topa, G ( 2002) "Socio-economic distance and spatial patterns in unemployment " *Journal of Applied Econometrics*, 17:303–327
70. De la fuente.A (2000) "Convergence across countries and regions : theory and empirics" SCIC W.P 447.00
71. De Long B., Summers L. (1991) "Equipment investment and economic growth", *Quarterly Journal of Economics*, 106, 445-502.
72. De Long, B. and Summers,L. (1993), "How Strongly do Developing Economies Benefit from Equipment Investment?," *Journal of Monetary Economics*, 32, 3, 395-415.
73. De palma, A; Ginsburgh, V and Thisse,J.F (1987) "On Existence of Location Equilibria in the 3-Firm Hotelling Problem" *Journal of Industrial Economics*; Vol. 36 Issue 2, p245-252
74. Démurger, S (2001)"Infrastructure Development and Economic Growth: An Explanation for Regional Disparities in China?" *Journal of Comparative Economics* 29, 95–117
75. Dixit A.K., Stiglitz J.E. (1977) Monopolistic competition and optimum product diversity, *American Economic Review*, 67, 297-308
76. Dosi, G (1988)"Sources, procedures, and microeconomic effects of innovation" *Journal of Economic Literature* Vol XXVI, 1120-1171
77. Dumais, D;Ellison,G;Glaser, E. (2002) "Geographic Concentration as a Dynamic Process" *Review of Economics and Statistics*, V OL.LXXXIV NUMBER 2
78. Duraton, G and Puga, D (2000) "Diversity and Specialisation in Cities: Why, Where and When Does it Matter?" *Urban Studies*, Vol. 37, No. 3, 533-555.
79. Durlauf S.N., Johnson P.A. (1995) "Multiple regimes and cross-country growth behaviour", *Journal of Applied Econometrics*, 10, 365-384.
80. Durlauf, S.N., Quah, D.T. (1998) "The new empirics of economic growth" W.P Centre for economic performance.
81. Durlauf,S.N; Johnson,P.A and Temple,J.R.W ( 2004) "growth econometrics" WP

82. Durlauf.S.N , Kourtellos.A et Minkin.A (2001) "The local solow growth model".  
European economic review 45
83. Eaton, J. and Kortum, S., (1999), "International Technology Diffusion: Theory and Measurement," International Economic Review, 40, 3, 537-570.
84. Egger P, Pfaffermayr M (2006) "Spatial convergence". Papers in Regional Science 85(2)
85. El Ouardighi, J (2002) "DÉPENDANCE SPATIALE SUR DONNÉES DE PANEL : APPLICATION À LA RELATION BREVETS-R&D AU NIVEAU RÉGIONAL"  
L'Actualité économique, Revue d'analyse économique, vol. 78, no 1
86. Ellison,G and Glaeser,E (1997) Geographic concentration in U.S. manufacturing industries: A dartboard approach, Journal of Political Economy, 105, 889 -927
87. Ellison,G and Glaeser,E (1999), The geographic concentration of an industry. Does natural advantage explain agglomeration, American Economic Association Papers and Proceedings 89,311-316
88. Englmann F.C., Walz U. (1995) Industrial centers and regional growth in the presence of local inputs, Journal of Regional Science, 35, 3-27.
89. Enright, MJ (1990) "Regional Clusters and firm strategy" in the dynamic firme, Oxford university press
90. Ertur, C. and W. Koch, 2007, Growth, technological interdependence and spatial externalities: theory and evidence, Journal of Applied Econometrics, forthcoming.
91. Ertur,C; Le Gallo, J and LeSage, J (2003)"Local versus Global Convergence in Europe: A Bayesian Spatial Econometric Approach" WP
92. Evans, Paul.(1996) "Using Cross-Country Variances to Evaluate Growth Theories." Journal of Economic Dynamics and Control, 20 (6-7), pp. 1027-49.
93. Fingleton B (1999) Estimates of time to convergence: an analysis of the regions of the European Union. International Regional Science Review 22: 5-34
94. Fingleton B (2000) Spatial econometrics, economic geography, dynamics and equilibrium: a 'third way'? Environment and Planning A 32: 1481-1498
95. Fingleton B (2001) Equilibrium and economic growth: spatial econometric models and simulations. Journal of Regional Science 41: 117-147
96. Fingleton B (2003) Externalities, economic geography, and spatial econometrics: conceptual and modelling developments. International Regional Science Review 26: 197-207
97. Fingleton B (2004) Regional economic growth and convergence: insights from a spatial econometric perspective. In: Anselin L, Florax R, Rey S (eds) Advances in Spatial Econometrics. Springer-Verlag, Berlin
98. Fingleton B, López-Bazo E (2003) Explaining the distribution of manufacturing productivity in the EU regions. In: Fingleton B (ed) European Regional Growth. Springer-Verlag, Berlin
99. Fingleton B, McCombie J (1999) Increasing returns and economic growth: some evidence for manufacturing from the European Union regions: a correction. Oxford Economic Papers 51: 574-575
100. Fotheringham, A. Stewart, Martin Charlton, and Chris Brunsdon 1997a. "Two Techniques for Exploring Non-Stationarity in Geographical Data," Geographical Systems, 4, 59-82

101. Fotheringham, A. Stewart, Martin Charlton, and Chris Brunsdon. 1997b. "Measuring Spatial Variations in Relationships with Geographically Weighted Regression," in Manfred M. Fischer and Arthur Getis (eds.), *Recent Developments in Spatial Analysis*. London: Springer-Verlag, pp. 60–82.
102. Fujita M, Krugman P, Venables A (1999) *The spatial economy: cities, regions, and international trade*. MIT Press, Cambridge MA
103. Fujita M. (1989) *Urban Economic Theory – Land Use and City Size*, Cambridge University Press, Cambridge
104. Fujita M., Thisse J.-F. (1997) *Economie géographique, problèmes anciens et nouvelles perspectives*, *Annales d’Economie et de Statistique*, 45, 37-87.
105. Fujita, M and Hu, D (2001)"Regional disparity in China 1985-1994: The effects of globalization and economic liberalization"*Annals of Regional Science*;Vol. 35 Issue 1, 3-35
106. Fujita, M; Ogawa, H and Thisse, J.F (1988) "A SPATIAL COMPETITION APPROACH TO CENTRAL PLACE THEORY : SOME BASIC PRINCIPLES"*Journal of Regional Science*;Vol. 28 Issue 4, p477
107. Fujita, M;Krugman, P and Mori, T(1999)"On the evolution of hierarchical urban systems"*European Economic Review*; Vol. 43 Issue 2, p209-251
108. Galor O. (1996) "Convergence ? Inference from theoretical models" *The economic journal*, 106.
109. Gelfand, A.E and Smith, A.F.M (1990)"Sampling-Based Approaches to Calculating Marginal Densities"*Journal of the American Statistical Association*, Vol. 85, No. 410, 398-409
110. Gelman, A ; Carlin, J.B; Stern, H.S; Rubin, D.B (1995)"*Bayesian Data Analysis. Texts in Statistical Science*" Chapman & Hall
111. Geweke, J (1993 ) "Bayesian treatment of the independent Student t Linear model" *Journal of applied econometrics*, Vol, 8 pp S19-S40.
112. Geweke, J (1998) "Using simulation methods for Bayesian econometrics models : inference, development and communication" W.P Univesity of Minnesota and federal reserve bank of Minneapolis.
113. Glaeser E.(2000) "The new economics of Urban and Regional Growth" In *The Oxford Handbook of Economic Geography*. Edited by : Gordon L. Clark, Maryann P. Feldman and Meric S. Gertter.
114. Glaeser E.L., Maré D.C. (2001) "Cities and skills" *Journal of Labor econometrics* 100, 1126-1152.
115. Glaeser, E (2000) "Urban and regional growth" In *The Oxford Handbook of Economic Geography*. Edited by : Gordon L. Clark, Maryann P. Feldman and Meric S. Gertter (2000).
116. Glaeser, E.,H.Kallal,J. Scheinkman, and A. Shleifer, (1992) "Growth in cities" *Journal of Political Economy* 100, 1126-1152.
117. Glaeser, E.,H.Kallal,J. Scheinkman, and A. Shleifer, (1995) "Economic growth in a cross-section of cities" *Journal of Monetary Economics* 36, 117-143.
118. Glaeser,E (1998), *Are cities dying?*, *Journal of Economic Perspectives*, 12, 139-160
119. Glaeser,E (1999) *Learning in cities*, *Journal of Urban Economics*, 46, 254-277

120. Glaeser.E (1997) "Geography concentration in US : Manufacturing industries A dashboard approach". *Journal of political economic* vol 105
121. Goetz, S. and D. Hu, (1996), "Economic Growth and Human Capital Accumulation: Simultaneity and Expanded Convergence Tests, *Economic Letters*, 51, 355-362.
122. Grossman G.M., Helpman E. (1991) *Innovation and Growth in the Global Economy*, MIT Press, Cambridge.
123. Grossman, G. M. and Helpman,E.(1994) "Endogenous Innovation in the Theory of Growth," *Journal of Economic Perspectives*, 8 (1), pp. 23-44.
124. Haining,R.P (2003)"*Spatial Data Analysis Theory and Practice*"Cambridge University Press
125. Hall, R. and Jones,C. (1999) "Why Do Some Countries Produce so Much More Output per Worker than Others?" *Quarterly Journal of Economics*, 114 (1), pp. 83-116.
126. Hall, S., D. Robertson, and M. Wickens, (1997), "Measuring Economic Convergence," *International Journal of Finance and Economics*, 2, 131-143.
127. Hall,R. and Jones, C. (1996). "The Productive of Nations", *National Bureau of Economic Research, Working Paper Series*, 5812
128. Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chain and their applications. *Biometrika*, 57, 97-109
129. Head, K. and T. Mayer (2000), "Non-Europe : The Magnitude and Causes of Market Fragmentation in Europe", *Weltwirtschaftliches Archiv* 136(2):285–314.
130. Head, K. and T. Mayer, (2004), "The empirics of agglomeration and trade", in: J.V. Henderson and J.-F. Thisse, eds., *Handbook of regional and urban economics*. Volume 4 (North Holland, Amsterdam) 2609-2669.
131. Helpman, E. and P. Krugman (1985), *Market Structure and Foreign Trade* (MIT Press, Cambridge).
132. Helsley,R.W and Strange,W.C (1990) "Agglomeration economies and matching in a system of cities", *Regional Science and Urban Economics*, 20, 189 -212
133. Henderson J.V (1986) " Efficiency of ressources usage and city size" *Journal of Urban Economics* 19, 47-70
134. Henderson J.V (1988) "*Urban Development: Theory, Fact and Illusion*" Oxford University
135. Henderson J.V (1994) " Where does an industry locate?" *Journal of Urban Economics* 35, 83-104
136. Henderson J.V (1997) " Externalities and Industrial development" *Journal of Urban Economics* 42, 449-470
137. Henderson J.V., Kuncoro and Turner (1995)"*Industrial Development in Cities*," *Journal of Political Economy* 103, 1067-1090.
138. Henderson V. (2000) "How urban concentration affects economic growth" *The World Bank, Development Research Group, Infrastructure and Environment*, April 2000.
139. Henderson V. (2003) "The Urbanisation process and economic growth : the so what question"*Journal of Economic Growth* 8, 47-71
140. Henderson.J.V. , Shalizi.Z et Venable .A.J (2000) "*Geography and development*" W.P

141. Howitt P (2000)"Endogeneous growth and cross-country income differences"  
American Economic Review 90, : 829-46.
142. Howitt, P (2008) " Endogenous growth"The New Palgrave Dictionary of  
Economics, 2nd Edition, edited by Steven Durlauf and Lawrence Blume, New York:  
Palgrave Macmillan, forthcoming
143. Islam, Nazrul. "Growth Empirics: A Panel Data Approach." Quarterly Journal of  
Economics, 110 (4), pp. 1127-70.
144. Jaffe, A.B, Trajtenberg,M and Henderson, R(1993) "Geographic Localization of  
Knowledge Spillovers as Evidenced by Patent Citations"The Quarterly Journal of  
Economics, Vol. 108, No. 3 pp. 577-598
145. Jalan, J. and Ravallion, M (1998a) Are there dynamic gains from a poor-area  
development program? Journal of Public Economics, 67, pp. 65 - 85.
146. Jalan, J. and Ravallion, M(1998b). Geographic Poverty Traps, Word Bank, May  
15 pp 1- 31.
147. Jones C.(1999) "Growth : With or Without scale effect?" New ideas on economic  
growth vol 89 n° 2
148. Kendrick,J.W (1976)" The formation and stocks of total capital" NBRE, New  
York (General series—National Bureau of Economic Research; no. 100)
149. Kim, S. (1995), "Expansion of Markets and the Geographic Distribution of  
Economic Activities: The Trends in U.S. Regional Manufacturing Structure, 1860-1987",  
Quarterly Journal of Economics 110(4):881–908.
150. Klenow, P.J and Rodríguez-Clare, A(1997) "Economic growth: A review essay"  
Journal of Monetary Economics Volume 40, Issue 3 597-617
151. Koop, G (2003) "Bayesian Econometrics" Wiley and sons, New York
152. Koranchelian,T (2005) "The Equilibrium Real Exchange Rate in a Commodity  
Exporting Country: Algeria's Experience" IMF Working Paper
153. Kosfeld.R. et Lauridsen.J (2004) "Dynamic spacial modeling of regional  
convergence processes" Empirical Economics Volume 29, Number 4.
154. Kpodar,K (2007) "Why Has Unemployment in Algeria Been Higher than in  
MENA and Transition Countries?" IMF Working Paper
155. Krugman P. (1981) "Trade, accumulation, and uneven development" Journal of  
Development Economics, 8, 149-161
156. Krugman P. (1991a) "Increasing returns and economic geography", Journal of  
Political Economy, 99, 483-499
157. Krugman P. (1991b) "Geography and Trade", MIT Press, Cambridge.
158. Krugman P., Venables A.J. (1995) "Globalization and the inequality of nations",  
Quarterly Journal of Economics, 110, 857-880
159. Krugman P., Venables A.J. (1996)" Integration, specialization and adjustment",  
European Economic Review, 40, 959-967
160. Lange, K.L, Little, R.J.A., Taylor, J.M.G (1989) "Robust statistical modelling  
using the t distribution" Journal of the American Statistical Association, Vol, 84, N° 408.
161. Le Gallo J, Ertur C, Baumont C (2003) "Aspatial econometric analysis of  
convergence across European Regions, 1980–1995". In: Fingleton B (ed) European  
Regional Growth. Springer-Verlag, Berlin

162. Lee, K., M. Pesaran, and R. Smith, (1997), "Growth and Convergence in Multi Country Empirical Stochastic Solow Model," *Journal of Applied Econometrics*, 12, 4, 357-92.
163. LeSage J(2005) "Bayesian estimation of spatial regression models" lecture note.
164. LeSage J., (2004). "A family of geographically weighted regression models", in L.Anselin, R.J.G.M. Florax and S. Rey (eds.), *Advances in Spatial Econometrics: Methodology, Tools and Applications*, Springer-Verlag, New York.
165. LeSage J., 1997. "Bayesian estimation of spatial autoregressive models", *International Regional Science Review*, 20: 113-129
166. LeSage J., 1999. "Spatial Econometrics: The Web Book of Regional Science", Regional Research Institute, West Virginia University, Morgantown, WV
167. Lichtenberg, F., (1992), "R&D Investment and International Productivity Differences," National Bureau of Economic Research Working Paper no. 4161.
168. Liu, Z. and T. Stengos, (1999), "Non-Linearities in Cross Country Growth Regressions: A Semiparametric Approach," *Journal of Applied Econometrics*, 14., 5, 527-38.
169. Loko,B; Kpodar, K and Diallo,O (2007) "Buoyant Capital Spending and Worries over Real Appreciation: Cold Facts from Algeria" IMF Working Paper No. 07/286
170. López-Bazo E, Vayá E, Mora AJ, Suriñach J (1999) Regional economic dynamics and convergence in the European Union. *The Annals of Regional Science* 33: 343–370
171. López-Bazo E., Vayá E. Artis M. (2004) "Regional externalities and growth : evidence from European regions" *Journal of regional science*, 44.1 : 43-73.
172. Lucas, R.E., (1988) "On the mechanics economic development" *Journal of Monetary Economics* 22, 3-42.
173. Magrini, S., (2004), "Regional (di)convergence", in: J.V. Henderson and J.-F. Thisse, eds.,*Handbook of regional and urban economics*. Volume 4 (North Holland, Amsterdam),2741-2796.
174. Mankiw G, Romer D, Weil D (1992) A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107: 407–437
175. Mankiw, N. G., (1995), "The Growth of Nations," *Brookings Papers on Economic Activity*, 275-310.
176. Martin P.J., Ottaviano G.I.P. (1999) Growing locations : industry location in a model of endogenous growth, *European Economic Review*, 43, 281-302.
177. Martin R., Tyler P. (2000) "Regional employment evolutions in the European Union: a preliminary analysis" *Regional studies*, 34.7: 601-616.
178. Martin, P and Rogers, C,An (2000)"Long-term growth and short-term economic instability" *European Economic Review*; Vol. 44 Issue 2, p359-381
179. Martin, R. (2001) " EMU versus the regions? Regional convergence and divergence in Euroland" *Journal of economic geography*, 1 : 51-80.
180. Matsuyama, K (1995)"Complementarities and Cumulative Processes in Models of Monopolistic Competition" *Journal of Economic Literature*, Vol. 33, No. 2, pp. 701-729
181. Maurel.F (1996) "Evolution locale de l'industrie 82 – 92 et convergence régionale" Document de travail MEF direction de la prévision 96-6

182. McCOMBIE, J.S.L and RIDDER, J.R (1984) "THE VERDOORN LAW CONTROVERSY: SOME NEW EMPIRICAL EVIDENCE USING U.S. STATE DATA" Oxford Economic Papers Volume 36 Number 2, 268-284
183. McMillen,D.P (1994) Vintage growth and population density: An empirical investigation,Journal of Urban Economics, 36, 333-352.
184. McMillen,D.P (1996) "One hundred fifty years of land values in Chicago : A non parametric approach " Journal of urban economics, 40,100-124
185. McMillen,D.P and McDonald,J.F (1991)" Urban land value functions with endogenous zoning, Journal of Urban Economics, 29, 14-27.
186. McMillen,D.P, McDonald.J.F (1997) "A nonparametric analysis of employment density in polycentric city" Journal of regional science, 37, 591-612
187. Metropolis,N, Rosenbluth,A, Rosenbluth,M, Teller, A, Teller, E (1953) " Equations of state calculations by fast computing machines" J.Chem.Phys. 21, 1087-1092
188. Meyn, S. P. and Tweedie, R. L. (1993). Markov chain and stochastic stability. London: Springer-Verlag.
189. Mion G and Naticchioni P (2005) Urbanisation externalities, market potential and spatial sorting of skills and firms. CEPR Discussion Paper N° 5172.
190. Mody,A and Wang, F. Y (1997)"Explaining Industrial Growth in Coastal China: Economic Reforms ... and What Else? The World Bank Economic Review Volume 11, Number 2, 293-325
191. Moreno, R., Trehan, B. (1997) " Location and the growth of nations" Journal of economic Growth, 2 : 399-418.
192. Murdoch,J(1997) "Towards a geography of heterogeneous associations" Progress in Human Geography, Vol. 21, No. 3, 321-337
193. Nelson, M. and R. Singh, (1994), "The Deficit-Growth Connection: Some Recent Evidence from Developing Countries," Economic Development and Cultural Change, 42, 167-191.
194. Nelson, R.R and Phelps,E.S (1966) "Investment in Humans, Technological Diffusion, and Economic Growth" The American Economic Review, Vol. 56, No. 1/2 , 69-75
195. Neven D., Gouyette C. (1995) "Regional convergence in the European Community", Journal of Common Market Studies, 33, 47-65.
196. Odland J. (1988) Spatial Autocorrelation, Scientific Geography Series, Volume 9, SAGE Publications, Newbury Park.
197. Ogawa H., Fujita M. (1980) Equilibrium land use patterns in a nonmonocentric city, Journal of Regional Science, Vol 20, 4, 455-475.
198. Ogawa H., Fujita M. (1989) Nonmonocentric urban configuration in a two-dimensional space, Environment and Planning A, 21, 363-374
199. Ottaviano G.I.P., Puga D. (1998) Agglomeration in the global economy : a survey of the "new economic geography", World Economy, 21, 707-731.
200. Ottaviano, G.I.P. and J.-F. Thisse, (2004), Agglomeration and economic geography, in: J.V. Henderson and J.-F. Thisse, eds., Handbook of regional and urban Economics. Volume 4 (North-Holland, Amsterdam) 2564-2608.

201. Ottaviano, G.I.P.( 2002) Models of ‘new economic geography’: factor mobility vs. vertical linkages, GIIS, mimeo.
202. Ottaviano, G.I.P., T. Tabuchi and J.-F. Thisse, (2002), Agglomeration and trade revisited, *International Economic Review*, 43, 409-436.
203. Ottaviano, G.I.P., Thisse, J.F( 2003)"Agglomeration and Economic Geography" CEPR Discussion Paper No. 3838
204. Pace R.K. and Barry R., 1997. Sparse spatial autoregressions, *Statistics and Probability Letters*, 33: 291-297.
205. Pace R.K. and Barry R., 1998. *Spatial Statistics Toolbox 1.0*, Real Estate Research Institute, Louisiana State University, Baton Rouge, LA.
206. Pace, R.K et LeSage, J (2004) "Spatial Autoregressive Local Estimation" in *Spatial Econometrics and Spatial Statistics De Arthur Getis, Jesús Mur, Henry Zoller, Western Regional Science*
207. Papageorgiou, C. and W. Masanjala, (2004), “The Solow Model with CES Technology: Nonlinearities with Parameter Heterogeneity,” *Journal of Applied Econometrics*, 19, 2, 171-201.
208. Papageorgiou, C., (2002), "Trade as a Threshold Variable for Multiple Regimes," *Economics Letters*, 71, 1, 85-91.
209. Parent.O et Riou.S (2003) "Bayesian analysis of knowledge spillovers in European regions" W.P CREUSET
210. Parente, Stephen L. and Edward C. Prescott. “Technology Adoption and Growth.” *Journal of Political Economy*, April 1994, 102 (2), pp. 298-321
211. Pesaran, M. H., (2004), “A Pair-Wise Approach to Testing for Output and Growth Convergence,” mimeo, University of Cambridge.
212. Prescott, E., (1998), “Needed: A Theory of Total Factor Productivity,” *International Economic Review*, 39, 525-551.
213. Puga D., Venables A.J. (1997) Preferential trading arrangements and industrial location, *Journal of International Economics*, 43, 347-368
214. Puga D., Venables A.J. (1999) Agglomeration and economic development : import substitution versus trade liberalisation, *Economic Journal*, 109, 293-311.
215. Qauh, D (2002)"Spatial Agglomeration Dynamics"CEPR Discussion Paper No. 3208
216. Quah, D., (1996), “Empirics for Economic Growth and Convergence,” *European Economic Review*, 40, 6, 1353-75.
217. Quah, D., (1997), “Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs,” *Journal of Economic Growth*, 2, 1, 27-59.
218. Rappaport J (1999b) « Local growth Empirics » CDI Working Paper No. 23, Centre for international development at Harvard University.
219. Rappaport J. (1999a) « Local growth theory » Mimeo, Centre for international development at Harvard University.
220. Ravallion, M. and Jalan, J. (1996), Growth Divergence Due Spatial Externalities, *Economic Letters* 53, 227-232.
221. Rey SJ (2001) Spatial empirics for economic growth and convergence. *Geographical Analysis* 33: 195-214

222. Rey SJ, Montouri BD (1999) US regional income convergence: a spatial econometric perspective. *Regional Studies* 33: 143-156
223. **Robelo, S.** (1991) "Long-run policy analysis and long-run growth", *Journal of Political Economy* 99 (1991), pp. 500–521. ...
224. Robert, C (2006) "Le choix Bayésien : Principe et pratique" Springer-Verlag France
225. Robert, C and Casella, G (2004) "Monte Carlo Statistical methods" Springer Verlag, New York, second edition
226. Romer P. (1986) Increasing returns and long run growth, *Journal of Political Economy*, 94,1002-1037.
227. Romer P. (1990) Endogenous technological change, *Journal of Political Economy*, 98, S71-S102.
228. Romer P. (1994) The origins of endogenous growth, *Journal of Economic Perspectives*, 8, 3-22
229. Romer, D., (2001), *Advanced Macroeconomics*, New York: McGraw-Hill.
230. Rosenthal S.S. and Strange W. (2004) "Evidence on the nature and sources of agglomeration economics" *Handbook of urban and regional economics vol. 4 Cities and Geography*
231. Rosenthal, S.S. and W.C. Strange, 2001, The determinants of agglomeration, *Journal of Urban Economics* 50, 191-229.
232. Ruppert, E. (1999) "The Algerian Retrenchment System: A Financial and Economic Evaluation" *THE WORLD BANK ECONOMIC REVIEW*, VOL. 13, NO. 1: 155–83
233. Sachs, J and A Warner, (1996), "Natural Resource Abundance and Economic Growth," *National Bureau of Economic Research Working Paper no. 5398*.
234. Sachs, J., (2003). "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income," *National Bureau of Economic Research Working Paper no. 9490*.
235. Sala-i-Martin X (1996) Regional cohesion: evidence and theories of regional growth and convergence. *European Economic Review* 40: 1325–1352
236. Satchi M., Temple J. (2005) "Growth and labor markets in developing countries" W.P.
237. Sergio J.R. et Montouri.B.D (1998) "US regional income convergence: A spatial econometric perspective". *Regional studies vol 33-2*
238. Solow, R., (1956), "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, 70, 65-94.
239. Tanner, M and Wong, W.H. (1987) "The Calculation of Posterior Distributions by Data Augmentation" *Journal of the American Statistical Association*, Vol. 82, No. 398 (Jun., 1987), pp. 528-540
240. Temple J (1999) "The new growth evidence" *Journal of economic literature*, vol XXXVII, 112-156
241. Temple J. (1998) Robustness tests of the augmented Solow model, *Journal of Applied Econometrics*, 13, 361-375.
242. Temple, J., (2000), "Growth Regressions and What the Textbooks Don't Tell You," *Bulletin of Economic Research*, 52, 3, 181-205

243. Temple, J., (2003), "The Long-run Implications of Growth Theories," *Journal of Economic Surveys*, 17, 3, 497-510.
244. Thisse J.-F. (1997) "L'oubli de l'espace dans la pensée économique", *Région et Développement*, 6, 13-39.
245. Trefler, Daniel.(1993) "International Factor Price Differences: Leontief Was Right!" *Journal of Political Economy*, 101 (6), pp. 961-87
246. Vayá E, López-Bazo E, Artís M (1998) Growth, convergence and (why not?) regional externalities. (Unpublished working paper E98/31, "Anàlisi Quantitativa Regional" Research Group, University of Barcelona
247. Venables D. (1996) "Equilibrium location of vertically linked industries", *International Economic Review*, 37, 341-359.

## Table des indexes

Figure 1 : Dynamique de croissance dans le modèle Solow-Swan : Economies identiques .....	15
Figure 2: Dynamique de croissance dans le modèle Solow-Swan : Economies différentes.....	15
Figure 3 : Le modèle AK .....	24
Figure 4 : Dynamique de la structure Centre – Périphérique.....	48
Figure 5 : Ensemble de polygones.....	58
Figure 6 : Croissance di PIB réel.....	90
Figure 7 : Productivité du travail dans les pays du MENA et des pays en transition 1989 - 2004.....	93
Figure 8 Graphe de Moran du taux de croissance .....	126
Figure 9 : Moyenne des estimations a posteriori de $\nu$ .....	126
Figure 10 : Distribution de rho en fonction de la taille m.....	143
Figure 11 Hétérogénéité dans la convergence des régions Algérienne .....	144
Tableau 1 : Sources de la croissance en Algérie .....	91
Tableau 2 : Elasticité emploi – production 1989 - 2006 .....	92
Tableau 3 Territoires Algérien en 1905.....	93
Tableau 4 : Départements Algériens en 1957 .....	94
Tableau 5 : .....	125
Tableau 6 : .....	128
Tableau 7 : .....	132
Carte 1 : Carte de l’Algérie entre 1974 et 984 .....	96
Carte 2 : Carte de l’Algérie depuis 1984 à nos jours.....	97
Carte 3 : Densité de la population en 1998.....	100
Carte 4 : Variation de la population entre 1987 et 1998 .....	101
Carte 5 : Taux d’activité de la population .....	102
Carte 6 : Taux d’emploi en 1998.....	103
Carte 7 : Taux d’actifs potentiel en 1998 .....	104
Carte 8: Part de la population dans les agglomérations secondaires et chefs-lieux en 1998 (%).....	105
Carte 9 : Variation de la population dans les agglomérations secondaires et chefs-lieux de 1987 à 1998 (%) .....	106
Carte 10 : Taux d’urbanisation en 1998.....	107
Carte 11 : Part de la population ayant un diplôme universitaire en 1998.....	108
Carte 12 : Pourcentage de la population analphabète en 1998 .....	109
Carte 13 : Taux d’occupation par pièce en 1998 .....	110
Carte 14 : KM de route nationale par habitant en 1998.....	111
Carte 15 Nombre d’élève par enseignant.....	112
Carte 16 : Taux de scolarisation en 1998 .....	113
Carte 17 : Nombre de médecins pou 1000 habitants en 1998.....	114

