

TRAVAUX ET RECHERCHES DE PROSPECTIVE

schéma
général
d'aménagement
de la France

emploi
et espace

49

éléments pour un modèle
de prévision

schéma
général
d'aménagement
de la France

emploi
et espace
49

éléments pour un modèle
de prévision

1

TRAVAUX ET RECHERCHES DE PROSPECTIVE

schéma
général
d'aménagement
de la France

**emploi
et espace**

DECEMBRE 1974

éléments pour un modèle
de prévision

1

AVANT-PROPOS

La localisation spatiale des créations et des suppressions d'emplois, les migrations des personnes qui les accompagnent ou les précèdent sont de toute évidence placées au coeur des préoccupations de l'Aménagement du Territoire.

D'abord à une échelle très fine, celle des bassins d'emplois la fermeture de certains établissements, la récession de certaines activités posent de façon constamment renouvelée de difficiles problèmes d'emploi, que l'on s'efforce de résoudre par les moyens cumulés du sauvetage d'entreprises en difficulté, de la décentralisation industrielle et tertiaire et du développement des entreprises locales.

Mais à une échelle plus vaste, celle des régions, les mouvements des personnes et des activités posent du point de vue de l'aménagement du territoire des questions en apparence très différentes : certaines régions sont-elles menacées d'une stagnation démographique, tandis que d'autres connaîtraient un engorgement nuisible et coûteux ? Certaines portions du territoire risquent-elles de tomber en déshérence, si l'on n'y prend garde ?

L'évolution des bassins d'emplois est dans la plupart des cas suivie dans le court terme : c'est au maximum un horizon de 3 ans que l'on se donne habituellement pour prévoir de façon précise les conversions industrielles ou les effets de l'ouverture de nouvelles usines dont la création vient d'être décidée. On est ici dans le domaine de l'actuel et de l'opérationnel.

Les grands mouvements qui façonnent la répartition des hommes et des activités sur le territoire sont, au contraire, étudiés dans le long terme. C'est aux échéances de 1985 et de 2000 que l'on s'efforce de prévoir l'impact des "tendances lourdes" et des "faits porteurs d'avenir". On est là dans un domaine relevant de la prospective, et défriché par plusieurs études qui ont déjà été publiées. (1)

(1) Voir notamment Travaux et Recherches de Prospective n° 20 (le Scénario de l'Inacceptable), n° 26 (La transformation du monde rural), n° 44 (Régions urbaines, régions de villes) publiés à la Documentation Française.

Pourtant, il est clair que ces deux domaines nettement séparés reposent en réalité sur des phénomènes de même nature : la création différentielle, selon les zones d'accueil, d'emplois nouveaux ; la récession inéluctable de certaines activités ; les migrations entre ville et campagne, de ville à ville.

Ainsi est né, en 1970, le projet consistant à jeter un pont entre les deux domaines, entre le court et le long terme, entre l'opérationnel très localisé et le prospectif très global.

*

* *

Ce projet s'est incarné dans la tentative de réalisation du modèle MURAT, lancée à l'initiative du Commissariat général du Plan dans le cadre des travaux du SESAME.

Il n'est pas question de décrire ici les caractéristiques de ce modèle dont une présentation résumée a déjà été faite par Rémy PRUD'HOMME (2). Il suffira de dire qu'il s'agissait essentiellement, sur une période de 15 ans, de répondre aux questions suivantes :

- a) Comment les évolutions globales prévisibles de notre économie vont-elles se traduire concrètement au niveau des bassins d'emplois ?
- b) De l'ensemble des ajustements des marchés locaux de l'emploi, résulte-t-il des mouvements généraux de population et d'activité, à l'échelle de grandes régions ?
- c) Quelles politiques peut-on recommander, soit pour se prémunir à l'avance contre les risques pesant sur telle ou telle famille des bassins d'emplois, soit pour contrecarrer des tendances d'ensembles jugées néfastes ?
- d) Les évolutions spontanées, ou infléchies par des politiques volontaires, de l'occupation du territoire ont-elles des conséquences sur le développement économique national ?

(2) Voir Rémy PRUD'HOMME - Un cadre pour l'analyse du développement inter-régional "l'actualité économique - Montréal 1972"

On voit que la tentative était ambitieuse, puisqu'il s'agissait de relier, par un faisceau de liaisons très nombreuses, des phénomènes perceptibles au niveau national, au niveau régional et au niveau local, et que de plus ces phénomènes n'étaient pas tous, loin de là, suffisamment bien connus.

La tâche prioritaire consistait évidemment à rassembler des connaissances, et à les accroître là où elles étaient suffisantes. Trois axes d'études furent alors sélectionnés :

- la localisation des emplois industriels relevant de décisions prises au niveau national,
- le fonctionnement du marché local de l'emploi,
- les interactions de l'aménagement du territoire et du développement économique national

cependant qu'était élaboré un pré-modèle destiné à incorporer les résultats de ces études et à explorer la manière dont les bassins d'emplois réagissent les uns sur les autres.

De sérieuses difficultés surgirent en 1971, notamment sur le plan de la disponibilité des statistiques, et il apparut alors que, si l'on voulait aboutir, les études préalables seraient nettement plus longues et plus lourdes que prévu : en définitive, le fonctionnement du modèle complet ne pouvait être garanti à brève échéance, et son élaboration ressortissait plus à de la recherche qu'à un travail administratif normal, l'abandon du projet, tel qu'il avait été initialement conçu, fut donc décidé à la fin de 1971.

*

*

*

Mais cet abandon du projet initial n'a pas signifié pour autant l'arrêt des réflexions et des travaux sur le sujet.

Les études sur le fonctionnement du marché local de l'emploi ont été en effet poursuivies par plusieurs équipes universitaires, notamment par l'Institut d'études de l'emploi de TOULOUSE, et par le centre d'analyses du développement (LILLE) pour le cas de DUNKERQUE.

Les réflexions sur le lien entre le développement différentiel des régions et le développement économique national ont été très tôt (dès 1971) reprises par Raymond COURBIS et Jean-Claude PRAGER (1), qui proposent de réaliser un modèle régional - national : "REGINA". Le modèle REGINA qui réalise une articulation interdépendante entre le modèle national FIFI et 5 sous-modèles régionaux, correspondant à 5 grandes régions, est actuellement en cours d'élaboration sous la direction du professeur Raymond COURBIS (2).

*
* *
*

Enfin Vincent GIARD, qui a été dès l'origine l'un des animateurs de l'équipe de conception de MURAT, a eu l'immense mérite de poursuivre, dans des conditions difficiles, le rassemblement d'une très importante base de données départementales, la mise au point économétrique de plusieurs sous-modèles, et enfin le test systématique de quelques unes des hypothèses faites sur les ajustements du marché de l'emploi.

Cet important travail s'est matérialisé par une thèse brillamment soutenue en novembre 1973 devant l'Université de Paris I. Vincent GIARD y expose les résultats de ses études, les justifie et les commente ; il présente également des réflexions et des recherches d'ordre méthodologique qui constituent un acquis considérable pour les études qui prendront la suite de celles-ci.

La recherche sur la prévision et la planification du développement régional à moyen et long terme continue à se développer dans la plupart des pays. En France d'importants travaux se poursuivent (modèle REGINA) et l'on peut d'ores et déjà prévoir qu'après le prochain recensement et à la suite de la préparation régionale du VIIe Plan, un regain d'intérêt se manifesterà pour la planification régionale et inter-régionale, en particulier chez les responsables des régions.

(1) Raymond COURBIS et Jean-Claude PRAGER : Analyse régionale et planification nationale : le projet de modèle "REGINA" d'analyse interdépendante - Collection de l'INSEE, série R, n° 12, mai 1973.

(2) Travaux du groupe d'analyse macroéconomique appliquée. (G. A. M. A.) - Université de Paris X - NANTERRE.

Aussi a-t-il paru indispensable de publier, dans la collection "Travaux et Recherches de Prospective" la thèse de Vincent GIARD sur l'emploi spatialisé, de façon à la rendre accessible aux chercheurs, hommes d'études, universitaires et étudiants appelés à travailler dans les prochaines années sur les problèmes de la spatialisation de l'emploi.

Michel HORPS
Chargé de Mission
à la D.A.T.A.R.

PREFACE

Le travail que nous présente aujourd'hui Vincent GIARD -issu d'une thèse brillamment soutenue en novembre 1973 devant l'Université PARIS I- est exemplaire à plusieurs égards. Il serait difficile de le résumer sans en trahir l'esprit ou d'en restituer fidèlement la substance sans le déflorer. Aussi se bornera-t-on à souligner ici la belle leçon qui s'en dégage au plan de la démarche.

L'intérêt du sujet relève de l'évidence. Compte tenu des données disponibles, il fallait une grande ambition associée à beaucoup de rigueur et de probité pour tester les possibilités - et les limites actuelles - de l'approche systémique dans "la prévision de l'emploi spatialisé en France". Mais ce sont là des qualités dont aucune ne fait défaut à Vincent GIARD.

La progression théorique de ce dernier se déroule selon trois axes.

Le premier consiste en une réflexion originale sur le problème du découpage spatial. Les techniques et indicateurs qu'il nous propose le conduisent à se référer au cadre départemental malgré la grande pénurie des données disponibles à ce niveau et les biais importants qui en résultent au plan des hypothèses testables comme à celui des indicateurs disponibles. Ce choix a donc pour corollaire un effort considérable afin de constituer une base de plusieurs milliers de données dont l'exploitation aura exigé le recours à quelques centaines d'heures d'ordinateur.

On soulignera en second lieu que ce travail est - à notre connaissance - le premier en France à utiliser la régression multiple optimale en se basant sur un algorithme défini en 1969 par KENDALL et BEALE. L'auteur

dégage - précision régulièrement négligée jusqu'ici - les hypothèses complémentaires dont doit s'accompagner l'utilisation de l'approche probabiliste et de la régression multiple pour des données spatialisées, échappant aux globalisations abusives. Il aborde le problème de la dissociation du modèle explicatif d'une variable endogène en plusieurs sous-modèles destinés à expliquer chacun une partie de cette variable : ainsi le tertiaire est-il décomposé en deux modèles relatifs l'un aux services finals, l'autre aux services fournis aux industries.

Enfin, il tire d'analyses théoriques relatives à la répartition spatiale des hommes le maximum d'hypothèses cohérentes susceptibles d'alimenter la confection des divers sous-modèles de son étude de "faisabilité". L'insuffisance générale de la réflexion concernant la prise en compte simultanée de l'ensemble des zones découpant le territoire national et le peu de portée pratique des analyses relatives à l'ajustement de l'offre et de la demande de force de travail lui imposent d'approfondir la notion de marché de l'emploi "défini par l'ensemble des trois caractéristiques suivantes qui réagissent d'ailleurs les unes sur les autres : un cadre temporel, un cadre spatial et un ensemble de mécanismes". Le petit nombre et le caractère partiel des études portant sur l'ensemble des départements le conduisent alors à proposer plusieurs études descriptives originales en tête de chacun des six chapitres consacrés à la définition des divers sous-modèles concernant le sous-système "demandes d'emplois" (immigrations intérieures, émigrations intérieures, taux d'activité) et le sous-système "offres d'emplois" (agriculture, tertiaire, B.T.P., secondaire).

Au terme de cet effort, il fallait à l'auteur une grande lucidité et aussi beaucoup de courage - mais nous avons parlé plus haut de sa rigueur et de sa probité - pour conclure de façon plutôt pessimiste sur les possibilités de l'approche systémique dans le domaine qu'il a choisi d'analyser. Mais en attendant, il aura dégagé (et partiellement comblé) un certain nombre de carences statistiques tout en nous apportant plusieurs études originales. Il aura posé des bases pour les développements futurs d'une méthode qui commence à peine à livrer ses fruits.

Nous pensons quant à nous qu'au-delà de la vieille analyse causale aujourd'hui contestée de toutes parts, l'application à l'économie de l'approche systémique sera l'instrument de bouleversements insoupçonnables.

Analyste méticuleux, Vincent GIARD se refuse par là-même les commodités des jongleries spectaculaires sur des concepts mal définis. Mais il ne succombe pas pour autant à cette autre facilité que constitue le refuge

dans la pseudo-neutralité du concret. Il prend parti au plan théorique comme à celui de l'action. Ses engagements - c'est un bel hommage à lui rendre - nous invitent à de nouvelles réflexions et à de futurs dépassements.

Au seuil de sa jeune carrière, il n'est point besoin d'être grand prophète pour prédire à Vincent GIARD un bel avenir d'économiste.

René PASSET
Professeur à l'Université PARIS I

AVERTISSEMENT

Guide de lecture

Les diverses parties de ce document peuvent présenter selon les lecteurs un intérêt plus ou moins grand, aussi a-t-il paru indispensable de lui fournir quelques indications sommaires pour lui permettre, dans la mesure où il ne disposerait que de peu de temps, d'en extraire le maximum d'informations utiles.

Pour placer cette étude dans son contexte général, il conviendrait de lire intégralement l'introduction, mais le lecteur pressé peut se contenter du § 1.32 de la page 27 ; du § 2.12 de la page 29 ; il ne saurait toutefois éviter la lecture du § 2.2. pages 30 à 36 où l'exposition de la démarche suivie situe assez bien chaque chapitre dans le cadre général de nos diverses préoccupations.

Il convient de signaler que les 7 premiers chapitres sont d'orientation théorique et qu'un certain nombre de points originaux y sont traités. Le lecteur ayant déjà une certaine culture en analyse spatiale et en techniques économétriques pourra donc facilement, au vu du plan détaillé des chapitres (pages 2 à 16) "écrèmer" les seules informations qui lui sont utiles. Néanmoins si les fondements théoriques de l'étude ne présentent guère d'intérêt pour lui, il pourra directement consulter la seconde partie (chapitres 8 à 14)

Enfin le détail des recherches que nous avons menés durant quatre années n'intéressant a priori que quelques centaines de spécialistes, il a été décidé de les livrer telles quelles, comme document de travail, malgré leur caractère parfois très technique. En conséquence cet ouvrage correspond, in extenso, au fascicule 1 de la thèse que nous avons soutenue le 13 novembre 1973, à l'Université de Paris. Le fascicule 2 regroupait une sélection d'une centaine de tableaux et d'environ 250 cartes, extraites de plusieurs milliers de tableaux et de cartes sorties sur ordinateur. Pour la présente publication, on a jugé préférable de limiter les annexes

à celles qui semblaient les plus significatives. Toutefois un double système de renvoi a été introduit par rapport au texte original : le premier est du type "Ax" où A signifie qu'il s'agit de l'annexe du présent document et x est le numéro de la page ; le second du type "Tx.y" où la lettre T renvoie au fascicule d'annexes (fascicule 2) de la Thèse, x désignant le numéro du chapitre et y la page du chapitre. Ce fascicule peut être consulté au service de documentation de la D.A.T.A.R. ainsi qu'à la bibliothèque de l'Université de Paris 1, rue Cujas.

Vincent GIARD

SOMMAIRE (*)

<i>Introduction</i> (1)	p	17
<i>Première Partie : Les Bases théoriques du modèle proposé</i>		
IA - Les fondements théoriques de l'analyse		
Techniques de choix d'un découpage spatial (2)	p	37
Techniques de modélisation (3)	p	61
Analyse théorique de la répartition des hommes entre les différentes zones du territoire (4)	p	85
IB - Adaptation de ces concepts au problème posé		
Le fonctionnement du marché local de l'emploi (5)	p	115
Le choix du découpage spatial (6)	p	137
Remarques sur certaines techniques économétriques retenues (7)	p	151
<i>Seconde Partie : La mise au point du modèle</i>		
IIA - Les sous-modèles proposés		
Le sous-modèle migrations intérieures (8)	p	175
Détermination de la demande Potentielle d'Emploi (9)	p	229
Le sous-modèle BTP (10)	p	257
Le sous-modèle agriculture (11)	p	281
Le sous-modèle tertiaire (12)	p	307
Le sous-modèle secondaire : étude des facteurs de localisation (13)	p	355
IIB - Résultats globaux - Conclusions (14)	p	369

(*) Les nombres entre parenthèses correspondent aux numéros de chapitre

INTRODUCTION

SECTION 1 - Les FONDEMENTS OBJECTIFS de l'AMENAGEMENT du TERRITOIRE

- 1.1 - *Les causes de disparités économiques entre zones*
 - 1.11 - Différence de dotations
 - 1.12 - Rôle de l'économie de marché
 - 1.13 - Le jeu des politiques
- 1.2 - *Nécessité d'une intervention*
 - 1.21 - La thèse non interventionniste
 - 1.22 - Critique de la thèse non interventionniste
- 1.3 - *La connaissance, condition permissive d'actions efficaces en aménagement du territoire*
 - 1.31 - Présentation sommaire des techniques d'analyse de système
 - 1.32 - L'analyse de système du SESAME

SECTION 2 - DEFINITION du CADRE de cette ETUDE

- 2.1 - *Définition de l'objet de cette étude*
 - 2.11 - Remarques préalables sur la disparité spatiale des revenus
 - 2.12 - Objet de l'étude
- 2.2 - *Démarche suivie*
 - 2.21 - Les bases théoriques du modèle
 - 2.22 - La phase de mise au point du modèle

CHAPITRE 2 - TECHNIQUES de CHOIX d'un DECOUPAGE SPATIAL

SECTION 1 - CONSTRUCTION d'une PARTITION de l'ESPACE

p 38

1.1 - Les données du problème

1.11 - Les concepts utilisés

1.12 - Découpages sans introduction de variable

1.2 - Partition du territoire suivant une variable

1.21 - Groupement des valeurs de la variable

1.22 - Autres méthodes de regroupement des u.s.

1.23 - Approche par consolidations successives

1.3 - Partition de l'espace suivant plusieurs variables

1.31 - Les indicateurs créés par techniques d'agrégation

1.32 - Appel à l'analyse factorielle

1.33 - Utilisation de l'indicateur pour la construction du découpage

SECTION 2 - COMPARAISON de DECOUPAGES DONNES

p 50

2.1 - Cas des coupages emboîtés

2.11 - Etude de la variation d'un indicateur

2.12 - Appel aux techniques purement statistiques

2.2 - Cas des coupages non emboîtés

2.3 - Précarité des résultats

CHAPITRE 3 - TECHNIQUES DE MODELISATION

SECTION 1 - GENERALITES sur les MODELES

p 62

1.1 - Modèles conceptuels et modèles mathématiques

1.11 - Modèle conceptuel

1.12 - Modèle mathématique

1.2 - La causalité dans les modèles mathématiques

1.21 - Causalité au niveau des variables

1.22 - Causalité entre sous-modèles

1.3 - Incidence des relations causales sur la structure du modèle mathématique

1.31 - Les différents types de système linéaire

1.32 - Décomposition des systèmes de Cramer

1.4 - Différentes techniques de détermination des variables endogènes

1.41 - Précision des résultats issus d'un modèle opérationnel

1.42 - Mode de variation des variables

1.43 - Solution analytique ou appel à la simulation

SECTION 2 - UNE TECHNIQUE PRIVILEGIEE des MODELES MATHÉMATIQUES : la REGRESSION MULTIPLE

p 72

2.1 - La régression multiple dans le cadre des modèles déterministes

2.2 - La régression multiple dans le cadre des modèles probabilistes

2.3 - Régression non linéaire

2.4 - Problème posé par la stratification d'une variable endogène

CHAPITRE 4 - ANALYSE THEORIQUE de la REPARTITION des HOMMES entre les DIFFERENTES ZONES du TERRITOIRE

SECTION 1 - MODIFICATION de la REPARTITION des HOMMES pour des RAISONS d'EMPLOI LIEES aux DECISIONS de LOCALISATION d'ENTREPRISES p 86

1.1 - L'analyse de localisation des entreprises dans l'optique des coûts

1.11 - Coûts d'obtention de facteur de production

1.12 - Coûts de distribution des produits

1.13 - Confrontation des coûts d'obtention de facteurs et de distribution de produits

1.14 - Evolution des données du problème

1.2 - Localisation induite

1.21 - Approche dans une optique descriptive

1.22 - Approche dans une optique prévisionnelle

1.23 - Le TEI interzonal

1.3 - Localisation autonome

1.31 - Sous-modèle d'implantation industrielle du Bass

1.32 - Le sous-modèle d'allocation industrielle du modèle du Bassin de la Susquehanna

1.33 - Le sous-modèle d'implantation industrielle du AAM

SECTION 2 - MODIFICATIONS de la REPARTITION des HOMMES LIEES aux DECISIONS INDIVIDUELLES p 101

2.1 - Modèles descriptifs

2.11 - Les modèles en terme de flux

2.12 - Modèles migratoires à la marge

2.2 - Modèles prévisionnels

2.21 - Utilisation des modèles descriptifs

2.22 - Modèles du type Markovien

SECTION 3 - BILAN PROVISOIRE

CHAPITRE 5 - Le FONCTIONNEMENT du MARCHE LOCAL de l'EMPLOI

SECTION 1 - ETUDE de la NOTION de MARCHE de l'EMPLOI

p 115

- 1.1 - Le marché de l'emploi dans la micro-économie
 - 1.11 - Le marché de l'emploi chez les classiques
 - 1.12 - Le marché de l'emploi chez les néoclassiques
 - 1.13 - Tendances récentes
- 1.2 - Le marché de l'emploi dans le cadre de la macro-économie
- 1.3 - Conception retenue du marché de l'emploi
 - 1.31 - Le cadre spatial
 - 1.32 - L'ensemble des mécanismes
 - 1.33 - Remarques sur l'unicité du marché de l'emploi

SECTION 2 - IMPLICATION de la DEFINITION du MARCHE de l'EMPLOI

p 125

- 2.1 - Remarques sur la notion d'état potentiel
 - 2.11 - Influence de la période retenue sur la notion d'état potentiel
 - 2.12 - Mesure de l'état potentiel
 - 2.13 - Interaction entre les concepts d'état potentiel et de mécanismes d'ajustement
- 2.2 - La notion de mécanisme d'ajustement
 - 2.21 - Les différents types de mécanismes d'ajustement
 - 2.22 - Recherche d'une typologie des mécanismes d'ajustement
- 2.3 - Problèmes posés par les délais d'ajustement
 - 2.31 - Connaissance des phénomènes et choix de la période
 - 2.32 - Problèmes posés par le choix d'un délai d'ajustement inférieur à la période intercensitaire

SECTION 3 - Les GRANDES OPTIONS du MODELE

p 133

- 3.1 - Le clivage entre emplois fondamentaux et emplois induits
- 3.2 - Les autres options

CHAPITRE 6 - Le CHOIX du DECOUPAGE SPATIAL

SECTION 1 - TRAVAUX EFFECTUES dans le CADRE du SESAME

p 138

1.1 - Etude d'un indicateur

1.11 - Fondements de l'étude

1.12 - Analyse des résultats

1.2 - Etude de plusieurs indicateurs

1.21 - Fondements de l'étude

1.22 - Analyse des résultats

SECTION 2 - IMPLICATIONS du CHOIX du DECOUPAGE DEPARTEMENTAL

p 150

CHAPITRE 7 - REMARQUES SUR CERTAINES TECHNIQUES ECONOMETRIQUES RETENUES

- SECTION 1 - LOI d'EVOLUTION de SPECIFICITES ZONALES OBSERVEES p 152
- 1.1 - *Quelques régularités d'évolution concevables*
 - 1.11 - Reconduction de la structure zonale observée
 - 1.12 - Evolution de la structure zonale observée
 - 1.2 - *Problèmes posés par les lois d'évolution de spécificités zonales*
 - 1.21 - Problème de cohérence
 - 1.22 - Signification des coefficients de spécificité
 - 1.23 - Techniques d'analyse du passé et techniques de prévision
- SECTION 2 - LOI d'EVOLUTION ECONOMIQUE : APPEL aux TECHNIQUES de REGRESSIONS MULTIPLES p 160
- 2.1 - *Cas des modèles déterministes*
 - 2.11 - Signification des écarts résiduels
 - 2.12 - Utilisation des coefficients résiduels
 - 2.2 - *Cas des modèles probabilistes*
 - 2.21 - Nécessité de l'introduction d'hypothèses complémentaires
 - 2.22 - Rejet de certaines variables exogènes
 - 2.23 - Signification des écarts résiduels
 - 2.3 - *Remarques*
 - 2.31 - Problèmes du choix entre modèles en taux et modèles en effectifs
 - 2.32 - Le programme de régression optimale utilisé
- SECTION 3 - LE PROBLEME de la COHERENCE NATIONALE p 168
- 3.1 - *Interaction entre le sous-modèle spatialisé et le sous-modèle national*
 - 3.11 - Le rôle des variables dans chaque sous-modèle
 - 3.12 - Signification des mécanismes de cohérence
 - 3.2 - *Quelques mécanismes de cohérence possibles*
 - 3.21 - Cadrage sur une donnée nationale
 - 3.22 - Cadrage sur plusieurs données nationales dépendantes

CHAPITRE 8 - Le SOUS-MODELE MIGRATIONS INTERIEURES

Section 1 - Mesure des migrations

p 176

1.1 - Problèmes posés par la définition du concept de migration

1.11 - Migrations et migrants

1.12 - Conséquences pratiques des simplifications adoptées sur la saisie concrète du phénomène migratoire

1.2 - Les données migratoires départementales

1.21 - Définition retenue

1.22 - Place des migrations départementales dans le temps

Section 2 - Etude des migrations interdépartementales 1962-1968

p 183

2.1 - Origine et destination des migrants

2.11 - Quelques grandes caractéristiques

2.12 - Direction et intensité des flux

2.2 - Influence de la structure par âge

2.21 - Etude globale

2.22 - Etude descriptive départementale de la structure par âge des migrants départementaux

2.3 - Activité des migrants interdépartementaux

2.31 - Remarques d'ensemble

2.32 - Etude départementale du niveau d'activité des migrants

2.33 - Etude départementale des migrants par CSP

2.34 - Synthèse par la technique de régression

2.4 - Influence du revenu sur les migrations d'actifs

Section 3 - Recherche d'un modèle explicatif des migrations apparentes intérieures

p 205

3.1 - Hypothèses testées et leurs indicateurs

3.2 - Détermination du sous-modèle immigrations intérieures et ses résultats pour 1968

3.21 - Etude des résultats de l'approche en effectifs

3.22 - Etude des résultats de l'approche en taux

3.23 - Choix du sous-modèle immigrations intérieures et ses résultats pour 1968

3.3 - Détermination du sous-modèle émigrations intérieures

3.31 - Etude des résultats de l'approche en effectifs

3.32 - Etude des résultats de l'approche en taux

3.33 - Choix du sous-modèle émigrations intérieures et résultats 1968

3.4 - Conclusions partielles

3.41 - Remarques sur les liaisons retenues

3.42 - Remarques sur les limites des techniques retenues

CHAPITRE 9 - DETERMINATION de la DEMANDE POTENTIELLE d'EMPLOI

SECTION 1 - LES MIGRATIONS EXTERIEURES

p 229

- 1.1 - *Remarque sur le problème de mesure des migrations avec l'étranger*
 - 1.11 - Les grandeurs observées
 - 1.12 - Détermination des émigrants extérieurs
- 1.2 - *Etude des immigrants extérieurs 1962/1968*
 - 1.21 - Etude globale
 - 1.22 - Etude départementale des migrants extérieurs
- 1.3 - *Recherche d'un sous-modèle des immigrants extérieurs*
 - 1.31 - Choix des hypothèses
 - 1.32 - Etude des résultats
 - 1.33 - Choix du sous-modèle et étude de ses résultats

SECTION 2 - DETERMINATION de la POPULATION POTENTIELLE en 1968

p 247

SECTION 3 - DETERMINATION de la DEMANDE d'EMPLOI POTENTIELLE

p 248

- 3.1 - *Causes de fluctuation du taux d'activité*
 - 3.11 - Remarque préalable sur les données de base
 - 3.12 - Analyse des écarts
- 3.2 - *Sous-modèle taux d'activité potentiel*
 - 3.21 - Choix d'hypothèses d'évolution
 - 3.22 - Mise en application du principe retenu
 - 3.23 - Biais introduits
 - 3.24 - Stratification de la population retenue
- 3.3 - *Analyse des résultats obtenus*

CHAPITRE 10 - Le SOUS-MODELE BTP

SECTION 1 - ETUDE DESCRIPTIVE du BTP

p 258

1.1 - *Les travaux effectués au niveau régional*

- 1.11 - L'influence de la population
- 1.12 - Répartition spatiale du marché
- 1.13 - Taille des entreprises et production
- 1.14 - Productivité dans le BTP

1.2 - *Les observations faites au niveau départemental*

- 1.21 - Evolution de l'emploi BTP entre 1954 et 1962
- 1.22 - Evolution de l'emploi BTP entre 1962 et 1968
- 1.23 - Evolution de l'emploi BTP et régression linéaire simple

SECTION 2 - RECHERCHE d'un MODELE EXPLICATIF DEPARTEMENTAL pour 1962

p 267

2.1 - *Hypothèse retenues pour test et leurs indicateurs*

- 2.11 - Ajustement de l'emploi à la demande de BTP
- 2.12 - Influence de l'urbanisation
- 2.13 - Influence de la présence d'une clientèle potentielle

2.2 - *Recherche de modèles portant sur des variables prises en valeur absolue*2.3 - *Recherche de modèle portant sur des variables prises en taux*

- 2.31 - Choix des dénominateurs
- 2.32 - Etude des résultats

2.4 - *Choix du modèle définitif et limite de ce modèle*

- 2.41 - Choix du modèle définitif
- 2.42 - Critique du modèle
- 2.43 - Etat potentiel 1968

CHAPITRE 11 - Le SOUS-MODELE AGRICULTURE

SECTION 1 - ETUDE DESCRIPTIVE de l'EVOLUTION de la POPULATION AGRICOLE

p 283

1.1 - Etude globale de l'évolution de la population active agricole entre 1954 et 1968

- 1.11 - Influence du sexe et du statut
- 1.12 - Influence de l'âge et du statut
- 1.13 - Migrations et emploi agricole

1.2 - Etude départementale de l'évolution de la population active agricole

- 1.21 - Entre 1954 et 1962
- 1.22 - Entre 1962 et 1968
- 1.23 - Synthèse par la technique de regression

SECTION 2 - RECHERCHE d'un MODELE d'EVOLUTION de l'EMPLOI AGRICOLE DEPARTEMENTAL

p 294

2.1 - Les modèles mathématiques existants

- 2.11 - Travaux de Klatzmann
- 2.12 - Le modèle agricole 1970 du ministère de l'agriculture

2.2 - Recherche d'un modèle

- 2.21 - Hypothèses retenues
- 2.22 - Modèle en valeur absolue
- 2.23 - Modèle en taux
- 2.24 - Modèle retenu

2.3 - Analyse des résultats du modèle pour 1968

- 2.31 - Les résultats obtenus
- 2.32 - Critique du modèle.

CHAPITRE 12 - L'ETUDE du SOUS-MODELE TERTIAIRE

SECTION 1 - APPROCHE THEORIQUE de la NOTION du TERTIAIRE p 308

SECTION 2 - DESCRIPTION de l'EVOLUTION du TERTIAIRE en FRANCE p 311

2.1 - L'évolution du tertiaire sur longue période et comparaisons internationales

2.2 - Analyse de l'évolution récente (1954 - 1968) du tertiaire par catégorie d'activité et département

2.21 - Description de l'évolution du tertiaire

2.22 - Test de quelques hypothèses

2.3 - Mobilité professionnelle et tertiaire

SECTION 3 - RECHERCHE d'un SOUS-MODELE TERTIAIRE p 341

3.1 - Approche en effectifs

3.11 - Hypothèses

3.12 - Recherche d'un sous-modèle pour la variation 54 - 62 du tertiaire

3.13 - Recherche d'un sous-modèle pour le tertiaire 1962

3.2 - Approche en taux

3.21 - Hypothèses

3.22 - Recherche d'un sous-modèle en taux

3.3 - Le sous-modèle tertiaire retenu

3.31 - Choix du sous-modèle

3.32 - Examen des résultats 62 du sous-modèle retenu

3.33 - Examen des résultats 68 du sous-modèle tertiaire

CHAPITRE 13 - LE SOUS-MODELE "SECONDAIRE ": ETUDE DES FACTEURS DE LOCALISATION

SECTION 1 - LES ANALYSES DE L'EMPLOI INDUSTRIEL ORIENTEES VERS LA PREVISION SPATIALISEE

p 355

1.1 - *La démarche démographique*

1.11 - Les bases de la démarche démographique

1.12 - Les travaux de l'Institut d'Etudes de l'emploi

1.2 - *L'analyse globale menées dans l'optique prévisionnelle*

1.21 - Les études descriptives préalables

1.22 - La recherche d'un modèle prévisionnel

SECTION 2 - RECHERCHE D'UN SOUS-MODELE DE DETERMINATION DES EMPLOIS INDUSTRIELS

p 365

2.1 - *Présentation de la démarche retenue*

2.11 - Stratification des emplois industriels

2.12 - Principes d'analyse retenus

2.2 - *Analyse des facteurs de localisation*

CHAPITRE 14 - RESULTATS GLOBAUX - CONCLUSIONS

SECTION 1 - RESULTATS d'ENSEMBLE	p 369
1.1 - Options complémentaires	
1.2 - Examens d'ensemble des résultats	
SECTION 2 - RECHERCHE d'une TYPOLOGIE d'AJUSTEMENT	p 372
2.1 - L'analyse systématique des écarts	
2.2 - L'appel à un modèle de référence	
SECTION 3 - CONCLUSIONS	p 377
3.1 - Remarques sur l'application des techniques économétriques utilisées	
3.11 - Techniques relatives au choix du découpage	
3.12 - Remarque sur les techniques de préhension des régularités	
3.2 - Remarques sur la modélisation	
3.21 - Critiques internes	
3.22 - Critiques externes	
3.3 - Vers un dépassement	

INTRODUCTION

"Au commencement était le chaos" affirment sous divers aspects, les grandes traditions de l'humanité. C'est dire que la notion d'hétérogénéité est inhérente à celle d'espace. Très tôt l'homme a cherché à "domestiquer" cet espace à son profit et chaque civilisation a marqué de son empreinte (défrichage, infrastructures routières et urbaines, grands monuments, ...) la conquête progressive de l'espace. Dans de très nombreux cas la conquête guerrière du territoire a constitué un substitut commode d'une plus grande maîtrise des sols et peut être considérée à ce titre comme l'un des mécanismes réducteurs des inégalités entre peuples, employés par l'humanité.

C'est avec l'avènement de l'espace fini pour les civilisations techniquement avancées, et donc de la fin de solutions de facilités, que, vers la fin du 19ème siècle, se produit l'émergence de la réflexion sur l'espace⁽¹⁾ et la prise de conscience de la nécessité de ce qu'il est convenu d'appeler l'aménagement du territoire. Là encore les impératifs politiques, générateurs d'une problématique nouvelle, ont constitué un préalable nécessaire à l'acquisition de processus cognitifs nouveaux et de schémas de pensée appropriés.

Les conséquences politiques et économiques de la crise des années 30 provoquent les premières mesures *coordonnées* à vocation spatiale ; aux Etats-Unis c'est la mise sur pied de la Tennessee Valley Authority qui cherche à résoudre le problème de l'emploi dans le Sud des E.U. en empiétant sur de multiples compétences territoriales et sectorielles ; en Grande-Bretagne c'est, durant la seconde guerre mondiale, le recensement des *depressed areas* afin de promouvoir des actions concertées d'ensemble ; enfin, en France, si l'on écarte les décentralisations effectuées pour raisons stratégiques pendant l'entre deux guerres, c'est à partir de 1945, sous l'impulsion de M. Claudius PÉTIT alors Ministre de la Reconstruction, qu'une vision d'ensemble de la reconstruction des villes, liée à une décentralisation industrielle se fait jour et constitue à ce titre le premier effort d'aménagement volontaire du Territoire.

(1) Voir par exemple le résumé de la genèse de l'analyse régionale écrit par J.R. Meyer intitulé "Regional Economics : a survey" dans l'American Economic Review, vol 53 (1963). Les précurseurs sont nombreux (voir Dockes : "L'espace dans la pensée économique du XVIe au XVIIIe siècle" (Flamarion - 1969). Certains travaux relativement anciens comme ceux de Von Thünen au siècle dernier restent des apports fondamentaux ce qui n'empêche pas ce courant de pensée d'être resté marginal jusqu'à la seconde moitié de ce siècle.

La conception de l'Aménagement du Territoire étant par essence politique, il convient, en se cantonnant à la France d'en rappeler les finalités telles que les voit le promoteur de la DATAR, M. GUICHARD qui écrivait⁽¹⁾ en 1965 :

Politique, l'aménagement du territoire l'est [donc] pleinement dans toutes ses dimensions, dans le quotidien et dans ses objectifs à plus long terme. Quels sont ces derniers ?

Objectifs d'ordre public au sens élevé du terme : les déséquilibres ne doivent pas s'accroître dangereusement ; dans certaines régions le vide serait insupportable.

Objectifs militaires : nous avons [déjà] évoqué le transfert de certaines industries de défense nationale.

Objectifs sociaux : il s'agit de veiller à ce que les différences de revenus ne soient pas trop importantes de région à région. Cette recherche de la parité, de plus en plus exigée par l'opinion publique, trouve cependant ses limites : elle ne doit pas conduire à une politique d'assistance ou à des interventions antiéconomiques qui figeraient les situations et, en dissimulant provisoirement le mal, le prolongeraient. Une reconversion permanente et une certaine mobilité géographique ou sociale sont indispensables.

Objectifs culturels : la puissance publique ne peut se désintéresser de cette mission à travers, par exemple, la régionalisation de la radiodiffusion et de la télévision, la décentralisation théâtrale et musicale, l'implantation universitaire.

Objectifs d'ordre économique enfin, car l'aménagement du territoire doit permettre une meilleure progression économique d'ensemble.

Le problème politique étant ainsi posé il est nécessaire d'examiner avec l'oeil de l'économiste ce qui motive les objectifs 1, 3 et 5 proposés par M. GUICHARD, abstraction faite de toutes considérations éthiques, et que l'on pourrait appeler les fondements objectifs de l'Aménagement du Territoire. Une fois ceci posé, il convient de se demander dans quelle mesure l'Économie peut se mettre au service du Politique pour lui permettre de réaliser ses finalités. Dans cette aide la connaissance, même sommaire, de certains phénomènes et de leurs interactions jouent un rôle considérable et la constitution de modèles économiques peut en aménagement du territoire constituer un outil précieux de réflexion comme cela est déjà le cas dans d'autres domaines de la politique économique. Les efforts entrepris dans ce sens sont récents en France et un certain nombre d'"études de reconnaissance" sont encore nécessaires avant d'aboutir. C'est dans ce cadre que s'inscrivent les travaux présentés ici. On examinera successivement ces différents points.

(1) O. GUICHARD "Aménager la France" (Laffont - Gauthier - 1965).

SECTION 1 - Les FONDEMENTS OBJECTIFS de l'AMENAGEMENT du TERRITOIRE

On trouve à la base des objectifs proposés, le fait qu'il existe des disparités qu'il convient de réduire. Il est nécessaire de regarder sur quoi repose ces 2 affirmations.

1.1 - Causes des disparités économiques entre les diverses zones du territoire

Elles peuvent être regroupées en trois catégories, lesquelles d'ailleurs dans le long terme interfèrent les unes sur les autres : les dotations de la région, le jeu de l'économie de marché et le jeu des politiques.

1.11 - Différence de dotations

Par dotation de la région il faut entendre l'ensemble des ressources naturelles (terres arables, ressources énergétiques et minières) et ce qui correspond en quelque sorte à l'héritage historique de la région (patrimoine immobilier et culturel, infrastructures de communications, équipements sociaux, structures démographiques et socio-professionnelles, capital industriel...).

Ces dotations présentent une grande inertie à la modification au cours du temps, mais toutes n'agissent pas avec la même force sur le problème de l'emploi.

Ces modifications sont dues dans les pays capitalistes au jeu de l'économie de marché mais aussi, l'Etat intervenant de plus en plus, au jeu des politiques.

1.12 - Rôle de l'économie de marché

Le système économique n'est pas neutre dans l'évolution des disparités spatiales. L'économie dite libérale par son système de prix, favorise une certaine réallocation des facteurs productifs. La répartition spatiale des dotations en capital peut changer aussi bien lors des investissements de remplacement que lors des investissements nouveaux tandis que les phénomènes migratoires sont générateurs de processus cumulatifs de déclin ou d'expansion différent selon les zones, à moins qu'ils ne soient provoqués par l'évolution des dotations en capital. Par ailleurs ces mouvements de facteurs productifs sont liés en grande partie à des croissances régionales différenciées.

1.13 - Jeu des politiques

Ces politiques sont en fait de deux types : les politiques globales à incidence spatiale et les politiques volontaristes d'aménagement du territoire.

- a) Les politiques globales à incidence spatiale les plus évidentes sont sectorielles ; toutes ont ceci de commun, qu'elles déterminent les conditions de survie des entreprises marginales et donc du maintien des populations actives s'y rattachant. On peut citer dans ce domaine la politique des prix d'intervention agricoles ou les différentes mesures douanières qui ne sont pas autre chose que la protection de larges secteurs d'activités (sidérurgie par exemple).

La politique fiscale n'est pas neutre non plus du point de vue spatial. Les modalités d'assiette et de taux peuvent engendrer une disparition plus ou moins rapide de certaines catégories d'emplois comme celle des petits commerçants ou des paysans.

Enfin, il faut noter l'influence des secteurs contrôlés par l'Etat. C'est ainsi que la détermination des tarifs de l'énergie (au coût marginal pour l'électricité par exemple) et celle des transports par fer, qui toutes deux relèvent d'entreprises nationalisées et donc contrôlées par l'Etat peuvent "ouvrir" de façon plus ou moins grande une région au marché national et influencer ainsi de façon décisive sur ses possibilités d'expansion.

- β) Les politiques volontaristes d'aménagement du territoire sont assez récentes en France⁽¹⁾. Elles sont l'oeuvre, au lendemain de la guerre, de la Direction de l'Aménagement du territoire du Ministère de la Reconstruction et de l'Urbanisme qui s'appuie sur une réglementation nouvelle des permis de construire, des aides financières (création en 1950 du Fond national d'aménagement du territoire, en 1953 du Fond d'adaptation et de développement, etc...) puis on assiste à une véritable prolifération d'organismes appelés à connaître tel ou tel aspect de l'aménagement du territoire. En 1955 c'est la création des programmes d'action régionales et en 1957 le ministère de la reconstruction et du Logement crée des plans régionaux d'aménagement du territoire. Tous deux fusionnent en 1957 pour donner naissance au "Plan régional de développement économique et social d'aménagement du territoire".

(1) Pour la période 1945 - 1955 voir : "Décentralisation industrielle, reconversion, aménagement du territoire" par J. LAJUGIE dans la Revue juridique et économique du Sud-Ouest 1956 n° 2 p 355.

Pour la période 1956 - 1963 voir : "La politique française de développement économique régionale de 1958 à 1963", INRJSO enfin pour la période récente voir les annexes régionalisées au projet de loi de finances (tome 1 intitulée "Les principaux éléments de la politique suivie pour l'aménagement du territoire et l'action régionale").

En 1960 est institué un comité interministériel permanent pour les problèmes d'aménagement du territoire et d'action régionale. Mais c'est en 1963 que l'administration se dote de moyens cohérents et adaptés pour aborder les problèmes dans leur ensemble avec la création de la Délégation à l'Aménagement du Territoire et à l'action régionale (DATAR) et de la Commission Nationale de l'Aménagement du Territoire (CNAT) tandis que la consultation locale s'effectue par l'intermédiaire des Comités d'expansion Economique Régionaux (CODER) créés l'année suivante à l'occasion de la préparation du 5ème plan. Et depuis c'est une marche hésitante vers une certaine régionalisation à laquelle on assiste sans que l'on puisse préjuger du partage final des pouvoirs et de la compétence entre les divers niveaux d'organisation de l'Etat. Quoiqu'il en soit depuis 25 ans un certain nombre d'actions sont menées par divers organismes.

Les actions d'aménagement du territoire sont soit de type ponctuel : aide au développement d'un type de transport (aérotrain, aquaplane), actions spécifiques sur une zone (aménagement de la Corse, Fos) soit de type général : aides à l'implantation industrielle à caractère systématique mais différencié suivant les régions. Cette aide sous forme de prime de développement industriel, d'adaptation industrielle, de décentralisation et d'allègements fiscaux⁽²⁾ peut représenter jusqu'à 20 % de l'investissement total.

La résultante de ces dotations et de ces politiques est une modification de la prospérité relative de chaque région, au cours du temps.

(2) "Aides au développement régional" (DATAR - Chambre de Commerce et d'Industrie de Paris 1969). Ces aides sont non déductibles sur le plan fiscal.

1.2 - Nécessité d'une intervention

Deux thèses s'affrontent sur cette importante question :

1.21 - a) La thèse non interventionniste

En se fondant sur l'analyse marginale classique et sans tenir compte de ses nombreuses hypothèses restrictives, l'école libérale, qui inspire encore certains courants politiques, estime que de telles interventions sont inutiles et même nuisibles.

Inutiles parce que dans une économie de marché des mécanismes régulateurs existent qui font que les zones défavorisées attirent par leurs bas salaires de nouvelles industries et que des actifs émigrent vers les zones de hauts salaires. De ce fait les disparités de prospérité doivent disparaître.

Ces interventions sont en outre nuisibles car elles pénalisent les zones en forte expansion, qui, elles, ne sont pas aidées par l'Etat.

1.22 - Critique de la thèse non interventionniste

NEEDLEMAN⁽¹⁾ avance deux arguments en faveur de l'interventionnisme :

- tout d'abord, on ne connaît pas d'exemple où les effets de ces mécanismes régulateurs aient laissé de traces tangibles, ou s'ils agissent, ils le font dans le très long terme et donc trop tard.
- ensuite, en se référant à l'analyse coût - bénéfice la mise au travail d'un chômeur dans une zone défavorisée, lorsque sa migration est difficile et donc ne se réalise pas entraîne par le biais de l'accroissement de demande solvable et de multiplicateur de type Keynésien, un accroissement de la quantité de biens disponibles par tête, et, par le biais fiscal (impôts directs et indirects), un allègement de l'effort fiscal par tête (ce raisonnement s'effectuant à un niveau global, pour l'ensemble du pays).

(1) NEEDLEMAN : "Regional Analysis" (Penguin - 1968).

A ces deux arguments nous pouvons en ajouter deux autres :

Il faut bien constater que les conditions requises par la théorie, nécessaires au bon fonctionnement de ces mécanismes sont loin d'être toutes réunies. L'hypothèse du très grand nombre d'employeurs et d'employés s'affrontant individuellement sur un marché, est détruite par l'existence d'oligopoles et de monopoles ainsi que celle d'accords professionnels nationaux, pour ne prendre que celle-là.

En outre, il ne faut pas oublier qu'un afflux de migration dans une zone, nécessite un accroissement du parc de logement, des équipements sociaux etc ... et entraîne dans des zones déficitaires un sous-emploi de ces mêmes "structures d'accueil"; ces déséconomies externes pour la société peuvent annuler l'intérêt du transfert d'actifs dans certaines zones, même s'il y a croissance du PNB, à cause de la mauvaise allocation des ressources pour l'ensemble de la collectivité, car cette croissance peut à la limite correspondre à une diminution de bien-être (le problème de la mesure de la croissance a donc aussi une dimension spatiale).

1.3 - La connaissance, condition permissive d'actions efficaces en aménagement du territoire

Les interventions de l'Etat sont limitées par le montant d'une enveloppe budgétaire difficile à augmenter, du moins en termes relatifs. Les dépenses prévues au Budget traduisent concrètement les priorités décidées par l'Etat et parmi elles, les préoccupations d'aménagement du Territoire. En pratique l'impact de ces diverses dépenses peut s'avérer non conforme aux volontés politiques sous jacentes et entraîner une distorsion de l'échelle des priorités par le jeu de répercussions en cascades imprévues ou parce que le type d'action engagée n'est pas celui qui a la plus grande efficacité, à dépense constante (ce qui entraîne une mauvaise allocation des ressources).

Le nombre et la complexité des mécanismes économiques mis en cause sont tels qu'une connaissance intuitive de l'ensemble de ces phénomènes est difficile et que la comparaison quantifiée de plusieurs interventions possibles est irréalisable d'emblée. Il faut d'abord démonter les mécanismes économiques mis en cause et quantifier l'impact de la variation d'un phénomène sur un autre.

Mais il faut bien constater que l'aspect spatial de la connaissance phénoménologique accuse en France un assez grand retard, parce que dans bien des cas, et en particulier pour les données en valeur, les filières classiques de production de l'information ont entraîné la perte de la dimension spatiale des phénomènes étudiés, de par les procédures adoptées d'obtention de grandeurs nationales. Il s'en suit qu'a fortiori la connaissance des mécanismes est fragmentaire voire nulle. C'est ainsi que l'impact régional d'une modulation spatiale des taux d'aides à la décentralisation est actuellement, impossible à déterminer(1)

D'ailleurs M. GUICHARD, particulièrement conscient de ces lacunes écrit à ce propos :(2)

L'aménagement doit créer ses propres instruments de mesure : ceux qui lui permettront de poser les problèmes en fonction de leurs localisations et des influences réciproques qu'exercent dans l'espace, les actions entreprises. Par exemple, ramenant cette préhension de la réalité à des schémas économiques, il faudra, un jour, dégager des paramètres - tarifs, mode de transport et distance - pour une estimation relative de l'espace géographique, apprécier le temps annuel d'ensoleillement d'une zone ou le rôle des forêts dans ses possibilités d'essor culturel et commercial.

L'appel aux disciplines classiques, pourvu qu'elles soient bien intégrées les unes aux autres, constituent l'arsenal dont dispose actuellement l'aménagement du territoire. Parmi elles, la statistique l'économétrie, les sciences humaines et sociales jouent un rôle essentiel car elles apportent, chacune pour sa part, les moyens d'une méthode de synthèse propre à l'aménagement".

C'est dans le cadre de cette ligne directrice qu'est créé en 1968 le Système d'Etudes du Schéma d'Aménagement de la France (SESAME) dont la vocation était de forger une méthodologie appropriée aux problèmes d'aménagement et de préparer des outils d'aide à la décision. L'aspect économétrique était bien entendu présent dès les débuts de cet entreprise et les travaux exploratoires du groupe modèles ont abouti à l'établissement⁽³⁾ d'un cadre de référence ambitieux d'un modèle spatial à moyen terme axé sur les problèmes d'emplois localisés. Une telle démarche est à bien des égards dangereuse, car génératrice de déceptions, si l'on s'attend à une représentation fidèle de la réalité mais elle n'en est pas moins nécessaire car comme le disait Valéry, "ce qui est simple est toujours faux, mais ce qui ne l'est pas est inutilisable".

(1) En fait on ignore l'incidence réelle de ces aides et il n'est pas sûr que la décision de très nombreuses décentralisations ait été sérieusement influencée par ces types d'incitation. On ne peut pas en dire autant de certaines mesures coercitives qui ont stoppé l'extension de certaines activités en Région Parisienne.

(2) Ouvrage cité.

(3) Voir le rapport de synthèse du groupe "Modèles" du SESAME intitulé "Un cadre pour l'analyse du développement interrégional" (mars 1971) à l'élaboration duquel nous avons participé.

Définissons brièvement l'approche par l'analyse de système avant de présenter celle qui a été menée dans le cadre du SESAME.

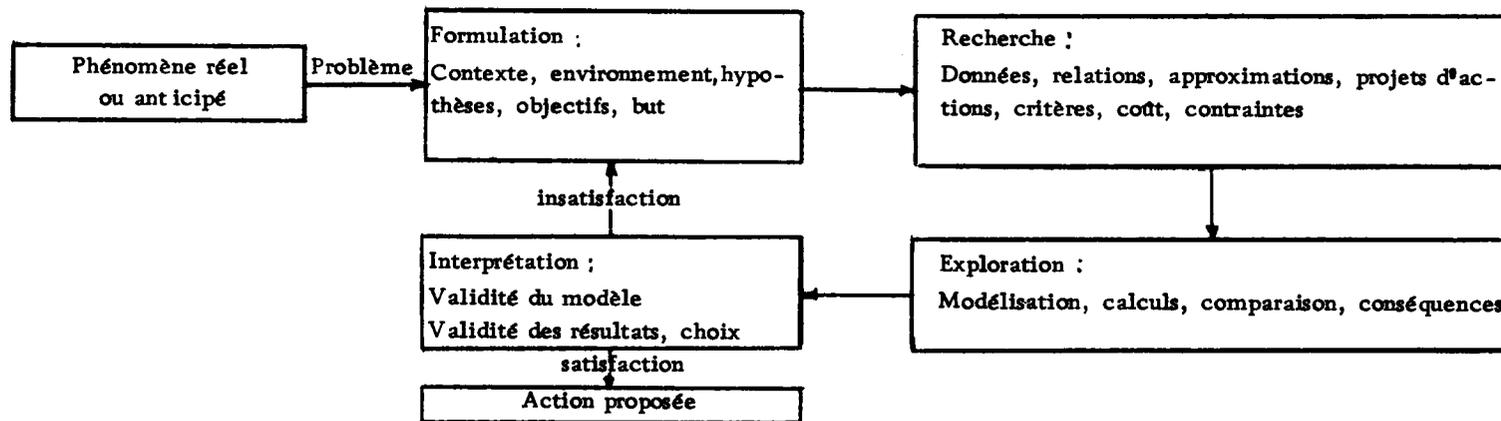
1.31 - Présentation sommaire des techniques d'analyse de système

a) Définition de l'analyse de système

Pour trouver des solutions à des problèmes d'une complexité sans cesse croissante de nombreuses disciplines ont été contraintes depuis la dernière guerre mondiale de procéder à un effort de formalisation sans cesse plus poussé. Cette systématisation de la démarche analytique sous l'impulsion en particulier de la Rand Corporation s'est opérée en faisant des emprunts plus ou moins importants à la théorie mathématique des systèmes⁽¹⁾, mais on peut dire que dans les disciplines économiques, l'apport a surtout consisté à mieux définir le rôle des techniques de modélisation dans la résolution des problèmes posés.

On peut définir un système comme un ensemble d'éléments reliés entre eux par des fonctions pour l'accomplissement d'objectifs, cet ensemble finalisé se voulant la représentation formelle d'une réalité tangible. Il faut souligner que dans l'analyse de système on ne s'intéresse pas à un élément en tant que tel mais avant tout aux relations qu'il entretient avec les autres éléments.

A partir de l'analyse générale effectuée par Quade⁽²⁾ on peut schématiser la démarche de l'analyse de système comme suit :



(1) Voir en particulier C. Ribeill "Théorie et analyse des systèmes" (Métra 1970) et le rapport du contrat SEMA-DGRST : "éléments pour une synthèse sur les systèmes à auto-organisation" par Fortet et Leboulanger (Métra 1967)

(2) E.S. Quade "Analysis for Military Decisions" chap. 8 (Rand Marc Nally & Co 1964) et surtout "Systems analysis and Policy planning, applications in defense" (Elsvier 1968) qui réunit une série de textes choisis par E.S. Quade et WI Boucher.

Selon la complexité de la représentation formelle retenue et donc le nombre de liaisons indépendantes agissant sur le ou les élément(s) que l'on désire contrôler, le système auquel on aboutit est un système à état, c'est-à-dire descriptif du jeu d'un certain nombre d'éléments, et fait donc appel à la simulation, ou alors, dans les cas les plus simples, le système auquel on abouti est un système à but et l'on applique alors la théorie du contrôle optimal. L'analyse de système du SESAME a conduit à la définition d'un système à état.

β) Apport de l'analyse de système

A la différence de la démarche scientifique classique en recherche fondamentale qui, visant à l'établissement de principes généraux, s'interdit de se lier par un objectif, l'analyse de système s'inscrit dans un cadre finalisé et ses enseignements n'auront pas vocation de généralité : il ne s'agit plus de comprendre ou prédire mais de préconiser une solution. A côté de cette spécificité de l'analyse de système on peut noter que l'emploi des techniques d'analyse a de tout temps fondé les progrès de l'humanité; aussi peut-on se demander si l'analyse de système présente un apport particulier. A notre avis l'apport essentiel se situe sur le plan heuristique dès lors que le problème abordé est complexe:

En reprenant la terminologie que l'on vient d'introduire, on peut dire que la complexité d'un problème tient d'une part à l'aspect combinatoire des relations entre éléments, et d'autre part à la possibilité de définir précisément certains de ces éléments.

En ce qui concerne les relations, l'analyse de système permet de procéder à une première simplification en éliminant toutes celles qui sont "sans importance" compte tenu de l'objectif que l'on s'est assigné. Une seconde simplification est opérée par l'utilisation systématique de graphes de causalité qui éliminent les relations de causalité indirecte et permettent de mettre clairement en évidence des boucles de causalité; il est bien certain que dans le cas de problèmes complexes la linéarité du discours empêche de faire comprendre clairement et rapidement la subtilité de certaines interdépendances (ce que permet l'emploi d'un graphe commenté).

Pour la définition des éléments deux commodités sont offertes par l'analyse de système : partir d'éléments plus ou moins agrégés et utiliser des "boîtes noires". L'analyse d'un problème est faite de tâtonnement et requiert plusieurs itérations, il peut alors être utile dans les premières phases de travailler à un niveau relativement agrégé (par exemple, l'emploi total sans aucune ventilation) pour dégrossir le problème, quitte à affiner l'analyse au cours de phases ultérieures. Quant aux boîtes noires, elles sont de la plus grande utilité dans la phase de recherche car elles permettent de ne pas bloquer l'analyse. La définition précise de la boîte noire (dont on connaît cependant les données d'entrées et de sortie) est reportée à une étape ultérieure.

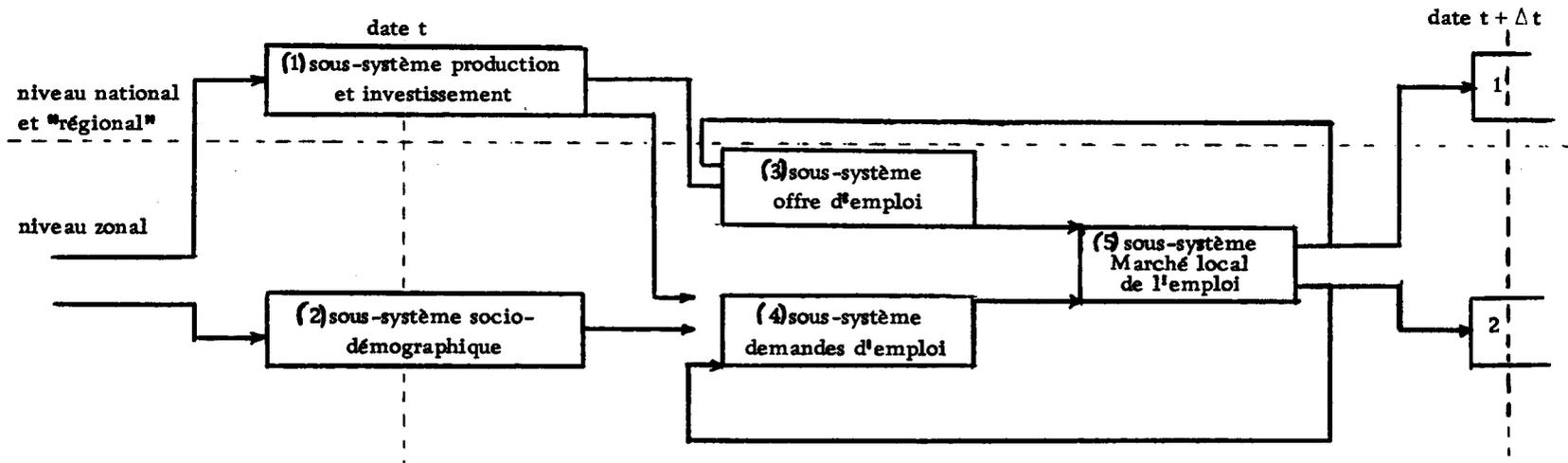
L'emploi des techniques d'analyse de système force donc le chercheur à savoir à tout instant où il en est très exactement dans la résolution du problème posé et lui permet de définir correctement les priorités à suivre. Deux remarques sont toutefois à faire. Tout d'abord lorsqu'un problème est peu complexe, l'apport réel de l'analyse de système est faible, voire nul, l'esprit humain étant alors capable de maîtriser simultanément tous les paramètres de ce problème.

Par ailleurs, lorsqu'un problème complexe est résolu, présenter la démarche suivie n'offre plus qu'un intérêt "historique" et le résultat final peut être présenté sous diverses formes. Il s'ensuit que l'intérêt de l'analyse de système, qui est nous l'avons dit, essentiellement heuristique, peut ne pas apparaître à certain.

1.32 - L'analyse de système du SESAME

Pour répondre aux objectifs définis par O. Guichard, une analyse de système semblait indispensable compte tenu de la complexité des phénomènes mis en jeu et de leur connaissance fragmentaire, voire inexistante. Parmi les différents objectifs, celui qu'on a privilégié est celui de l'emploi spatialisé et de la prévision à moyen terme des tensions des marchés locaux de l'emploi dans le cadre d'une interdépendance entre les diverses unités spatiales, hiérarchisées à trois niveaux (national, régional, zonal).

La première phase de l'analyse de système, celle de la formulation aboutissait à une proposition d'architecture de modèle de développement interrégional qui comportait une centaine d'éléments et quelques 250 relations. Une présentation succincte des divers sous-systèmes envisagés même au schéma suivant :



C'est une étude de faisabilité sur une version simplifiée de cette étude de système qui nous a été confiée et qui est à l'origine des travaux que nous présentons ici.

SECTION II - DEFINITION du CADRE de cette ETUDE

Nous examinerons successivement l'objet de cette étude puis la démarche suivie :

2.1 - Définition de l'objet de cette étude

Des trois objectifs à caractère économique proposés par O. GUICHARD, celui de l'intégration de la dimension spatiale dans la recherche d'une optimisation de la croissance est de loin le plus ambitieux et son analyse concrète ne relève pas d'étude exploratoire mais peut dans le cas le plus favorable constituer le faite de l'édifice théorique à bâtir et tester.

Des deux qui restent l'un est orienté vers la réduction des disparité spatiales de revenus, l'autre vers les problèmes d'emploi. Le problème des disparités spatiales des revenus est comme on va le voir un problème en grande partie sectoriel, il était donc naturel de mettre d'abord l'accent sur les problèmes de l'emploi spatialisé.

2.11 - Remarques préalables sur les disparités spatiales des revenus

La connaissance des revenus en France est de notoriété publique, fort mauvaise. Les seules données à peu près fiables dont on dispose concernent les salariés et résultent de l'exploitation des déclarations patronales (états 2460).

En 1968 le salaire annuel moyen France entière s'élevait à environ 13 150 F et sur 90 départements, 85 connaissent un salaire annuel moyen inférieur à cette moyenne nationale. Doit-on en déduire que tous ces départements sont pour autant défavorisés ? Rien n'est moins sûr : prenons le cas limite d'un département hypothétique où tous les travailleurs seraient ouvriers et gagneraient 13 000 F, les actifs de ce département seraient nettement favorisés par rapport aux autres ouvriers bien que gagnant moins que la moyenne France entière. Il n'est pas correct dans ce cas de parler de disparités d'origine spatiale sans faire référence aux disparités sectorielles. Ces dernières sont en grande partie à l'origine des disparités spatiales de revenus.

Pour cerner de plus près le phénomène nous avons mené une rapide étude à partir de données par sexe et catégories socio-professionnelles (en 6 postes) concernant les salaires annuels moyens en 1968 et les effectifs correspondants exprimés en personne-année-travail. Ils s'avère (voir les cartes et tableaux à la base de ces commentaires : A 18 T1.18 à 20) que si l'on applique aux salaires par sexe et CSP d'un département la structure d'effectifs (en PAT)⁽¹⁾ observée France entière, pour obtenir un salaire calculé éliminant cet effet structurel, l'écart de salaire par rapport à la moyenne baisse dans des proportions très importantes. La moyenne des écarts relatifs des salaires départementaux par rapport au salaire moyen France entière qui était d'environ - 15 % passe à - 9 % c'est-à-dire qu'en première approche l'effet de structure explique de l'ordre du tiers des disparités. On peut compléter cette indication globale en indiquant que plus de la moitié des départements désavantagés voient leurs écarts diminuer de plus d'un tiers et que le cinquième des départements voyant leur position s'améliorer (6 seulement ne sont pas dans ce cas) voient leurs écarts amputés de plus de moitié.

La notion de catégorie socio-professionnelle est insuffisante pour caractériser l'emploi et une étude plus approfondie faisant intervenir en outre la notion de branche améliorerait sans doute l'explication de ces disparités. Il faudrait d'ailleurs si l'on veut procéder à une étude complète des avantages comparatifs des salariés, intégrer les disparités spatiales de niveau des prix⁽²⁾ pour raisonner en termes réels.

Toutefois ces premières indications laissent penser qu'une politique sectorielle des revenus a, au moins dans un premier temps, une "sensibilité" plus grande qu'une politique spatiale des revenus, au demeurant fort difficile à mettre en place. C'est donc plutôt sur les problèmes d'emploi que l'effort doit d'abord porter.

2.12 - Objet de l'étude

L'étude de faisabilité que nous avons menée consiste en une recherche de modèle correspondant aux sous-systèmes 3, 4 et 5, en supposant connues, sans erreur, la plupart des données issues des autres sous-systèmes. Les phases de recherche, d'exploration et d'interprétation de l'analyse de système, telle que E.S Quadeles définit, sont donc entièrement à effectuer et l'étude de faisabilité consiste à examiner si le bouclage sur la phase de formulation s'impose ou non - Précisons les données du problème :

(1) P.A.T. : Personne Année Travail

(2) Actuellement l'indice des prix est une double somme pondérée faisant intervenir la notion de catégorie de dépense et celle d'espace (au niveau des Directions Régionales de l'INSEE) mais malheureusement la première sommation s'effectue sur l'indice spatial.

Il s'agit à l'aide de techniques économétriques d'élaborer des modèles dont les paramètres auront été déterminés pour un passé antérieur à 1968, et de procéder pour cette date à la création d'un état potentiel d'offre et de demande d'emplois dont la comparaison avec la réalité observée doit permettre de définir une typologie de mécanismes d'ajustements. L'état potentiel et les mécanismes retenus seront fonction d'un horizon à moyen terme et d'un découpage géographique pertinent avec la notion de marché de l'emploi. L'état potentiel sera calculé pour chacune des zones constitutives du découpage.

Ce problème étant posé examinons maintenant la démarche suivie pour tenter de le résoudre.

2.2 - Démarche suivie

La logique de l'analyse de système voudrait que l'on commence par les phases de formulation et de recherche, qui aboutissent à la définition des bases théoriques du modèle, avant d'aborder celles d'exploration et d'interprétation qui correspondent à la mise au point du modèle. La présentation que nous allons en faire est le fruit de plusieurs itérations sur les 4 phases de l'analyse de système.

2.21 - Les bases théoriques du modèle

Il semble que l'on peut grouper les problèmes rencontrés autour de deux thèmes : celui des données et celui des relations.

2.211 - Le problème des données spatialisées

La saisie chiffrée d'un phénomène se fait toujours en référence avec un cadre spatio-temporel. La présentation chiffrée d'un certain nombre de mesures est souvent faite sans référence à la notion d'espace, soit parce qu'elle est sans importance (cas habituel dans les sciences physiques), soit parce qu'elle est implicite et c'est souvent le cas dans les sciences économiques où fréquemment la dimension spatiale est l'ensemble du territoire national. Le problème posé de la définition d'un marché local de l'emploi nécessite une réflexion préalable sur la définition de la dimension spatiale des phénomènes, par ailleurs la démarche adoptée nous oblige à nous contenter du stock de données du passé existant.

α) Définition de la région

La région est une partie du territoire national composée de parcelles contigües. Elle nous intéresse ici en tant que cadre d'une mesure mais aussi en tant que siège d'un phénomène bien particulier, celui du marché de l'emploi. Une réflexion s'impose sur l'aspect épistémologique de la notion de région afin de voir son incidence sur la recherche d'un cadre de saisie de données pertinent au problème posé. Nous nous limiterons ici à l'aspect conceptuel sans chercher à déborder sur l'utilisation faite de ces concepts en analyse spatiale (et qui seront abordés au chapitre 4).

Comme le note J. LABASSE, "peu de concepts sont aussi ambigus que celui de région"⁽¹⁾. Pourtant les économistes régionaux font volontiers appel à une classification de ces diverses acceptions en trois thèmes : la région homogène, la région polarisée et la région-plan. Cette classification proposée par F. PERROUX⁽²⁾ est reprise par J.R. BOUDEVILLE qui en propose⁽³⁾ les définitions suivantes : "l'espace homogène se définit comme un espace dont chacune des parties constituantes présente des caractéristiques ou des propriétés aussi proches que possible les unes des autres", "l'espace polarisé se définit [lui] comme un espace hétérogène dont les diverses parties sont complémentaires et entretiennent entre elles et tout spécialement avec le pôle dominant plus d'échanges qu'avec la région voisine". Enfin l'espace plan se définit comme le cadre spatial d'une action volontariste et pour J.R. BOUDEVILLE, "l'espace homogène et l'espace polarisé n'ont eux-mêmes pour intérêt ultime que d'éclairer une politique et à aider à construire un espace plan (programme) le meilleur possible"⁽⁴⁾.

De ces diverses notions celle de l'espace polarisée est sans doute celle sur laquelle la réflexion s'est portée le plus tôt puisque les premiers travaux de Reilly et Christaller datent de l'avant guerre. Sans aborder le problème de l'analyse économique régionale en tant que tel, on peut indiquer en France que les premières études concrètes basées sur cette notion sont l'oeuvre de A. PIATIER⁽⁶⁾ qui met au point une méthode de mesure de la force d'attraction d'une ville. Cette recherche se concrétise quelques années plus tard par la confection d'atlas⁽⁷⁾ d'attraction urbaine⁽⁸⁾ couvrant l'ensemble du territoire, tandis que M.A. PROST recherche les liaisons hiérarchiques qui existent entre les différents pôles d'attraction.

(1) J. LABASSE - "L'organisation de l'espace" p. 397

(2) F. PERROUX - "Les espaces économiques" - Economie Appliquée 1950, p. 231

(3) J.R. BOUDEVILLE - "Les espaces économiques" PUF - 1970

(4) Ouvrage cité p. 25

(5) Voir chapitre 4 pages 89 et 102.

(6) A. PIATIER - "L'attraction commerciale des villes, une nouvelle méthode de mesure : l'enquête d'essai menée dans le Loir et Cher" dans Revue Economique et Juridique du Sud Ouest n° 4 - 1956 p. 575 et suivantes.

(7) M.A. PROST - "La hiérarchie des villes en fonction de leurs activités de commerce et de services" (Gauthier Villars 1965)

(8) Atlas d'attraction urbaine - collection TEM - Gauthier Villars (1965 - 1970)

L'espace homogène est considéré à quelques exceptions près comme une donnée de fait et la réflexion théorique s'oriente moins vers les problèmes que pose sa définition, que vers ceux de la non coïncidence de ses frontières avec celles de l'espace polarisé. P. Moran pense toutefois que ce problème est un faux problème, la vocation de ces deux concepts étant différente : "la région homogène fournit des coefficients de production et indique des comportement alors que la région polarisée précise la direction des attractions subies et exercées"⁽¹⁾.

Il semble, à la lumière de ces diverses considérations, impossible de dissocier la définition d'un cadre spatial de l'utilisation que l'on compte en faire. C'est sans doute la pluralité des vocations possibles d'une région qui fait écrire à A. PIATIER⁽²⁾, "si les régions existaient vraiment, on le saurait".

Il est donc indispensable, préalablement à tout travail de modélisation, de rechercher le découpage spatial qui semble le plus approprié à l'étude du marché local de l'emploi qui est la pierre angulaire de notre analyse de système. A notre avis c'est davantage en s'appuyant sur le concept de zone homogène que vers celui de zone polarisée qu'il faut se tourner et ce pour deux raisons. Tout d'abord parce que comme l'indique P. MORAN : "la polarisation générale de l'espace n'invite pas par elle-même à un découpage de ce dernier, puisqu'elle exprime des relations"⁽³⁾ or, il est difficilement envisageable d'élaborer un modèle spatialisé ne s'appuyant pas sur une partition de l'espace. Ensuite la connaissance du phénomène de polarisation ne peut s'obtenir que par enquête, chose impossible à réaliser sur le passé.

Les investigations sur les outils de recherche d'une partition de l'espace homogène n'ayant été jusqu'à présent que peu poussées, il nous faudra donc consentir à un effort préalable de réflexion théorique sur ce problème (*chapitre 2*) avant de trancher sur le choix du découpage (*chapitre 6*).

(1) P. MORAN - L'analyse spatiale en Sciences Economiques (Cujas 1966) P. 58

(2) Dans un article intitulé "Existe-t-il des régions en France" dans : Réforme de structure et action régionale 1966, cité par P. MORAN (Revue Juridique et Economique du Sud-Ouest 1969 n° 3).

(3) P. MORAN (article de la RJESO cité) poursuit en écrivant : "mais comme elle dégage une hiérarchie de centres, elle légitime, a priori, une articulation d'ensembles territoriaux, gravitant chacun autour d'un centre fonctionnel de niveau hiérarchique identique", insistant ainsi sur la vocation très particulière de cette notion.

β) Les données disponibles

Deux remarques fondamentales sont à faire sur le problème des données : d'une part elles sont toutes biaisées et d'autre part elles n'existent qu'en quantités limitées.

De tout temps et sous des formes plus ou moins affinées, l'information a été considérée comme un préalable à l'action et presque tous ceux qui ont transgressé durablement cette règle ont payé cher leur légèreté. La finesse et la qualité de l'information nécessaire est fonction de l'action que l'on entend mener et des moyens dont on dispose ; c'est ainsi que le premier recensement connu (Chine 23ème siècle avant J.C.) ne consistait qu'en un simple dénombrement. En contrepartie l'information gratuite n'existe pas dans les sciences humaines et si elle existe en sciences naturelles, ce n'est qu'à l'état potentiel et nécessite le passage par un "analyseur" humain ou matériel (datation au carbone 14 par exemple). Le biais systématique de l'information économique et sociale se traduit par une vision partielle d'un phénomène et son rattachement arbitraire à un cadre spatio-temporel, ces deux aspects étant d'ailleurs liés. Le cadre spatio-temporel est déterminé par les organismes responsables de cette collecte et par la raison de cette collecte : détermination d'assiette fiscale, enregistrement d'actes de gestion (registres paroissiaux, écritures, ...), ... Il s'ensuit que la présentation de l'information n'est pas souvent de nature à favoriser d'autres décisions que celles pour lesquelles l'information est directement collectée.

D'autre part il ne nous est possible dans notre travail que d'utiliser des données disponibles pour toutes les zones constitutives du territoire et pour *plusieurs* dates ou périodes du passé. Ces deux restrictions limitent très singulièrement le stock de données disponibles et apparaissent comme des *contraintes fondamentales* pouvant entraîner des biais importants au niveau du choix des hypothèses testables ainsi d'ailleurs que de la valeur d'indicateur attribuée à telle ou telle donnée. Il s'ensuit que notre démarche peut ne pas déboucher pour des simples raisons de manque de données sans que celle-ci ne soit pour autant fondée.

2.212 - Le problème des relations

Ce problème revêt deux aspects différents, celui de la forme et celui du contenu :

α) la forme des relations

La forme des relations, lorsque celles-ci font l'objet de quantifications, pose le problème de l'appel aux techniques de modélisation. Il nous a semblé nécessaire de faire un panorama succinct (*chapitre 3*) des outils dont nous disposons afin d'éclairer les choix effectués mais aussi parce que leur utilisation en analyse spatiale nécessite l'introduction d'hypothèses complémentaires et même à un certain nombre de remarques spécifiques. Ces considérations seront développées dans une seconde étape (*chapitre 7*).

β) Le contenu des relations

La quantification d'une relation n'est que la traduction dans un langage mathématique d'une liaison théorique postulée entre deux ou plusieurs éléments. Le domaine théorique n'est rien d'autre pour reprendre le langage des documentalistes que le "corpus" des liaisons connues postulées par un ensemble d'"auteurs". L'analyse spatiale, bien que récente, possède un "corpus" déjà impressionnant. Nous avons jugé indispensable d'en extraire les analyses qui affectent de près ou de loin la répartition des hommes dans l'espace. Cette présentation (*chapitre 4*) sera succincte car il ne s'agit que de rechercher si certaines démarches ou hypothèses pourraient être reprises dans notre approche du problème ; par ailleurs le choix des liaisons présentées a été guidé par le problème posé, ce qui explique que certains aspects importants de l'analyse spatiale soient éludés dans ce chapitre.

Le problème du marché spatial de l'emploi n'ayant que peu été abordé par les économistes jusqu'à maintenant, nous avons dû y consacrer un chapitre (*chapitre 5*), et ce d'autant plus que la démarche adoptée repose sur ce concept et que des hypothèses fondamentales pour le modèle doivent être introduites préalablement à toute phase opérationnelle.

2.22 - La mise au point du modèle

Pour la recherche de modèle satisfaisant il faut faire appel simultanément aux démarches inductives et déductives. Pour la dernière, nous nous contenterons des considérations développées dans le chapitre 4 puisque notre démarche exclut la création de modèles théoriques a priori, déconnectés des problèmes de données.

Sa démarche inductive se fonde sur un certain nombre d'analyses chiffrées. Malheureusement le niveau spatial finalement retenu a rebuté jusqu'alors de nombreux chercheurs de par la masse d'informations à manier. Les organismes producteurs de ces informations, lorsqu'ils les ont publiées (ce qui n'est pas le cas de toutes celles utilisées dans notre étude) l'on fait presque toujours sans y joindre une analyse même globale des renseignements fournis, se bornant ainsi à leur seul rôle de producteur. Ces lacunes ont alourdi considérablement notre travail et un certain nombre d'analyses originales ont dû être faites, lorsque la connaissance phénoménologique était trop faible.

Examinons successivement les trois sous-systèmes (3, 4 et 5) présentés dans le graphique de la page 27, pour voir comment ils sont analysés dans notre étude.

α) Le sous-système "demandes d'emploi"

Les demandes d'emploi d'une zone sont fonction de l'importance de la population vivant dans la zone et du taux d'activité désiré. Nous supposons résolu le problème du mouvement naturel et ne rechercherons que des sous-modèles de migrations intérieures (*chapitre 8*), de migrations extérieures et de taux d'activité pour définir une demande d'emploi potentiel dans chaque zone (*chapitre 9*).

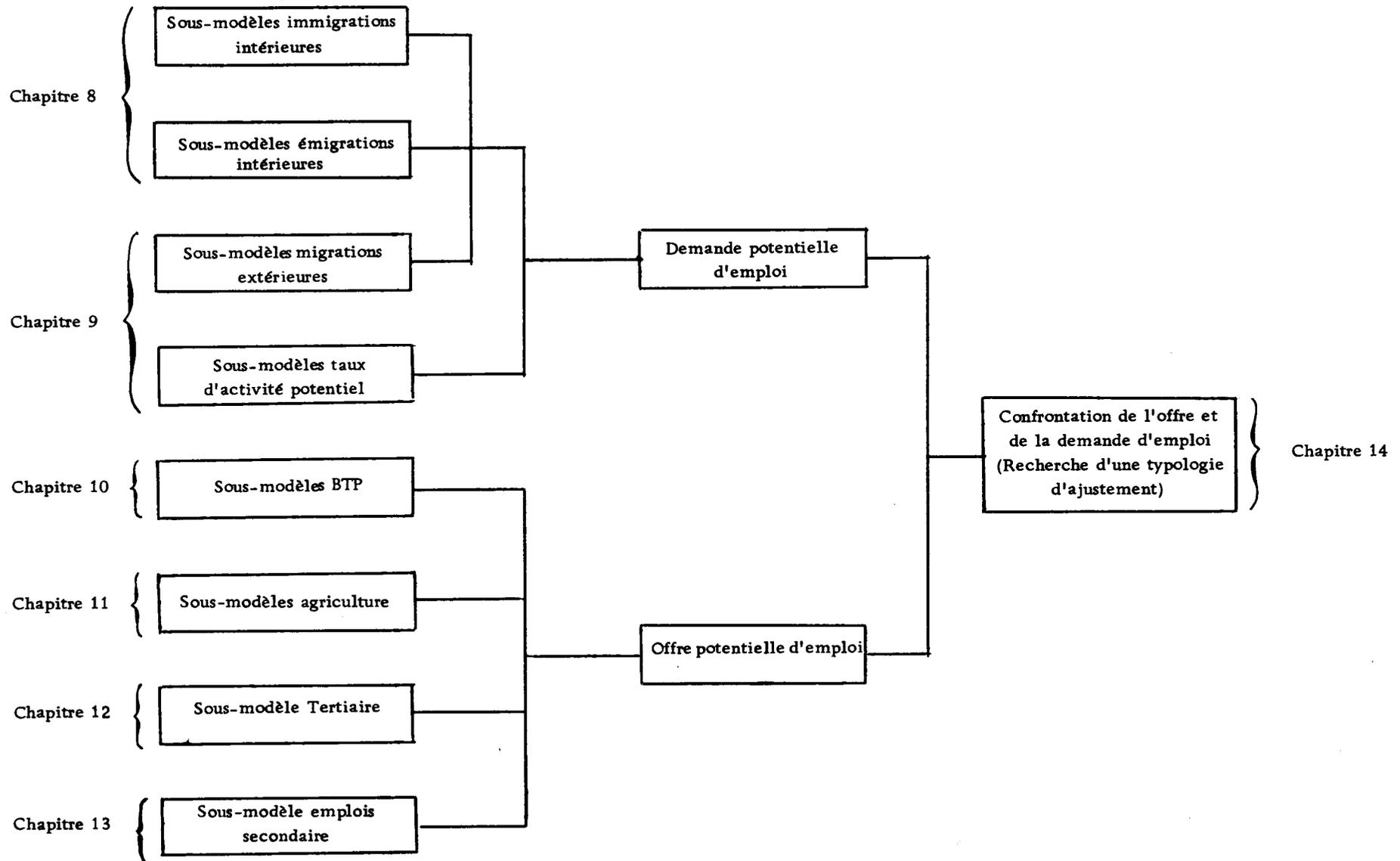
β) Le sous-système "offres d'emploi"

Nous avons séparé arbitrairement les offres d'emplois en 4 postes qui correspondent à autant de sous-modèles. Il s'agit de la catégorie d'activité économique "Bâtiment et Travaux Publics" (*chapitre 10*), des Agriculteurs (*chapitre 11*), du tertiaire (*chapitre 12*) et du secondaire (*chapitre 13*). Pour ce dernier chapitre, signalons tout de suite que des difficultés de traitements d'information nous ont empêché de mener à son terme l'effort entrepris.

γ) Le sous-système "marché local de l'emploi"

Enfin la confrontation de l'offre et de la demande d'emploi sera faite au cours du dernier chapitre (*chapitre 14*) où nous dégagerons les leçons tirées de cette entreprise.

On trouvera page suivante l'articulation des divers sous-modèles.



Techniques de choix d'un découpage spatial

2

CHAPITRE 2 - TECHNIQUES de CHOIX d'un DECOUPAGE SPATIAL

Nous avons vu que des trois conceptions classiques de l'espace, celle de l'espace homogène est celle que l'on devait s'efforcer de privilégier pour éviter, dans la mesure du possible, de biaiser l'approche tentée. Il y a bien sûr une certaine naïveté dans cette prise de position, d'une part parce que comme on l'a souligné, on est tributaire pour ne pas dire prisonnier du stock d'informations existantes et d'autre part parce que le concept d'homogénéité est un concept nécessairement flou du point de vue opérationnel.

Une réflexion théorique sur le problème posé (au chapitre 4 et 5) nous mènera compte tenu des données existantes, à prendre en considération séparément ou simultanément plusieurs indicateurs pour définir un découpage spatial pour l'étude. Ces indicateurs ne seront qu'en partie satisfaisants non seulement à cause des biais précédemment signalés mais aussi à cause des diverses hypothèses restrictives qui ont permis de les considérer comme des indicateurs. *Le problème est donc de passer d'une matrice d'observations à un découpage.*

Ce passage peut s'opérer de plusieurs façons et nous avons cherché dans ce chapitre à analyser a priori différentes démarches que l'on peut imaginer et dont quelques unes seulement ont fait déjà l'objet d'applications. Le terrain restait malgré tout assez neuf car cet effort est surtout lié à une remise en cause d'un cadre d'action, opération qui n'est presque jamais fondée sur une analyse objective.

Privilégier la notion de zone homogène revient à rechercher, à partir de données suffisamment fines, et sans se préoccuper des découpages existants une partition de l'espace appropriée; c'est à une analyse de ces diverses techniques qu'est consacrée la première section. Mais d'un point de vue pratique il est difficile de ne pas tenir compte de la notion d'espace-plan, celle-ci devenant prépondérante lorsque le Politique désire utiliser les résultats de modèles prévisionnels pour définir un ensemble d'actions. Nous avons cherché dans une seconde section à insérer, dans les processus d'obtention progressive de découpage, la possibilité de privilégier tel ou tel découpage existant pour chercher à les juger par rapport à une situation de référence ou à les comparer entre eux.

Enfin il importe de souligner que la présentation de nos réflexions ne dépend pas des techniques finalement retenues car il ne nous a pas été matériellement possible de balayer l'ensemble des possibilités. Il nous a semblé préférable d'exposer l'ensemble de ces réflexions au risque de faire apparaître un déphasage entre le niveau théorique et les applications qui ont été faites.

SECTION 1 - CONSTRUCTION d'une PARTITION de l'ESPACE

1.1 - Les données du problème

1.11 - Les concepts utilisés

Le territoire national peut être considéré comme un ensemble d'éléments que l'on appellera unité spatiale de base ou tout simplement unité spatiale (que l'on notera par commodité u.s.). La France, par exemple, peut être considérée comme l'ensemble des 38 000 Communes ou celui des 95 départements.

On définit la zone comme étant soit une u.s., soit la réunion de plusieurs u.s.

L'ensemble des zones (qui correspond à l'ensemble des parties de l'ensemble des u.s.) peut être muni des opérations de réunion, d'intersection et de complémentarité et constitue donc une algèbre de Boole.

A chaque u.s. peut être attribué soit un caractère soit plusieurs caractères. Un caractère peut être qualitatif (dépendance administrative d'une u.s. à la préfecture de Rouen) ou quantitatif (altitude moyenne d'une u.s.), dans les deux cas les diverses modalités du caractère permettent de partitionner la collection.

Les différents caractères donnés qui permettent d'effectuer une partition de l'ensemble des u.s. peuvent être représentés par un nombre, de nature cardinale, si les caractères sont quantitatifs, et de nature ordinale⁽¹⁾ si les caractères sont qualitatifs. L'ensemble de ces valeurs permet de définir une variable.

Si à chaque u.s. sont attribués plusieurs caractères, il y a plusieurs modes de partition de l'ensemble et donc plusieurs variables.

(1) ce qui signifie que le caractère est repéré par un numéro d'ordre et non qu'une relation d'ordre (au sens mathématique) est établie sur l'ensemble des caractères qualitatifs.

Indépendamment de toute caractérisation des u.s., tout couple d'u.s. possède ou non la propriété d'être un couple d'u.s. contigües, c'est-à-dire ayant une frontière commune. Cette relation de contigüité définie sur l'ensemble des u.s. peut être étendue à l'ensemble des zones. L'intérêt de cette extension est de permettre la création d'une zone par réunion de deux zones contigües données. On est donc en présence d'une technique de définition "récurrente" de la zone.

Il est toujours possible avec une seule variable d'effectuer une partition du territoire sans tenir compte des contraintes de contigüité et c'est même une pratique courante en géographie descriptive (par exemple, représentation du relief français par une carte à l'aide des couleurs différentes par tranches d'altitude).

Mais le plus souvent le nombre de zones est tel qu'il faut tenir compte des contraintes de contigüité et des différentes valeurs possibles d'une ou de plusieurs variables pour élaborer une partition de l'ensemble (de cardinal D) des u.s., formant un ensemble (de cardinal R) de zones toutes disjointes deux à deux. On appellera découpage une telle partition.

1.12 - Découpages sans introduction de variable

- a) Si l'on ne se donne pas de contrainte de contigüité et si l'on ne prend en considération aucune variable, le nombre de groupement des D u.s. en R, zones disjointes⁽¹⁾, risque d'être pratiquement infini⁽²⁾. Par exemple, le nombre de groupements possibles de 95 départements en 22 régions atteint le nombre qui dépasse les possibilités de double précision données par l'ordinateur (ce nombre est donc supérieur à 10^{75}).
- b) L'introduction de la contrainte de contigüité diminue de façon très sensible le nombre de découpages possibles. Il est possible sur des exemples simplifiés de trouver des méthodes récurrentes de dénombrement. Supposons, par exemple, que toutes les zones possèdent un même nombre K d'u.s. et que chaque u.s. soit contigüe à L autre u.s.

(1) Dans le reste de cette section on entendra par zones, des zones disjointes deux à deux, l'ensemble des zones correspondant à une partition des u.s.

(2) Il s'agit du nombre d'applications subjectives de l'ensemble des u.s. sur l'ensemble des zones. Le nombre $H(R, D)$ se détermine par récurrence en sachant que $H(1, J) = 1$ pour J variant de A à D et que $H(I, 1) = 1$.
 $[H(I-1, J-1) + H(I, J-1)]$ avec I variant de 2 à R et J de 1 à D (dans l'exemple, $D = 95$ et $R = 22$).

Le nombre des regroupements possibles des D u.s. en R zones (comprenant chacune K u.s.) et sans respecter la contrainte de contiguité est ⁽¹⁾ $\frac{D!}{K!R!}$ et la probabilité pour que 2 u.s. prises au hasard soient contigües

$(L/(D - 1))$. La relation de contiguité des u.s. d'une zone s'étudie sur un tableau carré de dimension K de variables booléennes valant 1 si les deux zones, représentées par la ligne et la colonne sont contigües et 0 dans le cas contraire. La relation de contiguité est non réflexive (i.e une zone n'est pas contigüe à elle-même) et symétrique, aussi l'étude s'effectue-t-elle sur le tableau triangulaire inférieur de dimension (K - 1) tiré du tableau précédent.

Pour qu'une zone soit compacte (c'est-à-dire que toute u.s. a au moins une frontière commune avec une autre u.s. de la zone), il suffit d'avoir au moins un élément non nul du tableau triangulaire de la zone dans chaque colonne.

La probabilité de trouver tous les éléments nuls dans la première colonne (laquelle comprend K - 1 variables booléennes) est, en utilisant le théorème des probabilités composées :

$$\frac{D - L}{D - 1} \cdot \frac{(D-1)-L}{D-2} \cdot \frac{(D-2)-L}{D-3} \dots \frac{D + 2 - K - L}{D - K + 1} = \prod_{I=1}^{K-1} \frac{D + 1 - I - L}{D - I}$$

Pour la seconde colonne, qui ne comprend que K - 2 variables, cette probabilité est :

$$\prod_{I=1}^{K-2} \frac{D + 1 - I - L}{D - I}$$

(1) Cette formule se trouve facilement à partir de l'exemple suivant : voici pour D = 80, K = 4 et R = 20 l'une des D (=80!) possibilités de découpage



l'ordre des régions importe peu (on peut permuter R₁ et R₂) ; on a donc R ! trop de découpages par rapport aux D ! initiaux. Mais sur ces D ! / R ! possibilités, l'ordre des u.s. à l'intérieur d'une région ne joue pas ; on a donc encore K ! trop de possibilités, d'où la formule.

Pour la J^{ème} colonne qui ne comprend que K - J variables cette probabilité est :

$$\prod_{I=1}^{K-J} \frac{D + 1 - I - L}{D - 1} \quad \text{z}$$

La probabilité pour qu'une zone composée d'u.s. tirées au hasard ne soit pas compacte est en faisant appel à l'axiome des probabilités totales :

$$\sum_{J=1}^{K-1} \left(\prod_{I=1}^{K-J} \frac{D + 1 - L - I}{D - I} \right)$$

En utilisant la probabilité qu'une zone soit acceptable (complémentaire de la probabilité calculée) et le théorème des probabilités composées car les R zones doivent être compactes, la probabilité qu'un découpage soit pertinent est :

$$\left[1 - \sum_{J=1}^{K-1} \left(\prod_{I=1}^{K-J} \frac{D + 1 - L - I}{D - I} \right) \right]^R$$

Le nombre de découpages pertinents de D u.s. en R zones comportant chacune K u.s. est :

$$\frac{D!}{K!R!} \cdot \left[1 - \sum_{J=1}^{K-1} \left(\prod_{I=1}^{K-J} \frac{D + 1 - L - I}{D - I} \right) \right]^R$$

Examinons successivement maintenant les problèmes de partition du territoire suivant une ou plusieurs variables

1.2 - Partition du territoire suivant une variable

On suppose dans ce qui suit, que l'on travaille sur des valeurs ponctuelles (centre de classes si la variable est continue, ou valeurs de variable discrète) et que ces valeurs ont une signification cardinale (c'est à dire qu'un caractère de type quantitatif a été attribué à chaque u.s.).

L'introduction simultanée de la contrainte de contiguïté et d'une variable, en l'absence de contrainte sur le nombre de zones du découpage, permet de trouver un découpage unique du territoire (dans lequel plusieurs zones auront une même valeur de variables⁽¹⁾).

Mais le nombre de zones ainsi déterminé est souvent trop grand pour que des observations intéressantes puissent être tirées de l'analyse d'une masse d'informations encore trop importante. Pour être opérationnel, le nombre de zones doit être inférieur à un seuil donné.

Restreindre le nombre de zones du découpage ne peut se faire qu'au détriment de la variable, c'est à dire qu'une zone aura des u.s. pouvant avoir plusieurs valeurs de la variable, car la relation de contiguïté est intangible et les zones, dans la catégorie de problème étudié, doivent être compactes. Il faut donc introduire une règle de regroupement des u.s. agissant sur la variable. On pourra, par exemple, décider qu'une zone peut comprendre des u.s. dont la majorité a la même valeur de la variable, les autres u.s. ayant soit la valeur immédiatement inférieure ou supérieure, à condition de ne pas être elle-même contigüe à une zone dont la majorité des u.s. est caractérisée par cette valeur supérieure ou inférieure de la variable.

Il apparaît donc que 4 éléments interviennent dans le choix du découpage :

- une valeur de la variable pour chaque u.s.
- une table de contiguïté entre les u.s.
- une contrainte sur le nombre de zones
- une règle de regroupement des u.s.

Cette règle de regroupement devra effectuer un choix entre :

(1) même valeur ponctuelle ou appartenance à la même classe.

- un nombre de zones fixé à l'avance, ce qui ne peut se faire qu'au détriment de l'homogénéité (mesurée par un critère du type variance) des différentes valeurs de la variable pour les u.s. constitutives de chaque zone.
- avec contrainte d'homogénéité des valeurs de la variable pour les u.s. constitutives de chaque zone, aux dépens du nombre de zones (celui-ci étant le plus souvent inférieur au seuil que l'on s'est fixé).

On peut classer ces règles de regroupement en deux : celles qui transforment le tableau de variables en rangeant en classe les valeurs proposées puis travaillent sur ces classes, et celles qui opèrent sans transformation préalable du tableau de variable.

1.21 - Groupement des valeurs de la variable

De nombreuses méthodes peuvent être imaginées qui permettent de créer de nouvelles classes. Toutes ont en commun la simplicité de la solution en ce sens que ne faisant apparaître qu'un découpage,⁽¹⁾ il n'y a pas de comparaison à effectuer.

Il faut souligner qu'aucune de ces méthodes ne peut assurer un découpage en un nombre de zones déterminé. Le plus souvent on retiendra le découpage ayant un nombre de zones inférieur ou égal à un seuil fixé à l'avance.

Examinons successivement quelques méthodes possibles, de façon à faire apparaître les divers problèmes auxquels se heurtent les méthodes par groupement des valeurs de la variable.

1.211 - Le critère de groupement porte sur les valeurs de la variable

Une première méthode consiste à partager l'étendue⁽²⁾ soit en 2 intervalles égaux, puis en 4 intervalles, puis avec le risque de voir rapidement un nombre de classes trop grand, et l'arrivée rapide de classes ne comprenant aucune valeur observée de variable.

Une autre possibilité consiste à faire 2 classes de valeurs définies d'intervalles égaux, puis 3 classes, puis 4... Cette seconde façon de procéder a l'avantage de faire progresser le nombre de classes selon une raison arithmétique et non une raison géométrique, mais en contre partie surviennent des discontinuités dans l'affinement des découpages.

(1) obtenu bien entendu, une fois les nouvelles classes de valeurs de la variable déterminées, en faisant jouer la contrainte de contiguïté et les différentes classes de valeurs à la différence des méthodes vues au 1.12

(2) l'étendue est l'écart entre les valeurs extrêmes observées.

En effet, dans la première méthode deux valeurs proches de la variable peuvent être affectées à deux classes différentes dès le début, mais jamais plus on ne les retrouvera dans la même classe, donc si au début les départements de l'Orne et du Calvados appartiennent à deux zones différentes, on ne les retrouvera pas dans la même zone lorsque le nombre de classes passera de 2 à 4. Il n'en n'est pas de même dans la seconde méthode qui peut permettre aux départements du Calvados et de la Manche, d'appartenir à 2 zones différentes dans le cas d'un regroupement en 2 classes, puis lorsque l'on passe à 3 classes, d'appartenir à la même zone et d'être de nouveau séparés lorsque l'on passe à 4.

Ces méthodes en négligeant l'aspect distribution des valeurs de la variable, aboutissent à attribuer à chaque classe un nombre très variable d'u.s. ; ceci importe peu dans la représentation du phénomène revenu par tête par exemple mais, peut y nuire dans d'autre cas.

1.212 - Le critère de groupement porte sur les

On peut adopter un regroupement en classe par rapport à la médiane, puis les quartiles, ... Cette méthode assure un nombre d'u.s. sensiblement égal dans chaque classe et éclate successivement chaque zone constituée en plusieurs nouvelles zones. Cette progression géométrique a là encore, l'inconvénient d'épuiser rapidement le nombre de valeurs.

On peut aussi travailler directement en quantiles, ce qui assure une croissance arithmétique du nombre de classes, mais engendre des discontinuités dans l'affinement des découpages.

Lorsque la distribution est relativement symétrique, il peut être intéressant de travailler en variable centrée réduite, mais cette méthode n'offre guère de changement par rapport aux précédentes puisque l'on remplace une correspondance "valeur de variable - fréquence cumulée" par "variable centrée réduite - probabilité cumulée". Le seul intérêt est, en effectuant certaines hypothèses de s'appuyer sur la théorie statistique pour éliminer (avec un risque de seconde espèce inconnu) les anomalies observées dans la distribution et qui seraient dues aux fluctuations d'échantillonnage⁽¹⁾.

1.22 - Autres méthodes de regroupement des u.s.

Ces autres méthodes ont en commun le fait de ne pas effectuer de groupement sur les valeurs de la variable et donc de travailler à la fois sur les valeurs et la contiguité. Une zone pourra donc comprendre des u.s. ayant des valeurs différentes. Il devient dès lors nécessaire de disposer d'un critère d'homogénéité de la zone et d'un critère de validité du découpage pour permettre le choix entre deux découpages.

L'application rigoureuse d'un critère d'homogénéité entraîne un nombre de comparaisons tel que, même en disposant de grandes facilités informatiques, le temps de calcul est prohibitif. Ces contraintes techniques nécessitent la recherche d'un cheminement permettant de trouver l'un des meilleurs découpages (au regard du critère choisi), mais sans pouvoir savoir si le découpage retenu par la méthode s'éloigne de façon sensible du découpage optimal.

(1) Cette méthode, on le verra au chapitre 10, pourrait être utilisée pour la représentation du taux d'administration par rapport à la population totale.

Deux types de cheminements sont concevables : on peut avoir une approche descendante, en partitionnant le territoire en deux zones, puis, en divisant l'une des zones ainsi créée en deux autres et ainsi de suite. A cette approche opérant par scissions successives s'oppose celle qui consiste à opérer par consolidations successives d'u.s. ou de zones.

1.221 - Approche par scissions successives

Cette approche consiste à rechercher tout d'abord le meilleur découpage possible du territoire en 2 zones. Pour trouver ce découpage optimum, un critère est nécessaire ; ce critère $K(R)$ est fonction du nombre R de zones du découpage.

On peut proposer par exemple, comme critère, la minimisation de la somme des écarts quadratiques des valeurs des u.s. I appartenant à la zone Z , écart mesuré par rapport à la moyenne $VAL(Z)$ de la zone Z qui contient $N(Z)$ u.s.

$$K(R) = \sum_{Z=1}^R \sum_I [VAL(I) - VAL(Z)]^2 \quad \left\{ \begin{array}{l} \text{avec } I \text{ indice d'u.s. appartenant à la zone } Z \text{ et} \\ VAL(Z) = \frac{1}{N(Z)} \left[\sum VAL(I) \right], \text{ où } N(Z) \text{ est le nombre d'u.s.} \\ \text{appartenant à la zone.} \end{array} \right.$$

Dans cette première itération $R = 2$

Il est dans la plupart des cas prohibitif d'effectuer la comparaison de tous les découpages possibles en 2 zones. Aussi cherchera-t-on à comparer quelques uns des meilleurs découpages (au sens du critère) sans pouvoir être certain que le meilleur découpage soit l'un de ceux testés.

On peut par exemple partir d'un découpage obtenu par un regroupement opéré sur les valeurs de la variable à l'aide de l'une des méthodes décrites au 1.11 de façon à partir d'un découpage en une quinzaine de zones au plus. L'homogénéité de chacune des zones de base est déjà garantie puisque les valeurs prises par la variable pour les u.s. d'une zone donnée sont en faible nombre et comprises dans un intervalle restreint.

Une fois le découpage en 2 obtenu, il ne reste plus qu'à recommencer notre opération sur la zone la moins homogène. Le critère d'homogénéité d'une zone peut être par exemple, la variance intra zone. En divisant en deux la zone dont la variance intra zone est la plus forte, le gain obtenu sur le critère de découpage sera normalement le plus important si les 2 zones ont sensiblement le même nombre d'u.s. et par ailleurs ce choix permet d'obtenir des découpages possédant des zones de plus en plus homogènes.

Cette méthode présente l'avantage d'obtenir un nombre de zones fixé à l'avance.

1.23 - Approche par consolidations successives

Dans un premier temps, tous les couples d'u.s. contiguës ayant la même valeur de variable donnent naissance à autant de zones.

Désormais deux zones contiguës ont toujours des valeurs différentes. Il faut donc définir un critère d'homogénéité pour choisir les nouveaux regroupements. On peut proposer, entre autre, deux méthodes :

- une méthode, proposée par JP. TRYSTRAM (1), consistant à travailler sur les étendues :

Si l'on admet que le meilleur regroupement cherché est celui qui veut différencier le plus possible les zones, les étendues intra-zones doivent être minimum et les étendues des couples de zones contiguës (étendue externe) maximum. On regroupe alors les zones contiguës dont l'étendue externe est la plus faible observée. Plusieurs couples de zones peuvent avoir la même étendue externe. Il y a alors constitution de plusieurs zones au cours de la même itération.

On recalcule ensuite les étendues externes des nouveaux couples de zones et l'on recommence l'opération. Le processus peut être réitéré soit jusqu'au moment de l'arrivée à la zone unique (territoire national) soit arrêté arbitrairement lorsqu'une étendue intra-zone dépasse une limite donnée ou que l'étendue externe descende en dessous d'une limite donnée, qui peut être par exemple l'intervalle inter-quartile.

Les essais de cette méthode ont montré que dans la plupart des cas l'on passait rapidement d'une partition en un trop grand nombre de zones, à une partition en un nombre de zone insuffisant. Le nombre de zones groupées à chaque itération étant variable, il n'est pas possible de garantir l'obtention d'un découpage ayant un nombre de zones fixé à l'avance.(2)

- Une seconde méthode possible consiste à travailler non sur les étendues mais sur un critère de dispersion des valeurs d'une zone. Le critère peut être par exemple la variance définie sur les parcelles de la zone(3) . On regroupe le couple de zones (réduites éventuellement aux u.s.) dont la variance intra-zones de la zone consolidée est minimum. Une fois le choix effectué les deux zones sont agrégées et l'on doit recalculer la variance des seuls couples constitués de la zone lors de l'itération précédente et des zones qui lui sont contiguës. Le découpage passe de D zones à 1 zone par pas de un. Le test d'arrêt peut être le nombre de zones ou un seuil d'un critère d'homogénéité du découpage conçu par exemple comme la somme des variances intra-zones pondérées par le nombre de parcelles, le seuil correspondant à une fraction de l'indicateur quadratique calculé pour le territoire "découpé" en une seule zone (= réunion de toutes les u.s.).

(1) Rapport de Synthèse du contrat 6601 Marc Bloch-DATAR sur le Système d'Accumulation des données (SAD) p. 34 et sq - septembre 1967

(2) Ce nombre de zones limite peut d'ailleurs servir de test d'arrêt pour le processus itératif.

(3) ou le coefficient de variation, ou ...

1.3 - Partition de l'espace suivant plusieurs variables

Le plus souvent pour appréhender un phénomène, il est nécessaire de prendre en considération plusieurs variables. Pour appréhender par exemple le phénomène circulation routière, il est nécessaire de prendre en considération par exemple le parc automobile, la consommation d'essence, l'état du réseau routier. Il est très vraisemblable que les découpages obtenus pour chacune de ces variables ne coïncideront pas. Il est dès lors nécessaire de construire une variable unique qui permettra de se resituer dans le cadre défini au 1.2

Cet indicateur n'aura pas forcément de signification cardinale, ce qui restreindra le champs des méthodes de découpage envisageables.

Ces indicateurs n'ont longtemps été basés que sur des procédures d'agrégation des variables mais depuis quelques décennies une nouvelle technique basée sur une analyse globale est apparue : l'analyse factorielle.

Examinons successivement ces deux approches.

1.31 - Les indicateurs créés par techniques d'agrégation

Lorsque les différentes variables décrivent une même donnée, mais à des dates différentes, et que c'est à l'évolution du phénomène que l'on s'intéresse, pour cela il suffit d'utiliser par exemple, le taux d'évolution sur la période ou le taux d'évolution moyen annuel. On est alors ramené au problème à une seule variable.

Lorsqu'une unité commune peut être trouvée (par exemple une valeur unitaire en franc) l'agrégation des différentes variables s'effectue sans difficulté par une somme pondérée des variables (prix par quantité par exemple). Il va de soi que l'on s'appuie sur la technique des indices lorsqu'il s'agit de variables décrivant diverses informations à des dates différentes et que c'est à l'évolution du phénomène que l'on s'intéresse.

Mais dans bien des cas, une unité commune ne peut être trouvée et l'agrégation des variables ne peut se faire de façon simple. Plusieurs techniques ont été élaborées pour résoudre cette difficulté.

1.311 - Utilisation de variables sans dimension

Le problème étant de prendre en compte des variables à unités différentes, il est possible par des transformations appropriées de les rendre comparables. Pour ce faire on peut par exemple travailler en variables centrées réduites ou encore en écart au minimum rapporté à l'étendue.

Une fois ces variables rendues comparables il faut les traiter de façon à obtenir un indicateur unique. On peut par exemple en faire la moyenne (ce qui suppose que toutes les variables ont la même importance) ou une moyenne pondérée (les coefficients de pondération sont obtenus par expert ou empiriquement par une méthode statistique), mais il s'agit dans un cas comme dans l'autre de technique empirique d'obtention de coefficient, ou l'arbitraire joue un grand rôle.

1.312 - Appel à l'analyse multicritère

Il ne saurait être question ici de "brosser" l'ensemble des techniques existantes d'autant qu'il existe quelques ouvrages et articles ⁽¹⁾ sur la question.

Ce qu'il importe de souligner c'est que ces différentes techniques par l'intermédiaire de métriques appropriées ("indices" de distance) permettent de transformer un ensemble d'informations ordinales (ou préordinales) pour chaque observation en une structure unique ordinale (ou préordinale). Les métriques définies sont arbitraires et leur emploi délicat sur le plan de l'interprétation non mathématique, mais une fois l'indicateur unique trouvé (et justifié), on est ramené au problème précédent et il est facile de définir une typologie.

1.32 - Appel à l'analyse factorielle

Cette méthode générale d'analyse des données ⁽²⁾ part de la représentation d'une matrice de Z observations pour N variables, dans un espace à Z dimensions ou dans un espace à N dimensions. La technique utilisée consiste à rechercher une droite de regression orthogonale sur ce nuage, que l'on appelle première composante principale, puis une autre droite de regression orthogonale correspondant au second minimum de l'indicateur d'écart quadratique (= seconde composante principale).

Puis on effectue une projection du nuage sur la plan défini par ces deux premières composantes principales. Ce travail est effectué pour les représentations du nuage dans les deux espaces ⁽³⁾ et les projections rapportées sur un même plan.

(1) Voir Kendall "The advanced theoriz of statistics" - Tome 3. Voir aussi article de synthèse de MM. BERNARD et BESSON "Douze méthodes d'Analyse Multicritère" -

Revue française d'Information et de Recherche opérationnelle n° V. 3 - 1971.

(2) Dans "Statistique et Informatique appliquées" (Dunod 1971) MM. LEBART et FENELON font une présentation générale de l'analyse des données de laquelle découle, par transformation du tableau de données l'analyse en composante principale (variable centrée) l'analyse en composante principale normée (variable centrée réduite) analyse des correspondances (tableau de dépendance).

(3) Il existe des relations simples entre les composantes principales des deux espaces.

L'indicateur quadratique utilisé dans la construction des composantes principales est sensible à l'amplitude comparée des variables, c'est pourquoi, pour éliminer ces influences, on peut travailler en variables centrées et si l'on désire éliminer l'hétérogénéité des moyennes, on peut travailler en variables centrées réduites.

Travailler au niveau de la variable ou de la variable centrée implique une grande sensibilité des résultats en fonction des unités : si les variables ne sont pas toutes de même nature (exemple parc automobile et consommation d'essence) l'unité adoptée (consommation d'essence en litres, ou en milliers de litres par exemple) influe de façon très importante en donnant un poids plus ou moins important à certaines variables dans le critère de distance retenu (le carré de l'écart dans la consommation d'essence des départements par rapport à la moyenne départementale est bien sur plus important si l'unité est le millier de litres au lieu du litre et la droite sera "attirée" par cette variable).

Travailler en variable centrée réduite, par contre, suppose que toutes les variables ont la même importance dans l'analyse que l'on effectue, si tel n'est pas le cas, un biais est introduit.

Cette technique est donc à manier avec infiniment de précautions, d'autant que l'interprétation en est délicate.

En effet, prenons l'exemple donné par MM. LEBART et FENELON⁽¹⁾ sur l'analyse des données départementales de causes de décès en 1962 pour montrer les possibilités de l'analyse de correspondance. L'interprétation du géographique ne peut que conduire à une typologie permettant de dire qu'un département est du type 1, 2 ou K.

Le tableau initial des données, après un traitement et une interprétation graphique aboutit donc à la définition d'une variable de nature ordinale, à moins de considérer comme un indicateur de nature quantitative la première composante principale. Dans ce dernier cas la cartographie est possible mais pour le moins délicate à interpréter.

1.33 - Utilisation de l'indicateur pour la construction du découpage

Une fois l'indicateur constitué il ne reste plus qu'à utiliser l'une des techniques décrites au 1.2

Il faut noter toutefois, que lorsque la variable créée a une signification ordinale (cas de l'analyse factorielle), il n'est pas possible de "travailler" les valeurs de la variable. Si le découpage obtenu comporte un trop grand nombre de zones, il faut repartir de l'étude de la constitution de la variable et voir si l'on peut effectuer d'autres groupements des informations.

(1) opus cité p. 240 et 241 - D'autres exemples peuvent trouver dans l'article de Brian J.L. BERRY "Grouping and Regionalizing : an approach to the problem using multivariate analysis" in quantitative geography part 1 (1967) où l'on trouvera en outre une liste des principaux travaux effectués en analyse factorielle dans cette optique spatiale, depuis trente ans.

SECTION 2 - COMPARAISON de DECOUPAGES DONNES

La construction d'une partition de l'espace permet de cerner au mieux la réalité géographique d'un phénomène. Son aide à la décision ne peut guère jouer qu'à un niveau très général : dépistage de zones à sous équipement hospitalier ou création d'un découpage de zones à primes d'équipement industriel différencié par exemple. Dans la pratique lorsqu'une prise de décision doit s'effectuer à un niveau spatial assez fin (répartition ponctuelle de crédits d'équipements scolaires par exemple), il devient nécessaire de tenir compte des découpages administratifs existants qui correspondent à des échelons géographiques effectifs de prise de décision, car il est très peu probable que l'importance du problème étudié justifie la création d'un découpage administratif approprié.

Le problème reste de savoir si un phénomène sur lequel la politique désire agir, a un sens au niveau d'un découpage donné ou, ce qui est un problème plus général, pour quel découpage géographique existant, ce phénomène est le plus significatif. Pour le phénomène "desserte scolaire", par exemple, le phénomène est-il mieux appréhendé au niveau de la Région de Programme, du département, du canton ou de la commune ; la réponse à cette question doit indiquer le niveau de décision le plus souhaitable.

Mais dans ce jugement de découpages, deux cas de figures peuvent se présenter. Le premier, le plus fréquent, est le cas de découpages emboîtés, c'est à dire que chaque zone d'un découpage est obtenu par agrégation de zones d'un autre découpage. Par exemple, le découpage en régions de programmes s'obtient par consolidations sur le découpage en départements. Le second cas est la comparaison de deux découpages quelconques.

2.1 - Cas des découpages emboîtés

Pour la description d'un phénomène donné, deux problèmes différents se posent.

- le premier est celui de la comparaison entre les deux découpages : lequel des deux est le meilleur pour rendre compte du phénomène ?

- le second est celui de savoir si le découpage retenu rend effectivement compte du phénomène que l'on veut décrire. En fait cette dernière question a surtout un sens lorsque l'on veut décrire un mécanisme. En effet, la notion de densité de population a un sens quelque soit la zone à laquelle on s'attache mais ce renseignement sera plus ou moins significatif ; tandis qu'une étude d'autofinancement des entreprises n'a pas grand sens à un niveau cantonal dans la mesure où une entreprise peut être composée d'établissements hiérarchisés fonctionnellement dans le cadre d'une concentration verticale et localisés dans plusieurs cantons.

La réponse à ce second problème, dans la mesure où il se pose, fera intervenir en plus une étude théorique du mécanisme économique envisagé pour établir des normes souhaitables pour certains indicateurs et permettre ainsi de confirmer ou d'infirmer le découpage retenu comme étant acceptable.

Pour tenter de résoudre le premier problème, ce qui est la préoccupation de cette section, on peut s'inspirer des différentes démarches envisagées pour la construction d'une partition de l'espace et fonder le choix à effectuer sur un jugement d'indicateur ou s'appuyer sur des analyses statistiques (comme celle de la variance).

2.11 - Etude de la variation d'un indicateur

Prenons, à titre d'exemple, et pour faciliter l'exposé, le cas des découpages départementaux et régionaux.

2.111 - Les données du problème

La comparaison entre ces découpages ne peut s'effectuer que si l'on dispose de données élémentaires à un niveau plus fin que le département (l'u.s. peut être par exemple le canton). En effet, la comparaison ne peut se faire que si l'on dispose d'un indicateur d'homogénéité pour chaque zone, lequel ne peut être bâti que si l'on dispose de données infra-départementales.

Il est nécessaire de porter un jugement global pour un découpage zonal donné, ce qui se fera par un indicateur de découpage tenant compte des indicateurs d'homogénéité de chaque zone.

Dans le cas de découpages emboîtés, il est plus que vraisemblable que le découpage le plus fin aura l'indicateur de découpage le meilleur, mais il n'empêche que, pour reprendre notre exemple, le découpage régional puisse être plus proche du découpage optimal de la France en 22 zones que le découpage départemental de celui de la France en 95 zones.

Pour tenir compte du nombre de zones dans la comparaison des découpages, il est donc nécessaire de passer par une référence externe qui peut être la valeur de l'indicateur du découpage optimal (lequel est une fonction du nombre de zones).

Pour illustrer ces considérations nous proposerons la méthode suivante :

2.112 - Méthode possible :

a) Constitution des découpages de référence :

En partant des données cantonales (C cantons existants) on utilisera la méthode de construction d'une partition de l'espace en partant de l'approche par consolidation successive, décrite au 1.122 :

- le nombre de zones de départ (après agrégation des cantons contigus ayant la même valeur) est par exemple de 200.
- on regroupe le couple de zones dont la variance intra zone de la zone consolidée est la plus faible. Pour cela on calcule pour chaque couple de zones contiguës, la valeur moyenne de la zone consolidée ; si cette valeur est un ratio, par exemple le revenu par tête, on repart des informations de base, ici la masse des revenus de chaque canton et la population de chaque canton, pour obtenir un ratio significatif. Dans certains cas, comme l'altitude moyenne de la zone, il y a intérêt à faire une moyenne pondérée (par la superficie de la zone). On a donc, en supposant pour simplifier que chaque zone est contiguë à 4 autres, et en tenant compte du fait de la symétrie de la relation de contiguïté, 400 indicateurs d'homogénéité intra-zone.
- lorsque, après consultation de ces 400 indicateurs, on effectue une consolidation, il reste 199 zones et seulement de 4 à 6 indicateurs d'homogénéité à recalculer sur les 398 à 400 indicateurs existants, suivant que le nombre de zones auxquelles la nouvelle zone créée est contiguë est de 4 à 6. L'itération ne demande donc que peu de calculs.

b) Choix de l'indicateur de découpage

Parallèlement, un indicateur de découpage (en R zones) est créé et qui est constitué d'une somme de carrés des écarts de la valeur de chaque canton, à la valeur de la zone à laquelle le canton appartient.

$$I (R) = \sum_z \sum_{C \in z} \left[\text{VAL} (C) - \text{VAL} (Z) \right]^2$$

Au début, il y a autant de zones que de cantons ; à la fin de l'opération, il n'y a plus qu'une seule zone, le territoire national. Cet indicateur varie donc entre 0 et la somme des écarts quadratiques à VAL (F). On montre facilement que I (R) est une fonction monotone (décroissante), en effet lorsqu'il y a regroupement.

$$\sum_{C \in Z_1} [\text{VAL}(c) - \text{VAL}(Z_1)]^2 + \sum_{C \in Z_2} [\text{VAL}(c) - \text{VAL}(Z_2)]^2 > \sum_{C \in Z_3} [\text{VAL}(c) - \text{VAL}(Z_3)]^2$$

avec $Z_3 = Z_1 \cup Z_2$

car les variances de Z_1 et Z_2 d'après notre critère sont forcément inférieures à celles de Z_3 .

Il est plus simple d'exprimer $I(R)$ en % de l'écart quadratique à la valeur $\text{VAL}(F)$.

On ne peut prétendre que cette courbe est la courbe optimale, en effet, il aurait fallu pour chaque valeur différente de R rechercher pour tous les découpages en R zones possibles, la valeur de l'indicateur $I(R)$ et retenir le découpage correspondant au minimum de $I(R)$, ce qui n'est pas envisageable pour des raisons de coût.

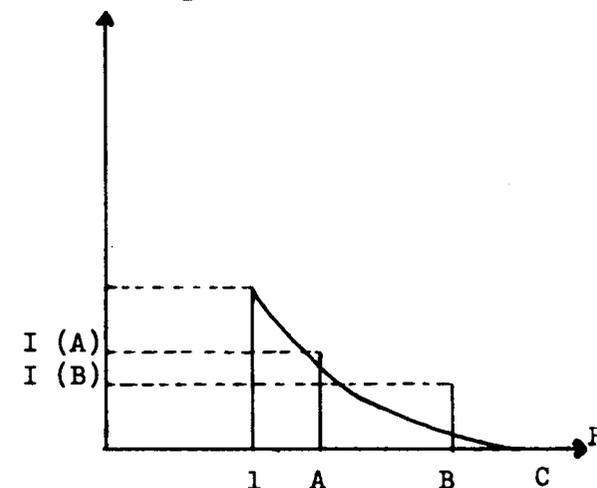
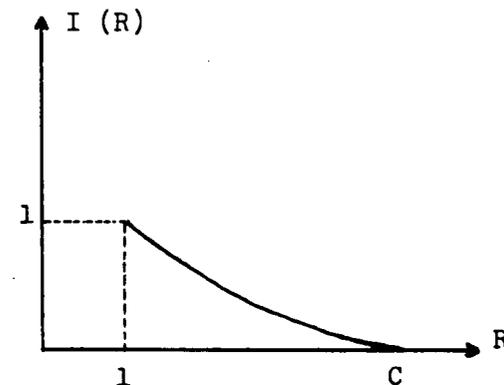
Mais il est vraisemblable que la méthode proposée fournit une approximation acceptable de l'évolution théorique optimale de l'indicateur $I(R)$.

e) Comparaison des découpages :

Il ne reste qu'à calculer la valeur de l'indicateur $I(D)$ pour les deux découpages à comparer $I(A)$ et $I(B)$.

Cette comparaison passe alors par les valeurs optimales (ou supposées proches) de l'indicateur pour les nombres de zones correspondants.

Le découpage A (ou B) se rapproche le plus du découpage optimal, si les écarts absolus et relatifs de A (ou B) par rapport à la norme D sont plus faibles que ceux de B (ou A). Si les écarts absolus sont d'amplitudes équivalentes, le découpage proposé ayant l'écart relatif le plus faible sera retenu. Enfin, si l'on observe un avantage absolu pour l'un des découpages et relatif pour l'autre, il est pratiquement impossible de conclure.



d) Remarques :

Il faut souligner que cette méthode permet de penser qu'un découpage existant est meilleur qu'un autre découpage existant pour décrire un phénomène donné, mais qu'elle ne permet pas d'établir si le phénomène donné est un mécanisme économique, que ce mécanisme a un sens pour le découpage retenu, ou autrement dit que l'on appréhende suffisamment bien ce phénomène à ce niveau.

2.12 - Appel aux techniques purement statistiques

Lorsque l'on s'occupe de découpages emboîtés, on peut considérer le premier niveau géographique, la région par exemple, comme une première variable et le second niveau géographique, le département par exemple, comme une seconde variable et construire un tableau du type tableau de dépendance. Il est dès lors possible d'adapter les techniques classiques de l'analyse de la variance⁽¹⁾ pour résoudre ce type de problème. Deux cas de figures se présentent selon que l'on veut traiter une seule ou plusieurs variables.

2.121 - Cas d'une seule variable

Soit VAL (I,J,K), la valeur prise par le Kⁱème canton du Jⁱème département, qui possède C (J) cantons, appartenant à la Iⁱème région qui possède D (I) départements, par la variable qui fait l'objet de notre étude et pour laquelle on cherche à savoir lequel du niveau départemental ou régional est le meilleur.

Soit VAL (...), la valeur moyenne calculée pour l'ensemble du territoire, qui possède R régions, D départements et C cantons, VAL (I,..) la valeur moyenne calculée pour la région I et VAL (I,J,..) celle du département J de la région I.

On peut écrire :

$$VAL (I,J,K) = VAL (I,J,..) + U (I,J,K)$$

où U (I,J,K) représente l'écart du canton par rapport à la moyenne VAL (I,J,..) du département auquel il appartient.

(1) D. MALKIN : "Découpage spatial et marché de l'emploi" 1970 DATAR ; note interne du Groupe Modèle du SESAME. D'autres techniques statistiques peuvent être utilisées, mais celle de l'analyse de la variance paraît être l'une des plus adaptées au problème.

On peut poser :

$$\text{VAL} (I, J, \cdot) = \text{VAL} (\cdot, \cdot, \cdot, \cdot) + V (I) + W (I, J) \quad \text{avec} \begin{cases} V (I) = \text{VAL} (I, \cdot, \cdot) - \text{VAL} (\cdot, \cdot, \cdot) \\ W (I, J) = \text{VAL} (I, J, \cdot) - \text{VAL} (I, \cdot, \cdot) \end{cases} \quad (2)$$

où $V (I)$ représente l'écart de la variable par région par rapport à l'ensemble du territoire, et $W (I, J)$ celui du département par rapport à la région. L'analyse de variance nous donnera le degré de signification de chaque niveau de découpage.

a) Ce système (II) comporte $D (1) + D (2) + \dots + D (22) = 95$ équations et $D (1) + D (2) + \dots + D (22) + R + 1 = 95 + 22 + 1$ variables.

Pour exprimer $\text{VAL} (\cdot, \cdot, \cdot)$, $V (I)$, $W (I, J)$ en fonction de $\text{VAL} (I, J, \cdot)$ il faut rajouter un certain nombre d'équations annexes.

$$\sum_{I=1}^R A (I) \cdot V (I) = 0 \quad \text{et} \quad \sum_{J=1}^{D(I)} B (I, J) \cdot W (I, J) = 0 \quad (3)$$

Ce qui signifie que la somme pondérée des déviations régionales et départementales est nulle. On peut, sans restreindre la généralité des hypothèses, considérer que ces coefficients de pondération satisfont la condition suivante :

$$\sum_I A (I) = 1 \quad \text{et} \quad \sum_{J=1}^{D(I)} B (I, J) = 1 \quad \forall I \quad (4)$$

Ces équations complémentaires (III) et (IV) permettent de résoudre le système (II). En multipliant par $B (I, J)$ et en sommant sur J , puis en multipliant par $A (I)$ et en sommant sur I on obtient un système d'où l'on sort :

$$\left[\begin{array}{l} \text{VAL} (\cdot, \cdot, \cdot) = \sum_{I=1}^R \frac{D(I)}{\sum_{J=1}^{D(I)} B (I, J)} A (I) \cdot B (I, J) \cdot \text{VAL} (I, J, \cdot) \\ V (I) = \frac{D(I)}{\sum_{J=1}^{D(I)} B (I, J)} B (I, J) \cdot \text{VAL} (I, J, \cdot) - \sum_{I=1}^R \frac{D(I)}{\sum_{J=1}^{D(I)} B (I, J)} A (I) \cdot B (I, J) \cdot \text{VAL} (I, J, \cdot) \\ W (I, J) = \text{VAL} (I, J, \cdot) - \frac{D(I)}{\sum_{J=1}^{D(I)} B (I, J)} B (I, J) \cdot \text{VAL} (I, J, \cdot) \end{array} \right. \quad (5)$$

On remarque que ces moyennes n'ont qu'un sens statistique et n'ont aucune raison de correspondre exactement à la moyenne observée.

Considérant les relations (I), (II) et (III) et en supposant que les $U(I, J, K)$ sont distribués de façon indépendante et suivent la même loi normale centrée, on peut trouver l'estimateur $\widehat{VAL}(I, J, .)$ de $VAL(I, J, .)$ qui minimise l'écart quadratique T .

$$T = \sum_{I=1}^R \sum_{J=1}^D(I) \sum_{K=1}^C(J) \left[VAL(I, J, K) - \widehat{VAL}(I, J, .) \right]^2$$

Cet estimateur est la moyenne des observations cantonales.

$$\widehat{VAL}(I, J, .) = \frac{1}{C(J)} \sum_K VAL(I, J, K) \quad (6)$$

En combinant (V) et (VI) on trouve les estimateurs de $VAL(., ., .)$, $V(I)$ et $W(I, J)$ en remplaçant $VAL(I, J, .)$ par $\widehat{VAL}(I, J, .)$.

b) Considérons maintenant les hypothèses :

$$H_D : V(I) = 0 \quad \forall I$$

$$H_R : W(I, J) = 0 \quad \forall (I, J)$$

H_D exprime le fait que dans (I) l'introduction du niveau départemental n'apporte rien à l'explication de la variabilité des observations à l'intérieur d'une même région. Si H_D et H_R sont vrais, l'introduction du niveau régional n'apporte lui non plus aucune explication. La confirmation ou l'infirmité de l'une ou des deux hypothèses est une réponse au problème du choix de découpage.

L'hypothèse H_D est indépendante du système de pondération choisi, en effet, elle nécessite pour être vérifiée que $VAL(I, J, .) = VAL(I, 1, .) \quad \forall J$, ce qui entraîne que $W(I, 1) = W(I, J)$, $\forall J$. La troisième relation de (V) reste alors vraie quelque soit le système de pondération sous réserve que $\sum_J B(I, J) = 1$.

L'hypothèse H_R dépend, elle, du système de pondération ; en effet, d'après la deuxième équation de (V), le fait que la déviation régionale par rapport à la valeur nationale soit nulle, implique :

$$V(I) = 0 \Leftrightarrow \sum_J B(I,J) \cdot VAL(I,J) = 0 \quad \forall(I)$$

Les conclusions que l'on tirera du niveau départemental seront indépendantes du système de pondération, alors que celles que l'on tirera pour le niveau régional en dépendront.

c) Soit les ensembles d'hypothèses suivants :

E = ensemble des hypothèses évoquées jusqu'ici,

$$E_D = E \cap H_D$$

$$E_R = E \cap H_R$$

Ces sous-ensembles définissent des sous-espaces de paramètres dans lesquels pour obtenir des estimateurs de paramètres, on minimise S . Ces S suivent des lois χ^2 et les tests à utiliser sont ceux de Fischer - Snedecor.

$$S = \sum_{I,J,K} [VAL(I,J,K) - VAL(I,J,.)]^2$$

Pour E , le minimum de S est, nous l'avons vu, T .

On peut écrire S sous la forme :

$$S = \sum_{I,J,K} \left[[VAL(I,J,K) - \widehat{VAL}(I,J,.)] + [\widehat{VAL}(I,J,.) - VAL(I,J,.)] \right]^2$$

et en tenant compte de la relation (II) :

$$S = \sum_{I,J,K} \left[[VAL(I,J,K) - \widehat{VAL}(I,J,.)] + [\widehat{VAL}(\dots) - VAL(\dots)] + [\widehat{V}(I) - V(I)] + [\widehat{W}(I,J) - W(I,J)] \right]^2$$

En effectuant les produits et en tenant compte de la nullité des sommes d'écart à la moyenne et en choisissant un système de pondération proportionnel au nombre de zones

$$\left[A(I) = D(I)/R \right] \quad \text{et} \quad \left[B(I,J) = C(J)/D(I) \right]$$

l'équation se transforme en :

$$= T + C \cdot \left[\widehat{VAL}(\dots) - VAL(\dots) \right]^2 + \sum_I D(I) \cdot \left[\widehat{V}(I) - V(I) \right]^2 + \sum_{I=1}^R \frac{D(I)}{\sum_{J=1}^R C(J)} \cdot \left[\widehat{W}(I,J) - W(I,J) \right]^2$$

Le minimum de S sous les diverses hypothèses est :

$$H \begin{cases} H_{\Omega} : S_E = T \\ H_R : S_{ER} = T + \sum_I D(I) \cdot \widehat{V}(I)^2 \\ H_D : S_{ED} = T + \sum_I \sum_J C(J) \cdot \widehat{W}(I,J) \end{cases}$$

Les F tests statistiques sont données par les rapports suivants :

$$H_R : F_R = \left[\frac{S_{ER} - S_E}{R - 1} \right] \left[\frac{\sum_I \sum_J [C(J) - 1]}{S_E} \right]$$

$$H_D : F_D = \left[\frac{S_{ED} - S_E}{\sum_I (D(I) - 1)} \right] \left[\frac{\sum_I \sum_J [C(J) - 1]}{S_E} \right]$$

et il ne reste plus qu'à faire la lecture de la table pour un coefficient de confiance α et les degrés de liberté utilisés dans le calcul du F.

2.122 - Cas de plusieurs variables

Dans ce cadre on dispose d'un vecteur de variables (repérées par l'indice L) pour chaque canton soit $VAL(I, J, K, L)$ ce vecteur.

On a maintenant :

$$VAL(I, J, K, L) = VAL(., ., ., L) + V(I, L) + W(I, J, L) + U(I, J, K, L)$$

- a) L'analyse de variance multivariable n'est possible que si l'on peut exprimer la matrice $M = VAL(I, J, K, L)$ comme le produit de deux matrices $M1.M2$ représentant la dépendance linéaire des observations par rapport aux paramètres $VAL(., ., ., L)$, $V(I, L)$, $W(I, J, L)$, plus un vecteur d'erreur μ .

$$M = M1 . M2 + \mu \quad (1)$$

Les hypothèses relatives aux niveaux d'homogénéité se traduisent par une matrice d'hypothèses $M3$ construites de façon que :

$$M3 . M2 . M4 = 0 \quad \forall M4 \quad (2)$$

En fait la contrainte (1) n'est jamais satisfaite du fait que le nombre de départements par région varie d'une région à l'autre.

- b) Que faire lorsqu'il paraît intéressant d'étudier la variabilité d'un vecteur de variable ? Deux solutions sont possibles :
- on peut effectuer une analyse du type de celle décrite au 2.121 pour chacune des variables. Si les conclusions (rejet ou acceptation des hypothèses) sont identiques pour toutes les variables, on conclueradans le même sens pour le vecteur.
 - s'il n'en est pas ainsi il faut définir un critère permettant d'effectuer une combinaison linéaire des variables, ce qui nous ramène au problème précédent. La justification de ce critère est d'ordre statistique : on prendra la combinaison qui fait perdre le moins d'informations sur la variabilité du vecteur, ou plus précisément qui rend compte du pourcentage maximal de la variance totale de N variables. La technique la mieux appropriée est celle de l'analyse en composantes principales. Le test de Fischer sera alors fait sur la première composante principale.
 - Si les unités des différentes variables diffèrent (ou si leurs significations diffèrent), une combinaison linéaire n'aura que peu d'intérêt et il sera sans doute plus intéressant de travailler sur la matrice de corrélation et sur les composantes centrées réduites.

2.2 - Cas des découpages non emboîtés

Le cas de découpages non emboîtés d'un point de vue pratique nous intéresse moins, car la structure française de l'administration générale est de type pyramidale et que cette hiérarchie a son exact dans la compétence géographique des décideurs.

Quelques secteurs, et non des moindres, il est vrai, ont des regroupements régionaux ne correspondant pas : la circonscription académique, la région PTT, ou la région militaire en sont les principaux exemples. Mais ces découpages à vocations sectorielles, ont tendance à s'harmoniser et la spécialité de leur action fait qu'ils peuvent de moins en moins prétendre rivaliser avec cette circonscription mise en place par le plan, la circonscription d'action régionale, dont la vocation de généralité s'étendra encore dans le cadre des réformes en cours.

On peut toutefois signaler que si l'appel aux techniques statistiques du type analyse de variance, n'est pas possible, l'étude de la variation d'un indicateur du genre de celle proposée reste faisable.

2.3 - Précarité des résultats

Le caractère empirique des méthodes proposées dans le cadre de cette section, même si de solides justifications peuvent être apportées, confère une valeur toute relative aux renseignements tirés de leurs applications.

L'absence de références théoriques pour l'étude des découpages, ainsi que les problèmes combinatoires qu'elle pose, nécessitera sans doute encore longtemps une approche très pragmatique de ces questions, qui ne peuvent de toute façon être ignorées dans le cadre d'une rationalisation des décisions à incidences spatiales.

Techniques de modélisation

3

CHAPITRE 3 - TECHNIQUES de MODELISATION

Après avoir montré dans le chapitre 1 la nécessité de faire appel aux techniques de modélisation, il est nécessaire de décrire, au moins sommairement la logique des modèles et les différents concepts qui s'y rattachent de façon à clarifier les choix que l'on est amené à effectuer dans la mise au point du modèle et à en souligner les implications. Une première section sera consacrée à ces problèmes.

De nombreuses techniques économétriques sont à la disposition du "modélisateur" mais nombre d'entre elles sont spécifiques à certains problèmes. Aussi n'est-il pas question d'effectuer un balayage exhaustif de ces techniques économétriques. On se bornera en fait à examiner brièvement l'une d'entre elles, la régression multiple, en soulignant les hypothèses sous jacentes qui lui sont propres d'une façon générale et l'on examinera ultérieurement⁽¹⁾ les hypothèses complémentaires qu'il est nécessaire d'introduire dans le cadre du modèle pour que cette technique conserve un sens. Dans cette seconde section nous développerons aussi nos réflexions sur un problème rencontré : celui de l'intérêt de la stratification d'une variable à expliquer en deux ou plusieurs variables à expliquer par autant de modèles explicatifs. Ce développement théorique ne pouvait que perdre sa place dans ce chapitre et renforce la nécessité d'une présentation sommaire de la régression multiple et de ses diverses hypothèses.

(1) Chapitre 7

SECTION I - GENERALITES SUR LES MODELES

Le modèle, qui "consiste en une présentation formelle d'idées ou de connaissances relatives à un phénomène"⁽¹⁾ est devenu de nos jours un outil de travail privilégié dans les sciences sociales, car il exige une rigueur implacable dans le déroulement du raisonnement. La démarche s'effectue en deux temps donnant naissance successivement au modèle conceptuel et au modèle mathématique.

1.1 - Modèles conceptuels et modèles mathématiques

1.11 - Modèle conceptuel

Un modèle conceptuel s'attache à décrire l'enchaînement logique d'événements ou d'idées que l'on peut représenter à l'aide de variables. Dans cette conception, un modèle se ramène donc à un ensemble de variables qui agissent les unes sur les autres par l'intermédiaire de relations non incompatibles.

Ces relations peuvent être le fruit de l'observation ou de la théorie⁽²⁾. L'intérêt de ces relations est de traduire de manière formalisée des causalités postulées.

Dès lors, deux types de variables peuvent être envisagées dans un modèle : les variables endogènes et les variables exogènes.

(1) Malinvaud . - Méthodes Statistiques de l'économétrie (1969)

(2) Il faut remarquer que toute théorie ne peut être mise sous la forme d'un modèle ; il faudra en effet que toutes les données de la théorie puissent être traduites par des variables et que celles-ci soient toutes reliées entre elles par des relations. Le fait pour une théorie d'être modélisable ne préjuge , d'ailleurs, en rien de sa validité, laquelle doit être jugée d'après son degré d'explication d'un ensemble d'observations.

Comme leur étymologie l'indique, les variables exogènes sont celles dont la détermination est étrangère au modèle, tandis que les variables endogènes sont déterminées par le modèle⁽¹⁾.

Un modèle conceptuel peut revêtir une forme littéraire (texte décrivant un ensemble de mécanismes) ou graphique (graphe de causalité dans lequel les variables sont symbolisées par des cases et les relations entre ces variables, par des flèches).

La forme graphique ne permet pas d'expliquer le choix des relations (ce que permet la forme littéraire), mais son avantage synoptique fait que l'on a bien souvent intérêt à l'utiliser comme complément de la forme littéraire. En outre cette forme graphique est bien souvent un intermédiaire commode pour le passage du modèle conceptuel au modèle opérationnel.

1.12 - Modèle mathématique

a) Définition

Le modèle conceptuel reste un outil d'analyse. Cette étape d'analyse est logiquement suivie d'une phase de quantification. On appelle modèle formalisé ou modèle mathématique, un modèle dans lequel la forme précise des relations est mathématiquement définie. A ce stade du raisonnement, les valeurs des paramètres et des coefficients peuvent n'être pas connues.

Un modèle mathématique revêt donc la forme d'un système d'équations (avec éventuellement des inéquations). Dans la pratique on utilise le plus souvent les seules équations linéaires. Cette restriction n'est pas très contraignante en fait. En effet, un modèle est une approximation et celle-ci est le plus souvent suffisante pour décrire certains phénomènes ; on effectue des transformations appropriées des variables : une variable à croissance exponentielle est transformée en logarithme pour obtenir une croissance arithmétique par exemple, où l'on restreint l'intervalle de variation des variables. Toutefois, cette observation est justifiée surtout pour les modèles à moyen et long terme (les modèles à court terme décrivent souvent des phénomènes cycliques).

...

(1) Cette terminologie introduite par des économistes (TINTNER et TINBERGEN) est utilisable dans d'autres disciplines scientifiques, ces deux concepts étant inhérents à la nature même de celui de modèle.

b) Types d'équation

On peut ranger les équations d'un modèle en quatre classes, les deux premières énoncent des propositions testables empiriquement, tandis que les deux dernières sont vraies de par leur nature :

- Les équations de comportement décrivent le comportement d'agents économiques (exemple : fonction de consommation).
- Les équations techniques décrivent les relations technologiques ou institutionnelles (exemple : fonction de production).
- Les identités sont des équations qui établissent une égalité qui est toujours vraie en raison des définitions introduites (exemple : la somme nationale des soldes migratoires interdépartementaux est nulle).
- Enfin, les conditions d'équilibre sont des équations vraies en raison de l'hypothèse particulière de l'équilibre postulé. Il faut souligner qu'un modèle opérationnel peut ne pas faire appel à tous les types d'équations.

c) Classification des modèles mathématiques en fonction de leurs finalités

Un modèle mathématique peut être élaboré pour répondre à trois besoins différents :

Le premier d'entre eux consiste à essayer de fournir une explication cohérente ou une synthèse commode, à une série d'observations. Le modèle est alors dit descriptif. La différence entre le modèle conceptuel et le modèle descriptif réside dans le fait que l'incidence causale d'un phénomène sur un autre est "chiffrée".

Mais lorsque l'adéquation d'un modèle descriptif au réel semble suffisante, il peut être intéressant de conserver les paramètres calculés et les hypothèses du modèle et :

- soit d'introduire de nouvelles valeurs de variables exogènes, le modèle est dit alors prévisionnel ;
- soit de se fixer des valeurs de variables endogènes (et éventuellement de certaines variables exogènes) puis de chercher des variables exogènes qui permettent d'atteindre cet objectif : le modèle est dit alors de planification ou normatif. Ces objectifs sont appelés variables cibles.

...

1.2 - La causalité dans les modèles mathématiques

Les relations de causalité jouent au niveau des variables mais un modèle peut décrire l'articulation de plusieurs mécanismes dotés d'une certaine indépendance et que l'on appelle sous-modèle.

1.21 - Causalité au niveau des variables

a) La notion de variables pré-déterminées

Dans la plupart des cas⁽¹⁾ il est possible de trouver pour la détermination des variables endogènes, une certaine hiérarchie dans la résolution des équations. Il s'en suit qu'une variable endogène peut être déterminée par les valeurs des variables exogènes et de variables endogènes déterminées dans des équations déjà résolues. On appelle variables pré-déterminées ces variables exogènes et endogènes. Il va de soi que le nombre de ces variables dépend étroitement de l'équation considérée.

b) L'Introduction du temps dans la causalité

La mesure de tout phénomène socio-économique s'effectue dans un cadre spatio-temporel. La conception du cadre temporel mène à trois types d'analyse et de ce fait à trois types de modèle.

Un modèle peut ne faire intervenir que des variables concernant la même date ou la même période. Ce type d'analyse où le temps n'intervient pas de façon explicite est appelé analyse statique.

Lorsque, dans un modèle, des équations techniques ou de comportement relient des variables observées pour plus d'une date ou d'une période, l'analyse est dite analyse dynamique⁽²⁾.

Avec Samuelson, un troisième type d'analyse a été développé, l'analyse statique comparée qui "se propose la recherche des variations survenues dans un système lorsque celui-ci passe d'une position d'équilibre à une autre sans se préoccuper du processus transitoire de cet ajustement"⁽³⁾. L'intérêt d'un

...

(1) Sauf cas de système de CRAMER indécomposable (cf. 1. 3. 2).

(2) Cette définition s'inspire de celle de Frisch (R. E. S. 1935) pour lequel "un système est dynamique si son comportement dans le temps est déterminé par des équations fonctionnelles où "des variables à des instants différents" interviennent de façon essentielle".

(3) Samuelson (Foundations) 1945.

tel type d'analyse est de renseigner sur la forme de certaines équations d'équilibre.

L'introduction de variables retardées par l'analyse dynamique pose un certain nombre de problèmes spécifiques⁽¹⁾ c'est pourquoi il est nécessaire de particulariser ce type d'analyse.

1.22 - Causalité entre sous-modèles

L'analyse des phénomènes économiques décompose le plus souvent, et au prix de simplifications parfois grossières, la réalité en un certain nombre de mécanismes économiques, articulés entre eux et dont chacun peut constituer à lui seul un modèle. On en arrive dès lors à la notion de sous-modèle.

a) La notion de sous-modèle

A la limite tout modèle est un sous-modèle d'un modèle plus vaste décrivant mieux la réalité et la notion de variable exogène s'estompe ou disparaît. Ce qui est variable exogène dans le sous-modèle peut être en fait, variable endogène dans le modèle plus général. C'est pourquoi la notion de variable pré-déterminée recouvre mieux l'aspect relatif de la causalité dans un modèle.

On entendra donc par sous-modèle un ensemble d'équations permettant de déterminer un ensemble de variables endogènes à l'aide d'un ensemble de variables pré-déterminées.

La notion de variable pré-déterminée permet d'introduire une notion de hiérarchie entre sous-système et donc de causalité, au sens large, entre sous-modèles. Un sous-modèle (A) est dépendant d'un autre sous-modèle (B) s'il existe au moins une variable pré-déterminée dans A qui soit variable endogène dans B. Bien entendu, aucune relation de ce type n'existe dans l'autre sens, auquel cas il y aurait inter-dépendance entre les sous-modèles A et B.

A moins de n'avoir qu'une juxtaposition de sous-modèles indépendants (c'est-à-dire n'ayant aucune position hiérarchique dans le modèle) ce qui ne présenterait guère d'intérêt théorique (et à la limite qu'un intérêt de compilation), tout sous-modèle est relié au moins à un autre par un lien de causalité.

...

(1) Par exemple : problème de l'autocorrélation des résidus dans les techniques de régression linéaire.

Pour s'assurer que l'on a effectivement affaire à un sous-modèle il est bien souvent nécessaire de faire appel aux concepts de forme structurelle et de forme réduite.

b) Forme structurelle et forme réduite

Lorsque les équations d'un modèle décrivent le fonctionnement des mécanismes tels qu'ils sont sensés se produire, on parle de forme structurelle du modèle.

Le plus souvent, cette forme structurelle est impropre à la résolution du système d'équations. Il est dès lors nécessaire d'effectuer des transformations pour mettre le système d'équation sous la forme explicite (au sens mathématique) qui en permettra la résolution. On parlera alors de forme réduite du modèle.

Dans un sous-modèle mis sous sa forme réduite on doit toujours être en mesure de trouver les valeurs des variables endogènes. C'est dans ce sens que la mise sous forme réduite permet de s'assurer que l'on a bien affaire à un sous-modèle.

Dans la pratique, les sous-modèles n'apparaissent pas toujours nettement et l'on procède souvent à l'analyse de la forme réduite du modèle pour en tirer les différents sous-modèles. C'est ce que l'on va analyser présentement.

1.3 - Incidence des relations causales sur la structure du modèle mathématique (1)

1.31 - Les différents types de système linéaire

Un système d'équations linéaires est dit système régulier de Cramer si :

- le nombre d'équations est égal au nombre de variables endogènes ;
- les équations sont indépendantes, c'est-à-dire qu'il n'existe pas d'équation se déduisant des autres ;
- il n'existe pas de couple d'équations qui soit en contradiction; c'est-à-dire, en pratique, qu'une variable endogène ne peut être déterminée par 2 équations de comportement (ou techniques) différentes.

...

(1) Voir KOGIKU : Introduction aux modèles macro-économiques, Sirey, 1971.

Un système est dit système indéterminé si le nombre d'équations est inférieur au nombre de variables endogènes. Dans ce dernier cas le but du modèle n'est pas complètement atteint puisqu'il n'y a pas détermination des variables endogènes par le modèle. L'intérêt d'un tel modèle n'est pas complètement négligeable, dans la mesure où il permet de décrire une partie des mécanismes.

1.32 - Décomposition des systèmes de Cramer

Pour l'étude des systèmes de Cramer, on fait en général appel à une matrice de variables muettes avec les variables endogènes repérées par les colonnes et les équations par les lignes : la variable muette $A(I, J)$ vaut 1 si dans l'équation I considérée la variable endogène J est présente, et 0 dans le cas contraire.

L'analyse d'un système de Cramer le classe dans l'une de ces trois catégories :

a) Système de Cramer décomposable.

Dans ce cas le système se partitionne en sous-système de Cramer. D'un point de vue technique cela entraîne que la matrice des variables muettes, par des permutations adéquates de lignes et de colonnes peut être transformée en matrice bloc-diagonale dont les éléments seront des matrices carrées (lesquelles représentent autant de sous-système de Cramer), réduites éventuellement à une seule variable muette et dont la valeur est 1.

b) Système de Cramer de Causalité

Dans ce cas le système comprend un ou plusieurs sous-systèmes de Cramer⁽¹⁾ mais les variables endogènes déterminées dans ces sous-systèmes sont utilisées dans la détermination du reste du système, on a donc une hiérarchie entre différents sous-systèmes.

Cette hiérarchie est repérée par un ordre, on parlera du sous-système d'ordre 1 etc. Les variables endogènes du sous-système d'ordre 0 sont, pour les sous-systèmes d'ordre 1 qui les utilise, des variables pré-déterminées. La notion de sous-système correspond donc à celle de sous-modèle et peut être un outil puissant dans la détermination des sous-modèles.

...

(1) Lorsque tous ces sous-systèmes de CRAMER ne sont que matrices réduites à un seul élément, le modèle est dit récursif.

Des permutations appropriées de la matrice des variables muettes permettent de la transformer en une matrice triangulaire dont les éléments de la diagonale sont 1 ou des matrices carrées (qui correspondent à des sous-systèmes de Cramer).

c) Système de Cramer indécomposable

Un système de Cramer est indécomposable lorsqu'il ne peut se subdiviser en systèmes de Cramer. Dans un tel système toutes les variables sont déterminées par l'interaction simultanée de toutes les équations du système. Un tel système (ou sous-système) est dit encore Système de Cramer minimal. Dans ce cas il n'est pas possible d'obtenir de matrice bloc-diagonale.

1.4 - Différentes techniques de détermination des variables endogènes

Connaître la liaison causale permettant de déterminer les valeurs des variables endogènes ne suffit pas, encore faut-il savoir si ces résultats sont définis de façon ponctuelle ou non et de savoir leur mode de variation, ces deux choix ayant des implications importantes du point de vue des techniques de modélisation.

1.41 - Précision des résultats issus d'un modèle opérationnel

Les variables endogènes peuvent être déterminées de deux façons différentes ; il s'en suit que deux types de modèles différents existent.

Le modèle déterministe, une fois déterminés les paramètres et coefficients des équations fait correspondre à un ensemble de valeurs des variables exogènes, une valeur et une seule pour chaque variable endogène.

Le modèle probabiliste (appelé encore aléatoire) lui, fait correspondre à cet ensemble de valeurs de variables exogènes, une loi de probabilité de forme définie à priori pour chaque variable endogène (et dont les paramètres sont déterminés par l'ensemble des observations).

Cette distinction implique, si le choix se porte sur un modèle probabiliste, l'introduction

...

d'hypothèses supplémentaires et auxquelles peuvent correspondre un certain nombre de contraintes.(1)

L'un des principaux intérêts du choix d'un modèle probabiliste est de permettre de juger certaines hypothèses sur lesquelles se fonde le modèle (degré de signification, au regard d'observations, d'une liaison causale postulée...)

1.42 - Mode de variation des variables

Les phénomènes qu'un modèle économique cherche à présenter peuvent dans la réalité varier d'une façon continue (comme le PNB), ou discrète (comme le taux de l'escompte), que l'analyse auquel le modèle fait appel soit dynamique ou statique, les variables, de par leur nature, ont un mode de variation(2).

En fait le concept de variation continue est une approximation de la réalité car celle-ci se heurte plus ou moins vite (souvent en fonction des unités adoptées) à la notion d'indivisibilité physique des phénomènes réels, mais l'appel à cette notion correspond à l'idée qu'il n'y a pas rupture dans la variation du phénomène que l'on étudie.

Un modèle, de par sa vocation, est une représentation simplifiée de la réalité. L'une des simplifications (motivée par des raisons de calcul ou de donné) peut être justement le mode de variations des variables.

a) Formalisation en continu

Les variables varient de façon continue. La conséquence de ce mode de variation est la possibilité de faire appel aux équations différentielles : ceci peut ne pas apparaître de façon explicite car on travaille souvent au niveau de la primitive(3).

...

- (1) Par exemple un modèle (ou un sous-modèle) déterministe est dit récursif si notre matrice de variables muettes est triangulaire, la diagonale ne comprenant pas de matrice carrée. Pour que ce modèle puisse être considéré comme aléatoire il faut en outre que la matrice des covariances des écarts entre les valeurs théoriques des variables endogènes et les valeurs observées soit diagonale, ceci quelle que soit la loi de probabilité de nos variables endogènes. Seul un calcul confirmant cette condition nécessaire peut nous autoriser à choisir un modèle probabiliste.
- (2) C'est le cas par exemple de la variation continue de l'Offre et de la Demande d'un bien en fonction du prix, dans les équations d'équilibre faisant appel à l'analyse statique qui en est faite par les économistes marginalistes.
- (3) C'est le cas par exemple des équations d'Offre et de Demande d'un bien en fonction du prix chez les marginalistes.

b) Formalisation en discret

Les variables varient de façon discrète. La conséquence en est que le modèle fait appel à des équations aux différences.

1.43 - Solution analytique ou appel à la simulation

Lorsque les paramètres des équations d'un modèle sont tous connus, on dispose de deux méthodes pour examiner les réactions du modèle : on peut utiliser une résolution analytique, ou les équations peuvent être simulées.

La résolution analytique permet par des formules littérales d'exprimer une variable endogène en fonction des autres variables et des paramètres. Elle s'impose par exemple lorsque l'on est en présence de système de Cramer indécomposable. Toutefois la recherche d'une solution analytique peut être extrêmement ardue (cas par exemple de certains modèles cherchant à optimiser une fonction objectif), ou ne présenter qu'un intérêt pratique, faible (formule très complexe).

Dans ce cas, et si cela est possible, on a intérêt à simuler le système, c'est à dire à rentrer les valeurs des variables exogènes pour obtenir la solution particulière des variables endogènes correspondantes. Lorsque le système de Cramer comporte un sous-système indécomposable, la solution analytique de ce sous-système doit être utilisée.

L'avantage d'un tel procédé est de permettre de mettre en évidence facilement la sensibilité d'une variable endogène aux variations d'une ou plusieurs variables exogènes.

Par ailleurs, l'appel aux techniques de simulation permet d'approximer de façon satisfaisante certaines solutions. C'est ainsi que l'on peut remplacer un système d'équations simultanées différentielles par un système d'équations aux différences, non simultanées ; ce qui suppose que l'on s'intéresse à la recherche de solutions particulières et non d'une solution générale.

Enfin dans le cadre de modèles dynamiques faisant intervenir un processus itératif au cours du temps, les techniques de simulation peuvent être à peu près les seules utilisables pour l'étude des conséquences de certaines hypothèses, le raisonnement intuitif étant incapable de fournir une réponse. C'est le cas par exemple de la recherche d'un état stationnaire dans certains modèles Markoviens.

SECTION 2 - UNE TECHNIQUE PRIVILEGIEE DES MODELES MATHEMATIQUES : LA REGRESSION MULTIPLE

Le spécialiste des questions socio-économiques a à sa disposition plusieurs moyens de "synthétiser" la masse d'information dont il veut tenir compte. Dans le cadre de relations fonctionnelles à mettre en évidence, l'une des techniques les plus couramment utilisées est celle de la régression multiple. Cette fréquence d'utilisation s'explique d'une part par la facilité d'utilisation de ces techniques, et d'autre part par le fait que beaucoup de phénomènes économiques peuvent être décrits par des liaisons linéaires, en faisant éventuellement appel à une transformation appropriée des variables (passage par les logarithmes en particulier).

L'utilisation de ces techniques implique l'acceptation de certaines hypothèses qui leur sont propres. Dans la pratique, ces hypothèses en entraînent d'autres qui dépendent du modèle étudié. Nous avons donc été contraints de résumer brièvement ces hypothèses de base, d'une part pour situer les hypothèses complémentaires que l'on sera amené à faire (chapitre 7 section 2) et d'autre part pour permettre de présenter succinctement la régression non linéaire, que nous prévoyons d'utiliser au chapitre 13 ainsi que pour un développement théorique que nous avons dû réaliser pour aborder correctement le problème de la stratification d'une variable endogène. Examinons successivement ces divers points.

2.1 - La régression multiple dans le cadre des modèles déterministes

a) La méthode

Seront envisagés ici les seuls modèles déterministes à une équation.

La représentation de n zones introduit l'espace vectoriel à n dimensions de toutes les variables observables sur les n zones.

Un modèle déterministe, à une équation, est défini par :

- les données d'une variable à expliquer (variable endogène) Y soit :

$$Y = [y_1, y_2, \dots, y_n]$$

- Les données de variables explicatives (ou exogènes) en nombre (k) soit :

$$X_1, X_2, \dots, X_k$$

appartenant à l'espace vectoriel à n dimensions.

Chaque zone fournit une valeur pour une variable X_j donnée :

$$X_j = \{ x_{1j}, x_{2j}, \dots, x_{ij}, \dots, x_{nj} \}$$

- un modèle, c'est à dire une théorie représentant Y en fonction des k variables exogènes, une partie U demeurant inexplicée :

$$Y = F(X_1, X_2, \dots, X_k) + U$$

Le vecteur U appartenant lui aussi à notre espace à n dimensions :

$$U = \{ u_1, u_2, \dots, u_n \}$$

est appelé vecteur d'écart ou résidu et est sensé représenter l'influence des variables qui n'ont pas été prises en compte dans le modèle.

En général le modèle ne spécifie que la forme F (linéaire par exemple) mais ne précise pas la valeur des divers paramètres qu'elle comporte. L'ajustement consiste à choisir, parmi toutes les fonctions F du type considéré, celle qui fournit un vecteur départemental théorique Y le plus proche possible du vecteur Y.

Pour certaines raisons que l'on verra ultérieurement on choisi :

- la fonction F linéaire :

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_k x_{ik} + u_i$$

et l'indicateur

$$\sum u_i^2$$

La méthode retenue consiste à minimiser cet indicateur d'écart quadratique et pour cette raison est appelée méthode des moindres carrés.

b) Résultats

La détermination des paramètres de l'équation matricielle

$$Y = X \cdot B + U$$

donne comme valeur au vecteur colonne B :

$$B = (X' \cdot X)^{-1} \cdot X' \cdot Y$$

Quels que soient les vecteurs de données pris en considération, la méthode conduit toujours à des valeurs de coefficients de régression sans que l'on puisse apprécier :

- la validité de la liaison posée comme hypothèse ;
- la précision avec laquelle sont déterminés les coefficients.

Il faut en outre souligner que la méthode des moindres carrés, mise au point pour résoudre ce problème :

- ne permet pas de choisir la théorie ;
- ne fait aucune référence explicite à la nature du vecteur U dans la détermination des paramètres.

Par ailleurs, l'indication d'écart quadratique permet de tester, l'amélioration de l'ajustement, lors de l'introduction de nouvelles variables exogènes. L'étude de ces améliorations mène à l'introduction de coefficients de corrélation. (Mais l'étude de coefficients de corrélation ne permet pas de tester la liaison postulée).

2.2 - La régression multiple dans le cadre des modèles probabilistes

Là encore ne seront envisagés que les modèles probabilistes à une équation. Un progrès important dans l'interprétation de la méthode des moindres carrés et la signification de ses résultats, peut être réalisé

...

en considérant les résidus u_i comme des variables aléatoires.

Cette hypothèse revient à supposer que l'effet des variables secondaires non retenues dans le modèle peut être traité comme une variable aléatoire.(1)

A un vecteur ligne donné de la matrice X (considérée comme certaine) correspond donc une distribution de probabilité des valeurs de y_i . Le vecteur colonne B des paramètres devient dans ce type d'approche une estimation du vecteur réel β non connu.

L'approche peut se faire en 2 temps en introduisant successivement 2 séries d'hypothèses.

a) Première série d'hypothèses

- (1) la matrice X est observée sans erreur et non aléatoire(2)
- (2) la loi de probabilité de u_i est indépendante de x_{ij} pour tout J
- (3) $E(u_i) = 0$ pour tout I
- (4) $E(u_i^2) = \sigma^2$ pour tout I (homoscedasticité)
- (5) $E(u_i \cdot u_j) = 0$ pour tout I différent de J (non auto-corrélation des résidus)
- (6) les moments du 1er et du second ordre tendent vers des limites finies quand la taille de l'échantillon tend vers l'infini.

Sous ces hypothèses on démontre que B est un estimateur convergent et sans biais de β .

Il s'en suit que nous pouvons associer maintenant à la valeur des coefficients de régression, leur écart type, ce qui facilite l'interprétation des résultats et permet de porter un premier jugement sur leur validité. Mais il faut souligner que seule une connaissance de la loi de distribution des paramètres et donc la constitution d'intervalles de confiance pour les paramètres, permet de juger en probabilité la valeur de Y , dans une utilisation en modèle prévisionnel.

(1) Et plus précisément dans le cadre de l'hypothèse 7 introduite au § 2.2b (cf infra), que ces variables aléatoires sont indépendantes, de lois quelconques mais en assez grand nombre pour que le théorème central-limite joue, et donc que la résultante de leurs actions conjuguées soit une variable normale.

(2) Cette hypothèse peut être levée ultérieurement.

b) Seconde série d'hypothèses

- (7) loi (U) = loi normale centrée

La connaissance de la loi de U permet alors de juger par intervalle de confiance :

- les paramètres de la régression
- le coefficient de corrélation de la population mère ;
- une valeur calculée de Y en utilisation prévisionnelle du modèle⁽¹⁾. Si la fourchette est trop large le modèle ne devient plus intéressant à utiliser. On se trouve alors en présence d'un test d'arrêt possible dans la recherche de variables explicatives pour le modèle.

Cette hypothèse (7) est testable par un χ^2 , pour le seul risque de première espèce.

2.3 - Régression non linéaire

La fonction F(cf en 2.1a) peut être choisie non linéaire. En effet si l'utilisation de transformation de variables en logarithme pour décrire certains phénomènes non linéaires n'introduit aucun biais en régression linéaire "unique" ; en régression linéaire multiple, la transformation d'une variable dépendante en logarithme amène à minimiser le poids dans l'indicateur d'écart quadratique retenu, des écarts importants, par rapport au trend de la variable considérée (dans l'hypothèse où la transformation en logarithme n'a pas été opérée sur toutes les variables).

La technique de régression non linéaire permet d'éviter ce biais. Par ailleurs, elle permet de traiter n'importe quelle fonction F.

On part d'un ensemble initial de valeurs des paramètres b_j , que l'on note b_{j0} . Cet ensemble est en général constitué de zéro ou de valeurs obtenues par une fonction linéaire simplifiée approximant la fonction retenue.

On a donc une première valeur théorique de la variation dépendante et un premier vecteur d'écart

$$Y_0 = F(X, B_0) + U_0$$

(1) On peut noter ce résultat peu connu mais nécessaire à l'établissement de la fourchette de Y : la variance de l'estimateur d'une valeur trouvée de la variable endogène, en appliquant à l'équation de régression obtenue un vecteur donné x_a de variables exogènes, est $(X \text{ étant la matrice des observations ayant servi à déterminer les coefficients de régression et } \sigma^2 \text{ la variance résiduelle obtenue sur ces mêmes observations})$: voir Kendall "The advanced theory of Statistics" tome 2, page 363 :

$$\left[\sigma^2 \left(1 + x_a' (X, X')^{-1} x_a \right) \right]$$

On définit alors la matrice suivante Z (à n et k dimensions) :

$$Z_0 = \frac{\partial (F(X, B))}{\partial B}$$

et l'on utilise la technique de la régression linéaire multiple pour la fonction

$$U_0 = Z C_0 + U_1$$

d'où l'on tire les paramètres C_0 puis B_1 en posant :

$$B_1 = B_0 + C_0$$

Ce processus est recommencé en remplaçant l'indice 0 par l'indice 1 et l'indice 1 par le 2 et ainsi de suite l'itération est arrêtée lorsque C_m devient négligeable. Le nombre d'itération m nécessaire (et donc le coût de l'opération) dépend des valeurs de départ retenues pour les paramètres.

2.4 - Problème posé par la stratification d'une variable endogène

Nous étudierons ce problème dans le cadre des modèles probabilistes (hypothèses 1 à 7). Toutefois, les résultats sont inchangés si l'on se place dans le cadre déterministe : seule l'interprétation des résultats diffère.

2.41 - Le problème posé

2.411 - L'intérêt du problème posé

Dans la recherche de modèles explicatifs satisfaisants il peut sembler intéressant dans certains cas de séparer une donnée que l'on cherche à expliquer en deux (ou plusieurs) composantes pour lesquelles un modèle explicatif, différent pour chaque composante, peut être trouvé. C'est ainsi par exemple que le tertiaire peut être a priori stratifié en deux : un tertiaire produisant des services **finals**, c'est-à-dire consommés par les particuliers et un tertiaire ne produisant que des services intermédiaires, c'est-à-dire concourant à la formation du PNB. On s'interrogera ici sur l'intérêt de cette stratification en s'appuyant sur l'exemple choisi.

2.412 - Définition des variables

Y_1 = vecteur départemental des effectifs du tertiaire produisant les services finals (l'indice Y_1 référencera toujours ce type de tertiaire).

Y_2 = vecteur départemental des effectifs du tertiaire produisant des services intermédiaires.

Y = vecteur départemental des effectifs du tertiaire. On a bien entendu $Y = Y_1 + Y_2$

X_1 = matrice départementale de l'ensemble des variables explicatives intervenant de façon satisfaisante pour expliquer Y_1 . Cette matrice est de dimension $n.k_1$, avec $n = 95$. On définit de même

X_2 ; X_1 et X_2 peuvent avoir plusieurs colonnes identiques (= mêmes variables explicatives).

X = matrice départementale de l'ensemble des variables explicatives intervenant soit dans X_1 soit dans X_2 . Sa dimension est $(95, k)$ avec : k_1 ou $k_2 < k < k_1 + k_2$

B_1, B_2, B sont les vecteurs de coefficients de régression (estimation de β_1, β_2 et β expliquant Y_1, Y_2 et Y). On a donc :

$$\left. \begin{array}{l} Y_1 = X_1 \beta_1 \\ Y_2 = X_2 \beta_2 \end{array} \right\} \text{modèle (1) avec } \hat{\beta}_1 = B_1 \text{ et } \hat{\beta}_2 = B_2$$

$$\left. \begin{array}{l} Y = X \beta \end{array} \right\} \text{modèle (2) avec } \hat{\beta} = B$$

γ_1 et γ_2 sont les vecteurs de coefficients de régression obtenus sur la matrice X (autant que de variables retenues dont certains éléments sont nuls). On a donc :

$$\left. \begin{array}{l} Y_1 = X \cdot \gamma_1 \\ Y_2 = X \cdot \gamma_2 \end{array} \right\} \text{modèle (3) avec } \hat{\gamma}_1 = C_1 \text{ et } \hat{\gamma}_2 = C_2$$

Le problème posé se décompose en deux sous-problèmes à résoudre successivement : quelles sont les liaisons entre les modèles 1 et 3 d'une part et 3 et 2 de l'autre. S'il s'avère que 1 est équivalent à 3 et 3 équivalent à 2 il n'y a aucun intérêt à faire de stratification. S'il n'en est pas ainsi on peut malgré tout chercher à calculer sous certaines hypothèses les biais introduits sur les estimateurs. Mais au préalable, il nous faut exprimer les coefficients de régression en fonction de la matrice des coefficients de corrélation.

2.413 - Coefficients de régression et coefficients de corrélation

Le modèle 2 est pris ici comme support du raisonnement. Désignons par X_0 le vecteur Y , par R la matrice des coefficients de corrélation entre les différentes variables et par V celle des variances - covariances. Ces deux dernières matrices sont de rang $k + 1$. On désignera par R^{lm} le mineur de R obtenu en supprimant la ligne l et la colonne m .

On définit les variables centrées-réduites suivantes :

$$z_i = (x_{0i} - \bar{x}_0) / \sqrt{v_{00}} \quad \text{et} \quad w_{ij} = (x_{ij} - \bar{x}_j) / \sqrt{v_{jj}} \quad \text{pour } j = 1, k$$

Dans le cadre du modèle en variable centrée réduite l'équation de régression devient :

$$Z = WD \quad (\text{modèle 4 bis})$$

d'où : $D = (W' W)^{-1} W' Z$ avec $(W W') = n R^{00}$ et $W' Z = n R_0$. On peut encore écrire D sous la forme :

$$D = (R^{00})^{-1} R_0$$

La correspondance entre les coefficients de régression d_j et b_j est donc :

$$b_j = d_j \sqrt{\frac{v_{00}}{v_{jj}}}$$

2.42 - Equivalence des modèles 3 et 2

Ces diverses propriétés que nous allons voir sont vraies qu'il y ait ou non indépendance entre les variables explicatives.

Dans le cadre du modèle probaliste on sait, indépendamment du fait que Y_1 et Y_2 soient indépendants ou non, que l'espérance mathématique de leur somme est égale à la somme de leurs espérances mathématiques

$$E(Y) = E(Y_1 + Y_2) = E(Y_1) + E(Y_2)$$

Il s'en suit, la matrice X étant une matrice certaine, que :

$$E(Y) = X.E(\beta) = X.[E(\gamma_1) + E(\gamma_2)] = X[C_1 + C_2]$$

Il s'en suit que du point de vue des estimations ponctuelles des coefficients de régression, on a *aucun intérêt* à opérer à une stratification de la variable endogène et que les modèles 3 et 2 sont strictement équivalents de ce point de vue.

Examinons maintenant ce qu'il en est de la variance des coefficients de régression, ce dernier point est important puisque les tests de rejet sont basés sur ces variances (hypothèse 7).

On sait par ailleurs que :

$$V(Y) = V(Y_1 + Y_2) = V(Y_1) + V(Y_2) + 2 \text{COV}(Y_1, Y_2)$$

et que si Y_1 et Y_2 sont indépendants leur covariance est nulle

On peut écrire :

$$V(Y) = V(X\beta) = X'.X.V(\beta)$$

$$V(Y_1) = V(X\gamma_1) = X'.X.V(\gamma_1)$$

$$V(Y_2) = V(X\gamma_2) = X'.X.V(\gamma_2)$$

$$\text{Or } V(Y) = V(Y_1 + Y_2) = V[X(\gamma_1 + \gamma_2)] = X'.X.V(\gamma_1 + \gamma_2) = X'.X[V(\gamma_1) + V(\gamma_2) + 2 \text{COV}(\gamma_1, \gamma_2)]$$

Il s'en suit que si Y_1 et Y_2 sont deux variables aléatoires indépendantes, on a :

$$V(\beta) = V(\gamma_1) + V(\gamma_2)$$

La qualité des estimateurs est donc inchangée dans le seul cas d'indépendance. Dans le cas contraire, on peut très bien avoir un rejet par le test de Student d'une variable donnée pour l'une ou l'autre des équations du modèle 3 et acceptation de cette même variable dans le cadre du modèle 2.

2.43 - Equivalence des modèles 3 et 2

2.431 - Cas d'indépendance totale entre les variables explicatives

a) Liaison entre coefficients de régression et coefficients de corrélation

L'indépendance totale entre les variables explicatives se traduit par le fait que la matrice des coefficients de corrélation est une matrice-unité. Il s'en suit alors que :

$$D = R_o$$

$$\text{d'où : } b_j = r_{oj} \sqrt{\frac{v_{oo}}{v_{jj}}} = \frac{v_{oj}}{\sqrt{v_{oo} \cdot v_{jj}}} \sqrt{\frac{v_{oo}}{v_{jj}}} = \frac{v_{oj}}{v_{jj}}$$

Ce résultat est une généralisation dans l'espace à k dimensions du résultat classique de la régression simple.

Désignons par un * les résultats du modèle 4 bis dans le cas d'indépendance des variables explicatives. On peut tout de suite indiquer les écarts imputables à la multicollinéarité :

$$(D - D^*) = [(R^{oo})^{-1} - I] \cdot R_o$$

ce qui permet facilement de calculer le vecteur d'écart $(B - B^*)$

β) Equivalence des modèles 1 et 3

La matrice des coefficients de corrélation reste par hypothèse une matrice unité - et les coefficients de corrélation entre Y et les variables de X_j non retenus sont nuls (ou voisin de zéro), on voit immédiatement que les coefficients de régressions correspondant à une même variable donnée dans les modèles 1 et 3, sont identiques que l'on parte de l'une ou de l'autre matrice.

Si les coefficients de corrélation entre Y et les X_i non retenus de X sont différents de 0, il est facile de faire un test de Student pour voir s'il y a vraiment lieu d'écartier la variable considérée (et ce d'autant plus facilement que la matrice des variances - covariances des coefficients de régression est une matrice diagonale).

2.432 - Cas de dépendance entre les variables explicatives

Faisons une partition de la matrice X en 2 blocs en effectuant les permutations nécessaires pour que l'on retrouve la matrice X_1 à gauche. Posons $X_a = X_1$

$$X = \begin{bmatrix} X_a & X_b \end{bmatrix} \quad \text{et} \quad B = \begin{bmatrix} B_a \\ B_b \end{bmatrix}$$

En raisonnement en variables centrées réduites, le vecteur D peut s'écrire :

$$D = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} \begin{bmatrix} R_{0a} \\ R_{0b} \end{bmatrix} \quad \text{avec} \quad \left\{ \begin{array}{l} a = R_{ij} \quad \text{et } i \text{ et } j = 1, n_1 \\ b = R_{ij} \quad \text{et } i = 1 + n_1, n ; j = 1, n_1 \\ c = R_{ij} \quad \text{et } i = 1, n_1 ; j = 1 + n_1 \\ d = R_{ij} = a \quad i \text{ et } j = 1 + n_1, n \\ R_{0a} = R_{ij} \quad \text{et } i = 0 ; j = 1, n_1 \\ R_{0b} = R_{ij} \quad \text{et } i = 0 ; j = 1 + n_1, n \end{array} \right.$$

D peut s'exprimer en fonction de a, b, c et d comme suit :

$$D = \begin{bmatrix} a^{-1} + a^{-1} b (d - ca^{-1} b)^{-1} ca^{-1} & - a^{-1} b (d - ca^{-1} b)^{-1} \\ - (d - ca^{-1} b)^{-1} ca^{-1} & (d - ca^{-1} b)^{-1} \end{bmatrix} \begin{bmatrix} R_{0a} \\ R_{0b} \end{bmatrix}$$

On constate alors que l'écart introduit par la colinéarité est :

$$\left[a^{-1} b (d - ca^{-1} b) ca^{-1} \right] R_{oa} - \left[a^{-1} b (d - ca^{-1} b)^{-1} \right] R_{ob}$$

Cet écart qui reflète la distorsion faite par l'approximation du modèle 1 par le modèle 3 peut être négligeable et doit s'apprécier en fonction des écart-types des divers coefficients de régression. On peut espérer toutefois qu'une solution théorique à ce problème puisse être trouvée en introduisant des hypothèses simplificatrices, on remarque en effet que si b (et donc c) est nul, la distorsion introduite est nulle elle aussi et que par ailleurs on démontre⁽¹⁾ facilement que l'inverse d'une matrice constituée d'éléments tous identiques à θ sauf les éléments diagonaux qui, eux sont égaux à 1, est une matrice symétrique dont les éléments non diagonaux sont tous égaux et les éléments diagonaux égaux entre eux. Il est peut être possible remplaçant a, b, c et d par des matrices que l'on puisse considérer comme des "bornes supérieures" des matrices réelles de trouver une solution théorique à ce problème de calcul de biais.

Nous devons souligner que notre impression de l'existence possible d'une solution analytique n'est fondée que sur ces seules considérations. Il nous semblerait utile de procéder ultérieurement à des recherches complémentaires.

$$\begin{aligned} (1) - \left[\begin{array}{cc} 1 & \theta \\ \theta & 1 \end{array} \right]^{-1} &= \frac{1}{D_n} \left[\begin{array}{cc} D_{n-1} & -D'_{n-1} \\ -D'_{n-1} & D_{n-1} \end{array} \right] \quad \text{avec} \begin{cases} D_n &= (1 - \theta)^{n-1} (1 + \theta(n-1)) \\ D'_n &= \theta(1 - \theta)^{n-1} \end{cases} \\ &= \frac{1}{(1 - \theta)(1 + \theta(n-1))} \left[\begin{array}{cc} 1 + (n-2)\theta & -\theta \\ -\theta & 1 + (n+2)\theta \end{array} \right] \end{aligned}$$

cette matrice inverse tend vers 0 avec θ .

Analyse théorique de la répartition des hommes
entre les différentes zones du territoire

CHAPITRE 4 - ANALYSE THEORIQUE de la REPARTITION des HOMMES entre les DIFFERENTES ZONES du TERRITOIRE

Notre propos ici est de passer succinctement en revue les diverses analyses théoriques qui ont une incidence sur la répartition spatiale des hommes et ce pour deux raisons. Tout d'abord il est nécessaire d'examiner si dans l'abondante littérature relative à l'analyse spatiale, certains aspects du problème qui nous préoccupe ont été abordés et ce tout particulièrement au niveau des mécanismes fondamentaux du modèle à élaborer. Ensuite pour chacun des sous-modèles il sera nécessaire de formuler un certain nombre d'hypothèses ; ces dernières résulteront soit de certaines considérations théoriques développées par des spécialistes de l'économie régionale, soit d'observations sur le passé.

Ces analyses théoriques comportant des impacts explicites sur la répartition spatiale des hommes sont de deux types :

- celles qui s'intéressent aux problèmes de localisation des entreprises (lesquelles ont évidemment un impact sur la main d'oeuvre nécessaire et donc sur la répartition des hommes)
- celles qui s'intéressent aux problèmes de migration.

Ces deux types d'analyse ne sont pas totalement indépendants, la première met l'accent sur le rôle de l'offre d'emploi tandis que le second s'attache aux motivations migratoires en général ou s'efforce d'expliquer la répartition spatiale des flux. C'est parce que le migrant est en même temps un demandeur de travail que cet aspect du problème ne peut être ignoré.

Nous limiterons notre "panorama" des diverses théories développées en fonction du but de notre modèle : c'est ainsi que l'on écartera les analyses de localisation ponctuelle ainsi que celles relatives à l'étude de phénomènes se produisant à un niveau spatial informel (modèles gravitaires par exemple) ou encore toutes les démarches explicitement tournées vers l'optimisation.

Nous ferons enfin dans une dernière section un très bref bilan.

SECTION 1 - MODIFICATION de la REPARTITION des HOMMES pour des RAISONS d'EMPLOI LIEES aux DECISIONS de LOCALISATION d'ENTREPRISES

Les problèmes de localisation des entreprises ont été jusqu'à présent envisagés sous deux angles. Le premier part de l'hypothèse de rationalité de comportement de l'entrepreneur, lequel est supposé vouloir minimiser ses coûts. Le second est basé sur des effets d'entraînement : une variation d'une catégorie de population pour une zone engendrant une variation d'une autre catégorie de population dans cette zone (ou dans une autre zone) ; il ne reste plus qu'à expliquer les variations motrices, ce que l'on verra dans un troisième paragraphe.

1.1 - L'analyse de localisation des entreprises dans l'optique des coûts

A la base de toutes ces approches on trouve l'idée que le phénomène de distance engendre un coût pour l'entrepreneur⁽¹⁾.

L'entreprise est ici conçue comme une entité économique qui se procure des facteurs de production (matières premières, produits semi-finis, main d'oeuvre...), pour les combiner dans le but de fournir au marché des produits. Dans cette optique le coût engendré par la distance peut intervenir d'une façon prépondérante, soit à la production, soit à la distribution (du moins pour certaines entreprises). Dans ce type d'analyse, implicitement l'entreprise est supposée soumise au marché, le prix de vente des produits est fixé de façon exogène (l'incidence des coûts dans la fixation monopolistique des prix n'étant pas du tout la même). On peut ajouter que ces analyses sont statiques et visent à définir, sous certaines hypothèses, une localisation théorique optimale.

1.11 - Coûts d'obtention de facteur de production

Certains facteurs de production, comme les matières premières sont, par essence, mobiles. Mais pour certains d'entre eux le coût de transport est important par rapport au coût de la matière première ; ce qui ne

(1) Voir : MORAN, *l'analyse spatiale en science économique* (Cujas, 1966) ; FRIEDMANN-ALONSO, *Regional development and Planning, A reader*, partie 2 (MIT Press, 1965) KARASKA-BRAMHILL, *Locational Analysis for manufacturing. A selection of Readings* (MIT Press, 1969) ; ISARD, *Methods of Regional Analysis* (MIT Press, 1960) ; et HAGGET, *Locational Analysis in Human geography* (Edward ARNOLD, 1965).

devient préoccupant que lorsque la part des coûts de transport de cette matière première dans le coût global du produit fini devient non négligeable. C'est le cas surtout des industries lourdes (minerais ou charbon pour la sidérurgie, électricité pour les industries de production de l'aluminium).

D'autres facteurs de production sont peu mobiles, voire fixes (le facteur travail en particulier). Pour se les procurer, l'entrepreneur doit payer un certain prix qui varie en fonction de la localisation de la main d'oeuvre et de sa qualification. Si la part du facteur travail dans le coût global du produit fini est importante, l'entrepreneur peut être amené à se localiser dans telle zone plutôt que dans telle autre, à moins de modifier la structure de la combinaison productive capital - travail (en fonction de leurs coûts relatifs). Un exemple peut être trouvé dans les industries textiles ou encore certaines industries de composants électroniques.

1.12 - Coûts de distribution des produits

Pour d'autres entreprises la production peut s'effectuer à un même coût pour diverses localisations. Toutefois, si le coût de transport du produit fini est important au regard de son prix de vente, l'entreprise a intérêt à se localiser près de son marché le plus important. C'est le cas par exemple de certaines industries de boisson (bière, limonade).

Lorsque le marché est vaste et connu, certaines techniques de recherche opérationnelle (programmation linéaire, théorie des graphes) peuvent aider à déterminer la localisation optimale.

1.13 - Confrontation des coûts d'obtention de facteurs et de distribution de produits

La confrontation de ces coûts a donné lieu à diverses réflexions théoriques concernant en fait la localisation d'établissements⁽¹⁾.

WEBER (1909) par exemple compare les prix du facteur travail et des coûts de transports de matières premières et des produits finis pour déterminer une implantation optimale en suivant une technique de barycentre, ou les diverses pondérations sont des prix (les centres de main d'oeuvre, tout comme les centres d'approvisionnement ou de distribution, étant ponctuellement définis) et aboutit à un découpage de l'espace, en triangle, pour les besoins de l'activité humaine. VON THUNEN (1875) en partant d'autres considérations aboutit à des cercles ; CHRISTALLER (1933) et LOSCH (1940, 1954) à des hexagones et une synthèse de ces différentes formes est proposée par W. ISARD.

(1) Voir W. ISARD (ouvrage cité), p. 222 ; HAGET (ouvrage cité) chapitres 2 et 5.

D'un point de vue opérationnel, l'utilisation de tels travaux ne peut mener qu'à constater l'existence d'une plus ou moins grande concordance entre une répartition théorique des hommes et une répartition observée. De toute façon ces différents travaux théoriques sont basés sur un certain nombre d'hypothèses restrictives qui font que la vérification de ces thèses n'a pu se faire que pour certains domaines particuliers et notamment l'agriculture.⁽¹⁾

1.14 - Evolution des données du problème

L'évolution technologique récente a modifié les données du problème, par la miniaturisation des machines et de certains produits, par une transformation de l'industrie de l'emballage (emballage perdu pour le liquide)..., accroissant ainsi la mobilité de certains facteurs et transformant les données des coûts de distribution. Elle implique par ailleurs une augmentation de la taille minimum des unités productrices et entraîne une diminution du nombre d'entreprises ; cette tendance oligopolistique s'accompagne d'une modification des comportements vis-à-vis des coûts.

1.2 - Localisation induite

A la différence de l'analyse précédente, celle de la localisation induite est dynamique et repose sur des bases théoriques très restreintes. Mais, dans la classe de préoccupations que nous avons, son intérêt opérationnel est plus grand.

L'induction d'emplois nouveaux par la demande a été étudiée dans deux optiques différentes : les multiplicateurs d'emploi et le tableau d'échanges inter-industriels. Ces deux optiques ont en commun la recherche des répercussions d'un accroissement de demande, sans se préoccuper des causes de cet accroissement.

Dans l'étude de ces concepts il faut bien distinguer l'utilisation qui en a été faite dans une optique descriptive, pour aboutir à des ratios, de celle faite dans une optique prévisionnelle qui aboutit à des multiplicateurs. On s'attachera tout particulièrement aux démarches empiriques de dissociation entre ces emplois induits et fondamentaux.

1.21 - Approche dans une optique descriptive

Elle part de la constatation que dans une zone donnée, on peut distinguer deux types d'entreprises :

(1) Travaux de Sarly concernant la localisation agricole en Israël

- celles dont la vocation est la production de biens et services à l'usage de la zone. Les emplois de ces entreprises seront dits induits ;
- et les autres dont la vocation est d'exporter tout ou partie de leur production de biens et services (laquelle est donc fixée d'une façon exogène à la zone). Les emplois de ces entreprises sont dits de base ou fondamentaux.

Il faut remarquer tout d'abord que la taille de la zone joue un rôle déterminant sur la signification du concept de base : en effet, l'agrégation de deux zones continues transforme une partie des emplois de base en emplois induits. La validité de cette analyse est donc plus grande pour des zones petites que pour des zones de grande taille où de nombreux feed-backs se produisent.

C'est l'étude de la taille des villes et de leur accroissement qui a mené à la distinction entre emplois destinés à la satisfaction des besoins spécifiques de la ville, et emplois desservant les besoins d'une zone d'influence plus ou moins vaste extérieure aux limites de la ville, ainsi qu'à leur détermination.

1.21 a - Utilisation de statistiques existantes

α) L'approche déterministe

Les premiers travaux sont l'oeuvre du géographe allemand CHRISTALLER⁽¹⁾, pour lequel l'influence qu'une ville exerce sur une certaine zone se limite à la frontière de l'influence exercée par la ville suivante pour la même activité du tertiaire. CHRISTALLER met en évidence la liaison entre la population d'une ville et la taille de l'aire d'influence desservie, pour aboutir à une typologie des unités urbaines pour le sud de l'Allemagne.

DICKINSON⁽²⁾ propose, sans y insister, quelques critères pour essayer de séparer les deux types d'emplois, mais il s'attache surtout à la recherche du processus de croissance des villes de Grande-Bretagne selon leur zone d'attraction et leur vocation spécifique (définie par une typologie en quatre groupes).

C'est semble-t-il avec AMERSFOOT (1949)⁽³⁾ que débutent les premiers travaux intéressants. Celui-ci, pour 80 villes hollandaises étudie le taux d'activité TAC (Z, B) défini pour chaque ville Z et activité B, à une date donnée (faisant donc appel à l'analyse statique). Ce taux est considéré comme la somme d'un taux induit IND (B) et d'un taux d'emplois fondamentaux FON (Z, B). Il considère que les villes n'ont sans doute

- (1) CHRISTALLER, Die Zentralen orte in Suddentschland (1933). Les thèses de CHRISTALLER sont exposées dans le livre de R.F. DICKINSON : City, Region and Regionalism (KEGAN Paul, TRENCH, TRUBNER and Co, 1947) ainsi que dans l'ouvrage de M.A. PROST "La hiérarchie des villes en fonction de leurs activités de commerce et de service (Gauthier(Villars, 1965).
- (2) DICKINSON (ouvrage cité) propose, par exemple, l'utilisation d'un indice de spécialisation défini par $(T1-T2)/T2$ où T1 est le taux d'emploi d'une catégorie dans la zone et T2 celui de l'emploi de cette catégorie au niveau national. Mais cet indice de spécialisation s'avère dépendant du découpage géographique retenu.
- (3) Cité par Gunnar ALEXANDERSSON, The industrial structure of American cities (University of Nebraska Press, 1956).

pas pour une activité donnée B, toutes des emplois fondamentaux et que donc, pour certaines d'entre elles FON (Z, B) = 0.

$$\text{TAC (Z, B) = IND (B) + FON (Z, B) \quad (\text{Equation I})$$

Aussi fixe-t-il le taux d'emplois induits (IND) (B) comme le taux minimum observé.

Cette méthode aboutit à une sous-évaluation du minimum IND (B) si la ville concernée est particulièrement desservie par une autre pour l'activité considérée. Aussi le suédois ALEXANDERSSON⁽¹⁾, dans son étude portant sur 846 villes des Etats-Unis, propose-t-il une fourchette pour IND (B), constituée par le premier et le cinquième centile, pour chaque activité. Par ailleurs, ce dernier remarque que les villes dont toutes les activités sont induites sont toutes situées dans des régions de haute densité urbaine.

MORRISSET (1958)⁽²⁾ perfectionne cette approche en effectuant une analyse par strates (coefficient valable pour un groupe de villes et un ensemble de régions).

ULLMANN et DACEY (1962)⁽³⁾ adoptent des méthodes voisines en définissant un taux minimum par strate (14 catégories d'activité, 6 tailles de villes) et perfectionnent l'approche en effectuant une régression linéaire simple (pour chacune des 14 catégories) reliant le taux minimum et la taille de ville, les écarts entre les valeurs observées et théoriques permettant de définir un indice de spécialisation⁽⁴⁾.

β) L'approche stochastique

Les méthodes déterministes proposées jusqu'alors présentent toutes une grande part d'arbitraire : pourquoi choisir le cinquième centile plutôt que le sixième, ou pourquoi prendre le minimum pour une strate donnée de villes. La stratification des villes susceptibles d'avoir des emplois fondamentaux n'est intéressante que si leur nombre est élevé, ce qui n'est le cas que de quelques pays. Partout ailleurs, on est conduit à prendre un taux minimum, ou déterminé par des techniques de quantile. On peut alors penser que IND (B) est sur évalué pour les activités à induction faible, c'est-à-dire celles pour lesquelles des emplois fondamentaux apparaissent pour à peu près toutes les villes (et donc le taux FON (Z, B) est ≠ 0, pour pratiquement tous les Z), et que le taux d'induction moyen de l'activité IND (B) est au contraire sous-évalué pour les activités à induction forte.

(1) Ouvrage cité, Une application des méthodes d'ALEXANDERSSON a été faite par G. LEGUEN (Annales de géographie, n° 734 ; Armand Colin, 1960) qui aboutit à une classification des villes selon leur activité fondamentale en 4 groupes assez nets de villes dont les fonctions sont analogues.

(2) MORRISSET, The economic structure of American Cities (P. A. R. S. A., 1958).

(3) ULLMAN and DACEY, The minimum requirements approach to the urban economic base (Land Studies in Geography, series B, Human Geography) cité par HAGGET, p. 134 (ouvrage cité).

(4) Cet indice de spécialisation est :

$$\sum_B \left[\frac{\text{TAC (B) - IND (B)}^2}{\text{IND (B)}} \right] / \left[\frac{\sum_B \left[\text{TAC (B) - \sum_B \text{IND (B)} \right]^2}{\sum_B \text{IND (B)}} \right]$$

Aussi pour parer aux fluctuations accidentelles du taux d'activité, une approche du taux d'induction fondée sur une analyse de droite de HENRY a-t-elle été développée par un bureau d'étude, la SOGREP⁽¹⁾. Ce dernier transforme l'équation (I) en :

$$\text{TAC}(Z, B) = \text{IND}(B) + U(Z, B) + \text{FON}(Z, B) \quad (\text{Equation II})$$

où $U(Z, B)$ est une variable aléatoire d'espérance mathématique nulle : $E(U(Z, B)) = 0, \forall B$

On admet, en outre :

- que $U(Z, B)$ suit une loi normale pour chaque B , ce qui semble raisonnable a priori et d'ailleurs semble vérifié pour certaines activités où l'emploi fondamental est pratiquement nul ;
- qu'il existe un nombre non négligeable de zones pour lesquelles $\text{FON}(Z, B)$ est nul, ce qui est justifié si le nombre de zones ou de villes est élevé (il est en effet peu vraisemblable que toutes soient exportatrices).

Pour estimer $\text{IND}(B)$ et l'écart-type $S(B)$ de $U(Z, B)$, on procède en raisonnant en deux temps :

- On suppose, dans une première étape, que $\text{FON}(Z, B) = 0$ pour tous les Z ; la fonction de fréquence cumulée observée⁽²⁾ est directement comparable à la fonction de fréquence cumulée théorique tirée de notre loi normale, c'est-à-dire que l'on peut comparer les quantiles observées⁽³⁾, $F(Z(J))$, aux valeurs $T(J)$ trouvées sur la table de GAUSS ; d'où l'on peut écrire, $M(S)$ et $S(B)$ étant inconnus :

$$\text{TAC}(Z(J), B) = M(B) + S(B) \cdot T(J)$$

$$\text{TAC}(Z(J+1), B) = M(B) + S(B) \cdot T(J+1)$$

D'où un système comportant deux inconnues, et autant d'équations que de villes ou de zones ; mais, compte tenu de l'hypothèse de normalité, toutes les équations sont des combinaisons linéaires de deux d'entre elles⁽⁴⁾.

En travaillant sur des distributions normales, les estimations de $\text{IND}(B)$ et de $S(B)$ faites aux extrémités sont sujettes à caution, compte tenu de l'importance des fluctuations d'échantillonnage. Aussi est-on en droit de penser que les estimations de $\text{IND}(B)$ et de $S(B)$ sont meilleures dans la résolution du système d'équations $J = 2$ et $J = 3$ que pour celles du système $J = 1$ et $J = 2$. En outre, compte tenu de la

(1) Société de Recherche et de Programmation (SOGREP) - L'induction, mesure et applications, juillet 67. Cette méthode est utilisée au niveau d'agglomération, de département et de ZAEI.

(2) Dans ce qui suit on supposera donc les valeurs $\text{FON}(Z, B)$ rangées par ordre croissant et l'on attribuera un n° d'ordre J aux zones étudiées, fonction de ce seul critère : $\text{FON}(Z, B)$ devient donc $\text{FON}(Z(J), B)$.

(3) On définit dans ce cas autant de quantiles qu'il y a d'observations.

(4) Ce qui se traduit graphiquement, en coordonnées gaussio-arithmétiques, par la droite de HENRY.

précision des valeurs de $T(J)$ et $T(J + 1)$, on observe des fluctuations relativement importantes dans les estimations de $IND(B)$ et de $S(B)$. Aussi retiendra-t-on comme estimateur, la moyenne des $J - 1$ premières estimations.

- On réintroduit, dans une seconde étape, le fait qu'un certain nombre de zones ont un emploi fondamental, ce qui se traduit par un taux d'emploi TAC (Z, B) plus élevé, puisque l'on y ajoute un FON (Z, B) non nul. On observe une constance de $IND(B)$, puis celle-ci fait un saut à partir du moment où des emplois fondamentaux apparaissent. Le taux d'induction moyen retenu correspondra à la valeur de $IND(B)$ avant cette rupture.

1.21 b - L'essai de saisie directe

Dès 1953, la dissociation entre des emplois induits et emplois fondamentaux a été tentée par enquête auprès de chaque entreprise implantée dans la zone étudiée, par ALEXANDER⁽¹⁾.

En France, des enquêtes effectuées sous la direction de CAHEN et PONSARD (1963)⁽²⁾ permettent de dissocier, pour 35 agglomérations, l'emploi urbain en plusieurs catégories basées sur la nature du marché desservi (local, régional, national, partie des emplois à vocation nationale satisfaisant les besoins locaux, partie des emplois régionaux satisfaisant les besoins locaux, emplois à fonction nationale nécessités par la croissance démographique de la ville, emplois non liés à l'économie de la ville mais implantés plus volontiers dans la ville), dépassant ainsi la dichotomie "emplois fondamentaux - emplois induits", ce qui peut être une amélioration dans l'utilisation en mode prévisionnel des résultats obtenus. A la différence des travaux précédents, aucune ventilation n'est faite selon un critère de branche ou de catégorie d'activité économique.

La difficulté de ce type d'approche réside dans la détermination pour les emplois mixtes de la part satisfaisant aux emplois locaux, avec le problème complémentaire des industries situées en amont des entreprises de base⁽³⁾.

1.21 c - Limite de ces approches

Ces différentes méthodes permettent toutes de calculer un emploi induit et un emploi fondamental (par différence entre l'emploi total et l'emploi induit ; mais la définition de cet emploi n'est guère satisfaisante, s'effectuant soit à partir d'indicateurs empiriquement déterminés, soit par enquête avec les problèmes de définition que l'on a vus.

(1) J.W.ALEXANDER, An economic Base Study of Madison, Wisconsin (University of Wisconsin Press, 1953) cité par W.ISARD (ouvrage cité).

(2) CAHEN et PONSARD, La répartition fonctionnelle de la population des villes et son utilisation pour la détermination des multiplicateurs d'emplois (Ministère de la construction, CEES, 1963).

(3) Voir les critiques faites sur la définition des emplois de base R.B. ANDREWS (Land Economics, vol. 29-31) repris par ISARD (ouvrage cité p. 194 et sq.).

Il nous semble que la vocation de ces méthodes reste surtout descriptive ; en effet :

- elles rendent impossible l'explication des distorsions spatiales du taux réel d'induction à une date donnée puisque l'on ne retient que des valeurs ponctuelles valables pour toutes les unités observées (minimum, fourchette ou espérance mathématique) ;
- l'application de cette méthode à deux dates espacées pourrait permettre l'explication de l'évolution de ce taux d'induction⁽¹⁾, mais celle-ci est sans doute vouée à l'échec car elle peut être menée :
 - soit sur le taux d'induction moyen, c'est-à-dire au niveau global, mais les disparités spatiales des facteurs explicatifs retenus devraient moduler les taux réels d'induction sans qu'il soit possible de procéder à une vérification ;
 - soit à partir du taux de l'activité concernée, ce qui n'est envisageable que pour les activités que l'on peut considérer comme presque totalement induites mais pour lesquelles la méthode proposée ne présente normalement aucun intérêt.

Par contre, ces méthodes semblent extrêmement intéressantes pour dresser une typologie des entités observées en les caractérisant par leurs emplois fondamentaux, c'est d'ailleurs l'utilisation principale qui en a été faite.

1.22 - Approche dans une optique prévisionnelle

L'implantation dans une zone de nouveaux emplois fondamentaux peut tout d'abord avoir comme effet un accroissement d'emplois dans des branches techniquement situées en amont. Cette complémentarité intersectorielle sera étudiée ultérieurement dans le cadre du tableau d'échanges inter-industriels.

Mais à côté de cette induction à caractère technique, les chercheurs qui ont mis en évidence l'existence d'emplois induits, ont essayé de mesurer l'impact d'une augmentation d'emplois fondamentaux sur le volume d'emplois induits par le biais d'un accroissement de population (nouveaux travailleurs et leurs familles), et de la satisfaction locale de certains besoins des entreprises "fondamentales", fournitures de bureaux, etc.).

Si l'on raisonne globalement et que l'on admette que la part H (Z) de l'emploi induit dans l'emploi total reste stable à court terme (ce qui repose sur plusieurs hypothèses que l'on envisagera ci-après), on peut écrire, PA (Z, BASE) étant le volume d'emploi de base dans la zone :

$$PA (Z) = PA (Z, BASE) + H (Z) \cdot PA (Z)$$

(1) Encore faut-il remarquer que pour la méthode de la SOGREP, l'emploi fondamental calculé est biaisé par le jeu de l'écart inconnu U (Z, I) entre le taux d'induction moyen et le taux d'induction réel, écart qui est reporté sur l'emploi fondamental.

$$\text{d'où : } PA(Z) = \frac{PA(Z, \text{BASE})}{1 - H(Z)} = M(Z) \cdot PA(Z, \text{BASE})$$

Ce coefficient $M(Z)$ peut jouer le rôle d'un multiplicateur de type keynésien dans une utilisation de cette relation en variation. En ajoutant des hypothèses de structure familiale, on obtient un multiplicateur de population. Ces formules de multiplicateurs sont à moduler en fonction des hypothèses concernant l'état du marché du travail (plein emploi, chômage déguisé, chômage déclaré)⁽¹⁾.

Lorsque l'on ne fait pas appel à l'observation directe⁽²⁾, on détermine la valeur de ces coefficients, soit à l'aide des données d'une seule année en assimilant coefficient moyen et coefficients marginaux (analyse statique), soit en comparant les données entre deux dates (analyse statique comparative).

La stabilité du multiplicateur repose sur plusieurs hypothèses :

- pas d'accroissement possible de productivité dans les emplois induits ;
- stabilité des prix relatifs dans la zone et comparativement aux autres régions ou, autrement dit, les avantages comparatifs de la région sont stables (pour éviter des substitutions par importations nouvelles de biens et services (fuites au sens de la théorie économique du commerce international).

Ce mécanisme passe par l'intermédiaire de l'accroissement global des dépenses et donc, des revenus de la zone, ce qui néglige :

- l'influence du type d'emploi créé (liaison catégorie socio-professionnelle - type de dépense et de structure familiale) ;
- des éléments importants intervenant dans la détermination des dépenses dans la zone (construction de logements, investissements industriels, dépenses gouvernementales et en particulier garnisons) ;
- des décalages (variables dans le temps) entre revenus et dépenses et entre arrivées de nouveaux travailleurs et celles de leur famille.

Pour toutes ces raisons, l'emploi de tels multiplicateurs est très délicat, valable en tout état de cause dans le seul court terme, et pour une période de la date des enquêtes ou recensements ayant servi aux calculs.

(1) Voir PAELINCK, *Theoretical Spatial Economics* (cours polycopiés de la Netherlands School of Economics, 1971), CAHEN et PONSARD (ouvrage cité).

(2) Les travaux de CAHEN et PONSARD retiennent 8 multiplicateurs statiques dont la vocation est avant tout descriptive et 3 multiplicateurs dynamiques (en fonction de 3 hypothèses d'emploi), caractérisant chacun un sous-ensemble différent de villes (8 retenues).

Toutefois, des travaux effectués conjointement par la SOGREP et le Commissariat Général du Plan (CGP) (voir méthode exposée au 1.21 β) ont été utilisés à titre indicatif dans la préparation du 6ème Plan. Le découpage utilisé est le découpage régional et les coefficients multiplicateurs étaient identiques pour l'ensemble des régions ($M(Z) = M(Z'), \forall Z'$). Les résultats obtenus différaient de ceux obtenus par la méthode dite de "l'esquisse" développée par le CGP et l'INSEE et qui ont été utilisés comme données d'encadrement pour le 6ème Plan, mais cette confrontation a permis de souligner l'importance de certaines hypothèses utilisées dans l'une ou l'autre approche.

1.23 - Le TEI interzonal

Les travaux de LEONTIEF sur la matrice d'échanges inter-industriels ont connu aussitôt après la seconde guerre mondiale des applications régionales⁽¹⁾.

A vrai dire, le TEI (Tableau d'Echange Inter-industriel) est conçu à partir de flux appréhendés monétairement. Mais des hommes sont nécessaires pour fabriquer ces flux et des ratios techniques (homme/production) permettent de mettre en évidence l'impact de l'expansion d'une branche donnée dans une zone donnée et donc la population active nécessaire. C'est la raison pour laquelle ce type d'approche est envisagé ici.

α) Les divers types de TEI interzonaux

Trois conceptions de TEI zonal peuvent être avancées :

- Dans la première, la zone est envisagée comme entité distincte et une simple transposition de la conception du TEI national est faite. Dans cette optique, les autres zones du pays sont assimilées à l'étranger, ou regroupées dans un poste unique "autres zones", généralement sans mention de la branche d'origine ou de destination. Il est donc nécessaire d'établir autant de TEI qu'il existe de zones dans le pays. Cette conception ne permet pas en fait d'études interzonales d'ensemble, mais est plutôt destinée à mesurer l'impact de l'implantation sur l'économie d'une zone (2).
- La seconde conception distingue l'origine et la destination géographique de chaque flux entre deux branches⁽³⁾. Il s'agit alors d'un véritable TEI interzonal unique pour le pays. Si (Card B) est le nombre de branches et (Card Z) le nombre de zones, la taille du tableau est :

$$\left[(\text{Card B}) \times (\text{Card Z}) \right]^2.$$

(1) Voir MEYER (article cité) et TIEBOULT : Regional and Interregional Input-Output Models : an appraisal 1957 (in Regional analysis de Needleman) Voir PAELINCK, ouvrage cité.

(2) Divers travaux d'ISARD en 1952-1953 s'inspirent de cette conception.

(3) Dite encore du FI BABY SAM (Full Information Baby Social Account Matrix). Elle a été utilisée par MOSE en 55 pour 3 régions et 11 branches (American Economic Review, décembre 1955) ; mais la formulation théorique remonte à 1950 (L. METZLER : A multiple Region Theory of Income and Trade in Econometrica, vol. 18, 1950).

- Enfin, diverses conceptions intermédiaires entre les deux premières sont possibles. Celle utilisée par PAELINCK⁽¹⁾ est la suivante : les consommations intermédiaires de la production des diverses branches d'une zone sont connues sans mention de leur origine (ce qui donne lieu à autant de sous-matrices bloc-diagonales de dimension Card (B) qu'il existe de zones) ; et les exportations (= consommations intermédiaires et finales) d'une zone, à destination d'une autre, sont connues sans ventilation par branche (ce qui donne lieu par couple de zones différentes à autant de matrices diagonales de dimension Card (B)). L'économie d'informations réalisée s'élève donc à :

$$\left[((\text{Card B})^2 - \text{Card B}) \cdot \text{Card Z} \right]$$

β) Utilisation des TEI interzonaux

La seconde conception du TEI interzonal de même que la troisième (L.I. BABY SAM) au prix d'hypothèses supplémentaires, permettent l'utilisation classique des TEI et en particulier de mesurer les répercussions d'une hausse de prix (régionale, sectorielle ou les deux) et de calculer l'impact d'une nouvelle implantation dans une zone.

γ) Limite des TEI interzonaux

Un certain nombre de critiques ont été formulées en ce qui concerne l'utilisation à un niveau national des TEI, sur la stabilité des coefficients en particulier.

Cette stabilité, à un niveau zonal assez fin, n'est pas sans poser quelques problèmes supplémentaires⁽²⁾ :

- représentativité incomplète des différentes activités d'une branche pour la zone ;
- stabilité des courants commerciaux entre zones postulée dans le temps (du moins pour les deux dernières conceptions de TEI) ;
- non prise en considération des modifications spatiales possibles de productivité (lesquelles modifient la structure relative des coefficients)⁽³⁾ ;
- non prise en considération de la distorsion des prix relatifs (et donc des élasticités de demande) ;
- existence de capacité de production excédentaire (capitaux et travail) et non saturation des capacités de transport entre zones.

(1) Celle du L.I. BABY SAM (Limited information). Elle est utilisée dans un modèle opérationnel de PAELINCK appliqué aux Pays-Bas.

(2) Certains sont évoqués par TIEBULT, Regional and Interregional Input-Output Models : an Appraisal (1957), repris dans le Regional analysis de NEEDLEMAN (op. cit).

(3) Dans la mesure où celle-ci se traduit par des distorsions spatiales, dans les matrices branche-qualification. On peut noter, d'après les données françaises, que non seulement l'évolution mais aussi les structures de départ, varient de région à région. Des études de la matrice branche-qualification en emplois, ont été menées par certains économistes étrangers, et sont cités dans le livre de BABAËU et HAYEC (Problèmes de planification) ; et KOMIYA et UCHIDA ont montré qu'une telle matrice n'était pas stable dans le temps, pour le Japon ; HOLLISTER, dans la revue internationale du travail (Les aspects économiques de la prévision en matière de main-d'oeuvre) souligne aussi cette difficulté.

Mais l'écueil majeur en France pour ce type d'approche est d'ordre statistique, les données souvent n'existent pas à un niveau régional, et même si elles existaient pour une année donnée, comme des mises à jour fréquentes sont nécessaires, le coût d'actualisation de cette information serait prohibitif, du moins avec les structures de collecte de l'information en place

1.3 - Localisation autonome

Les théories d'induction de l'emploi supposent donnée la variation d'une catégorie d'emploi considérée comme "motrice" dans une zone. Il semble donc intéressant de rechercher les raisons qui ont présidé à l'implantation d'une entreprise dans telle zone plutôt que dans telle autre zone.

En fait, on trouve des groupes d'industries qui, face à ces problèmes de localisation, ont le même comportement ; il est donc inutile de chercher à reproduire le processus décisionnel de chaque entité économique, et il vaut mieux s'efforcer de décrire le comportement d'un ensemble, homogène, de centres ponctuels de décisions, avec suffisamment de précision.

C'est dans cette optique que certains modèles américains ont cherché à établir une série de comportements types et à trouver une méthode empirique permettant de les simuler. Il nous a paru intéressant de présenter trois des principaux modèles multi-régionaux récents qui présentent la particularité (assez rare en économie régionale) d'avoir été opérationnels. Toutefois, nous n'avons pu avoir connaissance des enseignements tirés de ces applications et de la réelle plus value apportée par ces modèles.

Ces comportements peuvent s'analyser comme la prise en compte d'un certain nombre de facteurs de localisation, et le choix d'une échelle pour ceux-ci. La simulation permet, par l'intermédiaire d'une note, de choisir la zone où s'établira le nouvel établissement. Diverses méthodes empiriques ont été utilisées, elles n'ont pas de fondement théorique à proprement parler mais leur intérêt dans le cadre d'un modèle de simulation est indéniable.

1.31 - Sous-modèle d'implantation industrielle du Bass⁽¹⁾

Il s'agit d'un sous-modèle de simulation d'implantations de plusieurs activités industrielles, déterminées exogènement à ce sous-modèle, parmi les 742 zones de la Baie de San Francisco, en régime non interventionniste. Ces activités industrielles, au nombre de 20 (Code d'activités à deux chiffres) sont regroupées en

(1) Michael GOLBERG - Industrial Location model for the San Francisco Bay Area.

six groupes (notés B). Ces groupes sont définis par la nécessité absolue de trouver certains facteurs d'implantation, notés FAC (K, B) en nombre N (B) (on a donc $K = 1, \dots, N(B)$) ; deux groupes ne peuvent faire appel aux mêmes facteurs. Ces facteurs d'implantation sont au nombre de 8 (donc $N(B) \leq 8$) et qui sont : proximité de main d'oeuvre qualifiée, accès au rail, aux routes à grande circulation, aux transports en eau profonde, disponibilité en terrains, en eau, en moyens d'élimination des déchets industriels, impôts locaux. Pour des raisons de données, les facteurs socio-politiques n'ont pas été retenus.

On détermine avant toute itération le nombre d'unités d'emploi à affecter pour chaque branche. Cette unité varie en fonction de l'activité industrielle envisagée et peut concerner un seul emploi ou un groupe d'emplois pour tenir compte des seuils.

Ce modèle s'intéresse à la localisation de l'accroissement d'une activité industrielle, et sélectionne tout d'abord le sous-ensemble de 742 zones qui possède tous les facteurs essentiels au groupe auquel appartient l'activité industrielle. C'est sur ce sous-ensemble E (B) qui dépasse rarement une quinzaine de zones que l'on travaille et pour chaque facteur retenu, dans chaque zone du sous-ensemble, on définit l'indicateur suivant (qui permet de s'affranchir des problèmes d'unité) :

$$I(Z, K, B) = \frac{\text{FAC}(Z, K, B) - \text{minimum FAC}(Z', K, B)}{\text{maximum FAC}(Z', K, B) - \text{minimum FAC}(Z', K, B)} \quad \text{avec } Z' \in E(B)$$

Puis, pour chaque zone, on construit un coefficient d'attraction, lequel est la moyenne des I (Z, K, B).

$$\text{ATON}(Z, B) = \frac{1}{N(B)} \sum_K I(Z, K, B)$$

Puis on attribue une unité d'emploi de l'activité industrielle à la zone ayant la plus forte attraction, et l'on recalcule en tenant compte, si besoin est, de l'utilisation des facteurs, le nouveau coefficient d'attraction de la zone ; puis on recommence à attribuer une unité d'emploi. Ce processus itératif se poursuit jusqu'à épuisement des implantations à effectuer.

Ce sous-modèle a l'avantage de permettre d'aboutir à une solution unique pour l'implantation de plusieurs activités industrielles, car il est très peu vraisemblable qu'une zone possède les facteurs nécessaires à deux groupes d'activités industrielles données, et de surcroît possède pour ces deux groupes le meilleur coefficient d'attraction.

Par contre, il présente l'inconvénient de considérer sur un même plan les facteurs indispensables à un groupe d'activité donné. Ceci s'explique par la difficulté d'établir des coefficients de pondération alors que la mesure des facteurs est déjà loin d'être précise. On peut, par ailleurs, regretter que des difficultés d'ordre statistique aient empêché de prendre en considération un plus grand nombre de facteurs.

1.32 - Le sous-modèle d'allocation industriel du modèle du Bassin de la Susquehanna⁽¹⁾

Il s'agit d'un modèle de simulation à moyen terme sur 8 zones. La croissance de chacune des 4 branches (B) de base prises en considération est une modulation de la croissance de la branche ; Calculée comme une somme des taux de croissance nationaux des différentes activités constitutives de la branche, pondérée par la distribution de ces activités en début de période pour l'ensemble des 8 zones. On module ensuite ce taux par zone et branche.

Cette modulation du coût de croissance de la branche, qui joue le rôle d'une attractivité, est fonction d'une élasticité-coût : ELAS (B, ΔT) ; et d'un indicateur de coût relatif à la zone : CREL (B, Z, ΔT). Cette attraction peut s'inscrire (la période étant prise implicitement en compte) :

$$ATON (Z, B) = ELAS (B, \Delta T) \cdot (CREL (Z, B, \Delta T) - 1)$$

Le coût relatif est une fonction linéaire de 3 indicateurs de coûts (relatifs à la zone par rapport à la zone par rapport à la valeur nationale) des composantes du prix de revient, pondérés par la part constatée au niveau national du facteur correspondant dans le prix de revient unitaire de la branche. Ces facteurs pris en compte sont le niveau de salaire, le coût d'accès aux matières premières et le coût d'accès aux clients. La détermination de ces trois facteurs est endogénéisée dans le modèle.

L'influence de ces avantages comparatifs ne joue qu'en fonction de la sensibilité de la croissance d'une branche dans une zone à cet argument de coût relatif ; ce qui est mesuré par une élasticité⁽²⁾. Cette élasticité est calculée une fois pour toutes, à partir d'études empiriques : en particulier, une étude reliant le logarithme de l'investissement net de la période précédente (représentant la croissance de la branche) pour un certain nombre d'états représentatifs, à la marge bénéficiaire brute de la période précédente (représentant les conditions de coûts).

(1) System Simulation for Regional Analysis : An application to River Basin Planning, par HAMILTON et al. (MIT Press, 1969).

(2) Ce qui explique d'ailleurs que le modèle raisonne en variation de coûts relatifs. Le produit de cette variation par l'élasticité donne la variation du taux de croissance de la branche, et donc de sa population active.

1.33 - Le sous-modèle d'implantation industriel du AAM⁽¹⁾

Ce modèle est un modèle de simulation par périodes de cinq ans, pour 192 zones. Il comporte, entre autres, un sous-modèle d'implantation industriel, pour deux groupes d'activités industrielles, et un sous-modèle d'implantation d'activité du tertiaire pour 13 catégories d'activité. Examinons le sous-modèle d'implantation industrielle :

Les 19 activités industrielles sont regroupées en deux classes, d'après un critère de dispersion (coefficient de variation).

La variation entre T et T + 1 de la population active de l'un des groupes $\Delta PA (Z, B)$ est la résultante du jeu d'une variation relative identique pour la zone à celle prévue pour "l'ensemble du territoire", et de transferts d'activités de zone à zone.

Ces transferts sont fonction des populations actives concernées au début de la période, des terrains industrialisables en début de période, TI (Z, T) et d'un indice d'attraction, ATON (Z, B).

Pour définir l'attraction d'une zone, les auteurs du modèle sont partis de 60 facteurs explicatifs de la localisation et ont recherché ceux qui semblaient les plus significatifs (du point de vue de la corrélation) et affectaient le moins la stabilité des paramètres (phénomènes de colinéarité). L'analyse en composantes principales, utilisée pour cette sélection, a été relativement décevante.

Finalement, 6 facteurs explicatifs ont été retenus : pourcentage de ménages à hauts revenus, pourcentage d'emplois industriels de la zone, pourcentage de terrains disponibles, indice d'accessibilité à la zone, densité de population rapportée aux seuls terrains à vocation résidentielle, densité de voies navigables.

Ce coefficient d'attraction est une somme pondérée des facteurs de localisation FAC (Z, K, B) :

$$ATON (Z, B) = \sum_{K=1}^6 A (K, B) \cdot FAC (Z, K, B)$$

(1) AAM = Activities Allocation Model - Voir the construction of an Urban Growth Model par SEIDMAN, DELAWARE, Valley Regional Planning Commission, Technical Planning Commission, Technical Supplement (1969).

Un transfert d'activité de Z vers Z' ne se produit que si $ATON(Z) > ATON(Z')$, autrement dit la zone Z ne peut être qu'émettrice ou réceptrice par rapport à Z'. Posons :

$$C(Z, Z', B) = 1 \quad \text{si} \quad ATON(Z, B) > ATON(Z', B)$$

$$C(Z, Z', B) = 0 \quad \text{si} \quad ATON(Z, B) < ATON(Z', B)$$

Le transfert total d'activité de B au profit (ou au détriment de Z) est donc :

$$\sum_{Z'} \left[C(Z, Z', B) \cdot PA(Z, B, T) \cdot TI(Z, T) - C(Z, Z', B) \cdot PA(Z', B, T) \cdot TI(Z', T) \right] \left[ATON(Z, B) - ATON(Z', B) \right]$$

Etant donné que les $C(Z, Z', B)$ dépendent des coefficients d'attraction, il n'est pas possible d'estimer directement les coefficients de régression $A(K, B)$ par la méthode des régressions linéaires et il a fallu faire appel aux techniques de régression non linéaire⁽¹⁾.

SECTION II - MODIFICATIONS de la REPARTITION des HOMMES LIEES aux DECISIONS INDIVIDUELLES

Dans l'ensemble des études et théories envisagées dans la section I, une modification de la répartition des hommes s'imposait pour des raisons d'emploi, et implicitement l'ajustement de l'offre d'emploi à la demande s'effectuait automatiquement et sans problème. En fait, il n'en est pas toujours ainsi et l'échec d'un certain nombre d'études basées exclusivement sur des considérations d'offres d'emploi s'explique, en grande partie, par une demande qui ne s'ajuste pas à une modification de l'offre, ou même par une demande dont les variations ne s'expliquent pas par des considérations relatives à l'offre d'emplois de la zone.

Quelles que soient les raisons d'une migration, un candidat à la migration ne prendra de décision définitive qu'après avoir trouvé une place stable (à moins qu'il ne soit à la charge d'un autre migrant). Que cet emploi lui soit assuré avant son départ ou non importe peu. On constate donc des modifications de la répartition des demandes d'emploi. En fait, c'est plutôt au niveau du migrant en tant qu'individu et non en tant que travailleur que le raisonnement est effectué.

(1) Voir chapitre 3 page 3.16.

Le phénomène migratoire étant très complexe à appréhender, on a cherché très tôt à retracer, par des modèles descriptifs, les migrations passées. Ce n'est que depuis peu que des modèles prévisionnels ont été créés. Les modèles descriptifs ayant tout naturellement servi de base aux modèles prévisionnels, il a semblé utile d'en donner un rapide survol.

2.1 - Modèles descriptifs

L'analyse des phénomènes migratoires est ancienne. Dès 1885, RAVEINSTEIN⁽¹⁾ en Grande-Bretagne constate que la distance joue un rôle fondamental, mais ce n'est que depuis une trentaine d'années que des études systématiques ont été entreprises en ce domaine.

Deux approches des phénomènes sont concevables : il est possible d'étudier les flux migratoires pour chaque couple de zones possibles. L'étude de ces migrations s'effectue alors par l'intermédiaire d'une matrice de migrations. Il est, par ailleurs, possible de s'intéresser aux flux migratoires de chaque zone, sans distinction de provenance ou de destination. L'étude porte alors sur les marges de la matrice (dans cette seconde perspective, il est possible de ne s'intéresser qu'au solde). Ces deux démarches présentent de nombreuses analogies.

2.11 - Les modèles en terme de flux

A l'origine, ces modèles sont basés sur l'observation de RAVENSTEIN, et considèrent que la variable explicative est la distance. Cette notion de distance fut par la suite affinée pour devenir une "distance sociale". Et ce n'est que récemment que d'autres facteurs explicatifs sont venus compléter cet arsenal de modèles descriptifs en termes de flux.

2.11 a - Les modèles d'attraction mécanique des migrations

α) Explication par la distance et les populations des zones

YOUNG en 1928⁽²⁾ postule que le flux migratoire d'une zone vers une autre varie avec la force d'attraction de cette dernière et inversement avec le carré de la distance.

(1) RAVENSTEIN - The law of Migration, Journal of the Royal Statistical Society, 1885, vol. 48.

(2) YOUNG - The movement of farm population, Cornell Agricultural Experiment Station Bulletin, 1928, cité par ISARD (ouvrage cité) p. 62, REILLY à la même époque propose un modèle du même type mais en s'intéressant non pas aux migrations, mais à la notion de clientèle potentielle : W. J. REILLY, Methods for study of Retail Relationships (University of Texas, Bulletin n° 2 944 nov. 1929).

ZIPF⁽¹⁾ affine l'analyse en trouvant une expression pour cette force d'attraction, laquelle est fonction des emplois vacants de la zone de réception et de l'importance de la zone d'émission.

On désigne par IMINT (I, J) le flux migratoire des personnes de I vers J entre T et T + h ; par D (I, J), la distance entre I et J ; et par PT (I), la population de la zone I à la date T.

ZIPF formule les hypothèses suivantes :

- le flux migratoire entre I et J est d'autant plus important que la population de I est forte ;
- les emplois vacants dans J sont proportionnels à sa population ;
- le flux migratoire dépend d'une façon inversement proportionnelle de la distance entre I et J ;
- les échanges entre I et J sont donc indépendants des autres échanges.

Ces hypothèses se traduisent par la formule suivante :

$$\text{IMINT (I, J)} = h \cdot \frac{\text{PT (I)} \cdot \text{PT (J)}}{\text{D (I, J)}^n}$$

ZIPF⁽²⁾ a testé ultérieurement, avec succès, un modèle légèrement différent, du type :

$$\text{IMINT (I, J)} = h \cdot \left[\frac{\text{PT (I)} \cdot \text{PT (J)}}{\text{D (I, J)}} \right]^k$$

sur les mouvements de voyageurs entre 29 villes.

Des travaux empiriques ont montré qu'en Suède⁽³⁾ n variait, suivant les périodes, entre 0,4 et 3,3 avec des différences sensibles selon que les zones considérées sont de type rural ou urbain (dans ce dernier cas, n est plus faible).

(1) ZIPF - The P1/P2/D Hypothesis on the inter-city movement of persons, American Sociology Review, déc. 1946 - STEWARD à la même époque a développé ces formules (avec n = 2), en s'appuyant sur des analogies avec la physique Newtonnienne, pour jeter la base des modèles gravitaires.

(2) ZIPF - Human behavior and the principle of least effort (Addison - Wesley Press, 1949).

(3) HAGERSTRAND, Migration and Area, Lund Studie in Geography, 1957, p. 116.

En France, les travaux de D. COURGEAU⁽¹⁾ sur 12 départements n'ayant pas de fort centre d'attraction, et sur 4 recensements, fournissent des ajustements corrects avec n sensiblement égal à 2, et montrent un déplacement, constant dans le temps, des courbes observées.

Des travaux effectués par la SEMA⁽²⁾, et qui seront analysés ultérieurement, confirment la croissance de n en fonction du degré de ruralité des zones.

β) Explication par la distance sociale et les postes offerts

STOUFFER⁽³⁾ fait l'hypothèse que le flux migratoire est proportionnel au nombre de postes offerts en J et inversement proportionnel au nombre de postes offerts entre les zones I et J (c'est-à-dire la distance sociale), les zones J étant des couronnes centrées sur I. Ce concept de poste offert est mesuré par le nombre de migrants originaires de I et présents en J.

Ce modèle a été testé par FOLGER⁽⁴⁾, avec un pouvoir explicatif supérieur à celui de ZIPF.

Une application de ce modèle a été effectuée par D. COURGEAU⁽⁵⁾, pour quelques départements, et montre que les résultats fournis n'apportent pas plus de précision que ceux donnés par ZIPF.

Cette approche "concentrique" n'est pas directionnelle ; aussi STOUFFER a-t-il proposé⁽⁶⁾ un second modèle dans lequel le flux migratoire de I vers J est proportionnel aux immigrations dans J et aux émigrations de I, et inversement proportionnel, d'une part, au nombre de postes entre I et J (dans un cercle Ω tel que les barycentres des zones I et J se trouvent sur le même diamètre et que les zones I et J soient largement englobées, les postes offerts en J étant exclus), d'autre part, au nombre de migrants concurrents (dans un cercle Ω' inclu dans Ω) pour les postes offerts en J.

Nous avons alors la détermination suivante :

$$IMINT(I, J) = h \frac{\sum_J IMINT(I, J) \cdot \sum_I IMINT(I, J)}{\sum_{K \subset \Omega} IMINT(I, K \neq J) \cdot \sum_{K \subset \Omega'} IMINT(K \neq I, J)}$$

(1) D. COURGEAU - Les champs migratoires en France, cahier de l'INED n° 58, 1970.

(2) Modèles explicatifs des migrations intérieures en France, SEMA (1967) tome II p. 5.

(3) STOUFFER - Intervening opportunities : a theory relating mobility and distance, American Sociological Review, vol. 5, 1940.

(4) FOLGER - Some aspects of migrations in the Tennessee Valley, American Sociological Review, vol. 18, 1953.

(5) D. COURGEAU : op. cit.

(6) STOUFFER : Journal of regional sciences, vol. 2 (1960).

γ) Explication par les contacts existants et les postes offerts

HAGERSTRAND⁽¹⁾ pense que les migrants antérieurs forment la source principale d'information des migrants suivants, aussi propose-t-il⁽²⁾ un modèle où le flux migratoire prévu entre I et J est proportionnel au nombre de contacts entre I et J, lequel est appréhendé par le nombre de migrants de I vers J pour les n dernières années (n = 15 dans son étude) ; proportionnel aussi, comme dans le modèle de STOUFFER, au nombre de postes offerts dans la zone J, mesuré par le nombre total de migrants prévu pour cette période, et inversement proportionnel à la population de J.

Le modèle s'écrit alors :

$$IMINT (I, J, (T, T + 1)) = a \frac{\left[\sum_{K=0}^{T-m} IMINT (I, J (T - k - 1, T - K)) \right] \cdot \left[\sum_I IMINT (I, J (T, T + 1)) \right]}{PT (J, T)}$$

Ce modèle permet donc d'éclater selon l'origine, un flux d'immigrants prévu par tout autre méthode de prévision (dans le cas contraire, on retrouverait l'objection de "l'explication de la partie par le tout", déjà faite à STOUFFER).

2.11 b - Les modèles d'attraction socio-économique

- α) L'effort a porté jusqu'à présent sur les échanges migratoires et pas tellement sur leurs raisons. Or, la connaissance de ces motivations est indispensable à la prévision. HAGERSTRAND pense que deux types de migrants existent, dont les causes diffèrent : les migrations passives sont celles qui dépendent entièrement des migrations antérieures déjà effectuées ; les migrations actives sont indépendantes des migrations antérieures. L'accroissement de migrants présents dans une zone, durant un intervalle de temps donné, est égal à la somme des migrations actives et passives, auxquelles il faut retrancher les décès de migrants et les doubles migrants.

Il est à noter, à ce propos, que ce phénomène de migrations multiples est loin d'être négligeable. Les observations effectuées ne coïncident pas. Une enquête effectuée par POURCHER a établi que 33 % des migrants inter-communaux arrivés à Paris ont, depuis l'âge de 14 ans, migré plus d'une fois (en moyenne, 2,3 fois) avant d'arriver à Paris.

(1) Opus cit.

(2) Opus cit.

En France, l'enquête de l'INED de 1967 donne comme estimation du coefficient de migration multiple (nombre de migrations de rang supérieur ou égal à 2, rapporté au nombre de migrations tous rangs confondus) pour la période 1962 - 1967 et le découpage départemental, une valeur d'environ 0,31. L'étude de ces données effectuée par D. GOURGEAU⁽¹⁾ permet de mettre en évidence trois points importants (non infirmés d'ailleurs par une étude complémentaire portant sur des données suédoises et américaines) : tout d'abord, la probabilité de migration annuelle d'un individu quelconque est stable sur la période étudiée et dépend du découpage retenu ; ensuite, la probabilité qu'un individu d'une génération donnée effectue une migration de rang $n + 1$ dépend peu de n et de la génération considérée, mais varie selon le découpage retenu ; enfin, la probabilité qu'un individu d'une génération donnée effectue une migration de rang $n + 1$ à la date $t + k$, alors que la migration de rang n a été effectuée à la date t , est indépendante de k et du découpage et peu dépendante de l'âge. La connaissance de la stabilité de ces comportements devrait permettre aux modèles prévisionnels français de s'affranchir de la contrainte d'une amplitude de période imposée par les données du passé (qui auraient servi à déterminer les paramètres du modèle migratoire).

L'importance des migrations passives semble confirmée en France par les travaux de POURCHER⁽²⁾ (qui montre que 60 % des migrants à Paris connaissent quelqu'un capable de les aider et 60 % aussi estimaient avoir un emploi assuré à l'avance) et de BELTRAMONE⁽³⁾. Ce type de migration s'explique par la notion de contacts existants (HAGERSTRAND), ou plus simplement de structure d'accueil ; mais la raison profonde pour laquelle cette permanence du courant migratoire s'est instaurée, n'est pas pour autant donnée.

Le problème est donc de savoir ce qui a donné naissance à des migrations actives ; autrement dit, il faut rechercher de quoi dépend l'attraction d'une zone ?

β) Modèle de SOMERMEIJER

SOMERMEIJER, en 1961, dans son étude sur les migrations hollandaises de 1948 à 1957, essaie de mesurer l'attraction d'une zone, $ATON(Z)$, et teste un modèle de ZIPF amélioré :

$$IMINT(I, J) = \left[\frac{1}{2} K + C \cdot (ATON(J) - ATON(I)) \right] \cdot \frac{PT(I) \cdot PT(J)}{D(I, J)^n}$$

$$IMINT(J, I) = \left[\frac{1}{2} K - C \cdot (ATON(J) - ATON(I)) \right] \cdot \frac{PT(I) \cdot PT(J)}{D(I, J)^n}$$

$ATON(Z)$ est, pour SOMERMEIJER, une fonction du revenu par tête de la zone, du taux de croissance, du degré d'urbanisation et des possibilités de loisirs.

(1) D. COURGEAU : Cahier de l'INED, sous presse.

(2) POURCHER : Cahier de l'INED n° 43 : Le peuplement de Paris (1964).

(3) BELTRAMONE : La mobilité géographique d'une population (Gauthier - Villars).

γ) Premier modèle LOWRY⁽¹⁾

Aux USA, des travaux effectués sur le recensement de 1960 et sur un échantillon de 800 flux migratoires tirés des 8 010 flux interzonaux entre les 90 zones les plus peuplées, se sont inspirés du modèle de ZIPF (avec $n = 1$), l'adaptant à la population active non agricole (PA (Z, NA)), l'ont amélioré par introduction d'un coefficient d'attraction, fonction, d'une part du taux d'activité non agricole des zones, TAC (NA, Z) en faisant l'hypothèse que les gens se déplacent des marchés de l'emploi lâches vers les marchés de l'emploi tendus ; d'autre part, du salaire, SAL (Z), en faisant l'hypothèse que les gens se déplacent des zones de bas salaires vers les zones de hauts salaires.

Ce modèle se présente sous la forme :

$$IMINT (I, J) = K \frac{TAC (NA, I)}{TAC (NA, J)} \cdot \frac{SAL (J)}{SAL (I)} \cdot \frac{PA (NA, I) \cdot PA (NA, J)}{D (I, J)}$$

Les paramètres sont estimés par régression multiple après transformation en logarithme.

Dans cette formulation, ni les salaires, ni le taux d'activité de la zone émettrice ne semblent significatifs.

Lorsqu'on se limite aux migrations de moins de 60 ans, et si l'on observe la population non agricole civile, le salaire de la zone d'arrivée devient significatif, mais apparemment les conditions du marché de l'emploi de la zone d'origine jouent peu.

Il semble donc, sur l'exemplaire américain, qu'il y ait assymétrie entre les facteurs déterminant le départ et ceux déterminant l'arrivée ; les individus migrent de zones en état de crise, aussi bien que de zones prospères.

Quelques remarques sont à faire en ce qui concerne l'hypothèse du salaire, facteur explicatif des migrations. Une analyse marginaliste de ce phénomène en donne l'interprétation suivante : lorsque le niveau de salaire d'une zone J est supérieur à celui de la zone I, une partie de la population active de I se déplace en J ; les demandes d'emploi en J augmentant, le salaire de J diminue et inversement dans la zone I. Pour qu'un tel mécanisme puisse jouer, il faudrait que les individus désirent maximiser leur revenu, que leurs connaissances des possibilités d'emploi soient parfaites, que les travailleurs soient en grand nombre et d'égale qualification, et enfin qu'aucun obstacle d'ordre économique ou social ne s'oppose à leur migration. Il est évident que toutes ces conditions sont rarement réunies, il s'en suit que ce mécanisme ne joue en pratique que partiellement.

(1) LOWRY - Migration and Metropolitan Growth, 1966.

δ) Modèle de la SEMA

Dans ce modèle, le flux migratoire de la zone I, laquelle appartient à la strate K, à destination de la zone J, laquelle appartient à la strate L (L et K peuvent être identiques) durant la période 54-62 est désigné par IMINT (I (K), J (L)). Soit PT (I (K), 54), la population de la zone I. La stratification des zones se fait en trois ensembles : partie d'un département, partie urbaine appartenant à des agglomérations de moins de 50 000 habitants, partie urbaine du département appartenant à des agglomérations de plus de 50 000 habitants..

On désigne par PHI (I, (K), L) la probabilité qu'un individu de la zone I (K) migre dans une zone appartenant à la strate L..

On désigne par PSI (J (L)) la probabilité pour qu'un migrant issu de I (K), et migrant dans la strate L, soit présent en J (L) en 62.

Le modèle se présente sous la forme :

$$\text{IMINT (I (K) . J (L))} = \text{PT (I (K), 54)} . \text{PHI (I (K), L)} . \text{PSI (J (L))}$$

La fonction PSI (qui décrit une probabilité conditionnelle) se présente sous la forme :

$$\text{PSI (J (L))} = \text{H} . \text{RET (I, J)} . \text{PA (NA, J)} . (\text{SAL (J)})^m . \frac{1}{(\text{D (I, J)})^n}$$

où RET est un indicateur arbitraire de retention dans la zone I, des mouvements migratoires à destination de J.

La fonction PHI se présente sous forme :

$$\text{PHI (I (K), L)} = \sum_{\text{M (K; L)}} \left[a (\text{M (K, L)}) . (\text{FACTEUR (M (K, L))} \right]$$

où les facteurs explicatifs choisis varient en fonction des 9 couples de strates (K, L). Ils sont limités arbitrairement au nombre de 4, et les plus fréquents d'entre eux sont ceux qui interviennent dans la fonction PSI.

L'étude de la SEMA a porté sur les sous-modèles PHI et PSI et l'étude du modèle global n'a pas été menée à terme. Les résultats, compte tenu de l'effort fourni, sont très décevants.

2.12 - Modèles migratoires à la marge

Deux approches différentes sont possibles lorsque l'on raisonne sur les deux marges d'un tableau de migrations interzonales : dans la première, le raisonnement est mené de façon différente pour chacune de ces marges, dans la seconde, ces deux marges sont consolidées et l'on raisonne en terme de solde migratoire.

α) Raisonnement sur une marge

Mrs BLANCO⁽¹⁾ ne s'intéresse qu'aux immigrants (et donc à une seule des marges) et propose le modèle explicatif suivant pour une période donnée, où PA (Z, NA) est la population active non agricole, PT (Z, SCO) la population totale scolarisée, PT (Z, I18), la population âgée de plus de 18 ans, et SAL (Z) le salaire de la zone Z.

$$IMINT (Z) = a \cdot PT (Z, I18) + b \cdot \frac{\Delta PA (Z, NA)}{PA (Z, NA)} + c \cdot \frac{\Delta PT (Z, SCO)}{PT (Z, SCO)} + d \cdot \frac{\Delta SAL (Z)}{SAL (Z)} + U (Z)$$

Le modèle fonctionne en deux temps : dans un premier temps, les flux migratoires sont supposés être nuls et l'on calcule les différentes variables exogènes autres que le salaire, en fonction de l'évolution naturelle de la pyramide des âges et, dans un second temps, on explique les flux observés par la différence entre les valeurs "naturelles" et observées des variables exogènes.

Le coefficient de détermination trouvé est compris entre 0,83 et 0,86.

β) Raisonnement sur le solde

LOWRY, à la suite de ses travaux sur les flux migratoires (voir 3.1 a), s'est inspiré du modèle de BLANCO et a présenté le modèle suivant où NAC signifie non agricole civile et NAM non agricole militaire.

$$\frac{IMINT (Z) - EMINT (Z)}{PT (Z)} = a \cdot \frac{\text{accroissement naturel } PT (Z, 115 - 64)}{PT (Z)} + b \cdot \frac{\Delta PA (NAC, Z)}{PT (Z)} \\ + c \cdot \frac{\Delta PA (NAM, Z)}{PT (Z)} + d \cdot \frac{\Delta PT (Z, SCO)}{PT (Z)} + e \cdot (\text{revenu médian}) + f$$

(1) Mrs BLANCO - Review of Economics and Statistics, mai 1964 Prospective unemployment and interstate population.

Il s'avère, en fait, que seules les deux premières variables sont significatives ; mais si l'on prend le solde migratoire pour les 15 - 64 ans, toutes les variables sont significatives, à l'exception de l'accroissement de population scolaire.

On peut reprocher au modèle de BLANCO comme à ceux de LOWRY leur caractère tautologique : dans la mesure où le mouvement migratoire est expliqué essentiellement par ses composantes (et en particulier les 15 - 64 par la population active). LOWRY, très conscient de cette critique possible, donne l'explication suivante : l'accroissement de la population active correspond à une offre d'emploi, laquelle engendre une immigration avec tout une série d'ajustements mutuels, ce phénomène se reproduisant à plusieurs reprises durant l'intervalle de temps envisagé. Pour LOWRY, cette relation approxime un "équilibre" (au sens de la statistique comparative). Il est malheureusement très difficile de tester cette hypothèse.

Il faut souligner qu'à notre connaissance aucun modèle explicatif de chaque marge (immigration et émigration) n'a encore été tenté. Il est possible que cette voie permette de trouver une explication socio-économique satisfaisante des arrivées et des départs d'une zone, tout en ne faisant appel qu'à une masse de données restreinte.

2.2 - Modèles prévisionnels

2.21 - Utilisation des modèles descriptifs

L'une des utilisations des modèles explicatifs est leur utilisation sous une forme prévisionnelle, en conservant paramètres et hypothèses, implicites ou explicites, du modèle, et en introduisant de nouveaux vecteurs de données pour les variables exogènes.

Il semble, en fait, que peu d'utilisation prospective de modèles descriptifs opérationnels ait encore été effectuée. A côté de cette approche analytique du problème se développe depuis quelques années une approche basée sur des techniques de simulation.

2.22 - Modèles de type Markovien

2.22 a - Présentation et intérêt

Cette approche, développée par ROGERS⁽¹⁾, permet de trouver la conséquence de certaines hypothèses relatives aux flux migratoires dans des problèmes où l'obtention d'une solution analytique est trop coûteuse ou impossible.

(1) ROGERS - A Markovian Policy Model of interregional Migration in Papers, Vol. 17, 1966, Regional Association.

ROGERS,⁽¹⁾ dans sa première approche présente le modèle simplifié suivant : soit une matrice carrée de migrations interzonales, $IMINT(I, J, T(T+h))$, observée sur la période $(T, T+h)$, matrice dans laquelle la zone de départ est repérée par la ligne d'arrivée par la colonne.

Si l'on fait abstraction du mouvement naturel de la population (i.e. naissances et décès), la fréquence relative : $K(I, J) = IMINT(I, J, (T, T+h)) / PT(I, T)$

peut être supposée stable dans le temps et assimilée à une probabilité. Il n'y a donc, à la base, aucune étude causale, mais seulement la recherche des implications du choix de certaines hypothèses.

Il est possible de prendre en considération les décès en ajoutant une ligne et une colonne à cette matrice. Comme nul ne peut revenir des enfers, la probabilité de passage de cette zone aux autres est nulle (donc les éléments de cette ligne sont des zéros à l'exception de celui de la dernière colonne qui, bien entendu, vaut 1).

Lorsque les décès ne peuvent être pris en compte, on parle de modèle avec chaîne de MARKOV absorbante. Un modèle avec chaîne de MARKOV ergodique⁽²⁾ a été essayé par TARVER et GURLEY aux Etats-Unis⁽³⁾, et dont les résultats, qui n'ont pu être testés par comparaison avec le réel, paraissent être du domaine des possibles.

La connaissance de la répartition spatiale de la population pour une année de base et celle de la matrice de probabilité de transition permettent d'étudier, dans le cadre des hypothèses du modèle ergodique ou du modèle absorbant, un régime stationnaire vers lequel on tend.

Cet état stationnaire peut être modifié par une politique de transferts volontaristes. Il est possible⁽⁴⁾ de voir si une répartition spatiale de la population souhaitée comme régime stationnaire est mathématiquement possible et de trouver le vecteur de correction à apporter à chaque itération (= transferts volontaristes de population, plus, dans le cas du modèle absorbant, volume des naissances).

2.22 b - Limites et perspectives⁽⁵⁾ des modèles Markoviens

Un certain nombre de travaux ont montré que cette matrice de transition n'était pas stable dans le temps. L'une des raisons peut être trouvée dans l'incidence de la structure par âge dans la propension à émigrer. Une amélioration d'un tel modèle consisterait à éclater chaque information de la matrice suivant le

(1) ROGERS - A markovian Policy Model of interregional Migration in Papers, Vol. 17, 1966, Regional Association.

(2) Pour la présentation des chaînes de Markov, voir en particulier PLATIER, CAHUZAC et CHAMBADAL - Economie et mathématiques, tome 2, p. 249 et suivantes.

(3) TARVER-GURLEY - A stochastic Analysis of Geographic Mobility and Population Projection of the Census Division in the United States, Demography, n° 2 (1965).

(4) ROGERS : opus cit.

(5) On trouvera une description des dernières recherches en cours, dans la communication de MORRISSON à la Cornell Conference on human mobility (1968).

sexe et l'âge ; mais, si l'on ventile par tranches d'âge annuelles jusqu'à 100 ans, la taille de la matrice est multipliée par $(200)^2 = 40\ 000$ (!).

Une autre difficulté, dans le cas du modèle absorbant, réside dans l'arbitraire de l'introduction d'un vecteur de naissance. Une amélioration peut être trouvée par l'apport d'une matrice diagonale de taux de fécondité par âge qui, appliquée au vecteur obtenu en fin de période, permet de calculer la population de la première tranche d'âge, mais là aussi les problèmes de dimension de la matrice deviennent vite inextricables.

Une critique importante, quant au réalisme de cette approche, peut être formulée : le régime permanent n'est correctement approximé qu'au bout de nombreuses itérations. Il y a là un problème de convergence plus ou moins rapide vers le régime permanent et qui dépend des données. Or les hypothèses de constance structurelle sont d'autant moins réalistes que l'on s'éloigne dans le temps de la période initiale. La signification du régime permanent est donc très sujette à caution.

Un certain nombre de travaux⁽¹⁾ avaient montré que le risque de migrer était une fonction décroissante de la durée de présence de l'individu dans la zone envisagée. Certains auteurs⁽²⁾ ont voulu en tenir compte dans le "Cornell retention Model", ce qui a pour effet qu'il est impossible d'atteindre un état stationnaire.

BARTHOLOMEW⁽³⁾ propose de remplacer l'hypothèse de constance dans le temps de la probabilité de migrer de I vers J par une fonction de probabilité dépendant du temps (modèle de POISSON) dont le paramètre suivrait une distribution empirique à trouver.

Ces divers apports sont difficiles à prendre en considération, pour des raisons de données mais, aussi pour des raisons de traitement : compte tenu de la taille du tableau à traiter, il est vraisemblable que, pour des raisons de coûts, on ne pourra, en France, faire l'emploi de modèles Markoviens très élaborés.

(1) Ceux d'HAGERSTRAND en particulier.

(2) HENRY, NEIL, Mc GENNIS, TEGTMERGER - A finite Model of Mobility (1968) Communication au Congrès de la Population Association of America.

(3) BARTHOLOMEW - Stochastic Models for Social Process (1967).

SECTION 3 - BILAN PROVISOIRE

Nous devons constater que dans toutes les analyses rencontrées le problème de l'ajustement réciproque de l'offre et de la demande d'emploi est toujours éludé ou plus exactement dans les quelques cas où le problème est esquissé il est supposé réglé par le jeu des variations de salaires. Nous verrons dans les pages qui suivent ce qu'il faut penser de cette simplification.

Toujours est-il qu'il nous faudra poser un certain nombre d'hypothèses originales pour étayer notre démarche. Par contre certaines analyses seront des plus utiles pour définir nos différents sous-modèles.

LISTE des VARIABLES et INDICES UTILISES

I - INDICES

- Z référence une zone. On peut utiliser Z et Z' pour désigner 2 zones (remplacé par I, J et K si 3 zones utilisées) ;
- T référence le temps ;
- ΔT référence la période (T + 1, T) ;
- B référence la branche (où l'indice général B peut être remplacé par NA : non agricole, NAC : non agricole civile, NAM : non agricole militaire, BASE : emploi de base) ;
- SCO référence la population totale scolarisée ;
- I 18 référence la population totale âgée de plus de 18 ans ;
- I 15 - 64 référence la population totale âgée de 15 à 64 ans.

II - VARIABLES

- IMINT (I, J) nombre de migrants originaires de I en T et présents en J en T + 1 ;
- PT (Z) population totale de la zone ;
- PA (Z) population active de la zone (peut être, en outre, indiquée par B) ;
- ATON (Z, B) coefficient d'attraction de la zone pour la branche B ;
- FAC (Z, K) valeur prise pour le facteur de localisation K dans la zone Z (si les facteurs pris en compte sont différents de branche à branche, l'indice B intervient) ;
- I (Z, K) indicateur calculé de l'importance du facteur K dans la zone Z ;
- TI (Z, T) terrains industrialisables en T dans Z ;
- TAC (Z) taux d'activité ;
- IND (B) taux d'emploi induit ;
- FON (Z, B) taux d'emploi fondamental de la branche B dans la zone Z ;
- H (Z) coefficient d'induction (part de population active induite dans la population active) ;
- M (Z) multiplicateur d'emploi (utilisé avec une variation d'emploi fondamental) ;
- D (I, J) distance entre les zones I et J ;
- A (K, B) coefficient de pondération affecté au facteur de localisation I .

Le fonctionnement du marché local de l'emploi

5

CHAPITRE 5 - Le FONCTIONNEMENT du MARCHE LOCAL de l'EMPLOI

Il s'agit ici de définir les sous-modèles retenus (et qui seront examinés au cours des chapitres suivants) en fonction de l'objectif premier que l'on s'est fixé, la détermination à moyen terme de l'emploi spatialisé. Le cadre de cette détermination est appelé, d'une façon habituelle, marché de l'emploi. Il convient donc d'examiner cette notion de marché de l'emploi sur un plan théorique et ses implications avant de regarder les choix effectués au niveau de la recherche de sous-modèles permettant la détermination des agents intervenant sur ces marchés.

SECTION 1 - ETUDE de la NOTION de MARCHE de l'EMPLOI

A la base du concept de marché, se trouve l'idée d'une transaction, c'est-à-dire l'échange de prestations en nature (biens ou services) ou monétaires entre deux parties, et c'est d'ailleurs l'une des acceptions actuelles de ce terme. Par extension on a appelé marché, l'emplacement spécialement affecté à ces transactions. Puis lors de la genèse de l'économie politique une nouvelle acception de ce terme est donnée en considérant l'ensemble des échanges de prestations se produisant, ou pouvant se produire, au cours d'une période sur un territoire donné.

Les approches micro-économiques et macro-économiques aboutissent à des conceptions nettement différentes du cadre spatio-temporel. Il ne peut qu'en être ainsi, car pour reprendre une formule de R.G.D. Allen⁽¹⁾ . la macro-économie "s'applique à l'étude des relations entre de vastes ensemble économiques, par opposition aux processus de décision des individus et des firmes qui font l'objet de la micro-économie". Il importe d'examiner sommairement comment s'insère la définition du marché de l'emploi dans ces deux approches-type des problèmes économiques afin de situer correctement les bases de l'approche tentée face à ces deux démarches classiques de la pensée économique.

(1) RGD Allen : "Théorie Macro-économique" traduction française (Armand Colin 1969)

1.1 - Le marché de l'emploi dans le cadre de la micro-économie

On examinera successivement la position des auteurs classiques et néoclassiques puis des tendances récentes.

1.11 - Le marché de l'emploi chez les classiques

Dans le chapitre 10 du très célèbre "Recherche sur la nature et les causes de la richesse des Nations", Adam Smith a jeté les bases de l'étude de ce marché très particulier qu'est le marché du travail, après l'étude des autres marchés. Dans ce marché s'échangent une offre et une demande de travail⁽¹⁾. Voici ce que Smith écrit sur la détermination de la demande de travail :

"L'ensemble des avantages et désavantages qui sont attachés aux différents emplois de travail et de capital doivent, en un même voisinage, être égaux ou tendre continuellement vers cette égalité. Si, en un même voisinage, se trouvaient quelques emplois manifestement plus avantageux ou moins avantageux que les autres, on aurait tellement de candidats aux premiers et de défections aux seconds que leurs avantages respectifs auraient tôt fait de rejoindre le niveau de celui des autres emplois. Ceci du moins, serait le cas dans une société où les choses suivent leur cours naturel, où la liberté est totale, et où chaque individu est parfaitement libre à la fois de choisir l'activité qu'il estime lui convenir et de changer aussi souvent qu'il le désire. L'intérêt de chaque individu le poussera à rechercher les emplois avantageux et à éviter les emplois désavantageux".

A la base de cette conception se trouvent donc une hypothèse de comportement et la conséquence de cette hypothèse (l'égalité entre emplois de l'ensemble des avantages et désavantages qui leurs sont rattachés), sous la condition de réalisation de ce qu'il est convenu d'appeler, la concurrence pure et parfaite. Par ailleurs le marché est spatialement restreint, puisque Smith parle de "voisinage", et les ajustements se placent dans un cadre temporel à court ou moyen terme (l'analyse n'est donc pas statique). On peut ajouter que l'ajustement s'effectue par des fluctuations du couple avantage-désavantage, ce qui rend inutile la migration. Donc chez Smith non seulement le marché de l'emploi est local mais on peut ajouter que son imperméabilité ne mène pas à l'autarcie car les "avantages comparatifs" différents entre zones mènent chez cet auteur à une régulation des disparités par un échange de biens.

(1) Dans ce qui suit on parlera indifféremment de demandes de travail (sous-entendu par les particuliers) et de l'offre d'emploi (sous-entendu par les entrepreneurs). De même, les expressions demande d'emplois et offre de travail seront considérées comme équivalentes. La notion de "travail" est celle habituellement utilisée par les économistes (sous l'angle facteur de production) tandis que celles "d'emploi" l'est par les non-théoriciens économistes (c'est-à-dire que c'est celle qui est de loin la plus répandue).

L'offre de travail par les entrepreneurs résulte d'une hypothèse de comportement, la recherche du profit maximum pour chacun d'entre eux, et la demande de bien qui leur est faite. Si la recette procurée par une unité supplémentaire de bien est supérieure à son coût de production, cette unité supplémentaire sera produite. La quantité globale de biens à produire permet de définir le volume de travail nécessaire et donc l'offre d'emploi par les entrepreneurs. Cette quantité de bien à produire résulte, elle, de l'agrégation des demandes individuelles, solvables à un prix donné.

1.12 - Le marché de l'emploi chez les néoclassiques

La pensée de Smith est interprétée et développée par plusieurs auteurs de la fin du 19^{ème} siècle et du début du 20^{ème} et en particulier Mac Culloch, Stuart Mill et Senior. Rottenberg⁽¹⁾ résume leur position quant à la demande de travail en ces termes "il est clair que c'est l'avantage net total et non la rémunération qui est la pierre de touche du choix de l'activité et de son changement dans la théorie de ces économistes et c'est cet avantage net total et non la rémunération qui est dit égal entre toutes les activités". Il poursuit en disant "Récemment certains économistes qui ont développé un système plus élégant d'analyse que les classiques, font usage du prix comme allocation du travail entre divers emplois possibles, et leur courbe de demande est tracée en reliant la quantité à son prix. Les courbes quantités-prix sont tracées sous l'hypothèse du "toutes choses égales par ailleurs" mais depuis que des économistes ne sont plus répétitivement explicites sur ce point, cela est quelque fois oublié".

Cette approche semble intéressante tout d'abord parce que l'on dispose avec la rémunération d'une grandeur quantifiable et dont la variation est continue (ce qui n'est pas le cas de la plupart des autres facteurs explicatifs) et ensuite parce que l'utilisation de ce modèle analytique donne des résultats satisfaisants (on observe en effet sur le long terme un courant continu de migration de zones de bas revenus vers des zones de hauts revenus).

En fait comme l'indique Rottenberg diverses enquêtes semblent établir que les motivations de choix sont très complexes et qu'en outre la connaissance des avantages comparatifs nets (parmi lesquels le facteur sécurité d'emploi joue un grand rôle) entre diverses offres est le plus souvent sommaire et même dans certains cas impossible et que si cette connaissance existe, bien souvent le travailleur n'agit pas toujours dans le cadre de la logique proposée parce que l'intégration de toutes les données qui interviennent est délicate ou parce que certains coûts de mutation semblent prohibitifs (habitude).

Du côté de l'Offre de Travail l'accent est surtout mis sur l'hypothèse de comportement des entrepreneurs qui sont supposés agir dans le but de maximiser leur profit par l'intermédiaire de l'égalisation de leurs revenus et coûts marginaux, et à cet effet le salaire intervient. Une très abondante littérature est développée sur ce thème⁽²⁾ et une certaine radicalisation du comportement théorique de l'entrepreneur est postulée, ce qui mène à une remise en cause importante de la théorie marginaliste classique et de ses postulats à partir des années 40.

(1) S Rottenberg : "On choice in Labor Markets" (Industrial and Labor Relations Review volume 9 - 1956)

(2) Voir l'article de Lester "Shortcoming of Marginal Analysis for Wage-employment Problems" (American Economic Review - volume 36)

Lester⁽¹⁾ en particulier par l'intermédiaire d'enquêtes partielles établit que l'offre d'emploi est davantage fonction de la demande de produits et de son anticipation par l'entrepreneur que par le taux de salaire, que les chefs d'entreprises estiment que le coût variable unitaire est décroissant lorsque leur entreprise travaille de 70 % à 100 % de sa capacité (ce qui les rend moins sensibles à une hausse de salaire) que les firmes multirégionales à de rares exceptions près, n'ajustent pas leur ratio capital-travail pour compenser les distorsions de salaires.

En fait les difficultés pratiques de saisie directe des courbes d'offres et de demandes d'emploi sont telles qu'aucune construction empirique de ces courbes n'a été tentée et que de ce fait comme le souligne Machlups⁽²⁾ il est très difficile d'infirmier ou de confirmer la théorie marginaliste. On peut ajouter en outre que les processus d'ajustement postulés relèvent du court terme et que même dans certains cas procèdent d'une analyse statique et non dynamique ce qui complique encore le problème.

1.13 - Tendances récentes

Il semblerait⁽³⁾ que devant l'impasse dans laquelle se trouve l'approche concrète, les auteurs contemporains qui ne se sont pas attaqués directement au problème macro-économique, se soient attachés à analyser les données institutionnelles (en particulier l'influence du cadre syndical et la négociation du contrat de travail sous ses divers aspects), le rôle effectif des différences de salaire dans la réallocation des travailleurs. Il semble intéressant d'examiner brièvement quelques travaux récents sur ce dernier point⁽⁴⁾, car on touche là l'une des rares variables instrumentales, d'autant plus que ces travaux constituent un dépassement certain de la conception classique ou néo-classique de ce problème.

Les travaux très intéressants de Reddaway⁽⁵⁾ partent de la remise en question du principe d'allocation optimale des travailleurs en fonction de la libre flexibilité des salaires en faisant remarquer que l'on constate une certaine assymétrie entre la tendance à la baisse des salaires dans les activités en régression et celle à la hausse des activités en expansion et qu'une étude globale sur la liaison entre l'évolution de salaires relatifs d'une activité et de la variation relative correspondante d'emploi sur période de 5 ans et pour 111 types d'activités économiques en Grande Bretagne suggère que le lien est faible entre les deux phénomènes et relativement fort lorsque les activités sont groupées en 14 secteurs. Cette divergence de résultats peut s'expliquer en partie par une mobilité entre activités d'un même secteur mais cette compensation doit être relativement réduite compte tenu des implantations géographiques spécifiques de ces diverses activités, et s'explique surtout par les effets d'entraînement de croissance de salaire au sein d'un même secteur qui font que le salaire de certaines activités

(1) Article cité

(2) Machlups : "Marginal Analysis and Empirical Research" (American Economic Review - volume 36)

(3) L'essentiel du recueil de textes de Mac Cormick et Owen Smith intitulé "the Labor Market" (Penguin Book 1968) semble être celui des ouvrages de ce type paru ces dernières années, le plus orienté vers le problème qui est au coeur de cette étude, semble résumer assez bien ces tendances récentes.

(4) Dans le cadre du marché du travail (à la différence de certains travaux présentés dans le chapitre 4)

(5) Reddaway "Wage Flexibility and the Distribution of Labour" (Lloyd's Bank Review numero 54 - 1959)

croit indépendamment de l'évolution des effectifs. Parallèlement certains auteurs comme Reder⁽¹⁾ et Turner⁽²⁾ cherchent la cause du resserrement des disparités salariales ; pour le premier celui-ci est dû à une modification des critères de recrutement par les employeurs (la modification de ces critères étant un substitut de la modification possible de salaire offert, tandis que pour le second ce phénomène coïncide avec l'émergence du syndicalisme dans un contexte inflationniste et n'est que très partiellement imputable à l'accroissement du niveau général d'éducation et du progrès technique. Mais déjà avec cet auteur on quitte en partie le domaine de la micro-économie.

1.2 - Le marché de l'emploi dans le cadre de la macro-économie

On peut dire que si les problèmes d'emploi sont presque toujours sous jacents dans la théorie macro-économique c'est sous l'aspect facteur de production. On peut même ajouter que les travaux effectués jusqu'à présent ne s'intéressent que peu ou pas du tout à l'aspect demande de travail soit parce que, comme chez Keynes l'horizon est trop court (la demande est une donnée sur laquelle il convient d'essayer de se cadrer). Soit parce que dans les modèles à long terme on travaille sur des données d'emploi telles qu'elles résultent implicitement du jeu du marché.

Toutefois divers modèles empiriques nationaux à moyen terme, d'inspiration macro-économique (dont le modèle Fifi est une bonne illustration) ont bien été obligé de prendre en considération cet aspect demande. Mais c'est par l'application à une population calculée (éventuellement par strates) de coefficients empiriques (résultant d'hypothèse de stabilité statique ou dynamique de données observées sur le passé) que le problème est réglé. Les processus d'ajustement entre offre et demande sont partiellement normatifs (actions sur l'éducation, mise à la retraite ...) partiellement mécaniques (cadrage itératif vers un état stable) sans référence explicite à une vision complète et cohérente du marché de l'emploi.

Par ailleurs des études descriptives⁽³⁾ relatives aux diverses données d'emploi commencent à être entreprises, fournissant ainsi des matériaux importants, pouvant servir de support à un approfondissement théorique du problème.

(1) Reder "the Theory of Occupational Wage Differentials" (American Economic Review - volume 44 - 1955)

(2) Turner "the Theory of Wage determination" (Macmillan - 1957)

(3) Voir entre autre le manuel de MM. SELIER et TIANO : "Economie du Travail" (Themis - 1962)

1.3 - Conception retenue du marché de l'emploi

Dans l'approche qui a été tentée le marché de l'emploi est défini par l'ensemble des trois caractéristiques suivantes, qui réagissent d'ailleurs les unes sur les autres : un cadre temporel, un cadre spatial et un ensemble de mécanismes. Le cadre temporel est en fait imposé par le but assigné : c'est le moyen terme⁽¹⁾ il ne reste donc qu'à examiner les deux autres caractéristiques. Nous envisagerons ensuite le problème de l'unicité du marché de l'emploi.

1.31 - Le cadre spatial

Si l'activité humaine est suffisamment polarisée autour de centres bien distincts les uns des autres, (en faisant abstraction du cas bien particulier des agriculteurs tant que ceux-ci ne cherchent pas à changer d'emploi, car pour eux le domicile et le lieu de travail coïncident à peu de chose près), la zone idéale correspondant géographiquement à un marché de l'emploi est l'enveloppe des domiciles des gens travaillant dans ce centre. Si les centres sont suffisamment distants, ces enveloppes seront disjointes. La taille de cette zone doit être telle que des migrations définitives d'un point à un autre de la zone ne soient que rarement nécessaires, (cette dernière condition est indispensable sinon des zones comme la Grande-Bretagne où les migrations alternantes avec l'étranger sont inexistantes, répondraient à cette définition). C'est donc une conception spatiale proche de celle de Smith (voisinage) qui est retenue.

On constate dans les faits que ce cas idéal est rarement atteint, les zones d'influences des pôles économiques se chevauchent et le problème pratique à résoudre est d'essayer de définir un découpage qui minimise ces interpénétrations. C'est en partant de ces considérations et de celles développées au chapitre 2, que l'étude du choix du découpage (chapitre 6) a été menée.

On supposera dans ce qui suit que ce problème d'un découpage pertinent avec la notion de marché de l'emploi est résolu.

1.32 - L'ensemble des mécanismes

Ce que l'on peut observer, c'est un état de l'emploi, c'est-à-dire la structure observable à une date donnée et pour chaque zone, de la population active, occupée ou en chômage, par secteur, sexe, qualification,...

(1) On verra au § 2,3 qu'il est en théorie possible de s'affranchir de cette contrainte

Un certain nombre de tensions existent qui tendent à modifier l'équilibre réalisé. Des ajustements s'effectuent en fait quotidiennement (mutation spatiale, sectorielle, mise à la retraite, jeune travailleurs,...). A l'origine de ces tensions se trouvent un certain nombre de facteurs explicatifs : évolution des productivités, des élasticités de demande, élévation du niveau de vie, diminution du temps de travail, etc...

Si l'on parvient à isoler correctement l'ensemble de ces facteurs, il est possible de considérer un ajustement se réalisant de façon discrète et non continue ; c'est-à-dire que l'on cherchera à déterminer séparément la demande de travail de l'offre de travail en l'absence de toute intervention (et en particulier sans se préoccuper de leurs interaction) puis on recherchera les mécanismes permettant de passer de cet état potentiel, par définition déséquilibré, à un état réel en situation d'équilibre.

On peut indiquer tout de suite qu'à la différence des économistes classiques ou néo-classiques, la définition de l'offre et de la demande ne repose pas sur des hypothèses explicites de comportement alors que l'on sera obligé d'y avoir recours pour les mécanismes d'ajustement et que contrairement à l'approche macro-économique classique l'aspect "détermination de la demande d'emploi" est aussi importante que celle de son offre.

1.33 - Remarques sur l'unicité du marché de l'emploi

Il est bien évident que le marché de l'emploi est très hétérogène, ce qui résulte de la non substituabilité des individus, due à leurs aptitudes et à leurs préférences. Dans la réalité plusieurs marchés de l'emploi coexistent, avec leurs découpages spatiaux et leurs média spécialisés. A ce propos, il serait intéressant de faire une étude des annonces d'offres et de demandes d'emploi dans les journaux à vocation "locale", "régionale" ou "nationale"; toutefois, une consultation sommaire de la presse montre que l'étendue spatiale des marchés de l'emploi varie en fonction des catégories socio-professionnelles visées. C'est ainsi par exemple que le marché des Cadres supérieurs est en grande partie national (et même pour certains "profils" il est international)⁽¹⁾ Une autre dichotomie fondamentale est à faire selon qu'il s'agit de la recherche d'un premier emploi ou non ; et encore faut-il noter que l'age intervient de façon sensible pour cette dernière catégorie.

Pourtant il ne nous a pas été possible de tenir compte dans une première approche de l'hétérogénéité de ces marchés, car il aurait fallu partir de données départementales par catégorie d'activités économiques détaillées et catégories socio-professionnelles et sexe, et ce, pour trois recensements; ce qui nous aurait amené à intégrer

(1) De nos jours, au "brain - drain" maintes fois dénoncés, s'est ajouté ce que l'on pourrait appeler le "manager-drain", sous l'influence croissante des firmes multinationales.

à la base de données plus de 200 000 données, disponibles seulement sur support-papier pour les 2/3 d'entre elles. Cet investissement nous a paru se justifier qu'à la seule condition que la première étude de faisabilité soit couronnée de succès. Il semble toutefois intéressant d'indiquer nos réflexions sur ce qu'il aurait été possible de faire à partir de telles données :

Appelons $A(D,B,Q,T_0)$ l'élément du tableau d'emplois observés A , à la date T_0 , pour le département D , la branche B et la qualification Q (ces deux termes étant employés ici par commodité pour stratifier respectivement l'offre et la demande). Désignons par C le tableau de coefficients dont les éléments sont définis comme suit :

$$C(D,B,Q,T_0) = A(D,B,Q,T_0) / \sum_Q A(D,B,Q,T_0)$$

Supposons que le modèle calcule pour la date T_1 et pour chaque département un tableau d'offres potentielles d'emploi par branche $F(D,B,T_1)$ et un tableau de demandes potentielles d'emplois par qualification, noté $E(D,Q,T_1)$ dont la marge par rapport aux qualifications n'a aucune raison d'être identique à celle de F par rapport aux branches.

Examinons quel mécanisme d'ajustement on peut imaginer pour rendre ces données compatibles sans remettre en cause le calcul de l'état potentiel (c'est à dire pour reprendre notre terminologie⁽¹⁾ en refusant les mécanismes de rétroaction).

Dans un premier temps on peut penser qu'une partie des offres d'emploi présente un caractère fictif, dans la mesure où un certain nombre de travailleurs n'ont pas changé d'emploi. Il vient alors naturellement à l'idée de répartir la demande de chaque qualification conformément à sa structure par branche puis de sommer sur la qualification les emplois ainsi calculés (par branche et qualification) puis comparer le résultat avec l'offre c'est-à-dire :

$$\text{Comparer : } \sum_Q C(D,B,Q,T_0) E(D,Q,T_1) \text{ et } F(D,B,T_1)$$

Mais il est à peu près certain que cette comparaison révélera des écarts tantôt positifs, tantôt négatifs, ce qui signifie qu'à cause de ces derniers il n'est pas possible de conserver la structure branche-qualification observée en T_0 sans remettre en cause les calculs de demande potentielle. Pour arbitrer le conflit sans remettre en cause les données du tableau C (du moins dans une première étape) il faut introduire une fonction à optimiser sous contrainte .

(1) Voir page 127.

Celle ci pourrait être la recherche de X, tableau des emplois ajustés.

$$\forall D \left[\begin{array}{l} \text{Min} \left[\sum_B \sum_Q C(D, B, Q, T_0) \cdot X(D, Q, T_1) - \sum_Q E(D, Q, T_1) \right] \\ \text{sous contrainte}^{(1)}: X(D, Q, T_1) \leq E(D, Q, T_1) \end{array} \right]$$

Ce qui nous ramène pour chaque département à la résolution d'un programme linéaire classique. La fonction objectif de ce programme consiste à minimiser le volume des désajustements, mais il est bien évident que d'autres fonctions objectifs peuvent être intégrées. Par exemple si pour une raison particulière (politique sociale liée à des considérations de plus ou moins grande mobilité par exemple) on entend privilégier une catégorie, il sera facile d'en tenir compte en intégrant des coefficients de pondération $K(Q)$ dans la fonction-objectif qui devient :

$$\text{Min} \left[\sum_B \sum_Q K(Q) \cdot C(D, B, Q, T_0) X(D, Q, T_1) - \sum_Q K(Q) F(D, B, T_1) \right]$$

la contrainte restant bien évidemment la même.

Se trouve en fait posé avec cette définition de la fonction-objectif, toute la problématique de la politique de l'emploi. Cette politique différera suivant que la priorité sera accordée à la croissance maximum du PNB, ou de celle des revenus, ou à l'atténuation des disparités régionales, ou encore (l'hypothèse retenue ici) de celle du chômage minimal, ces objectifs dont la liste n'est pas limitative sont en grande partie concurrents.

A la fin de cette première itération, une partie importante des offres et demandes est donc satisfaite. On établit alors le tableau résiduel des demandes non satisfaites, noté \hat{E} et celui des offres noté F . L'ajustement de ces deux tableaux ne peut maintenant s'effectuer qu'au détriment de la constance du tableau C ou de celle des marges. Plusieurs techniques sont alors envisageables. Pour notre part nous avons pensé qu'il serait intéressant d'utiliser une technique de RAS.

(1) des contraintes supplémentaires peuvent éventuellement être introduites, par exemple pour respecter un minimum pour certaines branches.

Tout d'abord on calculera au niveau national (et donc après sommation) les écarts entre les emplois satisfaits par branche et ceux calculés exogènement par le modèle national de référence ; ce vecteur résiduel $N(B)$ sert alors de marge de référence. Par ailleurs un cadrage à la proportionnelle du tableau \hat{F} est fait pour que $\sum_D \sum_B \hat{F}(D,B) = \sum_B M(B)$. Un RAS⁽¹⁾ est alors effectué sur le tableau \hat{F} afin que les marges horizontales ($N(B)$) et verticales soient satisfaites. On peut imaginer que pour ce RAS certaines cases soient automatiquement bloquées en fonction des données disponibles pour T_0 pour éviter que la structure du tableau obtenu diffère de trop de celle du tableau original (par exemple on pourrait figer toute case qui au cours d'une itération du RAS donne un ratio (par exemple $F(D,B) / \sum_D \sum_B (D,B)$) différent de plus de 10 % de celui observé pour T_0)

La même opération est parallèlement menée pour le tableau par qualification et département. On dispose alors pour chaque département d'un vecteur d'offres résiduelles par branche et un autre de demandes résiduelles par qualification. De nouveau un programme linéaire peut être utilisé. Ceci suppose une fois de plus que la structure de qualification par branche est inchangée dans chaque département.

Une troisième et dernière itération est opérée après avoir calculé comme précédemment de nouvelles marges résiduelles. On peut alors par exemple déformer les coefficients de qualification par branche dans un département proportionnellement à la variation relative de ces coefficients pour la France entière, pondérée éventuellement par le poids relatif du département considéré dans les effectifs nationaux de la branche en T_0 . Un programme linéaire est utilisé alors avec nouveaux coefficients et le nouveau vecteur résiduel de demande d'emploi est mis en chômage.

Le mécanisme ainsi décrit correspond en fait à l'un des mécanismes d'ajustement possibles. Les hypothèses sous jacentes sont facilement justifiables même si à juste titre on peut les considérer comme arbitraires et restrictives.

Le choix auquel nous avons été contraint de ne pas pouvoir dissocier les demandes par qualification, prive cette étude exploratoire de mécanismes explicatifs essentiels des processus d'ajustements. La démarche adoptée sera de ce fait la recherche d'une typologie des états potentiels puis celle de mécanismes explicatifs, démarche lourde et peu sûre dans la mesure où les interactions des mécanismes explicatifs réels possibles risquent au niveau global de masquer ces derniers.

(1) Voir chapitre 7, § 3.22 - page 171.

SECTION 2 - IMPLICATIONS de la DEFINITION du MARCHE de l'EMPLOI

On examinera successivement quelques unes des implications les plus importantes de la définition retenue.

2.1 - Remarques sur la notion d'état potentiel

2.11 - Influence de la période retenue sur la notion d'état potentiel

A la base de ce concept d'état potentiel se trouve donc l'idée de rechercher ce qui se passerait si l'on n'intervenait pas dans l'évolution des phénomènes c'est-à-dire si les prises de décision des différents groupes d'agents n'étaient pas influencées les unes par les autres :

L'offre potentielle d'emploi résulterait donc de considérations sur l'évolution nationale (raisons techniques,...) et d'autres relatives à la zone (dynamisme propre,...) sans se préoccuper de savoir si ces offres pourront être toutes satisfaites et de connaître les éventuelles réactions chez les offreurs d'emploi d'autres zones, les divers marchés étant indépendants.

Du côté de la demande d'emploi seraient prises en compte des considérations d'ordre démosociologique (pyramide des âges , durée du travail, possibilité de travail féminin,...) et le maintien de certaines tendances économiques (exode rural,...).

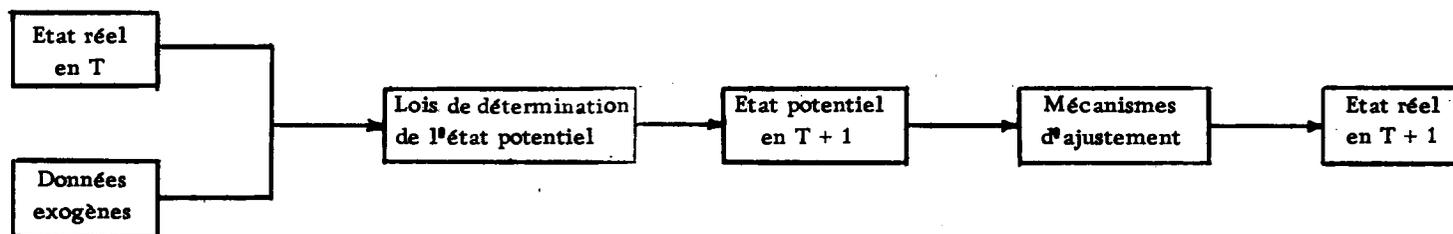
Cette conception d'un état potentiel basé sur *l'indépendance des décisions* nécessite que la période durant laquelle se crée l'état potentiel ne soit pas trop longue sous peine de faire perdre tout réalisme à ce concept (qui suppose en particulier que les décisions de début de période n'influent pas sur celles de fin de période).

1.12 - Mesure de l'état potentiel

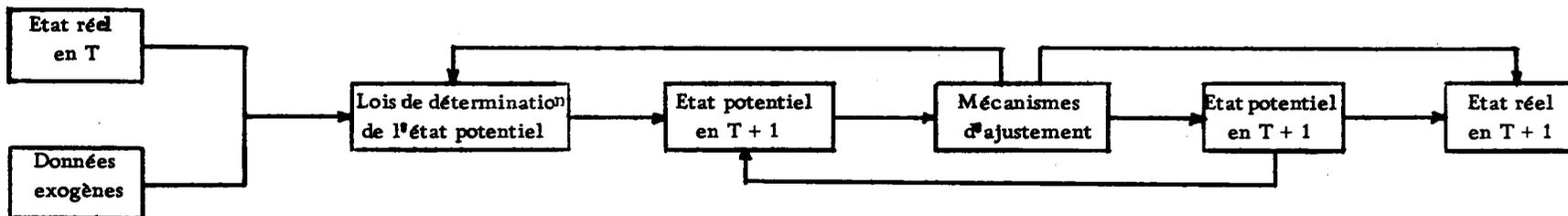
L'utilisation du concept de potentiel mène à la détermination de grandeurs qui, par définition, ne sont pas directement observables. Il est donc impossible de tester la justesse d'une grandeur potentielle. Il s'en suit qu'il n'est pas possible d'effectuer un choix parfaitement rationnel entre deux états potentiels possibles, obtenus par 2 méthodes différentes et que de ce fait il est *impossible de porter un jugement objectif* sur des données potentielles.

1.13 - Interaction entre les concepts d'état potentiel et de mécanismes d'ajustement

Le schéma logique de détermination d'un état d'emploi est le suivant :



Il est nécessaire de tester un tel mécanisme pour passer d'un modèle conceptuel à un modèle opérationnel. Il faut dès lors estimer les paramètres sur le passé. Le schéma de détermination se transforme par le jeu des mises au point inductives et déductives du modèle et devient :



Il est dès lors visible que la détermination d'un état potentiel ne peut être envisagée in abstracto mais dépend étroitement des mécanismes d'ajustement utilisés. On ne peut normalement porter de jugement que sur le doublet "état potentiel - mécanisme d'ajustement" et non sur l'une des deux composantes, abstraction faite de l'autre.

2.2 - La notion de mécanisme d'ajustement

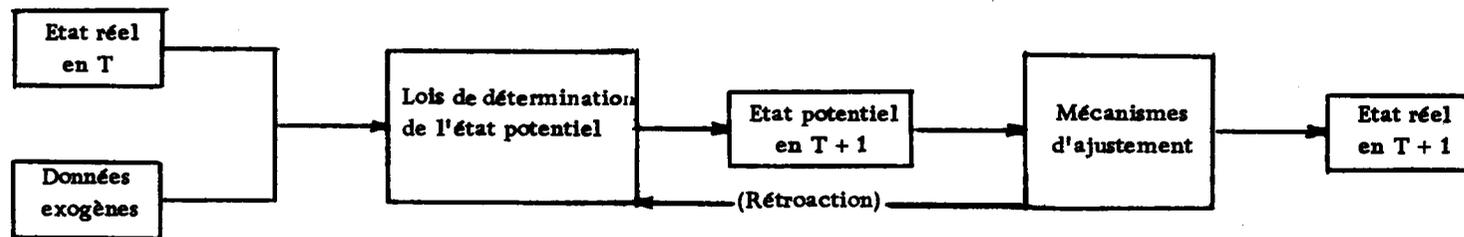
La conception de ces mécanismes d'ajustement affecte la nature même de la causalité entre sous-modèles aussi examinera-t-on en premier lieu les différentes conceptions de mécanisme d'ajustement face à ce problème de causalité entre sous-modèles. Il convient ensuite d'esquisser une démarche possible de recherche d'une typologie de mécanismes d'ajustement.

2.21 - Les différents types de mécanismes d'ajustement

Un mécanisme d'ajustement joue lorsque dans une zone donnée, le volume et la structure d'une offre potentielle ne correspondent pas au volume et à la structure d'une demande potentielle. Deux familles de mécanisme d'ajustement peuvent coexister :

2.211 - Les rétroactions

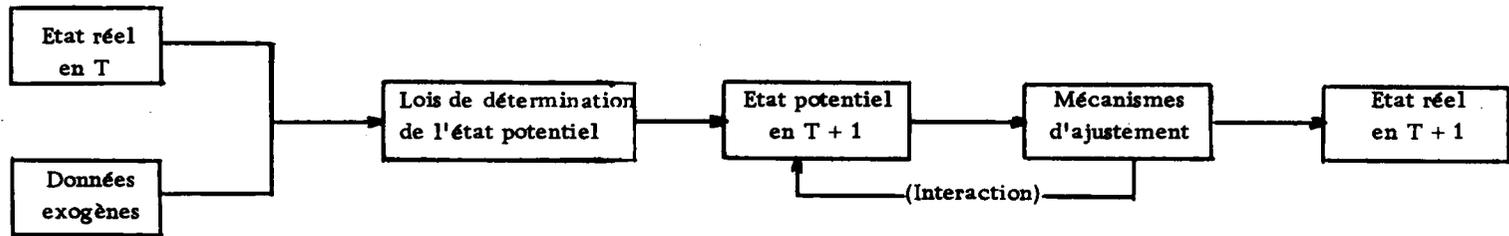
Il est tout d'abord possible d'envisager des mécanismes de rétroaction en entendant par rétroaction une modification des méthodes (engendrée au cours d'une itération précédente) d'obtention de l'état potentiel, en particulier des paramètres utilisés dans la détermination de l'état potentiel ou l'introduction de nouvelles valeurs de variables exogènes (en particulier instrumentales). Il y a donc le bouclage suivant (le nombre d'itérations pouvant être variable mais pas trop élevé sous peine d'ôter tout intérêt à la notion de potentiel) :



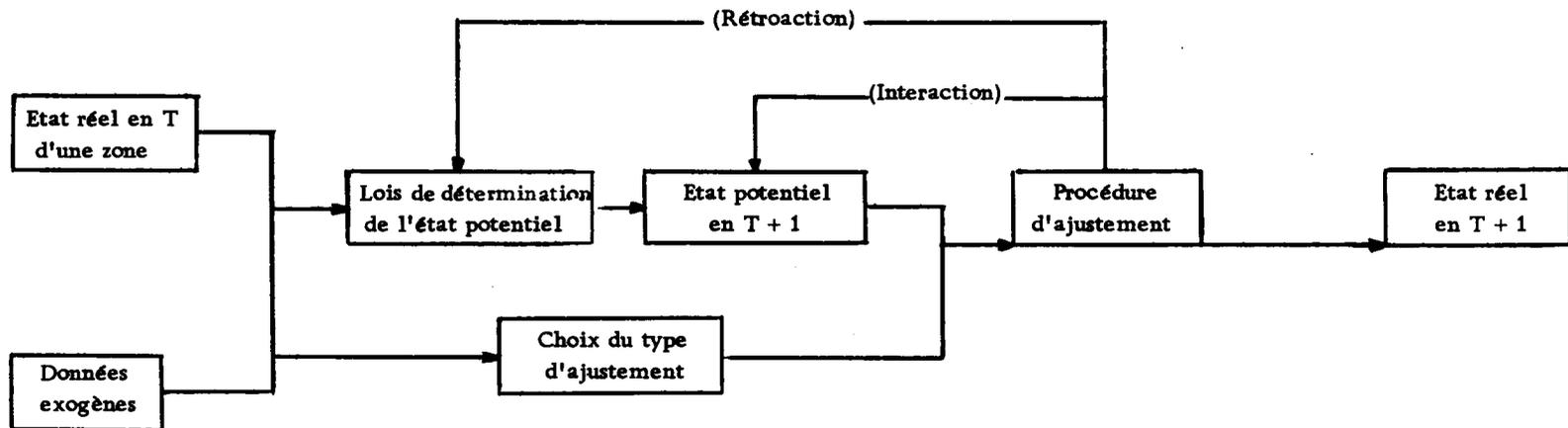
Les mécanismes de rétroaction n'entraînent pas nécessairement la cohérence interne de l'état potentiel et donc peut ne constituer qu'une partie des mécanismes d'ajustement, l'autre faisant appel aux mécanismes d'interaction.

2.212 - Les interactions

La seconde classe de mécanismes d'ajustement que l'on appellera mécanismes d'interaction consiste en une modification de données de l'état potentiel sans modification des lois de détermination de cet état potentiel.



En définitive la mise en jeu de ces deux types de mécanismes complémentaires mène au schéma de fonctionnement suivant où chaque boucle entraîne la définition d'un nouvel état potentiel (et où le nombre d'itération n'est pas fixé)



2.22 - Recherche d'une typologie des mécanismes d'ajustement

La recherche de mécanismes d'ajustement peut se concevoir de deux façons.

Il est tout d'abord possible d'envisager des mécanismes basés sur des considérations d'ordre théorique et qui visent à faire converger des résultats incompatibles au premier abord. Cette conception de "force de rappel" automatique est fréquemment utilisée dans un certain nombre de modèles macro-économiques complexes pour éviter que des évolutions calculées indépendamment pour plusieurs variables n'aboutissent au non respect de certaines identités.

Les développements exposés dans les pages 121 à 124 illustrent bien cette conception de mécanismes d'ajustement susceptible d'être mis en oeuvre.

Une seconde méthode, non concurrente de la première, est envisageable. On peut comparer à une date donnée un état potentiel calculé à l'aide d'un modèle défini sur des données du passé à un état réel observé pour cette même date. C'est d'ailleurs cette démarche qui est suggérée dans le chapitre introductif. Une analyse des écarts observés doit permettre d'"imaginer" des mécanismes de passage de l'état potentiel à l'état observé et d'en estimer les paramètres.

On peut supposer par exemple que pour une zone donnée, les causes principales de l'écart potentiel-réel soient les suivantes : le nombre d'agriculteurs a été sur-évalué, le taux d'activité sur-évalué, le taux d'émigration sous-évalué et les emplois non qualifiés du secondaire sur-évalué. On peut avoir eu en effet le type d'ajustement suivant : les créations d'emploi sur la période ont nécessité des ouvriers qualifiés et entraîné un décroissement des emplois non qualifiés, occupés en fait par des femmes d'agriculteurs ; ces dernières ne percevant pas ce revenu supplémentaire qui leur permettrait de survivre sur leur exploitation marginale, ont dû partir en plus grand nombre, cet exode n'a pu aboutir que hors de cette zone (qui n'offrait pas d'emplois non qualifiés), et comme dans cette zone le taux d'activité de la population agricole est supérieur au taux d'emploi global du département et qu'il n'y a pas que les femmes d'agriculteurs à avoir été touchées par cette crise, le taux d'emploi global en fin de période est inférieur à celui prévu.

Ce genre d'analyse doit être menée pour chaque zone et logiquement une typologie des ajustements doit se dégager. Chaque type d'ajustement doit correspondre à un ensemble bien particulier de caractéristiques de la zone : dans l'exemple imaginé, ce type d'ajustement pourrait intervenir lorsque la zone est caractérisée par un fort pourcentage d'exploitations agricoles marginales, la présence de petites industries en déclin (du type bonneterie) faisant appel à une main d'oeuvre non qualifiée essentiellement rurale et qu'au cours de la période, les emplois créés soient de type qualifié et que les suppressions d'emplois affectent les emplois non qualifiés.

2.3 - Problèmes posés par les délais d'ajustement

On examinera successivement l'incidence de la connaissance statistique des phénomènes et les problèmes posés par l'option possible du choix d'une période ne correspondant pas sur le passé à des dates d'observation des phénomènes.

2.31 - Connaissance des phénomènes et choix de la période

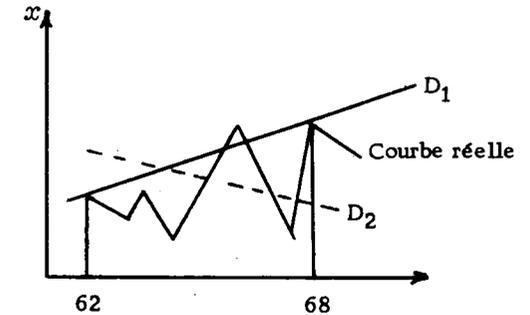
Il faut souligner tout d'abord que la situation réelle n'est connue en France qu'à des dates espacées du temps (recensements) et que bon nombre d'informations nécessaires, en particulier celles relatives aux flux, manquent ou ne sont que partiellement connues et que cette durée intercensitaire est importante (8 ans et 6 ans respectivement pour les trois derniers recensements). Ce problème de connaissance phénoménologique a deux incidences :

α) Incidence sur la significativité de l'état potentiel

Il est bien certain que la notion de potentiel est d'autant plus délicate à utiliser que la période retenue est longue, en effet les mécanismes permettant de résoudre les tensions deviennent alors de plus en plus difficile à déterminer et leur signification de moins en moins réaliste. Puisque l'économiste n'est pas maître du choix de la période (ce qui permettrait peut être de trouver un compromis optimal) on peut envisager de simuler un état réel intermédiaire (voir § 2.32).

β) Difficultés de préhension d'états intermédiaires satisfaisants

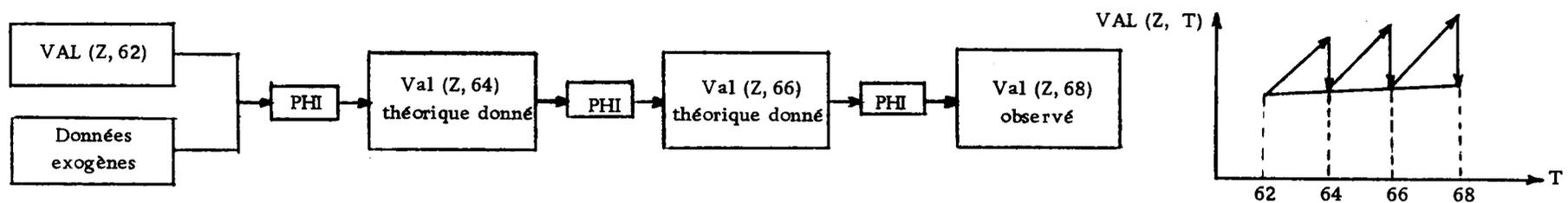
Le phénomène réel évolue dans le temps. Si l'on fait abstraction des variations saisonnières (par exemple en effectuant la saisie des données à des dates jugées comparables), il n'en reste pas moins que les observations faites peuvent correspondre à deux points de conjoncture radicalement opposés, l'élimination de fluctuation à court terme n'entraînant pas celle des fluctuations à moyen ou long terme. Le graphique ci-contre illustre d'ailleurs ce propos : les coefficients angulaires des droites d'interpolation (D_1) et de moindres carrés (D_2) étant de signes contraires. La droite d'interpolation constitue donc une approximation très grossière du réel que l'existence de trois observations ne permet pas d'améliorer. Un point réel théorique situé sur la droite d'interpolation ne semble donc pas a priori plus fondé qu'un autre (1).



2.32 - Problèmes posés par le choix d'un délai d'ajustement inférieur à la période intercensitaire

Pour faciliter l'exposé on utilisera un "opérateur" permettant le calcul d'un état potentiel et doté de mécanisme d'ajustement et que l'on appellera PHI. On peut supposer qu'il faille expliquer l'évolution entre 1962 et 1968 grâce à l'opérateur PHI et en se donnant 3 états potentiels, ou ce qui revient au même, en se fixant une période de 2 ans pour engendrer un état potentiel. Deux conceptions de PHI sont envisageables :

- α) L'état potentiel est calculé sur la période (2 ans), et les mécanismes d'ajustement ont pour contrainte de retrouver un état réel théorique donné (par exemple, par interpolation linéaire). On désignera par VAL (Z,T) un ensemble de données caractérisant la zone Z à une date T. On a alors le schéma suivant :



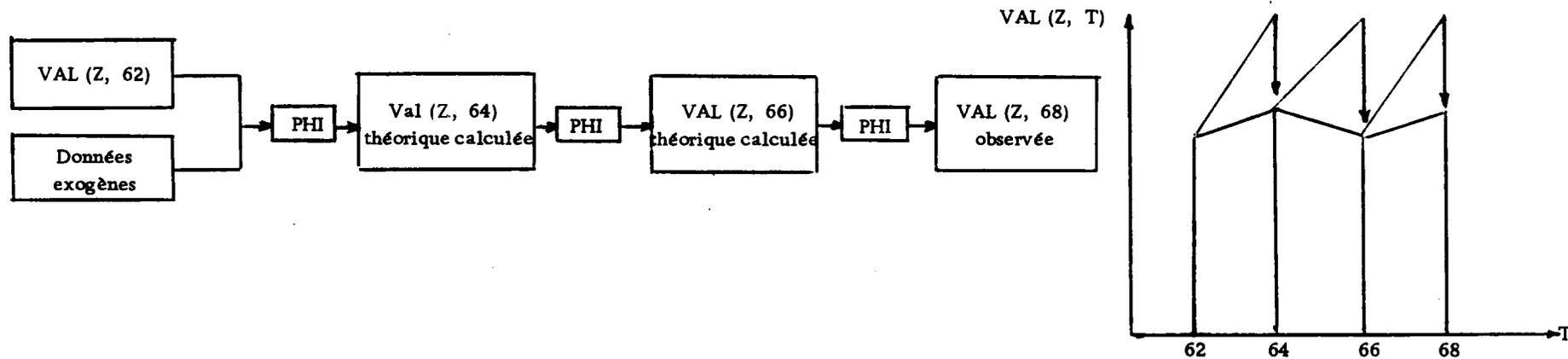
- (1) Certaines données peuvent a priori être considérées comme des proxis satisfaisants de l'évolution du phénomène auquel on s'attache mais les résultats obtenus peuvent s'avérer incohérents. A titre d'exemple nous avons tenté, en vain, avec M. ROCHEFORT du BETURE d'approximer des flux annuels de migrants intérieurs entre 1962 et 1968. Il n'a pas été jugé utile de rapporter cette étude dans le présent document mais il semble intéressant d'en indiquer la ligne directrice. La connaissance de l'évolution annuelle des statistiques scolaires des 6 - 14 ans (scolarité obligatoire) peut être considérée comme reflétant correctement l'évolution de cette classe d'âge de la pyramide ; l'application à la pyramide des âges de 1962 des tables de mortalité permet de connaître l'évolution naturelle hors migration de cette classe d'âge. La différence entre les taux d'évolution est donc normalement imputable aux phénomènes migratoires (dont la structure par âge pour la période 62-68 est parfaitement connue). Mais pour poursuivre plus loin (stabilité postulée de certains coefficients) il est nécessaire que le taux d'évolution annuel du solde migratoire de la classe d'âge ainsi calculé, soit constamment de même signe à moins de supposer que la répartition dans le temps des départs comme des arrivées se soit effectué par "vague" ce qu'il est impossible de connaître avec quelque exactitude. Sur les deux départements "testés" l'Isère et la Loire l'un d'entre eux ne présentait pas ce caractère.

β) L'état potentiel est calculé sur la période (2 ans) et les mécanismes d'ajustement ne sont pas soumis à des contraintes, sauf pour la dernière période (c'est-à-dire que la triple utilisation de l'opérateur PHI doit nécessairement aboutir à l'état observé pour 68).

Tout au plus fixerait-on un test obligeant un état théorique intermédiaire à se trouver dans une fourchette donnée à priori (en fonction par exemple de l'écart relatif maximum observé sur toutes les zones entre 1962 et 1968).

Mais il importe de remarquer qu'il s'agit d'un garde fou contre des états théoriques aberrants engendrés par PHI mais qu'en toute théorie il n'y en aurait pas besoin. Cette remarque en amène une autre : il est pour le moins délicat d'utiliser un point intermédiaire théorique comme estimation du réel : ceci n'est pas très grave pour le passé mais pour l'avenir peut poser des problèmes : en effet rien ne prouve qu'un tel opérateur n'engendrera pas de mouvements sinusoïdaux s'amplifiant au cours du temps.

Nous avons dans ce cas le schéma :



Il apparait donc, à la lumière de ces quelques remarques que le choix d'une période inférieure à la période intercensitaire, entraîne de graves difficultés quant au choix des mécanismes d'ajustement et qu'il soit préférable, au moins dans une première étape de se contenter de périodes de 6 ou 8 ans, malgré la moins grande "plausibilité" de l'hypothèse d'indépendance des décisions.

SECTION 3 - Les GRANDES OPTIONS du MODELE

Comme on a pu l'indiquer dans l'introduction la présentation de l'articulation entre les divers sous-modèles retenus est le résultat de nombreuses itérations et ne fait que traduire une étape empirique de recherche d'une architecture, à laquelle on s'est arbitrairement arrêté. Toutefois cet arbitraire a été limité, autant que faire ce peut, par la prise en compte de toutes les considérations développées dans les précédents chapitres. En outre il a été tenu compte de certaines réflexions développées par le groupe SESAME de la DATAR pour la détermination d'un modèle et pour lequel la présente étude ne constituait à l'origine que l'une des études de reconnaissance.

On exposera tout d'abord le clivage retenu en matière d'emploi avant d'examiner les grandes options retenues.

3.1 - Le clivage entre emplois fondamentaux et emplois induits

En s'appuyant sur les considérations théoriques développées au chapitre 4, le groupe Modèles du SESAME a décidé en 1970 de dissocier les emplois offerts en deux catégories : les emplois fondamentaux et les emplois induits. Cette dichotomie présente l'avantage, dans la mesure où elle est théoriquement fondée pour le niveau de découpage retenu, de dissocier assez nettement des emplois sur lesquels l'état peut par diverses incitations agir, d'emplois pour lesquels la part explicative de l'évolution du contexte socioéconomique est prépondérante et pour lesquels des actions d'inflexion sont difficilement envisageables.

Les emplois fondamentaux-type sont les emplois du secondaire, tandis que les emplois induits type sont ceux du tertiaire. La détermination des effectifs de BTP se rapproche de celle des emplois induits si l'on considère que ses données de production sont connues sans erreur. Enfin le secteur agricole est très particulier puisque c'est le seul secteur où par définition l'offre et la demande d'emplois coïncident.

Il a été décidé de considérer comme déterminés exogènement les effectifs du secondaire dans le cadre de cette étude. Toutefois nous avons entrepris avec M. HORPS, D. MALKIN une étude sur les facteurs explicatifs d'évolution du secteur secondaire à partir de données issues du traitement des fichiers entreprises et établissements pour la période 1961 - 1969 donnant les évolutions annuelles des effectifs par branche des grands établissements, en tenant compte de la localisation du centre de décision. Cette étude sera présentée dans le chapitre 13.

La détermination de l'état potentiel des autres variables d'emploi relève de la présente étude.

3.2 - Les autres options

On exposera sans ordre privilégié les principales options de principe retenues :

- a) Il convient tout d'abord de rappeler que dans l'objectif que l'on s'est fixé les variables endogènes sont exclusivement des variables de population et que la détermination des diverses variables exogènes est supposée faite par d'autres modèles donnant sans erreur les données voulues (en pratique on prendra toujours les variables observées, sauf s'il s'agit de données de population pouvant être fournies par l'un des sous-modèles retenus). Cette option permet de mieux cerner la difficulté dans l'analyse des écarts qui doit conduire à la définition d'une typologie des mécanismes d'ajustement.
- b) Les divers sous-modèles spatialisés que l'on déterminera fourniront des résultats qui seront confrontés à ceux fournis par un modèle national approprié et un "cadrage" sur cette donnée nationale sera effectué si on le juge nécessaire dans chacun des sous-modèles pour la recherche des typologies d'ajustement.

- c) On recherchera dans la mesure du possible des sous-modèles sur la base de "références théoriques" dont on soulignera les diverses hypothèses, mais il ne peut être question ici de rechercher l'adéquation de modèles théoriques mathématiquement définis et logiquement cohérents entre eux, à la réalité. En effet, un tel effort dépasse très largement le cadre de cette étude car il suppose la mise en place préalable pour chaque "secteur" de modèles conceptuels très élaborés alors que bien souvent la connaissance phénoménologique de base est légitime voire inexistante, sans compter qu'il est fort à parier que de tels modèles ne pourraient être "nourris" faute de données. Il s'agit plutôt ici de rechercher sur le passé un certain nombre de régularités à l'aide de variables exogènes dont beaucoup ne peuvent être que des indicateurs grossiers de facteurs intervenant dans la réalité. Il s'agit dans le fond, dans certains cas, d'une ébauche du premier cycle du processus "inductif - déductif" propre à toute recherche appliquée. On aura donc une utilisation en mode prévisionnel de modèles descriptifs.
- d) Ces modèles trouvés ne pouvant être considérés comme de véritables modèles explicatifs leur horizon est forcément limité au moyen terme (ce qui, soit dit en passant, est une limitation de l'objectif premier assigné au groupe Modèles de SESAME). Cette période retenue est celle du recensement afin d'éviter (au moins dans un premier temps) les problèmes soulevés au § 2.3.
- e) Les sous-modèles s'articuleront pour des raisons de commodité par l'intermédiaire de système de Cramer de causalité ce qui n'est pas très grave dans la mesure où les phénomènes d'interdépendance sont reportés au niveau des mécanismes d'ajustement. Par ailleurs les mécanismes d'ajustement seront, dans un premier temps, du type "interaction" et non "rétroaction".
- f) Enfin pour mémoire on rappellera que des options fondamentales ont été prises en ce qui concerne la conception du marché de l'emploi et de la séparation entre emplois fondamentaux et emplois induits.

Le choix du découpage spatial

6

CHAPITRE 6 - LE CHOIX DU DECOUPAGE SPATIAL

On a pu examiner, au cours du chapitre 2, diverses techniques possibles permettant à partir d'unités spatiales de base de constituer une partition de l'espace en fonction de certains critères. Les critères que l'on va retenir ici, conformément à ce qui a été vu au chapitre 5, seront supposés représentatifs d'un marché de l'emploi si toutefois celui-ci existe. Le groupe d'étude "Modèle" du SESAME (Système d'Etude du schéma d'aménagement du territoire) de la DATAR s'est d'abord penché sur ce problème de choix de découpage spatial.

Deux études⁽¹⁾ particulières ont été faites par MM. Delaunay et Malkin, études auxquelles nous avons participé dans la phase de réflexion critique des méthodes retenues. Une troisième étude plus ambitieuse, s'inspirant des principes développés au § 2.11 b du chapitre 2 et portant sur des données de la région Nord, n'a pu être menée à son terme par nous pour des raisons informatiques.⁽²⁾

On présentera donc l'essentiel de ces deux premières études avant d'examiner dans une seconde section les implications du choix du découpage départemental que l'on a finalement été amené à retenir. Cette présentation est par certains côtés plus approfondie que les rapports originaux : on s'est efforcé ici d'explicitier toutes les hypothèses sous jacentes et de mesurer de façon précise la portée des conclusions auxquelles on aboutit.

(1) Voir les rapports suivants :

"Découpage spatial et marché de l'emploi - Approche pour les migrations définitives" par D. Delaunay et D. Malkin mars 1970. Ainsi que

"Découpage spatial et marché de l'emploi - Une approche statistique" par D. Malkin septembre 1970

(2) Programmes d'exploitation écrits mais bande magnétique partiellement illisible à la suite de détériorations.

SECTION I - TRAVAUX EFFECTUES DANS LE CADRE DU SESAME (1)

Deux types d'approche ont été successivement tentées pour essayer de "juger", autant que faire se peut, différents découpages possibles et en particulier le découpage départemental. La première étude est basée sur l'étude empirique d'un indicateur, la seconde sur celle d'un ensemble d'indicateurs.

1.1 - Etude d'un indicateur

Il faut tout d'abord décrire les fondements de l'étude avant d'en analyser les résultats.

1.11 - Fondements de l'étude

a) Justification du choix de l'indicateur migration définitive

Certains découpages existants tel que celui des ZEDE sont basés en partie sur l'étude de migrations alternantes. Mais le marché de l'emploi de ces ZEDE, pour prendre cet exemple, en fait déborde une ZEDE donnée et porte sur une zone constituée de ZEDE agrégées et pour laquelle il existe une transparence satisfaisante des conditions d'offres et de demandes d'emploi.

Dans ces zones il est dès lors vraisemblable que des migrations définitives se substituent aux migrations alternantes dans un certain nombre de cas où ces dernières sont difficilement envisageables. Cette hypothèse implique d'une part que l'on admette que l'essentiel des migrations soient liées à l'emploi (ce qui est contestable comme on l'a vu dans le chapitre 4) et d'autre part que le coût du changement ne soit pas perçu économiquement et psychologiquement comme si important, qu'il soit un frein puissant à la mobilité spatiale dans la zone ; cette dernière considération mène à l'élimination des zones de trop grande dimension.

(1) voir références note page précédente : l'ordre de rapports correspond à celui des études développées dans chacune des sous sections.

Si pour un territoire donné un tel découpage en marchés de l'emploi hermétiques, existe, la totalité des migrants originaires de chaque zone se retrouveront dans une autre localité de leur zone de départ en fin de période. On dira alors qu'il y a rétention complète des mouvements migratoires par les différents marchés d'emploi.

Par contre, si aucun marché d'emploi n'existe dans le territoire on aura alors un mouvement brownien des migrations et une rétention plus ou moins importante des migrants par chaque zone.(1)

Il est en fait vraisemblable que les marchés hermétiquement fermés n'existent pas(2) et qu'il faille trouver empiriquement une cote mal-taillée.

b) Données utilisées

Les données de base de cette étude sont tirées d'une étude de CERES concernant la région Nord portant sur 44 zones d'études résultant du groupement de 433 zones élémentaires (données et découpage tirés du tableau migratoire MI 1/TE 3 du recensement de 1962). Ces zones d'études ont été constituées par agrégations successives en cherchant à chaque pas à grouper les deux zones dont l'intensité des échanges est maximum. Le découpage qui en résulte ne permet pas de retrouver les ZEDE, aussi c'est le découpage le plus proche de celui en ZEDE qui a été utilisé lorsque cela a été nécessaire.

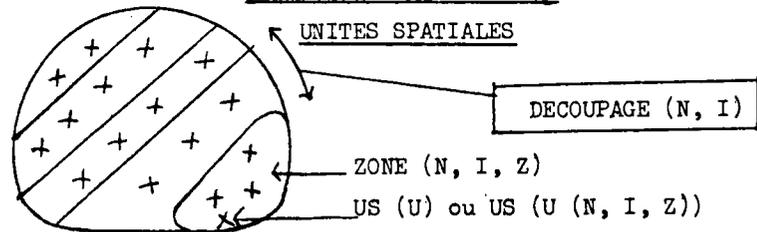
c) Le modèle théorique de référence

On trouvera page suivante l'ensemble des concepts et des variables que l'on utilisera.

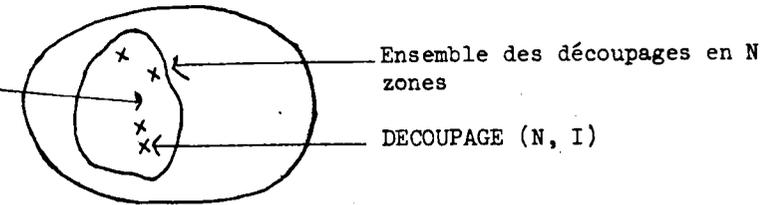
α - Le coefficient de rétention d'une zone peut varier entre 0 et 1. Il s'en suit qu'avec ce système de pondération adopté pour le coefficient de rétention du découpage, ce dernier peut lui aussi varier entre 0 et 1, et ceci quel que soit l'argument N du découpage. En pratique l'intervalle de variation ne sera pas aussi étendu : si l'un des découpages en N zones correspond à une partition en marchés d'emploi hermétiques, chaque coefficient de rétention de zone valant 1, le coefficient de rétention du découpage est donc égal à 1. Si ce cas de figure n'existe pas,

(1) ceci ne présume pas de l'intensité des échanges (voir les divers modèles migratoires du chapitre 4) puisque cette rétention s'exprime par un rapport de flux.

(2) aux problèmes de transparence du marché près, on pourrait retomber sur une application de modèle du type Stouffer avec forte transparence du marché de la zone d'émission et transparence imparfaite, ou coût de mutation élevé pour les zones de réception.

ENSEMBLE FONDAMENTAL Ω 

ENSEMBLE FONDAMENTAL DES DECOUPAGES POSSIBLES



ENSEMBLES

DEFINITION DES CONCEPTS SPATIAUX

US (U) = Unité spatiale élémentaire repérée par l'indice U ; $U \in [1,433]$ ou par son appartenance à une zone d'un zonage : $US (U(N, I, Z))$ où U (N, I, Z) repère le numéro d'ordre de l'U S dans la zone

ZONE (N, I, Z) = sous-ensemble constitué d'u.s. dont chacune est contigue à au moins une autre u.s. appartenant à la zone. Une zone est repérée par un numéro d'ordre Z dans un découpage donné DECOUPAGE (N,I); à chaque zone est associée une mesure appelée RETENTION (N,I,Z) et une mesure appelée ATTRACTION (N,I,Z)

DECOUPAGE (N,I) = partition de l'ensemble fondamental des u.s. en N sous-ensembles (= zones) disjoints ($N \in [1,44]$) dont le numéro d'ordre est I⁽¹⁾

MEILLEUR DECOUPAGE (N) est l'un (I = I') des découpages existant DECOUPAGE (N,I') et qui correspond à celui qui a la valeur la plus élevée possible de l'indicateur associé K (N,I)

Ensemble de zonages de taille N

DECOUPAGE (N,I); à ce découpage est associé un indicateur que l'on appelle coefficient de rétention du découpage et que l'on désigne par K (N,I)

DEFINITION DES VARIABLES

MIG (U,U*) = effectifs de migrants en provenance de U et à destination de U* (avec $U \neq U^*$). On peut encore écrire cette variable sous la forme $MIG (U (N,I,Z), U^* (N,I,Z'))$ où l'on peut avoir $Z' = Z(2)$. Par convention $MIG (U^*, U^*) = 0$

MIG (Z,Z*) = effectifs de migrants en provenance de Z et à destination de Z*. Si Z comprend au moins 2 unités spatiales on a : $MIG(Z^*, Z^*) \neq 0$. La relation de cette variable avec la précédente s'exprime par la relation suivante :

$$MIG (Z,Z^*) = \sum_{U(Z,N,I)} \sum_{U^*(Z^*,N,I)} MIG (U(Z,N,I), U^*(Z^*,N,I))$$

RETENTION (N,I,Z*) = $\frac{MIG (Z^*, Z^*)}{\sum_Z MIG (Z^*, Z)}$ Il s'agit donc de la part des migrants restés dans la zone, cet indicateur peut varier entre 0 et 1.

$$K (N,I) = \sum_{Z^*} RETENTION (N,I,Z^*) \cdot \frac{POPULATION (N,I,Z^*)}{\sum_{Z^*} POPULATION (N,I,Z^*)}$$

Ce coefficient de rétention du découpage est donc la moyenne des coefficients de rétention des zones, pondérée par la part de la zone dans la population totale. Le découpage pour lequel K (N,I) est maximum est appelé meilleur découpage.

(1) son domaine de variation est fonction de N et est difficilement calculable, voir chapitre 2

(2) seule la référence au même zonage présente un intérêt.

la borne supérieure n'est pas atteinte. Quant à la borne inférieure qui correspond à une rétention nulle pour toutes les zones, elle n'a que peu de chances d'être atteinte sauf pour le cas limite du découpage redonnant les unités spatiales de base. Que peut-on dire maintenant de l'évolution de cette fourchette en fonction de N ? S'il existe un découpage en marché optimal d'emplois hermétiques en N^* zones, il existe pour chaque découpage en un nombre inférieur de zones, au moins un découpage qui donne un coefficient de rétention de découpage égal lui aussi à 1 et qui est obtenu par agrégation de deux (ou plusieurs) zones du découpage optimal. Pour $N > N^*$ on a obligatoirement des échanges migratoires entre zones et donc ce coefficient est inférieur à 1. Peut-on alors penser que la décroissance de la borne supérieure de l'intervalle sera monotone. On ne peut l'affirmer sans introduire d'hypothèses complémentaires (en particulier sur les taux de migration, qui relie coefficients de pondération et de rétention).

En ce qui concerne l'évolution de la borne inférieure on ne peut pas davantage affirmer que sa décroissance est continue. Mais ces limites ne sont pas trop gênantes. En effet, en pratique les cas limites de marché hermétique et de mouvements browniens ne se rencontrent pas. Par ailleurs, l'ensemble des découpages d'argument N n'est pas intéressant à étudier si l'on doit tenir compte des remarques faites au 1.1) a) sur la dimension des zones. On a donc recherché une courbe empirique "satisfaisante" considérée comme borne supérieure opérationnelle (avec N en abscisse et $K(N)$ en ordonnée). On a retenu pour la construction de cette courbe le coefficient de rétention de découpage le plus élevé parmi les quelques découpages "réalistes" en N zones du territoire.

Cette courbe empirique peut avoir deux utilisations.

Si l'on observe un palier dans la décroissance, on peut penser que, si les découpages réalistes ont été bien choisis, on est en présence d'une certaine polarisation des marchés d'emploi. Le meilleur découpage pour les valeurs de N correspondant au palier observé est celui pour lequel N est le plus élevé (ce que l'on trouve en reprenant le raisonnement suivi pour le cas théorique des marchés).

En second lieu, elle peut servir de référence pour juger un découpage existant. Comme cette courbe est sans doute assez proche de celle de la borne supérieure, il est peu probable que le coefficient de rétention du découpage à tester soit au dessus de la courbe. Mais si tel est le cas, on peut alors modifier la courbe de référence et faire apparaître un palier, ce qui ramène à l'observation du paragraphe précédent. Si l'on se trouve en dessous de la courbe, on peut penser que le découpage testé coupe des zones de marché d'emploi et que donc ce découpage est moins bon. Enfin si le point trouvé est à peu près sur la courbe, on ne peut guère que conclure que ce découpage n'est ni meilleur ni plus mauvais que celui de référence et que cela est d'autant plus satisfaisant que la décroissance de $K(N)$ est moins rapide.

...

β - Un autre indicateur est proposé pour cette étude, il s'agit du coefficient d'attraction d'une zone. Il s'agit de la part de migrants intérieurs à la Région Nord se trouvant dans la zone considérée. On peut calculer un coefficient d'attraction de découpage comme une somme pondérée de ces coefficients de zones. Le fait que le dénominateur du coefficient de zone soit constant quelle que soit la zone, facilite l'étude de la variation du coefficient d'attraction du découpage. Mais l'interprétation de cet indicateur est délicate car il décroît même s'il existe un découpage correspondant à des marchés hermétiques. En conséquence les résultats de cette étude ne sont pas rapportés ici.

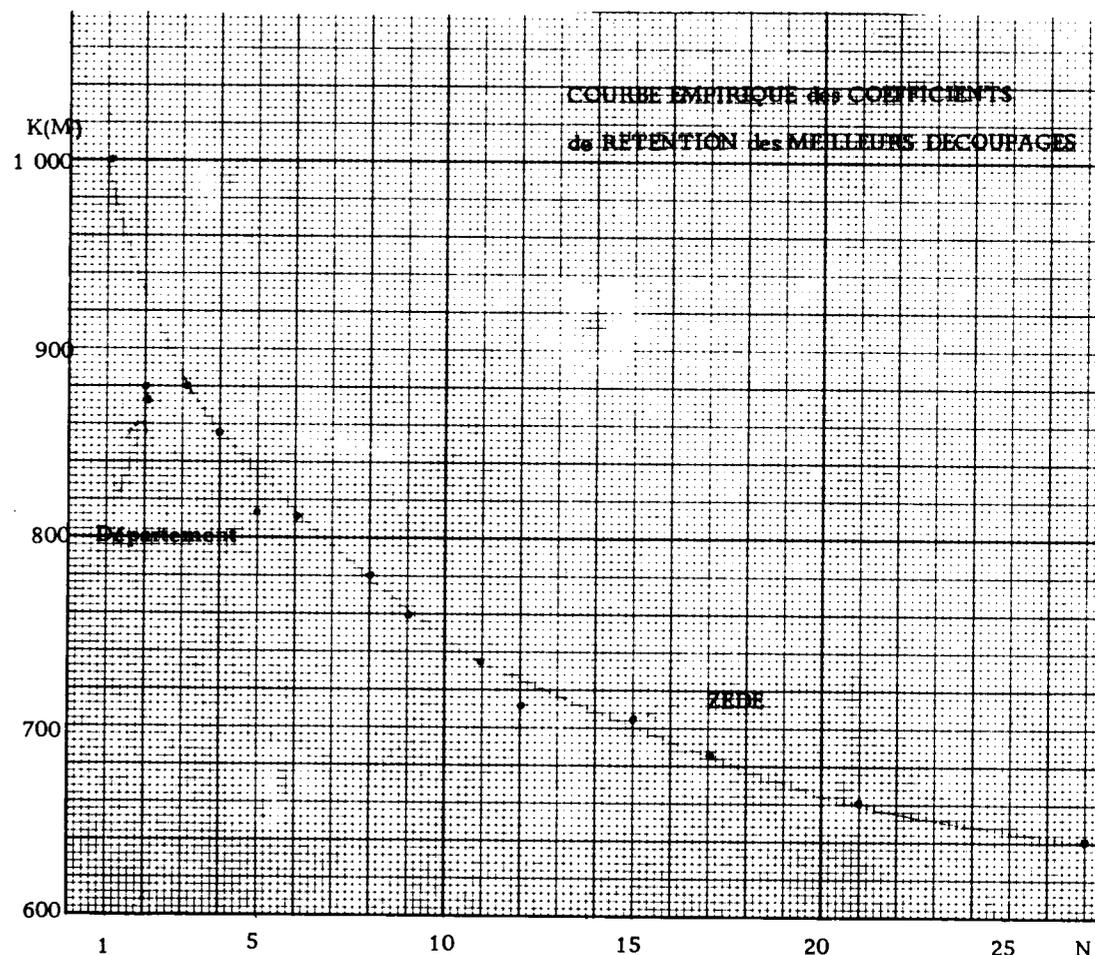
1.12 - Analyse des résultats

a) présentation des résultats

Les divers calculs effectués ont permis de tracer le graphique ci-contre.

La décroissance de la courbe est régulière ce qui laisse présumer que le choix des zones "réalistes" est satisfaisant dans l'ensemble. En effet un mauvais choix pour plusieurs arguments M de découpage se traduirait par une évolution en dents de scie.

A la lumière des considérations développées au 1.11 c, on peut penser que le découpage en deux zones (et donc le découpage départemental) n'est pas satisfaisant. Quant au niveau ZEDE, il n'est pas particulièrement privilégié.



b) critique des résultats

Ces résultats sont sans doute difficilement généralisables en raison de l'empirisme de la démarche suivie et des nombreuses hypothèses restrictives que l'on a été amené à faire : migrations exclusivement liées à l'emploi, approximation du découpage en ZEDE, incidence du système de pondération sur la sensibilité des résultats. En ce qui concerne le rejet du niveau départemental il est sans doute plus difficile de généraliser à partir de l'étude de deux départements seulement. La seconde étude basée sur une autre région et un autre type d'approche visait à lever certaines de ces objections.

1.2 - Etude de plusieurs indicateurs

On examinera successivement les bases de cette étude dont le but est non la constitution d'un découpage optimal mais la comparaison de deux découpages, puis ses résultats.

1.21 - Bases de l'étude

Les problèmes posés par l'utilisation de plusieurs indicateurs ont été vus au chapitre 2. En ce qui concerne la technique particulière utilisée ici on se reportera au § 2.12 du chapitre 2.

L'application de cette technique au problème posé suppose que plusieurs conditions soient remplies, ce que l'on examinera avant de donner les caractéristiques des données utilisées et d'en examiner l'incidence sur les découpages à comparer.

a) critères de choix des données à utiliser

Ces données doivent présenter deux caractéristiques, la première résulte de la recherche d'un découpage adéquat à la notion de marché de l'emploi, la seconde du type d'analyse.

α - Les données doivent concerner certaines procédures caractéristiques de résolution de tensions sur le marché de l'emploi. A ce propos deux remarques sont à faire ; il n'est pas nécessaire de disposer de renseignements relatifs à toutes ces procédures, puisqu'il suffit que pour un découpage donné les variables se comportent différemment dans les unités spatiales constitutives de chaque zone pour que l'on puisse rejeter l'hypothèse de pertinence du découpage à la notion qui nous intéresse.

D'autre part, si la notion de processus d'ajustement est par essence dynamique, il n'est pas pour autant nécessaire de disposer de données dynamiques. En effet si la zone est un marché de l'emploi homogène, le type de processus d'ajustement sera le même pour toutes les unités spatiales constitutives de la zone et en particulier la résolution des tensions doit se traduire par un type d'équilibre homogène pour l'ensemble des unités spatiales. Il suffit donc que les indicateurs testés reflètent certains types d'équilibre caractéristiques résultant de ces processus d'ajustement.

β - Ces données doivent être disponibles à un niveau spatial assez fin. En effet, la technique suppose que chaque zone constitutive d'un découpage donné soit constituée d'un nombre suffisant d'unités spatiales pour que les conclusions soient assez nettes.

b) Données de base de cette étude

Des données présentant les caractéristiques souhaitées étaient disponibles pour la région Rhône - Alpes. En effet, le CERAU pour les besoins d'une de ses études⁽¹⁾ avait été amené à élaborer une batterie d'indicateurs pour chacun des 219 cantons parmi lesquels on a retenu les indicateurs suivants :

- taux de chômage des 20 - 65 ans en 1962
- taux d'activité global des 20 - 65 ans en 1962
- taux d'activité féminin des 20 - 65 ans en 1962
- coefficient de stabilité rurale = (actifs agricoles 62/actifs agricoles 54)

Les trois premiers indicateurs sont de nature statique et reflètent bien l'équilibre réalisé à une date donnée des diverses tensions du marché de l'emploi.

(1) "Maquette de développement régional" CERAU - 1970 - Ajoutons que la note méthodologique du CERAU, écrite en décembre 1969 par R. Prud'homme, a inspiré une partie de l'analyse de système faite par le groupe du SESAME.

Le dernier indicateur est de nature dynamique et est dans une certaine mesure significatif du résultat de deux mécanismes d'ajustement : la mutation sectorielle et la mise en retraite.(1)

Une raison complémentaire militait en faveur de l'utilisation de ces données pour une étude de découpage, c'est que du point de vue des structures d'emploi la région Rhône-Alpes est représentative de façon satisfaisante de celles de l'ensemble de la France(2).

c) Incidence des données disponibles sur le choix des découpages à comparer

Le but de cette étude est de comparer deux découpages qui semblent intéressants a priori : le découpage départemental et celui en bassin d'emploi. Les données disponibles ont nécessité l'adoption d'un découpage légèrement différent du découpage départemental.

Tout d'abord les bassins d'emploi (au nombre de 23) ont été choisis de préférence aux ZEDE. En effet, la notion de bassin d'emploi, définie conjointement par les services régionaux du Ministère du Travail et de l'INSEE à partir d'études d'armature urbaine, est centrée sur celle de zone d'influence et d'attraction d'une agglomération, pôle privilégié de migrations alternantes et donc coeur d'un marché local de l'emploi. Par ailleurs, les ZEDE à la différence des bassins d'emploi ne forment pas une partition de la Circonscription d'Action Régionale.

Le département n'étant pas une réunion de bassins d'emplois, il a fallu trouver un découpage proche à soumettre au test. Il faut à ce propos souligner que, puisque le découpage soumis est constitué de zones réputées homogènes du point de vue de l'emploi, une conclusion de rejet de ce découpage sera a fortiori valable pour le découpage départemental. La recherche du découpage à prendre en considération a mené à retenir 7 départements économiques au lieu de 8 départements administratifs, avec une fusion des départements de l'Ardèche et de la Drome (due au fait que les bassins d'emploi de Valence et de Montélimar sont à cheval sur le Rhône). Chacun de ces 7 départements comporte un nombre inégal de bassins d'emploi (de 2 à 4).

1.22 - Résultats de l'étude

L'analyse de variance a été utilisée pour tester le rejet des hypothèses de la pertinence des deux découpages face aux problèmes du marché de l'emploi. Comme il s'agit de découpages emboîtés, le découpage le plus agrégé doit être nécessairement pertinent pour que l'autre puisse l'être, dans le cadre de l'analyse de variance retenue(3).

(1) voir analyse faite au chapitre 11 pour l'aspect agricole de ce problème

(2) rapport cité p. 22

(3) s'il y a rejet du niveau agrégé on peut alors mener une analyse de variance de type classique sur la seule variable découpage fin.

Ces hypothèses sont :

- HD = les différences observées entre départements ne sont pas significatives, et donc l'introduction de cette partition des cantons n'amène aucun renseignement supplémentaire ;
- HZ = les différences observées entre bassins d'emploi d'un même département ne sont pas significatives et donc il y a homogénéité des cantons dans le département en ce qui concerne les problèmes d'emploi et le découpage en bassins d'emploi n'amène rien.

Du fait que les départements ne comportent pas le même nombre de bassins d'emploi, l'analyse de variance multivariante n'est pas possible. (1) Cette contrainte étant impossible à lever on ne peut que mener une analyse de variance pour chaque variable et regarder s'il y a homogénéité dans le rejet ou le non rejet des hypothèses testées, ou chercher à synthétiser l'ensemble des variables par une variable unique. Ces deux voies ont été successivement explorées.

a) comparaison des résultats de plusieurs analyses de variance

α - les résultats

L'équation type du modèle d'analyse de variance testée pour chacun des 4 indicateurs

$$X(D, Z, C) = K + KD(D) + KZ(D, Z) + E(D, Z, C)$$

- . où D est l'indice de département, Z l'indice du bassin dans le département D, et C l'indice du canton dans le bassin Z du département D ;
- . et où X est l'un des indicateurs testés, K une constante représentative de l'effet régional (Rhône Alpes). KD est représentatif de l'effet départemental, KZ de l'effet bassin d'emploi et E un terme aléatoire.

Les deux hypothèses à tester s'écrivent :

$$HD : KD(D) = 0 \quad \forall D$$

$$HZ : KZ(D, Z) = 0 \quad \forall Z \in D$$

Si HD est vérifiée, HZ est ramenable alors au test d'hypothèse nulle $KZ(Z) = 0 \quad \forall Z$ (c'est à dire que les degrés de liberté du F seront 22 et 218 au lieu de 15 et 218). Si elle ne l'est pas HZ est alors à interpréter par rapport à une référence départementale.

(1) voir chapitre 2 page 59

VARIABLES	D	K	KD (D)	KZ (D, 1)	KZ (D, 2)	KZ (D, 3)	KZ (D, 4)
taux de chômage des 20 - 65 ans	1	26,77	- 4,19	2,64	- 2,05	0,22	-
	2		2,45	4,40	- 5,31	0,19	- 3,23
	3		- 7,74	- 0,36	- 4,93	5,96	-
	4		- 5,67	- 3,92	3,89	- 0,60	-
	5		4,64	- 9,09	2,48	-	-
	6		0,09	5,53	- 5,86	- 2,36	-
	7		15,30	- 2,28	14,41	- 9,41	6,41
taux d'activité global des 20 - 65 ans	1	63,80	- 4,97	- 2,11	3,51	- 1,86	-
	2		0,17	- 1,50	1,10	0,82	0,54
	3		2,35	- 3,02	3,43	0,21	-
	4		1,66	4,81	- 5,09	1,36	-
	5		2,14	1,98	0,54	-	-
	6		2,61	- 3,85	3,06	2,84	-
	7		0,52	2,01	2,52	0,52	- 4,14
taux d'activité féminin des 20 - 65 ans	1	21,07	- 4,54	- 0,75	1,94	- 1,29	-
	2		0,29	- 1,99	0,88	0,77	2,34
	3		3,86	0,81	1,96	- 3,26	-
	4		2,20	4,36	- 4,27	0,55	-
	5		1,53	1,05	- 0,28	-	-
	6		0,79	- 3,66	2,13	3,63	-
	7		- 0,78	0,30	- 0,04	1,93	- 3,12
coefficient de stabilité rurale	1	67,40	- 5,87	0,61	- 0,37	- 0,04	-
	2		5,47	3,05	7,87	- 16,06	10,31
	3		0,56	- 2,23	6,38	- 4,11	-
	4		3,52	6,58	- 8,46	4,85	-
	5		- 2,62	8,18	- 2,23	-	-
	6		2,11	- 1,44	6,20	- 4,84	-
	7		- 7,15	2,89	- 9,89	5,76	- 4,46

VARIABLE	Niveau de découpage	Somme des carrés	degré de liberté	F calculé (1)
Taux de chômage des 20 - 65 ans	Département	10266,2	6	7,75 ⇒ rejet
	Bassin	4986,9	15	1,50 ⇒ acceptation
	Canton	48128,1	218	-
Taux global d'activité des 20 - 65 ans	Département	1857,2	6	11,26 ⇒ rejet
	Bassin	1683,9	15	4,08 ⇒ rejet
	Canton	5991,7	218	-
Taux d'activité féminin des 20 - 65 ans	Département	1797,5	6	12,78 ⇒ rejet
	Bassin	1166,4	15	3,32 ⇒ rejet
	Canton	5106,7	218	-
Coefficient de stabilité rurale	Département	5287,7	6	4,33 ⇒ rejet
	Bassin	9180,6	15	3,00 ⇒ rejet
	Canton	44367,5	218	-

(1) Pour 6 et 218 degrés de liberté et un risque de première espèce (rejet à tort de l'hypothèse nulle) de 5 %, la valeur critique de F est 2,14 ; et au risque de 1 % cette valeur est 2,90 ; Pour 15 et 218 degrés de liberté et un risque de 5 % la valeur critique de F est 1,72, et de 2,15 au risque 1 %.

L'examen des résultats du tableau de la page 11 montre qu'en ce qui concerne les trois derniers indicateurs la conclusion est nette : le découpage en départements ainsi que celui en bassins d'emploi dans le département sont des éléments importants à prendre en compte pour l'explication de disparités spatiales des indicateurs. Par contre, il semblerait qu'en ce qui concerne les taux de chômage dont le coefficient de variation est nettement plus élevé que pour les autres variables, le niveau départemental soit pertinent mais pas celui des bassins d'emploi, encore que le risque d'acceptation à tort de cette dernière hypothèse (risque de seconde espèce) n'est peut être pas négligeable.

Cette analyse implique que l'on puisse adopter à coup sûr, le niveau départemental et sans doute le niveau ZEDE.

β - critique de ces résultats

Il faut tout d'abord rappeler que les découpages testés ne sont pas exactement les découpages réels et qu'un biais est peut être de ce fait introduit. Il est toutefois vraisemblable que compte tenu des valeurs du F les conclusions resteraient les mêmes avec le découpage réel.

Le fait que les deux niveaux améliorent l'explication peut signifier que des mécanismes différents coexistent aux deux niveaux mais aussi qu'il existe peut être un niveau de découpage intermédiaire ou plus fin qui rende compte de façon satisfaisante de l'intégralité de ces distortions spatiales d'indicateurs. Enfin, l'hypothèse selon laquelle le terme résiduel suit une loi normale de moyenne nulle est contestable en cas de désajustements structurels sur le marché de l'emploi.

C'est donc avec beaucoup de prudence qu'il faut manier ces conclusions.

b) Analyse de variance d'un indicateur unique

Les unités de divers indicateurs (%) étant comparables et se rapportant au même phénomène d'ajustement sur le marché de l'emploi, la combinaison linéaire retenue est celle fournie par l'analyse en composantes principales

...

qui donne :

$$X(D, Z, C) = 0,755.TC(D, Z, C) + 0,038.TAG(D, Z, C) + 0,040.TAF(D, Z, C) + 0,653.STA(D, Z, C)$$

Cette première composante explique 54,8 % de la variance, ce qui est satisfaisant. L'analyse de variance donne les résultats suivants :

niveau de découpage	somme des carrés	degré de liberté	F calculé
départemental	12913,0	6	8,86
bassin	10852,4	15	2,98
cantons	52907,9	218	

La encore, l'hypothèse de nullité est à rejeter et cette analyse d'ensemble confirme donc les conclusions des analyses partielles, mais les limites de cette conclusion restent bien entendu les mêmes.

SECTION 2 - LES IMPLICATIONS DU CHOIX DU DECOUPAGE DEPARTEMENTAL

Les conclusions des deux études précédentes sont en partie contradictoires mais les différences de techniques et hypothèses sousjacentes ne permettent pas de trancher de façon décisive en faveur ou en défaveur du découpage départemental. On peut faire trois observations relatives au choix à effectuer.

α) Il faut tout d'abord rappeler que ce jugement de pertinence est fondé sur l'étude d'indicateurs reflétant un fonctionnement théorique du marché de l'emploi. Le découpage peut donc être en fait pertinent mais les indicateurs retenus, mauvais, ou encore la théorie fautive, parce que non adaptée à la réalité de par ses hypothèses et ses mécanismes (1). Si tel est le cas, on peut espérer que l'étude entreprise pour l'ensemble de la France permettra de trouver de meilleurs indicateurs ou d'infirmar, par ses observations, la théorie erronée.

β) Si le marché de l'emploi correspond à un découpage infra-départemental, l'analyse développée au niveau départemental peut d'un point de vue opérationnel s'avérer satisfaisante. En effet, les lois mises en évidence ne traduiront pas le fonctionnement réel du marché de l'emploi mais on se trouvera en face de la classique clause "tout se passe comme si ..." qui permet, du moins durant un certain temps, de pouvoir prévoir les réactions du système. Ce temps de validité est fonction de la permanence de grandeurs caractéristiques déterminées par des mécanismes réels (inconnus) opérant à un niveau infra-départemental (2). Si cette situation est frustrante pour l'économiste, il faut bien reconnaître qu'elle satisfait l'homme d'action. Dès lors que l'économiste est amené à rechercher un modèle opérationnel prévisionnel, il est conduit à sacrifier certaines préoccupations théoriques à l'objectif d'opérationnalité. Mais ce compromis comme beaucoup d'autres est souvent douloureux !

γ) Enfin il faut souligner que si le marché de l'emploi existe à un niveau infra-départemental, il est certain que des modèles théoriques élaborés ne pourront, pendant encore quelques décennies, être testés (du moins pour l'ensemble de la France). Dans ce cas on est face au dilemme suivant : ou rechercher un modèle satisfaisant que l'on ne pourra pas tester et donc ne servira pas, ou rechercher un modèle non satisfaisant mais que l'on pourra alimenter en données en espérant que la remarque du paragraphe précédent pourra s'appliquer.

Toutes ces considérations font que l'on retiendra le découpage départemental et qu'à la fin de cette étude il faudra essayer de faire un bilan.

(1) Voir à ce propos le chapitre 3 section 1

(2) D'une certaine façon le problème posé s'apparente à celui du passage de la micro-économie à la macro-économie.

Remarques sur certaines techniques économétriques retenues

CHAPITRE 7 - REMARQUES sur CERTAINES TECHNIQUES ECONOMETRIQUES RETENUES

Le passage d'un modèle conceptuel à un modèle opérationnel implique le choix de la forme mathématique des relations entre les variables, du moins en ce qui concerne les équations de comportement et les équations techniques.

Dans le cas d'un modèle spatialisé prévisionnel, les zones peuvent être étudiées en comparaison avec une référence exogène (le plus souvent l'évolution d'une valeur nationale). Le modèle opérationnel postule alors un certain type d'évolution de chaque zone comparativement à celle de la référence exogène. Cette étude s'effectue par l'intermédiaire de coefficients de spécificités zonales, que l'on peut définir comme le rapport de la valeur de chaque zone à celle de la valeur de référence. Il n'y a pas dans ce cas de véritable liaison causale décrite au niveau de chaque zone mais observation de régularité d'évolution et hypothèse de leur maintien au cours du temps.

Un second type d'approche est également possible et qui consiste à décrire un mécanisme économique. La description des mécanismes théoriques se fait alors au niveau de la zone.

Ces deux types d'approche permettent de trouver, pour une date donnée, un vecteur de valeurs zonales pour une variable endogène donnée. Si, quel que soit le type d'approche retenu, l'on suppose l'existence d'un modèle national fournissant pour la date considérée une valeur pour cette variable endogène, il est plus que vraisemblable que l'information fournie par le modèle national ne coïncidera pas avec la même information obtenue par sommation des valeurs zonales. Ce problème de cohérence sera évoqué dans la dernière section de ce chapitre.

SECTION 1 - LOI d'EVOLUTION de SPECIFICITES ZONALES OBSERVEES

Une variable donnée à un instant du temps présente des valeurs différentes pour les différentes zones de l'étude. La structure du vecteur de données est plus facile à appréhender lorsque l'on rapporte toutes les valeurs zonales à une même valeur (la valeur France entière par exemple). C'est un tel rapport que l'on définira comme étant la spécificité zonale.

Les informations fournies au chapitre 1 sur les disparités du revenu par tête selon les départements fournissent un exemple de spécificité zonale.

Les techniques que l'on peut imaginer peuvent être utilisées soit pour des données absolues, (par exemple, nombre de véhicules particuliers, par département), soit pour des ratios (par exemple, nombre de véhicules particuliers par habitant, par département).

Dans le cas de ratios le vecteur zonal du dénominateur de l'année finale est supposé connu, dans l'exemple choisi, la population de chaque département pour l'année considérée. Lorsqu'il s'agit de données elles-mêmes calculées, il faut être bien conscient que les erreurs de prévision peuvent se cumuler⁽¹⁾.

Soit $VAL(Z, T(I))$ la valeur de la variable endogène de la zone Z à la date $T(I)$. On utilisera F pour la valeur France entière qui sert de référence. Les raisonnements qui vont suivre restent inchangés, l'on remplace la notion de date par celle de période, $T(I)$ représente alors la période $(T(I), T(I + K))$.

1.1 - Quelques régularités d'évolution concevables

1.11 - Reconstitution de la structure zonale observée

Il est possible tout d'abord de considérer que la structure des disparités zonales reste inchangée au cours du temps. Cette hypothèse de structure constante revient à admettre un taux d'évolution identique pour toutes les zones. Le coefficient de spécificité est alors supposé constant dans le temps, une telle technique est sans doute utilisable si l'on envisage des phénomènes d'évolution lente et dans le cadre d'une utilisation à court ou à moyen terme.

(1) Cette remarque est du reste valable pour l'analyse en terme de loi d'évolution économique.

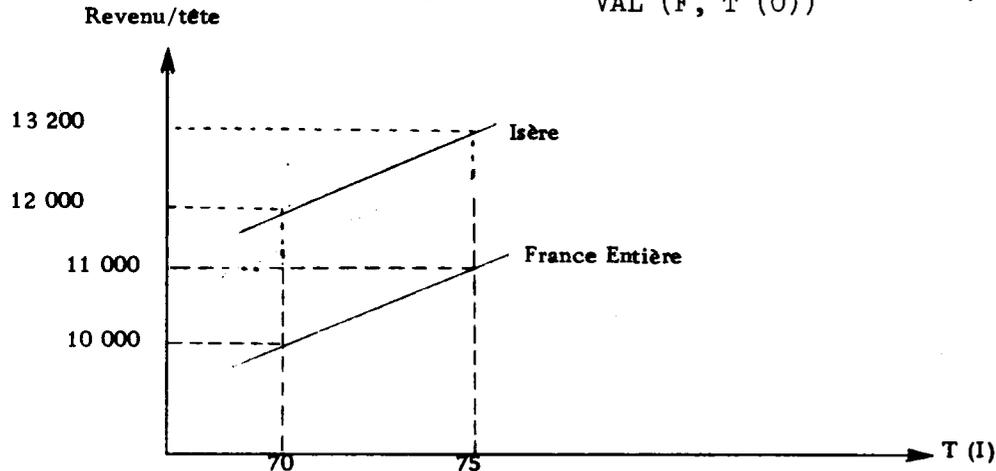
Le coefficient de spécificité est alors supposé constant dans le temps, une telle technique est sans doute utilisable si l'on envisage des phénomènes d'évolution lente et dans le cadre d'une utilisation à court ou à moyen terme.

Mathématiquement, nous avons alors :

$$\frac{\text{VAL}(Z, T(0))}{\text{VAL}(F, T(0))} = \frac{\text{VAL}(Z, T(I))}{\text{VAL}(F, T(I))}$$

ou encore :

$$\text{VAL}(Z, T(I)) = \frac{\text{VAL}(Z, T(0))}{\text{VAL}(F, T(0))} \cdot \text{VAL}(F, T(I))$$



Exemple - Supposons que le revenu annuel par tête soit en moyenne en France de 10 000 en 1970 et de 12 000 pour le département de l'Isère. L'utilisation de l'hypothèse de la constance des coefficients de spécificité (ici 1,2 pour le département de l'Ain) et la connaissance exogène du revenu national annuel moyen par tête par exemple de 11 000 pour 1975 donne, pour cette même année, pour le département de l'Ain, un revenu de $1,2 \cdot 11\ 000 = 13\ 200$ F.

Il faut souligner que la constance de coefficient de spécificité ne signifie pas forcément permanence de toutes les données structurelles de l'ensemble décrit. En effet, lorsque les valeurs envisagées portent sur une période $(T(I), T(I + K))$ (par exemple l'accroissement (absolu ou relatif) de population totale de chaque département entre 1962 et 1968) la population active obtenue pour 1974 (par exemple) est telle que le poids relatif de chaque département dans la population totale est changé : il n'y a donc pas permanence de cette donnée structurelle.

Zones	Sud	Nord	France
PT (62)	450	550	1 000
PT (68)	600	600	1 200
PT (68) - PT (62)	150	50	200
Coefficient de spécificité	0,75	0,25	
PT (74) - PT (68)	75	25	100
PT (74)	675	625	1 300

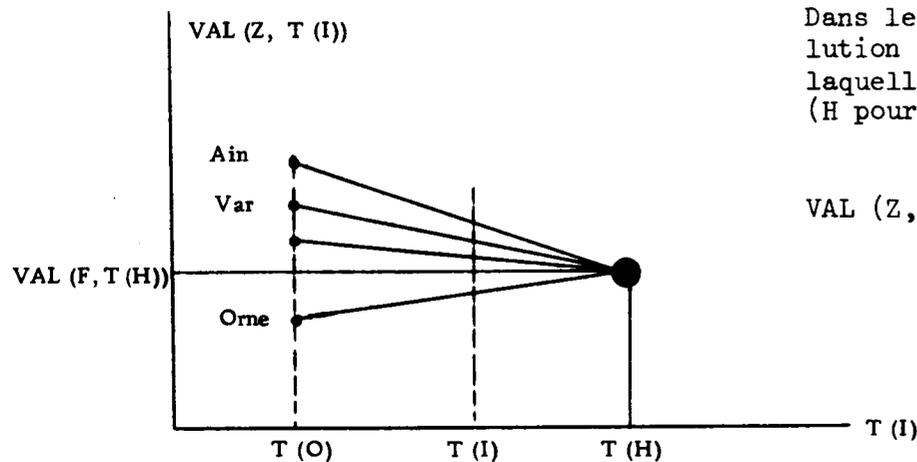
Le tableau ci-contre illustre ce fait. Compte tenu des observations pour 1962 et 1968 ainsi que de la donnée exogène population totale 1974 et en utilisant l'hypothèse de constance des coefficients de spécificité, on obtient une répartition finale différente de 50 %.

1.12 - Evolution de la structure zonale observée

Il peut y avoir un certain nombre de raisons pour penser que tôt ou tard, pour certains phénomènes (durée de travail par exemple, ou taux d'équipement automobile des ménages), il y aura une certaine uniformisation. Les disparités zonales d'une variable donnée vont tendre à s'estomper dans le temps.

La valeur la plus plausible dans le long terme est la valeur moyenne nationale VAL (F, T (H)). Pour l'année de base T (0), on dispose pour chaque zone d'une valeur VAL (Z, T (0)).

a) Evolution linéaire des coefficients de spécificité



Dans le graphe ci-contre, on a figuré quelques zones et une loi d'évolution linéaire. L'année étudiée est T (I), tandis que l'année pour laquelle nos disparités se seront complètement estompées est T (H) - (H pour horizon).

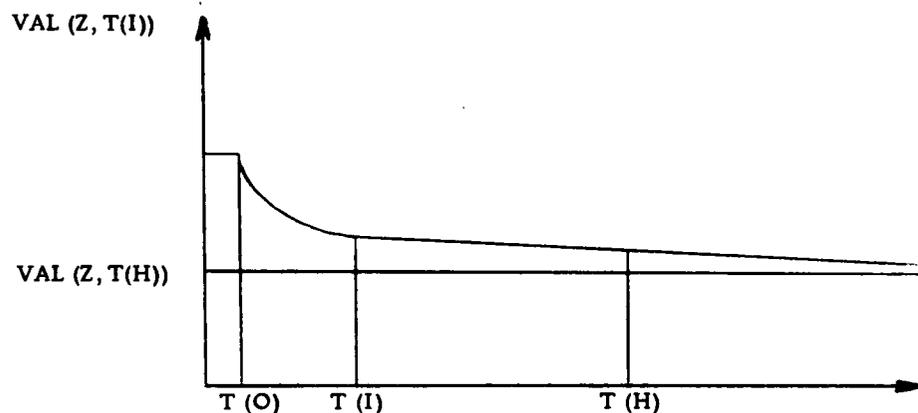
$$VAL(Z, T(I)) = \frac{T(I) - T(0)}{T(H) - T(0)} \cdot VAL(F, T(H)) + \frac{T(H) - T(I)}{T(H) - T(0)} \cdot VAL(Z, T(0))$$

b) Evolution non linéaire des coefficients de spécificité

- α - Il est possible de compliquer la loi d'évolution, en supposant connue une évolution dont le taux doit en être fixé exogènement, on suppose que la valeur limite est connue pour une date éloignée (ce qui permet de raisonner, puisqu'il s'agit d'une asymptote horizontale, comme si cette date était rejetée à l'infini).

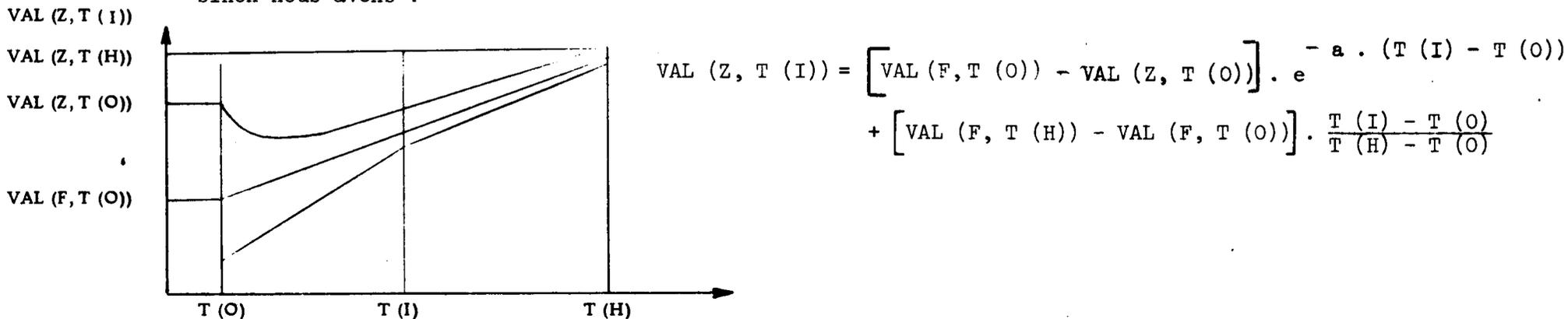
Mathématiquement, nous avons :

$$\text{VAL}(Z, T(I)) = \left[\text{VAL}(F, T(H)) - \text{VAL}(Z, T(O)) \right] \cdot e^{-a \cdot (T(I) - T(O))} + \text{VAL}(F, T(H))$$



a étant le taux de décroissance, ou de croissance, fixé par expert pour chaque zone. Notons que si nous disposons de 2 points et d'une valeur asymptotique, il y a une exponentielle et une seule passant par ces deux points (et donc une valeur fixe de a).

- β - Il est possible aussi d'envisager un amortissement exponentiel avec asymptote oblique (laquelle peut être le trend de la valeur France entière). Ce procédé est intéressant si VAL(Z, T(O)) est inférieur à VAL(F, T(H)), sinon nous avons :



γ - Ces deux types d'évolution ne sont donnés qu'à titre d'exemple, il est possible en effet d'en envisager beaucoup d'autres. Leur seul intérêt est d'essayer de quantifier certaines hypothèses et en permettre ainsi la vérification. On pourrait faire appel à l'une de ces deux formulations pour décrire des phénomènes de rattrapage (pour l'évolution de taux d'activité des femmes d'une tranche d'âge donnée, par exemple).

Mais ces techniques sont à utiliser avec prudence : il faut, par exemple, penser que la méthode du β implique une décroissance puis une croissance des zones dont la valeur initiale est supérieure à celle de la valeur France entière.

1.2 - Problèmes posés par les lois d'évolution de spécificités zonales

L'utilisation des coefficients de spécificités posent deux types de problèmes : un problème d'ordre comptable et un de signification des coefficients de spécificité.

1.21 - Problème de cohérence

La technique des coefficients de spécificité n'assure pas de cohérence au niveau national :

- si les valeurs ainsi traitées ne sont pas sommables, c'est-à-dire si la somme des informations par zone n'a pas de signification, ce qui est le cas des ratios, indices, etc.
- ou si le numérateur (ou le dénominateur) de la valeur non sommable calculée pour l'année de prévision permet de trouver une information sommable homogène avec l'information "France Entière" prévue.

Le nombre de degrés de liberté du tableau est, en fait, insuffisant. L'exemple qui suit illustre ce problème de cohérence.

	Zones	Nord	Sud	France
1970	Population totale	1 000	1 000	2 000
	Population active	600	500	1 100
	Taux d'activité	0,6	0,5	0,55
	Coefficient de spécificité	1,09	0,91	1
1975	Population totale	1 050	1 100	2 150
	Population active			1 204
	Taux d'activité	0,6109	0,509	0,56

Dans le tableau ci-contre, les informations relatives à la population des zones pour 1975 sont supposées connues, d'où l'on tire la population de la France en 1975. On connaît, par ailleurs, la population active de la France en 1975, d'où l'on tire les taux d'activité des zones pour 1975, en supposant constants les coefficients de spécificité. Mais le produit de ces taux d'activité zonaux par la population correspondante ne permet pas de retrouver le volume de population active initiale :

$$0,6109 \cdot 1\,050 + 0,509 \cdot 1\,100 = 1\,201 \neq 1\,204.$$

Des solutions à ce problème seront envisagées dans la troisième partie de ce chapitre.

1.22 - Signification des coefficients de spécificité

L'utilisation des coefficients de spécificité repose essentiellement sur des hypothèses d'évolution de comportement, non quantifiables dans le cadre du modèle (parce que d'ordre institutionnel ou sociologique par exemple). L'analyse causale sous-jacente ne nécessite pas, ou ne permet pas, l'introduction de variables explicatives. La seule variable exogène nécessaire est la valeur nationale de la variable.

Il faut bien reconnaître que l'utilisation de telles techniques ressemble, dans une certaine mesure, à un aveu d'impuissance : par exemple, l'emploi des coefficients de spécificité pour calculer des taux d'activité départementaux par tranche d'âge, permet d'éviter d'étudier, au niveau départemental, l'influence de l'abaissement de l'âge de la retraite, de l'allongement du temps des études, des modifications relatives au service militaire, etc.

1.23 - Techniques d'analyse du passé et techniques de prévision

On peut trouver quelques analogies entre certaines techniques d'analyse des données passées et certaines régularités d'évolution utilisables pour la prévision. Prenons, par exemple, la technique utilisée BEAUD⁽¹⁾ pour la comparaison entre les recensements de 1962 et de 1954 pour certains phénomènes :

(1) M. BEAUD : "Une analyse des disparités régionales de croissance", Revue Economique, janvier 1966, repris dans Espace Economique Français, 1962, Tome 2.

Soit une variable, par exemple la population active agricole, que l'on partitionne suivant un critère, par exemple l'âge, indicé par K. Le problème est d'isoler dans l'évolution de la valeur d'une zone VAL (Z, ...) entre deux instants du temps⁽¹⁾ T (0) et T (1), l'influence du caractère de partition de la variable (ici l'âge) de ce qu'il appelle un effet régional.

Pour ce faire, on éclate l'écart entre le taux d'évolution de la zone et de la France

$$\left[\frac{\text{VAL (Z, T (1))}}{\text{VAL (Z, T (0))}} - \frac{\text{VAL (F, T (1))}}{\text{VAL (F, T (0))}} \right] \text{ en une somme de deux écarts :}$$

- l'écart entre le taux d'évolution observé pour la zone et un taux fictif obtenu en rapportant l'effectif théorique calculé en appliquant à chaque classe (VAL (Z, K, T (0))) les coefficients d'évolution correspondants pour la France (VAL (F, K, T (1)) / VAL (F, K, T (0))) à la valeur globale de la zone pour l'année de base. Cet écart est significatif des distorsions du taux global d'évolution dues aux différences structurelles de la zone avec la France :

$$\frac{\text{VAL (Z, T (1))}}{\text{VAL (Z, T (0))}} - \left[\sum_K \text{VAL (Z, K, T (0))} \cdot \frac{\text{VAL (F, K, T (1))}}{\text{VAL (F, K, T (0))}} \right]$$

- et l'écart entre ce taux fictif et le global France Entière, significatif de l'effet zonal :

$$\sum_K \left[\text{VAL (Z, K, T (0))} \cdot \frac{\text{VAL (F, K, T (1))}}{\text{VAL (F, K, T (0))}} \right] - \frac{\text{VAL (F, T (1))}}{\text{VAL (F, T (0))}}$$

Il est à noter que la formule permettra de retrouver les effectifs théoriques utilisés dans le calcul du taux fictif soit :

$$\sum_K \left[\text{VAL (Z, K, T (0))} \cdot \frac{\text{VAL (F, K, T (1))}}{\text{VAL (F, K, T (0))}} \right]$$

ressemble fort à l'utilisation de coefficients de spécificité avec reconduction de la structure zonale observée⁽²⁾; en effet, cette formule s'écrit encore :

$$\sum_K \left[\frac{\text{VAL (Z, K, T (0))}}{\text{VAL (F, K, T (0))}} \cdot \text{VAL (F, K, T (1))} \right]$$

(1) La technique pourrait être conservée si l'on prenait deux périodes au lieu de deux dates.

(2) Voir le 1.1 a de ce chapitre.

Or, la technique des coefficients de spécificité permet de calculer une valeur théorique VAL (Z, K, T (I))

$$\frac{\text{VAL (Z, K, T (0))}}{\text{VAL (F, K, T (0))}} = \frac{\text{VAL (Z, K, T (1))}}{\text{VAL (F, F, T (1))}} \implies \frac{\text{VAL (Z, K, T (0))}}{\text{VAL (F, K, T (0))}} \cdot \text{VAL (F, K, T (1))} = \text{VAL (Z, K, T (1))}$$

Mais cette technique sert à analyser une évolution passée alors que celle des coefficients de spécificité sert à la prévision, en faisant des hypothèses de régularités économiques.

On peut remarquer que l'effet régional peut être considéré comme une spécificité zonale, jouant additivement, et que l'on pourrait en envisager des lois d'évolution de la même façon que l'on a envisagé des lois d'évolution pour les coefficients de spécificité (qui eux jouent de façon multiplicative)⁽¹⁾. Il faudrait toutefois faire jouer, en outre, l'effet structurel sur l'évolution du coefficient global de la zone.

(1) Il est toujours possible de transformer une loi portant sur un écart en loi sur un coefficient de spécificité, et réciproquement.

SECTION 2 - LOIS d'EVOLUTION ECONOMIQUES : APPEL aux TECHNIQUES de REGRESSION MULTIPLE

L'application des techniques de régression multiple à l'analyse spatiale pose un problème particulier dans la mesure où l'on s'intéresse à la fin du processus, individuellement à chaque zone du territoire.

L'économètre peut, à partir de techniques de régression multiple, donner le montant le plus probable de la quantité consommée, par ménage, d'un bien donné, en connaissant par exemple le montant du revenu disponible par tête et le prix du bien considéré. Ce qui l'intéresse, en fait, dans sa démarche, c'est une prévision globale de la consommation de ce bien et non la prévision de la consommation de chaque ménage en particulier.

Pour l'économiste régional, la précision dans la prévision de la répartition spatiale d'une grandeur a au moins autant d'importance que celle de son montant exact, car son propos est l'étude de disparités spatiales.

La spécificité du problème spatial mène à des remarques sur les techniques de régression multiple différentes selon que l'on se trouve dans le cadre de modèles déterministes ou de modèles probabilistes.

2.1 - Cas des modèles déterministes

Comme on l'a vu au chapitre 3, la technique de régression multiple, dans le cadre des modèles déterministes, n'implique aucune hypothèse restrictive : elle permet simplement d'établir une liaison fonctionnelle "synthétisant" un tableau de données croisant les observations, c'est-à-dire ici des zones (département, par exemple) et des variables.

2.11 - Signification des écarts résiduels

L'écart entre la valeur observée d'un département et celle théorique déterminée par la régression multiple, représente, aux erreurs d'observation près, la part des autres facteurs explicatifs non pris en considération dans la régression. La régression multiple ne s'intéresse plus à la personnalité de l'unité observée; une fois calculés les différents paramètres (les résultats obtenus étant des résultats "d'ensemble"). Or pour le spécialiste des problèmes régionaux les zones observées ne sont pas interchangeables ; il faut donc formuler des hypothèses explicites sur la nature des écarts.

Pour une utilisation dans un modèle prévisionnel, on peut faire :

- soit l'hypothèse d'une modification importante des facteurs explicatifs non pris en compte, et ce pour chaque département. Dans ce cas, la valeur calculée à l'aide des paramètres de la régression et des nouvelles valeurs des variables endogènes est certainement celle qu'il y a lieu de retenir ;
- soit celle d'une modification lente des facteurs explicatifs non pris en compte. Dans ce cas, l'écart constitue en quelque sorte une spécificité départementale dans la mesure où l'ensemble des valeurs de ces autres facteurs n'est pas le même d'un département à un autre.

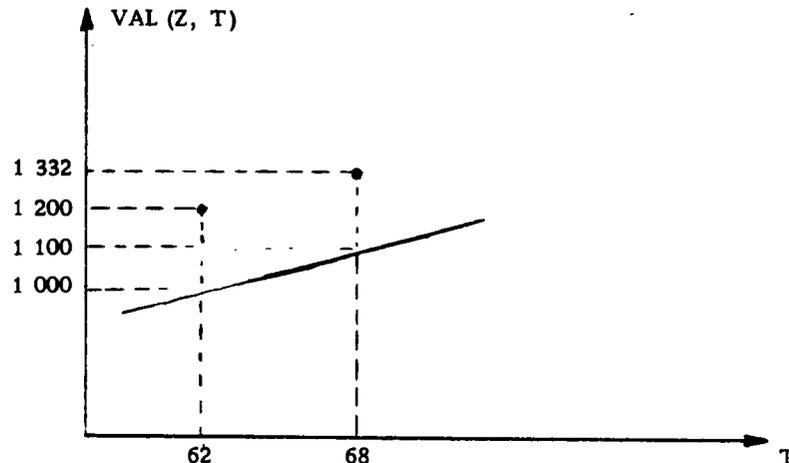
L'adoption de cette dernière hypothèse nous a permis d'introduire une nouvelle technique dans la recherche des divers sous modèles (utilisée concurremment avec l'approche probabiliste (cf § 2.2)).

2.12 - Utilisation des coefficients de spécificité de régression

Si l'écart entre la valeur observée et la valeur calculée caractérise chaque département pour la date ou la période d'observation utilisée par le modèle, dans une utilisation en prévisionnel il est nécessaire de poser une loi d'évolution de cet écart.

On peut proposer, entre autres, comme loi d'évolution de cet écart :

- constance dans le temps de l'écart, pris en valeur absolue ;
- constance dans le temps du rapport entre la valeur observée et la valeur théorique (nous appellerons un tel rapport un coefficient résiduel). Ce qui repose sur l'hypothèse de stabilité de la part non-expliquée par la régression dans l'ensemble des facteurs explicatifs.



Le graphique ci-contre illustre pour un département la constance du coefficient résiduel dans le temps. Il va de soi que seule une confrontation avec la réalité peut permettre d'adopter de telles hypothèses : par exemple, les résultats calculés pour 1968 à l'aide des coefficients résiduels par exemple devront être comparés aux valeurs observées pour cette date et si le résultat est jugé satisfaisant, il y aura lieu de retenir cette méthode.

2.2 - Cas des modèles probabilistes

Dans les modèles probabilistes, le vecteur résiduel $U(I)$ est considéré comme une variable aléatoire ; il s'en suit que la variable endogène $Y(I)$ est elle-même une variable aléatoire.

Il convient tout d'abord d'expliquer pourquoi cette série d'hypothèses faisant de $Y(I)$ une variable aléatoire est justifiée dans le cadre de l'analyse spatiale avant d'en envisager les implications au niveau de l'adéquation du modèle au réel.

2.21 - Nécessité de l'introduction d'hypothèses complémentaires

Le fait que le vecteur $Y(I)$ observé est l'une des réalisations possibles d'un vecteur de variables aléatoires peut trouver deux explications :

- Le vecteur observé est l'une des solutions d'une série d'états potentiels, fonctions de facteurs accidentels plus ou moins permanents (tels que l'étiquette politique du député, un été "désastreux", ...) ou de facteurs non accidentels (tels que la distribution des revenus dans la zone...) mais ne caractérisant que secondairement (dans l'étude) la zone étudiée. Ces facteurs secondaires sont susceptibles de variations dont l'impact sur le sens et l'amplitude de l'écart résiduel n'est pas prévisible.
- Une seconde explication peut être avancée pour justifier le fait que la variable $Y(I)$ soit une variable aléatoire. A chaque instant du temps, chaque zone se caractérise par un ensemble de valeurs des variables qui évolue au cours du temps (le nombre de personnes résidant dans un département se modifie chaque jour, par exemple). Supposons que la loi que l'on veuille établir le soit à partir d'une coupe synchronique (par exemple au 31 mars 1962), la démarche effectuée implique que cette loi reste inchangée sur une période temps (autrement dit, les paramètres trouvés auraient été les mêmes si l'on avait pris les informations au 15 mars ou au 30 décembre). Autrement dit, l'observation de 95 données (si l'on est au niveau départemental), à un instant du temps procède des techniques d'échantillon.

Ces deux explications avancées, qui, il faut le souligner, peuvent être complémentaires, permettent de se situer dans le cadre probabiliste. Le problème reste de savoir dans quelle mesure l'échantillon peut être considéré comme aléatoire.

On pourrait considérer que l'aléa n'est pas réalisé pour deux raisons :

Si l'on considère que les lois que l'on va étudier sont valables sur la période 1954 - 1985, le sondage n'étant possible en fait que sur la période 1954 - 1970, n'est réalisé en fait que sur une sous-population. Cet argument peut être contre-balancé par l'introduction de l'hypothèse de l'identité des caractéristiques de la distribution des valeurs dans les deux sous-populations. Cette hypothèse, très restrictive, peut être réduite à la constance de la seule variance (1).

Une coupe synchronique nous oblige à prendre 95 données. Cet argument, valable en soi, peut amener à effectuer un sondage de taille quelconque sur l'ensemble possible des unités spatio-temporelles (Ain 1957, Loiret 1963, Calvados 1954, etc.). Mais l'hypothèse de la grande probabilité d'une similitude structurelle entre des données d'une coupe synchronique et des données d'un tirage aléatoire ne paraît pas dénuée de fondement.

L'acceptation des hypothèses faisant du modèle (ou du sous-modèle) un modèle probabiliste, permet de tester son adéquation au réel.

2.22 - Rejet de certaines variables exogènes

L'un des intérêts du cadre probabiliste est le fait de pouvoir associer à chaque paramètre estimé dans la régression son écart-type, si la loi de distribution des résidus n'est pas connue, et un intervalle de confiance si cette loi est normale.

En conséquence de quoi un jugement peut être porté pour savoir si un paramètre donné est significativement différent de zéro ou non. Dans la négative, la variable explicative correspondante doit être écartée. Il faut toutefois noter que le modèle conceptuel peut être valable et le modèle opérationnel mal choisi, les relations causales étant altérées par des phénomènes parasites tels que l'effet de dimension (qui peut être éliminé en travaillant en ratios et non en valeurs absolues)...

Si les différentes hypothèses, intrinsèques aux techniques de régression multiple dans le cadre probabiliste, ou nécessitées par l'application de ces techniques à l'analyse spatiale, sont justifiées, on dispose d'un puissant outil d'analyse permettant de cerner le réel.

(1) Cette critique d'ailleurs porte plus sur l'utilisation des relations fonctionnelles obtenue que sur l'établissement de ces relations.

2.23 - Signification des écarts résiduels

L'amplitude et le sens de l'écart résiduel étant dans cette approche parfaitement aléatoire, il ne saurait être question d'utiliser la technique des coefficients résiduels décrite au 2.1 b de ce chapitre.

2.3 - Remarques

Nous avons été amenés, au cours de nos recherches, à réfléchir sur les deux problèmes théoriques suivants : celui du choix entre un modèle en taux et un modèle en effectifs, et celui de l'intérêt de la stratification d'une variable endogène, ces deux problèmes, à notre connaissance, n'ayant pas encore été abordés. Le dernier a fait l'objet du § 2.4 du chapitre 2 (page 77 et suivantes). Il ne reste plus qu'à indiquer les fondements du programme de régression optimale utilisé et qui est de loin supérieur aux autres programmes existants en France.

2.31 - Problème du choix entre modèles en taux et modèles en effectifs

- a) Si l'on se place dans le cadre de modèle déterministe, le critère que l'on utilisera naturellement est le coefficient de corrélation : on choisira alors le modèle ayant le coefficient de corrélation le plus élevé. Par contre, si l'on se place dans le cadre de modèle probabiliste, cette règle de choix est à manier avec beaucoup de prudence, car les fluctuations de la distribution d'échantillonnage des coefficients de corrélation peuvent amener des écarts non significatifs.
- b) Le fait qu'un modèle "en valeur absolue" ait un coefficient de corrélation supérieur à un modèle "en taux" ne doit pas conduire à écarter d'emblée ce dernier. Supposons, en effet, que le dénominateur utilisé pour définir le taux (ici la population active du département) soit connu sans erreur, l'erreur relative commise sur les effectifs calculés grâce au modèle en taux est la même que celle commise sur le taux. Il est alors facile de calculer l'écart entre les effectifs observés et calculés par cette méthode. La moyenne quadratique de ces écarts a la même signification que la variance résiduelle obtenue dans le modèle en effectifs et lui est donc directement comparable. On choisira alors le modèle à plus faible variance résiduelle⁽¹⁾.

(1) Rappelons que le coefficient de détermination est le complément à 1 du rapport de la variance résiduelle à la variance totale de la variable expliquée, il s'en suit que, dire que l'on choisit le modèle ayant le plus fort coefficient de détermination (ou de corrélation) ou dire que l'on choisit celui ayant la plus faible variance résiduelle revient au même.

2.32 - Le programme de régression optimale utilisé

Nous ferons appel ici aux conventions d'écriture adoptées au chapitre 3, section 2.

2.321 - Le problème posé

Très souvent, le chercheur qui utilise les techniques de régressions multiples, dispose de plusieurs indicateurs pour tester un ensemble d'hypothèses. La recherche de k variables, prises dans un ensemble de p variables et qui fournissent la meilleure régression à k variables, passe par la résolution d'un problème combinatoire (C_p^k) qui nécessite la comparaison d'un nombre de solutions possibles, très rapidement prohibitif lorsque p croît.

Divers programmes existants se contentent, une fois que $(k - 1)$ variables sont retenues, d'introduire comme K ième variable celle qui a le coefficient de corrélation le plus élevé avec le vecteur résiduel $(Y - Y^*)$ calculé sur la régression à $(k - 1)$ variables ou, ce qui est déjà mieux, de rechercher la régression la meilleure sur un ensemble de k variables comprenant obligatoirement au moins $(k - 2)$ des variables déjà intégrées dans la régression retenue à $(k - 1)$ variables. Or, aucune de ces deux démarches n'est satisfaisante car elles ne garantissent pas l'obtention de la meilleure régression possible puisque celle-ci, du fait des colinéarités, peut, à la limite, n'intégrer aucune des variables retenues dans le sous-ensemble de $(k - 1)$ à condition, bien entendu, que l'on ait $p > 2(k - 1)$.

2.322 - Les critères de choix

Une équation de régression est, en approche probabiliste⁽¹⁾, jugée sur son coefficient de corrélation et, pour un k donné, la meilleure régression est celle qui a le meilleur coefficient de corrélation. L'introduction d'une nouvelle variable n'est pas forcément utile car, au coefficient de corrélation de la meilleure régression à $(k + 1)$ variables et qui est plus élevé que celui à k variables, peut être associé un intervalle de confiance moins satisfaisant, ce qui s'apprécie par un test de FISCHER où ρ_k^2 est le coefficient de détermination optimal pour k variables explicatives et F_k la valeur associée.

(1) Voir chapitre 3, page 75.

$$F_k = \frac{n - k - 1}{k} \cdot \frac{\rho_k^2}{1 - \rho_k^2}$$

Le test sur F_k permet de trouver l'optimum optimorum.

Avant de donner les principes du programme retenu, il est nécessaire de rappeler quelques résultats fondamentaux, un peu différents de ceux donnés au chapitre 3, § II, 4.1 c, mais qui partent des mêmes conventions d'écriture⁽¹⁾.

On peut écrire les coefficients de détermination ρ^2 et de régression b_j comme suit :

$$b_j = \frac{|R^{oj}|}{|R^{oo}|} \cdot \sqrt{\frac{v_{oo}}{v_{jj}}}$$

$$\rho^2 = 1 - \frac{|R|}{|R^{oo}|}$$

2.323 - Principes de la méthode

Trois auteurs américains⁽²⁾ ont montré que l'application de quelques règles simples permettaient de s'affranchir en partie du problème combinatoire. Ces règles sont au nombre de 3.

- Lorsque l'on désire sélectionner sur les $(n - k + 1)$ variables non prises en considération dans une régression à $k - 1$ variables la variable qui maximise $[(\rho_k - \rho_{k-1})]$, sans remettre en cause les $(k - 1)$ variables déjà intégrées, on démontre que cette variable Q est celle qui maximise l'expression $[|R^{oq}| / |R^{qq}|]$.

(1) Yamane "Statistics" (Harper International Edition), p. 819.

(2) Beale, Kendall et Mann "The discarding of variables in multivariate Analysis" (Biometrika, 54 (3) - 1969).

- On associe à chaque variable Q un seuil non conditionnel T_q qui est l'écart-type résiduel obtenu par la régression effectuée sur toutes les n variables sauf Q . N'importe quel sous-ensemble de variables qui ne contient pas Q ne peut avoir d'écart-type résiduel (calculé sur la régression sur cet ensemble) inférieur au seuil T_q . Il s'en suit que si à une étape quelconque on trouve qu'une variable donnée a un seuil non conditionnel plus grand ou égal au plus petit des écarts-types résiduels calculés sur n'importe lequel des sous-ensembles à k variables considérés jusqu'à présent, alors cette variable doit se trouver dans les sous-ensembles qui fournissent un écart-type résiduel meilleur et on peut concentrer son attention sur de tels sous-ensembles.
- On associe en outre à chaque variable Q des seuils conditionnels par rapport aux variables X_i et X_j et que l'on note $S_{q/i, j}$ et qui est l'écart-type résiduel calculé dans la régression obtenue en utilisant toutes les n variables sauf X_i et X_j . Si l'écart-type résiduel obtenu sur un sous-ensemble donné est supérieur à $S_{q/i, j}$, on portera son attention sur les sous-ensembles qui incluent Q et excluent I et J .

Le programme utilisé⁽¹⁾ s'appuie sur ces trois considérations pour rechercher progressivement à partir de la matrice de corrélation des $(p + 1)$ variables de départ, y compris donc la variable endogène (à cause du calcul de ρ_k^2) et en utilisant la méthode du pivot pour rentrer (ou sortir) des variables pour définir le sous-ensemble optimal. Pour éviter un temps de calcul prohibitif dans le cas où n est grand (> 20) on peut arbitrairement limiter le nombre des essais, ce qui fait que l'on n'obtient pas la solution optimale, mais que l'on en est proche puisque chaque itération a pour effet d'améliorer le sous-ensemble. Dans cette étude, la solution optimale a toujours été trouvée avant que l'on ait atteint le nombre d'itérations limite que l'on s'était fixé.

(1) Ecrit par MM. BOYCE, FARHI et WEISCHEDEL - Voir "A computer program for optimal regression analysis" (1969).

SECTION 3 - Le PROBLEME de la COHERENCE NATIONALE

Un modèle spatialisé peut avoir intérêt à s'appuyer sur un modèle national. Le modèle général se décompose alors en deux sous-modèles.

L'intérêt d'un sous-modèle national se justifie, entre autres, pour les raisons suivantes :

- La description de certains mécanismes économiques n'a un sens qu'au niveau national, parce qu'à ce niveau, on peut prendre en considération, par exemple, la concurrence internationale, l'impact technique de la productivité du capital de certaines branches, l'évolution de prix relatifs...
- Par ailleurs, on peut préférer un modèle conceptuel élaboré à un niveau national, à un modèle plus frustré à un niveau zonal (les contingences d'encombrement de données et la longueur des calculs intervenant nécessairement au niveau d'un modèle opérationnel).
- Certaines données n'existent qu'au niveau national.

Entre le sous-modèle national et le sous-modèle spatialisé existent nécessairement plusieurs liaisons⁽¹⁾ qui s'expriment par le fait que certaines variables interviennent dans les deux sous-modèles.

La détermination d'une variable peut se faire simultanément dans les deux sous-modèles : dès lors, il n'est pas certain que l'information du sous-modèle national coïncide avec la somme des valeurs régionales de l'autre sous-modèle. Il s'agit là de l'un des problèmes les plus évidents de cohérence. Pour mettre en évidence ces problèmes de cohérence, il est nécessaire d'étudier tout d'abord les interactions entre les sous-modèles. On examinera par la suite les moyens disponibles permettant d'assurer cette cohérence.

3.1 - Interaction entre le sous-modèle spatialisé et le sous-modèle national

3.11 - Le rôle des variables dans chaque sous-modèle

Ces variables, que l'on retrouve au niveau des deux sous-modèles, peuvent intervenir de trois façons différentes :

(1) Sinon, il y aurait une simple juxtaposition de modèles indépendants.

- α) Une variable est endogène dans le sous-modèle national et exogène dans le sous-modèle spatialisé. La donnée nationale sert d'input au modèle spatialisé où elle est éclatée à un niveau zonal.
- β) Une variable est endogène dans le sous-modèle spatialisé et exogène dans le sous-modèle national. La somme des données zonales sert d'input au modèle national. On peut imaginer, par exemple, que les besoins en main-d'oeuvre étrangère soient calculés au niveau de chaque zone et que le cumul de ces besoins fournisse le montant de l'immigration nécessaire.

Il faut remarquer que ce type de détermination exogène peut engendrer des incohérences dans le sous-modèle national. Il est possible, par exemple, que le volume de main-d'oeuvre étrangère nécessaire soit trop important pour des raisons de temps de mise en place de structure d'accueil (logements, écoles, ...) ou d'accentuation de déséquilibre de la balance des transferts.

- γ) Une variable peut être endogène dans les deux sous-modèles. Il est nécessaire d'assurer une cohérence entre la valeur calculée par le sous-modèle national et la somme des valeurs des différentes zones. Cette cohérence ne peut être assurée qu'en cadrant la valeur zonale à la valeur calculée par le sous-modèle national, ou en imaginant une série d'ajustements réciproques, convergeant vers une valeur unique comprise entre ces deux valeurs.

Le plus souvent, le rôle de l'un des deux sous-modèles est de fournir une contrainte (plus ou moins rigide) à l'autre sous-modèle. La croissance d'une branche peut ne pas pouvoir s'effectuer parce que la répartition spatiale des actifs correspondants ne correspond pas aux équipements existants ou créés⁽¹⁾, par exemple.

Un dernier cas est celui d'une variable définie à un niveau zonal, et pour laquelle il existe une valeur implicite donnée pour la valeur nationale, laquelle n'intervient pas au niveau du sous-modèle national⁽²⁾. C'est le cas, par exemple, de la somme des soldes migratoires intérieurs qui doit être nulle ou encore de la somme de part de chaque zone dans la nation qui doit être égale à 1. Ces valeurs implicites (ici 0 et 1) ne sont pas le plus souvent atteintes, lorsque l'on applique successivement à chaque zone la même chaîne de traitement (par exemple, calcul du solde migratoire de la zone en fonction de paramètres d'un sous-modèle migratoire valable pour toutes les zones et de valeurs variables exogènes pour la zone).

(1) L'une des utilisations d'un modèle spatialisé couplé avec un modèle national peut être donc de faire apparaître des contraintes spatiales de croissance.

Dans un tout autre ordre d'idée, l'économiste américain FRANKEL a construit un modèle de croissance mettant en évidence des blocages sectoriels à la croissance ; ce type de couplage n'est donc pas nouveau.

(2) Il s'agit, en fait, d'une relation comptable implicite.

Après avoir examiné les différents rôles possibles d'une variable commune aux sous-modèles nationaux et zonaux, il est nécessaire de préciser la nature des mécanismes de cohérence envisageables.

3.12 - Signification des mécanismes de cohérence

Les deux conceptions sont possibles :

- Les écarts constatés entre les deux sous-modèles sont significatifs et révèlent des contradictions entre des tendances nationales et celles des particularismes zonaux. Il est alors nécessaire de faire appel à un sous-modèle d'ajustement chargé de réduire ces tensions.
- Les écarts constatés entre les deux sous-modèles ne sont pas significatifs en soi et résultent de deux modes de calcul différents. Le mécanisme de cohérence a surtout un rôle de rigueur comptable.

C'est en fait à la première catégorie de problème de cohérence que s'adresse le paragraphe suivant.

3.2 - Quelques mécanismes de cohérence

Les mécanismes de cohérence que l'on va étudier ont ceci de commun que celle-ci s'effectue au détriment des valeurs calculées dans le sous-modèle zonal.

Il faut distinguer le cas d'une seule variable, ou de plusieurs variables indépendantes, de celui où l'on s'occupe simultanément de plusieurs variables dépendantes (c'est-à-dire liées par une ou plusieurs équations comptables).

3.21 - Cadrage sur une donnée nationale

C'est le cas le plus fréquent sans doute, deux cas peuvent se présenter : l'information est sommable (par exemple, nombre d'agriculteurs) ou elle ne l'est pas (par exemple, taux d'activité).

- α) Si l'information est sommable, le problème est facile à régler : en effet, on peut effectuer une comparaison entre la valeur nationale du modèle national et la valeur nationale du modèle zonal, obtenue comme la somme de toutes les valeurs zonales. Il suffit ensuite de multiplier chaque valeur zonale VAL (Z) par le rapport VAL (F) / Σ VAL (Z). A la suite de cet ajustement proposé (car d'autres sont possibles), la cohérence est assurée ; en effet, nous avons :

$$\sum_Z \left[\text{VAL} (Z) \cdot \frac{\text{VAL} (F)}{\sum \text{VAL} (Z)} \right] = \text{VAL} (F)$$

Il est possible de compliquer la méthode en bloquant les valeurs (considérées comme "sûres") de certaines zones en effectuant le cadrage proportionnellement au rapport de la valeur nationale diminuée de celle des cases bloquées, à la somme des valeurs zonales non bloquées.

- β) Par contre, si l'information n'est pas sommable, le problème est plus délicat, il faut en effet supposer tout d'abord que le dénominateur "population totale", du ratio envisagé⁽¹⁾, "taux d'activité" est connu pour chaque zone et que la valeur nationale calculée correspondante est cohérente avec la valeur nationale, s'il n'en est pas ainsi, il faut assurer la cohérence par une méthode du type de celle décrite précédemment.

Une fois cette première condition satisfaite, on peut calculer la population active de chaque zone en multipliant la population totale de la zone (dénominateur) par le ratio, ce qui permet après sommation de trouver le taux national du sous-modèle zonal et l'on se ramène au problème précédent, à ceci près que la cohérence n'est pas forcément assurée, pour des raisons identiques à celles exposées au 1.21. La réutilisation de cette méthode sur le nouveau taux d'activité ne garantit pas la convergence du taux national du sous-modèle zonal vers le taux du sous-modèle national.

3.22 - Cadrage sur plusieurs données nationales dépendantes

C'est le cas, par exemple, de la confrontation à une date donnée T (I) de calculs de la population active par branche effectués par un modèle national (M2 (F, B)) et de calcul donnant la population active par zone (M1 (Z)). Ces deux marges sont supposées compatibles, c'est-à-dire que la marge des marges est la même, c'est-à-dire que :

$$\sum_B \text{M2} (F, B, T (I)) = \sum_Z \text{M1} (Z, T (I))$$

(1) Pour se cantonner au cas le plus simple, icar il est évident que le problème est infiniment plus complexe pour le cas d'indicateurs composites.

Supposons qu'en outre l'on dispose d'une matrice d'emploi branche - zone compatible avec la marge des marges obtenue, soit par un modèle zonal, soit encore en conservant une structure observée pour l'année d'origine T (0) c'est-à-dire que l'on a :

$$PA (Z, B, T (I)) = \left[\sum_B PA (F, B, T (I)) \right] \cdot \frac{PA (Z, B, T (0))}{\sum_B \sum_Z PA (Z, B, T (0))}$$

Selon toute vraisemblance, nous aurons :

$$M1 (F, B, T (I)) \neq \sum_Z PA (Z, B, T (I))$$

$$M2 (Z, T (I)) \neq \sum_B PA (Z, B, T (I))$$

Ses marges M1 et M2 sont intangibles, le problème est d'essayer de distordre la matrice PA de telle sorte que les inégalités se transforment en égalité, tout en essayant de s'écarter le moins possible de la matrice d'origine. La technique du RAS que nous allons examiner rapidement a été élaborée pour résoudre ce problème.

On commence par corriger proportionnellement toutes les lignes en bloquant leurs marges à celles fournies exogènement :

$$PA_1 (Z, B, T (I)) = PA_0 (Z, B, T (I)) \cdot \frac{M1 (Z, T (I))}{\sum_B PA (Z, B, T (I))}$$

puis l'on corrige proportionnellement toutes les colonnes en bloquant leurs marges à celles fournies exogènement :

$$PA_2 (Z, B, T (I)) = PA_1 (Z, B, T (I)) \cdot \frac{M2 (Z, T (I))}{\sum_B PA_1 (Z, B, T (I))}$$

On réitère le processus :

$$PA_n (Z, B, T (I)) = PA_{n-1} (Z, B, T (I)) \cdot \frac{M1 (Z, T (I))}{\sum_B PA_{n-1} (Z, B, T (I))}$$

jusqu'à ce que l'écart entre les marges calculées et les marges exogènes soit, pour chaque zone ou chaque branche, inférieur à un seuil donné.

La convergence est en général assez rapide.

Cette technique a été utilisée, par exemple, pour la confection de la Fresque Régionale n° 2⁽¹⁾ dans le cadre d'une étude menée conjointement par l'INSEE, la DATAR et le Plan, et est actuellement utilisée dans les procédures de cadrages des propositions régionales d'équipements collectifs lors de la préparation du budget.

(1) Rapport du groupe informatique de la DATAR, Décembre 1969.

Le sous-modèle migrations intérieures

8

CHAPITRE 8 - LE SOUS-MODELE "MIGRATIONS INTERIEURES"

Le propos de ce chapitre est de rechercher un modèle migratoire explicatif satisfaisant pour la période 1954 - 1962 ce qui permettra en gardant les paramètres obtenus sur le modèle explicatif d'obtenir pour chaque département des volumes d'émigrants et immigrants intérieurs potentiels pour la période 1962 - 1968.

Mais cette recherche de modèle explicatif ne peut se faire qu'à partir d'hypothèses résultant de modèles théoriques de comportement ou d'une connaissance typologique du phénomène. Cette dernière, au niveau départemental est quasi inexistante, aussi avons nous dû procéder à une analyse du phénomène pour compléter le corps des hypothèses tel qu'il peut résulter du chapitre 4. Mais au préalable il est nécessaire de s'interroger sur la mesure du phénomène.

SECTION 1 - MESURE DES MIGRATIONS

1.1 - Problèmes posés par la définition du concept de migration

On examinera ces problèmes d'abord sous l'angle conceptuel⁽¹⁾ avant de voir les implications des définitions retenues dans les statistiques disponibles, sur l'appréhension du phénomène.

1.11 - Migrations et migrants

Tout individu est repérable au cours de son existence par ses coordonnées spatio-temporelles. L'étude globale des déplacements d'un ensemble d'individus ne peut se faire sur l'ensemble de ces informations ponctuelles. En pratique, on doit procéder simultanément à deux simplifications :

La première consiste à considérer comme équivalents un certain nombre de points de l'espace où se meut l'individu et donc à le référencer dans l'espace par rapport à une zone. La définition et l'étendue de cette zone est fonction de la finalité de l'étude entreprise. C'est ainsi par exemple que pour l'établissement d'un plan de transport en commun optimal d'une ville, il y aura intérêt à faire un découpage en îlots.

La seconde simplification consiste à ne prendre les informations qu'à des dates espacées t_1 et t_2 , quitte à reconstituer à l'aide de diverses techniques, telles que l'interpolation, les mouvements théoriques effectués par les divers individus au cours de la période. L'amplitude de cette période doit, là aussi, être fonction de l'objet de l'étude entreprise. Dans l'exemple du plan de transport celle-ci devra être de l'ordre de quelques minutes, ou quelques dizaines de minutes au plus.

(1) Voir travaux de D Courgeau cités au chapitre 4 et "Note Méthodologique sur les Migrations" de Calot, communication pour le 4ème Colloque national de démographie (avril 1973). Ces documents ne concernent que les problèmes de définition des migrants et migrations et les problèmes d'additivité lors de la consolidation de zones.

Au prix de ces deux simplifications il est possible de faire une matrice de passage où chaque date est référencée par l'une des deux dimensions de la matrice (par exemple date de départ en ligne et celle d'arrivée en colonne) tandis que les différentes zones constitutives du territoire étudié sont repérées de la même façon sur chacune des deux dimensions. Est considéré comme migrant tout individu se trouvant en t_2 dans une zone différente de celle où il se trouvait en t_1 . Chaque migrant a effectué dans la période au moins une migration.

1.12 - Conséquences pratiques des simplifications adoptées sur la saisie concrète du phénomène migratoire

Il nous semble que chacune de ces simplifications est génératrice d'une série de problèmes.

a) Corrections de l'attachement d'un individu à une zone

Dans de nombreux cas, l'objet de l'étude amène à modifier pour l'une des dates données l'attachement de l'individu à une zone. C'est ainsi que si le phénomène étudié est le changement durable de résidence de l'individu on lui affectera à chacune des deux dates, quelque soit l'endroit où il se trouve, son lieu de résidence habituelle. Cette modification d'attachement peut ne pas être la même à chacune des dates, afin de permettre l'étude de certains phénomènes socio-économiques. C'est ainsi que pour l'étude du déplacement des travailleurs on sera amené à prendre en considération la zone de résidence habituelle et celle du lieu de travail, avec une période d'une demi-journée par exemple.⁽¹⁾ Il s'ensuit que les informations résultant de cet attachement ne sont utilisables que pour une catégorie donnée de problèmes.⁽²⁾

Par ailleurs, lorsque l'attachement est le lieu de domicile habituel, se posent des problèmes particuliers en ce qui concerne les personnes présentes à l'une des deux dates dans des institutions (hôpitaux, prisons, internats) mais disposant d'une résidence propre. En France les recensements imputent à cette population particulière dite "comptée à part" le lieu de résidence habituelle (s'il existe), dans les calculs de migrations.

b) Distorsions imputables à l'espacement des observations

α) L'amplitude de la période ne permet pas dans certains cas de se faire une idée précise des mouvements réels qui importent au regard de l'étude, les diverses techniques d'interpolations déformant trop le phénomène.

(1) Si l'on définit une période plus longue, on pourra étudier des migrations saisonnières ; ou si entre deux dates assez éloignées l'attachement s'effectue sur le lieu de résidence habituel et celui de résidence temporaire on peut étudier des migrations de vacances, etc ...

(2) Les problèmes pour lesquels l'attachement et la période conviennent mais pas le découpage spatial, ne sont pas traitables avec les statistiques dont on dispose si le découpage requis est plus fin. Ils peuvent l'être si le découpage retenu est obtenu par agrégation des zones du premier découpage. Si cette dernière condition est satisfaite, l'étude n'est possible que si l'on dispose encore de la matrice interzonale des flux élémentaires, si l'on désire travailler sur des flux ou des marges de cette matrice ; tandis que si l'on désire travailler sur des balances migratoires, il suffit de posséder celles du découpage le plus fin (en effet seules les balances migratoires possèdent la propriété d'additivité dans les agrégations de zones).

C'est ainsi que l'étude, par les recensements, des changements de lieu de résidence habituelle, ne peut prendre en compte que des migrations apparentes, plusieurs migrations pouvant être en fait intervenues au cours de la période pour un même individu. Il est évident que plus l'amplitude de la période est grande plus importante est la sous-estimation du nombre de migrations réelles si l'on ne prend en compte que les migrants tels qu'ils résultent de la matrice d'échange.

β) - Un autre problème résulte de cette simplification de l'appréhension du temps : en effet, on a indiqué au début de cette section que tout individu était repérable *au cours de son existence* par ses coordonnées spatio-temporelles, mais l'une des dates d'observation (voire dans certains cas les deux) peut être extérieure aux références temporelles d'un individu donné soit parce que celui-ci est décédé soit parce qu'il n'est pas encore né. La solution pratique de ce problème dépend étroitement des méthodes d'observation.

L'utilisation d'un fichier de personnes, à condition que les mises à jour ne détruisent pas les données anciennes, permet de connaître exactement des migrations intervenues après la naissance ou avant le décès et de les comptabiliser à part.

La technique de recensement peut être utilisée de deux façons pour connaître les migrations. Tout d'abord on peut comparer⁽¹⁾ des données individuelles aux deux dates t_1 ou t_2 et en déduire pour la liste d'individus commune aux deux recensements les divers changements d'état et en particulier les flux de migrations, la liste résiduelle de t_1 correspond aux individus décédés entre t_1 ou t_2 ou ayant migré hors du territoire de référence ou encore n'ayant pas été recensés tandis que celle de t_2 correspond, aux naissances intercensitaires (ce que l'on peut trouver d'ailleurs immédiatement si l'on possède pour t_2 la date de naissance de chaque individu). La seconde façon consiste à intégrer dans le second recensement une question relative au lieu de résidence en t_1 , il n'est alors pas possible de connaître les décès puisque l'on ne connaît que la population survivante en t_2 , par contre à l'aide de la question relative au lieu et date de naissance, si elle existe, il est possible de connaître les naissances intercensitaires et si celles-ci ont eu lieu avant ou après la migration de la mère. Dans un cas comme dans l'autre les décès migrants échappent à la connaissance puisque par hypothèse on ne dispose pas de fichier de personnes.

(1) C'est une technique de ce genre qu'utilisera l'INSEE dans l'exploitation du panel au 3/365 constitué lors du dernier recensement et qui ne fournira de renseignements qu'à partir des prochains recensements.

Pour certaines raisons (en particulier de coût) il peut sembler plus intéressant de ne pas exploiter (si elle existe) l'information relative du lieu de naissance mais de procéder à un attachement arbitraire, c'est ainsi qu'en France on considère les enfants de migrants, nés entre t_1 et t_2 , comme ayant migré avec leur mère, même si la naissance est intervenue après la migration de celle-ci (1).

Il importe donc de souligner que le phénomène des migrations n'est que très imparfaitement appréhendé par les données relatives aux migrants, du moins au niveau des recensements. Les enquêtes effectuées par l'INED permettent de mieux saisir le phénomène mais avec une finesse de détail moindre, compte tenu des coûts de sondage et de la précision de l'information que l'on désire avoir.

1.2 - Les données migratoires départementales

Les seules données exhaustives existant sur la période considérée sont celles du recensement de 1962 qui ne fournit que des données concernant les migrants (et non les migrations). Comme des problèmes de définition sont communs avec ceux des échanges migratoires avec l'étranger(2) on abordera la saisie statistique des migrants avant de voir comment se situent les migrations interdépartementales par rapport à d'autres migrations interzonales.

1.21 - Définition retenue

Rappelons tout d'abord que pour la population comptée à part, la résidence du recensement, pour l'étude des migrations, est celle du domicile légal.

Examinons maintenant à l'aide du tableau carré (donné page suivante) qui s'attache aux seuls phénomènes démographiques affectant la zone I, la définition pratique retenue ainsi que certaines des distorsions qu'elle engendre ou qu'entraîne dans certains cas l'utilisation des statistiques légales de mouvement naturel. Ses renseignements concernant les personnes présentes en J en 1962 ne figurent pas, I et J jouant un rôle symétrique.

(1) On verra au 2.23 qu'en pratique les distorsions introduites peuvent être importantes (P 197)

(2) Voir chapitre 9, section 1 p 229.

		PRESENTS EN I	PRESENTS EN J OU DECEDES EN J APRES MIGRATION	A L'ETRANGER OU DECEDES A L'ETRANGER APRES MIGRATION	DECES DOMICILIES EN I ENTRE 1954 et 1962
individus nés avant 1954 se trouvant en 1954	en I	SEDENTAIRES SURVIVANTS (ou migrants multiples revenus dans leur zone de départ) (11)	EMIGRANTS INTERIEURS (12)	EMIGRANTS EXTERIEURS (13)	DECES DE SEDENTAIRES (14)
	en J	IMMIGRANTS INTERIEURS SURVIVANTS (EVENTUELLEMENT MIGRANTS MULTIPLES) (21)	MIGRANTS MULTIPLES AYANT EU ENTRE 1954 ET 1962 LEUR RESIDENCE HABITUELLE EN I (22, 23, 32, 33)		DECES D'IMMIGRANTS INTERIEURS (EVENTUELLEMENT APRES MIGRATIONS MULTIPLES) (24)
	à l'étranger	IMMIGRANTS EXTERIEURS SURVIVANTS (EVENTUELLEMENT MIGRANTS INTERIEURS MULTIPLES) (31)			DECES D'IMMIGRANTS EXTERIEURS (EVENTUELLEMENT APRES MIGRATIONS INTERIEURES MULTIPLES) (34)
individus nés entre 1954 et 1962 en I de parents se trouvant en 1954	en I	NAISSANCES SURVIVANTES DE LA POPULATION SEDENTAIRE (41)	NAISSANCES EN I D'EMIGRANTS INTERIEURS (42)	NAISSANCES EN J D'EMIGRANTS EXTERIEURS (43)	MORTALITE INFANTILE SEDENTAIRE DE PARENTS SEDENTAIRES (44)
	en J	NAISSANCES SURVIVANTES SEDENTAIRES DE PARENTS IMMI- GRANTS INTERIEURS (éventuelle- ment migrants multiples) (51)	NAISSANCES SEDENTAIRES DE MIGRANTS MULTIPLES DOMICILIES EN I LORS DE LA NAISSANCE (52, 53, 62, 63)		MORTALITE INFANTILE SEDENTAIRE D'ENFANTS DE PARENTS IMMIGRANTS INTERIEURS (54)
	à l'étranger	NAISSANCES SEDENTAIRES SURVIVANTES DE PARENTS IMMIGRANTS EXTERIEURS (61)			MORTALITE INFANTILE SEDENTAIRE D'ENFANTS DE PARENTS IMMIGRANTS EXTERIEURS (64)
individus nés entre 1954 et 1962 en J ou à l'étranger de parents se trouvant en 1954	en I	IMMIGRATION INFANTILE SUI- VIVANTE DE MIGRANTS MULTIPLES REVENUS DANS LEUR ZONE DE DEPART (71)	X		MORTALITE INFANTILE D'IMMIGRANTS MULTIPLES INTERIEURS (74)
	en J	IMMIGRATION INTERIEURE INFANTILE SURVIVANTE (81)			MORTALITE INFANTILE D'IMMIGRANTS INTERIEURS (84)
	à l'étranger	IMMIGRATION EXTERIEURE INFANTILE SURVIVANTE (91)			MORTALITE INFANTILE D'IMMIGRANTS EXTERIEURS (94)

Compte tenu du fait que l'INSEE a retenu le concept de migrant et que par convention les enfants nés entre 1954 et 1962 sont réputés migrer avec leur mère, on travaille sur les définitions suivantes de migrants par rapport au département I :

$$\text{Immigrants intérieurs} = (21) + (51) + (81)$$

Le décompte des immigrations intérieures intégrerait la partie migration multiple de (11) et leurs enfants (71), ainsi que les migrants multiples survivants c'est à dire : $[(22) + (23) + (32) + (33) + (52) + (53) + (82) + (63)]$ et les immigrants décédés en I depuis leur immigration $[(24) + (74) + (84)]$

$$\text{Emigrants intérieurs} = (12) + (42)$$

Le décompte des émigrations intérieures intégrerait le même facteur correctif de migrations multiples (mais cette fois ci survivantes ou décédées) que celui intégré pour les immigrations.

On dispose par ailleurs des statistiques de mouvement naturel qui donnent pour les naissances domiciliées en I l'ensemble de population définie dans les lignes 4 à 6 du tableau, et pour les décès domiciliés en I l'ensemble de population défini dans la colonne 4.

De ces diverses définitions il résulte qu'il n'existe pas de passage direct entre les statistiques de migrations, de migrants, et de mouvement naturel et que ceci tient autant au mode de collecte qu'aux définitions retenues.

1.22 - Place des migrants interdépartementaux

On peut tout d'abord comparer globalement le volume de migrants dans le découpage départemental avec celui d'autres découpages du territoire métropolitain :

Source : recensements INSEE

Nombre de migrants ⁽¹⁾	Période	
	1954 - 1962	1962 - 1968
intercommunaux	11 418 135	13 741 388
interdépartementaux	4 939 140	5 374 748
interrégionaux	3 285 600	3 221 072

Les travaux de Courgeau⁽²⁾ permettent d'avoir une indication sur le taux annuel moyen de migration, c'est-à-dire en tenant compte des migrations multiples et retours intervenant durant la période intercensitaire et de l'évolution de la population soumise au risque migratoire :

Source : enquête INED

Découpage spatial (définition 1968)	taux annuel moyen de migrations (‰)		croissance du taux (%)
	années 1954 à 1961	années 1962 à 1967	
logement	648	725	+ 12
commune	428	473	+ 11
département	226	253	+ 12
région	134	156	+ 16

On constate donc que globalement le phénomène migratoire s'accélère quel que soit le découpage considéré. Par ailleurs, les distorsions spatiales engendrées par ce phénomène peuvent être relativement importantes par rapport au mouvement naturel, le taux migratoire annuel moyen est plus de trois fois plus important que le taux annuel moyen d'accroissement naturel (64 ‰ en moyenne sur 1953 - 1955).

Cette importance du phénomène et ses incidences sur les problèmes spatiaux de l'emploi justifient donc la place toute particulière faite à cette étude des migrations internes dans l'ensemble de cette étude.

(1) les données départementales 54 - 62 dans le découpage 68 ont été prises dans un cahier de l'INED à paraître "Etudes sur les migrations" "Mesure de la mobilité 5 études sur les migrations internes" Cahier n° 67 Travaux et document de l'INED. Pour le niveau départemental le découpage 62 pour la période 54-62 et le découpage 68 pour la période 62-68 ont été utilisés.

(2) voir chapitre 4, page 106.

SECTION 2 - ETUDE DES MIGRATIONS INTERDEPARTEMENTALES 1962 - 1968

Nous avons traité pour les besoins de cette étude les données du tableau MI 501 du recensement de 1968 donnant quelques 600.000 données relatives à la structure des migrants selon 14 critères (la liste en est donnée dans le tableau de la page 191). Ces données rendues homogènes dans leur forme avec les autres données possédées ont pu être comparées à d'autres données possédées et notamment celles de revenu.

Cette étude a permis de dégager des renseignements intéressants sur quelques unes des caractéristiques migratoires mais d'autres très importantes, tel le sexe, ne figurant pas dans le tableau n'ont pu être considérées compte tenu des délais impartis. On examinera successivement l'origine et la destination des migrants, l'âge des migrants, les naissances intercensitaires des migrants et enfin l'influence du revenu sur les migrations.

2.1 - Origine et destination des migrants

On dégagera tout d'abord quelques caractéristiques globales avant d'examiner l'importance des différents flux.

2.11 - Quelques grandes caractéristiques

a) migrations à destination de zones de forte urbanisation

L'examen du tableau migratoire interdépartemental montre que sur les 5 375 000 migrations apparentes observées, 35% sont à destination de départements de la Région Parisienne (région de programme), dont 13% seulement sont originaires d'un département n'appartenant pas à la Région Parisienne. C'est à dire que le cinquième des

...

migrations interdépartementales s'est produit dans une zone où l'étendue du marché du travail est telle que ces migrations s'apparentent plus, de par leurs motivations, à une mobilité intra-urbaine.

Par ailleurs, on peut remarquer que le quart des migrations n'arrivant pas dans la région parisienne se dirige vers des départements où se trouvent des unités urbaines de plus de 200 000 habitants en 1968,⁽¹⁾ c'est à dire que les facteurs "région parisienne" et grandes agglomérations expliquent à eux seuls plus de la moitié des migrations interdépartementales.

D'une façon générale d'ailleurs, l'urbanisation croissante du territoire s'est faite davantage par les migrations (intérieures et extérieures) que par le mouvement naturel comme d'ailleurs l'attestent les deux dernières colonnes du tableau ci-après où l'on voit que systématiquement le bilan des migrations intérieures avec l'ensemble des catégories inférieures est positif alors que le bilan avec l'ensemble des catégories supérieures est négatif.⁽²⁾ Pour comparer les deux périodes, il faut se reporter à la première partie du tableau qui sans pouvoir mettre en évidence la hiérarchie migratoire que l'on vient de rappeler met nettement en évidence que la croissance des diverses catégories d'unité urbaine est imputable davantage au solde migratoire (intérieur et extérieur) qu'au mouvement naturel, d'où la déformation constante de la structure de population municipale vers une urbanisation accrue.

Ces observations que nous venons de faire au niveau départemental tendent à montrer que dans ce processus d'urbanisation croissante, l'attraction urbaine, pour les migrations, dépasse, du moins pour les plus grosses d'entre elles, la seule sphère d'influence de leur département.

(1) On peut ajouter que les autres flux d'amplitude supérieure à 50 000 migrants (soit 10% du total) sont tous à destination de départements ayant des unités urbaines de plus de 100 000 habitants (en 1968).

(2) Une analyse détaillée du phénomène de l'exode rural a été menée par MERLIN dans le Cahier de l'INED n° 59 (PUF 1971)

Source (2)

Catégorie de communes (délimitation 1968 pour les colonnes 1 à 8 et 1962 pour les colonnes 9 et 10)	Taux de variation annuel moyen par cause de variation (2) (unité : 100)				Structure de la population municipale				% des survivants résidant dans la catégorie de communes en 1954 des balances migratoires avec l'ensemble des catégories de communes (définition 1962)	
	1954 - 1962		1962 - 1968		(définition 1954)		(définition 1962)		de taille inférieure	de taille supérieure
	Excédent naturel	solde migratoire	Excédent naturel	solde migratoire	1954	1962	1962	1968		
Communes rurales										
de moins de 1 000 habitants	+ 0,5	- 1,1	+ 0,3	- 0,8	2 770	2 437	2 453	2 226	-	-
de 1 000 à 1 999 habitants	+ 0,5	- 0,3	+ 0,4	+ 0,1	886	818	823	786	-	-
Ensemble des communes rurales	+ 0,5	- 0,9	+ 0,3	- 0,6	3 656	3 255	3 276	3 012	-	- 5,7
Communes appartenant à unités urbaines de										
2 000 à 4 999 habitants	+ 0,6	+ 0,4	+ 0,6	+ 0,7	529	526	528	533	+ 5,7	- 4,6
5 000 à 9 999 habitants	+ 0,7	+ 0,6	+ 0,7	+ 0,8	466	473	474	485	+ 5,8	- 3,8
10 000 à 19 999 habitants	+ 0,8	+ 1,0	+ 0,9	+ 0,9	414	436	436	453	+ 6,6	- 3,1
20 000 à 49 999 habitants	+ 0,8	+ 1,3	+ 0,9	+ 1,3	655	710	708	756	+ 6,9	- 2,2
50 000 à 99 999 habitants	+ 0,9	+ 1,3	+ 0,9	+ 1,1	563	611	609	640	+ 6,9	- 1,3
100 000 à 199 999 habitants	+ 1,0	+ 1,4	+ 1,1	+ 1,2	687	759	756	809	+ 4,5	- 0,5
200 000 habitants et plus	+ 0,8	+ 1,2	+ 0,8	+ 1,2	1 485	1 588	1 580	1 664		
Agglomération de Paris	+ 0,7	+ 1,1	+ 0,8	+ 0,5	1 545	1 642	1 633	1 648	+ 3,8	
Ensemble de la France	+ 0,7	+ 0,4	+ 0,7	+ 0,5	10 000	10 000	10 000	10 000		

(2) Source : Ville et agglomérations urbaines (délimitation en 1968) - Evolution démographique 1962 - 1968 et 1954 - 1962 - INSEE 1970 - Renseignements tirés ou calculés d'après les données du tableau des pages 696 - 697 - Pour les colonnes 9 et 10, la source est l'Espace Economique Français II p. 57 (INSEE 1967).

b) Migrations vers des départements contigus

L'étude du tableau de migrations interdépartementales révèle en outre la très grande importance relative des flux d'un département à destination des départements qui lui sont contigus⁽¹⁾. L'importance de ces flux s'élève à 41 % du total des migrations et si l'on retranche tous les cas de contiguïté avec ou à l'intérieur de la région parisienne le taux se ramène à 19 % ce qui fait, compte tenu des migrations vers la région parisienne que 54 % des migrations sont prises en compte par ces seuls deux facteurs ce qui est considérable

unité : 1000

	Immigrants vers un département de la Région parisienne				Immigrants vers un département contigu		
	dont originaires d'un département			Ensemble	dont		Ensemble
	de Rég. Par.	contigu à Rég. Par.	Autres		vers Rég. Par.	Autres	
Effectifs	1 154	30	692	1 876	866	1 335	2 201
Part	215	5	129	349	161	248	409

Le tableau de la page suivante montre⁽²⁾ de façon très nette que l'importance de la mobilité des ruraux est assez faible entre 1962 et 1968, aussi bien de par son poids dans les migrations (22%) que par rapport à la population rurale puisque le taux de migration est deux fois moins important pour les ruraux que pour les urbains d'une agglomération de plus de 50.000 habitants. La taille des ménages migrants en provenance de commune rurale est plus forte qu'en moyenne (3,70 personnes contre 3,42) bien que la taille des ménages ruraux en 62 soit plus faible qu'en moyenne (3,30 personnes). Ce gonflement de la taille des ménages ruraux migrants s'explique en partie par la présence de très jeunes (0 - 5 ans) puisque leur part dans les migrants est plus forte qu'en moyenne (17,5 contre 17,1%).

(1) On trouvera la ventilation départementale de ces données dans le tableau de la page T8. 60.

(2) On commentera ultérieurement les renseignements autres que ceux relatifs à l'origine des migrants. On peut compléter ces informations en indiquant que sur 1 000 personnes recensées en 1962 n'habitant pas le même département en 1954, 95 habitaient la Région Parisienne en 1954, 301 un département contigu de celui de leur nouvelle résidence, 335 un autre département et 265 se trouvaient hors métropole en 1954 (sont exclus du champ 2,6 % de personnes pour lesquelles les réponses sont incomplètes ou la question relative à la résidence antérieure, sans objet) ces données ont été calculées à partir des informations de l'"Espace Economique Français" tome 1 (INSEE, 1965).

MIGRATIONS INTER-DEPARTEMENTALES 1962/68⁽¹⁾Unités : effectifs en milliers⁽¹⁾, taux en %

		Résident en 1962 dans un autre département			
		Commune rurale	Unité urbaine (définition 1968)		Ensemble
			< 50 000	> 50 000	
Ensemble des métropolitains	Migrants intérieurs	1 181	1 149	3 046	5 375
	Population de départ ⁽²⁾	15 299	10 265	21 804	47 367
	Taux de migration	7,7	11,2	14,0	11,3
Personnes nées avant 1962	Migrants intérieurs	975	954	2 524	4 453
	Population de départ ⁽²⁾	1 375	9 168	19 625	42 540
	Taux de migration	7,1	10,3	13,0	10,5
Chefs de ménage (en 1968)	Migrants intérieurs	319	315	938	1 572
	Population de départ ⁽²⁾	4 632	3 176	7 356	15 164
	Taux de migration	6,9	10,0	12,8	10,4
Actifs migrants (en 1968)	Migrants intérieurs	524	467	1 259	2 250
	Population de départ ⁽²⁾	6 356	4 005	9 057	19 418
	Taux de migration	8,2	11,6	13,9	11,6
Agriculteurs et salariés agricoles (en 1968)	Migrants intérieurs	48	11	14	72
	Population de départ ⁽²⁾	565	300	134	2 999
	Taux de migration	0,19	3,7	10,5	0,24

(1) Source traitement du tableau MI 501, Pour des raisons d'arrondi la cohérence en ligne n'est pas toujours assurée.

(2) Définition de la population de départ : ensemble des personnes métropolitaines en 1962, survivantes en 1968.

C'est en fait, puisque l'on ne connaît que les migrants survivants, la population soumise au risque migratoire (pour les données départementales, la résidence retenue est celle de 1962).

Les migrants d'origine rurale, quelle que soit la catégorie envisagée, sont toujours les derniers quant au taux d'émigration. Mais pour cette même donnée, ainsi d'ailleurs que pour l'importance dans les migrations intérieures, ce sont les émigrants en provenance d'agglomérations de plus de 50 000 habitants qui, et de loin, l'emportent tandis que le comportement des migrants des agglomérations de moins de 50 000 habitants reflète assez fidèlement le comportement migratoire interdépartemental moyen.

Il faut bien souligner que ce qui vient d'être dit sur les migrants ruraux n'est exact que dans le cadre des migrations interdépartementales car de très nombreuses migrations rurales se font à l'intérieur de département ; il est donc possible que, dans les migrations intercommunales, le poids des ruraux soit nettement plus important.

Donc le migrant interdépartemental actuel "modal" vient d'une zone assez fortement urbanisée pour se rendre dans une autre zone urbanisée de préférence proche ou située dans la région parisienne.

2.12 - Direction et intensité des flux

Les flux migratoires entre deux zones quelconques sont très rarement symétriques mais les balances dans de nombreux cas sont, de par leur amplitude, négligeables. Les données possédées étant issues d'un sondage au 1/4 et les erreurs commises sur chacun des flux pouvant se reporter sur le solde, on a considéré comme non significativement différents de zéro tous les soldes positifs inférieurs à 50. Ce seuil est arbitraire et ne résulte pas d'une analyse théorique par ailleurs impossible car les techniques d'échantillonnage utilisées ne permettant pas de définir de loi de probabilité, on peut ajouter que les fluctuations d'échantillonnage peuvent très bien nous amener à considérer qu'un solde est négatif alors que dans la réalité celui-ci est positif, et réciproquement. Pour rendre pratiquement nul le risque d'erreur cette "barre" a été portée à 1 000 dans une seconde étape de cette étude mais en procédant ainsi il est bien évident qu'une partie importante de l'information est perdue.

a) Nombres de balances positives

On doit tout d'abord remarquer que le coefficient de remplissage du tableau d'échange interdépartemental est relativement faible puisque sur les $(95-1)$. $(95-1)/2 = 4418$ couples de départements, on n'en trouve que 2530 à avoir une balance migratoire significativement différente de zéro ce qui représente un coefficient de remplissage d'environ 57 %, ce qui est relativement peu.

Un examen plus poussé (voir tableau et carte pages T8.61 et 62) montre qu'un département sur deux a moins de 20 balances migratoires intérieures en sa faveur et que 10 seulement dépassent le cap des 60 balances positives avec en tête les départements de l'Essonne et des Alpes maritimes. Par ailleurs, l'importance du montant des balances migratoires favorables à un département n'est que faiblement liée au nombre de ces balances, il serait donc dangereux de prendre ce paramètre comme indicateur d'attractivité. D'ailleurs, cette discordance s'explique par l'importance de quelques types de flux privilégiés, signalée précédemment. La carte du nombre de ces balances montre que la répartition spatiale de ce phénomène est loin d'être aléatoire, et en particulier tous les départements de forte attraction selon ce critère sont à une exception près groupés sur deux zones : la région parisienne et une zone qui regroupe la vallée du Rhône et le rivage méditerranéen.

b) Recherche de l'existence de filières migratoires au niveau départemental

Il a semblé intéressant d'examiner s'il était possible de mettre en évidence l'existence de groupes de départements systématiquement émetteurs ou récepteurs vis-à-vis d'autres et s'il existait une hiérarchie entre ces groupes. Pour déceler ce phénomène, on a dressé un tableau booléen de ces balances, où le département I, repéré en ligne I se voit attribuer la valeur 1 à l'intersection de cette ligne et d'une colonne J repérant un autre département J si les migrations de J vers I sont supérieures à celles de I vers J et la valeur 0 dans le cas contraire. S'il existe des filières migratoires de type transitif, c'est à dire qu'un département donné de rang n est émetteur vis à vis de tous les départements de rang supérieur et récepteur vis-à-vis de tous les départements de rang inférieur, il est possible, par des permutations appropriées de lignes et de colonnes, d'obtenir un tableau triangulaire inférieur composé exclusivement de 1. Deux obstacles s'opposent en fait à l'obtention d'un tel tableau : tout d'abord l'inexistence de relations privilégiées entre certains couples de départements, et qui on l'a vu affecte plus de 4 couples de départements sur 10 et l'existence d'effets Condorcet. Si ces derniers sont en grand nombre, on s'écartera très sensiblement du tableau théorique de référence par la présence dans le tableau triangulaire supérieur de nombreux 1.

La présence inévitable d'effets Condorcet pose le problème du critère de réorganisation du tableau permettant d'obtenir une hiérarchie des départements (ou de groupes de départements) qui minimise l'importance des balances favorables aux départements de rang inférieur. On pourrait même, si l'on travaille sur la détermination de groupes de départements, compliquer le problème en introduisant la notion de contiguïté. La recherche d'une solution optimale passe alors par l'intermédiaire de la théorie des graphes. Mais on s'est contenté ici

...

d'une solution proche de la solution optimale⁽¹⁾ et qui consiste à réordonner le tableau en rangeant les départements (ordre identique en ligne et en colonne) par ordre croissant du nombre de balances favorables que ceux-ci connaissent. On trouvera page A2 le tableau booléen réordonné.⁽²⁾ On trouve au dessus de la diagonale 424 couples de départements ayant la valeur 1, c'est à dire que dans une première approche on peut estimer qu'environ 82% des balances migratoires entre deux départements sont conformes à la filiaire migratoire théorique proposée. On n'a pu malheureusement déterminer l'importance relative des flux qui suivent cette règle et des études complémentaires sont nécessaires pour pouvoir cerner nettement ces filiaires dont l'existence d'après cette première étude, semble pouvoir être présumée.

Si l'on élève la barre de la balance migratoire entre deux zones à 1000 personnes, pour mettre en évidence que certains départements ont tendance à se vider au profit d'autres départements, on obtient seulement 149 balances. La carte de ce résultat (page T8. 03) met en évidence que ces balances concernent ou des départements contigus (ou pour une dizaine de cas, séparés par un seul département), ou un département de province et un département de la région parisienne (qui a autant d'échanges favorables que défavorables (25) avec la Province), et on assiste seulement à cinq cas de flux à longue distance : le Var vers le Finistère (ce qui s'explique par des transferts de marins de Brest à Toulon), le Nord vers le Rhône et les Bouches-du-Rhône, la Meurthe et Moselle vers le Var et le Rhône vers les Alpes maritimes.

Examinons maintenant le phénomène migratoire selon le critère de l'âge et de l'activité. Il sera pour cela fréquemment fait référence au tableau synoptique de la page suivante.

(1) D'après le spécialiste des graphes consulté, Michel HORPS. On pourrait sans doute utiliser avec profit les techniques de diagonalisation d'un tableau mises au point par J. BERTIN dans "Semiologie Graphique" p. 168 pour trouver empiriquement une solution satisfaisante, sans faire appel à la théorie des graphes.

(2) ainsi qu'aux pages T8. 56 à 59.

Catégories de population	(1)	(2)	(3 = 1 + 2)	(4)	(5)	(6)
	Migrations intérieures	Population stable	Population métropolitaine en 1962 survivante en 1968	Rapatriés d'Algérie	Etrangers	Population métropolitaine en 1968
Ensemble (2)	5 375	41 992	47 367	929	1359	49 655
Personnes nées avant 1962	4 453	38 087	42 540	790	1171	44 500
Chefs de ménage	1 572	13 592	15 164	273	325	15 763
Actifs	2 250	17 168	19 418	339	605	20 398
Ensemble	1 556	9 600	11 156	266	369	11 791
par						
âge						
en						
1968						
0 à 14 ans	1 556	9 600	11 156	266	369	11 791
15 à 24 ans	994	6 627	7 621	127	255	8 003
25 à 34 ans	1 152	4 335	5 487	137	347	5 971
35 à 64 ans	1 318	15 238	16 557	315	356	17 227
65 ans et plus	333	6 212	6 546	84	33	6 662
Ensemble	72	2 927	2 999	10	39	3 048
par						
catégorie						
1968						
agriculteurs et salariés agricoles	72	2 927	2 999	10	39	3 048
patrons de l'industrie et du commerce	164	1 754	1 918	25	13	1 955
professions libérales, cadres sup. et moyens, employés	1 058	4 673	5 731	153	113	5 996
ouvriers, personnel de service et autres	955	7 815	8 770	152	476	9 398
personnes non actives	3 104	24 844	27 948	590	719	29 256
	(7 = 1/3)	(8 = 1/2)	Part dans la population métropolitaine 1968 des			
	Taux de migrations intérieures	migrants intérieurs / population stable	(9 = 2/6)	(10 = 1/6)	(11 = 4/6)	(12 = 5/6)
			stables	migrants intérieurs	rapatriés	étrangers
Ensemble (2)	113	128	846	108	19	27
Personnes nées avant 1962	105	117	856	100	18	26
Chefs de ménage	104	116	862	100	17	21
Actifs	116	131	843	110	17	30
Ensemble	140	162	814	132	23	31
par						
âge						
en						
1968						
0 à 14 ans	140	162	814	132	23	31
15 à 24 ans	130	150	828	124	16	32
25 à 34 ans	210	266	726	193	23	58
35 à 64 ans	80	87	884	77	18	21
65 ans et plus	51	54	932	50	13	5
Ensemble	24	25	960	24	3	13
par						
catégorie						
1968						
agriculteurs et salariés agricoles	24	25	960	24	3	13
patrons de l'industrie et du commerce	85	93	896	84	13	7
professions libérales, cadres moyens et sup., employés	185	226	780	176	25	19
ouvriers, personnel de service et autres	109	122	831	102	16	51
personnes non actives	111	125	849	106	20	25

(1) Ces données ont été arrondies ce qui fait que la cohérence entre certaines lignes peut n'être pas assurée.

(2) Ces données sont issues d'un traitement effectué sur les bandes du tableau migratoire M1-501 sondage au 1/4, dans lequel 19296 émigrants du département de la Seine à destination des régions Auvergne, Bourgogne, Franche-Comté, Rhône-Alpes et Provence-Côte d'Azur n'ont été comptés. Cette erreur a été redressée pour la seule population totale et ne peut l'être pour les autres données (la population stable de la Seine étant alors gonflée d'autant) mais l'erreur relative est faible. L'écart entre la population totale sans double compte du sondage du 1/4 et celle résultant des statistiques légales, soit 124000 individus résulte des techniques utilisées.

2.2 - Influence de la structure par âge

Nous examinerons tout d'abord les données d'ensemble avant d'étudier les disparités spatiales de ce phénomène sous l'angle descriptif puis sous l'angle statistique.

2.21 - Etude globale

Il faut noter que l'histogramme des âges des migrants intérieurs diffère très sensiblement de celui de la population totale. En effet, la courbe de répartition par âge des migrants inter régionaux est bimodale : un premier mode à un an et le second à 26 ans pour les hommes et 25 pour les femmes (âge atteint en 1968); l'écart entre les deux modes correspond à celui d'une génération; par ailleurs la population migrante est nettement plus jeune que le reste de la population. Les observations faites sur cette répartition par âge des migrants inter régionaux sont d'après les travaux de Courgeau indépendantes du découpage géographique de référence et donc transposables aux migrations interdépartementales (2).

Examinons maintenant la répartition de ces derniers par tranches d'âge :

âge en 1968	Part des effectifs de la tranche d'âge considéré dans l'ensemble				rapport 2/1
	(1) des migrants intérieurs 62 - 68	(2) métropolitains en 62 survivants en 68	(3) non migrants entre 62 - 68 survivants en 68	(4) métropolitains en 1968	
0 - 14	251	236	228	238	1,234
15 - 24	186	161	158	161	1,154
25 - 34	215	116	103	120	1,857
35 - 64	246	349	363	347	0,704
> 64	62	138	148	134	0,450
Ensemble	1 000	1 000	1 000	1 000	-

Il ressort du tableau ci-dessus que 69,2 % des migrants avaient moins de 35 ans en 1968 alors que la

(1) D.Courgeau in Population - Janvier 1973

(2) A notre connaissance la pyramide annuelle de l'ensemble France entière des migrants interdépartementaux n'a pas été établie

population de ces tranches d'âge regroupe 51,3 % de la population qui était métropolitaine en 1962.

Cette plus grande jeunesse des migrants se traduit par un taux de migration de la tranche d'âge 0 - 34 ans de 15,6 % contre 11,3 % pour l'ensemble de la population "soumise au risque" et atteint 21 % pour la tranche 25 - 34, soit près du double du taux moyen observé. On peut remarquer en outre que l'apport extérieur qui explique la différence entre les colonnes 2 et 4 a pour effet de rajeunir la population métropolitaine en 1968.

Examinons maintenant les différences spatiales de ce phénomène.

2.22 - Etude descriptive départementale de la structure par âge des migrations départementales

a) Choix d'un indicateur

Les tableaux des pages T8. 64 à 70 donnent une vue d'ensemble des comparaisons de la structure par âge des émigrations avec celle de la population de départ (= population stable + émigration) ainsi que de la structure par âge des immigrations avec celle de la population intérieure d'arrivée (= population stable + immigration).

Il s'agit d'un tableau de coefficients de spécificité du rapport "part d'une tranche d'âge des migrants" d'un département à celle de la population de référence. Prenons pour bien fixer les idées l'exemple de la première ligne, seconde colonne (il y a bien entendu autant de colonnes que de classes d'âge) :

émigrants de 15 - 24 ans en provenance de l'ain/émigrants en provenance de l'ain	0,225		
population survivante de 15 - 24 ans en 68 dans l'ain en 62/population survivante 68, dans l'ain en 62	0,163	1,385	
			= 1,200
migrants interdépartementaux de 15 - 24 ans/migrants interdépartementaux	0,186		
population survivante de 15 - 24 ans en 68 métropolitaine 62/population métropolitaine 62 survivante 68	0,161	1,154	

Cet indicateur fait apparaître les écarts d'un département aux observations moyennes faites France

entière ; s'il est supérieur à 1 c'est que pour la tranche d'âge considérée, compte tenu de sa pyramide d'âge et comparativement à l'ensemble de la France, ce département est plus émetteur de migrants (ou fait moins de rétention), tandis que s'il est inférieur à 1 ce département est moins émetteur ou plus rétentif.

De même en ce qui concerne les immigrations, un coefficient de spécificité supérieur à 1 signifiera que le département est relativement plus attractif pour la tranche d'âge et moins attractif ou répulsif dans le cas contraire.

On a fixé arbitrairement et pour des raisons cartographiques la rétention accentuée pour les valeurs inférieures à 0,7, la rétention pour des valeurs comprises entre 0,7 et 0,9, l'émission moyenne pour les valeurs comprises entre 1,1 et 1,3 et l'émission accentuée pour les valeurs supérieures à 1,3.

b) Etude des émigrations

L'examen du tableau et des cartes (pages Pb. 67 à 71) que l'on en a tirées pour les cinq premières tranches d'âge, qui seules réellement importent dans les migrations mène aux réflexions suivantes :

L'émigration des 0 - 14 ans est un phénomène particulièrement stable puisque la quasi totalité des départements ont le même comportement que le comportement moyen. Toutefois, certains départements normands (Seine maritime, Calvados, Manche, Eure, Sarthe) font une rétention tandis que quatre départements méditerranéens (Pyrénées Orientales, Aude, Herault et Var) ainsi que la Creuse et la Haute Vienne sont émetteurs et que la Seine a un comportement d'émission accentuée.

Pour les 15 - 24 ans, peu de régularités se dégagent si ce n'est qu'il y a peu de coefficients de spécificité de valeur inférieure à 1 et qu'il s'agit à quelques rares exceptions près des départements les plus urbanisés, qui font donc une rétention de jeunes.

La France se sépare de nouveau en deux pour les 25 - 34 ans, qui rappelons-le, migrent deux fois plus que la moyenne, l'ouest et le sud-ouest (à l'ouest d'une ligne Avranches St-Etienne Montpellier) sont relativement plus émetteurs, le reste du pays, à l'exception des départements périphériques de la Seine (avec prolongement vers le Val d'Oise et l'Eure) qui ont un comportement de rétention, est homogène avec le comportement moyen.

...

Pour les 35 - 64 ans on retrouve à peu près le même découpage que le précédent mais l'Ouest et le Sud-Ouest sont cette fois-ci en état de rétention et la couronne parisienne (réduite à sa seule couronne) en état d'émission avec en outre les Bouches du Rhône. D'autre part, la zone de rétention se prolonge vers le nord et sur une ligne Chateauroux - St Dié.

c) Etude des immigrations

Pour les immigrations aucune régularité simple ne semble pouvoir se dégager des cartes dressées mais il est vrai que la comparaison qui est faite de la structure par âge des immigrants avec celle de la population d'accueil n'a qu'une valeur de renseignement, en particulier sur les distorsions des pyramides des âges imputables au phénomène migratoire, alors qu'implicitement la même comparaison entre émigrants et population d'origine fait appel à la notion de strate de population soumise au risque.

d) Examen de la liaison statistique entre structure par âge des migrants et populations de référence

Trois types de fonctions linéaires simples ont été testées (voir tableau de résultats page suivante). Deux d'entre elles, la première et la dernière ne font rien d'autre que de reprendre sous une autre forme l'étude qui vient d'être faite des coefficients de spécificité. Mais pour l'étude des immigrations un biais est introduit par le fait que la taille de la population d'arrivée n'est pas indépendante du volume des immigrants. Aussi pour confirmer les résultats on a remplacé dans l'équation II la population d'arrivée par celle de départ.

D'une façon générale, la part d'une tranche d'âge dans l'émigration est d'autant plus forte qu'elle l'est dans la population de départ, mais pour les 25 - 34 ans ce renseignement n'est pas significatif. Sans doute, comme on l'a observé, faudrait il tenir compte du phénomène urbain et faire deux strates, pour que cette donnée redevienne significative. Par ailleurs, la classe pour laquelle le coefficient de corrélation est le plus élevé (0 - 14 ans) est celle dont on avait noté le comportement uniforme.

La répartition des immigrants par âge est assez conforme à celle des départements d'arrivée à l'exception de la tranche des 15 - 24 ans dont la répartition semble assez aléatoire.

...

ETUDE de la STRUCTURE par AGE des MIGRANTS INTERDEPARTEMENTAUX 62 - 68
 COMPAREE à CELLE des POPULATIONS d'ORIGINE ou d'ARRIVEE METROPOLITAINE en 62 et SURVIVANTE en 68

Trois types d'équations ont été testées :

$$\begin{aligned} \text{I} \quad \frac{\text{émigrants (Age } i, \text{ Dép)}}{\text{émigrants (Dép)}} &= a \cdot \frac{\text{Population d'origine (Age } i, \text{ Dép)}}{\text{Population d'origine (Dép)}} + b \\ \text{II} \quad \frac{\text{immigrants (Age } i, \text{ Dép)}}{\text{immigrants (Dép)}} &= a \cdot \frac{\text{Population d'origine (Age } i, \text{ Dép)}}{\text{Population d'origine (Dép)}} + b \\ \text{III} \quad \frac{\text{immigrants (Age } i, \text{ Dép)}}{\text{immigrants (Dép)}} &= a \cdot \frac{\text{Population d'arrivée (Age } i, \text{ Dép)}}{\text{Population d'arrivée (Dép)}} + b \end{aligned}$$

Résultats (R est le coefficient de corrélation) :

Paramètre	Classe d'age Type d'équation	0 - 14	15 - 24	25 - 34	35 - 64	65 et plus
		a	I	0,676	1,015	0,326
	II	0,758	- 0,034	1,975	0,667	0,820
	III	0,648	- 0,005	1,582	0,416	0,732
a/σ _a	I	13,201	2,823	1,677	5,312	4,878
	II	9,801	- 0,094	13,117	4,601	11,776
	III	8,273	- 0,018	9,630	3,343	10,717
b	I	0,129	0,056	0,177	0,051	0,019
	II	0,123	0,176	- 0,020	0,023	- 0,045
	III	0,149	0,171	0,024	0,110	- 0,032
R	I	0,804	0,278	0,170	0,479	0,447
	II	0,709	- 0,009	0,802	0,427	0,770
	III	0,647	- 0,002	0,702	0,324	0,730

2.23 - Etude des distorsions introduites par les naissances intervenues entre 1962 et 1968

Il convient tout d'abord de rappeler que toute personne née entre 1962 et 1968 est considérée du point de vue migratoire comme résidant en 1962 au domicile de la mère et que d'autre part dans les statistiques légales, les naissances sont enregistrées au domicile de la mère au moment où elles se produisent, indépendamment bien sûr de l'ancienneté de résidence de la mère à son domicile. Le suivi de la population d'une zone au moyen de ces statistiques légales⁽¹⁾ est donc biaisé par le fait que, entre deux dates données, certains jeunes enfants enregistrés comme nouveaux habitants de la zone sont partis tandis que d'autres nés ailleurs sont arrivés. Ces distorsions sont sans importance si les flux d'entrées et de sorties sont de pyramides d'âges et d'importance identique, or ce n'est pas toujours le cas. L'étude de ces très jeunes migrants présente donc un double intérêt :

- permettre de déceler si la population d'une zone a tendance à vieillir ou rajeunir du fait des migrations intérieures ;
- se faire une idée de l'importance des distorsions introduites par les migrations dans la composante naissances de l'accroissement naturel d'une zone. (ce qui est indispensable dans le choix de modèle prévisionnel spatial de population).

Globalement le phénomène est important puisque 16,8 % des migrants interdépartementaux sont nés depuis le précédent recensement, alors que l'on ne trouve que 9,3 % de personnes nées après 1962 dans la population n'ayant pas changé de département.

Comme on pouvait s'y attendre avec les constatations faites précédemment pour la tranche 0 - 15 ans, les disparités spatiales de cette proportion sont plus réduites chez les émigrants (extréma 11,8 - 19,7 %) que chez les immigrants (extréma 10,1 - 21,4 %) alors que la fourchette est assez resserée dans la population stable (6,7 - 11,9 %).

D'une façon générale, l'écart entre ces proportions chez les immigrants et émigrants n'est pas très fort mais peut quand même atteindre 6,1 points (%), ce qui laisse présumer l'existence de biais non négligeable dans le calcul de certains accroissements naturels.

(1) Rappelons que de nombreuses estimations annuelles de population d'une zone sont basées sur les données du recensement précédent, auquel on ajoute l'accroissement naturel durant la période tel qu'il résulte des statistiques légales et un solde migratoire moyen (qui reconduit en fait les observations faites entre les deux recensements précédents).

Pour mesurer l'importance de ces distorsions on a fait appel à trois indicateurs (voir tableau page T8. 72 et cartes des pages T8. 73 et A3)

$$I_1 = \frac{\text{Emigrants nés après 1962} - \text{Immigrants nés après 1962}}{\text{Emigrants nés après 1962}} \cdot 1000$$

$$I_2 = \frac{\text{Emigrants nés après 1962} - \text{Immigrants nés après 1962}}{\text{Population stable} + \text{Emigrants}} \cdot 1000$$

$$I_3 = \frac{\text{Emigrants nés après 1962} - \text{Immigrants nés après 1962}}{\text{Population stable née après 1962} + \text{Emigrants nés après 1962}} \cdot 1000$$

Le premier indicateur permet de dire que pour 1000 départs, le département a enregistré une perte nette d'effectifs correspondant à I_1 . Il s'agit en quelque sorte d'un indicateur de dynamisme démographique puisque sa stabilité dans le temps entraîne s'il est positif un vieillissement de la population du département imputable au phénomène migratoire. La gamme des résultats observée est très large puisque I_1 s'étend de -2790 à 920. Deux grandes zones semblent nettement avantagées : le Bassin Parisien à l'exception de la Seine et les régions de Provence Côte d'Azur et Rhônes Alpes (voir carte page T8. 73)

Le second indicateur mesure l'incidence de la variation d'effectifs de la classe d'âge sur la population de départ du département (survivante en 1968). Seul le département de la Seine connaît un exode sensible, correspondant à plus de 3 % de la population de départ et c'est d'ailleurs lui qui pour 1000 départs enregistrait une perte nette de 920, mais il est vrai que le rapport de la balance migratoire intérieure à la population de départ est d'environ -14 %. Par contre les autres départements de la région parisienne, à l'exception de l'Essonne, enregistrent

des gains pour cette tranche d'âge, qui représentent de 1 à 5 % de la population survivante de départ. Il est probable que les problèmes de logements (liés à la rareté et à l'importance des loyers) de ce département freine l'immigration des jeunes ménages qui s'installent alors en banlieue et provoque l'émigration de ménages venant d'avoir un enfant. Cet indicateur mesure d'une certaine façon la distorsion introduite par les techniques d'estimation annuelle de la population d'un département, par le biais de la composante "naissance" du mouvement naturel, dans la population totale. On peut dire que sauf pour la région parisienne cette distorsion est faible.

Le dernier indicateur corrige d'une certaine façon le premier en permettant de voir que pour 1000 naissances qu'aurait enregistré le département en l'absence de migrations, celui-ci observe en fait une perte nette d'enfants de 0 à 6 ans se montant à I3 (voir carte page A3). Par rapport à l'indicateur précédent, I3 mesure d'une certaine façon pour un département la distorsion introduite par les migrations intérieures au niveau de la classe d'âge 0 - 6 ans (alors que I2 mesurait cette distorsion pour la population totale). Ces distorsions sauf dans le cas de la région parisienne et du département des Basses Alpes, ne dépassent pas 10 % ce qui est peu si l'on se cantonne au moyen terme et beaucoup si l'on ne devait travailler qu'à l'aide des statistiques légales sur le long terme.

2.3 - *Activité des migrants interdépartementaux*

Les statistiques dont on dispose permettent de distinguer quatre catégories socio-professionnelles (CSP) chez les migrants actifs mais avant de passer à l'étude détaillée de ces catégories au niveau départemental on examinera globalement l'activité des migrants puis celle-ci au niveau départemental sans distinction de CSP.

2.31 - Remarques d'ensemble

a) niveau d'activité des migrants

L'étude du tableau de la page 191 montre que le taux de migration des actifs (11,6 %) est légèrement supérieur à celui de la population totale (11,3 %). On peut encore décrire le phénomène en disant que 42,0 % des migrants intérieurs sont actifs contre 41,0 % pour la population soumise au risque migratoire (= population stable survivante + immigrants intérieurs = population de départ) et 40,8 % pour la population stable.

Mais la comparaison des taux d'activité départementaux des immigrants et de la population stable⁽¹⁾ révèle que seuls 19 départements, dont tous ceux de la région parisienne, ont un taux d'activité migratoire supérieur au taux d'activité de la population stable.⁽²⁾

Ces deux observations ne sont pas contradictoires : il suffit en effet de se rappeler que l'on avait observé que les migrants étaient nettement plus jeunes que les non migrants. Mais la répartition spatiale de ce phénomène est loin d'être homogène : un migrant sur deux (50,5 %) arrivés dans un département n'appartenant pas à la région parisienne a moins de 25 ans (contre 41,1 % pour la région parisienne, 47,7 % en moyenne et seulement 38,6 % dans la population stable), c'est-à-dire que le poids des jeunes inactifs est nettement plus fort chez les immigrants à destination de la province et conséquemment le taux d'activité de ces départements est plus faible.

Il convient donc de dissocier, du point de vue de l'activité, les migrants à destination de la région parisienne qui sont plus actifs et un peu plus jeunes que l'ensemble de la population stable, des autres migrants qui sont moins actifs et nettement plus jeunes que dans l'ensemble de la population stable.

b) catégories socio-professionnelles

Catégories	Part de chaque CSP dans l'ensemble			(4 = $\frac{1}{3}$)
	(1) des migrants intérieurs	(2) des non migrants	(3) de la population totale 1968	
Agriculteurs + salariés agricoles	32	171	149	0,208
Patrons de l'industrie et du commerce	73	102	96	0,738
Professions libérales + cadres + employés	470	272	294	1,593
Ouvriers + autres	425	455	461	0,940
Ensemble	1000	1000	1000	-

...

(1) Les observations sont les mêmes si l'on prend la population stable + immigrants intérieurs voir colonnes 3 et 6 du tableau de la page 17.

(2) On trouve de ce fait un taux d'activité de migrants inter-régionaux inférieur (40,6 %, à celui des migrants interdépartementaux, ce qui atteste une fois de plus l'incidence du découpage sur les observations que l'on peut faire.

Le tableau donné à la page précédente complète les indications de celui donné page 191.

Deux faits sont saillants :

- On observe une très grande disparité catégorielle des résistances aux migrations, puisque le taux de migration varie dans un rapport de 1 à presque 8. Il convient toutefois de tempérer cette remarque en rappelant que les données possédées ne permettent pas de connaître les mutations professionnelles (par exemple un migrant agricole en 1962 et ouvrier en 1968 est décompté comme migrant ouvrier) il s'ensuit que toute migration suscitée par une mutation accroît le nombre de migrants de la catégorie d'arrivée. De même, toute migration suscitée par la recherche d'un premier emploi est comptabilisée comme migration d'un actif de la catégorie. La définition utilisée a donc tendance à surestimer les taux de migrations de la troisième catégorie qui est en pleine expansion, et à un degré moindre de la dernière, tandis que les deux premières catégories, en récession, ont très normalement moins tendance à migrer en restant dans la CSP.
- Les actifs du tertiaire sont de loin les plus mobiles (la part des employés cadres et professions libérales migrants dépasse de 75 % celle des non migrants) tandis que la part des ouvriers dans les actifs migrants est très légèrement inférieure à celle de cette catégorie dans la population active stable ou observée en 1968, mais l'apport relativement important de l'extérieur (4,4 % des actifs 68 de la catégorie) vient compenser cette mobilité géographique légèrement plus faible.

2.32 - Etude départementale du niveau d'activité des migrants

On examinera successivement les taux d'émigration puis d'immigration⁽¹⁾ d'actifs, dont on a dressé deux cartes en même échelle (voir page A4 et T8. 74 à 75).

Il est frappant de constater que le taux d'émigration d'actifs est très homogène pour l'ensemble des départements : on observe seulement 10 départements ayant un taux d'émigration active extérieur à une fourchette définie à ± 10 % du taux moyen de migration active départementale : six départements sont au-dessus de la fourchette supérieure et sont groupés autour du Golfe de St-Malo (Manche, Côtes-du-Nord, Ile-et-Vilaine, Orne, Mayenne, Sarthe) et quatre départements pratiquent une rétention d'actifs, deux à l'est (Moselle, Haut-Rhin) et deux dans le midi (Vaucluse, Var)

(1) Le taux d'émigration est défini comme précédemment au rapport à la population soumise au risque (population stable + émigrants) et le taux d'immigration par rapport à la population d'arrivée (métropolitaine en 1962 (population stable + immigration)).

Par contre, et avec les mêmes paliers que ceux utilisés pour les émigrants, on constate une très grande diversité spatiale du taux d'immigration active. Ce phénomène, on l'a vu au paragraphe précédent, s'explique en grande partie par le jeu conjugué de la répartition spatiale des immigrants (1/3 en région parisienne) et le poids important des jeunes inactifs chez les migrants à destination de la Province. Seule, la Seine et ses trois départements limitrophes (Hauts de Seine, Seine Saint-Denis, Val de Marne) sont au dessus de la borne supérieure à 10 % (=46,1 %) tandis que 48 départements sont en dessous de la borne inférieure à 10 % (= 37,1 %). On observe que plus de la moitié des départements compris entre ces deux bornes sont dans le Bassin Parisien mais les distorsions imputables aux différences de structure par âge sont trop importantes pour que la représentation cartographique ait une valeur autre qu'indicative. Il faut descendre en fait au niveau des CSP pour essayer d'y voir plus clair.

2.33 - Etude départementale par CSP

Il faut tout d'abord souligner que la CSP considérée est celle connue en 1968. Les migrations qui se sont accompagnées d'une mutation professionnelle ne sont donc pas distinguées des autres. De ce fait il y a sans doute une légère surestimation des migrants de la C.S.P. "cadres et employés" dont la croissance s'est effectuée au détriment des autres C.S.P. dans une population active n'ayant que peu augmenté sur la période considérée. Par ailleurs, il importe de remarquer que les stratifications par âge et par CSP ne sont pas indépendantes et que de ce fait certaines constatations faites dans ce paragraphe et au § 22 peuvent du fait des liaisons entre ces deux phénomènes, être similaires.

On a limité cette étude aux deux dernières CSP ("cadres et employés", ouvriers et personnel de service") qui représentent à elles seules près de 90 % des migrants actifs.

Pour l'étude des phénomènes de migrations d'actifs par catégorie socio-professionnelle, il a semblé plus particulièrement intéressant d'étudier les balances migratoires plutôt que chacun des flux car pour l'économie d'un département ce n'est pas tant la mobilité de telle ou telle catégorie d'actifs qui importe (encore qu'à ce facteur soit lié le degré d'adaptabilité d'une population aux mutations nécessitées par la croissance) que ce qui lui reste une fois les migrations effectuées. Mais pour comparer entre les départements l'importance relative du phénomène, il a fallu rechercher une pondération pour les diverses balances, ce qui mène à un problème d'indicateur.

a) Choix des indicateurs

Deux indicateurs ont été retenus et calculés pour chacune des deux CSP étudiée (ces indicateurs sont du même genre que ceux définis page 198).

$$I1 (CSP (I), DEP) = \frac{IMMIGRANTS (CSP (I), DEP) - EMIGRANTS (CSP (I), DEP)}{EMIGRANTS (CSP (I), DEP)}$$

$$I2 (CSP (I), DEP) = \frac{IMMIGRANTS (CSP (I), DEP) - EMIGRANTS (CSP (I), DEP)}{STABLES (CSP (I), DEP) + EMIGRANTS (CSP (I), DEP)}$$

avec :

I = 1 \Rightarrow CSP (1) = Professions libérales, Cadres moyens et supérieurs, employés.

I = 2 \Rightarrow CSP (2) = Ouvriers, personnel de service, divers.

Le premier indicateur I1 mesure l'importance relative du gain net d'actifs du fait des migrations par rapport à l'importance des départs, c'est à dire par rapport au phénomène migratoire lui-même. Cet indicateur peut encore s'écrire :

$$I1 (CSP (I), DEP) = \frac{IMMIGRANTS (CSP (I), DEP)}{EMIGRANTS (CSP (I), DEP)} - 1$$

Pour permettre de comparer les deux indicateurs I1 et I2 dont les pondérations répondaient à deux objectifs différents, il a semblé préférable d'éviter de prendre directement le rapport des immigrants aux émigrants qui, à priori, semblait un meilleur indicateur de "vitalité" mais qu'il est possible de retrouver facilement par une simple transformation(1).

(1) On trouvera pages T8. 76 à 77 bis les divers ratios possibles pour chaque CSP retenue

En rapportant dans le second indicateur I2 la balance migratoire à la population d'origine survivante on mesure de façon correcte pour chaque département la variation des effectifs de la CSP imputable aux seules migrations intérieures puisque les décès d'actifs,⁽¹⁾ les retraites et entrées de jeunes⁽²⁾ sont pris en compte dans les flux migratoires comme dans la population stable.

Le signe de I2 est bien entendu le même que celui de I1. Il faut à ce propos souligner que le fait que I2 (ou I1) soit négatif n'implique absolument pas, même en l'absence d'immigrants extérieurs que la variation 62 - 68 d'actifs de la catégorie soit négative, d'une part parce que le solde migratoire défavorable peut être compensé dans la population stable par des rentrées de jeunes dans la profession supérieures aux retraites et décès et à'autre part, parce que certains émigrants n'étaient pas actifs en 1962. Il ne s'agit donc que d'un indicateur destiné à étudier les phénomènes migratoires.

Un raisonnement symétrique est bien entendu à tenir lorsque I2 (ou I1) est positif.

On trouvera page T8. 78 à 80 le tableau de ces indicateurs et les cartes qui en ont été tirées.

b) Etude des indicateurs retenus

A 9 exceptions près, tous les départements ont la même évolution pour les deux CSP. Pour les exceptions on se trouve toujours près d'une "opération blanche" pour les échanges migratoires, chaque solde visé représentant moins de 5 % de la population d'origine. Près de 2/3 des départements connaissent une dégradation simultanée pour les deux CSP. Il y a donc une assez forte dissymétrie dans les échanges interdépartementaux par CSP, ce qui se retrouve d'ailleurs dans l'étude des bilans migratoires de population totale ou active.

La région parisienne, à l'exception du département de la Seine est favorisée par le processus et on ne trouve que dans ces départements, les Alpes Maritimes et la Haute Savoie, des soldes migratoires positifs.

Un autre point d'ordre général mérite d'être souligné, c'est qu'il ne semble pas que la présence de grandes agglomérations dans le département joue un rôle important au niveau des gains relatifs (I2) sensibles imputables aux migrations intérieures dans chacune des CSP étudiées.

(1) Il s'agit de population survivante en 1968

(2) L'activité est définie en 1968

Le tableau de la page 191 montre que les cadres et employés ont davantage tendance à migrer que les ouvriers et personnels de service (avec un taux de migrations apparentes de 18,5 % contre 10,9 %) tout en représentant sensiblement la même part d'actifs migrants. L'examen du tableau de la page T8. 78 montre qu'en revanche la répartition spatiale des premiers est plus homogène que celle des seconds, en effet I2 (qui correspond à un taux de migrations intérieures nettes) est en général inférieur pour la CSP cadres et employés que pour la CSP ouvriers et personnels de service, ce qui est d'autant plus significatif que le taux de migration est près de 2 fois plus important pour les premiers que les seconds. La dissymétrie des échanges interdépartementaux est donc nettement plus forte pour les ouvriers et personnels de service que pour les cadres et employés.

L'examen des cartes faites à partir de I2 n'amène que peu d'observations nouvelles par rapport à celles faites sur le tableau ; pour les cadres et employés 18 des 22 départements les plus défavorisés (soldes défavorables représentant plus de 5 % de la population d'origine) se trouvent à l'ouest d'une ligne Rouen - Marseille. En ce qui concerne la dernière CSP le nombre de départements défavorisés (au sens où nous l'avons défini) est nettement plus élevé (42) de ce fait la répartition est moins concentrée, toutefois on peut noter qu'en général on ne trouve pas ces départements dans le bassin Parisien, le Rhône-Alpes et la Provence côte d'Azur.

On complètera cette approche par la vision d'ensemble donnée par la technique de régression.

c) Synthèse statistique par les régressions

On trouvera page suivante le tableau synoptique des résultats. D'une façon générale les taux d'immigrants sont mieux "expliqués" par les modèles proposés que les taux d'émigrants ; pour les CSP 2 (patrons de l'industrie et du commerce) et 4 (ouvriers, personnel de service, autres) il semble même qu'il n'y ait aucune liaison entre les structures de migrants et celles du département de départ, ce qui recoupe d'ailleurs l'observation de dissymétrie précédemment faite. Ces analyses ont été faites sur la base des données du tableau MI 501.

ETUDE DE LA STRUCTURE PAR CSP DES MIGRANTS INTERDEPARTEMENTAUX, COMPAREE A

CELLE DES POPULATIONS D'ORIGINE ET D'ARRIVEE METROPOLITAINE EN 1962 ET SURVIVANTE EN 1968

$$\begin{aligned}
 \text{I} \quad & \frac{\text{EMIGRANTS (CSP (I), DEP)}}{\text{EMIGRANTS (ACTIFS, DEP)}} = \text{a.} \frac{\text{POPULATION (CSP (I), DEP, 62)}}{\text{POPULATION (ACTIVE, DEP, 62)}} + \text{b} \\
 \text{II} \quad & \frac{\text{IMMIGRANTS (CSP (I) DEP)}}{\text{IMMIGRANTS (ACTIFS, DEP)}} = \text{a.} \frac{\text{POPULATION (CSP (I), DEP, 68)}}{\text{POPULATION (ACTIVE, DEP, 68)}} + \text{b} \\
 \text{III} \quad & \frac{\text{IMMIGRANTS (CSP (I), DEP)}}{\text{IMMIGRANTS (ACTIFS, DEP)}} = \text{a.} \frac{\text{POPULATION (CSP (I), DEP, 62)}}{\text{POPULATION (ACTIVE, DEP, 62)}} + \text{b}
 \end{aligned}$$

Paramètre	Type d'équation	Catégorie socio-professionnelle			
		Exploitants agricoles et salariés agricoles	Patrons de l'industrie et du Commerce	Professions libérales, cadres moyens et supérieurs employés	Ouvriers, personnel de service, autres
a	I	+ 0,134303	- 0,069682	0,380808	+ 0,097027
	II	+ 0,302457	+ 0,728417	0,579435	+ 0,405898
	III	+ 0,316746	+ 0,561284	0,492871	+ 0,266819
a/σ _a	I	+ 9,35	- 0,64	+ 7,36	+ 1,89
	II	+ 13,94	+ 6,66	+ 12,79	+ 8,39
	III	+ 13,65	+ 4,43	+ 13,19	+ 5,71
b	I	+ 0,014578	+ 0,070686	+ 0,351733	+ 0,402431
	II	- 0,004224	+ 0,010803	+ 0,272169	+ 0,260339
	III	- 0,005383	+ 0,029368	+ 0,294163	+ 0,321051
k	I	+ 0,692342	- 0,065035	0,600788	+ 0,189880
	II	+ 0,819539	+ 0,564166	0,795434	+ 0,652420
	III	+ 0,813923	+ 0,413821	0,804258	+ 0,505356

2.4 - Influence du revenu sur les migrations

Certaines études⁽¹⁾ ont mis en évidence le rôle du revenu dans les migrations internes de certains pays. A la base de cette liaison on postule que les migrants sont d'autant plus nombreux à partir, que leur niveau de rémunération dans la zone est faible et qu'ils iront de préférence dans les endroits où celui ci sera élevé. Ce type de liaison repose sur l'hypothèse que les migrants ont une bonne connaissance du marché de l'emploi. De plus, une certaine illusion monétaire est postulée car des différences spatiales importantes de salaires ne se traduisent pas forcément par des différences importantes de pouvoir d'achat.

En France, les seules données fiables dont on dispose au niveau départemental sont celles relatives aux salaires. Aussi est on parti des données déjà recueillies dans la rapide étude de disparité spatiale des salaires menée dans le chapitre d'introduction. On avait alors vu que les disparités spatiales de salaire moyen s'expliquaient en grande partie par des différences spatiales de structure d'emploi. Aussi a t-on mené cette étude pour chacune des deux catégories salariées dont on disposait ; la part "professions libérales" dans l'ensemble des migrants "cadres moyens et supérieurs" étant faible, on peut supposer qu'aucun biais significatif n'est introduit dans le calcul des taux d'émigrations ou d'immigrations de la catégorie).

On a calculé le salaire de ces deux catégories comme une somme pondérée des salaires des 12 catégories de base dont on disposait⁽²⁾ La structure de ces salaires 1968 est considérée comme assez stable pour être représentative du phénomène sur la période migratoire.

L'étude de ces données (voir tableau page T8. 81) à l'aide de régressions linéaires simples donne les résultats suivants (on trouvera en-dessous des coefficients de régression, leur valeur réduite, qui nous permettra de juger l'influence du salaire sur la variable expliquée).

$$\frac{EMIG (C1,D)}{POP DEP (C1,D)} = \frac{1,3625}{100000} \cdot SAL (C1,D) + 0,4120 \quad (\text{Coefficient de corrélation} = - 0,475)$$

(-5,09)

(1) Voir chapitre 4.

(2) Voir chapitre 1, p 28.

$$\frac{\text{EMIG (C2, D)}}{\text{POP DEP (C2, D)}} = - \frac{1,4473}{100000} \cdot \text{SAL (C2, D)} + 0,25255 \quad (\text{Coefficient de corrélation} = - 0,368)$$

(- 3,73)

$$\frac{\text{IMMIG (C1, D)}}{\text{POP AR (C1, D)}} = + \frac{1,4316}{100000} \cdot \text{SAL (C1, D)} - 0,06921 \quad (\text{Coefficient de corrélation} = 0,500)$$

(5,45)

$$\frac{\text{IMMIG (C2, D)}}{\text{POP AR (C2, D)}} = + \frac{1,7181}{100000} \cdot \text{SAL (C2, D)} - 0,06736 \quad (\text{Coefficient de corrélation} = 0,487)$$

(5,26)

Où :

C1 est la catégorie : Cadres moyens + cadres supérieurs + employés (+ professions libérales)

C2 est la catégorie : Ouvriers (+ autres salariés)

EMIG est l'émigration intérieure ;

IMMIG est l'immigration intérieure ;

POP DEP est la population de départ c'est à dire les métropolitains en 1968 présents en 62 dans le département D ;

POP AR est la population d'arrivée c'est à dire la partie de population 1968 du département, métropolitaine en 1962.

Les liaisons sont assez nettes : plus le salaire d'une catégorie est élevé dans une zone, plus forte est son attraction sur les salariés de cette catégorie en provenance d'autres départements, et plus faible est l'exode de ces salariés.

Ce qu'il est intéressant de noter c'est que cette liaison qui est nette au niveau de catégorie socio-professionnelle est masquée lorsque l'on essaye de relier le taux global d'émigration ou d'immigration au salaire moyen du département.

SECTION III - RECHERCHE d'un MODELE EXPLICATIF des MIGRATIONS APPARENTES INTERIEURES pour 1954 - 1962

L'étude descriptive n'a pu être faite que sur la période 1962 - 1968 ce qui biaise d'une certaine façon la démarche effectuée, car certaines hypothèses introduites pour test ne l'auraient peut être pas été si l'étude avait été menée sur la seule période 1954 - 1962. De toute façon certaines d'entre elles découlent des analyses théoriques présentées au chapitre 4 ou d'hypothèses de comportement, car l'étude descriptive est insuffisante pour alimenter en hypothèses les sous modèles.

L'étude descriptive ayant révélé une certaine dissymétrie des caractéristiques des flux d'immigrants et d'émigrants il a semblé préférable de rechercher un modèle pour chacun de ces flux plutôt qu'un modèle de solde migratoire, ce qui ne signifie pas que ce type d'approche soit sans intérêt comme le montre les travaux de l'équipe du CETEM (1).

Plusieurs des hypothèses formulées sont communes aux deux sous-modèles avec un effet postulé inverse dans un cas par rapport à l'autre. On formulera ces diverses hypothèses et leurs indicateurs avant d'examiner le choix de chaque sous modèle ainsi que les résultats obtenus pour la période de 1962 - 1968.

3.1 - Hypothèses testées et leurs indicateurs

On peut regrouper ces hypothèses autour de six thèmes : le poids du passé, l'exode rural, les revenus, l'urbanisation, l'emploi, l'influence de l'âge.

Les hypothèses à tester étant dans l'ensemble les mêmes qu'il s'agisse de régression en effectifs ou en taux, on indiquera, si nécessaire, le dénominateur utilisé pour obtenir un indicateur exprimé dans une unité comparable à celle de la variable endogène.

(1) qui ont été publiées après que cette étude soit terminée - voir "Les Migrations" collection TEM-Espace 1972 n° 3.

3.11 - Le poids du passé

L'importance sociologique des structures d'accueil ("colonies de bretons à Paris, ...) dans les phénomènes migratoires à maintes fois été soulignée dans diverses monographies. On peut donc penser que les habitants d'un département auront d'autant plus de facilités à quitter celui-ci qu'ils connaissent des personnes l'ayant déjà quitté.(1)

La variable testée est le nombre de personnes nées dans un département et n'y résidant pas en début de période. Lorsque l'on travaillera en taux d'émigration cette donnée sera rapportée à l'ensemble des personnes nées dans le département. Cette donnée ne joue que pour les migrations de rang 1 mais l'importance de celle-ci reste grande(2) (50 % pour les migrations 54 - 62). Cette variable devrait jouer positivement pour les émigrations et n'a aucune raison d'intervenir pour les immigrations.

En ce qui concerne les immigrations on peut espérer qu'elles seront d'autant plus importantes que les "structures d'accueil" dans le département seront solidement implantées. La variable retenue pour le test est le nombre de personnes qui habitent dans un département sans y être nées. Si cet indicateur est relativement élevé, on peut s'attendre à ce que le courant d'immigration se poursuive. En ce qui concerne les émigrations, on peut penser que l'influence de cet indicateur est, (si elle existe) positive. En effet, certains travaux (2) ont établi que la probabilité de migration de rang supérieur à un est plus forte que celle de rang un, il s'en suit que si le poids de cette population est relativement plus important qu'en moyenne, les émigrations devraient être plus fortes. Pour les régressions en effectifs, ce sont les effectifs de résidents non nés dans le département en 1954 qui seront pris comme indicateur tandis que pour les régressions en taux, c'est la part de cette population dans la population totale qui sera prise.(3)

3.12 - L'exode rural

On verra dans le chapitre 11(4) que la variation d'actifs agricoles était en grande partie imputable aux cessations d'activités et aux non entrées dans la profession. Il s'en suit que les enfants d'agriculteurs doivent, en tout état de cause, quitter leur milieu rural pour trouver un emploi non agricole. La nécessité de ce déplacement devrait les prédisposer plus que d'autres à migrer dans un autre département. En conséquence de quoi deux indicateurs ont été testés.

(1) Voir les travaux d'Hagerstrand (chapitre 4, page 105)

(2) Voir D. Courgeau (op. cit.)

(3) Attention dans ce cas, le dénominateur est l'ensemble des présents alors que pour le premier il s'agissait de l'ensemble des nés survivants

(4) Voir page 287 du chapitre 11.

Tout d'abord la variation d'actifs agricoles sur la période (ou le taux de variation pour les régressions en taux) dont la majeure partie résulte des décès et surtout des retraites. Les émigrations devraient être d'autant plus faibles que cet indicateur est important, la propension migratoire des retraités étant faible.

Pour le second indicateur, ne disposant pas de données fines sur les enfants d'agriculteurs pouvant entrer dans la profession durant la période (1), on a pris (pour les deux types de régression) le taux d'actifs agricoles dans la population active en supposant la part de ces migrants potentiels assez stable par rapport aux actifs agricoles en place. Une raison militait en faveur de cette solution, c'est le fait déjà signalé que près de la moitié des nouveaux agriculteurs abandonnaient la profession dans les années qui suivaient leur insertion. Il faut toutefois souligner que l'interprétation de cette donnée est délicate, car à une constante près, elle représente la part d'emplois non agricoles dans la population active, c'est-à-dire qu'une influence positive de ces derniers se traduit par un coefficient de régression négatif pour le taux d'emplois agricoles.

3.13 - Revenu

Plusieurs modèles migratoires ont mis en évidence l'influence des revenus. La méconnaissance des revenus non salariaux est telle que l'on doit se contenter des seuls revenus salariaux ce qui n'est pas très grave puisqu'on a pu voir que dans la période 1962-1968 près de 9 migrants actifs sur 10 étaient salariés, et que l'ordre de grandeur est sans doute le même pour la période 54-62. L'étude descriptive a mis en évidence l'influence du salaire au niveau de deux grandes CSP. Il n'est pas possible de descendre à un tel niveau de détail pour la période 54-62 aussi a-t-on testé les variables suivantes : salaire annuel moyen pour l'ensemble des salariés, salaire moyen des employés et salaire moyen des ouvriers. Le rejet de l'indicateur ne constitue pas pour autant une infirmation de l'hypothèse (confirmée par ailleurs page 33) mais peut traduire le fait que l'intervention du revenu dans les mécanismes explicatifs est masqué lorsqu'on travaille au niveau global. D'après les mécanismes néo-marginalistes sous jacents à cette hypothèse, à une forte valeur de cet indicateur correspondra une forte attraction migratoire et une sensible rétention des émigrations (et inversement dans le cas contraire).(2)

(1) dans la base de données possédée car on aurait pu, au prix d'un travail complémentaire mais long et coûteux, intégrer les données plus représentatives du phénomène mais la plus value obtenue aurait sans doute été disproportionnée avec l'effort fourni. Il est certain que dans l'optique d'une approche moins grossière cet effort serait à entreprendre.

(2) voir chapitre 4 et 5

3.14 - Urbanisation

Pour tenir compte du fait que l'urbanisation croissante est en grande partie imputable aux phénomènes migratoires (voir remarques du 2.11 a du présent chapitre) et que l'étude descriptive a montré sur trois strates que la propension à migrer croît sensiblement avec la taille de commune, on a introduit la taille moyenne urbaine du département dans l'ensemble des variables à tester.⁽¹⁾ L'immigration comme l'émigration devrait être d'autant plus forte que l'indicateur est élevé.

3.15 - Opportunités d'emplois

L'introduction de ce type de données est ambiguë dans la mesure où il est difficile de distinguer à posteriori les emplois offerts avant migrations, de ceux créés à cause des migrations⁽²⁾. On peut penser toutefois que, quels que soient les mécanismes qui dans la réalité interviennent, il est intéressant de tester certains indicateurs, quitte à ne pas les retenir en dernier ressort (s'ils sont statistiquement non significatifs), dans une utilisation en mode prévisionnel du modèle "descriptif". Dans le cas contraire, comme certains types d'emploi (services par exemple) sont fonction de la population du département et donc indirectement des migrations, il faudrait résoudre un système de Cramer indécomposable⁽³⁾, ce qui poserait quelques problèmes complémentaires dans le cadre de cette étude exploratoire.

(1) Cet indicateur qui sera présenté de façon complète dans le chapitre 10 page 269 est une moyenne pondérée de la taille des unités urbaines de département par la part de population départementale qui y vit. Le calcul a été effectué de façon très précise pour 1962 et 1968 pour 1954 on a utilisé une extrapolation linéaire dans une première étape se réservant la possibilité si la variable était retenue et avait un "poids" important dans la régression de faire les calculs dans la définition 1954 pour 1954 et 1962 et d'appliquer un coefficient de variation correct à la valeur 1962 dans la définition de population 1962. On peut indiquer tout de suite que ce calcul n'a pas été rendu nécessaire.

(2) à ce propos il est difficile de raisonner en terme de migrations nettes dans la mesure où de par leur structure on a pu voir que les immigrants et émigrants pouvaient être assez différents et donc ne pas être substituables du point de vue de leurs besoins.

(3) voir chapitre 3

On a donc introduit pour les régressions en effectifs la variation d'emplois non agricoles qui seuls peuvent faire l'objet d'une offre, ainsi que la variation du seul secteur tertiaire qui est le secteur en pleine expansion. Pour les régressions en taux on a introduit les variations relatives ainsi que le pourcentage de tertiaire dans la population active qui est une donnée structurelle de nature à favoriser certaines immigrations.

Par ailleurs, le marché de l'emploi peut être plus ou moins structurellement "tendu" ce qui peut expliquer certaines migrations. Si la pression de l'offre d'emploi est forte, certains départs en retraite seront retardés et certaines insertions dans la vie active avancées tandis que l'activité féminine sera plus intense. Ces tensions doivent donc se traduire par des disparités spatiales des taux d'activité des 20-65 ans par sexe. Toutefois, le caractère conjoncturel de certaines fluctuations de ces taux, souligné par tous les experts des problèmes d'emploi oblige à beaucoup de prudence ; mais la grande disparité observée pour le taux féminin qui s'étend de 25,4 % à 56,1 dépasse très certainement le cadre conjoncturel. On a aussi introduit le coefficient de variation des taux d'activités masculin et féminin bien que cette donnée combinant les effets conjoncturels de deux dates soit sans doute encore plus sujette à caution que les données ponctuelles de la date de départ. Ces quatre variables ont été testées dans chacune des approches.

3.16 - Influence de l'âge

Enfin, on a introduit des données de dynamisme démographique par l'intermédiaire de la population de moins de 35 ans présente dans le département (en effectifs et en taux selon les besoins).

Cette variable peut difficilement passer comme variable explicative au niveau des mécanismes mais, elle peut traduire une certaine stratification des comportements dans le temps.

3.2 - Choix du sous-modèle immigration intérieure

On trouvera aux deux pages suivantes les tableaux synoptiques de résultats relatifs aux choix de deux sous-modèles migratoires. Ces résultats ont été obtenus par *régression optimale* sur l'ensemble des indicateurs proposés. On a groupé dans une seule colonne tous ceux qui ne sont jamais intervenus. Enfin en-dessous des coefficients de régression, on a porté leur valeur réduite, qui suit (sous certaines hypothèses vues au chapitre 3) une loi de Student Compte tenu du nombre de degrés de liberté que l'on a dans les relations retenues, lorsque la valeur réduite, prise en valeur absolue, est supérieure à 2 il y a moins de 5 % de chance que la variable ne soit pas significative, 1 % pour 2,6 et 1 %/100 pour le seuil de 3, 4. Cette présentation sera reprise pour les différents sous-modèles. Ajoutons que les régressions, pour lesquelles une variable n'était pas significative au risque de 5 %, n'ont pas été reportées sur ce tableau synoptique. Examinons successivement les approches en taux et en effectifs.

	Nombre de variables de la regression optimale	coefficient de détermination	taille moyenne urbaine 1954	salaire moyen ouvriers 1955	part des nés dans le dép. n'y résidant plus 1954 ‰	variation relative 54-62 d'actifs agricoles ‰	part d'actifs agricoles de pop. active 1954 ‰	variation relative 54-62 d'emplois non agricoles ‰	part de résidents non nés dans le département ‰	taux d'activité masculin 20 - 65 ans en 1954 ‰		part des moins de 35 ans dans la population totale 1954 ‰	variation relative 54-62 des moins de 35 ans ‰	autres variables	constante ‰
TAUX D'EMIGRANTS INTERIEURS 1954 - 1962 (par rapport à la population départementale 1954)	2	0,67086			0,33417 (12,93)	0,12146 (3,55)									52,92
	3	0,69509			0,32770 (13,03)	0,12516 (3,78)			0,04191 (2,60)						43,00
	4	0,74924			0,26984 (10,68)		0,07382 (5,25)		0,11805 (5,80)		0,22007 (4,87)				-135,81
	5	0,78448			0,25473 (10,65)		0,09338 (6,60)	-0,12176 (-3,68)	0,17378 (7,16)		0,25611 (5,92)				-162,48
	6	0,81266	0,00001 (3,69)		0,22787 (9,56)		0,06448 (4,71)		0,15252 (6,80)		0,27983 (6,77)	-0,11562 (-3,75)			-157,37
	7	0,82845	0,00001 (4,11)		0,24410 (11,23)		0,11093 (8,29)	-0,10870 (-3,60)	0,13459 (5,71)	-0,27498 (-3,14)	0,27063 (6,89)				86,62
	8	0,84086	0,00001 (4,54)		-0,1656 (-2,50)	0,24810 (11,75)		0,09189 (6,11)	-0,11520 (-3,92)	0,13983 (6,10)	-0,25729 (-3,02)	0,28218 (7,36)			108,42

	Nombre de variables de la regression optimale	coefficient de détermination	salaire moyen employé 55	part des nés dans le dep. n'y résidant plus (1954) ‰	part d'actifs agricoles de pop. active (1954) ‰	variation relative d'emplois tertiaires (54-62) ‰	variation relative d'emplois non agricoles (54-62) ‰	part des résidents non nés dans le département ‰	taux d'activité des 20 - 65 ans en 1954 ‰		coefficient de variation 54-62 du taux d'activité féminin 20-65 ans ‰	variation relative 54-62 des actifs agricoles ‰	variation relative 54-62 des - de 35 ans ‰	variation relative 54-62 des emplois agricoles ‰	autres variables	constante ‰
									Homme	Femme						
TAUX D'IMMIGRANTS INTERIEURS 54 - 62 (par rapport à la population 54)	2	0,72928					0,28837 (5,03)	0,28183 (9,20)								-5,95
	3	0,79432		0,20306 (5,68)		0,25784 (6,26)		0,28573 (11,34)								-73,54
	4	0,82971		0,25188 (6,83)	0,08666 (4,24)			0,22060 (7,05)				0,38221 (8,01)				-81,25
	5	0,84414		0,26401 (7,50)		0,19123 (4,44)		0,21611 (7,22)	0,08642 (3,33)			0,18717 (4,05)				-106,75
	6	0,85567	-0,20607 (-2,56)	0,24772 (7,15)		0,1715 (4,04)		0,21445 (7,40)	0,06728 (2,57)			0,24070 (4,87)				-35,73
	8 (7 : non optimal)	0,86668		0,20089 (5,18)	0,09535 (3,49)	0,11209 (2,41)		0,23278 (8,14)	0,0669 (2,47)	0,10909 (2,48)	-0,12127 (2,68)	0,30500 (5,34)				-250,95
	10 9 non optimal	0,87731		0,18518 (4,82)	0,16948 (4,03)	0,14175 (2,90)	-0,17464 (-2,03)	0,24462 (7,57)	-0,28607 (-2,07)	0,08526 (3,02)	0,17348 (3,37)	-0,18457 (-3,43)	0,42202 (5,61)	-0,12127 (-2,60)		-95,68

RECHERCHE DE MODELE DESCRIPTIF DES MIGRATIONS INTERIEURES EN EFFECTIFS (1954 - 1962)

	Nombre de variables rentrant dans la regression optimale	coefficient de détermination	nés non résidents en 1954	résidents non nés dans le département 1954	variation 54-62 d'actifs agricoles	Population 1954	population âgée de moins de 35 ans	variation 54 - 62 de la population agée de - de 35 ans	variation 54 - 62 de tertiaire	variation 54 - 62 d'emplois non agricoles	taille moyenne urbaine en 1954	Autres variables	Constante
EMIGRANTS INTERIEURS (interdépartementaux)	2	0,98518	0,37787 (17,25)								0,04236 (20,07)		1762
	3	0,99129					0,14561 (22,63)	- 0,43333 (- 13,47)			0,05918 (34,58)		14370
	4	0,99470	0,25732 (15,13)	0,14397 (10,66)				- 0,25279 (- 9,68)			0,02346 (8,22)		1889
	5	0,99594		0,18842 (9,52)		- 0,22165 (- 7,85)	0,48982 (10,12)	- 0,62592 (-19,05)			0,04106 (13,60)		12185
	6	0,99711	0,13342 (5,75)	0,18700 (11,12)		- 0,17921 (- 7,15)	0,36618 (7,90)	- 0,49862 (- 14,00)			0,03300 (11,29)		6660
	7	0,99745	0,13320 (6,07)	0,18318 (11,50)		- 0,17290 (-7,27)	0,34792 (7,87)	- 0,61925 (-12,43)	0,64593 (3,28)		0,02740 (8,44)		5128
	IMMIGRANTS INTERIEURS (interdépartementaux)	2	0,98187							2,99094 (16,20)	1,03590 (7,61)		
3		0,98585							1,65584 (5,20)	1,32777 (9,84)	0,02008 (4,89)		9740
4		0,98842				- 0,17854 (- 4,63)	0,36841 (5,22)			1,66035 (18,81)	0,04451 (10,50)		11460
6 (5 non optimal)		0,98951			- 0,6357 (- 2,61)	- 0,15186 (- 3,90)	0,28424 (3,78)	0,40918 (2,74)		1,03779 (4,33)	0,05040 (10,98)		11328

3.21 - Etude des résultats des modèles d'immigrations intérieures en effectifs

Il faut tout d'abord souligner que le coefficient de détermination est très élevé puisqu'il dépasse dans tous les cas de figure 0,98. Sur la vingtaine de variables testées, aucune donnée en taux n'intervient et les régressions optimales trouvées ne possèdent pas plus de 6 variables explicatives, encore faut-il remarquer que la meilleure régression à 5 variables n'est pas optimale (l'un des coefficients de régression n'étant pas significatif au seuil retenu).

L'influence des salaires n'est pas perceptible, mais on a déjà souligné qu'un éventuel rejet statistique n'entraînait pas pour autant une infirmation de l'hypothèse.

L'interprétation des deux premières régressions optimales est ambiguë car ont été simultanément sélectionnées la variation d'emploi tertiaire et celle d'emplois non agricoles (qui inclue la précédente). Sans doute l'effet de variation des emplois tertiaires est-il plus marqué que celui des emplois secondaires, ce qui expliquerait la composition particulière de la population active migrante mais l'interprétation n'en reste pas moins délicate. On rappellera ici qu'il est difficile pour le tertiaire de dissocier la cause de l'effet, c'est à dire un accroissement de tertiaire induit par un accroissement de population, de celui qui est autonome⁽¹⁾.

L'influence de l'urbanisation est assez nette pour cette période, mais son influence est faible (moins de 5 % de la variable calculée).

L'influence de la population est, elle aussi, délicate à interpréter. En effet, interviennent toujours simultanément la population totale en 1954 du département d'arrivée et celle de moins de 35 ans. Compte tenu des signes des coefficients de régression (négatif pour la population totale et positif pour celle de moins de 35 ans) et de leur valeur respective, la résultante du jeu de ces deux variables est négative dans la plupart des cas, et le sera d'autant moins que la part des plus de 35 ans sera faible. Le résultat final n'est le plus souvent, que moyennement sensible à cette stratification par l'âge, car si 35 ans est l'âge médian, le coefficient de la population totale est égal à environ la moitié de celui de la population de moins de 35 ans.

(1) appelé par certains auteurs le quaternaire. Voir le chapitre 12 consacré aux problèmes tertiaires

Enfin intervient une dernière variable, la variation d'actifs agricoles, avec malheureusement un signe négatif, ce qui est contraire à l'hypothèse formulée. Mais cette régression arrive après une régression non optimale (rang 5) et avec une amélioration négligeable du coefficient de détermination ce qui laisse penser que cette combinaison linéaire de variable améliore le critère du coefficient de détermination mais n'est pas pour autant significative.

Il découle de toutes ces considérations que la régression la plus satisfaisante est celle obtenue avec 4 variables explicatives, l'interprétation des deux dernières étant trop ambiguë et le choix de la suivante ne se justifiant pas du fait du très faible gain obtenu pour le coefficient de détermination.

3.22 - Etude des résultats des modèles d'immigrations en taux

Une première remarque s'impose : les régressions optimales retenues sont nettement moins bonnes que celles obtenues dans l'approche en effectifs, mais comme on a pu le montrer⁽²⁾ ceci n'implique pas le rejet de l'approche en taux.

Ces régressions font toutes apparaître l'importance du poids du passé puisque les deux variables de "trend" (part des résidents non nés dans le département et part des nés dans le département n'y résidant plus) interviennent dans toutes les régressions de rang supérieur à 2. On peut remarquer que ces deux variables expliquent avec la constante, plus de 90 % des disparités expliquées par les régressions. Cette constante est toujours négative et importante pour les régressions de rang supérieur à 2 (il faut la comparer à la moyenne des taux d'immigration qui est 107,20 ‰)

(1) voir dans le chapitre 11 l'étude des taux de variation agricole (p 290 et suivantes)

(2) voir chapitre 7 & 2. 31

Contrairement à ce que l'on pourrait attendre, le taux d'immigration croît avec la part d'actifs travaillant dans l'agriculture en 1954. Or, l'introduction de cette variable dans la régression optimale de rang 3 se traduit par un saut important du coefficient de détermination qui passe de 0,73 à 0,79. Une interprétation de ces résultats basée sur l'idée que ce que l'on appréhende ici, est le complément à 1 000 de cette variable, c'est-à-dire le nombre d'actifs non agricoles pour 1 000 actifs (avec report sur la constante de l'équation de régression de l'influence de la constante d'écart 1 000) n'est guère plus satisfaisante car le coefficient de régression devient alors négatif avec ce nouveau système de référence et l'interprétation de cette liaison inverse est tout aussi délicate à réaliser. On peut ajouter qu'en général ce taux varie en sens inverse de la taille urbaine, laquelle n'intervient dans aucune des régressions. Cette liaison doit masquer une autre relation. Il s'agit là, fort heureusement, du seul résultat totalement contradictoire avec l'hypothèse formulée rencontrée au cours de toute cette étude.

A partir de là, il est relativement vain d'examiner ce qu'il en est des résultats obtenus par les autres variables. On peut quand même noter que l'immigration semble favorisée par les disponibilités d'emplois féminins, les accroissements relatifs d'emplois agricoles ainsi que par un certain dynamisme démographique naturel (% de moins de 35 ans). Par contre, l'influence négative des salariés moyens de la CSP employée est pour le moins difficile à interpréter.

Il s'agit donc là de résultats très décevants d'autant plus qu'on est en présence de deux faits contradictoires : une amélioration progressive très nette du coefficient de détermination qui passe de 0,73 à 0,87 et des résultats apparemment inexplicables (l'hypothèse est peut être mal traduite par l'indicateur retenu).

On ne peut donc retenir comme meilleure régression en taux que celle qui fait intervenir trois variables explicatives dont deux représentent l'inertie du passé et la dernière, une donnée dynamique (la variation relative d'emplois tertiaires).

3.23 - Choix du sous modèle immigrations intérieures et résultats 1968

3.231 - Choix du sous-modèle

On est amené à choisir entre les deux modèles suivants

$$(I) \text{ (Immigrants intérieurs } 54 - 62) = - 0,17854 \text{ (Population totale } 54) + 0,36841 \text{ (Population de moins de 35 ans } 54) \\ + 1,66035 \text{ (variation } 54-62 \text{ d'emplois non agricoles)} + 0,04451 \text{ (taille moyenne urbaine } 54) + 11 \text{ } 460$$

$$(II) \left(\frac{\text{Immigrants intérieurs } 54-62}{\text{Population totale } 1954} \right) = 0,20306 \left(\frac{\text{nés et ne résidant plus dans le département } 54}{\text{nés dans le département } 54} \right) \\ + 0,28573 \left(\frac{\text{résidents non nés dans le département}}{\text{population } 54} \right) - 0,074$$

Conformément aux remarques du chapitre 7 on a comparé la variance résiduelle du modèle en effectifs avec celle calculée en transformant en effectifs les résultats du sous-modèle en taux. Celles-ci sont respectivement de 122.10^6 et 919.10^6 ce qui mène à retenir l'approche en effectifs.

3.232 - Examen des résultats 54-62 du sous-modèle retenu

On trouvera pages T8. 84 à 86 les cartes et tableaux qui sont à la base de ces commentaires⁽¹⁾ Les résultats sont assez décevants comme le laissait prévoir l'importance de la variance résiduelle.

Un département sur deux connaît par ce modèle une erreur relative extérieure à la fourchette $[-15\%, +15\%]$ ce qui est considérable d'autant plus que l'amplitude de cette erreur est de $[48\%, 108\%]$. La plupart des départements où il y a sous-estimation sont groupés dans une zone située au sud de Paris et s'étendant jusqu'à Bordeaux et on trouve dans le nord-est de la région Rhône-Alpes la zone la plus étendue de surestimation.

L'approche probabiliste semble exclue pour ce type de modèle ; l'intervalle de confiance à 95 % est en effet si large qu'il perd toute portée pratique. Ceci s'explique par l'importance de l'écart-type résiduel qui dépasse 11 000 (soit environ 10 % de la moyenne départementale des immigrations), alors que l'écart-type d'une variable calculée lui est toujours supérieur⁽²⁾

(1) Les cartes les plus intéressantes sont données aussi page A 5.

(2) voir note en bas de la page 12 du chapitre 3

3.233 - Résultats 1962-1968 obtenus

Avec un tel modèle, il ne fallait pas s'attendre à de bons résultats avec une utilisation en mode prévisionnel de ce modèle descriptif (utilisation dont on a déjà souligné les limites).

α - les données utilisées :

Pour obtenir les résultats on a remplacé les données ponctuelles de 1954 par celles de 1962 et celles de la variation 54-62 d'emploi non agricole par celle observée entre 1962 et 1968⁽¹⁾ avec un facteur multiplicatif correctif de 8/6 pour tenir compte des différences d'amplitude des périodes. Le flux migratoire ainsi calculé, représente celui de la période 62-70. Il a, en conséquence, été multiplié par 0,75 pour le ramener à un flux qui corresponde à la période plus courte qui est imposée par l'étude⁽²⁾.

β - Etude des résultats :

Globalement les immigrations hors région parisienne sont sous-estimées de près de 11 % (et 8 % si on utilise la technique des coefficients de spécificité de la régression présentée au chapitre 7. La dispersion des erreurs correspond à 1/3 de surestimation et 2/3 de sous-estimation (ce qui est d'ailleurs assez logique compte tenu de l'observation précédente) ; alors qu'avant les deux tendances s'équilibraient numériquement.

Une partie des erreurs se compensant, il faut compléter cette information en indiquant que la variance résiduelle, toujours hors Région Parisienne, est de $110 \cdot 10^6$ ce qui est très important car l'écart type correspondant représente 25 % des effectifs moyens d'immigrants hors Région Parisienne. L'utilisation de la technique des coefficients de spécificité de régression n'améliore guère la précision, la variance résiduelle n'étant diminuée que de 1 %. De plus la sous-estimation n'explique que partiellement ces résultats car les variances calculées encadrant, à la proportionnelle, ces deux vecteurs calculés sur l'ensemble observé des immigrants interdépartementaux hors Région Parisienne, ne tombent respectivement qu'à $93 \cdot 10^6$ et $99 \cdot 10^6$.

L'examen des cartes établies est très décevant, aucune régularité nette ne se dégageant.

3.3 - Détermination du sous-modèle émigrations intérieures

Comme pour le sous-modèle d'immigrations intérieures, on étudiera successivement les résultats de l'approche en effectif puis de celle en taux avant de choisir le sous-modèle définitif et d'en examiner les résultats pour 1968.

- (1) dans un modèle prévisionnel il serait nécessaire de prendre des valeurs calculées et non observées, mais il a paru plus simple de ne pas reporter les erreurs commises dans d'autres sous-modèles de façon à n'observer que les erreurs imputables au sous-modèle d'immigrations intérieures en fin de parcours.
- (2) En fait la valeur de ce facteur correctif importe peu dans la mesure où l'on utilisera ultérieurement des données départementales cadrées sur des données "France Entière" supposées être données par un modèle national (et en fait observées en 1968). Ce coefficient ne tient pas compte de l'observation faite de l'accélération du phénomène migratoire sur la période 62 - 68.

3.31 - Etude des résultats de l'approche en effectifs

Les résultats obtenus sont encore meilleurs que ceux des essais effectués pour la recherche d'un modèle d'immigrations intérieures en effectifs puisque, à partir du rang 3, on dépasse 0,99 pour le coefficient de détermination. Trois facteurs seulement se dégagent dans l'explication de ces émigrations ce qui a contrario constitue autant d'infirmités statistiques des hypothèses non retenues (mais qui sont peut être fondées dans la réalité, comme on l'a déjà souligné).

Le premier d'entre eux est le poids du passé puisque dans toutes les régressions, au moins l'un des deux facteurs "nés non résidant" et "résidents non nés" intervient et ce avec un poids très important (d'environ 90 % dans l'explication).

L'influence de l'urbanisation est elle aussi très sensible puisque l'indicateur choisi est retenu dans toutes les régressions optimales, mais son poids ne dépasse pas 5 % dans l'explication du phénomène.

Reste le facteur démographique. La population totale et celle des moins de 35 ans interviennent de la même façon que pour les immigrants intérieurs. L'interprétation délicate de cette conjonction de facteurs est la même que celle faite précédemment. Intervient en plus une variable dynamique, la variation de la population des moins de 35 ans. Cette liaison n'est pas explicative et ne fait que constater une covariation des deux variables si on la considère seulement sous l'aspect démographique. Par contre, si l'on peut considérer cette variable comme un indicateur de dynamisme socioéconomique la liaison causale est rétablie. Cette liaison repose sur l'idée qu'il existe un ajustement dans le long terme de la population d'une zone aux possibilités de son économie. Dans ce cas, et à la différence des économistes classiques, le principal mécanisme de l'ajustement n'est pas la mortalité mais la non-activité provisoire ou définitive⁽¹⁾, ou la migration, laquelle s'opère surtout chez les moins de 35 ans. On peut espérer une certaine inertie de ces mécanismes mais peut être pas de leur indicateur pour des raisons d'ordre démographique, affectant l'ensemble du pays⁽²⁾.

(1) voir chapitre 9 section 3

(2) La part des moins de 35 ans dans la population totale est stable (53 %) sur les trois derniers recensements mais cette stabilité résulte d'un accroissement relatif des femmes en âge d'avoir des enfants et d'une baisse du taux de fécondité. Cette inertie est donc à apprécier en terme relatif et non absolu. On peut ajouter que ce ratio est de 43 % en 1946 et que cette stabilité relative a été facilitée par l'immigration étrangère.

Une dernière variable intervient, la variation du tertiaire. Mais elle le fait dans la régression optimale de rang 7 avec une amélioration peu importante du coefficient de détermination et un signe contraire à celui auquel logiquement on pouvait s'attendre. Mais comme dans toutes les régressions multiples de rang élevé on aboutit à des combinaisons linéaires qui ne veulent pas dire grand chose dans la mesure où une variable peut être introduite en abaissant le poids d'une variable déjà utilisée et avec laquelle elle est faiblement corrélée, ce qui à cause de quelques observations peut faire baisser la variance résiduelle de la régression sans que la liaison statistique soit probante d'une quelconque liaison fonctionnelle.

Toutes ces diverses considérations nous mènent à retenir la régression à quatre variables explicatives.

3.32 - Etude des résultats de l'approche en taux

Comme pour les immigrants l'approche en taux donne systématiquement des coefficients de détermination inférieurs à ceux obtenus dans l'approche en effectifs. Par ailleurs, les augmentations assez nettes de valeur de ce coefficient font qu'il est vraisemblable que les liaisons statistiques mises en évidence traduisent bien des liaisons significatives.

Là encore, le poids du passé est prépondérant et les deux indicateurs retenus interviennent dans toutes les régressions de rang supérieur à 2.

Du point de vue des données de l'emploi le taux de variation de l'emploi agricole est significatif pour les premières régressions, et, conformément à ce que l'on pouvait espérer, le taux d'émigration est d'autant plus faible que la variation relative d'actifs agricoles (dont le signe est négatif) est importante. A partir des régressions de rang 4 la variation relative d'emplois agricoles est remplacée par une donnée statique, la part de population active travaillant dans l'agriculture. On peut noter au passage que la liaison statistique entre ces deux phénomènes est faible (coefficient de corrélation de 0,10).

La variation relative d'emplois non agricoles n'intervient qu'assez tardivement et comme on pouvait s'y attendre, le taux d'émigration est d'autant plus faible que ce supplément d'emplois est relativement important. Le taux d'activité masculin des 20 - 64 ans intervient lui aussi vers la fin et le signe de cette variable indique

...

que plus cette activité est intense (ce qui reflète une certaine tension sur le marché de l'emploi), plus faible est le besoin d'émigrer, observation qui cadre parfaitement avec certains mécanismes théoriques classiques.

A partir des régressions optimales de rang 4 l'indicateur démographique qu'est le pourcentage de moins de 35 ans joue avec un poids tel que la constante change de signe et varie de près de 180 ‰. La variation relative de cette strate de population, dont l'interprétation ambiguë a déjà été soulignée, ne rentre en ligne de compte qu'une seule fois.

On peut enfin souligner que la variable taille moyenne urbaine n'intervient pas avant les régressions optimales de rang 6 et ce avec un poids dérisoire et que dans la dernière régression optimale on trouve une donnée de revenu avec le salaire moyen de la CSP ouvrier et que le signe du coefficient de régression est négatif, ce qui signifie qu'à de forts salaires de la CSP considérée correspond une forte rétention des émissions migratoires.

De toutes ces régressions il semble que l'on puisse retenir, compte tenu de la progression du coefficient de détermination les régressions à 5 ou à 7 variables explicatives. Cette dernière régression intègre en plus, par rapport à la précédente, la taille moyenne urbaine dont l'incidence est négligeable, et le taux d'activité masculin qui modifie complètement la constante et a un poids important. Le caractère conjoncturel souligné pour cette donnée fait que, dans le cadre d'une utilisation en mode prévisionnelle⁽¹⁾ des paramètres estimés, il semble plus judicieux de se contenter de la régression à 5 variables explicatives.

3.33 - Choix du sous-modèle émigrations intérieures et résultats 1968

3.331 - Choix du sous-modèle

On est donc amené à choisir entre les deux modèles suivants qui semblent les meilleurs pour chacune des deux approches :

$$(I) \text{ (Emigrants intérieures 54-62) } = 0,25732 \text{ (nés non résidant 54) } + 0,14397 \text{ (résidents non nés 54) } + 0,02346 \text{ (taille moyenne urbaine 54) } - 0,25279 \text{ (variation 54 - 62 de moins de 35 ans) } + 1889.$$

...

(1) pour une simple utilisation descriptive la régression multiple à 7 variables explicatives est sans doute préférable.

$$\begin{aligned}
 \text{(II)} \quad & \left(\frac{\text{émigrants intérieurs } 54 - 62}{\text{Population totale } 1954} \right) = 0,25473 \left(\frac{\text{nés non résidant } 54}{\text{nés } 54} \right) + 0,09338 \left(\frac{\text{actifs agricoles } 54}{\text{actifs } 54} \right) \\
 & - 0,12176 \frac{\Delta 54 - 62 \text{ Actifs non agricoles}}{\text{actifs non agricoles}} + 0,17378 \frac{\text{résidents non nés}}{\text{résidents } 54} \\
 & + 0,25611 \left(\frac{\text{Pop de moins de 35 ans}}{\text{Pop tot } 54} \right) \\
 & - 0,16248
 \end{aligned}$$

La variance résiduelle de la régression (I) est de 44.10^6 alors que celle calculée en transformant en effectifs les résultats du sous-modèle (II) est de 431.10^6 ce qui mène à préférer l'approche en effectifs. Il faut malgré tout rapporter que dans le cas de la régression en taux à 7 variables explicatives et que l'on n'a pas retenue, on obtient une variance résiduelle, après transformation des résultats, qui est inférieure à celle de l'approche en effectifs, ce qui illustre bien l'insuffisance d'un choix basé sur le seul coefficient de détermination.

3.332 - Examen des résultats 54 - 62 du sous-modèle retenu

On trouvera pages T8. 88 à 93 cartes et tableaux servant de base à ce commentaire⁽¹⁾ L'erreur relative s'étend de - 37 % à 29 % (moyenne des erreurs relatives absolues de 11 %) et un département sur trois connaît une erreur relative extérieure à la fourchette (-15 % + 15 %), ce qui est meilleur que pour le sous-modèle précédent.

...

(1) les cartes les plus significatives sont portées page A6.

La cartographie des écarts montre une surestimation importante des départs calculés sur l'ensemble du Massif Central, et une sous-estimation forte sur l'Ouest de la Seine Maritime aux Charentes à l'exception des trois départements les plus à l'ouest de la Bretagne et une zone à prévision correcte sur le reste du Bassin Parisien. Mais les erreurs relatives restent fortes, une trentaine de départements se situant à l'extérieur de la fourchette $\pm 15\%$.

L'amplitude des fourchettes dans l'approche probabiliste est moins forte que pour les immigrants intérieurs, ce qui est normal puisque le coefficient de détermination est supérieur, mais il n'en reste pas moins que celle-ci reste trop importante pour que l'on puisse envisager sérieusement une utilisation probabiliste de ce sous-modèle.

3.333 - Résultats 62 - 68 obtenus

α - Les données utilisées

Les données statistiques sont bien entendu celles observées en 1962. Par contre, la variation des effectifs de moins de 35 ans pose un double problème : il faudrait disposer d'un modèle démographique prévisionnel, ensuite ce modèle prévisionnel devrait avoir un horizon à 8 ans, ce qui mène à l'année 1970. En l'absence d'un tel modèle, deux attitudes sont possibles. On peut tout d'abord prendre une interpolation de la variation observée (ce qui inclut les mouvements migratoires observés). La seconde solution qui consiste à reconduire les variations 54 - 62 en se basant sur le fait que le poids des deux premières variables, ainsi que leurs variations respectives sont tels que les distorsions engendrées par cette approximation sont du second ordre par rapport aux écarts engendrés par les deux premières variables, et qu'ensuite le problème de la cohérence nationale de ces données⁽¹⁾ est compliquée par la nécessaire identité au niveau global des immigrants intérieurs et émigrants intérieurs. C'est cette seconde option qui a été choisie en s'appuyant en outre sur le fait qu'il ne s'agit là que d'un exercice grossier destiné à tester des méthodes et non à faire un véritable modèle prévisionnel.

Il a fallu ici aussi multiplier les flux migratoires calculés par le facteur correctif 6/8 pour tenir compte des différences d'amplitude intercensitaire (correction appliquée aussi pour les immigrants intérieurs)

β - Etude des résultats

Globalement les flux d'émigrants 62 - 68 hors région parisienne sont sous estimés d'environ 376000, sur les

(1) voir chapitre 7 section 3

3520000 émigrants, ce qui correspond à une erreur de prévision globale d'environ 11 % par le modèle. Une partie des erreurs se compensant, il faut compléter cette information en indiquant que la variance résiduelle, toujours hors région parisienne est de $65,10^6$ ce qui est très important car l'écart type correspondant représente près de 20 % de la moyenne. L'utilisation de la technique des coefficients de spécificité de régression n'améliore guère la précision puisque l'on observe une erreur relative d'environ 10 % au niveau global et que la variance résiduelle ne descend qu'à 59.10^6 .

L'importance de cette variance résiduelle résulte en partie de la sous estimation globale. En effet, si l'on cadre à la proportionnelle sur la donnée observée pour l'ensemble des émigrations hors Région Parisienne les deux vecteurs précédemment calculés, les variances résiduelles correspondantes ne tombent respectivement qu'à 37.10^6 et 35.10^6 .

Par rapport aux prévisions d'immigrants intérieurs, ces résultats sont nettement meilleurs en ce qui concerne l'indicateur de variance résiduelle, mais les compensations plus importantes chez les immigrants intérieurs ont conduit à un écart global sensiblement identique (et même inférieurs dans le cas d'utilisation de coefficients de spécificité).

L'examen des cartes d'écarts relatifs est très décevant car peu de régularités se dégagent. On peut seulement signaler que l'Ouest, à l'exception de la Bretagne, connaît une sous estimation, comme dans la période précédente, mais que la correction par les coefficients de spécificité de régression mène pour de nombreux départements à une surestimation notable. En fait, les seuils cartographiques retenus empêchent de déceler des régularités de comportement pour d'autres grandes zones, mais la remise en question de ces seuils ne s'imposait pas dans la mesure où le supplément d'informations obtenu par la différenciation spatiale des erreurs les plus importantes ne pouvait modifier les divers choix effectués.

3.4 - Conclusions partielles

Les résultats obtenus sont décevants pour chacun des deux sous-modèles. Mais compte tenu des statistiques de base dont on dispose il semble difficile d'intégrer de nouvelles variables représentatives des mécanismes

...

explicatifs classiques. Ceci mène à s'interroger sur l'utilisation faite des méthodes proposées, mais aussi sur l'adéquation de ce type de méthode au problème particulier des migrations. Ces remarques ne concernent bien entendu que ces deux sous-modèles et ne présument pas du jugement d'ensemble que l'on s'efforcera de porter au chapitre 14.

3.41 - Remarques sur les liaisons retenues

Deux points méritent d'être soulignés car ils expliquent sans doute en partie la médiocrité relative des résultats (mais qui ne semblent pas plus mauvais que pour les quelques autres modèles ayant dépassé le "stade du papier" et dont nous avons pu prendre connaissance).

Tout d'abord, les variables endogènes sont de nature dynamique alors que les variables exogènes retenues sont pour les 3/4 d'entre elles de nature statique. Or il est bien évident que la recherche de liaisons entre des flux et des variables de stock est sujette à caution lorsque la période est grande surtout si les mécanismes mis en jeu sont complexes. Par ailleurs les contraintes statistiques n'ont pas permis d'intégrer dans l'ensemble des variables à tester certaines variables relatives à la période précédente. Toutefois, dans un certain nombre de cas, l'introduction de l'exploitation de flux par des flux de la même période peut ne rien expliquer du tout si ces flux sont imputables en partie à la variable endogène.

Le choix de la variable endogène, dicté par les statistiques disponibles est contestable. En effet le centre de décision migratoire est le ménage et il est certain, comme on a pu l'indiquer au chapitre 4 que les problèmes d'emplois sont prépondérants pour le chef de ménage, les autres migrants du ménage pour les ménages de plus d'une personne étant en fait des migrants induits du point de vue décisionnel. Il est vraisemblable que pour obtenir un meilleur modèle il faudrait sans doute dissocier les flux migratoires en deux ou trois CSP pour obtenir des résultats sensiblement meilleurs, encore qu'un certain nombre de migrations s'accompagnent d'une mutation professionnelle. Il faudra sans doute attendre l'exploitation au prochain recensement du pannel constitué en 1968 pour que la connaissance phénoménologique et donc la constitution de modèle progresse.

3.42 - Remarque sur les limites des techniques utilisées

Ce que l'on peut appréhender statistiquement est le résultat de la mise en branle d'un grand nombre de mécanismes complexes souvent hors de portée de toute connaissance quantifiée précise et dont les poids respectifs varient dans le temps et l'espace. Il en résulte que les diverses variables (dynamiques et statiques) que l'on

peut introduire ne peuvent qu'être des indicateurs grossiers du jeu de ces mécanismes et peuvent donc être difficilement considérées comme explicatives, stricto sensu.

A la limite, s'il existait des mécanismes que l'on pouvait appréhender directement⁽¹⁾ et que les variables qui jouent dans ces mécanismes soient toutes comprises dans l'ensemble des variables à tester, les techniques utilisées ne permettraient pas forcément de déceler leur poids réel dans ces mécanismes ou même leur influence, car il faudrait que les autres mécanismes mis en jeu soient correctement pris en compte par l'ensemble des autres variables.

Une seconde remarque est importante à faire : les phénomènes de polarisation et d'influence de la distance mis en évidence dans l'étude descriptive ne sont qu'imparfaitement pris en compte par les diverses variables de stocks (taille urbaine moyenne, nés non résidents ou résidents non nés) car les dits phénomènes expriment des relations entre des couples de départements. Mais l'étude, par les techniques de régression, d'un ensemble de flux migratoires (ou par conséquent l'origine et la destination importe) se heurte à la non prise en compte du caractère concurrentiel de ces flux, qui tous ne pourront aboutir de façon privilégiée à un endroit (ou en partir) et du caractère ponctuel de cette concurrence, que des coefficients de régression à vocation globale ne peuvent prendre en compte.

(1) On ne peut à proprement parler appréhender un mécanisme mais ce que l'on veut dire ici c'est que ce type de mécanisme doit être simple c'est à dire ramenable à une équation linéaire (ce qui exclut les systèmes de Cramer indécomposables) dont toutes les variables sont exogènes, pour que leur détermination ne soit pas entachée d'erreur.

Détermination de la demande potentielle d'emploi

CHAPITRE 9 - DETERMINATION de la DEMANDE POTENTIELLE d'EMPLOI.

Dans le chapitre précédent, on a cherché à déterminer l'une des composantes de l'évolution de la population départementale au cours du temps. Deux autres composantes interviennent pour modifier la population du département durant la période considérée : les migrations extérieures et le mouvement naturel. On traitera à part le problème des migrations extérieures, dans une première section. La solution retenue pour la détermination du mouvement naturel étant simple, on l'examinera au cours d'une seconde section consacrée essentiellement à la détermination de la population potentielle en 1968. La demande d'emploi potentielle résulte de la conjonction de la population potentielle et d'un taux d'activité potentiel. Une troisième section sera consacrée à ces problèmes.

SECTION 1 - Les MIGRATIONS EXTERIEURES

La connaissance des migrations entre la France et l'étranger est, du fait des techniques de saisie d'information adoptées, aussi bonne que celle des migrations intérieures en ce qui concerne les immigrants en provenance de l'étranger, et nulle en ce qui concerne les émigrants vers l'étranger.

Nous examinerons tout d'abord nos réflexions sur le problème de saisie des données.

En ce qui concerne les immigrants extérieurs, la période récente a été marquée par la fin de la décolonisation qui a fait coexister deux courants migratoires de caractéristiques très différentes : les rapatriés et les étrangers⁽¹⁾. Nous présenterons dans une seconde sous-section notre étude de caractéristiques de ces deux populations d'après les données possédées du recensement de 1968.

(1) Les données d'immigrants extérieurs sont légèrement fausses dans leur répartition entre rapatriés et autres immigrants dans la bande de l'INSEE. Ces erreurs sont sans conséquences sur l'étude des structures.

Enfin, on recherchera un modèle explicatif satisfaisant des immigrants intérieurs non rappatriés.

1.1 - Remarque sur le problème de mesure des migrations avec l'étranger

On a vu au cours du chapitre précédent que l'on ne peut pas appréhender toutes les migrations intervenues durant la période intercensitaire, mais les seuls migrants survivants. C'est ainsi que l'on obtient dans un même tableau les données relatives aux immigrants intérieurs et celles relatives aux immigrants en provenance de l'étranger. Notons que les conditions d'habitat de ces derniers laissent présumer que le taux de non recensés ou de non réponse est plus fort pour cette catégorie qu'en moyenne. Peut on essayer de déterminer le volume d'émigrants à destination de l'étranger à l'aide des statistiques de recensements et de mouvement naturel. On pourrait le penser en tenant compte du fait que logiquement, le solde migratoire extérieur s'obtient à partir de l'équation fondamentale : (variation de population) - (excédent naturel) - (solde migratoire intérieur) - (solde migratoire extérieur) = 0 et que l'on a donc :

$$\text{Emigrants extérieurs} = (\text{excédent naturel}) + (\text{solde migratoire intérieur}) + \text{immigrants extérieurs} - (\text{variation de population}).$$

Pour voir ce qu'il en est reportons nous au tableau de la page 180 du chapitre 8 pour voir à quoi correspondent exactement les divers concepts utilisés puis pour examiner les distorsions introduites par cette approche

1.11 - Les grandeurs observées

La population du département I en 1954 (définition recensement 1954) est celle définie par la première ligne du tableau :

$$X(11) + X(12) + X(13) + X(14) + (\text{décédés en J où à l'étranger})$$

Cette population en 1962 est celle définie par la première colonne du tableau :

$$X(11) + X(21) + X(31) + X(41) + X(51) + X(61) + X(71) + X(81) + X(91)$$

En 1962, les individus nés en I entre 1954 et 1962 (lignes 4 à 6) sont ou vivants (colonnes 1 à 3) ou morts (colonne 4).

Les naissances domiciliées se définissent donc comme :

$$X (41) + X (42) + X (43) + X (44) + X (51) + X (52) + X (53) + X (54) + X (61) + X (62) + X (63) + X (64)$$

Les décès domiciliés sont définis dans la quatrième colonne du tableau :

$$X (14) + X (24) + X (34) + X (44) + X (54) + X (64) + X (74) + X (84) + X (94)$$

Sont considérées au recensement comme migrantes, les personnes ayant migré entre 1954 et 1962, ainsi que leurs enfants, quelque soit le lieu de leur naissance. Il s'en suit que :

$$\text{Immigrants intérieurs survivants} = X (21) + X (51) + X (81)$$

$$\text{Emigrants intérieurs survivants} = X (12) + X (42)$$

$$\text{Immigrants extérieurs survivants} = X (31) + X (61) + X (91) + X (34)$$

$$\text{Emigrants extérieurs 54-62 (survivants ou non)} = X (13) + X (43)$$

1.12 - Détermination des émigrants extérieurs

Transcrivons l'équation de détermination des émigrants extérieurs dans le tableau de la page suivante où les diverses variables sont repérées par leurs initiales avec leur signe tel qu'il résulte de l'équation de détermination. Pour obtenir les émigrations réelles on doit faire intervenir une variable d'ajustement qui "balance" les cases autres que X (31) et X (43). Pour obtenir une balance générale nulle il faut prendre comme signe implicite pour les émigrants extérieurs le signe moins. On appellera émigrants extérieurs apparents la somme des émigrants extérieurs réels et de la variable d'ajustement.

			Individus se trouvant en 1962			décès domiciliés en I entre 1954 et 1962
			présents en I	présents en J ou décédés en J après migration	à l'étranger ou décédés à l'étranger après migration	
			1	2	3	4
Individus nés avant 1954 se trouvant en 1954	en I	1	(PT62 -) (PT54 +)	(EI -) (PT 54 +)	ⒺⒺ (PT54 +)	(D -) (PT54 +)
	en J	2	(PT62 -) (II +)			(D -) A +
	A l'étranger	3	(PT62 -) (IE +)			(D -) A +
Individus nés entre 1954 et 1962 en I, de Parents se trouvant en 1954	en I	4	(PT62 -) (N +)	(EI -) (N +)	ⒺⒺ (N +)	(D -) (N +)
	en J	5	(PT62 -) A - (II +) (N +)	A - (N +)	A - (N +)	(D -) (N +)
	à l'étranger	6	(PT62 -) A - (IE +) (N +)	A - (N +)	A - (N +)	(D -) (N +)
Individus nés entre 1954 et 1962 en J ou à l'étran- ger de Parents se trouvant en 1954	en I	7	(PT62 -) A +			(D -) A +
	en J	8	(PT62 -) (II +)			(D -) A +
	à l'étranger	9	(PT62 -) (IE +)			(D -) A +

PT62 : population totale définition 1962

PT54 : population totale définition 1954

EI : émigrants intérieurs (définition recensement)

EE : émigrants extérieurs (définition recensement)

II : immigrants intérieurs (définition recensement)

IE : immigrants extérieur (définition recensement)

N : naissances

D : décès

A : ajustement

Comme on peut le constater grâce au tableau, la liaison à corriger si l'on utilise la méthode de décomposition des variations est :

Ajustement :

$$\begin{array}{l}
 + \left\{ \begin{array}{l}
 X (24) : \text{décès d'immigrants intérieurs} \\
 X (34) : \text{décès d'immigrants extérieurs} \\
 X (74) : \text{mortalité infantile d'immigrants multiples intérieurs revenus en I} \\
 X (84) : \text{mortalité infantile d'immigrants intérieurs} \\
 X (94) : \text{mortalité infantile d'immigrants extérieurs} \\
 X (71) : \text{immigration infantile survivante de migrants multiples revenus en I}
 \end{array} \right. \\
 \\
 - \left\{ \begin{array}{l}
 X (51) : \text{naissances survivantes sédentaires de parents immigrants intérieurs} \\
 X (52) : \text{naissances en I d'enfants de migrants multiples intérieurs hors de I en 1954 et 1962} \\
 X (53) : \text{naissances en I d'enfants de migrants multiples en J en 1954 et à l'étranger en 1962} \\
 X (62) : \text{naissances en J d'enfants de migrants multiples à l'étranger en 1954 et en J en 1962} \\
 X (63) : \text{naissances en I d'enfants de migrants multiples à l'étranger en 1954 et 1962} \\
 X (61) : \text{naissances sédentaires survivantes de parents immigrants extérieurs}
 \end{array} \right.
 \end{array}$$

Donc le volume correct d'émigrants extérieurs (vivants ou décédés après migration) se définit comme suit :

$$\text{Emigrants extérieurs} = \text{émigrants extérieurs apparents} + \text{Ajustement}$$

La méconnaissance complète des variables qui interviennent dans l'ajustement et l'impossibilité de pouvoir affirmer qu'il y a compensation à peu près satisfaisante entre les variables intervenant positivement et négativement dans l'ajustement, fait qu'il ne nous est pas possible d'étudier les émigrations départementales vers l'étranger.

1.2 - Etude des immigrants extérieurs 1962 - 1968

On étudiera globalement le phénomène ayant d'en examiner les principaux aspects spatiaux. Pour des commodités d'exposition on appellera *rapatriés* les immigrants extérieurs en provenance d'Algérie et *étrangers* les immigrants extérieurs en provenance des autres pays (bien que certains d'entre eux soient de nationalité française).

1.21 - Etude globale

On a vu que les statistiques tirées de l'exploitation du tableau MI - 501 ne fournissent que des migrations apparentes. Dans le cas des migrations intérieures on peut penser que les "fuites" hors frontières des migrations ne sont pas trop importantes, par contre les migrations d'étrangers sont pour beaucoup d'entre elles, loin d'être définitives et les retours sont assez importants.⁽¹⁾

Il faut tout d'abord souligner que 4,6 % des métropolitains en 1968 se trouvaient à l'étranger en 1962 ce qui est beaucoup si l'on songe que 9,6 % des métropolitains en 1968 sont nés depuis 62 de parents présents sur le territoire français en 1962. C'est donc environ 1/3 de l'accroissement net (2) de population qui est imputable aux migrations extérieures. Dans cet accroissement la part des rapatriés d'Algérie est importante puisqu'elle représente 40 % de l'apport extérieur.

On examinera successivement les différences de structures par âge et d'activité de ces nouveaux arrivants par rapport à la population métropolitaine, et ceci au niveau global.

a) structure par âge

On peut compléter les informations du tableau de la page 192 du chapitre 8 par celles du tableau suivant.

- (1) Sans pouvoir quantifier ces retours on peut avoir une idée de l'importance du phénomène en regardant le chiffre d'affaire des deux ou trois sociétés spécialisées qui organisent par centaines de trains complets ou charters les retours périodiques au pays de travailleurs étrangers.
- (2) net puisque les données d'où sont tirés les renseignements correspondent aux personnes survivantes en 1968 et que pour obtenir l'accroissement brut il faudrait tenir compte des décès. Mais la pondération reste la même que l'on raisonne en terme d'accroissement brut ou net.

REPARTITION DE 1000 PERSONNES PAR TRANCHES D'AGE

Catégorie de population	0 - 14 ans		15 - 24 ans	25 - 34 ans	35 - 64 ans	65 ans et plus	Ensemble
	ensemble	dont 0-6ans					
rapatriés	286	(150)	137	147	339	90	1000
étrangers	271	(139)	187	255	262	24	1000
mi-grants intérieurs interdépartementaux]	291	(172)	186	215	246	62	1000
population stable	228	(93)	158	103	363	148	1000
population 68	238	(104)	161	120	347	134	1000

L'apport extérieur se traduit par un rajeunissement de la population française ; en l'absence de cet apport on aurait observé que 51,9 % au lieu de 51,3 % de la population était âgée de moins de 35 ans. Mais ceci est davantage dû aux étrangers qui sont responsables de 65 % de ce rajeunissement (= part des étrangers dans les immigrants extérieurs de moins de 35 ans). On peut ajouter que l'âge médian des étrangers doit se situer un peu en dessous de 30 ans car 45,8 % d'entre eux ont moins de 25 ans et 71,3 % moins de 35 ans.

Le faible poids des plus de 65 ans chez les rapatriés doit s'expliquer par le retour en France des retraités dès leurs droits acquis ; une partie de ces retours est antérieure à 1962.

Enfin, une autre idée généralement répandue est infirmée par ces données : les étrangers arrivés en France ne sont pas exclusivement des travailleurs, 28,6 % d'entre eux sont âgés de moins de 14 ans. Toutefois, sur le plan familial, cette donnée n'est pas directement comparable au 23,6 % de personnes âgées de moins de 14 ans en

1968 et métropolitaines en 1962, dans la mesure où le nombre moyen d'enfants par ménage est plus fort chez les étrangers. Du point de vue de la scolarisation ce phénomène est loin d'être négligeable car il représente en moyenne environ 3 % de la population scolaire et que la distribution géographique du phénomène fait que dans certains départements (comme celui de la Seine) ce taux fait plus que doubler.

b) Activité des immigrants extérieurs

Les renseignements globaux tirés de l'exploitation du MI - 501 sont réunis dans le tableau synoptique suivant :

unité : 1000

catégories de population	Répartition de 1000 actifs par catégorie professionnelle en 1968				taux d'activité
	Agriculteurs et salariés agricoles	Patrons de l'industrie et du commerce	Professions libérales Cadres moyens et supérieurs employés	ouvriers, personnel de service et autres	
rapatriés	30	72	451	447	365
étrangers	101	25	163	711	471
migrants intérieurs	32	73	470	425	420
population stable	171	102	272	455	409
population 1968	149	96	294	461	411

En ce qui concerne les rapatriés leur niveau d'activité est particulièrement bas et ce phénomène n'est pas imputable à la structure par âge puisque le poids de la classe d'âge 25 - 64 ans est plus fort chez les rapatriés que chez les métropolitains (68 comme 62). Il faut sans doute rechercher l'explication de ce phénomène dans l'importance des anticipations, volontaires ou non, de retraite et une activité féminine plus réduite. Par ailleurs, il semblerait que les migrations de rapatriés se soient accompagnées de nombreuses mutations au profit essentiellement d'emplois tertiaires.

...

On n'est pas surpris de constater que les étrangers sont plus actifs qu'en moyenne. Deux facteurs semblent expliquer cette supériorité : la structure par âge qui fait que 51,7 % des étrangers ont entre 25 et 64 ans contre 46,7 % en moyenne ce qui explique environ le tiers de la différence observée des taux d'activité ; et l'importance du travail féminin chez les étrangers qui doit sans doute expliquer l'essentiel du restant. Ces immigrants se trouvent à plus de 80 % dans des emplois manuels ce qui n'étonne guère quand on sait que dans certaines branches consommatrices de main-d'oeuvre peu qualifiée comme la branche automobile, près du quart des postes sont tenus par des étrangers.

1.22 - Etude départementale des migrants extérieurs

On n'examinera que l'implantation spatiale de ces migrants car les différenciations spatiales d'âge et de structure d'activité de ces migrants présentent un intérêt moindre dans le cadre de cette étude. On trouvera pages T9. 29 à 43 les cartes et tableaux qui ont servi de base à ces commentaires (ainsi d'ailleurs que des tableaux et cartes de structures par âge et CSP), deux de ces cartes ont été reproduites page A 7.

En ce qui concerne les migrants étrangers leur volume diminue par rapport à la période précédente puisqu'il passe de 1807000 pour 54 - 62 à 1359000 pour 62 - 68 mais ce premier flux intègre les premiers retours liés à la décolonisation et donc n'est pas directement comparable. Toutefois cette diminution n'est pas répartie de façon homogène dans l'espace : la Région Parisienne accueille à peu près le même volume d'étrangers sur les deux périodes (respectivement 479000 et 462000) ce qui compte tenu des différences d'amplitude des périodes correspond à une accélération du phénomène alors que dans le restant du territoire c'est l'inverse qui se produit.

La distribution spatiale est telle que l'on trouve plus de 50 % des étrangers dans 13 départements (qui sont par ordre décroissant : Seine, Haut de Seine, Bouches du Rhône, Seine St Denis, Rhône, Val de Marne, Nord, Yvelines, Alpes Maritimes, Moselle, Val d'Oise, Essonne, Isère) ce qui correspond d'ailleurs par rapport à la période précédente à une diminution de la concentration car alors 9 départements seulement groupaient plus de 50 % des étrangers. En ce qui concerne les rapatriés 50 % d'entre eux sont groupés dans 14 départements.

...

Cette distribution spatiale des effectifs absolus doit être complétée par une étude du poids relatif de ces deux catégories dans la population 1968. Les cartes établies des ratios⁽¹⁾ de chaque catégorie sont à peu près identiques qu'il s'agisse de la part dans la population totale ou dans la population active. On constate que les étrangers comme les rapatriés ont un poids relatif important dans la Région Parisienne et au sud d'une ligne Bordeaux - Genève à l'exception des départements du Massif Central et que dans le Sud d'Ouest le poids relatif des étrangers n'est relativement important que sur les départements frontaliers.

Donc par rapport au phénomène de migrations intérieures, les migrations extérieures se caractérisent spatialement par une très grande polarisation spatiale.

(1) les seuils correspondent à 70 % , 90 % , 110 % et 130 % du ratio moyen France Entière

1.3 - Recherche d'un sous-modèle des immigrants extérieurs

On présentera tout d'abord les hypothèses formulées avant d'examiner ce qu'elles deviennent à l'épreuve des faits et les résultats du sous-modèle retenu pour 1968.

1.31 - Choix des hypothèses

On retrouve tout naturellement certaines des hypothèses testées pour les sous-modèles de migrations intérieures et d'autres qui découlent de caractéristiques spécifiques à ces migrants extérieurs mises en évidence dans la sous section précédente.

a) hypothèses reprises de l'étude des migrations intérieures

On retrouve tout d'abord l'hypothèse relative à l'influence de l'urbanisation sur l'immigration. En matière d'immigration étrangère celle-ci joue un rôle sans doute légèrement différent : diverses études de ce type de filières migratoires ont montré que dans de très nombreux cas ces migrants arrivent sans contrat de travail et se dirigent donc vers des zones susceptibles de leur procurer rapidement une embauche c'est à dire de grandes agglomérations. Bien souvent donc, en dépit des textes réglementaires, l'offre d'emploi précède la demande. On a complété cet indicateur par la population départementale en 1954 pour voir si des effets de polarisation se dégagèrent.

On peut là encore penser que l'attraction des zones de forts salaires sera plus importante que dans les autres. Toutefois, on ne peut guère s'avancer sur la relative transparence du marché pour les premières migrations d'étranger ni sur la pertinence des indicateurs de salaire dont on dispose, dans l'étude de ce phénomène bien particulier.

En ce qui concerne les tensions d'emploi on a repris malgré leurs défauts, les taux d'activités des 20 - 65 ans par sexe.

Ces indicateurs ont été testés tels quels dans les essais de régression en taux comme en effectifs.

....

b) hypothèses spécifiques aux immigrants extérieurs

L'influence des structures d'accueil joue un rôle surement plus important en ce qui concerne les immigrants étrangers dont le déracinement est souvent total, que pour les immigrants intérieurs. Il s'en suit que de nouveaux indicateurs ont été introduits : le nombre d'étrangers, le nombre de naturalisés ainsi que leur somme (ce dernier indicateur a été utilisé exclusivement des deux autres dans l'un des tests effectués)⁽¹⁾. Le nombre d'étrangers présente l'inconvénient d'intégrer la variable endogène, mais si ceci est exact pour la France entière, il n'est pas sûr que les immigrants extérieurs arrivés après 1954 ne soient spatialement répartis comme ceux arrivés avant 1954. On a conservé en outre le nombre de résidents non nés dans le département qui inclut les précédents et indique une certaine prédisposition au "brassage" de population. Dans l'étude des régressions en taux ces indicateurs ont tous été rapportés à la population départementale en 1952.

Enfin pour tenir compte de la catégorie très particulière d'emploi occupée par la main d'oeuvre étrangère en France, on a introduit le nombre de manoeuvres présents dans le département en 1954 (rapporté à la population active dans l'approche en taux). L'importance de ces effectifs reflète normalement l'existence d'un plus ou moins grand marché pour ce type d'emploi qui est contraint à une très grande mobilité d'embauche et débauche du fait d'une grande interchangeabilité des individus, liée à la très faible qualification requise.

Ce dernier indicateur est statique aussi a-t-on introduit la variation d'emplois dans le secondaire pour tenir compte du fait que la fixation définitive de ces migrants dépend en grande partie de l'évolution d'offres d'emplois par l'industrie. Mais comme on l'a remarqué pour les migrants intérieurs, l'interprétation d'une bonne liaison statistique avec cet indicateur est ambiguë car dans certains cas, l'offre d'emploi ne s'est faite que parce qu'une demande potentielle existait préalablement.

1.32 - Etude des résultats

On trouvera page suivante le tableau synoptique des résultats obtenus. On peut faire deux remarques préalables avant d'examiner les résultats de chaque approche.

- l'approche en effectif semble, si l'on s'en tient au seul critère du coefficient de détermination, nettement

(1) Le groupe de variables testées par la technique de la régression optimale est définie à l'aide d'une croix de Saint André, dans le tableau synoptique.

Par ailleurs, les conventions décrites page 214 sont reprises dans ce tableau, comme elles le seront d'ailleurs chaque fois que les techniques de régression optimales seront utilisées.

RECHERCHE d'un SOUS-MODELE EXPLICATIF en EFFECTIFS

Nombre de variables de la régression optimale	Coefficient de détermination	Salaire moyen employés 1955	Salaire moyen ouvriers 1955	Salaire moyen 1955	Résidents non nés dans le départ. résidents	Taux d'activité 20 - 65 ans hommes (1954)	Taux d'activité 20 - 65 ans femmes (1954)	Δ 54 - 62 secon. secondaire 54	Etrangers + natu. population totale	Naturalisés population totale	Autres variables	Constante
Groupe de variables		X	X	X	X	X	X	X	X		X	
2	0,71402				0,00009 (4,72)				0,00037 (8,91)			- 0,01144
4 (3 non optimale)	0,74875	- 0,00031 (- 3,88)	- 0,00070 (- 4,52)	0,00084 (5,90)					0,00046 (12,76)			0,03735
5	0,77167	- 0,00028 (- 3,66)	- 0,00062 (- 4,12)	0,00077 (5,59)				0,00005 (2,89)	0,00043 (11,93)			0,02694
7 (6 non optimale)	0,50063	- 0,00032 (- 4,33)	- 0,00091 (- 5,18)	0,00090 (5,84)		0,00007 (3,40)	- 0,00009 (- 2,77)	0,00006 (3,82)	0,00039 (9,98)			0,05462
Groupe de variables		X	X	X	X	X	X	X	X	X	X	
2	0,80385				0,00007 (4,19)					0,00096 (12,45)		- 0,00427

RECHERCHE d'un SOUS-MODELE EXPLICATIF en EFFECTIFS

Nombre de variables de la régression optimale	Coefficient de détermination	Taille urbaine en 1954	Salaires annuels en 1955	Salaires moyens en 1955	Δ 54 - 62 secondaire	Etrangers et naturalisés en 1954	Naturalisés en 1954	Autres variables	Constante
Groupe de variables		X	X	X	X	X		X	X
2	0,94355	0,01512 (10,88)				0,45023 (13,89)			3388
3	0,95633	0,01276 (9,69)			0,38500 (4,99)	0,47549 (16,33)			1275
5 (4 non optimale)	0,96133	0,01201 (9,36)	- 306,068 (- 3,27)	240,23 (2,88)	0,36034 (4,55)	0,46825 (14,04)			9973
Groupe de variables		X	X	X	X	X	X	X	
2	0,96375	0,01504 (14,08)					1,01061 (18,66)		4517
3	0,97299	0,01331 (13,56)			0,32471 (5,39)		1,04125 (21,99)		2936

meilleure que l'approche en taux. Ce qui n'amène pas pour autant à rejeter d'emblée cette dernière approche comme on a déjà eu l'occasion de le souligner⁽¹⁾ Le coefficient de détermination est très élevé, mais la variance de la variable expliquée est elle aussi très importante.

- les régressions faisant intervenir les naturalisés sont toujours meilleures que celles faisant intervenir l'ensemble des naturalisés et étrangers ; la stratification est donc justifiée statistiquement. Peut être faut il en chercher la raison dans le fait que l'importance des naturalisés dans un département traduit l'existence de courants migratoires permanents et stables déjà assimilés par le pays alors que la répartition spatiale des étrangers est, elle, sans doute davantage sujette aux aléas conjoncturels.

a) Etude des résultats de l'approche en effectifs

On ne trouve pas de régression optimale de rang supérieur à 3 lorsque l'on introduit les seuls naturalisés. Par contre, lorsque l'on travaille avec l'indicateur "étrangers + naturalisés" il existe des régressions optimales de rang 5 et 6 mais dont l'interprétation économique est pour le moins délicate; aussi a t'on laissé la seule régression de rang 5 à titre d'information sans avoir pu en trouver d'explication cohérente (dans cette régression) le signe du coefficient de régression des salaires moyens ouvriers est négatif tandis que celui du salaire moyen toutes CSP est positif ; par ailleurs la valeur de ces coefficients conduit à multiplier la constante par 8).

Quelque soit l'indicateur retenu du "poids du passé", c'est toujours la donnée de taille urbaine qui intervient avec lui dès la régression de rang 2. La variation d'emplois secondaires intervient toujours comme troisième variable explicative, alors que des données à priori plus appropriées comme le nombre de manoeuvres n'interviennent pas. Peut-être aurait il fallu, pour cette dernière variable, introduire la variation 54-62 de manoeuvres dans le groupe de variables testées pour déceler l'influence de la catégorie d'emploi sur les migrations extérieures encore que rien ne soit moins sûr dans la mesure où le turn-over des immigrés est important et que d'autre part la variation de ce type d'emploi est très certainement fortement corrélée à celle des emplois secondaires (et donc implicitement pris en compte).

b) Etude des résultats de l'approche en taux

On ne trouve pas de régression optimale de rang supérieur à 2 pour l'approche tentée sur le groupe de variables incluant le taux de naturalisés mais cette régression donne un meilleur coefficient de détermination que toutes les régressions optimales intégrant le taux de naturalisés étrangers. Dans le cadre de cette

(1) page 164

dernière approche on ne trouve pas de régression optimale de rang 3 et 6 et l'interprétation des 3 variables de salaires, qui sont constamment présentes, semble soulever des difficultés insurmontables. Ceci mène à délaisser ces régressions malgré les gains importants effectués au niveau des coefficients de détermination.

La seconde variable qui intervient dans les deux régressions de rang 2 est la part de résidents dans un département et qui y sont nés et pour laquelle, on l'a vu⁽¹⁾ les nés hors-métropole interviennent en moyenne pour 26 %. L'hypothèse formulée reçoit donc une confirmation statistique mais rappelons qu'il reste tout à fait possible que le mécanisme implicite sous-jacent n'existe pas (ce "risque de seconde espèce" est d'ailleurs inéluctable pour toutes les interprétations retenues).

1.33 - Choix du sous-modèle et étude de ses résultats

a) Choix du sous-modèle

Il découle des diverses considérations développées ci-dessus que les deux modèles en concurrence sont les suivants :

$$(I) \text{ (immigrants extérieurs 54-62) } = 0,01331 \text{ (taille urbaine 54) } + 0,32471 \text{ (variation 54-62 d'emplois secondaires) } \\ + 1,04125 \text{ (naturalisés 54) } + 2936$$

$$(II) \frac{\text{immigrants extérieurs 54-62}}{\text{population totale 54}} = 0,00007 \frac{1000 \text{ (résidents non nés 54)}}{\text{population totale 54}} + 0,00096 \frac{1000 \text{ (naturalisés 54)}}{\text{population totale 54}} - 0,00427$$

. (1) Voir la note 2 de la page 186

Le modèle (II) donne après transformation en effectifs des taux calculés une variance résiduelle de 1169.10^6 (en comparaison avec le vecteur observé) tandis que cette variance dans le cadre de l'approche en effectif est seulement de 42.10^6 , ce qui reste malgré tout considérable car l'écart type correspondant représente en gros le tiers de l'effectif moyen départemental. C'est donc le modèle 1 qui sera retenu.

b) Etude des résultats de la période 54 - 62

On trouvera pages A8 et T9. 44 à 48 les tableaux et cartes qui sont à l'origine de ces commentaires.

Les résultats sont assez mauvais : la moyenne des écarts absolus relatifs est de 36 % et l'étendue des erreurs du modèle descriptif est importante : 5 départements sont sous estimés de plus 100 % (sans que l'écart soit supérieur à 9000) et 9 sont surestimés de 50 à 61 % (avec un écart supérieur à 8000 dans un seul cas). On peut difficilement imputer ces résultats à la non-homogénéité constatée dans la répartition spatiale des immigrants extérieurs puisque l'on observe à peu près autant d'écarts positifs que d'écarts négatifs dans l'ensemble des départements comme dans le groupe des départements les plus récepteurs. En valeur absolue ces écarts restent toutefois acceptables. 24 départements connaissent un écart de plus de 5000 et 9 seulement un écart supérieur à 10000.

Compte tenu de l'importance relative des écarts, les cartes qui ont été établies à l'aide des seuils standards ne sont pas très nuancées. On peut relever toutefois que la zone comprenant le Bassin Parisien et le Nord de la France est surestimée tandis que la couronne limitrophe est sousestimée ainsi que les départements côtiers atlantiques et méditerranéens, à l'exception de la Vendée, l'Aude et les Alpes maritimes.

Compte tenu du fait que l'écart type résiduel est important (le tiers de la moyenne départementale) et de la concentration du phénomène sur quelques départements, il n'est pas surprenant de constater que l'approche probabiliste fournisse des résultats sans intérêt pratique : près des 2/3 des départements ont une limite inférieure de la fourchette définie du coefficient de confiance à 95 % qui vient buter sur la valeur zéro.

c) Etude des résultats de la période 62 - 68

On examinera les données utilisées dans le modèle prévisionnel avant d'en examiner les résultats.

...

α - Les données utilisées

Les deux informations ponctuelles utilisées (nombre de naturalisés et taille urbaine) ont été prises à la date de 1962 tandis que la variation d'emplois secondaires est celle observée sur la période, corrigée par un facteur de 8/6 pour tenir compte des différences d'amplitude des périodes. Le flux migratoire ainsi calculé est relatif à la période 62 - 70 puisque les coefficients de régression ont été calculés pour une période de 8 ans, en conséquence le résultat obtenu a été multiplié par 0,75 pour fournir un flux migratoire correct.

β - Etude des résultats 62 - 68

Deux phénomènes sont frappants :

Tout d'abord le flux des immigrants extérieurs est surestimé par le modèle pour l'ensemble de la Province d'environ 60 % ce qui est considérable ; 8 départements seulement sur les 87 sont sous estimés alors que pour la période précédente il y avait à peu près autant de cas de surestimation que de cas de sous estimation. En outre, 34 départements sont surestimés de plus de 100 %.

La répartition des écarts importants dans le cas de surestimation est loin d'être aléatoire. Si l'on retient le seuil de 10000, sur les 14 départements concernés (et qui ensemble font les 2/3 de la surestimation globale), 10 sont situés au sud de la Loire. On peut ajouter que pour 6 d'entre eux on observe un changement important de comportement puisque d'une période sur l'autre le flux migratoire qui était supérieur à 50000 dans la première période, diminue de plus de moitié au cours de la période suivante.

Il semblerait que l'afflux des rapatriés a contrecarré l'évolution naturelle de l'immigration étrangère. D'autres arguments militent en faveur de cette interprétation : tout d'abord les rapatriés se sont massivement installés dans les départements qui attirent proportionnellement le plus les immigrants extérieurs et quelque 153000 emplois ouvriers ont été pris par eux (ce qui s'est fait en partie au détriment des immigrés extérieurs).

...

Ce renversement de tendance n'ayant pas eu lieu de façon homogène sur l'ensemble du territoire il est normal que l'utilisation de la technique des coefficients de spécificité de régression⁽¹⁾ n'apporte aucune amélioration. Pour la même raison il n'est pas étonnant de constater qu'après cadrage du vecteur calculé sur l'ensemble observé des immigrants extérieurs hors Région Parisienne dans l'ensemble l'amélioration soit sensible (8 départements seulement surestimés de plus de 100 %) mais qu'à cause de quelques départements le gain réalisé par la variance résiduelle n'est pas aussi important qu'il aurait pu l'être ; en effet celle ci, en ce qui concerne les résultats sans coefficient de spécificité, passe avec ce cadrage de 16.10^5 à 3.10^5 .

(1) décrite au chapitre 7 § 2.12 page 161

SECTION 2 - DETERMINATION DE LA POPULATION POTENTIELLE EN 1968

On dispose maintenant de trois informations calculées ; immigrants intérieurs, immigrants extérieurs et émigrants intérieurs. Le phénomène des rapatriés correspond à un accident politique et ne relève pas de la prévision économique, on doit donc considérer que les données de rapatriés sont fournies exogènement au modèle (malgré l'incidence soulignée sur les immigrations d'étrangers dans le sud de la France).

La connaissance de la population en 1962 étant par définition acquise, il ne reste plus en théorie qu'à calculer l'accroissement naturel de la population stable, puisque l'accroissement naturel des migrants est implicitement pris en compte dans leur détermination (avec toutes les erreurs imputables aux déformations de pyramide des âges que cela suppose). La solution consiste alors à appliquer les tables démographiques aux "populations soumises aux risques" démographiques (principalement fécondité et mortalité). Mais l'usage de telles techniques étant répandu on a aucun intérêt dans le cadre de cet exercice à les utiliser.

En conséquence, l'accroissement naturel de la population stable est considéré comme donné et la population potentielle se définit comme suit :

$$\begin{aligned} (\text{Population Potentielle 1968}) &= (\text{Population observée 1968}) - (\text{solde migratoire intérieur observé}) \\ &\quad - (\text{immigrants extérieurs observés}) + \text{immigrants intérieurs calculés} \\ &\quad - \text{émigrants intérieurs calculés} + \text{immigrants extérieurs calculés} \end{aligned}$$

Cette technique intègre implicitement le volume d'émigrants extérieurs et des différences de décompte de population imputable aux différences de saisie de données et dont on a déjà eu l'occasion de parler.⁽¹⁾ Par contre son avantage est de bien cerner les erreurs qui ne sont imputables qu'aux seuls 3 sous-modèles utilisés.

(1) au § 1.12 du présent chapitre (voir page 233)

SECTION 3 - DETERMINATION DE LA DEMANDE D'EMPLOI POTENTIELLE

Maintenant que la population potentielle est déterminée il ne reste plus qu'à trouver le taux d'activité potentiel à lui appliquer pour trouver la demande potentielle.

Avant de définir le sous-modèle potentiel et d'examiner les résultats, il est nécessaire d'examiner les causes récentes de fluctuation du taux d'activité en France, complétant ainsi les aspects théoriques développés dans le chapitre 5. (Cette analyse de type descriptif n'avait pas sa place dans ce chapitre d'orientation théorique).

3.1 - Causes de fluctuation du taux d'activité

Ces diverses causes de fluctuation ont été analysées de façon descriptives pour la Commission de l'emploi du 6^e Plan et l'on rapportera ici quelques unes des observations les plus intéressantes, Celles-ci sont orientées sur la prévision de l'emploi pour la période 1968 - 1976 mais elles illustrent bien les causes de variation ; aussi les a-t-on rapportées telles quelles. Mais au préalable, il convient de faire observer que les causes dont il est question dans cette section ne sont pas des causes premières mais plutôt leurs effets et que l'on se place ici dans une optique d'analyse d'écart, et pas, sauf exception⁽¹⁾ de leurs causes profondes (et en particulier de leurs relations avec le marché de l'emploi).

On trouvera page suivante le tableau synoptique de ces écarts, qui constitue la base de notre analyse. Cette analyse dépend étroitement des statistiques dont on dispose, c'est pourquoi on en parlera brièvement avant de procéder à l'analyse des écarts.

(1) Celles d'origine institutionnelle, encore que l'on puisse s'interroger sur leurs motivations profondes.

DECOMPOSITION des VARIATIONS ANNUELLES de la POPULATION ACTIVE

249

 DEPUIS 1954 et PREVISIONS pour 1975⁽¹⁾

unité : 1000

		1 - 1 - 1954 1 - 1 - 1962	1 - 1 - 1962 1 - 1 - 196	1 - 1 - 1968 1 - 1 - 1976
Incidence de l'évolution naturelle de la population	14 - 24 ans	- 33,0 (1)	+ 173,3 (2)	+ 41,7 (3)
	25 - 54 ans	- 16,4 (1)	- 71,0 (2)	+190,4 (3)
	55 et +	+ 69,1 (1)	+ 29,4 (2)	- 61,3 (3)
Incidence totale de l'évolution naturelle.....		+ 19,7 (1)	+ 131,7 (2)	+ 170,8 (3)
Incidence de la prolongation de la scolarité				
	14 - 24 ans	- 45,5	- 76,9	- 88,0 (4)
Incidence de la variation des taux d'activité des femmes	14 - 24 ans	+ 20,0	+ 13,1	+ 11,8 (4)
	25 - 54 ans	- 3,1	+ 25,3	+ 90,7 (4)
Incidence de la variation des taux d'activité des hommes	14 - 24 ans	+ 6,3	- 0,5	-
	25 - 54 ans	- 2,6	- 5,1	- 4,7 (4)
Incidence de la variation des taux d'activité aux âges élevés	55 ans et +	- 32,2	- 83,9	- 36,9 (4)
Incidence totale des variations de taux d'activité		- 57,1	- 128,0	- 27,1 (4)
Solde migratoire d'actifs.....		+ 65,8	+ 135,8	+ 83,1
dont rapatriés.....		+ 19,4	+ 62,5	
Variation totale de la population active.....		+ 28,4	+ 139,5	+ 226,8
Variation de l'effectif du contingent et des militaires hors métropole		- 37,5	+ 63,3	+ 1,3
Variation des disponibilités en main d'oeuvre		- 9,1	+ 202,8	+ 228,1
(1) Taux d'activité au 1. 1. 1954 (2) Taux d'activité au 1. 1. 1962 (3) Taux d'activité au 1. 1. 1968 (4) Le calcul de l'incidence de la variation du taux d'activité a été fait dans l'hypothèse d'une conjoncture de l'emploi moyenne au 1 janvier 1976 (300.000 personnes disponibles à la recherche d'un emploi)				

3.11 - Remarque préalable sur les données de base :

L'étude de la variation du taux d'activité se heurte d'abord et surtout à l'imprécision du matériel statistique dont les chercheurs disposent. En effet, la projection de la population active se fait par application du taux d'activité constaté lors des recensements. Or ces derniers ont coïncidé avec des périodes conjoncturelles contrastées, caractérisées soit par une détente marquée sur le marché du travail (1954 et 1968), soit par une tension exceptionnelle (Mars 1962). D'où la nécessité pour les statisticiens de recourir à des calculs de corrections pour atténuer ce phénomène de "flexion".

La Commission Emploi note à ce propos : "Si l'existence de ce phénomène de flexion des taux d'activité en fonction de la conjoncture, que l'on observe dans tous les pays pour lesquels il existe des statistiques appropriées, ne peut être mis en doute, l'estimation de son incidence repose sur des méthodes économétriques qui, par la nature des hypothèses qu'elles supposent et la qualité des ajustements obtenus, laissent une large part d'incertitudes, notamment en ce qui concerne la part entre les retraits spontanés de l'activité et le chômage".

3.12 - Analyse des écarts

Les principales causes d'évolution de la population active sont surtout liées aux catégories pour lesquelles les attitudes à l'égard du travail peuvent se modifier (jeunes, femmes, personnes âgées).

En effet ces groupes représentent près de la moitié de la population active française. Chacun pris séparément peut modifier le taux d'activité par une cause qui lui est propre :

α - Les jeunes avec la prolongation de la scolarité

La Commission Emploi du 6^e plan fait notamment remarquer en se basant sur le passé que "La prolongation de l'âge de fin de scolarité obligatoire est entrée en vigueur à la rentrée d'octobre 67... Il a été admis que cette mesure atteindrait son plein effet dès le début du plan pour les enfants de 14 ans et dès la rentrée de 72 pour ceux de 15 ans. En outre, on a jugé que la prolongation à 16 ans de l'âge de scolarité obligatoire constituerait une incitation pour un certain nombre de jeunes, à poursuivre des études au-delà de cet âge".

Cet état de fait entraîne une augmentation du taux de scolarité bien que celui-ci reste soumis à l'influence de la conjoncture : à conjoncture médiocre correspond une incitation à la poursuite des études et à l'inverse, une forte expansion avec une politique de hauts salaires incite à l'entrée immédiate dans la vie active. Ce facteur sera à l'origine d'une diminution annuelle de 105 000 actifs sur la période 68 - 76.

β - Les personnes âgées avec l'avancement de l'âge de la retraite

La réduction de l'activité aux âges élevés s'explique par la diminution relative de la population agricole et de l'ensemble des non salariés chez lesquels les taux d'activité sont en moyenne plus élevés. L'influence de la conjoncture peut être assez importante. On estime par exemple que la croissance (nouvelle) de la population active réduira cette tendance à un chiffre annuel de 37 000 pour la période 68 - 76 (alors que le chiffre de la période précédente était de 84 000), les possibilités de conversion des travailleurs indépendants âgés se réduisant. On notera également le rôle d'incitateur à la baisse du taux d'activité de certaines actions politiques comme l'accord national interprofessionnel du 10/2/1969 sur la sécurité de l'emploi relatif à l'assouplissement de l'âge de la retraite ou encore la politique des indemnités viagères de départ menée par le Ministère de l'Agriculture.

γ - Les femmes avec leur importance accrue dans la vie active

Comme le note la commission de l'emploi : "La prolongation de cette tendance à moyen terme, dégagée après correction des taux observés de l'influence de la conjoncture, conduit, par rapport au niveau d'activité des femmes en 1968, à l'accroissement des disponibilités en main d'oeuvre en 1975 d'environ 800.000 personnes. L'augmentation de la participation des femmes à l'activité économique apparaît ainsi comme un des principaux facteurs d'accroissement de la population active au cours du 6^e plan". Ici particulièrement la modification des comportements évolue sous l'effet de facteurs structurels socio-économiques complexes (changement des attitudes à l'égard du travail, développement de la formation professionnelle et surtout diminution de la durée du travail). D'après la Commission de l'Emploi, l'augmentation du nombre des femmes actives passerait de 100.000 par an pour la période 68 - 76 contre 40.000 par an pour 62 - 68.

3.2 - Sous-modèle taux d'activité Potentiel

Il ressort de l'analyse précédente que les facteurs explicatifs de l'évolution des taux d'emploi sont difficiles à intégrer dans un modèle explicatif départemental du type régression linéaire. C'est plutôt vers des techniques de coefficients de spécificité⁽¹⁾ utilisées pour diverses tranches d'âge, qu'il faut se tourner. Ces techniques impliquent, il faut le rappeler, une parfaite connaissance de la situation au niveau national. Dans le cadre de l'exercice proposé, on supposera qu'un modèle national nous aura fourni les données souhaitées et intégré les causes majeures de variation du taux d'activité et sert de référence à l'évolution départementale. On prendra en pratique les données observées au niveau France entière en 1968. Dès lors il ne reste plus qu'à poser des hypothèses d'évolutions départementales comparées à celle de l'ensemble du pays avant de les traduire concrètement par des formules appropriées, de souligner certains biais inévitables et de choisir les classes d'âge. Examinons donc notre analyse du phénomène.

3.21 - Choix d'hypothèses d'évolution

Sur la première période on peut classer les évolutions départementales d'une catégorie âge-sexe donnée en deux groupes : celles qui convergent vers l'évolution France entière et celles qui divergent.

Or dans le long terme, en l'absence d'informations précises contradictoires, le taux d'activité le plus probable pour un département est celui prévu pour l'ensemble du territoire⁽²⁾ Il semble donc logique d'essayer à moyen terme de resserrer l'éventail des valeurs prises par notre variable. En conséquence, lorsqu'il y a convergence, on maintiendra celle-ci à moins que le maintien de cette évolution comparée ne mène à une divergence, et lorsqu'il y aura divergence, on limitera l'évolution départementale à celle observée pour l'ensemble du pays.

3.22 - Mise en application du principe retenu

On s'inspirera de l'analyse faite dans l'étude de l'évolution du tertiaire⁽³⁾ pour définir les phénomènes de rattrapage, d'accélération, de retard accentué et d'uniformisation.

(1) Voir chapitre 7 section 1

(2) et si l'on pouvait considérer que la distribution des taux est aléatoire on pourrait même parler de meilleur estimateur (encore que le taux moyen diffère de la moyenne des taux)

(3) Voir chapitre 12 page 322. Ce chapitre a été écrit avant le chapitre 9.

Définissons les variables suivantes :

- TAC (I, J, C) = taux d'activité de I à la date J pour une catégorie de population C donnée ;
- K (D, J, C) = TAC (D, J, C) / TAC (N, J, C)
- L (D, C) = K (D, 62, C) / K (D, 54, C)

avec I = D ou N : D pour un département D donné et N pour l'ensemble de la nation
et J = 54 ou 62 ou 68 (années d'observation)

On peut dresser alors sur la période 1954 - 1962 la typologie suivante :

	L (D, C) < 1	L (D, C) > 1
K (D, 54, C) < 1	retard accentué	rattrapage
K (D, 54, C) > 1	uniformisation	accélération

La mise en oeuvre du principe énoncé au paragraphe précédent diffère selon que l'évolution du taux d'activité de la catégorie donnée France entière se poursuit d'une période sur l'autre ou s'il y a renversement de tendance.

α - Si la tendance se poursuit, il faut ne conserver que les situations de rattrapage et d'uniformisation à condition que l'utilisation des coefficients de spécificité des évolutions n'amène pas un renversement de tendance c'est à dire que K (D, 54, C), K (D, 62, C) et K (D, 68, C) doivent simultanément être inférieurs ou supérieurs à 1. S'il n'en était pas ainsi, on serait amené à prendre comme valeur pour TAC (D, 68, C) celle donnée pour l'ensemble de la nation, TAC (N, 68, C). Dans le cas d'évolution divergente sur la première période (accélération et retard accentué), il faudrait restreindre l'évolution du département à celle observée France entière.

β - S'il y a renversement de tendance de l'évolution France entière, le maintien des coefficients de spécificité des coefficients d'accroissement mène dans la plupart des cas à l'inverse de ce que l'on recherche. Aussi subsistera-t-on dans ce cas une vision statique à une vision dynamique, pour atteindre l'objectif de resserrement des disparités que l'on s'est fixé. Si l'évolution France entière sur la période 62 - 68 est décroissante on postulera la stabilité du coefficient de spécificité observé en 1962 ($K(D,62,C) = K(D,68,C)$) ce qui revient à appliquer le même coefficient de variation (ici inférieur à 1) à tous les départements et donc à l'écart entre la valeur départementale et celle France Entière. Lorsque l'évolution sera croissante on postulera la constante de l'écart observé en 1962 entre les valeurs nationales et départementales ($TAC(D,62,C) - TAC(N,62,C) = TAC(D,68,C) - TAC(N,68,C)$), car l'application de la règle précédente mène à accroître ces écarts.

3.23 - Biais introduits

α - La méthode utilisée

Les hypothèses introduites sont totalement arbitraires et déconnectées des mécanismes réels et en particulier de l'évolution des particularismes locaux de comportement ou de structures économiques. Ce que l'on suppose implicitement c'est une grande inertie du système tant du point de vue des facteurs d'emploi que des mécanismes. Si le modèle d'ensemble est mené à son terme de façon satisfaisante ceci n'est pas trop grave car les mécanismes d'ajustements mis en valeur sont en partie destinés à modifier l'impact de cette inertie. En tout état de cause la démarche est d'autant plus contestable que l'horizon s'éloigne.

β - Le problème des différences de conjonctures dans le temps et l'espace

Les tensions de marché de l'emploi signalées à un niveau global au § 3.11 non seulement diffèrent aux deux dates de référence mais de plus dans l'espace. Les données des deux recensements, ne sont donc pas directement comparables. Les études préparatoires du 6^e plan tenaient compte de ce fait en s'intéressant à la population active disponible, c'est à dire titulaire d'un emploi ou à la recherche de celui-ci et c'est bien entendu à partir de ces données que nos travaux ont été menés. Il n'en reste pas moins que la déclaration de recherche d'emploi telle qu'elle résulte du recensement est fonction du marché de l'emploi ; il est en effet certain qu'un certain nombre d'individus dans des zones en crise seraient éventuellement disposés à travailler (ce qui correspond bien à la notion de potentiel) mais, certains de ne rien pouvoir trouver, ne se déclarent pas candidats, le partage entre les retraits spontanés et les

chômeurs est donc largement entaché d'incertitudes. Par ailleurs, l'immigration étrangère dont le taux d'activité est plus important se dirige vers des zones de tension d'emploi. Ces phénomènes sont sans doute de faible importance si l'on raisonne au niveau global mais les biais peuvent être sensibles lorsqu'on descend à un niveau géographique assez fin. Il est hélas impossible d'aller plus loin avec les statistiques dont on dispose.

γ - Les taux d'activité utilisés

Il n'est malheureusement pas possible ici de procéder à une analyse par génération. C'est donc l'analyse transversale qui a été utilisée. Certaines des causes de variation des taux sont donc d'origine démographique, d'autant plus qu'il n'a pas été possible de descendre en dessous de la classe quinquennale.

3.24 - stratification de la population retenue

Les considérations développées au § 3.1 et 3.23 obligent à définir avec précaution la stratification à retenir. Un premier critère s'impose, le sexe. Le second qui est l'âge nécessite de définir un certain nombre de classes.

Le découpage a été choisi en fonction de l'homogénéité de comportement des classes au regard de l'évolution du taux d'activité France entière entre 1962 et 1968, et au regard de l'évolution de la structure par âge pour la France entière sur la même période, ce qui aboutit, compte tenu des contraintes informatiques⁽¹⁾ à la stratification suivante :

SEXE MASCULIN		SEXE FEMININ	
15 - 19 ans	55 - 59 ans	15 - 19 ans	40 - 44 ans
20 - 24 ans	60 - 64 ans	20 - 24 ans	45 - 54 ans
25 - 29 ans	65 - 69 ans	25 - 29 ans	55 - 59 ans
30 - 49 ans	70 - 74 ans	30 - 34 ans	60 - 64 ans
50 - 54 ans	+ de 75 ans	35 - 39 ans	+ de 65 ans

(1) qui limitent le nombre de classes

Pour obtenir un taux d'activité global on est contraint d'effectuer une somme pondérée de ces divers taux par la part que chaque strate occupe dans la population totale du département. Les pondérations retenues sont celles observées en 1962, ce qui est assez logique tant qu'un modèle démographique prévisionnel par âge n'est pas envisagé. Ce système de pondération est malheureusement sensible aux "accidents démographiques" de la pyramide des âges, ainsi que des échanges migratoires avec l'étranger, mais il est difficile de faire autrement dans une première étape.

Il ne reste plus qu'à examiner les sorties du sous-modèle "taux d'activité" ainsi défini.

3.3 - Analyse des résultats obtenus

On trouvera pages A9 et T9. 49 à 50, les tableaux et cartes servant de support au présent commentaire.

Les résultats sont dans l'ensemble satisfaisants dans la mesure où l'erreur relative, prise en valeur absolue, n'est supérieure à 3 % que dans 16 départements et que les écarts arithmétiques ne dépassent pas 2 points (unité 100). Ces écarts sont donc tous compatibles avec l'hypothèse d'ajustement possible par le chômage. En outre, tous les départements où la sousestimation par le modèle dépasse 1 point sont groupés autour de la Région Parisienne, tandis que les quelques cas de surestimation de plus d'un point sont assez regroupés (quatre zones : la frontière belge, la Bretagne Atlantique, la Rive gauche du Rhône, le sud des régions Midi-Pyrénées et Aquitaine).

Toutefois les diverses limites que l'on a donné incitent à beaucoup de prudence dans l'interprétation de ces résultats.

Le sous-modèle B.T.P.

CHAPITRE X - Le SOUS MODELE BTP

"Quand le bâtiment va, tout va" disait-on couramment jusqu'à une époque très récente. Cette coïncidence constatée qui fait du bâtiment le baromètre de la santé, de l'économie d'un pays, traduit imparfaitement la réalité économique. En effet, la production de cette branche (bâtiment, routes, travaux de génie civil), contrairement à celle de la plupart des autres activités économiques a une longue durée d'utilisation ; sa demande est donc fonction de l'entretien ou du renouvellement d'un stock existant, d'ou accroissement de ce stock lié à l'accroissement de la population totale ou de ses besoins privés et collectifs ainsi qu'aux impératifs de croissance économique.

Ces trois types de demande peuvent s'exprimer avec plus ou moins de retard par rapport aux besoins objectivement ressentis, en fonction des capacités de financement des agents économiques concernés et être satisfaits plus ou moins rapidement en fonction des disponibilités de capacité productive. Mais à côté de ces demandes induites peut exister une demande autonome (portant principalement sur des équipements collectifs), liée à une politique de grands travaux inspirée ou commandée par l'état, que ce soit dans une optique keynésienne ou non. Cette hétérogénéité de la demande (et des élasticités par rapport aux besoins) fait donc qu'il est a priori difficile de faire du bâtiment l'indicateur privilégié de la santé économique du pays.

Toutefois, cette branche contribue à près de 60 % de la formation brute de capital fixe en France et il n'en reste pas moins qu'à moyen terme une production de BTP, en volume suffisant, est une condition permissive de la croissance. Aussi la branche BTP a-t-elle toujours fait l'objet d'une attention particulière de la part des pouvoirs publics.

Nous avons isolé cette activité des autres car elle ne peut être considérée comme totalement induite ou totalement autonome. Examinons successivement la description du phénomène, puis la recherche d'un modèle explicatif.

SECTION I - ETUDE DESCRIPTIVE du BTP

Le BTP a été étudié jusqu'à présent en France presque exclusivement dans une optique globale et productive⁽¹⁾ alors que seuls les aspects d'emploi spatialisé nous intéressent.

On ne retiendra ici que quelques études spatialisées, sans insister sur l'aspect de la prévision de la demande, qui ne relève pas, dans une première étape du moins, de la problématique retenue⁽²⁾. Il est toutefois évident que la connaissance de certaines données relatives à la production est indispensable à la compréhension des structures de l'emploi dans le BTP.

Ce n'est que récemment que la dimension spatiale du phénomène BTP a été prise en compte et encore au niveau de la région de programme⁽³⁾. Malheureusement, il semble hors de question de pouvoir reconstituer à un niveau départemental les informations fines dont on dispose au niveau régional depuis 1961 (structure de l'appareil productif, structure de la demande en valeur, ...). Aussi semble-t-il intéressant d'examiner rapidement les principales conclusions des travaux effectués au niveau régional sur la période 1961 - 1969 avant d'étudier l'évolution de cet emploi sur les trois derniers recensements.

- (1) voir en particulier "Principe pour une méthode sommaire de calcul de la demande solvable en logement" Ministère de l'Équipement et du Logement - STCAU (février 1968) et "Contribution à l'élaboration d'un modèle de développement spatial : un modèle de localisation des logements neufs" CERAU 1968.
- (2) l'étape ultérieure si l'approche directe de l'emploi semble être satisfaisante consisterait à trouver des modèles de prévision satisfaisants de certaines données (telles que la production BTP), considérées comme variables exogènes dans la première étape.
- (3) voir en particulier les études effectuées par le Bureau d'Information et de Prévisions économiques (BIPE) et tout particulièrement le rapport de synthèse intitulé "Évolutions régionales de la production et de l'activité des entreprises de bâtiment et génie civil et des industries de matériaux de construction - 1, éléments rétrospectifs de 1961 à 1969" juillet 1969.

1.1 - Les travaux effectués au niveau régional

Les travaux de reconstitution du BIPE ont permis de connaître le montant en valeur de la production de chaque région et donc de connaître la disparité spatiale de cette production. L'explication de ces disparités a été recherchée dans quatre directions : différences de population , répartition spatiale du marché, taille des entreprises, différence de productivité.

1.11 - L'influence de la population

Pour comparer les productions entre les régions, il faut tout d'abord tenir compte de la population desservie , dont dépend en grande partie la demande. On peut noter, outre l'influence déjà citée des stocks BTP existants, celle du découpage spatial retenu (plus ou moins grande économie d'échelle d'infrastructures routières, ...) sur laquelle on reviendra, ainsi que celle du relief (qui se traduit par une plus ou moins grande densité d'ouvrages d'art au km²). L'indicateur que l'on retiendra sera celui de la production globale par tête (voir tableaux et cartes page suivante).

La comparaison de deux dates éloignées (1961 et 1968) montre une assez grande évolution de cet indicateur, ce qui en souligne bien, d'ailleurs, les limites explicatives. En 1968 le Sud-Est (Rhône - Alpes, Languedoc, Provence Côte d'Azur) et le Centre au sens large (Région Parisienne, Centre, Bourgogne) sont les régions qui connaissent la plus forte production par tête.

Par ailleurs, on peut remarquer que les régions qui ont le plus faible pourcentage d'entreprises de moins de 5 salariés sont celles qui en plus ont la plus forte production au km². Il s'agit de la Région Parisienne, Haute Normandie, Nord, Lorraine, Alsace, Rhône - Alpes et Provence - Côte d'Azur. Ces régions sont de plus celles qui ont le plus fort taux d'urbanisation.

DENSITE REGIONALE de la PRODUCTION
en FRANCS CONSTANTS 1961 par HABITANT

Source BIPE ⁽¹⁾

	1961	1968
Région Parisienne	891	1 497
Champagne	910	1 313
Picardie	888	1 203
Haute-Normandie	825	1 303
Centre	842	1 360
Basse-Normandie	786	1 210
Bourgogne	754	1 467
Nord	793	1 139
Lorraine	1 036	1 267
Alsace	974	1 305
Franche-Comté	730	1 177
Pays de la Loire	764	1 243
Bretagne	730	1 211
Poitou - Charentes	682	1 122
Aquitaine	920	1 352
Midi - Pyrénées	915	1 340
Limousin	708	1 275
Rhône - Alpes	941	1 432
Auvergne	814	1 214
Languedoc	958	1 587
Provence - Côte d'Azur	1 157	1 592
France entière	866	1 337

(1) Rapport cité

1.12 - Répartition spatiale du marché

Une enquête effectuée en 1966 par le Commissariat aux entreprises de Travaux Publics et de Bâtiment a permis d'établir un tableau carré dont est extrait le tableau de la page suivante donnant la ventilation de la production en valeur par localisation régionale des chantiers (en ligne) et siège régional de la société responsable du chantier (en colonne). Il ressort de ces informations que toutes les régions à l'exception de la Picardie et de la Haute-Normandie font appel pour plus de 80 % du volume de travaux BTP effectués chez elles à des entreprises de la même région (en moyenne 84 %) et pour 10 % en moyenne à des entreprises de la Région Parisienne (qui exporte 32 % de sa production).

1.13 - Taille des entreprises et production

Quelles que soient les régions et l'année d'observation, plus de 90 % des entreprises comprennent moins de 21 salariés : toutefois l'importance des petites entreprises est relativement faible dans la production globale, comme le montre le tableau suivant établi pour l'année 1962⁽¹⁾.

Classe d'effectifs	% entreprises	% production
0 - 10	91,2	23,5
11 - 100	7,9	34,8
> 100	0,9	41,7

1.14 - Productivité dans le BTP

Dans la branche BTP⁽²⁾, la productivité apparente du travail est sans doute fonction de la taille des chantiers pour des raisons d'économie d'échelle, mais aussi de la conjoncture car les licenciements interviennent après les retournements de conjoncture et dépendent de l'élasticité d'anticipation des entrepreneurs (pour cette dernière hypothèse, le coefficient de corrélation entre gains de productivité annuels et la variation relative de la production est significative pour plus de la moitié des régions). Il faut noter aussi que la répartition de la demande entre travaux de génie civil et travaux de bâtiment proprement dits influe sensiblement sur la productivité globale de la branche (la mécanisation possible dans le génie civil permettant des productivités moyennes et marginales du travail nettement plus fortes).

(1) Etudes et Conjonctures décembre 1968

(2) BIPE, étude citée

VENTILATION de la PRODUCTION REGIONALE en 1966

Unité : 100

	Origine de l'entreprise de chantiers			Localisation des chantiers de siège de l'entreprise		
	Même région	Région parisienne	Autre région	Même région	Région parisienne	Autre région
Nord	88,2	10,3	1,5	92,7	1,3	6,0
Picardie	70,9	20,8	8,3	88,6	4,1	7,3
Région Parisienne	93,8	-	6,2	68,1	-	31,9
Centre	80,9	13,6	5,5	89,8	4,5	5,7
Haute-Normandie	70,4	20,4	9,2	80,3	9,2	10,5
Basse-Normandie	85,0	9,9	5,1	92,2	3,5	4,3
Bretagne	83,5	12,5	4,0	97,2	0,9	1,9
Pays de la Loire	84,3	11,8	3,9	86,3	2,6	11,1
Poitou - Charentes	85,6	7,4	7,0	90,6	2,8	6,6
Limousin	86,8	6,6	6,6	88,2	2,7	9,1
Aquitaine	83,9	9,3	6,8	88,4	3,8	7,8
Midi - Pyrénées	89,2	5,7	5,1	86,2	3,2	10,6
Champagne	81,0	11,3	7,7	88,6	5,4	6,0
Lorraine	88,8	7,9	3,3	88,6	2,7	8,7
Alsace	86,9	6,3	6,8	98,2	0,1	1,7
Franche-Comté	86,8	5,5	7,7	96,9	0,5	2,6
Bourgogne	78,8	14,4	6,8	87,6	1,6	10,8
Auvergne	86,7	4,1	9,2	93,7	0,4	5,9
Rhône - Alpes	87,2	9,2	3,6	91,9	0,9	7,2
Languedoc	79,8	9,7	10,5	93,2	0,2	6,6
Provence - Côte-d'Azur	87,1	8,3	4,6	96,0	0,3	3,7
France entière	83,8	10,3	5,9	83,8	1,5	14,7

Source : Enquête du Commissariat aux Entreprises des Travaux Publics et de Bâtiments

Il est intéressant enfin de souligner l'influence du recrutement de main d'oeuvre dans le BTP sur sa productivité : pendant très longtemps une partie non négligeable des mutations agricoles a transité par le BTP (qui permettait un cumul d'emploi) avant le passage définitif dans le secteur secondaire ; par ailleurs, l'appel de plus en plus important à une main d'oeuvre étrangère non qualifiée a sans doute techniquement ralenti les progrès de productivité du travail (la substitution capital - travail qui s'est effectuée dans ce secteur est sans doute en partie un effet de cette modification de recrutement).

1.2 - Les observations à un niveau départemental

Aucune étude globale n'ayant été effectuée au niveau départemental, il a semblé intéressant d'étudier l'évolution des disparités spatiales de ce secteur, ou ce qui revient au même, de tester le modèle "simpliste" basé sur l'hypothèse du maintien des disparités spatiales observées en t à la date t' , étant entendu que les prévisions France entière pour t' sont supposées correctes (la donnée prévisionnelle utilisée est en pratique la valeur observée pour la France entière en t').

1.21 - Evolution de l'emploi BTP par département entre 1954 et 1962

a) Evolution en termes d'effectifs

Soit pour chaque département, X_1 les effectifs employés dans le BTP en 1962 et X_2 les effectifs employés dans le BTP en 1954. Sa valeur calculée pour 1962 pour un département s'obtient en écrivant :

$$X_1 = a.X_2$$

avec comme valeur pour le scalaire a , le rapport $62/54$ des valeurs observées France entière (la valeur observée 62 étant prise comme donnée prévisionnelle, c'est-à-dire que l'on suppose non entachée d'erreur la prévision faite au niveau France entière).

On trouvera dans les pages T 10.26 à 31 les écarts entre valeurs observées et calculées et l'erreur relative commise par l'adoption de ce mode de calcul (ou si l'on préfère la distorsion entre les deux dates de la répartition spatiale) ainsi que leur représentation graphique.

Il faut tout d'abord noter l'importance des distorsions (= erreur relative) qui s'étendent de - 48 % à + 29 % avec un écart-type de 18 % ; les écarts entre valeurs calculées et observées s'étendent eux de - 11 000 à + 18 000, avec un écart-type de 4 900.

Cela étant, la cartographie de ces écarts révèle une surestimation systématique au nord d'une ligne Grandville - Belfort (à l'exception de la Région Parisienne) ainsi qu'une sous estimation systématique au sud d'une ligne Bordeaux - Genève (à l'exception des Hautes-Pyrénées d'une part, et de l'Ardèche, du Rhône et de la Savoie de l'autre).

L'étude en terme de taux (= part des actifs du BTP dans la population totale) va permettre de corriger ces premières impressions en tenant compte de l'influence des variations de population entre ces deux dates.

b) Evolution en terme de taux (voir cartes page A 10)

L'utilisation du même modèle en prenant pour X1 et X2 la part des actifs du BTP dans la population active a permis l'établissement du tableau et des cartes des pages 26, 27, 28. Les erreurs relatives sont moins importantes puisque les extremums observés d'erreurs relatives sont - 40 % et + 25 % mais l'écart-type reste le même (18 %) toutefois la dichotomie entre la France du Nord et la France du Sud s'accroît puisque les effectifs BTP de tous les départements situés au nord de la ligne Grandville - Belfort sont surestimés (sauf pour la Région Parisienne dont la sous-estimation précédemment notée s'explique par un fort accroissement de population totale).

Au sud de cette ligne à l'exception de 14 départements où la prévision est satisfaisante (erreur relative inférieure à 5 %), tous les départements sauf 5 sont sous-estimés. Ces exceptions sont : la Savoie et les Hautes-Pyrénées déjà notées, ainsi que les trois départements bretons de la façade atlantique surestimés déjà dans le modèle en effectif (ce fait n'avait pas été souligné car la bande intermédiaire présentait une grande hétérogénéité).

On peut sans doute trouver une explication de ce phénomène dans le fait que le nord de la France, de loin le plus éprouvé par la guerre, n'avait pas terminé sa reconstruction en 1954, et l'emploi dans le BTP y était de ce fait à son maximum. L'effort de reconstruction terminé, la pression du Nord a décru et le sud, dont la satisfaction des besoins avait été freinée, a bénéficié de transfert d'actifs du BTP en provenance du nord puisque

la Région Parisienne a absorbé à elle seule sur cette période davantage d'emplois de BTP que ne le permettait l'accroissement global d'emploi BTP sur l'ensemble du territoire (cette remarque ne présume pas d'ailleurs du type de filière adoptée pour effectuer ces transferts).

Il faut par ailleurs rappeler la très grande différence de situation conjoncturelle aux deux dates de recensement, mais on peut penser que les distorsions engendrées sont notablement plus faibles que celles observées, car la forte tension existant en 1954 sur le marché du BTP n'a sans doute pas été très affaiblie par la basse conjoncture de mai 1954.

1.22 - Evolution de l'emploi BTP par département entre 1962 et 1968

On trouvera en annexe⁽¹⁾ les cartes et tableaux de chiffres relatifs à l'évolution des effectifs employés dans le BTP ainsi que du taux correspondant, entre 1962 et 1968. Deux traits sont marquants :

On observe peu de concordance entre les cartes de distorsion des deux périodes, c'est-à-dire qu'en dynamique l'évolution même n'est pas stable. Il faut noter d'ailleurs que l'erreur relative sur le taux s'atténue puisque si l'on retire la Région Parisienne dont les délimitations ont changé, celle-ci s'étend de - 22 % à + 15 % avec un écart-type de 13 % (contre 18 % pour la période précédente).

Il n'en reste pas moins que peu de départements changent radicalement de tendance. Si l'on écarte la frange comprise entre + 5 % les seuls bouleversements (= passage d'écart relatif supérieur à 5 % inférieur à moins 5 % et réciproquement) affectent la façade atlantique de la Bretagne (Finistère, Morbihan) et la Marne qui de sous estimés deviennent surestimés et le sud-est pour lequel le bouleversement est inverse: l'Isère, l'Aveyron, la Drôme, les Bouches-du-Rhône, les Alpes de haute Provence et les Alpes Maritimes.

On peut donc conclure que aussi bien en statique comparative qu'en dynamique les données structurelles de l'emploi BTP varient dans des proportions trop importantes pour que l'on puisse raisonnablement utiliser l'hypothèse de stabilité de coefficients structurels au cours du temps. Cette constatation renforce l'intérêt de recherche d'un modèle explicatif.

(1) Voir pages T 10.29 à 31, ainsi que la carte de la page A 10 qui concerne l'évolution 62-68

1.23 - Evolution de l'emploi BTP et régression linéaire simple

L'absence de distorsion dans l'évolution départementale des données (homothétie) se traduit dans le cadre de régression linéaire simple par le fait que le produit des coefficients des deux régressions est égal à 1. Le fait de s'écarter de cette liaison linéaire se traduit par un rapprochement vers zéro du coefficient de corrélation⁽¹⁾. On peut donc compléter les observations faites dans les paragraphes précédents par l'appel aux techniques de régression.

Les résultats obtenus pour la période 54 - 62 sont les suivants :

$$(\text{effectifs BTP } 62) = 1,30367 \cdot (\text{effectifs BTP } 54) - 1064 \quad (\text{coefficient de corrélation} = 0,979315)$$

(45,6)

$$\left(\frac{\text{BTP } 62}{\text{pop. Act. } 62}\right) = 0,739366 \cdot \left(\frac{\text{BTP } 54}{\text{pop. Act. } 54}\right) + 0,0371 \quad (\text{coefficient de corrélation} = 0,653013)$$

(8,13)

et entre 1962 et 1968 en excluant les départements de la région parisienne en raison des modifications de découpage :

$$(\text{effectifs BTP } 68) = 1,12219 \cdot (\text{effectifs BTP } 62) + 1771 \quad (\text{coefficient de corrélation} = 0,98513)$$

(54,00)

$$\left(\frac{\text{BTP } 68}{\text{pop. Act. } 68}\right) = 0,907465 \cdot \left(\frac{\text{BTP } 62}{\text{pop. Act. } 62}\right) + 0,0267 \quad (\text{coefficient de corrélation} = 0,902541)$$

(19,55)

Ces résultats sont meilleurs pour les régressions en effectifs que pour celles en taux ce qui résulte de divergences spatiales de l'évolution de la population active, par ailleurs globalement stable sur la période, et de celle de l'emploi BTP, en croissance depuis 1954. En outre, les valeurs des paramètres obtenus confirment la relative permanence du phénomène noté plus haut.

(1) dont la valeur absolue est égale à la racine carrée du produit des coefficients des deux régressions linéaires possibles (et donc égale 1 en cas de fonction linéaire), cas extrême de la liaison stochastique, c'est-à-dire que l'on se place dans le cas de fonction réciproque).

SECTION 2 - RECHERCHE D'UN MODELE EXPLICATIF DEPARTEMENTAL POUR 1962

On s'efforcera à l'aide d'indicateurs, qui sont supposés traduire correctement un ensemble d'hypothèses pertinentes avec les observations faites, de rechercher des modèles explicatifs satisfaisants pour 1962 en effectifs, et en taux, afin de choisir un modèle définitif à utiliser en mode prévisionnel pour 1968.

Il faut noter que tous les modèles pour lesquels la variable explicative était non pas l'état à une date donnée mais la variation de l'état entre deux dates se sont avérés nettement moins bons que les premiers (1). Pour cette raison les résultats n'ont pas été mentionnés.

2.1 - Les hypothèses externes pour test et leurs indicateurs

2.11 - Ajustement de l'emploi à la demande de BTP

On peut tout d'abord supposer que dans ce secteur où les garanties d'emploi sont faibles, le volume de population active s'ajuste à la demande de travaux à effectuer.

- Il semble tout d'abord naturel de dissocier l'activité de construction de logements de celle de bâtiments à usage industriel et travaux publics non seulement parce que les demandeurs sont différents mais aussi parce que les techniques utilisées le sont également. Et en ce qui concerne le logement on a dissocié le secteur commandé par l'état (H.L.M) du secteur qui dépend des particuliers.

...

(1) Cette observation diffère de celle effectuée par le CERAU sur un découpage spatial infra départemental (23 bassins d'emploi) dans la Région Rhone Alpes ce qui est tout à fait possible car le degré de signification des variables est étroitement dépendant en fait du découpage géographique retenu.

Les variables introduites sont donc :

- nombre total de logements construits (l'utilisation de cette variable est exclusive des deux suivantes)
- nombre de logements terminés, du secteur HLM
- nombre de logements terminés autres que ceux du secteur HLM
- montant des travaux de construction industrielle et des travaux publics.

Les actifs du BTP recensés en mars 1962 ont participé à la construction de logements terminés dans le courant des années 1962 et suivantes compte tenu des décalages variables entre autorisation, commencement des travaux et livraison. On a retenu la moyenne des logements terminés en 1962 et 1963 en supposant que cette part de travaux est représentative de la production de logements à l'époque du recensement.

En ce qui concerne les activités autres que celles des logements, les seules données dont on dispose sont celles établies par le BIPE au niveau régional. On a donc calculé le montant départemental des travaux publics et constructions industrielles en éclatant ces données régionales au prorata de l'importance des effectifs BTP du département dans leur région d'appartenance en 1962. On a, en outre, pris la moyenne des travaux 61 et 62 pour tenir compte du fait que le nombre d'actifs du BTP en mars 62 est imputable en partie au niveau d'activité en bâtiments industriels et travaux publics enregistrés en 1961 (compte tenu des délais d'ajustement entre les variations de demande et celles d'emploi).

2.12 - L'influence de l'urbanisation

La construction en milieu urbain implique un certain nombre de contraintes et de servitudes diverses qui, pour beaucoup d'entre elles, allongent les délais de constructions (par rapport à celle effective en rase campagne) et le plus souvent ne sont compensées que par la mise en oeuvre d'outillage perfectionné. Aussi est-il possible de penser que pour un niveau de construction donné, les taux de BTP seront plus importants dans les départements fortement urbanisés.

L'appréhension au niveau départemental de ce phénomène n'est guère aisée puisque par définition le phénomène urbain est ponctuel. Aussi plusieurs définitions ont été essayées pour définir cette variable.

La première a été la taille de commune moyenne du département, c'est-à-dire le rapport de la population départementale au nombre de communes ; comme on pouvait s'y attendre avec le découpage communal français, cette définition n'a donné aucun résultat satisfaisant.

Une seconde définition a été tentée, celle de la "commune en agglomération médiane", c'est-à-dire celle qui est la taille de la commune ou unité urbaine telle que 50 % de la population vit dans une agglomération ou une commune de taille identique ou inférieure et 50 % du reste de la population dans une agglomération ou une commune de taille identique ou supérieure ; cette définition présente l'inconvénient des discontinuités fortes imputables aux définitions des agglomérations et de la grande variation de leur importance.

L'essai d'une troisième définition a consisté à faire une moyenne calculée à partir de la taille de chaque commune ou agglomération, pondérée par la part de population du département qui y vit. (1) Les données présentées le sont dans cette dernière définition.

(1) La définition spatiale des départements et unités urbaines est celle de 1968. Il a fallu tenir compte des agglomérations situées sur plusieurs départements.

Aussi la définition opérationnelle retenue est :

$$\text{TAILLE (D)} = \sum_I \text{POP AG (D,I)} \cdot \frac{\text{POP AG (D,I)}}{\text{POP TOT (D)}} + \sum_J X (J) \cdot \frac{\text{POP AG (D,J)}}{\text{POP TOT (D)}} + \frac{\text{POP RUR (D)}}{N (D)} \cdot \frac{\text{POP RUR (D)}}{\text{POP TOT (D)}}$$

ou - POP TOT (D) est la population totale du département D

- POP AG (D, I) est la population de l'unité urbaine dont toutes les communes sont dans le même département D (l'unité urbaine se définit au seuil de 2000 habitants)
- X (J) est la population d'une agglomération dont les communes appartiennent à 2 départements au moins (ou à un département et un pays étranger)
- POP AG (D, J) est la population de X(J) vivant dans le département D (population nulle le cas échéant)
- POP RUR (D) est la population de l'ensemble des communes rurales du département, communes en nombre N (D)

Il est intéressant de noter que si l'on exclut la Région Parisienne cette grandeur n'est que moyennement corrélée avec la population totale du département comme on peut le voir avec les résultats suivants :

$$\text{POP AG (D, 62)} = 0,259 \cdot \text{POP TOT (D, 62)} - 43\,268 \quad R = 0,54$$

(5,98)

$$\text{POP AG (D, 68)} = 0,283 \cdot \text{POP TOT (D, 68)} - 50\,722 \quad R = 0,57$$

(6,41)

alors que les coefficients de corrélation sont aux alentours de 0,90 si l'on inclut la région Parisienne mais le poids de la constante (qui reste négative) est tel que la régression perd tout son intérêt. Par ailleurs on peut souligner que la croissance de l'urbanisation en province avec cet indicateur est supérieure à celle de l'accroissement de population totale hors région parisienne (+ 16 % contre +6 %) passant de 70 887 en 1962 à 82 377 en 1968 (dans le découpage 68).

...

2.13 - L'influence de la population totale

L'ensemble des hypothèses et de leurs indicateurs étant posé, il ne reste plus qu'à les tester à travers la recherche de modèle prévisionnel satisfaisant.

Enfin on a voulu prendre en considération la population totale(1) pour deux raisons. La première c'est que les dépenses d'équipements collectifs sont peut-être approximées de façon non satisfaisante par la variable travaux publics telle qu'elle résulte de diverses approximations successives, or il est de notoriété publique que la population totale est le critère le plus souvent utilisé dans la répartition des crédits publics d'équipements collectifs susceptibles d'être spatialisés. En second lieu on a voulu tenir compte du fait souligné par diverses études qui semblent établir l'importance des anticipations des entrepreneurs sur le volume de leurs effectifs souhaités, on peut en première approximation penser que cette demande potentielle est proportionnelle à la population du département.

2.2 - Recherche de modèles explicatifs portant sur des variables prises en valeur absolue

On examinera successivement les principaux résultats obtenus avant d'en faire le commentaire.

2.21 - Présentation des résultats

La technique utilisée est celle de la régression optimale dans le cadre de modèle probabiliste. On trouvera page suivante un tableau synoptique des principaux résultats, d'où ont été écartés :

- les essais où la variable endogène est la variation d'effectifs BTP. La juxtaposition de variations positives et négatives ne pouvait d'ailleurs dans un modèle valable pour l'ensemble des départements (et donc sans strati-

(1) Rappelons que l'on vient de voir dans la précédente note que la population totale du département n'est que moyennement corrélée avec la taille moyenne des communes et agglomérations, ces deux variables ne font donc pas double emploi.

- fication a priori) que donner des résultats médiocres, voire franchement mauvais;
- les transformations de certaines variables en logarithme afin de leur faire jouer un rôle multiplicatif dans le modèle explicatif, ces modèles donnant de moins bons résultats;
 - les régressions où tous les coefficients de régression n'étaient pas significatifs au coefficient de confiance de 95 % ; en pratique on a rejeté les variables dont le rapport entre le coefficient de régression et l'écart type de ce dernier était compris entre -2 et +2. C'est la valeur de ce rapport qui figure entre parenthèses en dessous de celle du coefficient de régression dans le tableau synoptique.

nombre de variables retenues dans la régression optimale	variables coef de détermination	nombre de logements terminés HLM	nombre de logements terminés non HLM	Production de travaux publics	taille urbaine moyenne	population totale	nombre de logements construits	effectifs BTP en 54	Constante
GROUPE 1		X	X	X	X	X			
2	0,99268		1,77114 (15,88)			0,82614 (41,43)			735
3	0,99352		1,53257 (11,99)	4,70740 (3,31)		0,02130 (13,51)			941
GROUPE 2		X	X	X	X	X		X	
2	0,99268		1,77114 (15,88)			0,82614 (41,43)			735
3	0,99456		1,69090 (17,27)			0,01875 (12,74)		0,29967 (5,41)	210
4	0,99608		1,52653 (17,28)		0,00191 (5,72)	0,01367 (8,90)		0,40505 (7,99)	1194
GROUPE 3				X	X	X	X		
2	0,99134					0,02378 (27,92)	1,48825 (14,13)		1095
3	0,99231			5,09957 (3,27)		0,01896 (11,29)	1,26139 (10,38)		1264
GROUPE 4				X	X	X	X	X	
2	0,99134					0,02378 (27,92)	1,48825 (14,13)		1095
3	0,99348					0,01596 (9,65)	1,42221 (15,34)	0,31979 (5,29)	519
4	0,99465				0,00170 (4,29)	0,01179 (6,57)	1,27631 (14,00)	0,41402 (6,98)	1361

2.22 - Examen des résultats

L'examen du tableau synoptique de la page précédente mène aux remarques suivantes :

- a) Dans tous les cas de figures, le coefficient de détermination est exceptionnellement élevé. Il faut toutefois tempérer cette remarque par le fait que la variance de l'emploi BTP départemental est près de 30 000 fois plus importante que la moyenne.
- b) La stratification des logements en HLM et autres qu'HLM amène une amélioration certaine dans la qualité de l'ajustement (si l'on prend comme critère le coefficient de corrélation).

Toutefois seul le secteur non HLM est significatif. La faible signification du secteur HLM s'explique sans doute par le fait que les techniques utilisées et la taille des chantiers mis en oeuvre sont moins "consommatrices" de main d'oeuvre (indépendamment des fuites signalées à la page 261) que les logements individuels dont la construction relève encore en grande partie de techniques artisanales fortement consommatrices de main d'oeuvre.

- c) La population totale intervient dans tous les cas de figure sauf un et même avant les effectifs BTP 54 lorsque ces deux variables sont en concurrence. Cette variable intervient pour près d'un tiers dans les régressions de rang supérieur à 2 dans l'explication du BTP 62 (au barycentre du nuage de point), ce qui est un peu gênant pour une variable à vocation de proxis, mais contrebalance la faible représentativité de la variable travaux publics dans les diverses régressions.

- d) La variable "montant des travaux publics" n'intervient que lorsque la variable "BTP 54" n'est pas retenue dans le groupe de variables retenues. Il est vraisemblable que le mode de calcul retenu, l'influence des disparités de productivités et les "exportations de production" expliquent en grande partie cela et que les disparités de population totale et d'actifs du BTP soient des proxis suffisants de cette partie de l'activité BTP. Ce résultat est malgré tout décevant dans la mesure où il prive le modèle d'une éventuelle variable instrumentale.

2.3 - Recherche d'un modèle explicatif portant sur des variables prises en taux

Il a d'abord été nécessaire de définir les dénominateurs à prendre en considération avant de pouvoir faire l'étude proprement dite.

2.31 - Choix des dénominateurs

La variable effectif BTP a été rapportée à la population active employée dans le département correspondant. Ce dénominateur a été retenu de préférence à la population totale (encore que, compte tenu de la très faible dispersion des taux d'emploi, les disparités spatiales doivent être sensiblement les mêmes) parce que ce taux représente d'une certaine façon le degré de priorité accordé par le département à l'effort construction.

Les logements terminés ont été rapportés à la population totale. Il aurait été possible de calculer une population à desservir en priorité en tenant compte d'indice de surpeuplement, mais comme on le sait en la matière, la demande objective et la demande solvable sont loin de coïncider. Aux deux principaux circuits de financement de la construction correspondent en fait 2 types de catégorie de demandeur : pour les HLM il s'agit des ouvriers et employés, pour les autres logements il s'agit principalement des cadres moyens et supérieurs ainsi que des professions libérales. Ces deux populations ont donc été respectivement retenues pour la définition de ces deux taux.

On a conservé la taille urbaine moyenne telle quelle puisqu'elle ne pouvait qu'être rapportée à une valeur moyenne France entière (ce qui revient à la multiplier par un scalaire et ne fait que multiplier le coefficient de régression que l'on aurait obtenu sans cette multiplication par l'inverse de ce scalaire). Il en est de même pour la population totale du département. Toutefois pour présenter plus facilement les résultats ces grandeurs ont été divisées par 100 000.

Les essais effectués avec un volume de travaux publics rapporté à la population totale ou à la population active ayant donné les mêmes résultats (rejet) on a groupé ces deux taux dans la même colonne du tableau synoptique. D'ailleurs à partir du moment où il y a rejet assez net avec l'un des taux, on pouvait s'attendre

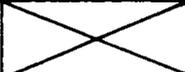
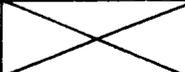
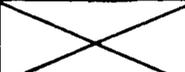
avec la relative stabilité du taux d'emploi, à un rejet automatique de l'autre taux.

2.32 - Etude des résultats

On trouvera page suivante le tableau synoptique des résultats obtenus. Ceux ci amènent les observations suivantes :

- a) les coefficients de détermination sont nettement inférieurs à ceux obtenus dans les essais en valeur absolue puisqu'ils ne dépassent pas 0,75. Mais ces résultats restent malgré tout satisfaisants, puisque ces régressions expliquent quand même environ les 3/4 de la variance totale de la variable endogène.
- b) L'activité en travaux publics n'est plus du tout explicative des disparités départementales des taux de BTP et ce, que le dénominateur soit la population totale ou la population active.
- c) Lorsque l'on introduit le taux de BTP 54, la stratification des logements terminés en deux secteurs n'apporte pas d'amélioration très sensible dans l'explication.
- d) Le poids relatif de la population du département dans le peuplement de la France a bien une influence, mais celle ci contrairement à ce que l'on pouvait attendre, est négative. Toutefois, contrairement au modèle en effectifs l'influence de ce facteur est faible puisqu'il n'intervient que pour environ 1 % de l'explication du barycentre dans la régression. Il se produit donc, compte tenu d'un poids plus important des autres variables, un effet inverse à celui d'une polarisation qui fait apparaître dans les départements les moins peuplés l'existence d'un minimum du taux d'emploi BTP ou si l'on préfère, dans les départements les plus peuplés, l'existence d'économies d'échelle. Cette interprétation n'est pas contradictoire avec celles faites précédemment puisque l'effet population est en grande partie déjà intégré dans le dénominateur du taux.

RECHERCHE D'UN MODELE EXPLICATIF DE L'EMPLOI DEPARTEMENTAL BTP (TAUX)

nombre de variables retenues dans la ré- gression optimale	variables coef de détermination	HLM terminés	non HLM terminés	logements	taille	population	production de TP /	effectifs	constante
		(ouvriers employés)	(cadres et prof. libérales)	terminés 1;100 population totale	urbaine moyenne / 100.000	totale/ 100.000	(population totale ou active)	BTP 54 Pop active 54	
GROUPE 1		X	X		X	X	X		
2	0,63394	0,67537 (2,10)	0,65066 (12,07)						0,04908
GROUPE 2		X	X		X	X	X	X	
2	0,72957		0,50847 (9,82)					0,42806 (6,03)	0,03253
3 (non optimal)	0,74052	0,51829 (1,89)	0,51471 (10,07)					0,41513 (5,91)	0,02764
4	0,75121	0	0,50354 (10,01)		0,00455 (2,06)	-0,00101 (2,61)		0,45406 (6,50)	0,03506
GROUPE 3				X	X	X	X		
2	0,63090			5,66024 (12,06)		-0,00064 (3,15)			0,05857
GROUPE 4				X	X	X	X	X	
2	0,70181			4,18786 (8,11)				0,43020 (5,72)	0,03309
3	0,73663			4,43829 (9,851)		-0,00053 (3,35)		0,41577 (5,84)	0,03556

2.4 - Choix du modèle définitif et limites de ce modèle

Quelques remarques préalables sur les règles de choix sont à faire pour expliquer celui-ci, avant d'en faire la critique et de comparer les résultats du modèle choisi pour 1968 avec les observations.

2.41 - Choix du modèle définitif

Si l'on se place dans le cadre d'un modèle probabiliste en valeur absolue, la première remarque du paragraphe précédent nous amène à considérer comme à peu près équivalents les 11 modèles possibles puisque la différence entre les extremums de coefficients de détermination est de l'ordre de 5 pour 1000 ce qui est négligeable. Toutefois compte tenu de la très grande dispersion du phénomène⁽¹⁾ on retiendra la régression correspondant au meilleur coefficient de détermination soit :

$$(\text{effectifs BTP 62}) = 1,52653 \cdot (\text{logements terminés non HLM}) + 0,00191 \cdot (\text{taille urbaine moyenne}) + 0,01367$$

$$(\text{Population totale}) + 0,40505 (\text{effectifs BTP 54}) + 1194$$

L'approche en terme de taux donne, on l'a vu, un coefficient de corrélation nettement moins important, mais cette solution comme on l'a signalé au paragraphe précédent n'est pas à écarter d'office pour autant. Dans ce second type d'approche la "qualité" des différents modèles trouvés, au regard du coefficient de détermination, est nettement plus différenciée, aussi le modèle le plus approprié semblerait être le suivant:

$$\left[\frac{\text{effectifs BTP 62}}{\text{population active 62}} \right] = 0,50354 \cdot \left[\frac{\text{HLM terminés 62}}{\text{ouvriers et employés 62}} \right] + 0,00455 \cdot \left[\frac{\text{taille urbaine moyenne 62}}{100\ 000} \right]$$

$$- 0,00101 \left[\frac{\text{Population totale 62}}{100\ 000} \right] + 0,45406 \left[\frac{\text{effectif BTP 54}}{\text{population active 54}} \right] + 0,03506$$

On a calculé comme d'habitude la variance résiduelle correspondant à la transposition en effectifs du modèle en taux. La comparaison défavorable dans un rapport de 1 à 10 de cette dernière donnée avec la variance résiduelle trouvée dans le modèle en effectifs nous amène à retenir le modèle en effectifs (voir les données utilisées page T 10.32).

Examinons maintenant les limites de ce modèle.

(1) Rappelons que la variance est 30 000 fois plus importante que la moyenne.

2.42 - Critiques du modèle retenu

On examinera les résultats sous l'angle quantitatif avant de donner les limites d'un tel modèle.

a) Critiques des résultats du modèle

α - dans le cadre d'un modèle déterministe

La moyenne des erreurs relatives, prises en valeur absolue, est pour 1962 de 6,7 % avec toutes les erreurs relatives prises en valeur absolue inférieures à 20 %, à l'exception du plus petit département, le territoire de Belfort qui lui, atteint 28 %. Ce résultat est assez satisfaisant si l'on songe que 26 départements dépassaient ce cap des 20 % dans le cadre du modèle "homothétique" vu au 1.21.

La carte des écarts entre valeurs observées et valeurs calculées pour 1962, comparée à celle du "modèle homothétique" révèle deux similitudes marquantes: une sous-estimation sur les régions Rhône-Alpes et Provence-Côte d'Azur et une surestimation sur le nord de la France ; il y a donc rétrécissement des zones de franche surestimation et sous-estimation

β - dans le cadre d'un modèle probabiliste

Par ailleurs on a poursuivi jusqu'au bout la logique de l'approche probabiliste en tenant compte des remarques faites au chapitre 3 (1). On a donc calculé pour 1962 la valeur de l'écart-type de la valeur calculée pour chaque département et donné un intervalle approximatif ($t = 2$), avec un coefficient de confiance à 95 % des valeurs prises par la variable endogène. Les résultats trouvés (voir tableau page T 10.33) présentent un caractère inquiétant du fait de la forte amplitude de la fourchette. Mais celle ci n'a rien d'étonnant :

En effet la prévision du modèle dépend à un coefficient près (valeur du t de Student pour le coefficient de confiance donné), du rapport de l'écart type de la valeur calculée, à la valeur observée. (voir carte page T 10.34)

(1) Voir page 76, note de bas de page : ajoutons que ces calculs ont été faits pour tous les sous-modèles tant au niveau du "calibrage" du modèle qu'à celui de la comparaison des prévisions aux réalisations.

Cet écart-type, toujours supérieur à l'écart-type résiduel (ici 1461) a un poids d'autant plus grand que la valeur observée du département est petite. Or la modulation de cet écart-type est faible puisque celui ci dépasse rarement 1500 pour atteindre seulement 2000 dans le département de la Seine, il s'ensuit donc que la précision dans le cadre du modèle probabiliste est, en gros, inversement proportionnelle à l'observation faite pour 1962. La dispersion des valeurs départementales est telle que l'approche probabiliste, à moins de stratifier les observations départementales et faire autant de modèles (ce qui risque de mener à un arbitraire assez important), désavantage les distributions fortement dissymétriques. C'est le cas de la distribution départementale, et dans le cas présent, par exemple, 65 des 89 départements métropolitains ont des effectifs BTP inférieurs à la moyenne départementale. Il est donc inéluctable que la précision d'un modèle probabiliste soit franchement mauvaise dès que l'écart-type résiduel du modèle est important au regard de la plupart des valeurs observées (ou calculées).

b) Critiques du choix du modèle

Si le modèle est relativement satisfaisant du point de vue descriptif, il n'en reste pas moins que du point de vue explicatif, deux variables sur les quatre n'ont guère qu'une valeur de proxis et une d'inertie des structures, ce qui interdit une utilisation en mode prévisionnel à long terme de ce modèle d'une part et d'un point de vue opérationnel le rend peu intéressant du fait du manque de variables instrumentales d'autre part.

2.43 - Etat Potentiel 1968

Faire "tourner" le modèle pour 1968 présente des difficultés de données que l'on examinera avant d'étudier les résultats de l'état 1968 ainsi généré.

a) Les données intégrées dans le modèle prévisionnel

Le problème de changement de définition des agglomérations entre 1962 et 1968 ne se posait pas puisque les données avaient été calculées dans la définition 1968. Le seul problème dans ce modèle dynamique restait celui de la non coïncidence du décalage temporel entre une variable endogène et une variable exogène (en l'occurrence ici les effectifs BTP). Ce décalage de 8 ans utilisé pour la détermination des paramètres est ramené à 6 ans dans l'utilisation prévisionnelle. Plusieurs solutions sont possibles pour résoudre ce problème :

- la première consiste à calculer le coefficient de régression en prenant une valeur 1956 interpolée des valeurs observées 1954 et 1962. Cette solution a l'inconvénient d'intégrer une partie de la variable expliquée dans la variable explicative.

- La seconde consiste à considérer que ce qui importe c'est avoir une approximation de la structure effectif BTP de départ que les autres variables du modèle sont censées modifier, en conséquence de quoi on peut considérer que le vecteur observé en 1962 est, à un coefficient de spécificité près, représentatif de celui que l'on aurait observé en 1960, ce qui permet d'utiliser le coefficient de régression défini pour une période de 8 ans. Ce coefficient de spécificité est égal au rapport de la valeur interpolée France entière 1960 à la valeur observée France entière en 1962 soit : $(1676 - 0,25 (1676 - 1358))/1676$.

C'est cette dernière méthode qui a été retenue.

b) Etat potentiel 1968

Les résultats de l'utilisation du modèle pour calculer un état potentiel 1968 (voir tableau et cartes pages T. 10.35 à 36⁽¹⁾) sont d'une certaine façon assez décevants puisque ce modèle sous estime d'environ 9 % (~ 156000) l'emploi BTP des 87 départements métropolitains non compris dans la Région Parisienne et que la sous-estimation est à peu près systématique (73 départements). On peut ajouter que l'utilisation de la technique des coefficients de spécificité de régression donne un écart-type résiduel légèrement plus important et que le cadrage sur la valeur observée pour l'ensemble des effectifs BTP de province n'améliore l'écart-type résiduel d'environ 50 % et un peu moins dans le cadre de l'utilisation des coefficients de spécificité.

Ce fait s'explique par une très faible croissance du BTP non parisien (environ 19 %) alors que l'augmentation des diverses variables endogènes n'est que moyenne. Il faut bien reconnaître qu'une telle mutation est difficile à prévoir dans un modèle de ce genre. Face à un tel résultat deux attitudes sont possibles : ou considérer que cette discordance est inhérente à la notion d'état potentiel, ou cadrer le modèle départemental sur un modèle national, considérant que le modèle départemental n'apporte que des informations sur la structure spatiale du phénomène⁽²⁾ On verra ultérieurement quelle attitude a été choisie.⁽³⁾

(1) Voir aussi page A 12. Le cas du B T. P. a été retenu pour illustrer (dans la sélection restreinte de cartes) un cas de prévision médiocre, à l'aide de ces techniques économétriques.

(2) Voir chapitre 5.

(3) Voir chapitre 13

Le sous-modèle agriculture

11

CHAPITRE 11 - Le SOUS MODELE AGRICOLE

La production agricole est la seule à être liée à un facteur de production par essence immobile : la terre ; il en découle des problèmes spécifiques d'évolution très différents de ceux observés dans d'autres secteurs. Ces problèmes résultent de la surabondance de cette production et leur importance ressentie est liée à deux facteurs politiques dans tous les pays occidentaux : d'une part à la répercussion des prix agricoles sur le niveau moyen des prix⁽¹⁾ et de son incidence sur le consommateur-électeur, et d'autre part à la pression à la baisse des effectifs agricoles techniquement nécessaires et les tensions sociales qui en découlent, avec un impact plus grand encore sur l'agriculteur-électeur. La dimension et la complexité de ces problèmes sont partout telles que dans tous les pays occidentaux le Ministère de l'Agriculture est considéré comme l'un des ministères stratégiques.

En France ce n'est qu'à partir de 1945 que le secteur agricole s'est mis à bouger⁽²⁾ : après une stabilité de 20 ans de la part de la population agricole à 33 % de la population active, celle-ci décroît de plus de moitié dans les 23 années suivantes pour atteindre 15 %. Parallèlement le nombre d'exploitations diminue très sensiblement puisque plus du quart d'entre elles (26,1) disparaissent entre 1955 et 1967 et que la superficie moyenne s'accroît de 27 % pour atteindre 18 hectares en 1967 ; (toutefois les 3/4 des exploitations ont constamment une superficie inférieure à la surface moyenne. Cette évolution trouve sa cause non seulement dans le progrès technique mais aussi dans l'influence croissante de facteurs socio-économiques (mode de vie, évolution de la consommation) et s'accompagne d'une intégration verticale croissante du secteur agricole dans le reste de l'économie).

(1) 31 % du budget moyen des ménages français est consacré à des achats alimentaires (dans l'indice des 295 postes actuellement en vigueur).

(2) Pour une étude sur une très longue période voir Poutard "Les disparités régionales dans la croissance de l'agriculture française" (Gauthier - Villars 1965) qui présente une description très intéressante de ces disparités depuis 1840 et un essai de recherche de leurs causes d'évolution.

Si l'on se réfère aux travaux préparatoires du VIème Plan les problèmes agricoles risquent de tenir la vedette encore longtemps⁽¹⁾ aussi est-il intéressant de faire un point rapide des diverses observations que l'on peut faire sur l'évolution de l'emploi agricole avant de rechercher un modèle explicatif satisfaisant.

(1) Extrait du "Rapport Général de la Commission sur l'avenir à long terme de l'agriculture française 1968 - 1985 (juin 1969) dit "Rapport Vedel" du nom de son rapporteur : "La Commission fonde son diagnostic sur trois points :

- compte tenu des hypothèses portant sur la demande alimentaire et le progrès des techniques que la Commission a retenu dans une première approche, il apparaît que le surplus relatif des productions agricoles tendra à s'aggraver d'ici à 1985, à moins qu'une politique agricole nouvelle ne soit imaginée et efficacement appliquée.
- les perspectives démographiques ne sont pas telles qu'elles puissent améliorer spontanément les conditions même de l'activité des agriculteurs dont un tableau de la situation présente et prévisible est dressé
- aucune des solutions préconisées habituellement ne peut tenir lieu de panacée, car aucune d'elles n'est, à elle seule, à la mesure du problème. Toutes méritent sans doute d'être reprises ou améliorées mais dans le cadre d'une révision plus fondamentale de la politique agricole".

SECTION 1 - ETUDE DESCRIPTIVE DE L'EVOLUTION DE LA POPULATION AGRICOLE

On étudiera globalement l'emploi agricole avant d'en étudier les différences spatiales :

1.1 - Etude globale de l'évolution de la population active agricole entre 1954 et 1968

Les caractéristiques de l'emploi agricole qui ont le plus souvent retenu l'attention sont : l'âge, le sexe, et le statut qui du reste constituent à peu près les seules données que l'on peut retirer des recensements pour cette étude. L'autre source possible pour étudier certaines caractéristiques de cette évolution est l'utilisation des enquêtes de 1955, 1963 et 1967, mais celles-ci présentent l'inconvénient de fournir des renseignements non homogènes avec les données des recensements⁽¹⁾. On réutilisera donc les données tirées de ces enquêtes que dans la seule mesure où celles-ci ne peuvent être tirées des recensements.

La comparaison des données fournies par les trois derniers recensements n'a pas à notre connaissance été menée par sexe et statut, au niveau global aussi avons nous jugé utile d'examiner les tendances qui se dégagent de ces données. Nous avons ensuite complété les analyses longitudinales effectuées par le SEDES par une analyse de structure qui fournit une vision d'ensemble qui, nous semble-t-il explique certains aspects de l'évolution agricole. Enfin, nous avons procédé à une étude complémentaire des données migratoires partiellement analysées dans les chapitres 8 et 9 pour examiner le phénomène des migrants actifs agricoles en 1968.

1.11 - Etude de l'influence du sexe et du statut

On trouvera page suivante un tableau synoptique de l'évolution structurelle des effectifs agricoles entre les trois derniers recensements.

L'évolution est particulièrement rapide puisque pour la première période les effectifs agricoles diminuent d'un quart et d'un cinquième pour la seconde période ce qui correspond en fait à une accélération du phénomène d'une période sur l'autre puisque le taux moyen de décroissance annuel passe de - 3,43 % à - 3,76 %.

(1) où l'unité de base est le logement et non l'exploitation comme dans les enquêtes.

Les caractéristiques de cette baisse sont assez différentes d'une période à l'autre. Dans un premier temps c'est une véritable purge qui s'opère chez les aides familiaux qui constituent plus de la moitié des départs et portent sur le tiers d'entre eux, pour les départs de cette catégorie le poids des femmes est prépondérant ce qui s'explique par leur importance relative en 1974, mais l'exode relatif des hommes (environ 2 sur 5) est de loin celui le plus important. Le reste est surtout imputable aux salariés agricoles masculins et aux indépendants et exploitants masculins, toute fois pour l'importance de la part de cette dernière catégorie dans la variation globale résulte de son poids dans les structures en 1954 car son exode relatif est le plus bas observé.

ACTIFS AGRICOLES : STRUCTURE par SEXE et STATUT

Effectifs d'actifs agricoles (unité : millier, définition catégorie d'activité économique)

1954 : 5 142
1962 : 3 888
1968 : 3 087

STRUCTURE par SEXE et STATUT pour 1000 ACTIFS AGRICOLES aux
RECENSEMENTS de 1954, 1962, 1968

	Indépendants et employeurs			Aides familiaux			Salariés agricoles			Ensemble		
	54	62	68	54	62	68	54	62	68	54	62	68
Homme	316	372	398	137	104	97	191	191	179	644	667	674
Femme	55	59	55	267	248	247	34	26	24	356	333	326
Ensemble	371	431	453	404	352	344	225	217	203	1000	1000	1000

STRUCTURE par SEXE et STATUT de la DIMINUTION de 1 000 ACTIFS AGRICOLES
entre les RECENSEMENTS de 1954, 1962 et 1968

	Indépendants et employeurs		Aides familiaux		Salariés agricoles		Ensemble	
	54 - 62	62 - 68	54 - 62	62 - 68	54 - 62	62 - 68	54 - 62	62 - 68
Homme	149	272	234	133	192	238	575	643
Femme	41	76	327	247	57	34	425	357
Ensemble	190	348	561	380	249	272	1 000	1 000

VARIATION RELATIVE (unité 100) de la POPULATION ACTIVE AGRICOLE

	Indépendants et employeurs		Aides familiaux		Salariés agricoles		Ensemble	
	54 - 62	62 - 68	54 - 62	62 - 68	54 - 62	62 - 68	54 - 62	62 - 68
Homme	- 11,8	-15,0	- 42,7	- 26,2	- 25,1	- 25,6	- 22,3	- 19,8
Femme	- 18,7	- 26,5	- 30,5	- 20,5	- 41,9	- 26,5	- 29,8	- 22,0
Ensemble	- 12,8	- 16,6	- 34,7	- 22,2	- 27,6	- 25,7	- 25,0	- 20,5

TAUX ANNUEL MOYEN de DEPART⁽¹⁾

(unité = 100)

	Indépendants et employeurs		Aides familiaux		Salariés agricoles		Ensemble	
	54 - 62	62 - 68	54 - 62	62 - 68	54 - 62	62 - 68	54 - 62	62 - 68
Homme	- 1,56	- 2,68	- 6,73	- 4,94	- 3,54	- 4,81	- 3,10	- 3,61
Femme	- 2,56	- 5,00	- 4,45	- 3,75	- 6,57	- 5,01	- 4,33	- 4,18
Ensemble	- 1,7	- 2,98	- 5,18	- 4,09	- 4,18	- 4,83	- 3,43	- 3,76

(1) Ce taux x est tel que x^n , où n est égal à 6 ou 8 selon les périodes, est égal au rapport de la valeur de fin de période sur celle de début.

Dans un second temps l'importance relative des trois statuts s'équilibre à peu près dans la variation d'effectifs, ce qui résulte d'une transformation importante des comportements migratoires : l'exode masculin s'accélère tandis que l'exode féminin fléchit (sauf dans la catégorie indépendants et employeurs), tout en ayant comme pour la période précédente, le poids le plus faible dans la variation totale et la ponction relative la plus importante. Par ailleurs, l'exode relatif des salariés agricoles devient plus fort que celui des aides familiaux mais ces deux catégories continuent à connaître (quel que soit le sexe) un taux de départ supérieur à la moyenne, ce qui fait que les exploitants agricoles ont tendance à être de plus en plus seuls.

Ces évolutions se traduisent par un accroissement sensible du "poids" des exploitants masculins dans la structure observée de population active agricole, au détriment des salariés agricoles et surtout des aides familiaux (où la baisse de poids est deux fois plus importante chez les hommes que chez les femmes). Mais comme on a pu le constater, cette relative stabilité dans l'évolution des structures s'est faite dans des conditions très différentes sur les deux périodes, il s'en suit que toute démarche basée sur l'extrapolation de tendance de ces modifications de structures semble vouée d'avance à l'échec.

1.12 - Etude de l'influence de l'age et du sexe

Il semblerait qu'à cause de l'hécatombe de la première guerre mondiale et du freinage de l'exode rural imputable à la crise des années 30, la population des agriculteurs ait été anormalement jeune au lendemain de la seconde guerre mondiale⁽¹⁾. Le vieillissement récent de la pyramide des ages des actifs agricoles a facilité indubitablement, par l'importance des départs, une partie des problèmes de l'adaptation des effectifs agricoles aux contingences de modernisation de l'agriculture française et explique dans une certaine mesure le succès apparent de la politique de transformation des structures agricoles en France.

(1) voir note provisoire ronéotée du Ministère de l'agriculture "Perspective d'évolution à long terme de la population active agricole masculine" p 13 (1970)

L'appel aux techniques d'analyse démographique (en particulier, analyse longitudinale) esquissées à l'occasion du recensement de 1962⁽¹⁾ devait porter pleinement ses fruits dans des travaux consécutifs au recensement de 1968, dont la maîtrise d'oeuvre appartenait au SEDES⁽²⁾ et dont on examinera rapidement la méthode dans le paragraphe suivant.

Les variations de cohortes d'actifs agricoles entre 1962 et 1968 sont imputées à des mutations pour les cohortes de moins de 45 ans en 1962 et aux retraités et décès pour les cohortes plus âgées. Un relèvement de 5 ans de cette barre ne modifie d'ailleurs pas les ordres de grandeur des conclusions de cette analyse. L'application des tables de mortalité agricole permet de séparer les décès des retraités (définie résiduellement). En ce qui concerne les entrées des jeunes dans l'agriculture on a considéré que la cohorte des 15 - 20 ans en 1968 était constituée exclusivement d'enfants de ménages agricoles ; on peut noter à ce propos que 24 % seulement des enfants d'agriculteurs rentrent dans la profession, pourcentage particulièrement faible car la moitié d'entre eux abandonneront l'agriculture dans les quelques années qui suivent.

Le tableau résumant les résultats de cette analyse (voir page suivante) révèle que les entrées dans la profession compensent à peine les mutations et que l'essentiel de la diminution est imputable aux retraités ($\sim 2/3$) et aux décès ($\sim 1/3$). Un modèle prévisionnel⁽³⁾ basé sur la constance des divers taux et des pyramides d'âge a conduit certains experts à pousser un cri d'alarme demandant de cesser de favoriser la non entrée des enfants d'agriculteurs dans la profession, prévoyant d'ici 15 ans des problèmes de recrutement sérieux dans l'agriculture.

(1) voir "Espace Economique Français" tome II p 20 et 32

(2) SEDES "Etude démographique de la Conversion" mai 1972

(3) voir conclusion de la note du Ministère de l'Agriculture précitée.

VARIATION des COHORTES d'ACTIFS AGRICOLES entre 1962 et 1968⁽¹⁾

unité : milliers

âge en		Hommes	Femmes	Ensemble
1962	1968			
9 - 14 ans	15 - 20 ans	+ 187	+ 81	+ 268
15	21	- 17	- 18	- 35
16	22	- 29	- 40	- 69
17	23	- 41	- 50	- 91
18	24	- 33	- 53	- 86
19	25	- 41	- 52	- 93
20	26	+ 17	- 47	- 30
21	27	+ 166	- 41	+ 125
22	28	+ 163	- 33	+ 130
23	29	- 21	- 27	- 48
24	30	- 24	- 19	- 43
25	31	- 20	- 13	- 33
26	32	- 17	- 9	- 26
27	33	- 17	- 4	- 21
28	34	- 14	- 3	- 17
29	35	- 13	- 2	- 15
30 - 44 ans	36 - 50 ans	- 99	- 4	- 103
45 - 59 ans	51 - 65 ans	- 185	- 115	- 300
60 - 64 ans	66 - 70 ans	- 203	- 102	- 305
65 - 68 ans	71 - 74 ans	- 100	- 55	- 155
69 et plus	75 et plus	- 153	- 72	- 225

COMPOSITION de la VARIATION de 1 000 ACTIFS AGRICOLES entre 1962 et 1968⁽²⁾

unité : milliers

	Hommes	Femmes	Ensemble
Métropolitains entrant dans la profession	+ 220	+ 81	+ 301
Rappatriés	+ 10	+ 1	+ 11
Autres immigrés	+ 45	+ 4	+ 49
Total des entrées	+ 275	+ 86	+ 361
Mutations	- 286	- 91	- 377
Retraités	- 386	- 275	- 661
Décès	- 263	- 70	- 323
Total des sorties	- 936	- 435	- 1 361
Solde	- 650	- 350	- 1 000

(1) source : RGP 1962 tableau D3 fascicule France entière, et RGP 1968 sondage au 1/4 tableau mécanographique PA9/Q

(2) Renseignements dérivés de l'analyse longitudinale effectuée par le SEDES (Etude démographique de la conversion) l'étude est menée en CSP d'où le léger écart avec la composition par sexe donnée dans le tableau précédent.

Il aurait été intéressant de disposer d'une étude montrant l'importance du facteur âge dans l'évolution par statut et sexe des actifs agricoles depuis 1954, mais malheureusement celle-ci ne semble pas avoir été faite à ce jour.

1.13 - Migrations et emploi agricole

L'exploitation faite⁽¹⁾ du tableau MI-501 du recensement de 1968 permet de se faire une idée sur la mobilité spatiale des agriculteurs (voir tableau et carte pages T 11. 26 à 28).

Il importe tout d'abord de souligner que ces données regroupent les exploitants et salariés exerçant une activité agricole en 1968 mais que l'on ignore si le migrant exerçait une activité en 1962 et dans l'affirmative laquelle ? Il est toutefois permis de penser, compte tenu des rigidités des filières d'accès à la profession, que les migrations vers la profession d'exploitant agricole sont faibles.

Une première remarque s'impose au vu des résultats : les migrations agricoles sont très faibles puisque seulement 2,5 % des agriculteurs contre 13,1 % des actifs ont changé de département au cours de la période 62 - 68. Mais si l'on examine la provenance de ces migrants agricoles, on constate que 15 % d'entre eux avaient leur résidence dans une agglomération de moins de 50 000 habitants et 19 % dans une agglomération de plus de 50 000 habitants. On peut penser que la majeure partie de ces migrants ne résidant pas habituellement "à la campagne" sont des salariés non pas d'exploitants agricoles mais d'entreprises à vocation agricole (organismes divers de sécurité sociale, mutuelle, organismes techniques, ...). Il semble donc qu'il faille retenir, comme ordre de grandeur, qu'environ 50 000 exploitants agricoles ont changé de département, ce qui représente une mobilité comparable à l'apport d'agriculteurs et salariés agricoles en provenance de l'extérieur.

Du point de vue spatial les immigrants en provenance de communes rurales sont d'autant plus mobiles qu'ils sont originaires d'une zone comprenant le Bassin Parisien et s'étendant au sud jusqu'à la Loire et à l'ouest aux limites de la Bretagne.

(1) On trouvera au chapitre 8 les commentaires de cette exploitation réalisée pour les besoins de notre étude

On peut ajouter en outre qu'entre les départements on assiste à un véritable mouvement brownien puisque seulement 89 soldes sont significativement non différents de zéro(1) et que le nombre maximum observé de balance positive est de quatre. On peut ajouter que les deux tiers des départements ont au plus une balance positive et que l'importance de ces balances est très faible puisque la plus importante (Deux-Sevres vers la Vienne) est de 204 seulement.

1.2 - Etude départementale de l'évolution de la population agricole

On s'intéressera à chacune des deux périodes avant d'examiner la "synthèse" apportée par les techniques de régression multiple.

1.21 - Evolution 1954 - 1962

a) Evolution des effectifs départementaux

Il faut d'abord souligner la grande homogénéité de l'évolution départementale de l'emploi agricole. Celui-ci décroît d'environ un quart sur la période et l'on pourrait s'attendre à une assez grande dispersion : en fait 61 départements connaissent par rapport à l'évolution moyenne (France entière(2)) un écart relatif, pris en valeur absolue, inférieur ou égal à 5 % et seulement 11 départements ont un écart relatif pris en valeur absolue supérieur ou égal à 10 % (dont un seul dépasse 20 %).

Une résistance à la baisse est assez nette sur la façade maritime du Nord-Ouest et concerne quatre régions de programme : Nord, Haute et Basse Normandie, Bretagne. La baisse des effectifs agricoles est supérieure à la baisse générale sur une bande Bordeaux-Genève (sauf le Rhône et l'Ain), sur l'ensemble du massif des Alpes, ainsi que dans la couronne Parisienne (et en particulier à l'est de celle-ci : Aisne, Seine et Marne, Yonne).

Examinons maintenant si le statut explique certains des écarts.

(1) seuil défini à 50 compte tenu de la précision due aux techniques déchantillonnage (voir p 188)

(2) c'est-à-dire que l'on compare les effectifs observés en 1962 avec ceux calculé pour cette date, homothétiques de ceux observés en 1954, le rapport d'homothétie étant le rapport des effectifs de la France entière entre 62 et 54 (voir tableaux et cartes p. T11. 29 à 34, ainsi que la page A 13).

b) Influence du statut

On a vu que la variation d'actifs agricoles est imputable pour un peu plus de 55 % aux aides familiaux et 26 % aux salariés agricoles. Du point de vue de la distribution spatiale du phénomène on peut noter que plus de la moitié des départs apparents sont imputables aux salariés agricoles dans les départements du Bassin Parisien et que la gamme de l'importance relative de ces salariés est très étendue puisqu'elle va de 3 % (Savoie) à 81 % (Oise). Pour les aides familiaux leur prépondérance n'intervient que pour 53 départements seulement sur les 89 et une rétention importante s'opère sur la façade maritime du nord-ouest de la Bretagne à la Région Nord où la tendance à la baisse est nettement inférieure à la tendance moyenne (ce que l'on avait observé déjà pour l'ensemble des actifs agricoles), ce qui n'empêche pas la Bretagne de voir sa variation d'actifs agricoles imputable aux aides familiaux dans une proportion nettement supérieure à la moyenne. On peut ajouter que là où les salariés agricoles sont partis en force il y a eu une assez forte rétention d'aides familiaux.

Si l'on descend un peu dans le détail, il semble que la relative homogénéité des départs d'agriculteurs résulte du jeu conjugué d'évolutions par statut, très hétérogènes dans l'espace. Cette divergence entre ces phénomènes s'explique sans doute par la relative substitution des aides familiaux et des salariés agricoles et il semblerait qu'il faille, au niveau de la prévision, tenir compte de ce fait.

c) Evolution du taux d'emploi agricole (par rapport à la population active)

On observe tout d'abord que 63 départements sur 89 ont un taux sous estimé alors que 39 seulement avaient des effectifs sous estimés mais tous ces derniers départements (hormis le Rhône, la Seine et les Yvelines, département de très fort peuplement) sont aussi sous estimés quant aux taux.

D'autre part, tous les départements à taux sous estimés à l'exception de 7 (dont 3 seulement n'ont pas d'agglomération de plus de 200 000 habitants) connaissent un taux d'emploi agricole supérieur à la moyenne nationale.

On observe donc une résistance à la baisse plus forte dans les départements restés proportionnellement les plus agricoles mais ceci n'empêche pas (compte tenu de l'évolution de l'emploi global départemental) qu'en général la répartition de la baisse d'effectifs se fasse en gros au prorata des effectifs agricoles de chaque département. Ce phénomène se comprend plus facilement si l'on se rappelle que 2/3 de la variation globale est imputable aux retraites et décès (1) .

(1) entre 1962 et 1968 mais compte tenu des pyramides d'âges agricoles l'ordre de grandeur pour 1954 - 1962 est le même (voir taux de diminution de cohortes d'actifs agricoles de "L'espace Economique Française" (op - cit.))

1.22 - Evolution 1962 - 1968

On trouvera pages T 11. 29 à 34 les cartes et tableaux retraçant l'évolution comparée des départements avec celle de la France entière. Une étude au niveau régional a été effectuée par l'INSEE⁽¹⁾ mais rien, à notre connaissance n'a encore été publié comme étude au niveau départemental.

a) Evolution des effectifs départementaux

La dispersion des taux d'évolution est encore plus faible que dans la période précédente puisque 11 départements seulement sur 87 (hors Région Parisienne) ont un écart relatif, pris en valeur absolue, supérieur à 10 % (et toujours inférieur à 15 %), il y a donc une assez grande homogénéité spatiale du phénomène comparable à celle observée sur la période précédente, ce qui est d'autant plus remarquable que le rythme de décroissance s'est encore accru.

On constate d'une période sur l'autre, une certaine permanence de la zone à décroissance nettement moins forte que la moyenne mais celle-ci en retréci (Manche, Calvados, Orne, Côtes-du-Nord, Ile et Vilaine, Sarthe et Mayenne) et l'écart relatif n'est important que pour l'Orne et la Sarthe tandis que la Provence connaît le même phénomène, amorcé il est vrai pour deux départements au cours de la période précédente.

b) Evolution des taux d'emploi agricole

Il ne reste plus qu'une zone compacte importante de baisse nettement moins forte que la moyenne, du taux d'emploi agricole et qui se situe dans le massif central avec une extension vers les départements limitrophes de l'ouest et du sud. Or, à une ou deux exceptions près l'évolution en effectifs est analogue à celle de la moyenne de la France entière, quand elle ne lui est pas supérieure. C'est donc dans une évolution défavorable du nombre total d'actifs qu'il faut rechercher l'origine de cette disparité de taux.

Quant à l'ouest il ne subsiste que quelques départements où le taux d'emploi agricole est lui aussi anormalement élevé par rapport à la moyenne (Manche, Côtes-du-Nord, Orne, Mayenne).

Tous les départements fortement ruraux, c'est-à-dire ayant plus de 30 % d'actifs dans l'agriculture (alors que la moyenne française est de 15,6 %) ont, à l'exception des Landes et de la Vendée, une baisse de taux d'emploi agricole inférieure à la moyenne, ce qui révèle une grande inertie du phénomène.

(1) Voir étude de Muet "Mouvements inter-régionaux de population et passage de l'emploi agricole à l'emploi non agricole" publiée dans "Economie Rurale" n° 84 page 69 (1970).

Mais d'une façon générale, il semble que l'évolution des emplois agricoles ne soit dans la comparaison entre deux dates, que peu liée à celle de la population active ce qui donne à l'étude du taux d'emploi agricole davantage une valeur d'observation, que de recherche d'explication.

1.23 - Synthèse statistique

Comme dans l'étude du BTP, nous ferons appel à la technique de régression simple pour synthétiser les informations dont on dispose mais sans nous intéresser, pour des raisons qui découlent de ce qui vient d'être vu, au taux d'emploi agricole.

On obtient, dans le découpage 62 :

$$\begin{aligned} (\text{emploi agricole } 62) &= 0,771810. (\text{emploi agricole } 54) - 832 & R &= 0,992576 \\ & (77,02) \end{aligned}$$

$$\begin{aligned} (\text{emploi agricole } 68) &= 0,797885. (\text{emploi agricole } 62) - 456 & R &= 0,989276 \\ & (63,92) \end{aligned}$$

Si l'on exclut pour des raisons de modifications de découpage, la Région Parisienne on obtient

$$\begin{aligned} (\text{emploi agricole } 62) &= 0,772651. (\text{emploi agricole } 54) - 903 & R &= 0,992347 \\ & (74,99) \end{aligned}$$

$$\begin{aligned} (\text{emploi agricole } 68) &= 0,792057. (\text{emploi agricole } 62) + 13 & R &= 0,993354 \\ & (80,55) \end{aligned}$$

La faiblesse de l'emploi agricole dans la Région Parisienne explique que les coefficients de régression dans le nouveau découpage soient plus faibles que dans l'ancien. On peut remarquer que les coefficients de régression sont très voisins des ratios de variation des effectifs observés sur chacune des périodes, ce qui s'explique par la très faible dispersion que l'on a observé des évolutions départementales par rapport à l'évolution globale ; la constante étant faible par rapport aux grandeurs mises en jeu (respectivement 2 % et 1,2 % de la moyenne de la variable endogène dans les deux premières régressions).

SECTION 2 - RECHERCHE d'un MODELE d'EVOLUTION de L'EMPLOI AGRICOLE

Après avoir dégagé d'un point de vue descriptif quelques observations sur l'évolution de l'emploi agricole, il semble intéressant d'examiner quelques modèles opérationnels existants donnant des informations sur la répartition de l'emploi agricole dans l'espace même si, et c'est le cas général, cet emploi est envisagé comme variable d'ajustement plutôt que comme variable objectif. Nous introduirons ensuite un corps d'hypothèses à tester et dans une deuxième partie nous examinerons si ces hypothèses semblent confirmées ou infirmées avant de choisir le sous-modèle définitif.

2.1 - Les modèles mathématiques de répartition spatiale de l'emploi agricole

Les premiers économistes classiques, vivant dans une économie en quasi totalité rurale n'ont pas manqué en soulevant les problèmes de la valeur et des mécanismes régulateurs du marché, de rechercher les causes de fluctuation de l'emploi agricole dans des modèles conceptionnels à long terme. A cette approche globale succèdent les premiers travaux théoriques de localisation spatiale⁽¹⁾, où les problèmes agricoles servent largement de support à ces réflexions. Mais celles-ci pour aboutir à des éléments chiffrés devaient attendre l'existence de données suffisantes, la mise au point d'un arsenal de techniques statistiques et économiques élaborées et ce n'est que depuis les années 50 que des modèles opérationnels ont pu voir le jour⁽²⁾.

Les principaux d'entre eux (modèle de HEADY et EGLERT aux Etats-Unis⁽³⁾ et de BIROWO et REMBORG⁽⁴⁾ en Suède) sont des modèles statiques de simulation où l'emploi agricole est une variable d'ajustement, l'aspect le plus important étant l'adéquation des offres régionales à la demande nationale de production agricole. Une tentative de dynamisation de ce type de modèle est essayée aux Etats-Unis avec le modèle récursif du Département de l'Agriculture des Etats-Unis⁽⁵⁾, mais là encore c'est l'aspect de la production qui prime.

(1) voir chapitre 4 (2) voir le panorama qui en est fait par . Farhi et J Vercueil dans 2 recherches pour une planification cohérente : le modèle de prévision du Ministère de l'Agriculture" (CNRS 1969) p 82 et sq et les divers articles parus dans un rapport de l'OCDE intitulé "Programmation interrégionale en agriculture : problèmes méthodologiques" (OCDE 1963) (3) HEADY et EGLERT "Programming regional adjustments in grain productivity eliminate surpluses (journal of Farm Economics vol 41 - 1959) et "Modèles de programmation linéaire pour déterminer les systèmes régionaux de production dans l'agriculture des Etats-Unis cahier de l'ISEA - AG2 -n° 135 -1963) + rapport OCDE cité (4) Programmation internationale de la production agricole en Suède in rapport de l'OCDE cité (5) SCHALLER et DEAN "Predicting Regional Crop Production" et SCHALLER" A national Model of Agricultural Production Adjustment and Supply Response" USDA 1966.

Il semblerait donc que l'accent n'ait jamais été mis, jusqu'à présent, sur la prévision régionale de l'emploi agricole en tant que tel, dans les pays étrangers. Les travaux français ont eux aussi été faits dans la même optique. On examinera donc seulement ces derniers.

2.11 - Les travaux de KLATZMANN

C'est KLATZMANN⁽¹⁾ qui le premier en France s'est intéressé aux problèmes spatiaux de l'emploi agricole. Le but du modèle des "courbes de potentialités" qu'il a développé est de rechercher la répartition spatiale d'une production globale agricole en valeur entre zones, une fois défini l'objectif économique retenu (maximisation du revenu agricole moyen ou égalisation des revenus agricoles par région par exemple), l'ajustement se faisant par le niveau d'emploi agricole dans chaque région.

La résolution du problème est graphique et s'effectue par itération sur des courbes régionales de production et de revenu décrivant les aptitudes naturelles de ces régions où les variables "densité de production à l'hectare" et "revenu à l'hectare", ont fonction de la densité de travail à l'hectare. La difficulté de ce type d'approche consiste à définir des zones homogènes et à trouver assez de points pour définir les courbes, et l'intérêt de ce type d'approche est limité par l'absence de prise en considération des divers types de production agricole et de la diversité des avantages régionaux comparés pour un type de production agricole donné. Mais ces travaux précurseurs devaient sensibiliser des équipes de recherche à ce type de préoccupation.

2.12 - Le modèle agricole 1970 du Ministère de l'Agriculture

Une équipe de chercheurs gravitant autour du Service d'Etude et de Synthèse du Ministère de l'Agriculture a mis au point un modèle spatialisé⁽²⁾ statique décrivant une situation d'équilibre et dont le but était de mesurer les incidences prévisibles sur le plan de la production du revenu et de l'emploi agricole, de variation de la politique des prix agricoles et de celle des structures. Une situation agricole pour 1970 se distinguera d'une autre, dans le modèle agricole, dans les outputs dans chaque région d'une quinzaine de produits agricoles, par les prix de ces produits sur le marché national, par la population active employée dans les régions et son revenu, enfin par la répartition du sol dans chaque région entre quelques grands types d'exploitation⁽³⁾

(1) voir en particulier "La localisation des cultures et productions animales en France" (Imprimerie Nationale 1955) "La planification interrégionale dans l'agriculture, un modèle simplifié" (cahier de l'ISEA AG n° 1 sup 122 - 1962

(2) "Recherche pour une planification cohérente : le modèle de prévision du Ministère de l'Agriculture" par L. FARHI et J. VERCUEIL (Monographie du Centre d'économétrie - CNRS 1969)

(3) Rapport cité page 20.

Il ne peut être question de résumer en quelques lignes un modèle aussi complexe que celui-ci aussi se bornera-t-on à quelques remarques portant sur l'aspect spatialisé de la détermination de l'emploi agricole.

Ce modèle se décompose en sous-modèles, en interactions les uns, les autres, et qui tous se caractérisent par la définition des variables dans l'un des trois découpages emboîtés suivants : national, régional (21 régions de programmes) subrégional (100 zones).

- au niveau national sont définies les fonctions "quantités demandées - prix" pour 17 types de produits.
- au niveau régional est défini le volume des effectifs agricoles (salaire ou familial) en fonction du revenu perçu, ainsi que des fourchettes de l'importance relative dans la surface offerte d'une région, de classes d'exploitation, définies par leur taille (ou parfois par des considérations techniques).
- un sous-modèle au niveau subrégional détermine l'utilisation des terres par classe d'exploitation et système de production, donnant la réponse du système agricole à une demande nationale (quantités et prix fixés par produit), une offre régionale de travail agricole et certaines contraintes d'utilisation des sols. Ce sous-modèle fait appel à un programme linéaire où la fonction-objectif à maximiser est le bénéfice brut dégagé par l'ensemble de la profession agricole.

Trois ajustements sont simultanément mis en oeuvre et reposent sur l'hypothèse d'une recherche indépendante par chaque producteur de la maximisation de son profit brut. Le bénéfice brut dégagé va agir sur le volume d'agriculteurs disposés à rester dans chaque région et ce faisant sur les prix et donc sur la demande nationale et la répartition régionale de la production, des structures et systèmes de production des exploitations ; les contraintes du programme linéaire sont alors modifiées, d'où une nouvelle mise en oeuvre du programme linéaire, il s'en suit alors un nouveau bénéfice brut et une nouvelle répartition de l'utilisation des sols et il ne reste qu'à remettre en branle les trois ajustements jusqu'à obtention d'une solution stable. Le modèle sous certaines conditions converge vers une solution d'équilibre (statique) et permet de tester les effets sur l'emploi agricole et l'utilisation des terres de modifications de politiques des prix ou de structure.

2.2 - Recherche d'un modèle explicatif départemental pour 1962

Il n'est pas pensable ici de procéder à un modèle explicatif intégrant l'évolution de la demande et des modes d'utilisation des sols et des techniques utilisées mais on supposera que ces divers facteurs explicatifs ont une évolution et une importance relatives sensiblement stables à moyen terme et que ce que l'on peut mettre en évidence c'est un certain nombre de régularités qui expliquent de façon satisfaisante l'évolution de l'emploi agricole du moins à moyen terme⁽¹⁾.

Le problème reste alors de pouvoir endogénéiser certains de ces indicateurs dans un modèle définitif, problème qui ne se pose pas dans le cadre de l'exercice proposé.

2.21 - Hypothèses retenues pour test et leurs indicateurs

a) Influence des revenus des professions non agricoles

L'exode agricole est fondamentalement lié à un problème de revenu, sous le double aspect de son niveau absolu (minimum psycho-sociologique) et de son niveau relatif (comparaison avec la rémunération offerte par les autres activités), mais l'influence de ce facteur diffère selon le statut de l'agriculteur (exploitant propriétaire ou non, aide familial, salarié agricole) et se complique encore par le fait qu'un certain nombre d'agriculteurs ne réussissent à se maintenir dans des exploitations infra-marginales que grâce à l'appoint de la rémunération d'une activité complémentaire exercée à temps partiel à domicile ou non. On estime qu'un cinquième des chefs d'exploitation sont dans ce cas.

Si la majeure partie de l'exode agricole se fait dans une zone proche de l'exploitation, le salaire à prendre en compte est le salaire de la zone. Trois variables ont été testées : le salaire moyen, toutes catégories confondues, le salaire moyen ouvrier et le salaire moyen employé, ces données étant prises en début de période (1954).

(1) Si l'on pouvait tout intégrer correctement (les facteurs explicatifs de demande, données techniques etc.) on pourrait espérer obtenir un modèle prévisionnel à long terme d'effectifs (et non de variations)

(2) D'après MM. BRUN, LACOMBE et LAURENT dans l'article publié dans le Cahier de l'ISEA série AG n° 10 - 1972 et intitulé "La place des agriculteurs à temps partiel en France", les chiffres actuellement utilisables montrent que les exploitations dont le chef a une activité extérieure occupent une place croissante 18 % en 1963, 18,5 % en 1967 et 22,4 % en 1970.

b) Influence de structures d'accueil éloignées

L'exode rural s'est en grande partie effectué dans des zones éloignées de l'exploitation. Certains courants migratoires étaient favorisés par l'existence de "colonie" d'anciens agriculteurs implantée dans telle ou telle ville. Les données par catégorie socio-professionnelle n'existent pas, on a choisi de prendre un indicateur global qui est le pourcentage de gens nés dans un département et qui n'y résident pas. Il faut noter que si cet indicateur est rejeté par les méthodes utilisées l'hypothèse n'est pas pour autant confirmée par ce résultat car il est possible que la structure de l'indicateur s'éloigne assez de celle de la variable que l'on veut prendre en compte.

c) Influence de l'âge et du sexe

On a vu que la variation d'actifs agricoles est imputable en grande partie à des départs non compensés par des arrivées. On a donc cherché à lier cette variation aux effectifs de vieux agriculteurs en début de période. Comme l'exode féminin est plus fort que l'exode masculin les variables retenues pour le test ont été les actifs agricoles par sexe et pour 2 tranches quinquennales (60-64, 65-69) ainsi que les divers regroupements (soit 9 variables au total) en 1954.

d) Influence de la taille des exploitations

Il faut tout d'abord souligner que la notion de taille d'exploitation revêt un caractère différent selon le type d'utilisation faite du sol : c'est ainsi que 10 hectares de vignoble en Champagne permettent largement à un viticulteur de vivre alors que l'agriculteur de la Corrèze nanti de ces mêmes 10 hectares sera assez rapidement acculé à la faillite. Par contre, un certain nombre d'exploitations non viables parce que trop petites ne subsistent que parce que l'exploitant agricole exerce une activité complémentaire ailleurs. Une estimation faite sur la base du recensement agricole de 1970 donne 375 000 exploitations de complément (temps partiel ou retraite) et 520 000 exploitations marginales à faible revenu (production finale par exploitant inférieur à 35 000 F). On a donc testé comme variables explicatives, outre la taille moyenne des exploitations agricoles, le nombre d'exploitations agricoles par tranches de taille (taille < x, x étant croissant).

e) Influence de la taille du marché de l'emploi voisin

Les opportunités de trouver du travail étant plus fortes en cas de concentration urbaine il y a lieu de penser que l'exode agricole sera facilité dans les départements ayant une forte taille urbaine. Cette dernière variable est celle utilisée dans le sous modèle BTP : il s'agit de la moyenne des tailles d'agglomérations urbaines et de communes rurales, pondérée par l'importance relative de la population du département qui y vit.

f) Influence des actifs agricoles n'ayant pas le statut d'exploitant

L'exode rural ayant touché surtout les salariés agricoles et les aides familiaux, il a semblé intéressant de tester chacune de ces deux variables, ainsi que la variable "salariés agricoles, aides familiaux".

g) Accessibilité à un marché d'emplois peu qualifiés

L'enquête sur la mobilité du travail⁽¹⁾ a montré que 2/3 des agriculteurs mutants se dirigeaient vers des emplois manuels de faible qualification. On a fait l'hypothèse que l'importance des postes offerts était proportionnelle au nombre de ces postes déjà existants dans le département. Les variables prises en compte dans le test ont été le nombre de manoeuvres et le nombre d'ouvriers par département.

h) Importance des effectifs agricoles au départ

La variation relative du nombre d'agriculteurs n'est pas trop dispersée dans l'espace, aussi a-t-on testé l'incidence du volume des effectifs agricoles de l'année de départ sur leur variation au cours de la période. Toutefois, si cette variable était retenue, elle ne pourrait l'être simultanément avec la variable "salariés agricoles + aides familiaux" qui constitue 63 % des actifs agricoles en 1954.

Examinons maintenant ce que donnent les diverses hypothèses à l'épreuve des faits.

2.12 - Recherche d'un modèle en valeur absolue

On a donc soumis les quelque trente variables retenues à l'épreuve du programme de régression optimale, comme on l'avait fait pour la recherche du sous-modèle BTP, se réservant de modifier le corps des variables retenues si certaines variables logiquement incompatibles étaient retenues simultanément⁽²⁾.

Il n'existe pas de régression optimale ayant plus de trois variables significatives comme le montre le tableau ci-après.

(1) Bâtiment (19 %), manoeuvre (30 %), conducteurs (14 %) in "Mobilité professionnelle en France entre 1959 et 1964" - Etudes et Conjoncture (act. 1966)

(2) Cas des actifs agricoles par âge et sexe ainsi que les exploitations par tranche de taille.

RECHERCHE d'un MODELE EXPLICATIF des VARIATIONS DEPARTEMENTALES d'EMPLOIS AGRICOLES entre 1954 et 1962

Nombre de variables retenues dans la régression optimale	coef. de variables de détermination	Aides familiaux + salarié agricoles en 1954	Actifs agricoles de 65-69 ans de sexe féminin en 1954	Actifs agricoles de 65-69 ans en 1954	Nb d'exploitations de moins de 5 ha en 1955	Autres variables	Constante
2	0,91785	- 0,19004 (- 10,24)	- 5,76696 (- 9,78)				- 1559
3	0,93131	- 0,19189 (- 11,24)		- 3,20390 (- 11,39)	+ 0,20433 (+ 4,19)		+ 364

Ce résultat est, bien sûr, assez décevant car il infirme la plupart des hypothèses formulées ou plus exactement si celles-ci sont fondées les variables chargées de les traduire déforment par trop la réalité.

Cela étant, les deux modèles possibles peuvent être considérés comme assez satisfaisants avec leur coefficient de détermination supérieur à 0,9. On y trouve dans les deux cas le rôle de l'âge mis en évidence par les études longitudinales de population agricole qui montraient le rôle prépondérant des retraites et décès, ainsi que le poids souligné depuis longtemps des agriculteurs non exploitants dans la variation globale d'actifs agricoles. Quant aux exploitations de moins de 5 hectares leur coefficient de régression est positif c'est-à-dire que cette variable joue un rôle de rétention. Ce phénomène est très certainement lié au double revenu perçu par les propriétaires de ces exploitations inframarginales (à l'exception de certaines exploitations viticoles ou de poly-culture intensive).

Examinons maintenant l'approche en taux.

213 - Recherche d'un modèle en taux

Il a fallu tout d'abord définir des dénominateurs pour chacune des variables représentatives des diverses hypothèses formulées :

a) choix des dénominateurs

Toutes les données relatives à la population active agricole totale ou par statut, et en particulier la variable endogène, ont été rapportées à la population totale malgré les inconvénients résultant des remarques faites dans la première section. En fait il n'y avait guère d'autre choix possibles car la division par un scalaire tel que la population agricole totale de la France entière revenait à se ramener au problème précédent. On peut ajouter que le quotient de population active agricole par la population totale du département n'a vraiment de signification que dans le cadre d'une économie agricole de type autarcique.

Lorsque les actifs agricoles sont repérés par âge et sexe, le dénominateur retenu a été la population active agricole par sexe (ou totale pour les données de la marge).

Les données d'emploi non agricole ont été rapportées à l'emploi départemental non agricole.

Les données relatives aux strates d'exploitation ont été rapportées logiquement au nombre total d'exploitations.

Les données par nature non sommables comme les salaires moyens et le ratio de personnes ne vivant pas dans le département où elles sont nées par département de naissance ont été laissées telles quelles.

Examinons maintenant les résultats obtenus.

b) étude des résultats

Comme précédemment il n'existe pas de régression optimale intégrant plus de trois variables significatives comme le montre le tableau suivant :

RECHERCHE d'un MODELE EXPLICATIF des VARIATIONS RELATIVES DEPARTEMENTALES de l'EMPLOI AGRICOLE en 1954 et 1962

Nombre de variables retenues dans la régression optimale	variables coef. de détermination	(salariés agricoles + aides familiaux)/ (population active agricole)	(agriculteurs hommes de 60-69 ans)/popu- lation active agr.)	(agriculteurs 65-69 ans)/ (population active agricole)	(aides familiaux)/ (population active agricole)	(salariés agricoles) (population active agricole)	Autres variables	Constante
2	0,69980	- 0,63251 (-11,46504)	- 6,10122 (- 11,69)					+ 0,49227
3	0,77184			(- 5,12939) (- 11,60)	- 0,78485 (- 14,60)	- 0,61160 (- 12,04)		+ 0,51477

Ces régressions sont moins satisfaisantes dans la mesure où les coefficients de détermination sont nettement moins élevés que dans l'approche en effectifs, mais comme on l'a souligné au chapitre 10 cette remarque n'implique pas qu'il faille automatiquement préférer l'approche en effectifs. Par ailleurs comme précédemment l'effet statut et l'effet âge interviennent et ce qui frappe c'est que la stratification en salariés agricoles aides familiaux apporte un gain d'explication appréciable par rapport à la non stratification.

2.24 - Modèle retenu

La démarche suivie est identique à celle suivie jusqu'à maintenant.

...

a) choix du modèle

Dans un premier temps, on a choisi de comparer les meilleurs modèles des approches en taux et en effectifs choisis respectivement d'après le critère du plus fort coefficient de détermination. Il est vrai qu'aucun modèle n'avait plus de trois variables explicatives et que le passage de deux à trois variables amenait un accroissement non négligeable du coefficient de détermination ; la conjonction de ces deux faits limitait singulièrement les choix possibles.

La comparaison de la variance résiduelle du modèle en taux après transformation en effectifs et de celle du modèle en effectifs mène sans conteste à retenir ce dernier. Toutefois le poids très important dans l'explication de la régression, des actifs agricoles de 65 à 69 ans et le fait que cette génération (1911 - 1916) correspondait d'une certaine façon à un "accident démographique" dans la pyramide des âges des actifs agricoles, on a été amené à retenir, pour diminuer l'importance de ce biais, une régression ayant pratiquement le même coefficient de détermination, mais non optimale par rapport à l'ensemble des variables explicatives testées. Celle-ci présente comme seule différence le choix de la tranche d'âge des 60-69 ans à la place de celle des 65 - 69 ans. Ce modèle en effectifs finalement retenu, dont toutes les variables sont évidemment significatives, ne met pas en cause le rejet de l'approche en taux ; en effet, la variance résiduelle, bien qu'un peu plus forte que dans le modèle optimal, reste très sensiblement inférieure à celle obtenue avec le meilleur modèle en taux transformé en effectifs.

L'équation du sous-modèle agricole retenu est la suivante :

$$\begin{aligned} (\text{POP ACT (Agriculture 62)} - \text{POP ACT (Agriculture 54)}) = & - 0,139517 \cdot \text{POP ACT (Agriculture, aides familiaux + salariés 54)} \\ & - 1,68155 \cdot \text{POP ACT (Agriculture, 60-69 ans, 54)} \\ & + 0,238872 \cdot \text{EXPLOITATIONS (< 5 hectares, 55)} \\ & + 367 \end{aligned}$$

On peut noter qu'avec un écart type de 474, la constante n'est pas significativement différente de zéro.

b) analyse des résultats

On trouvera les tableaux et cartes de résultats pages T 11. 37 à 41 et A 14.

Si l'on écarte deux départements bien particuliers la Seine et le Territoire de Belfort, l'erreur relative commise par ce modèle sur la variation 54-62 s'étend de - 31 % à 30 % avec une moyenne de ces erreurs relatives prises en valeur absolue de 13 %.

Les erreurs relatives par rapport aux effectifs 1962 calculés par le modèle ne dépassent que dans 4 cas 10 %, alors qu'il y a 9 départements dans ce cas dans la comparaison des effectifs 1962 avec ceux calculés homothétiques de 1954. La comparaison des cartes obtenues dans ces deux approches montrent qu'en gros celles-ci semblent également satisfaisantes. Cela signifie que si les relations mises en évidence dans le modèle sont stables, une approximation correcte des effectifs agricoles 1968 peut-être obtenue sans passer par un modèle national et qu'au niveau départemental on peut espérer obtenir peu d'erreurs relatives importantes.

On peut faire en outre la même remarque que celle faite pour le sous-modèle BTP en ce qui concerne l'approche probabiliste qui là aussi semble condamnée.

2.3 - Analyse des résultats 1968 du sous-modèle

Examinons tout d'abord les résultats obtenus avant de faire la critique du sous-modèle.

2.31 - Les résultats obtenus

L'analyse des résultats des tableaux et cartes des pages T 11.42 à 44 permet de noter plusieurs faits intéressants.

Globalement tout d'abord la diminution d'actifs agricoles (hors Région Parisienne) est surestimée d'environ 85 000 individus, c'est-à-dire d'un peu plus de 10 % ce qui n'est pas mauvais car cela représente une erreur relative sur la prévision d'effectifs agricoles en 1968 d'environ 2,5 %.

Si l'on rentre un peu dans le détail deux faits sont saillants. Seulement trois départements (si l'on excepte le Rhône dont les délimitations ont changé pour 1968) ont une erreur de prévision supérieure à 10 % - et plus de la moitié des départements ont une erreur relative de prévision inférieure à 5 %. Le sous modèle a tendance à sous estimer partout (sauf dans 15 départements) les effectifs agricoles ce qui signifie que ce phénomène souligné sur le plan global se répartit de façon assez homogène dans l'espace.

Enfin on peut ajouter que l'utilisation des coefficients de spécificités se traduit par un accroissement de l'écart type résiduel, ce qui est contraire à ce que l'on pouvait espérer, tandis que le cadrage sur l'observé hors région parisienne n'améliore sensiblement que l'approche qui n'utilise pas ces coefficients.

2.32 - Critique du sous-modèle agricole retenu

Ce modèle n'est pas un modèle explicatif dans la mesure où l'interprétation des liaisons obtenues statistiquement ne permet pas de mettre en évidence des mécanismes économiques. Il s'agit plutôt ici de "régularités observées" puisque la variation d'actifs agricoles est avant tout fonction de données structurelles de l'année de départ. Le seul mécanisme sous-jacent postulé est celui des revenus d'appoint mais le poids de la variable explicative "exploitations de moins de 5 hectares" est très faible dans la régression obtenue.

Dans ce cadre, les liaisons mises en évidence ne peuvent être considérées comme bonnes que pour autant que la distorsion des structures n'est pas trop forte, ce qui n'est pas le cas dans l'évolution de l'emploi agricole. Il s'en suit que ce type de modèle ne peut être utilisé que pour des "prévisions" à moyen terme.

Rappelons toutefois que la démarche suivie ne visait pas à proposer un modèle explicatif satisfaisant du point de vue théorique (ce que pourrait être dans une version plus orientée vers l'emploi, le modèle 1970 du Ministère de l'Agriculture) mais plutôt d'essayer de trouver des constances résultant de la stabilité d'un certain nombre de mécanismes démographiques ou économiques.

Les difficultés rencontrées dans la prévision au niveau global de l'emploi agricole à long terme ne permettent pas de penser que pour le problème encore plus délicat de la spatialisation de cette prévision on puisse dans un avenir proche disposer de modèle satisfaisant.

Le sous-modèle tertiaire

12

CHAPITRE 12 - L'ETUDE DU SOUS-MODELE TERTIAIRE

Contrairement à une idée communément répandue, le secteur tertiaire, est chronologiquement le second à voir le jour. A Piatier⁽¹⁾ souligne en effet que dans les premières civilisations "lorsque pour une raison quelconque, la production par tête dépasse ce besoin minimum, le surplus de production sert à entretenir, puisque la population se développe, des personnes qui peuvent être employées ailleurs que dans la "production élémentaire" : personnel d'organisation du groupe (au sens large) : chefs politiques et religieux, personnel militaire et administratif, producteurs de services et commerçants". Ce n'est qu'avec le progrès technique que le secteur secondaire "décolle" vers la fin du 18^e siècle dans les pays européens.

A la révolution industrielle a succédé la révolution tertiaire et ce secteur est en passe de devenir le secteur prépondérant dans la population active de tous les pays pour lesquels la "première révolution" s'est effectuée au siècle dernier.

Ce phénomène, lié à l'évolution urbaine, est à l'origine de certains travaux théoriques sur la répartition des hommes⁽²⁾, mais la notion de tertiaire, en tant que telle, n'a jamais beaucoup intéressé les économistes régionaux. Aussi semble-t-il préférable de présenter très succinctement quelques uns des travaux théoriques récents relatifs à la notion de tertiaire pour guider le choix à effectuer dans la définition opérationnelle du ou des secteurs tertiaires dont on se servira.

La démarche s'effectuera donc en 3 étapes : une approche théorique de la notion de tertiaire, une approche descriptive de l'évolution du tertiaire d'après trois derniers recensements et enfin le test de quelques modèles.

(1) A. Piatier - "Les inégalités du développement démographique et économique" in Encyclopédie française, tome 11 p. 11. 14. 13

(2) Voir chapitre 4.

SECTION 1 - APPROCHE THEORIQUE DE LA NOTION DE TERTIAIRE

La répartition de l'activité humaine en trois catégories remonte à l'antiquité puisqu'on la retrouve déjà chez Platon et Aristote⁽¹⁾ puis sous une forme déjà moderne chez Antoine de Montchrestien. Chez les économistes classiques la notion de production de service, si elle est reconnue en tant que catégorie, reste longtemps (en particulier, chez les physiocrates, Smith, Say, Saint-Simon) le fruit d'un travail improductif et ce n'est qu'à partir de la seconde moitié du 19^{ème} siècle (avec Stuart Mill, Bastiat, Dunoyer en particulier) qu'elle obtient ses lettres de noblesse.

La formulation moderne des trois secteurs remonte à Fisher⁽²⁾ qui le premier utilise le terme de secteur tertiaire et le définit comme regroupant toutes les activités consacrées à la production de biens immatériels, puis Colin Clark caractérise ce secteur par sa faible productivité et Fourastié⁽⁴⁾ donne une définition dynamique des trois secteurs, non en fonction de leur objet, mais suivant leur taux d'accroissement de la productivité (rapproché d'une élasticité de consommation variant en sens inverse), le secteur tertiaire se caractérisant par un progrès technique faible.

Lengellé⁽⁵⁾ en s'inspirant de la comptabilité analytique propose 4 définitions théoriques du tertiaire car la définition résiduelle utilisée opérationnellement par les économistes présente l'inconvénient d'une très grande hétérogénéité et les comparaisons temporelles ou internationales du tertiaire n'ont guère de sens si elles sont faites au seul niveau global.

L'accord sur la définition même de tertiaire est donc loin d'être réalisé et trouve d'ailleurs son pendant en comptabilité nationale où se pose le problème de l'inclusion à la production (au sens large) des services.

(1) Voir le livre de Maurice Lengellé : "La révolution tertiaire" (Genin 1966)

(2) Fisher "The Clash of Progress and Security" 1935

(3) Colin Clark : "Les conditions du progrès économique" 3^e édition traduction Française 1957

(4) J. Fourastié : Le grand espoir du XX^{ème} siècle PUF (1948)

(5) ouvrage cité.

Il s'avère (comme Colin Clark l'avait déjà noté) que dans les pays industrialisés, le taux de tertiaire est en règle générale d'autant plus important que le revenu par tête est élevé(1), mais la dispersion n'est pas négligeable et à deux pays très en avance sur la France du point de vue du revenu par tête : les Etats-Unis (1er) et la Suède (3ème), correspondent deux taux très différents de tertiaire. L'engellé après étude des tendances récentes pense que la France a choisi de suivre la voie des Etats Unis plutôt que celle de la Suède.

Aussi semble t-il intéressant d'examiner les analyses faites par V.R. FUSCHS sur les causes d'évolution du tertiaire aux Etats-Unis.

L'étude de FUSCHS porte sur la période 1929 - 1965 ; celui-ci note qu'aux Etats-Unis la part du tertiaire dans le PNB, exprimé en dollars constants, reste la même ($\approx 48\%$), au cours de la période alors que la population active du tertiaire passait de $40,4\%$ à $54,8\%$. Pour essayer d'expliquer ce déplacement, FUSCHS a testé trois hypothèses fréquemment avancées :

1 - Ce déplacement au profit du tertiaire est imputable à une croissance très rapide de la demande finale de service, faisant appel à l'évolution de la structure de consommation sous l'influence de la croissance du revenu (Lois d'Engel). Des mesures d'élasticités - revenu de la consommation finale de tertiaire sur les séries temporelles portant sur 20 ans, ou en coupe synchronique, sont concordantes et montrent que celles-ci sont voisines de 1 et de très peu supérieures (de 5 à 10 %) aux élasticités de consommation des biens non alimentaires; mais la hausse plus forte des prix du service a contre-balancé l'effet revenu, ce qui a permis à la part du tertiaire dans le PNB d'être constante. Cette hypothèse ne peut donc être retenue.

2 - Ce déplacement est dû à un accroissement relatif plus fort de la demande intermédiaire de service, résultant en quelque sorte d'une division accrue du travail. L'examen des tableaux d'échange interindustriels de 1947 et 1958, montre effectivement une tendance dans ce sens, mais on ne peut guère imputer plus de 10 % du changement à cet effet.

(1) V. R. FUSCHS, dans "The Service Economy" (NBER - 1968) sur des données récentes portant sur 20 pays de l'OCDE, trouve (P. 37) un ajustement satisfaisant sur une portion courbe de type $y = -\frac{a}{x} + b$, définie dans le quadrant supérieur droit où y est le taux de tertiaire et x le revenu par tête.

3 - Enfin la dernière hypothèse attribue ce déplacement à une croissance plus lente de la productivité des services que celle des autres secteurs(1). Il ressort en fait des séries 1929 - 65 une croissance annuelle de la productivité de 3,4 % pour l'agriculture, de 2,2 % pour l'industrie et de 1,1 % pour les services. Cet écart de croissance entre l'industrie et le tertiaire s'explique pour plus de moitié par 3 causes : baisse relative plus grande dans le tertiaire de la durée de travail (0,1) augmentation relative plus grande dans l'industrie de la qualité de travail (0,5 %) ; et enfin une croissance relative plus grande dans l'industrie du capital par tête (incidence non chiffrée, les données n'étant pas assez sûres).

(1) Ce qui suppose en outre (ce que FUSCHS ne souligne pas) que la demande s'est aussi constamment exercée dans le sens d'un maintien de la part des services dans le PNB. En effet la croissance de la productivité est une condition permissive de la satisfaction d'une demande croissante de services, compte tenu d'une croissance techniquement possible de population active travaillant dans ce secteur et d'un niveau de prix donné. C'est d'ailleurs un aspect que soulignait déjà Fourastié (opus cité).

SECTION 2 - DESCRIPTION DE L'EVOLUTION DU TERTIAIRE EN FRANCE

Pour cette description et sans préjuger de la définition qui sera retenue finalement dans le sous-modèle tertiaire, on peut prendre dans une première approche la définition retenue par l'INSEE et qui porte sur les catégories d'activité économique (CAE) 61 à 94 (voir signification des codes dans le tableau page 8).

Cette définition repose sur le concept d'activité collective c'est-à-dire que toutes les personnes concourant à la fourniture de services par un établissement (privé ou administratif) sont considérées comme appartenant à l'activité collective tertiaire, quel que soit le métier individuel.

2.1 - L'évolution du tertiaire sur longue période et comparaisons internationales

De nombreux ouvrages retracent l'évolution du tertiaire en France⁽¹⁾, et malgré un léger biais, introduit par une modification de définition de la population active à partir de 1954, on peut estimer qu'en cent ans la part du tertiaire dans la population active a doublé passant de 23,1 % en 1876 à 44,7 % en 1968, mais que près de la moitié de cette évolution s'est effectuée entre 1954 et 1968⁽²⁾. Il faut toutefois noter de profonds bouleversements dans la hiérarchie des diverses activités du tertiaire⁽³⁾ sous l'influence notamment d'une transformation des fonctions de l'Etat (Sécurité Sociale, PTT,...) d'un accroissement des fonctions de distribution liées à la croissance importante des consommations intermédiaires (division du travail) ou finales et d'un changement important du niveau de vie (congés payés, travail de la femme) facilité par le progrès technique (substitution partielle de l'électro ménager aux services domestiques..). En outre on peut remarquer⁽⁴⁾ que les villes sont en moyenne d'autant plus tertiaires qu'elles ont une population importante.

(1) voir en particulier L. Cahen : "migrations professionnelles, données statistiques sur l'évolution en divers pays" - cahier de l'INED n° 31 (1957) ; "Bleton ; les hommes des temps qui viennent, essai sur les classes moyennes" (Editions ouvrières 1957) ; Lengellé ouvrage cité ; Fusch ouvrage cité ; M. Praderie : "ni ouvrier, ni paysans : les tertiaires" (Le Seuil - 1968)

(2) La part du tertiaire a augmenté de 8 points (unité 100) entre 1954 et 1968

(3) voir en particulier les très intéressants tableaux publiés dans l'ouvrage cité de Bleton (p72 et 74).

(4) Prost : "la hiérarchie des villes en fonction de leurs activités de commerce et de service" (Gauthier Villars 1965).

Sur le plan des comparaisons internationales les données réunies par l'OCDE⁽¹⁾ permettent de situer la France, dans la mesure où les informations sont comparables par rapport aux principaux pays développés ; en 1966 l'estimation de la part du tertiaire dans la population active était de 41,6 pour la France, de 60,4 pour les USA, de 49,9 pour la Grande Bretagne, de 48,7 pour les Pays-Bas, de 39,9 pour la R.F.A. et de 34,4 pour l'Italie.

2.2 - Analyse de l'évolution récente (1954 - 1968) du tertiaire par catégorie d'activité et département

Cette étude sera menée en 2 temps. On s'efforcera tout d'abord d'avoir une vision détaillée de l'évolution en restant sur un plan exclusivement descriptif. Dans une deuxième phase un essai d'interprétation commencera à être effectué en cherchant à tester quelques hypothèses simples ou en reprenant l'étude descriptive de l'évolution de regroupements en fonction de critères adoptés par certains des auteurs qui se sont penchés sur l'étude du tertiaire en France.

2.21 - Description de l'évolution du tertiaire

Cette description se fera en 3 étapes : tout d'abord une analyse France Entière portant sur chacune des catégories d'activité, puis une analyse départementale de l'évolution globale du tertiaire et enfin une tentative de synthèse de ces 2 approches.

2.211 - L'évolution France Entière

Cette étude se fera à l'aide des tableaux 1 et 2 qui fournissent les effectifs absolus et la part dans la population active totale de chacune des 18 catégories d'activité économique (CAE) constitutives du tertiaire, puis l'importance de chacune d'entre elles dans la population active dans l'évolution du tertiaire ainsi que leurs taux d'évolution absolu et relatif (par rapport aux taux d'évolution moyen du tertiaire) pour chacune des périodes.

(1) Statistiques de la population active - OCDE 1968.

TABLEAU 1

Catégorie d'Activité économique Détaillée (nomenclature INSEE)	1954			1962			1968		
	effectifs ⁽¹⁾	moyenne ⁽²⁾	écart type ⁽²⁾	effectifs	moyenne	écart-type	effectifs	moyenne	écart-type
61 Transports terrestres	648 866	7 290	14 180	649 621	7 299	14 777	708 424	6 606	7 852
62 Tran. flu. mar. aériens	125 089	1 405	3 721	149 804	1 683	4 590	145 796	1 300	3 436
71 Com. Agric. alimentaires	224 877	2 526	2 964	255 899	2 875	3 549	277 956	2 855	2 270
72 Com. agr. alim. détail	573 825	6 447	9 251	593 780	6 671	9 783	611 072	6 065	5 292
73 hôtellerie débit boisson	536 547	6 028	11 273	510 671	5 737	11 304	525 880	5 289	6 646
74 autres commerces gros	218 541	2 455	6 863	303 929	3 414	8 282	350 220	3 235	4 572
75 autres commerces de détail	644 529	7 241	13 753	801 948	9 010	17 069	979 032	9 592	10 199
76 commerces mal désignés	75 099	843	5 771	18 358	206	639	37 928	376	1 872
77 Interm. du com. industriel	105 112	1 181	3 584	99 970	1 123	4 008	114 224	995	2 470
78 banques et assurances	244 970	2 752	10 362	313 354	3 520	12 827	406 864	3 431	67 73
79 spectacles sédentaires	43 878	493	2 564	39 907	448	2 068	37 640	323	11 26
81 services aux entreprises	95 402	1 071	4 165	150 721	1 693	7 186	226 564	1 880	5 249
82 services domestiques	573 612	6 445	12 147	533 647	5 995	11 834	496 476	5 004	7 490
83 services rend. aux part.	1 031 511	11 589	25 689	1 240 917	13 942	30 020	1 566 148	15 288	20 637
91 eau, gaz, électricité	136 787	1 536	3 095	156 551	1 758	3 514	184 704	1 622	1 737
92 transmissions et radios	245 962	2 763	5 592	301 361	3 386	7 440	359 140	3 412	4 779
93 administrations publiques	970 241	10 901	21 760	1 220 313	13 711	24 846	1 581 208	15 160	15 410
94 défense nationale	275 908	3 100	5 440	369 858	4 155	5 981	317 932	3 129	3 093
ENSEMBLE DU TERTIAIRE	6 770 206	76 075	158 433	7 710 609	86 635	176 195	8 927 208	85 570	104 682

TABLEAU 2

	Part de la CAE dans la Pop Active						Taux d'évolution ⁽⁴⁾			Part dans l'évolution		
	%	1954	rang	%	1962	rang	%	1968	rang	54 - 62	%	62 - 68
83 Serv. rendus aux Part.	54	1	65	1	78	2	20	26	52	223	267	248
93 Adm. publiques	51	2	64	2	79	1	25	29	63	265	295	283
75 autres com. de détail	34	4	41	3	49	3	24	22	52	167	150	155
61 transports terrestres	34	3	34	4	35	4	0	9	9	1	48	28
72 com. Agr. alim. détail	30	5	31	5	30	5	3	2	6	21	14	17
82 services domestiques	30	6	28	6	24	7	-6	-6	-14	-42	-30	-36
73 hôtellerie débit boissons	28	7	26	7	26	6	-4	2	-2	-28	12	-5
94 Défense Nationale (3)	14	8	19	8	15	11	34	-14	15	100	-42	+19
78 banques et assurances	13	10	16	9	20	8	27	29	66	73	76	75
74 autres commerces gros	11	12	15	10	17	9	39	15	60	91	38	61
92 transmissions et radios	13	9	15	11	17	10	22	19	46	59	47	52
71 Commerces agric. alim.	11	11	13	12	13	12	13	8	24	33	18	25
91 Eau, gaz, électricité	7	13	8	13	9	14	14	17	35	21	23	22
81 services aux entreprises	5	16	7	14	11	13	57	50	137	59	62	61
62 transp fluv, maritimes, aériens	6	14	7	15	7	15	19	-2	17	26	-3	10
77 Interm. du commerce industriel	5	15	5	16	5	16	-4	14	9	-5	12	4
79 spectacles sédentaires	2	18	2	17	1	17	-9	-5	-14	-4	-2	-3
76 commerces mal désignés	3	17	0	18	1	18	-75	108	-49	-60	16	-17
ENSEMBLE DU TERTIAIRE	360	-	405	-	447	-	13	15	32	-	-	-

(1) Les effectifs sont ceux observés sur l'ensemble du territoire métropolitain (Corse exclue en 1954 et 1962). Les effectifs 1954 ont été redressés par l'intégration dans les diverses CAE des "non déclaré" et ce au prorata de la CAE dans les actifs du département.

(2) Les moyennes et écarts type sont calculés pour les trois années sur les départements 1 à 19 et 21 à 90, ce qui exclut la Corse et les nouveaux départements de la Région Parisienne, de ce fait les données en valeur absolue ne sont pas exactement comparables en raison des modifications territoriales des départements 75 et 78, ainsi d'ailleurs que celles intervenues sur les départements 38, 42 et 69.

(3) 100.000 militaires de carrière ont été rapatriés en métropole entre 54 et 62.

(4) On rappelle que le taux d'évolution de la population totale entre 1954 et 1962 est de 8 % et qu'entre 1962 et 1968 il a été de 7 %.

Entre 1954 et 1962, alors que la population active est restée stable (une population active en augmentation de seulement 150.000), le taux d'évolution du tertiaire a été de + 13 % et entre 1962 et 1968, avec une population active en augmentation (d'environ 900.000 personnes), ce taux sur une période plus courte de 2 ans s'est encore accru pour passer à 15 %. La montée du tertiaire est donc assez impressionnante.

Cette évolution globale masque des disparités importantes puisque sur 18 des catégories, on n'en trouve que 9 ayant eu sur les 2 périodes une croissance supérieure à celle de la population et 11 à avoir été en croissance constante sur la période, comme le montre le tableau synoptique suivant.

TABLEAU 3

Type d'évolution de la CAE	Evolution de même type pour 54 - 62 et 62 - 68	1954 - 1962	1962 - 1968	
Croissance supérieure à celle de la population	83 Services rendus aux particuliers	94 Défense Nationale	77 intermédiaires du commerce et de l'industrie	
	93 Administrations publiques			
	75 Autres commerces de détail	62 Transports fluviaux maritimes et aériens	76 Commerces mal désignés	
	78 Banques et assurances			
	74 Autres commerces de gros			
	92 Transmissions et radios			
	71 Commerces agricoles et alimentaires (gros)			61 Transports terrestres
	91 Eau - gaz - électricité			
81 Services rendus aux entreprises				
Croissance inférieure à celle de la population	72 Commerces agricoles et alimentaires (détail)	61 Transports terrestres	73 Hotellerie - débits de boisson	
Décroissance	82 Services domestiques	77 Intermédiaires du commerce et de l'industrie	94 Défense Nationale	
	79 Spectacles sédentaires	76 commerces mal désignés	62 transports fluviaux maritime et aériens	
		73 Hotellerie débits de bois		

Encore faut il noter que seules 5 catégories ont en 1968 un effectif inférieur à celui de 1962 ; il s'agit, par ordre d'importance décroissante en 1962 : des services domestiques (82), de l'hôtellerie - débits de boisson (73), des transports fluviaux - maritimes et aériens (62), des spectacles sédentaires (79) et des commerces mal désignés (76).

On peut remarquer en outre que, entre 1954 et 1968 :

- les administrations publiques (93) prennent la première place au détriment des "services aux particuliers" (83) qui passent en 2ème position.
- les "autres commerces de détail" (75) prennent la 3ème place aux "transports terrestres" (61)
- les "services rendus aux entreprises" (81) gagnent 3 places
- le secteur bancaire (78) et les "autres commerces de gros" (74) gagnent 2 places

Ces 6 catégories représentent 20,8 % de la population active en 1954 et 25,5 % en 1968 ; près de 90 % de l'évolution leur est imputable, (87,9 % entre 1952 et 1962 et 88,8 % entre 1962 et 1968) et plus de 50 % aux seuls administrations et services aux particuliers (93 et 83).

2.212 - L'évolution globale du tertiaire au niveau départemental

L'évolution du tertiaire est loin d'être semblable dans tous les départements, ce qui est d'ailleurs normal dans la mesure où les situations sont très différentes selon les départements. Examinons brièvement les situations départ avant d'étudier les évolutions (voir pages T12. 55 à 60 et A 15).

a) Etude des disparités départementales des situations de départ :

La disparité en terme de taux de tertiaire par rapport à la population totale est très grande puisque lorsque l'on prend 100 comme valeur France entière, l'éventail, en excluant la Seine⁽¹⁾ s'étend en 1954 de 55 à 145 en 1962 : 62 à 134 et en 1968 de 66 à 126, et qu'un raisonnement en taux par rapport à la population active ne

(1) L'indice pour la Seine est pour les 3 derniers recensements pour le taux par rapport à la population totale respectivement de 180, 174 et 189 et pour le taux par rapport à la population active respectivement de 153, 140 et 150 mais rappelons qu'un biais dans les comparaisons est introduit par la modification des délimitations de ce département entre les 2 derniers recensements.

modifie pas grand chose (éventail 1954 : 51 - 156 ; éventail 1962 : 57 - 146 ; éventail 1968 : 67 - 134). Il est remarquable de constater que l'écart entre les extrêmes (il s'agit d'ailleurs toujours des mêmes départements le Gers et les Alpes Maritimes) s'est resserré de façon très sensible sur une période relativement courte et ce malgré une très forte croissance de tertiaire pour l'ensemble de la France. On étudiera ultérieurement (point b) ce resserrement des disparités spatiales du tertiaire.

La courbe de distribution des taux de tertiaire est fortement dissymétrique à droite puisque seulement 13 départements en 1962 et 18 en 1968 ont un taux de tertiaire par rapport à la population totale supérieur au taux moyen national et que pour le taux de tertiaire par rapport à la population active, le nombre de ces départements est de 17 en 1954 et de 18 en 1968.

On étudiera au 2.213 la spécialisation de ces départements privilégiés. Enfin, on constate que le tertiaire au niveau départemental n'est pas systématiquement lié à l'urbanisation. car sont absents de ce tableau plusieurs départements ayant une agglomération de plus de 150.000 habitants. Il s'agit par ordre décroissant de Lille, Nantes, St-Etienne, Sens, Grenoble, Valenciennes, Mulhouse, Clermont-Ferrand, Rennes et Tours.

Plusieurs cartes ont été dressées (T 12 55 à 58) donnant la situation du tertiaire à chacun des recensements, en terme de coefficient de spécificité des taux de tertiaire, par rapport à la population totale et par rapport à la population active. On y constate que les différences entre cartes en taux par rapport à la population active ou par rapport à la population totale sont minimales (ce qui est d'ailleurs assez normal, la distribution des taux d'activité étant très resserrée). On raisonnera donc ultérieurement indifféremment sur l'un ou l'autre critère.

La situation de départ se caractérise par un très petit nombre de départements à taux supérieur à la moyenne, avec une concentration sur la côte méditerranéenne, une zone très marquée de sous-tertiarisation accentuée qui comprend le Massif Central et s'étend vers le sud-ouest, et dans l'ouest quelques îlots de sous-tertiarisation accentuée (Mayenne, Manche, Morbihan, Vendée, Deux-Sèvres). L'évolution de la situation se caractérise par un fort resserrement de disparité et une disparition progressive des zones sous-tertiarisées.

b) Etude de l'évolution

L'étude de l'évolution du département peut être menée, en valeur absolue ou en taux (par rapport à la population totale ou active) comparativement à l'évolution nationale⁽¹⁾ et avec ou sans référence à la situation de départ. La combinaison de ces deux types d'approche est indispensable pour nuancer les conclusions que l'on peut

(1) On prendra par exemple la moyenne arithmétique des effectifs départementaux comme référence si l'on désire travailler en valeur absolue.

tirer des observations. Encore faut-il pouvoir disposer d'une méthode systématisant la référence de l'évolution à la situation de départ. On en proposera une avant de passer à l'analyse proprement dite de l'évolution globale du tertiaire.

α - Choix d'une méthode de référence d'une évolution à la situation de départ

Il est nécessaire de travailler avec une double référence : une référence par rapport à une valeur nationale de départ et une référence de l'évolution relative du département par rapport à celle de l'ensemble du pays. Cette méthode doit donc conduire au choix de 2 critères permettant de dresser un tableau à double entrée afin de porter globalement un jugement. Ce tableau peut être plus ou moins affiné et donc comporter en ligne (comme en colonne) plus de 2 classes, mais il semble raisonnable pour la clarté de l'interprétation de se limiter à 2 ou 3 classes pour chacune des dimensions du tableau.

La position de départ du département peut être appréciée en écart ou en rapport (= coefficient de spécificité). Quant à l'évolution, elle peut se décomposer en combinaison de l'évolution nationale et d'un résidu caractérisant l'évolution propre du département.

Plusieurs méthodes peuvent être proposées.

- Etude des écarts

soit $K(1) = K(D,1) - K(N,1)$ l'écart à la date 1 entre la valeur départementale et la valeur nationale, on définit de même $K(2)$

soit $L(D) = K(2) - K(1)$

On peut dresser alors le tableau suivant :

	$L(D) < 0$	$L(D) > 0$
$K(1) < 0$	le département accentue son retard (retard accentué)	le département rattrape son retard (rattrapage)
$K(1) > 0$	le département perd son avance (uniformisation)	le département accentue son avance (accélération)

Notons que ce tableau ne permet pas de dire pour la première colonne s'il y a eu croissance ou décroissance ; selon le problème étudié il peut y avoir intérêt ou non à introduire cette distinction.

D'autres méthodes peuvent être utilisées et donnent le même résultat.

- Etude de l'évolution des coefficients de spécificité (ou des coefficients d'évolution départementaux et nationaux)

$$\text{Posons : } K(1) = K(D,1)/K(N,1)$$

$$K(2) = K(D,2)/K(N,2)$$

$$L(D) = \left[K(2)/K(1) = K(D,2)/K(D,1)/K(N,2)/K(N,1) \right]$$

	$L(D) < 1$	$L(D) > 1$
$K(1) < 1$	retard accentué	rattrapage
$K(1) > 1$	uniformisation	accélération

- Etude des taux d'évolution

$$K(D) = (K(D,2) - K(D,1)) / K(D,1)$$

$$K(N) = (K(N,2) - K(N,1)) / K(N,1)$$

$$K(3) = K(D,1) - K(N,1)$$

$$L(D) = K(D) / K(N)$$

	$L(D) < 1$	$L(D) > 1$
$K(3) < 0$	retard accentué	rattrapage
$K(3) > 0$	uniformisation	accélération

Ces 3 méthodes permettent également de dresser une carte rangeant chaque département dans l'un de ces 4 cas de figures; toutefois la première présente l'avantage d'être plus immédiatement compréhensible puisque L (D) représente dans ce cas la part de l'évolution qui est spécifique au département ; en effet on peut écrire :

$$\begin{aligned}
 K(D,2) &= K(2) + K(N,2) \\
 &= L(D) + K(1) + K(N,2) \\
 &= \underbrace{K(D,1)}_{\text{situation de départ}} + \underbrace{(K(N,2) - K(N,1))}_{\text{évolution nationale}} + \underbrace{L(D)}_{\text{évolution propre du département}}
 \end{aligned}$$

En choisissant la première méthode on s'offre la possibilité de constituer une cinquième catégorie correspondant aux valeurs de L (D) très peu différentes de zéro, et pour lesquelles on considérera la situation comme sans changement significatif. C'est la raison pour laquelle on utilisera cette méthode de préférence aux autres (pour lesquelles il faudrait faire référence aux valeurs de départ et d'arrivée pour déterminer si la situation est sans changement).

8 - Etude de l'évolution des effectifs

Entre 1954 et 1962, les effectifs du tertiaire ont augmenté dans tous les départements (voir cartes pages T 12 - 59 et 60), tandis qu' au cours de la période suivante, deux départements ont vu leurs effectifs tertiaires décroître, l'Indre et la Meuse (on ne tiendra pas compte de la Seine et Oise et de la Seine dont les découpages ont varié).

Entre 1954 et 1962, les zones de croissance supérieure à la croissance nationale sont localisées dans les départements périphériques de la Seine (avec le Calvados en extension), en Bretagne dans le Nord-Est, sur la rive gauche du Rhône, et dans le Languedoc au Sud du Bassin Aquitain, dans le Poitou-Charentes et le Sud-Ouest de la région Centre.

Entre 1962 et 1968 il est frappant de constater que l'évolution des effectifs est supérieure à la moyenne nationale, massivement dans le Sud (Régions Rhône-Alpes, Provence-Côte-d'Azur, Midi-Pyrénées, Aquitaine) et que le développement du tertiaire du Bassin Parisien se fait à l'est et au nord.

γ - Etude de l'évolution du taux de tertiaire

Si l'on module les observations précédentes en tenant compte de l'évolution de la population totale, il ne reste plus que 13 départements à avoir une croissance supérieure à la croissance moyenne avec 3 départements accélérants leur avance : la Haute-Garonne, le Bas-Rhin et la Côte d'Or, les autres ont une croissance inférieure à la croissance nationale (uniformisation) et même pour 4 d'entre eux il y a décroissance (Alpes-Maritimes, Bouches du Rhône, le Rhône et le Var).

On ne trouve que trois départements qui accentuent leur retard (le Doubs, la Haute-Marne et les Ardennes). C'est donc à un fort resserrement des disparités des taux de tertiaire auquel on assiste.

Entre 1962 et 1968 seul le département de la Haute-Garonne maintient son accélération, et la Seine-Maritime après une phase d'uniformisation reprend de l'avance. Les départements ayant déjà un taux de tertiaire par rapport à la population active supérieur à la moyenne nationale en 1954, mais pas par rapport à la population totale, possèdent en fin de période un taux de tertiaire supérieur à la moyenne nationale pour les deux taux. Chose curieuse ces départements (voir tableau p.321) se situent tous en Provence-Côte d'Azur et en Languedoc. Les départements dont le taux était inférieur à la moyenne nationale ont eu tous, à l'exception du département de la Meuse, une croissance positive qui se traduit même pour beaucoup d'entre eux par un rattrapage.

On peut résumer ces observations par le tableau synoptique suivant :

Type d'évolution (taux/Population totale)	nombre de départements	
	1954 - 1962	1962 - 1968
Perte de vitesse par rapport à l'évolution nationale mais la situation se - dégrade	0	1
- s'améliore	11	11
- rattrapage	65	63
- uniformisation	9	13
- accélération	4	1
	89	87

DEPARTEMENTS							Agglomérations de plus de 150 en 1962 (unité 1000) (1)
	totale			active			
	54	62	68	54	62	68	
Hautes Alpes			x		x	x	(Gap)
Alpes Maritimes	x	x	x	x	x	x	Nice (311)
Bouches du Rhône	x	xx	x	x	x	x	Marseille (808)
Côtes d'Or	x	x	x	x	x	x	Dijon (155)
Gard			x		x	x	(Nimes)
Haute Garonne	x	x	x	x	x	x	Toulouse (330)
Eironde	x	x	x	x	x	x	Bordeaux (462)
Hérault			x	x	x	x	(Montpellier)
Loiret		x					(Orléans)
Marne	x	x	x	x	x		(Reims)
Meurthe et Moselle				x	x	x	Nancy (209)
Pyrénées Orientales			x	x	x	x	(Perpignan)
Bas Rhin	x	x	x	x	x	x	Strasbourg (303)
Rhône	x	x	x	x	x	x	Lyon (886)
Seine	x	x	x	x	x	x	Paris
Seine Maritime	x	x	x	x	x	x	Rouen (327) - Le Havre (223)
Seine et Marne	x	x	x	x	x	x	(Melun)
Seine et Oise	x	x	x	x	x	x	(Versailles)
Savoie			x				(Chambery)
Var	x	x	x	x	x	x	Toulon (223)
Vaucluse			x	x	x	x	(Avignon)

(1) Sont donc absentes comme agglomérations de plus de 150 000 en 1962, par ordre d'importance décroissante : Lille, Nantes, Saint-Etienne, Lens, Grenoble, Valenciennes, Mulhouse, Clermont-Ferrand, Rennes, Tours. Les agglomérations entre parenthèses ont moins de 150 000 habitants.

2.213 - L'évolution du tertiaire, par département et CAE

L'introduction des divers composants du tertiaire, doit permettre d'affiner les remarques faites jusqu'à présent et d'établir, dans la mesure du possible une typologie de l'évolution, avec toutefois le risque de trouver presque autant de cas d'espèce que de départements.

Il a semblé intéressant de mettre en évidence tout d'abord les traits caractéristiques des départements dotés d'un tertiaire supérieur à la moyenne avant de caractériser les évolutions.

a) Etude de la situation de départ par CAE des départements "tertiaires".

Dans le tableau synoptique suivant on a caractérisé les départements ayant un taux de tertiaire, par rapport à la population totale, supérieur à la moyenne nationale à l'un des 3 recensements (à l'exception du Loiret, qui ne fait qu'une apparition en 1962) par 2 critères définis pour l'année 1962 :

- leur degré de spécialisation : on s'est pour cela intéressé aux CAE dont le coefficient de spécificité (= rapport du taux départemental au taux national) du taux par rapport à la population totale était supérieur à 1 en 1962 et l'on a indiqué ces CAE par ordre décroissant de leur coefficient de spécificité. Ce sont en effet ces activités qui sont responsables du fait que le taux de tertiaire départemental est supérieur à la moyenne nationale.
- on a retenu ensuite parmi les CAE retenues précédemment, celles qui représentent en 1962 plus de 2 % de la population active du département c'est-à-dire celles qui, par le poids de leurs effectifs, comptent réellement. Ces CAE sont mentionnées par ordre décroissant d'importance de leurs effectifs dans ce département.

Il est frappant de constater que les CAE importantes dans la spécialisation tertiaire (au sens défini plus haut) des départements considérés appartiennent dans plus de la moitié des cas à une ou plusieurs des 5 CAE suivantes : Administrations publiques (93) et défense nationale (94) commerces de détail (72 et 75), et transports terrestres (61). Par rapport au classement établi dans le tableau de la page 313 on se rend compte alors que 4 de ces CAE étaient dans les 5 premières places (93, 75, 61, 72) mais que les services aux particuliers (83) n'y figuraient pas (bien qu'ayant la première place) alors que la défense nationale, qui n'est qu'en 8ème position y figure.

Départements	Coeffi- cient de spécifi- cité ⁽¹⁾ en 1962	Catégorie d'activité économique retenue	
		Degré de spécialisation	Importance de la spécialisation
Hautes Alpes	98	91, 73, 71, 92, 83, 93, 76, 94	83, 93, 73, 94, 92
Alpes Maritimes	134	76, 73, 79, 82, 81, 91, 75, 83, 71, 72, 74, 92, 93	83, 73, 93, 75, 82, 72
Bouches du Rhône	124	62, 76, 81, 77, 79, 91, 61, 71, 74, 94, 75, 93, 83, 72	93, 83, 61, 75, 62, 72, 94, 74, 71
Côtes d'Or	111	71, 94, 61, 92, 93, 74, 75, 82	93, 61, 75, 94, 82, 71, 92
Gard	83	76, 91, 61, 94,	61, 94, 91
Haute Garonne	107	74, 92, 93, 75, 94, 77, 79, 91, 72	93, 75, 72, 74, 94, 92
Gironde	114	62, 77, 71, 94, 74, 82, 79, 75, 72, 93, 61	93, 75, 61, 82, 72, 94, 71, 74
Herault	95	76, 93, 92, 72, 75, 83,	93, 83, 75, 72, 92
Marne	101	72, 94, 61, 82,	72, 61, 94, 82
Meurthe et Moselle	96	94, 61, 72, 74,	94, 61, 72
Pyrénées Orientales	90	76, 71, 93, 61,	93, 61, 71
Bas Rhin	104	77, 76, 74, 83, 61, 79, 81, 92, 78, 75, 91,	83, 61, 75, 74,
Rhône	117	77, 76, 74, 83, 61, 81, 79, 92, 78, 75, 72, 91, 93,	83, 93, 75, 61, 72, 74,
Seine Maritime	112	62, 74, 82, 61, 75, 71, 73, 93,	93, 62, 75, 61, 82, 73
Seine et Marne	105	94, 91, 76, 61, 78, 72, 82, 93,	93, 61, 72, 82, 94, 78,
Seine et Oise	124	78, 81, 62, 91, 77, 61, 79, 93, 94, 83, 74, 92, 76, 75, 72, 82,	93, 83, 61, 75, 78, 72, 82, 94
Var	108	94, 76, 73, 83, 72	94, 83, 72, 73
Vaucluse	96	71, 76, 91, 61, 74, 94, 75,	75, 71, 61, 94

(1) Ce coefficient est égal à 100 (taux départemental/taux national).

b) Etude de l'évolution

On a vu que sur les 2 périodes, presque tous les départements avaient une évolution positive que ce soit en effectifs (seuls les départements de la Meuse et de l'Indre ont une évolution négative, entre 1962 et 1968) ou en taux par rapport à la population totale (les seules exceptions sont entre 1954 et 1962, les départements des Bouches du Rhône, Rhône et Alpes Maritimes qui sont en situation d'uniformisation, et entre 1962 et 1968 celui de la Meuse qui est en situation de retard accentué). Aussi ne s'intéresse-t-on qu'aux composantes du tertiaire qui expliquent la croissance du tertiaire, sans chercher à étudier les causes d'évolution négative (ou inférieure à l'évolution moyenne nationale).

On s'est tout d'abord intéressé aux seules CAE, responsables d'au moins 10 % de l'évolution du tertiaire du département dans la période considérée (c'est à dire que parmi les 18 CAE on ne s'est intéressé qu'à celles qui avaient un certain poids dans l'évolution). Pour éviter des interprétations erronées on a retiré de l'étude des 2 ou 3 départements ayant eu un accroissement de moins de 1.000 actifs dans le tertiaire (Cantal, Yonne et Haute-Saône entre 1954 et 1962, et Indre, Meuse entre 1962 et 1968).

Parmi les CAE retenues dans ce premier stade, on a sélectionné celles qui étaient responsables d'une accélération ou d'un rattrapage au sens défini à la page 317. On a retenu pour définir la situation "sans changement significatif" le seuil de 1 pour 1.000, c'est-à-dire que, si la variation d'écart du taux de tertiaire entre le taux départemental et le taux national de la CAE considérée est inférieure à 1 point (pour 1.000), on a considéré que l'évolution du département était parallèle à celle de la CAE pour l'ensemble du tertiaire.

On trouvera pages T 12.61 à 69 les tableaux qui ont permis d'établir la typologie retenue. On peut remarquer que le nombre de cas de rattrapage ou d'accélération est beaucoup plus important lorsque l'on travaille en taux par rapport à la population active que lorsque l'on travaille en taux par rapport à la population totale, mais que, à quelques rares exceptions les tableaux en taux par rapport à la population totale se déduisent des tableaux précédents par suppression de cas retenus.

Le tableau synoptique suivant permet une rapide comparaison entre les résultats selon les deux calculs possibles de taux.

TABLEAU 6

C. A. E. D.	1954 - 1962						1962 - 1968					
	Taux (Pop. active)			Taux (Pop. totale)			Taux (Pop. active)			Taux (Pop. totale)		
	Rattra- page	Accélération		Rattra- page	Accélération		Rattra- page	Accélération		Rattra- page	Accélération	
		Nb de dépar- tements	Numéro dép.		Nb de dépar- tements	Numéro dép.		Nb de dépar- tements	Numéro dép.		Nb de dépar- tements	Numéro dép.
61 Transports terrestres	0	1	73	id	0	-	3	1	73	id	id	id
62 Transp. fluv., marit., aériens	0	2	29, 76	id	id	id	0	0	-	id	id	id
71 Commerces agric. alimentaire	6	3	3, 14, 71	2	2	14, 22	1	2	16, 32	0	3	16, 32, 53
72 Com. agri. alim. détail	0	3	10, 51, 76	id	id	id	0	1	10	id	id	id
73 Hôtellerie débit de boisson	0	1	5	0	id	id	3	2	5, 73	id	id	id
74 Autres commerces gros	11	3	31, 67, 90	4	id	id	4	0	-	id	id	id
75 Autres commerces de détail	26	4	14, 31, 34 83	10	2	31, 34	33	9	...	20	1	...
76 Commerces mal désignés	0		0	id	id	id	0	0	-	id	id	id
77 Interm. du commerce industriel	0	1	69	id	id	id	0	0	-	id	id	id
78 Banques et assurances	1	2	72, 75	0	id	id	7	2	72, 77	4	id	id
79 Spectacles sédentaires	0		0	id	id	id	0	0	-	id	id	id
81 Services aux entreprises	0	1	75	1	id	id	3	1	69	id	id	id
82 Services domestiques	1		0	id	id	id	0	0	-	id	id	id
83 Services rendus aux particuliers	20	6	5, 34, 44, 64, 74, 83	8	4	5, 6, 75, 78	14	9	...	11	id	...
91 Eau, gaz, électricité	1	0		id	id	id	2	1	4	1	id	id
92 Transmissions et radios	0	0		id	id	id	1	2	48, 87	id	3	14, 48, 87
93 Administrations publiques	31	10	...	9	7	...	17	11	...	19	8	...
94 Défense Nationale	17	17	...	5	15	...	0	2	40, 70	id	id	id

Les résultats de cette étude ont été cartographiés (voir page A16⁽¹⁾) et il a semblé intéressant de rapprocher ces cartes des cartes d'évolution en effectifs et en taux pour les périodes correspondantes. On essaiera de dégager quelques caractéristiques sur l'évolution de zones définies ayant connu une croissance supérieure à celle de la moyenne.

- En Bretagne (Finistère, Morbihan, Ille et Vilaine, Loire Atlantique), la croissance du tertiaire est imputable pour 54 - 62 partout aux "autres commerce de détail" (75), et aux "services rendus aux particuliers" (83), auxquels il faut ajouter pour le Morbihan la CAE 71 (Commerces Agricoles et alimentaires de gros), et pour l'Ille et Vilaine administration publique (93) et les commerces agricoles et alimentaires de gros (71). Cette croissance du tertiaire s'interrompt pendant la période suivante et seul l'Ille et Vilaine conserve pour la période suivante une croissance supérieure à la moyenne (en valeur absolue comme en taux), grâce à la croissance des administrations publiques (93).
- Pour le Bassin Parisien, entre 1954 et 1962, le tertiaire se développe sur les départements périphériques de la Seine, plus le Calvados, essentiellement sous l'impulsion d'activités liées à des services de consommation finale ("autres commerces de détail" : 75 et "services rendus aux particuliers" 83), sauf en ce qui concerne le Loiret, alors que l'administration n'intervient que pour le Loiret et le Calvados.

Entre 1962 et 1968, les départements de l'Ouest de la couronne parisienne (à l'exception du Calvados) ne sont plus concernés tandis qu'une poussée s'effectue à l'Est avec adjonction de la Marne, de l'Yonne et de l'Aube (dont on ne tiendra pas compte du fait de son évolution très peu supérieure à l'évolution moyenne nationale). Ces nouveaux départements se caractérisaient par une accélération des commerces agricoles et alimentaires de détail durant la première période (avec en plus pour la Marne, les autres commerces en gros), cette tendance se maintient dans la période suivante pour l'Aube. L'évolution de tous les départements se caractérise par le développement des autres commerces de détail (avec même une accélération dans la Marne), les banques et assurances se développent dans le Calvados et la Seine-et-Marne, et les services aux particuliers se développent dans l'Yonne. En résumé ce qui caractérise essentiellement l'évolution de cette couronne parisienne c'est le développement de services consommés par les particuliers.

- Le Nord-Est : (Meuse, Meurthe et Moselle, Moselle, Bas-Rhin) cette région ne connaît une croissance qu'entre 1952 et 1962, et se trouve en perte de vitesse (sauf la Moselle) par rapport à l'évolution globale entre 1962 et 1968, avec même une croissance négative pour la Meuse, imputable essentiellement à une baisse absolue des services aux particuliers et de l'administration non compensée par la croissance de la défense nationale (94). La forte croissance de la Moselle en 1954 - 62 (28 % des effectifs du tertiaire) s'explique par une croissance

(1) ou T 12. 61

importante de sa population totale, qui a cru de 20 % sur la période. Entre 54 et 62 la seule évolution significative affecte le Bas-Rhin (autres commerces de gros et administrations publiques) tandis qu'entre 62 et 68, seul le tertiaire de la Moselle croît plus que la moyenne (et encore seulement en taux) sous l'effet de l'administration et des autres commerces de détail. On assiste donc à une extrême hétérogénéité dans les croissances et leurs explications.

- Le Sud-Est : entre 54 et 62 tous les départements situés sur la rive gauche du Rhône, à l'exception des Bouches du Rhône et du Var voient leurs effectifs tertiaires croître plus que la moyenne et en taux l'évolution est la même sauf pour les 2 départements précités auxquels s'ajoutent les Alpes Maritimes, connaissant la même évolution favorable, le Gard et l'Hérault qui jouxtent cette région. Ces 3 départements sont par ailleurs les seuls à avoir un taux de tertiaire supérieur à la moyenne nationale ; viennent s'ajouter pour la période suivante, outre les Bouches du Rhône et le Var, l'Ardèche et le Rhône pour une croissance des effectifs supérieure à la moyenne nationale, mais seul l'Ardèche a une croissance du taux de tertiaire supérieure à la moyenne. Pour la première période, tous les départements concernés, à l'exception de la Drôme et de la Haute Savoie, ont une forte croissance de la CAE 93 (administration). On constate en outre que les départements frontaliers du massif des Alpes ainsi que le Var et l'Hérault, tous départements à vocation touristique, voient une augmentation sensible de la CAE 83 (services rendus aux particuliers). Il y a la même accélération pour la Savoie et les Hautes Alpes. Parallèlement le Vaucluse rattrape son retard en ce qui concerne la CAE 74 (autres commerces de gros).

Pour la période 62 - 68, cette région affirme sa vocation tertiaire, puisqu'en plus des Bouches du Rhône du Var, du Rhône et des Alpes Maritimes, les Hautes Alpes, le Gard, l'Hérault la Savoie et le Vaucluse rentrent dans la catégorie des départements ayant un taux de tertiaire par rapport à la population totale supérieur à la moyenne nationale (soit en tout 9 départements sur 15 à taux supérieur à la moyenne 68). Notons toutefois, qu'en taux par rapport à la population active, tous ces départements, à l'exception de la Savoie étaient déjà privilégiés. La croissance de l'activité 93 se poursuit dans les départements déjà concernés auparavant (à l'exception du Gard) et pour 3 d'entre eux (Basses et Hautes Alpes et Hérault l'accélération de cette CAE se poursuit et affecte les Bouches du Rhône, le Var et l'Ardèche.

On constate aussi une progression des services rendus aux particuliers (83) sur les départements côtiers (sauf l'Hérault) et les départements du Centre (Drôme, Hautes Alpes, Isère) ainsi que pour le Rhône.

Fort accroissement également des "autres commerces de détail" (75) sur l'ensemble des Alpes, le Vaucluse et le Var. Enfin l'Ardèche rattrape partiellement son retard en matière bancaire (CAE 78) le taux de service de cette CAE par rapport à la population totale n'était que le tiers du taux moyen en 1962.

- Le Sud-Ouest : La croissance des effectifs tertiaires n'est supérieure à la moyenne entre 54 et 62 que dans 4 départements (Gers, Haute-Garonne, Pyrénées Atlantiques et Landes). La CAE 93 (administration) est responsable de cette évolution pour tous les départements sauf les Pyrénées Atlantiques, même chose pour les "autres commerces de détail" (75) sauf pour le Gers. Si on considère maintenant les taux de tertiaire par rapport à la population totale, on constate que les départements les plus défavorisés de cette région (Corrèze, Dordogne, Lot, Aveyron, Lozère, Tarn, Tarn et Garonne, Ariège) connaissent une croissance très forte et rattrapent leur retard.

Cette forte croissance est pourtant imputable sauf dans l'Aveyron à l'Administration et sauf dans le Tarn et Garonne et le Cantal aux autres commerces de détail. Pour la période 62 - 68, il est difficile de caractériser l'évolution du tertiaire. On peut toutefois noter qu'elle est imputable essentiellement à une ou plusieurs des catégories suivantes : administration, services aux particuliers, et autres commerces de détail avec une accélération des services aux particuliers dans les Hautes et Basses Pyrénées et en Haute Garonne. Ce dernier département, le seul globalement en accélération sur les 2 périodes, accentue aussi son avance entre 62 et 68 en ce qui concerne les administrations.

L'ensemble de cette zone, à l'exception des Basses Pyrénées et de la Haute-Garonne reste sous-tertiarisée, mais une bonne partie du très fort handicap initial est comblé. Aussi ne faut-il pas s'étonner de la prépondérance de tous les services consommés par les particuliers dans l'évolution du tertiaire, le retard à rattraper étant très grand.

Rattachons enfin à cette zone, les Pyrénées Orientales, qui se distinguent par une accélération de l'Administration (93) sur les 2 périodes et un rattrapage puis une accélération des autres commerces de détail (75).

- Le Centre-Ouest : (Indre, Indre et Loire, Vienne, Deux-Sèvres, Charente) connaît pour son département une croissance supérieure à 2 fois la croissance nationale, en effectif comme en taux par rapport à la population totale (à l'exception de l'Indre et Loire) entre 1954 et 1962. Cette forte croissance est imputable aux autres commerces de détail (75) et aux administrations publiques (93) (sauf pour les Deux-Sèvres), tandis que l'Indre et Loire rattrape son retard en "autres commerces de gros" et les Deux-Sèvres dans le secteur Banques et Assurances. Au cours de la période suivante le tertiaire de l'Indre décroît (mais moins que sa population totale) et la croissance ne reste forte, pour les effectifs que dans les Deux-Sèvres et l'Indre et Loire. On a très exactement pour la seconde période les secteurs explicatifs de la croissance au cours de la première période (à l'exception de la disparition de la CAE 75 pour l'Indre et Loire). Et viennent s'ajouter pour trois des départements une ou plusieurs CAE suivant les services aux particuliers (83) les autres commerces de gros (74) et l'hôtellerie débit de boisson (73) et pour la seule Charente, les transports terrestres (61).

2.22 - Test de quelques hypothèses

Plusieurs hypothèses qui semblaient a priori intéressantes ont été testées. Il s'agit tout d'abord puisqu'un certain resserrement des disparités a été constaté, de voir l'influence de la situation de départ sur l'évolution. Ensuite on a essayé de voir dans quelle mesure il était possible de dissocier un tertiaire induit d'un tertiaire fondamental. Enfin, il a semblé utile de reprendre à un niveau départemental l'étude de 2 types de regroupements thématiques en 4 classes.

2.221 - L'influence de la situation de départ du département sur l'évolution des CAE.

On a procédé successivement à une étude typologique à partir d'indicateurs et à une synthèse statistique par régression.

a) typologie à partir d'indicateurs

Le tableau synoptique suivant permet de constater à partir d'indicateurs statistiques le resserrement des disparités déjà constaté précédemment. Les conclusions étant identiques, que l'analyse soit menée en terme de taux par rapport à la population totale ou à la population active, on ne rapportera que l'une d'entre elles.

L'indicateur retenu, l'écart type des taux de tertiaire départementaux a été calculé par rapport à une moyenne qui diffère assez sensiblement du taux moyen France Entière. Ceci n'est pas étonnant puisque l'on a déjà constaté que moins de 20 % des départements avaient un taux supérieur au taux moyen France Entière T, c'est-à-dire que si l'on considère T comme une somme pondérée de taux départementaux, le poids de ces quelques départements est très nettement supérieur à la moyenne. Aussi sur les deux tableaux synoptiques suivants a-t'on porté la moyenne des taux départementaux qui permet d'affiner le jugement sur l'évolution des écarts-type et calculer le coefficient de variation (= moyenne / écart-type) qui permet de juger l'évolution de l'écart-type par rapport à celle de la moyenne.

On peut définir les situations suivantes :

- 1 - L'écart-type décroît tandis que le taux augmente ou encore l'écart-type décroît tandis que le taux diminue, mais moins que ce dernier (le coefficient de variation alors augmente) il y a alors un resserrement des disparités absolues et relatives.

ETUDE du RESSERREMENT DISPARITES des TAUX de TERTIAIRE

(par rapport à la population totale)⁽¹⁾

C. A. E.	Taux moyen France Entière (1) ‰			Moyenne des taux départe- mentaux (2) ‰			écart-type (3) ‰			Coefficient variation (= 2/3)		
	1954	1962	1968	1954	1962	1968	1954	1962	1968	1954	1962	1968
61 transports terrestres	151	140	142	128	118	120	51	44	40	2,51	2,68	3,00
62 tran. flu. mar. aériens	29	32	29	16	17	14	34	35	28	0,47	0,49	0,50
71 com. agric. alimentaires gros	52	55	55	52	55	58	17	16	16	3,06	3,44	3,63
72 com. agr. alim. détail	134	128	122	126	121	118	21	20	20	6,00	6,05	5,90
73 hôtellerie débit boisson	125	110	105	115	102	102	40	36	35	2,88	2,83	2,91
74 autres commerces gros	51	65	70	33	48	53	21	23	21	1,57	2,09	2,52
75 autres com. de détail	150	172	196	131	152	178	35	31	30	3,74	4,90	5,93
76 commerces mal désignés	17	3	7	4	2	3	13	5	8	0,31	0,40	0,38
77 interm. du com. industriel	24	21	22	17	12	13	11	10	11	1,55	1,20	1,18
78 banques et assurances	57	67	81	35	42	56	22	24	27	1,59	1,75	2,07
79 spectacles sédentaires	10	8	7	4	4	3	5	4	4	0,80	1,00	0,75
81 services aux entreprises	22	32	45	12	16	24	10	15	23	1,20	1,07	1,04
82 services domestiques	134	115	99	119	103	90	35	31	32	3,40	3,32	2,81
83 services rend. aux part.	241	267	314	204	230	277	55	57	71	3,71	4,03	3,90
91 eau, gaz, électricité	31	33	37	27	29	30	12	12	15	2,25	3,93	2,00
92 transmissions et radios	57	64	72	54	89	64	12	15	19	4,50	3,93	3,37
93 administrations publiques	226	263	317	202	243	293	41	39	47	4,93	6,23	6,23
94 défense nationale	64	79	68	61	81	62	61	61	39	1,00	1,32	1,59
ENSEMBLE TERTIAIRE	1.582	1.663	1.793	1.347	1.440	1.569	309	284	302	4,36	5,07	5,20

(1) Voir note bas de page 313 du tableau 1.

- 2 - L'écart-type décroît tandis que le taux diminue, mais plus que ce dernier (le coefficient de variation diminue), il y a resserrement des disparités absolues et accroissement des disparités relatives.
- 3 - L'écart-type croît tandis que le taux diminue ou encore l'écart-type croît tandis que le taux augmente mais moins que ce dernier (le coefficient de variation alors augmente). Il y a accroissement des disparités absolues et relatives.
- 4 - L'écart-type croît tandis que le taux augmente mais plus que ce dernier (le coefficient de variation alors diminue), il y a alors accroissement des disparités absolues et diminution des disparités relatives.

Il est possible de moduler cette classification pour tenir compte d'une quasi-stabilité de certains paramètres en introduisant une catégorie supplémentaire (stabilité) à celles proposées (resserrement, accroissement). Pour classer les diverses catégories on fera appel au tableau suivant qui donne les signes des variations entre 2 dates, respectivement du taux, de l'écart-type, et du coefficient de variation avec les conventions suivantes pour chaque triplet :

- + évolution positive
- évolution négative
- o stabilité
- . paramètre n'intervenant pas

		Evolution absolue		
		resserrement	Stabilité	accroissement
Evolution relative	resserrement	+ - . - - +	+ o -	+ + -
	Stabilité	- - o	+ o o - o o	+ + o
	accroissement	- - -	- o . + o +	- + . + + +

On peut alors dresser le tableau synoptique suivant :

TABLEAU 8

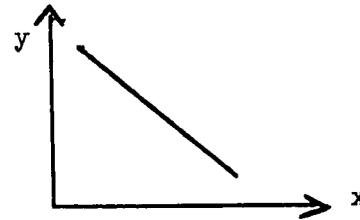
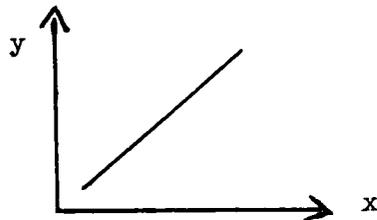
EVOLUTION DES DISPARITES DE TAUX DE TERTIAIRE DEPARTEMENTAUX
(RAPPORTES A LA POPULATION TOTALE)

		1954 - 1962			1962 - 1968		
		Evolution absolue			Evolution absolue		
		resserrement	stabilité	accroissement	resserrement	stabilité	accroissement
Evolution relative	resserrement	61, 62, 75, 76, 93, en- semble terti- aire		81, 92	61, 74, 94,	77	76, 81, 83 91, 92
	Stabilité		72			62	93
	accroissement	73, 82,	71, 79, 91, 94, 77	74, 78, 83		72, 73, 75, 79, 82, 71,	78, ensemble tertiaire

On peut constater sur ce tableau que six catégories seulement ont gardé d'après les critères retenus un comportement stable sur les 2 périodes, il s'agit des catégories 61, 71, 78, 81, 92 et que huit catégories ont changé radicalement de comportement (passage de resserrement à accroissement ou l'inverse) : les CAE 74, 75, 76, 77, 83, 91, 93, 94. Ce qui est regrettable, sur le plan des techniques de prévision, c'est que 2 d'entre elles (83 : services aux particuliers et 93 administrations publiques) sont responsables de plus de 50 % de l'évolution du tertiaire.

b) Synthèse par la technique de régression

On peut compléter cette description des resserrements par une autre étude qui consiste à examiner l'influence de la situation de départ sur l'évolution.



x = taux de départ

y = variation du taux

- la variation du taux est en général d'autant plus importante, que le taux de départ est élevé. Dans cette situation un grand nombre de départements sont en accélération ou en perte de vitesse. Il y a en quelque sorte polarisation
- la variation du taux est, en général d'autant plus faible que le taux de départ est élevé. Cette situation se caractérise par de nombreux rattrapages et uniformisations de départements. On a donc une homogénéisation de la répartition spatiale du taux.

On fera donc appel pour tester cette hypothèse à un modèle linéaire et l'on examinera le signe de la pente de la droite. L'utilisation d'un tel modèle amène un certain nombre de remarques.

- pour tester si le signe de la pente est significatif il est nécessaire d'accepter les hypothèses sous-jacentes à ce type de technique et déjà discutées à la page 162.

	61	62	71	72	73	74	75	76	77	78	79	81	82	83	91	92	93	94		
	Transports terrestres	Transports fluviaux, maritimes et aériens	Commerce agricole et alimentaire (gros)	Commerce agricole et alimentaire (détail)	Hôtellerie débits de boissons	Autres commerces de gros	Autres commerces de détail	Commerces mal désignés	Intermédiaire du commerce, industrie	Banques et assurances	Spectacles sédentaires	Services rendus aux entreprises	Services domestiques	Services rendus aux particuliers	Eau, Gaz, Electricité	Transmissions et radios	Administrations publiques	Défense Nationale	Ensemble tertiaire (CAE 61-94)	
y	54 - 62	- 0,0004	+ 0,0003	+ 0,0017	+ 0,0010	- 0,0013	+ 0,0041	+ 0,0073	- 0,0003	- 0,0009	+ 0,0023	+ 0,0000	+ 0,0011	- 0,0019	+ 0,0098	+ 0,0010	+ 0,0023	+ 0,0136	+ 0,0059	+ 0,0457
	62 - 68	+ 0,0014	- 0,0005	+ 0,0009	- 0,0001	+ 0,0006	+ 0,0016	+ 0,0078	+ 0,0002	+ 0,0004	+ 0,0037	- 0,0000	+ 0,0021	- 0,0027	+ 0,0136	+ 0,0006	+ 0,0018	+ 0,0144	- 0,0043	+ 0,0417
	54 - 68	+ 0,0010	- 0,0002	+ 0,0026	+ 0,0009	- 0,0006	+ 0,0057	+ 0,0151	- 0,0000	- 0,0005	+ 0,0060	- 0,0001	+ 0,0032	- 0,0046	+ 0,0235	+ 0,0016	+ 0,0040	+ 0,0280	+ 0,0017	+ 0,0874
x	54	+ 0,0299	+ 0,0040	+ 0,0122	+ 0,0292	+ 0,0265	+ 0,0078	+ 0,0305	+ 0,0010	+ 0,0040	+ 0,0081	+ 0,0010	+ 0,0028	+ 0,0275	+ 0,0469	+ 0,0063	+ 0,0124	+ 0,0468	+ 0,0145	+ 0,3114
	62	+ 0,0295	+ 0,0043	+ 0,0138	+ 0,0302	+ 0,0252	+ 0,0118	+ 0,0378	+ 0,0007	+ 0,0030	+ 0,0104	+ 0,0011	+ 0,0039	+ 0,0255	+ 0,0567	+ 0,0072	+ 0,0147	+ 0,0604	+ 0,0204	+ 0,3571
y	54 - 62	+ 0,0029	+ 0,0015	+ 0,0026	+ 0,0024	+ 0,0027	+ 0,0020	+ 0,0040	+ 0,0027	+ 0,0019	+ 0,0010	+ 0,0003	+ 0,0014	+ 0,0029	+ 0,0042	+ 0,0021	+ 0,0012	+ 0,0040	+ 0,0074	+ 0,0154
	62 - 68	+ 0,0026	+ 0,0017	+ 0,0016	+ 0,0020	+ 0,0027	+ 0,0018	+ 0,0023	+ 0,0013	+ 0,0013	+ 0,0015	+ 0,0002	+ 0,0024	+ 0,0036	+ 0,0047	+ 0,0019	+ 0,0016	+ 0,0048	+ 0,0102	+ 0,0165
	54 - 68	+ 0,0047	+ 0,0020	+ 0,0031	+ 0,0034	+ 0,0045	+ 0,0023	+ 0,0044	+ 0,0022	+ 0,0024	+ 0,0022	+ 0,0003	+ 0,0033	+ 0,0045	+ 0,0070	+ 0,0033	+ 0,0025	+ 0,0070	+ 0,0088	+ 0,0241
x	54	+ 0,0128	+ 0,0083	+ 0,0042	+ 0,0061	+ 0,0094	+ 0,0047	+ 0,0087	+ 0,0028	+ 0,0027	+ 0,0049	+ 0,0012	+ 0,0024	+ 0,0080	+ 0,0121	+ 0,0029	+ 0,0030	+ 0,0108	+ 0,0154	+ 0,0760
	62	+ 0,0114	+ 0,0089	+ 0,0041	+ 0,0056	+ 0,0086	+ 0,0056	+ 0,0080	+ 0,0014	+ 0,0023	+ 0,0053	+ 0,0010	+ 0,0035	+ 0,0072	+ 0,0129	+ 0,0033	+ 0,0036	+ 0,0110	+ 0,0164	+ 0,0702
a	54 - 62	- 0,1310	+ 0,0563	- 0,2191	- 0,1536	- 0,1154	+ 0,1203	- 0,1799	- 0,8621	- 0,3879	+ 0,0706	- 0,1916	+ 0,3722	- 0,1532	+ 0,0050	- 0,1132	+ 0,1189	- 0,0427	- 0,0507	- 0,0931
	62 - 68	- 0,1010	- 0,1685	- 0,1020	- 0,1017	- 0,0795	- 0,1173	- 0,0493	+ 0,0122	+ 0,0430	+ 0,0841	- 0,0071	+ 0,4184	- 0,1425	+ 0,1825	+ 0,2086	+ 0,2216	+ 0,1342	- 0,4843	- 0,0351
	54 - 68	- 0,2315	- 0,1269	- 0,2344	- 0,2668	- 0,2126	- 0,0004	- 0,2076	- 0,5614	- 0,3969	+ 0,1425	- 0,1704	+ 0,9084	- 0,2567	+ 0,1788	- 0,0174	+ 0,3094	+ 0,0572	- 0,4253	- 0,1294
a	54 - 62	+ 0,0191	+ 0,0186	+ 0,0615	+ 0,0379	+ 0,0279	+ 0,0431	+ 0,0444	+ 0,0505	+ 0,0603	+ 0,0210	+ 0,0224	+ 0,0472	+ 0,0352	+ 0,0365	+ 0,0747	+ 0,0418	+ 0,0391	+ 0,0508	+ 0,0191
	62 - 68	+ 0,0220	+ 0,0096	+ 0,0414	+ 0,0370	+ 0,0325	+ 0,0327	+ 0,0306	+ 0,1020	+ 0,0604	+ 0,0287	+ 0,0277	+ 0,0511	+ 0,0511	+ 0,0341	+ 0,0592	+ 0,0412	+ 0,0436	+ 0,0413	+ 0,0246
	54 - 68	+ 0,0299	+ 0,0211	+ 0,0736	+ 0,0526	+ 0,0452	+ 0,0522	+ 0,0484	+ 0,0568	+ 0,0836	+ 0,0444	+ 0,0232	+ 0,1076	+ 0,0528	+ 0,0586	+ 0,1213	+ 0,0820	+ 0,0683	+ 0,0399	+ 0,0307
b	54 - 62	+ 0,0035	+ 0,0001	+ 0,0043	+ 0,0055	+ 0,0018	+ 0,0031	+ 0,0129	+ 0,0006	+ 0,0006	+ 0,0018	+ 0,0002	+ 0,0001	+ 0,0023	+ 0,0096	+ 0,0017	+ 0,0008	+ 0,0157	+ 0,0067	+ 0,0747
	62 - 68	+ 0,0043	+ 0,0002	+ 0,0011	+ 0,0030	+ 0,0026	+ 0,0030	+ 0,0097	+ 0,0002	+ 0,0003	+ 0,0028	- 0,0000	+ 0,0004	+ 0,0009	+ 0,0032	- 0,0008	+ 0,0014	+ 0,0063	+ 0,0056	+ 0,0542
	54 - 68	+ 0,0079	+ 0,0003	+ 0,0055	+ 0,0087	+ 0,0050	+ 0,0057	+ 0,0214	+ 0,0005	+ 0,0011	+ 0,0049	+ 0,0001	+ 0,0007	+ 0,0024	+ 0,0151	+ 0,0017	+ 0,0002	+ 0,0253	+ 0,0078	+ 0,1277
R ²	54 - 62	- 0,5890	+ 0,3062	- 0,3533	- 0,3951	- 0,4020	+ 0,2838	- 0,3948	- 0,8753	- 0,5634	+ 0,3349	- 0,6721	+ 0,6408	- 0,4189	+ 0,0146	- 0,1587	+ 0,2885	- 0,1150	- 0,1053	- 0,4587
	62 - 68	- 0,4376	- 0,8804	- 0,0308	- 0,2795	- 0,2508	- 0,3649	- 0,1685	+ 0,0127	+ 0,0753	+ 0,2961	- 0,2715	+ 0,6555	- 0,2836	+ 0,4931	+ 0,3495	+ 0,4953	+ 0,3103	- 0,7795	- 0,1495
	54 - 68	- 0,6339	- 0,5371	- 0,3200	- 0,4734	- 0,4462	- 0,0007	- 0,4135	- 0,7237	- 0,4494	+ 0,3223	- 0,6151	+ 0,6669	- 0,4583	+ 0,3077	- 0,0152	+ 0,3712	+ 0,0885	- 0,7492	- 0,4076
t	54 - 62	- 6,8759	+ 3,0345	- 3,5633	- 4,0576	- 4,1415	+ 2,7921	- 4,0541	- 17,0790	- 6,4332	+ 3,3535	- 8,5633	+ 7,8753	- 4,3522	+ 0,1382	- 1,5164	+ 2,8427	- 1,0917	- 0,9993	- 4,8704
	62 - 68	- 4,5917	- 17,5150	- 0,2912	- 2,7463	- 2,4442	- 3,5815	- 1,6126	+ 0,1202	+ 0,7122	+ 2,9244	- 0,2562	+ 8,1877	- 2,7904	+ 5,3471	+ 3,5197	+ 5,3790	+ 3,0795	- 11,7400	- 1,4268
	54 - 68	- 7,7313	- 6,0062	- 3,1860	- 5,0703	- 4,7043	- 0,0068	- 4,2844	- 9,8923	- 4,7457	+ 3,2125	- 7,3592	+ 8,4437	- 4,8636	+ 3,0504	- 0,1432	+ 3,7718	+ 0,8384	- 10,6720	- 4,2107

Equation : $y = a \cdot x + b$ où : y = variation du taux de l'activité (= actifs de la CAE dans le département/ Population active du département) sur la période x = taux de l'activité à la date de départ

$$\bar{x} = \frac{1}{89} \left[\sum_{i=(1,19),(21,90)} (x_i) \right] \quad \sigma_x^2 = \frac{1}{89} \left[\sum_{i=(1,19),(21,90)} (x_i - \bar{x})^2 \right]$$

 R^2 = coefficient de corrélation $t = \frac{b}{\sigma_b}$ suit une loi de Student sous certaines hypothèses

- Il n'y a aucune relation entre un coefficient de corrélation faible et le fait que le coefficient de régression soit significatif ou non (dans le sens contraire il y en a bien évidemment une), il suffit de se rappeler que l'introduction de variables explicatives nouvelles significatives (du point de vue du coefficient de régression) ont, dans une régression multiple où les variables exogènes sont normalement faiblement corrélées, un faible coefficient de corrélation avec la variable expliquée.

On peut tirer du tableau de résultat précédent le tableau suivant :

		1962 - 1968		
		tendance à la polarisation	aucune tendance significative	tendance à l'homogénéisation
1954 - 1962	tendance à la polarisation	74, 78, 81, 92		62
	aucune tendance significative	83, 91, 93		94
	tendance à l'homogénéisation		71, 75, 76, 77, 79 ensemble tertiaire	61, 72, 73, 82

Ce tableau n'est pas comparable directement au tableau synoptique de l'étude du resserrement des disparités qui relève de la statistique comparative. Pour qu'il le soit il serait nécessaire que les coefficients de corrélation soient élevés et l'on aurait alors coïncidence entre la notion de resserrement et d'homogénéisation et celle d'accroissement et de polarisation. Il faut donc interpréter avec la plus grande prudence ces résultats. Il est toutefois intéressant de constater qu'entre les 2 périodes, 10 des 18 CAE changent de catégorie (ainsi que l'ensemble du tertiaire) mais qu'une seule change fondamentalement (62). S'il n'y a pas vraiment stabilité de comportement il n'y a pas malgré tout, bouleversement complet.

2.222 - Etude de l'existence d'une induction

Il a semblé intéressant d'examiner dans quelle mesure il était possible de séparer un tertiaire induit d'un tertiaire fondamental en partant des approches théoriques définies au chapitre 4. Parmi

les diverses approches concrètes possibles, deux semblaient particulièrement intéressantes : les travaux de Cahen et Pousard et ceux de la SOGREP. La méthode des enquêtes directes n'est pas envisageable parce que, indépendamment des contraintes de coût, le découpage adopté ne s'y prêtait pas et que de plus cette étude remonte trop loin dans le temps. Il ne restait plus qu'à essayer la méthode utilisée par la SOGREP. (1)

Par rapport aux autres aspects de l'étude menée on peut observer que :

- si tout l'emploi est induit, le resserrement des disparités ne joue que sur l'écart-type de la distribution normale (indépendamment de l'évolution de la moyenne).
- avoir la totalité de l'emploi induit n'exclut pas une polarisation des valeurs fortes sur quelques départements. Cette polarisation peut être stable dans le temps, mais on ne peut d'après la théorie adoptée considérer que cette stabilité implique l'existence d'emplois fondamentaux sur ces zones.

Cette étude est donc complémentaire des études précédentes mais les résultats seront à interpréter avec beaucoup de précaution. Certains emplois sont, de par leur nature, totalement induits mais les techniques utilisées peuvent mettre en évidence un emploi fondamental, factice parce que certains départements disposent de services notablement sous-développés en ce qui concerne la CAE considérée.

On trouvera en annexe les résultats obtenus en utilisant cette méthode. On a tiré de ces documents le tableau synoptique où pour certaines catégories une fourchette de taux d'induction a été donnée devant l'impossibilité de déterminer ponctuellement le taux d'induction.

Toutes les variations à la hausse ou à la baisse correspondent à une croissance ou une décroissance relative de la CAE par rapport à la population.

Il semblerait que dans certaines CAE, après une période de rattrapage on en arrive à une distribution normale des taux: c'est le cas des autres commerces de détail à la fin de la seconde période et des Banques et Assurances à la fin de la première, mais pour cette dernière CAE la libération des implantations d'agences bancaires au début de la seconde période modifie très sensiblement la physionomie des réseaux bancaires.

Enfin et c'est la conclusion la plus intéressante la CAE "Administration Publique" a une répartition spatiale de son taux par rapport à la population totale, constamment normale. La moyenne trouvée comme paramètre dans cette loi normale est constamment sous-estimée d'environ 30 points (sur 1000) ce qui correspond sans doute à la constance du poids de la région Parisienne.

(1) Voir chapitre 4 - page 89

TAUX D'INDUCTION (= effectifs CAE / Population totale)

unité 1000

C . A . E . D . (1)	1954	1962	1968
61 Transports terrestres	88 - 90	86 (2)	90 - 1
71 Commerces agricoles et alimentaires gros	40	42	51
72 Commerces agricoles et alimentaires détail	111	109 - 110 (2)	110 (3)
73 Hotellerie et débits de boissons	90 - 95	90 - 91 (2)	89 - 91 (3)
74 Autres commerces de gros	18 - 19	36	40
75 Autres commerces de détail	117 - 120	144	normalité
77 Intermédiaires du commerce et de l'industrie	11	6 (2)	9
78 Banques et assurances	28	normalité	46 - 47
79 Spectacles sédentaires	2	2 (2)	2 (3)
81 Services rendus aux entreprises	7	8 - 9	16
82 Services domestiques	105 - 106	90 (2)	75 (3)
83 Services rendus aux particuliers	181	196	240
91 Eau, gaz, électricité	18 - 20	20 - 21	23
92 Transmission et radio	49 - 50	54 - 55	normalité
93 Administrations Publiques	normalité	normalité	normalité
94 Défense Nationale	25 - 30	46 - 48	38 - 41 (3)
ENSEMBLE DU TERTIAIRE	1.240	1.285	1.300

(1) Les CAE 62 transports fluviaux maritimes et aériens et 76 commerces mal désignés n'ont pas été pris en compte du fait de leur caractère très spécial.

(2) Le rapport de cette CAE dans la population totale est en regression sur la période 54 - 62.

(3) Le rapport de cette CAE dans la population totale est en regression sur la période 62 - 68.

2.223 - Etude globale de l'évolution des consommations intermédiaires de tertiaire

Comme on l'a vu dans la section 1 de ce chapitre certains économistes expliquent le déplacement des emplois vers le tertiaire, par un accroissement relatif plus fort de la demande intermédiaire de services, résultant en quelque sorte d'une division accrue du travail. Il a semblé intéressant d'examiner cette hypothèse bien qu'on ne puisse en aucune façon être "spatialisé" dans l'état actuel de nos connaissances statistiques.

On trouvera page suivante des données extraites des tableaux d'échange interindustriels entre 1957 et 1971 (T.E.I. en 28 branches).

Des remarques analogues à celles faites par Fuschs peuvent être faites : la part du tertiaire dans les consommations intermédiaires, pour des données exprimées en francs constants, est en très légère hausse, tandis que des comparaisons menées en francs courants montrent une hausse sensible.

Ce décalage entre observations tirées des données en francs courants et constants, laisse à penser qu'en plus d'un effet de division de travail, se sont ajoutés un effet - qualité et un ajustement par les prix (généralisé par une demande en valeur en croissance plus rapide, que celle possible en volume).

Il est donc probable que comme aux Etats-Unis la croissance de la demande intermédiaire ne joue qu'un rôle limité.

	1959		1962		1967		1971	
	Prix 1963	Prix courants						
Consommation intermédiaire								
- branche 16 (Services)	19 848	16 299	25 545	24 164 (*)	35 612	42 589	48 432	76 158
- branche 14 A (Transports)	<u>11 460</u>	<u>10 065</u>	<u>13 612</u>	<u>12 985</u>	<u>17 727</u>	<u>19 344</u>	<u>22 221</u>	<u>32 282</u>
(1) Total	31 308	26 364	39 157	37 149	53 339	61 933	70 653	108 440
(2) Somme des consommations intermédiaires	171 869	152 103	214 994	205 567	290 166	311 622	377 334	502 852
(3) Valeur ajustée	281 404	241 254	339 874	322 408	448 859	503 533	561 320	783 504
(4) Production des branches au prix réel à la production 4 = 2 + 3	453 271	393 357	554 868	527 975	739 025	815 155	938 654	1 286 356
(5) Production intérieure brute	285 306	244 956	346 743	329 177	460 227	515 111	580 857	803 516
(7) = (1)/(2) unité 10 000	1 822	1 733	1 821	1 807	1 838	1 987	1 872	2 156
(8) = (1)/(4) unité 10 000	691	670	706	704	722	760	753	843

Source - Rapport sur les comptes de la nation 1971 (TEI en annexes) et les comptes de biens et services (collection INSEE : C 10) - TEI en 28 branches.

(*) En 1962, la ventilation de ce poste est la suivante : (en partant du TEI en 78 branches en tenant compte des problèmes de consolidation inhérents aux TEI) : 16 829 services rendus aux entreprises, 3 453 services des artisans mécaniciens, 304 services de santé et 3 578 de services rendus principalement aux particuliers (pour leur partie consommation intermédiaire).

2.3 - Mobilité professionnelle et tertiaire

Une enquête faite sur la mobilité professionnelle en France⁽¹⁾ entre 1959 - 1964 donne le tableau des transferts d'un secteur à un autre (le secteur étant obtenu par agrégation de C.S.P.) pour cette période :

1.1.59 1.1.64	primaire	secondaire	tertiaire	TOTAL
Primaire	886	97	17	1000
Secondaire	8	940	52	1000
Tertiaire	2	28	970	1000

On constate que les travailleurs du tertiaire sont les plus stables. Par ailleurs une analyse fine⁽²⁾ révèle que l'arrivée des hommes vers le tertiaire est essentiellement due aux promotions d'ouvriers vers les emplois de cadres et à des glissements du statut d'ouvriers ayant statut d'indépendant, tandis que les femmes deviennent surtout employées (40 %) ou personnel domestique (20 %). Quant aux nouveaux venus sur le marché on constate que davantage de femmes (64 %) que d'hommes (56 %) s'orientent vers le tertiaire.

(1) Etude et Conjoncture oct. 1966 (cette enquête a donné lieu à d'autres articles jusqu'en mars 1972 dans cette revue)

(2) Voir à ce propos M. PRADERIE (ouvrage cité).

SECTION 3 - RECHERCHE D'UN SOUS-MODELE TERTIAIRE

Comme dans les chapitres précédents, on établira d'abord un certain nombre d'hypothèses que l'on désire tester à l'aide d'indicateurs supposés représentatifs. La recherche d'un modèle satisfaisant s'effectuera successivement dans le cadre des approches en effectifs et en taux. Enfin dans une dernière sous-section on examinera les résultats du sous-modèle retenu pour 1962, ainsi que la prévision 1968 obtenue grâce à ce sous-modèle.

Au préalable il convient de définir de façon précise la variable endogène. Deux approches peuvent être tentées.

La première consiste à scinder le secteur tertiaire en plusieurs catégories, correspondant chacune à une vocation spécifique, c'est ainsi que l'on peut définir un tertiaire à vocation industrielle et une catégorie de services rendus aux particuliers, on reprendra des groupements plus fins en croisant des critères de finalité de prestation et d'homogénéité de comportement à l'aide de diverses typologies avancées dans la section 2 du présent chapitre. Cette démarche se heurte en pratique à deux écueils : d'une part il faut disposer pour chaque catégorie d'un corps d'hypothèses assez différenciées, fondées sur des mécanismes théoriques différents, et pour lesquels des indicateurs soient disponibles au niveau départemental ; d'autre part il est nécessaire que les variables endogènes soient non équivoques, ce qui signifie par exemple pour l'administration (dans l'exemple de séparation du tertiaire en 2 catégories) que l'on soit capable de dissocier les services rendus aux particuliers de ceux rendus aux entreprises, soit en descendant à un niveau de détail plus fin soit en disposant de clés de répartition appropriées et correctes.

Une seconde approche peut être envisagée, et qui consiste à prendre le secteur tertiaire dans son ensemble en s'appuyant sur les trois considérations suivantes. La première découle de la très grande difficulté de pouvoir suivre correctement la première démarche. Ensuite on peut arguer du fait que si l'approche des mécanismes réels se heurte à de très grandes difficultés ceci n'est pas trop grave si des "régularités" peuvent être mises en évidence, et dans ce cas pourquoi ne pas le faire au niveau global. Enfin on peut tenir compte de l'importante remarque

...

faite au chapitre 3 page 73 sur la stratification des variables endogènes.

C'est finalement la seconde approche qui a été retenue ici. Avant de procéder à la recherche du sous-modèle tertiaire, il faut d'abord préciser le champ de définition du secteur tertiaire. Une catégorie d'activité économique du tertiaire, la défense nationale (CAE 9.4) a été retirée de la définition du tertiaire compte tenu de son caractère très particulier et du fait que dans ce cas les prévisions sont toujours réalisées. Le tertiaire défini ici comprend donc les CAE 6.1 à 9.3.

3.1 - Approche en effectifs

Deux approches ont été simultanément tentées, la première dans laquelle la variable endogène est la variation 54 - 62 de tertiaire, la seconde cherche à déterminer directement les effectifs tertiaires en 1962. Avant d'examiner chacune des démarches il faut examiner le corps d'hypothèses et leurs indicateurs que l'on va chercher à tester.

3.11 - hypothèses

On peut grouper celles-ci autour de quatre grands thèmes. Tous les indicateurs introduits seront testés pour les deux variables endogènes, même si a priori certains ne concernent que l'une d'entre elles.

α - poids du passé

Il n'est pas question ici de rechercher un modèle qui soit explicatif de façon satisfaisante mais de trouver un modèle descriptif satisfaisant. Dans ce cadre, on peut supposer qu'un certain nombre de facteurs ou de mécanismes explicatifs ont une assez grande stabilité et qu'en conséquence, la reprise de la donnée "effectifs tertiaire 1954" rentrera de façon satisfaisante dans les régressions descriptives du tertiaire 1962. On court le risque en introduisant cette variable de masquer certaines relations explicatives.

...

Par contre, cette donnée a peu de chances d'intervenir pour la variation de tertiaire, car on a pu voir dans l'étude descriptive que les disparités départementales des taux du tertiaire se sont resserrées au cours du temps ; pour qu'il y ait une certaine proportionnalité entre la variation et la situation de départ, il faudrait que des conditions bien particulières de liaison entre l'évolution du tertiaire et celle de la population totale soient en général remplies.

β - la consommation de service par les particuliers

Si l'on se rapporte aux lois d'Engel(1) la consommation de service par les particuliers devrait être d'autant plus importante que les revenus sont élevés. Les indicateurs introduits sont les salaires moyens en 1955 pour l'ensemble de la population et les deux CSP les plus représentatives : les ouvriers et les employés. On ne peut malheureusement pas disposer de revenus pour cette date comme on l'a déjà indiqué dans un précédent chapitre et en ne prenant qu'une partie des revenus perçus on introduit un biais important. En outre la liaison transite par des consommations en valeur. Il est donc vraisemblable que la recherche d'une liaison entre des effectifs tertiaires et des salaires ne mène à rien.(2)

On a en outre introduit dans l'ensemble d'indicateurs à tester la population totale en 1954, ainsi que la variation 54 - 62 de population totale, cette dernière variable n'ayant de sens que pour l'étude de la variation de tertiaire.

γ - les services aux entreprises

Une partie importante du tertiaire a pour vocation de fournir certaines prestations de service au secteur secondaire. Cette complémentarité comme on a déjà eu l'occasion de le dire est délicate à saisir, et dépend en outre de l'importance locale de l'auto-fourniture de service (puisque la définition retenue ici est celle de l'activité collective et non du métier). On a assimilé ici l'armée au secteur secondaire car la Défense Nationale

...

(1) Voir page 309

(2) Sauf si la disparité départementale des revenus est bien prise en compte par celle des salaires et que la production service par actif du tertiaire (concerné) soit assez homogène sur l'ensemble du territoire, ce qui n'est pas a priori le cas.

est aussi consommatrice de services produits localement. Pour l'étude de la variation de tertiaire on a ajouté comme indicateur la variation 54 - 62 du secteur secondaire + armée.

λ - le mode de vie et l'éducation

La consommation de service par les particuliers est liée à une certaine division du travail que l'on trouve surtout en milieu urbain (cantines, transports en commun,...) On a donc introduit comme indicateurs la taille moyenne urbaine (déjà plusieurs fois utilisée). Par ailleurs, les comportements face à ce type de consommation diffèrent avec l'âge, aussi a-t-on introduit la population de moins de 35 ans comme indicateur de couche de population à plus forte propension consommatrice de tertiaire. On peut ajouter en outre qu'une fraction importante de cette population se trouve être scolarisée et que les salariés de l'éducation nationale représentent la catégorie la plus importante (en nombre) du tertiaire. On a en conséquence introduit aussi la population à scolarisation obligatoire, celle des moins de 14 ans, dans l'ensemble d'indicateurs à tester. Pour l'étude de la variation de tertiaire on a ajouté la variation 54 - 62 de chacun de ces deux indicateurs.

3.12 - recherche d'un sous-modèle pour la variation 54 - 62 du tertiaire

On trouvera page suivante le tableau synoptique de résultats obtenus.

Ce coefficient de détermination est très élevé. Parmi les hypothèses infirmées statistiquement on notera celles de l'influence des revenus et de la situation du tertiaire en 1954.

Deux facteurs jouent un rôle prépondérant, la taille moyenne urbaine et la variation 54 - 62 de population totale qui expliquent constamment plus de 60 % de la variable endogène. La variation de population totale inclut en fait une variation de population active ; il n'est en fait pas possible de dissocier la part de services aux particuliers de celle imputable au secondaire ; la variation de population totale intègre très certainement les deux influences.

Dans la régression optimale de rang 3 intervient la population totale. On peut sans doute interpréter ceci comme l'incidence de l'accroissement de consommation de services par individu.

Nombre de variables rentrant dans la régression optimale	coefficient de détermination	Tertiaire 54	Actifs du secondaire + armée 1954	population totale 1954	population de moins de 35 ans 1954	taille moyenne urbaine 1954	variation 54 - 62 de population totale	Autres variables	Constante
2	0,98755					0,01162 (30,9)	0,00038 (22,4)		3178
3	0,99270			0,00606 (7,7)		0,00937 (22,5)	0,11544 (26,8)		1077
4	0,99448		- 0,04468 (- 5,2)	0,01620 (7,8)		0,00993 (26,1)	0,11335 (29,9)		- 200

RECHERCHE D'UN MODELE EXPLICATIF DU TERTIAIRE 62

Nombre de variables rentrant dans la régression optimale	coefficient de détermination	Tertiaire 54	Actifs du secondaire + armée 1954	population totale 1954	population de moins de 35 ans 1954	taille moyenne urbaine 1954	variation 54 - 62 de population totale	Autres variables	Constante
2	0,99962	1,07709 (316,3)					0,15387 (24,3)		- 886
3	0,99966	1,03050 (237,7)				0,00776 (12,2)	0,12697 (28,7)		1710
4	0,99987	1,00727 (102,2)			0,00903 (2,6)	0,00926 (11,0)	0,11557 (18,9)		1224
5	0,99991	1,02297 (112,9)	- 0,04606 (-5,3)		0,02354 (5,8)	0,00940 (12,8)	0,11503 (21,6)		200

(1) tertiaire = CAE 6.1 à 9.3

Dans la régression suivante une dernière variable intervient avec un signe négatif. Il s'agit du secondaire + armée. L'interprétation de ce signe est délicate car cette variable est une partie de la population totale (environ 1/8) ; le fait que le coefficient de régression de la population totale fasse plus que doubler peut laisser penser que l'accroissement de tertiaire imputable au secondaire est moindre que celui imputable à la population totale, mais il s'agit là d'une hypothèse non vérifiable.

Il semble donc qu'il soit préférable de retenir la régression optimale de rang 3 dont l'interprétation n'est pas équivoque.

3.13 - recherche d'un sous-modèle pour le tertiaire 1962.

Les coefficients de détermination sont meilleurs que dans l'approche en variation mais la variance de la variable endogène est ici nettement plus importante que dans le cas précédent (le coefficient de variation étant, lui, du même ordre).

Comme on pouvait s'y attendre, le tertiaire 1954 joue un rôle prépondérant dans toutes les régressions et le coefficient de régression est toujours légèrement supérieur à 1. La variation de population totale est, elle aussi, présente dans toutes les régressions.

La régression de rang 3 correspond approximativement à la somme du tertiaire 54 et de la variation 54 - 62 de tertiaire telle qu'elle est définie dans la régression optimale de rang 2 dans l'approche précédente (seule la constante diffère mais sa baisse est compensée par le fait que le coefficient de régression de la variable tertiaire 54 est très légèrement supérieure à 1).

Le parallélisme se poursuit dans les régressions suivantes à ceci près que la population totale est remplacée par la population de moins de 35 ans ce qui s'explique sans doute par le fait que l'autre catégorie de population est déjà prise en compte dans le tertiaire 54. Dans la dernière régression optimale on retrouve les actifs du secondaire + armée avec le même signe négatif du coefficient de régression.

Pour la même raison que précédemment on s'arrêtera à la régression optimale de rang 4.

...

On ne s'étonnera pas, compte tenu de l'observation précédente de constater que les écarts résiduels entre les deux régressions retenues sont très voisins. Il s'en suit que la variance résiduelle des deux régressions est du même ordre de grandeur.

Il semble donc qu'il n'y ait pas de raisons particulières de retenir l'approche en variation plutôt que celle en effectif. Toutefois, comme il faut bien retenir l'une des deux régressions pour pouvoir comparer l'approche en effectif et celle en taux on prendra arbitrairement la seconde (ce qui est sans importance puisque le critère de décision retenu jusqu'ici est la variance résiduelle des effectifs).

3.2 - Approche en taux

Le dénominateur retenu ne peut guère être que la population totale, parce que directement ou indirectement (dans le cas de services aux autres activités économiques) tout peut se ramener à la population totale, dont les actifs constituent un sous-ensemble, alors qu'inversement privilégier la population active ou une partie de celle-ci revient implicitement à négliger la part, très importante dans le tertiaire, des services rendus aux particuliers. La prise en compte du facteur population dans la variable endogène oblige à modifier certaines hypothèses ou certains indicateurs. On examinera donc ces modifications avant d'examiner les résultats obtenus.

3.21 - Hypothèses

Examinons successivement ce que deviennent les quatre familles d'hypothèses envisagées précédemment.

α - le poids du passé

On peut introduire, comme pendant aux effectifs tertiaire 54, le taux de tertiaire 54. Mais il semble plus judicieux d'introduire ici le phénomène de resserrement des disparités de taux noté plus haut et qui correspond d'une certaine façon à une homogénéisation du comportement. Le taux de tertiaire introduit comme variable exogène est alors le taux observé en 1962 pour l'ensemble de la France auquel on a retranché l'écart observé en 1954 entre la valeur France entière et la valeur départementale. Comme le taux de tertiaire pour l'ensemble de la

...

France croit sur la période, l'éventail des coefficients de spécificité des taux départementaux se restreint par le jeu du mécanisme choisi.

β - la consommation de service par les particuliers

Les données de salaires sont conservées, et les limites d'interprétation déjà soulignées demeurent.

Par ailleurs, on peut considérer que dans une zone donnée, le revenu d'un ménage (dont dépend la consommation de service) sera d'autant plus élevé que les femmes ont la possibilité de travailler dans cette zone, ainsi que les franges inférieures et supérieures de la population masculine en âge de travailler.⁽¹⁾ On a donc intégré les taux d'activités masculins et féminins des 20 - 64 ans en 1954, malgré l'importance déjà signalée de la conjoncture. Comme la variable endogène est datée en 1962 on a intégré en outre la variation 54 - 62 de ces deux taux d'activités afin de prendre en compte, malgré l'imperfection de cet indicateur, l'influence de l'évolution de cette activité sur l'évolution des revenus.

Enfin la population totale a été maintenue (après multiplication par un scalaire) pour tenir compte d'éventuelles économies d'échelle.

γ - les services aux entreprises

On a rapporté les emplois secondaires et de l'armée à la population active en 1954.

λ - le mode de vie et l'éducation

La taille urbaine est bien entendu laissée telle quelle et les populations de moins de 14 ans et de moins de 35 ans ont été rapportées à la population totale.

A la taille moyenne urbaine un autre indicateur a été ajouté, le pourcentage d'agriculteurs dans la population active en 1954. En effet, les départements fortement agricoles sont ceux dont le retard en consommation de services est le plus important. L'évolution de cet indicateur s'accompagne d'un rattrapage qui doit être

(1) Voir chapitre 9 section 3.

d'autant plus important que la sous-consommation de service (que traduit ce ratio) l'est. Il s'agit donc d'une sorte d'indicateur de décollage alors que la taille moyenne urbaine traduit plutôt, pour une autre strate de population, la poursuite d'un mouvement.

3.22 - Recherche d'un-sous modèle en taux

On trouvera page suivante le tableau synoptique des résultats obtenus. Il s'avère que dans l'approche en taux les coefficients de détermination sont légèrement inférieurs à ceux de l'approche en effectifs. Ils n'en sont pas moins satisfaisants puisque tous gravitent aux alentours de 0,97 ; on peut d'ailleurs noter que les gains obtenus en ajoutant des variables exogènes sont médiocres puisqu'entre les régressions optimales à 2 et 5 variables le gain n'est que de 0,005.

Les deux variables qui interviennent dans toutes les régressions sont le taux de tertiaire calculé, comme on pouvait s'y attendre, et le taux d'agriculteurs ce qui était moins prévisible. Il semblerait, si notre hypothèse est bien fondée, que le phénomène de "take-off" joue un rôle important dans l'évolution du tertiaire.

Ensuite intervient la taille moyenne urbaine mais avec un poids assez faible dans l'explication de la variable endogène. Elle permet toutefois d'avantager les départements possédant les quelques plus grosses agglomérations et dans lesquelles s'installe de préférence le "quaternaire".

A partir de la régression de rang 4, les données de taux d'activité sont prises en compte mais avec un signe qui infirme totalement l'hypothèse avancée.

On retiendra donc comme meilleure régression en taux celle à 3 variables explicatives.

...

RECHERCHE D'UN MODELE EXPLICATIF DU TAUX DE TERTIAIRE 1962⁽¹⁾

nombre de variables rentrant dans la régression optimale	coefficient de détermination	taux de tertiaire 62 calculé ⁽²⁾ ‰	taille moyenne urbaine en 1954	Taux d'activité hommes 20-65 ans 1954 ‰	Taux d'activité femmes 20 - 65 ans 1954 ‰	coefficient de variation 54-62 du taux d'activité hommes ‰	Population totale 1954 (/1000)	actifs agricoles 1954/Population active ‰	Autres variables	Constante ‰
2	0,96965	0,97735 (40,3)						0,01351 (3,0)		- 2
3	0,97121	0,94348 (33,1)	0,000001 (2,1)					0,01260 (2,8)		2
4	0,97296	0,93098 (32,9)	0,000001 (2,7)	- 0,07523 (- 2,3)				0,01576 (3,5)		71
5	0,97440	0,95603 (35,4)			- 0,01610 (- 2,3)	0,08436 (2,8)	0,00334 (2,9)	0,02483 (4,7)		- 79

(1) taux de tertiaire = $\frac{\text{tertiaire}}{\text{Population totale}} = X$; tertiaire = CAE 6.1 à 9.3

(2) $X(D, 62) = X(F, 62) - (X(F, 54) - X(D, 54))$

3.3 - Le sous-modèle tertiaire retenu

On examinera successivement le choix de ce sous-modèle, les résultats obtenus pour 1962 puis ceux obtenus pour 1968.

3.31 - choix du sous-modèle

On est donc amené à choisir entre les deux modèles suivants :

$$(I) \text{ (tertiaire 62)} = 1,00727 \text{ (Tertiaire 54)} + 0,00903 \text{ (Population 54 } \langle 35 \text{ ans)} \\ + 0,00926 \text{ (taille moyenne urbaine 54)} + 0,11557 \text{ (variation 54 - 62 POP TOT)} \\ + 1224$$

$$(II) 1000 \frac{\text{tertiaire 62}}{\text{POP TOT 62}} = 0,94348 \text{ (Taux de tertiaire calculé . 1000)} + 0,000001 \text{ (taille moyenne urbaine 54)} \\ + 0,01260 \left(1000 \frac{\text{Agriculteurs 54}}{\text{Population active 54}} \right) + 2$$

Pour comparer l'approche en effectifs et celle en taux, on a calculé comme pour les sous-modèles précédents la variance résiduelle du sous-modèle en taux après transformation en effectifs (la référence étant constituée par les effectifs observés du tertiaire en 1962). L'approche en effectifs donne une variance résiduelle de $4 \cdot 10^6$ et celle en taux de $6 \cdot 10^6$ - ce qui est dans les deux cas excellent puisque l'écart-type correspondant représente moins de 2 % de l'effectif moyen départemental du tertiaire. D'après ce critère, que l'on a retenu jusqu'alors, c'est le modèle en effectifs que l'on devrait choisir. Toutefois, ce dernier modèle présente comme variable explicative la variation de population totale et une utilisation en mode prévisionnel nécessite la connaissance de la variation 62 - 70 de la population totale. Deux attitudes sont alors possibles : la première consiste à prendre une extrapolation 62 - 70 de la variation observée 62 - 68, afin de pouvoir effectuer une analyse des écarts imputables au seul sous-modèle tertiaire, ce qui limite l'interdépendance des modèles et appauvrit la notion de potentiel ; la seconde consiste à prendre la variation calculée par le sous-modèle de détermination de la population totale, ce qui mène à un système de Cramer de causalité de ces sous-modèles mais rend l'analyse des écarts plus délicate.

Compte tenu de la faible différence de "qualité" des deux approches, il semble plus simple de prendre l'approche en taux qui n'intègre pas de variables relatives à la période 62 - 68 autre que la prévision France entière du Tertiaire.

3.32 - Examen des résultats 62 du sous-modèle retenu

On trouvera pages T 12.48 à 54 cartes et tableaux de résultats obtenus, ainsi que page A 17.

Les erreurs relatives commises pour 1962 sont toujours inférieures à 9 % (2,6 % en moyenne) et l'on trouve 11 départements seulement ayant une erreur relative extérieure à la fourchette + 5 % ce qui explique la "platitude" des cartes confectionnées à l'aide des seuils habituels. Les départements sous estimés se trouvent dans les Alpes du Nord (Savoie, Haute Savoie, Isère et Hautes Alpes) et s'y ajoutent l'Oise et la Haute Garonne dont on a déjà souligné le comportement spécial.

L'écart-type résiduel étant faible (pour une fois!) l'approche probabiliste garde un caractère opérationnel plausible. On peut ajouter que la variable exogène 1 explique toujours plus de 90 % de la variable endogène, la variable exogène 2 ne dépasse 1°/oo que pour 14 départements et 5 °/oo pour 3 seulement, la variable 3 explique de 0 à 77 °/oo tandis que la constante voit son poids osciller entre 7 et 21 °/oo.

3.33 - Examen des résultats 68 du sous-modèle tertiaire

Les données utilisées sont celles relatives à l'année 1962 pour les variables exogènes 2 et 3. La variable exogène 1 dont le poids est prépondérant dans la régression a été calculée pour 1968 et non 1970, et les coefficients de régression laissés tels quels. Cette façon de procéder ne devrait pas être génératrice de trop de distorsion, dans la mesure où comme on l'a vu la variable 1 explique en moyenne 95 % de la variable endogène et que le coefficient correcteur qui aurait dû porter sur le coefficient de régression l'est en fait sur la valeur intégrée sur la variable exogène, ce qui globalement ne biaise pratiquement pas le résultat. On trouvera pages 67 à 70 les tableaux et cartes à la base de ces commentaires.

Les résultats obtenus sont assez satisfaisants⁽¹⁾ car on observe que 5 départements ayant une erreur relative extérieure à l'intervalle + 5 % et dans le pire des cas cette erreur est de 8 % (surestimation dans la Mayenne et sous estimation dans la Savoie). La transformation de ces résultats de taux en effectifs en y appliquant la population totale observée en 1968 mène à une variance résiduelle légèrement inférieure à celle trouvée

(1) Voir cartes page A 18

sur la période précédente ($6 \cdot 10^6$) mais lorsque l'on utilise les taux calculés pour 1968 après correction par les coefficients de spécificité, cette variance résiduelle double. Ceci ne doit pas tellement étonner dans la mesure où le resserrement des disparités noté dans l'étude descriptive va à l'encontre de la stabilité postulée par l'utilisation de ces coefficients.

Là encore, l'examen des cartes établies ne révèle pas grand chose si ce n'est que l'utilisation des coefficients de spécificité transforme de nombreuses disparités faibles en disparités accentuées.

**Le sous-modèle secondaire:
étude des facteurs de localisation**

CHAPITRE 13 - LE SOUS MODELE SECONDAIRE : ETUDE DES FACTEURS DE LOCALISATION

Comme nous avons déjà eu l'occasion de le signaler,⁽¹⁾ cette partie de notre travail n'a pu être menée à son terme, parce que l'élaboration des données de base (traitements simultanés des fichiers des établissements et des entreprises de l'INSEE) n'a pas été achevée pour des raisons indépendantes de notre volonté. Il nous a semblé malgré tout intéressant de présenter sommairement ce que nous avons l'intention de faire, en, en situant notre démarche par rapport à quelques travaux-type accomplis dans ce domaine. Nous n'avons pas voulu, compte tenu de l'objectif limité que l'on s'est fixé, allourdir cette présentation par un panorama complet des études effectuées dans ce domaine (et en particulier à l'étranger). Examinons donc successivement quelques travaux anciens et notre approche du phénomène.

SECTION 1 - LES ANALYSES DE L'EMPLOI INDUSTRIEL ORIENTÉES VERS LA PREVISION SPATIALISEE

Toute analyse des emplois industriels passe, du moins implicitement, par la notion d'établissement qui permet de stratifier ces emplois, par secteurs d'activité. Il est bien certain que l'évolution des effectifs d'un secteur d'activité dans une zone ne peut se comprendre qu'à partir d'une référence explicite à l'employeur qui est l'établissement ; or c'est dans le secondaire que les effectifs s'ajustent le plus strictement aux besoins, sous la pression du jeu de l'économie de marché. Cette référence à l'employeur peut se faire de deux façons différentes. Il est tout d'abord possible de n'accepter de s'intéresser à l'emploi qu'en référence aux établissements. Dans cette optique, l'unité de base est l'établissement, et l'ensemble de ces établissements est stratifié par tranche d'effectifs. La connaissance de l'évolution de cette distribution d'effectifs et de celle du nombre d'établissements permet de faire de la prévision ; c'est comme nous le verrons une approche que l'on peut qualifier de démographique. La seconde démarche possible n'introduit l'établissement que par le biais des facteurs de l'évolution et c'est en fait, à l'ensemble des établissements d'une zone que l'on s'attache ; pour cette raison nous qualifierons cette démarche de globale. Ces deux démarches, différentes dans leur esprit, se sont forgées des outils de prévision différents eux aussi.

1.1 - La démarche démographique

A la différence des trois études de localisation ponctuelle que nous avons vues au chapitre 4⁽²⁾, on s'intéresse ici à l'ensemble des établissements se trouvant dans une zone au cours d'une période, et non aux seules

(1) Chapitre 5 page 134.

(2) Page 97 et suivantes

implantations nouvelles. Il nous semble intéressant de poser le problème dans sa généralité car, pour un certain nombre de raisons, les études qui procèdent de cette démarche, n'ont abordé la question que de façon restrictive.

1.11 - Les bases de la démarche démographique

La démarche est d'inspiration démographique : l'analyse porte sur une population d'établissements qui naissent, vivent un certain temps puis disparaissent. Ce qui caractérise la démographie c'est la très grande inertie des phénomènes étudiés et la possibilité de jouer sur la loi des grands nombres. De ce fait, la prévision s'en trouve singulièrement simplifiée car elle peut s'opérer à l'aide de paramètres judicieusement choisis tels que les pyramides d'ages, les tables de mortalités et de fécondité.

La stabilité observée de ces paramètres permet la prévision et il importe peu, de ce point de vue, de connaître de façon détaillée les raisons de cette stabilité.

Ici l'approche se fonde sur l'étude de la stabilité des fréquences de passage d'une classe d'établissements caractérisés par un ensemble de modalités i à la date t , à un état caractérisé par l'ensemble de modalités j en $t + 1$. On peut en particulier avoir un caractère intitulé "age" dont les modalités seraient des classes d'age mais possédant une modalité "pas encore créé" et une modalité "cessation d'activité". On peut désigner cette matrice de changement d'état (E) par son terme général, qui s'analyse comme une fréquence conditionnelle :

$$P (E_{j, t + 1} / E_{i, t}) = f_{j/i} \quad \text{avec} \quad \sum_j f_{j/i} = 1 \quad \forall_{i, t}$$

S'il s'avère que l'on puisse déterminer les lois d'évolution des $f_{j/i}$ (ou que ceux-ci soient relativement stables) et que l'on connaisse, par une quelconque méthode, les flux annuels prévisionnels de créations d'établissements, la prévision se ramène à la résolution d'un problème markovien probabilisé (ou non), à condition bien sûr de disposer d'une distribution de fréquence.

Il faut toutefois souligner que pour que cette démarche conserve un caractère réaliste, il faut que le nombre d'établissements soit suffisamment important au regard du nombre de modalités retenues. Cette contrainte implique que le nombre de modalités soit nécessairement faible ; par ailleurs, elle exclut du champ de cette technique, l'étude de l'emploi secondaire dans une zone de petite ou moyenne importance (où les établissements sont en petit nombre). Enfin, elle écarte de l'analyse les très grands établissements du fait de la faiblesse du nombre des établissements, même si l'étude porte sur une zone importante.

On peut espérer que l'extension des facilités informatiques permettront d'explorer à fond ce type

d'analyse qui peut, malgré l'absence de modèles explicatifs, fournir des résultats prévisionnels intéressants. Examinons maintenant les travaux effectués sur la base de cette démarche.

1.12 - Les travaux de l'Institut et Etudes de l'Emploi

La constitution par J. VINCENS de fichiers d'établissements de la région Midi-Pyrénées, sur longue période, ont permis aux chercheurs de l'Institut d'Etudes de l'Emploi⁽¹⁾ de tester l'approche démographique du problème. Nous nous intéresserons ici à l'étude effectuée par P. KELLER et P. SIMULA en 1970.

Ce dernier a établi les distributions en 15 classes des variations de l'effectif des établissements créés avant 1954 et toujours existant en 1967, c'est à dire en excluant du champ de l'étude les entrées et sorties. L'étude porte donc sur 12 tableaux de contingence, reliés entre eux par l'identité de la marge horizontale du tableau n et de celle de la marge verticale du tableau n + 1.

Quelques régularités apparaissent nettement. Tout d'abord, la somme du flux de créations d'emploi et du flux de suppression d'emplois varie très peu, et même beaucoup moins que le solde (qui mesure la croissance de l'effectif global.) D'autre part, on assiste selon la conjoncture, à des permutations des fréquences marginales, par rapport à la classe de la variation nulle. L'expansion (ou la récession) ne se caractérise donc pas par une translation de la distribution.

L'étude des tableaux de contingence montre que quelque soit la conjoncture, il existe une forte liaison entre la variation d'effectifs pour un établissement au cours d'une période et celle de l'année précédente : en effet, dans la quasi totalité des cas l'établissement poursuit sa croissance (ou décroissance), on retrouve un effectif voisin de celui qu'il avait antérieurement.

Enfin, on peut remarquer que 70 % des distributions marginales sont reconduites d'une année sur l'autre, cette distribution minimum est appelée distribution de base tandis que la distribution résiduelle, liée à la conjoncture, est appelée distribution additionnelle.

P. Simula résume ainsi⁽²⁾ le passage de la description à la prévision "Ainsi, la conjoncture ne semble pas avoir une influence marquée sur le volume des variations : elle agirait uniquement sur le sens de la

(1) Revue économique n° 6, 1967 et n°1, 1968 . Voir surtout "Recherche sur la prévision à long terme, analyse de la transformation d'une population d'établissement industriels" Institut d'Etudes de l'Emploi janvier 1970.

(2) "Recherche sur la prévision à long terme" p. 101

variation. En réunissant l'ensemble des résultats établis, la relation entre deux variations consécutives pourrait s'exprimer de la façon suivante. Soit un ensemble d'établissements ayant subi au cours d'une année donnée une variation d'effectif donnée. Au cours de l'année suivante ces établissements vont se répartir en deux sous-ensembles : le premier sera composé des établissements, qui tendent à poursuivre leur mouvement de croissance, le deuxième sous-ensemble sera composé des établissements qui tendent, soit par un mouvement de recul, soit par un mouvement de hausse à recouvrer leur effectif initial. Le volume relatif de ces sous-ensembles varie sous l'influence de la conjoncture".

A partir de ces principes P. Keller et P. Simula ont ajusté deux distributions théoriques aux observations pour passer à la prévision. L'une donne la distribution des variations absolues des effectifs des établissements en croissance sur la période précédente, cette distribution est conditionnelle car elle dépend de la taille de l'établissement. L'autre distribution concerne les établissements qui étaient en décroissance. On associe à chaque distribution, une probabilité de signe, laquelle dépend de la situation conjoncturelle.

Par rapport aux techniques markoviennes citées au 1.11 le remplacement d'une matrice par des lois de probabilité facilite grandement la prévision mais il est vraisemblable que l'horizon de la prévision en soit quelque peu raccourci; de plus, ici, le problème des entrées et sorties est laissé de côté. L'utilité d'un tel instrument reste malgré tout indéniable mais un certain nombre d'approfondissements (fonction en partie de la disponibilité des données) seraient encore indispensables pour que l'on puisse disposer ainsi de techniques de prévision raffinées.

Examinons maintenant l'autre type de démarche.

1.2 - L'analyse globale menée dans l'optique prévisionnelle

Nous avons vu au chapitre 4, section 1, un certain nombre de développements théoriques consacrés en fait à l'approche globale. Certains outils utilisés, tel que le tableau inter-régional d'échanges inter-industriels, ne sont pas envisageables à cause de l'inexistence en France, des données requises. C'est donc à un niveau agrégé et sans prise en compte d'effets d'entraînement que le problème a été abordé et son caractère relativement frustré est surtout imputable au manque de données fines.

La démarche prévisionnelle se fonde alors sur l'analyse du passé et une formalisation théorique des causes d'évolution. Cette dernière était absente, nous l'avons vu dans la démarche démographique, ou plus

exactement l'hypothèse implicite était celle de la stabilité des causes d'évolution. Il nous semble intéressant de brosser rapidement ces études préalables à l'analyse prévisionnelle, avant d'examiner à titre d'exemple de la démarche d'une étude menée par le Centre d'étude et des techniques économiques modernes (CETEM).

1.21 - Les études descriptives préalables

1.211 - Les analyses descriptives de type global

Il est nécessaire dans l'analyse de l'évolution passée, de dissocier ce qui est imputable à l'évolution d'ensemble d'une branche de ce qui est spécifique à la région. Depuis de nombreuses années les spécialistes de l'économie régionale comparent la croissance d'une branche dans une région à celle de la branche pour l'ensemble du pays. En outre, certains affinent l'analyse en partant de secteurs d'activité détaillés, constitutifs de la branche et font apparaître un effet "structural" et un effet "régional". Le premier effet mesure la variation de croissance imputable à une répartition des activités B(I) de la branche B différente dans une zone à celle de l'ensemble du territoire, en supposant une croissance identique partout de ces activités B(I). Le second effet mesure le dynamisme relatif de la zone par rapport à l'ensemble du pays. Ce qui peut s'écrire comme suit en reprenant nos notations du chapitre 7 :

$$\left[\frac{X(D,B,T(1))}{X(D,B,T(0))} - \frac{X(F,B,T(1))}{X(F,B,T(0))} \right] = \left[\frac{X(D,B,T(1))}{X(D,B,T(0))} - \sum_I \frac{X(D,B(I),T(0))}{X(D,B,T(0))} \cdot \frac{X(F,B(I),T(1))}{X(F,B(I),T(0))} \right] \\ + \left[\sum_I \frac{X(D,B(I),T(0))}{X(D,B,T(0))} \cdot \frac{X(F,B(I),T(1))}{X(F,B(I),T(0))} - \frac{X(F,B,T(1))}{X(F,B,T(0))} \right]$$

Cette relation n'est pas sans rappeler du point de vue formel, celle de l'analyse de la variance.

(1) méthode préconisée par E. S. DUNN dans son étude sur l'évolution de l'économie américaine entre 1939 et 1954, voir "Regional Economic Growth in the United States" par PERLOFF, DUNN, LAMPARD & MUTH (John Peking University Press)

Un certain nombre d'études basées sur ces principes ou sur la technique des coefficients de spécificités⁽¹⁾ ont été menées en France et en particulier par l'INSEE⁽²⁾ au niveau des régions de programme. Le CETEM⁽³⁾, en ne s'intéressant qu'à un écart entre la croissance régionale et la croissance nationale d'un branche donnée, a bâti un certain nombre d'indicateurs assez sophistiqués pour mesurer globalement la mobilité des activités industrielles et l'on dispose avec cette étude des seuls travaux descriptifs détaillés portant à la fois sur le niveau départemental et sur les trois derniers recensements.

Le commentaire de cette étude descriptive ne s'impose pas compte tenu de l'objectif limité que l'on s'est fixé, aussi passerons nous à un résumé succinct des facteurs de localisation.

1.212 - La recherche des facteurs de localisation

Un certain nombre de facteurs sont pris en compte dans les diverses approches théoriques du problème de la localisation optimale⁽⁴⁾. Par ailleurs, un certain nombre d'enquêtes sur le sujet ont été lancées dès les années 40 et l'on dispose désormais dans certains pays d'une connaissance relativement précise des divers facteurs, mais leurs pondérations respectives restent mal connues. Ces résultats ne semblent pas très différents d'un pays à l'autre et il est vraisemblable qu'au niveau de l'énumération des facteurs, les résultats soient transposables à la France.

Pour présenter ces facteurs tels qu'ils apparaissent dans l'abondante littérature qui leur est consacrée⁽⁵⁾, il est nécessaire d'effectuer des groupements thématiques forcément arbitraires et divers auteurs proposent leur classement. Nous reprendrons ici l'un d'entre eux, proposé par le CETEM⁽⁶⁾, et qui groupe ces facteurs en 4 thèmes :

- la demande (tant pour les biens que pour les services) et les prix (jeu des élasticités)
- le progrès technique (qui modifie les fonctions de production)
- la structure de l'industrie et surtout le facteur taille (qui joue sur l'horizon économique et les économies d'échelles)
- Le comportement des agents économiques (inertie au changement,...)

(1) Voir chapitre 7, page 152 et suivantes.

(2) Espace économique Français fascicule II (Imprimerie Nationale - Presses Universitaires de France 1967) page 37 et suivantes

(3) "La mobilité des activités industrielles" p70 et suivantes (CETEM 1970)

(4) Voir chapitre 4 section 1 et chapitre 5, section 1.

(5) Voir en particulier SMITH "Industrial Location" (John Wiley & Sons - 1971)

(6) Voir l'analyse proposée par MM. AYDALOT, BESSON, GUIGOU et REBOULLET dans "la mobilité des activités industrielles"

Le problème de fond qui reste est celui de la pondération de ces divers facteurs. Un système de pondération peut être trouvé par les coefficients budgétaires afférents à chaque poste sur lequel le facteur joue et permet de mesurer des "sensibilités" comparées entre divers facteurs. En fait, ces critères peuvent ne pas correspondre à la réalité, le caractère de "condition permissive d'implantation" qui peut être attaché à un facteur donné présente un caractère bernouillien qui échappe totalement à la notion de prix. Aussi est-il sans doute préférable de rechercher comment s'expriment objectivement ces comportements sur un ensemble d'établissements pour retrouver des coefficients de pondération utilisables en mode prévisionnel. C'est cette démarche de détermination des facteurs et de leur pondération que nous allons examiner à présent.

1.22 - La recherche d'un modèle prévisionnel

Nous examinons tout d'abord les diverses démarches analytiques utilisées jusqu'ici afin de mieux situer celle de l'étude du CETEM.

Depuis quelques années, le développement simultané de l'informatique et des techniques d'analyse des données a modifié de façon assez sensible la démarche des économistes dans la recherche des facteurs explicatifs de la localisation industrielle. On peut distinguer sur le plan méthodologique trois types de démarches qui, chacune d'entre elles, utilisent la technique de régression multiple pour traduire des relations économiques.

α) La première consiste à passer au crible d'une analyse factorielle,⁽¹⁾ un ensemble de variables jugées explicatives, de préférence en excluant la variable que l'on cherche à expliquer.

Diverses techniques sont appliquées selon les problèmes, mais sont en fait les cas particuliers d'un modèle général.⁽²⁾ L'analyse en composantes principales s'opère à partir d'une matrice de variables centrées, ou centrées réduites et la distance retenue est une distance euclidienne classique, l'analyse des correspondances porte sur un tableau de correspondance et la distance est une distance du χ^2 . L'analyse factorielle en facteurs communs et spécifiques, qui pose, elle, un modèle à priori, ne se ramène pas au modèle général, mais elle ne présente ici, pas d'intérêt. Nous laissons aussi de côté diverses autres méthodes utilisées en analyse des données (celle des nuées dynamiques, l'analyse logit, l'analyse probit entre autres) car à notre connaissance, elles n'ont pas encore été utilisées dans l'étude du problème qui nous intéresse.

(1) "Statistique et informatique appliquées" par L. Lebart et JP. Fenelon (Dunod 1971)

(2) Voir L. Lebart et JP. Fenelon page 196.

Les axes factoriels dont la représentativité se mesure par le pourcentage d'inertie expliqué, sont des combinaisons linéaires de toutes les variables (ou des observations). Les variables qui ont le plus de poids dans cette combinaison linéaire ont une représentation graphique sur les plans factoriels, bien particulière : leurs abscisses par rapport à l'origine des axes (centre de gravité du nuage) sont les plus importantes⁽¹⁾. Par ailleurs, les variables corréllées sont voisines sur le plan factoriel. Il est facile à partir de ces deux propriétés de "sortir" de l'ensemble de variables celles qui provoquent la plus grande dispersion du nuage de points. Cette extraction, il faut le souligner ne peut que s'effectuer à partir de critères arbitraires dont le bien fondé dépend du "doigté" de l'analyste. Ces variables sont alors intégrées dans une régression multiple de type classique.

Deux remarques s'imposent : tout d'abord lorsque les phénomènes que l'on cherche à étudier sont représentés par un nombre très inégal dans la matrice de départ, les phénomènes sur-représentés attireront à eux les axes factoriels provoquant ainsi un biais qui peut être important dans la sélection des variables. Ensuite on peut dire que s'il est vraisemblable que les variables qui se trouvent très près du centre de gravité n'ont aucune chance d'appartenir à la régression multiple optimale, il n'y a par contre *aucune raison théorique* pour que les variables retenues par l'analyse factorielle parmi l'ensemble des n initiales correspondent au sous-ensemble optimal du point de vue du critère de régression multiple⁽²⁾: en effet certaines variables qui avaient moins de poids dans la "réduction" du nuage peuvent être en fait mieux explicatives dans la régression⁽³⁾. Le passage de l'analyse descriptive à l'analyse explicative est donc loin d'être rigoureusement assuré et l'utilisation conjointe de ces méthodes d'analyses de données et de régression nécessite un grand doigté et malheureusement la non prise en considération de ce "no-bridge" a biaisé de façon rédhibitoire un certain nombre d'études.

β) La seconde consiste à utiliser les techniques de régression multiple optimale ou, à défaut, de techniques plus frustrées telle que celle du stepwise. Ces techniques sont assez récentes et coûteuses en temps-machine ce qui explique que la première approche soit celle qui ait été la plus utilisée.

γ) La troisième consiste à expliquer notre variable endogène par une régression factorielle sur les premiers axes définis par une analyse en composantes principales⁽⁴⁾ en ne retenant finalement que les axes les plus significatifs pour la régression. Si l'on donne aux composantes principales une valeur d'indice on s'arrête là, sinon on opère une itération complémentaire en utilisant une technique de régression multiple optimale sur le sous-ensemble des variables expliquant le mieux les axes retenus. La différence avec la première méthode réside dans la détermination des axes puisque l'on ne retient pas obligatoirement les premiers.

(1) Cette abscisse est proportionnelle au cosinus-directeur de l'axe factoriel par rapport à l'axe de la variable.

(2) Voir chapitre 7 page 165 et suivantes.

(3) On peut ajouter que la distance retenue dans l'analyse en composantes principales et celle en facteurs communs est euclidienne, ce qui rapproche ces analyses de la régression orthogonale (au sens français du terme). Il existe des liens entre la régression classique et la régression orthogonale par l'intermédiaire des tangentes des angles formés par les hyperplans de régression et les axes des variables explicatives.

(4) Ce qui ramène au cas simple décrit au chapitre 3 page 21, § 2.431.

1.222 - Un exemple de l'approche globale : l'étude du CETEM

Une équipe du CETEM, sous la direction de P. AYDALOT s'est attaquée à l'étude⁽¹⁾ de la mobilité des activités industrielles en se fixant deux objectifs. Il s'agit tout d'abord de mesurer l'intensité et l'orientation spatiale du déplacement des 18 secteurs d'activités du secondaire sur diverses périodes intercensitaires, en prenant pour référence la situation fictive d'une croissance d'un secteur, identique dans tous les départements. Le second objectif est de chercher à expliquer ces déplacements.

Atteindre ce dernier objectif suppose que des éléments de réponse soit apportés à deux types de questions : comment se détermine le niveau de la mobilité industrielle et quelles sont les conséquences de cette mobilité sur la structuration de l'espace. Du point de vue méthodologique la démarche suivie est la première de nos trois démarches avec utilisation de l'analyse factorielle des correspondances et sans utilisation des techniques de régression optimale ou de stepwise. Toutefois, cette étude présente certaines particularités que nous soulignerons en temps utile. Examinons succinctement les deux parties de cette étude en nous bornant à l'aspect méthodologique et aux conclusions.

a) les facteurs explicatifs de l'intensité de la mobilité

Une première analyse factorielle des correspondances a été menée sur une matrice branches x départements (18 x 89) d'un indicateur Q'_{bd} qui se définit comme un coefficient de spécificité départemental du coefficient de variation d'une branche B, pondéré par la part des effectifs "France Entière" de cette branche dans le secondaire. Ces premiers axes factoriels prennent alors la signification d'indicateur du phénomène et la variable définie par les abscisses des projections des 18 points-variables est considérée comme variable endogène dans la seconde étape de l'analyse. L'explication des trois premiers axes factoriels pour chacune des périodes pour un ensemble de variables explicatives spécifiques pour chaque branche et chaque facteur (la dimension spatiale ayant disparu) permet de leur donner une interprétation économique.

Ces variables explicatives ont été retenues à partir de l'analyse théorique des facteurs de localisation et sont de trois types : facteurs de production, comportement spatial et structures techniques.

Le premier axe s'explique par des données de croissance et le second par des données structurelles.

Dans une troisième étape, on recherche à affiner les orientations grossières fournies précédemment en cherchant des régressions satisfaisantes pour deux indicateurs : le premier est la marge obtenue sur la matrice

(1) rapport cité

Q' par sommation sur les départements, le second, qui revêt une réalité plus tangible, s'analyse comme la part des travailleurs de la branche B qui ne travaillent pas dans le département D, où ils auraient été occupés si la croissance de la branche avait été la même partout.

Les auteurs de cette analyse résument leur conclusion ainsi : "La mobilité paraît liée à la croissance, au déséquilibre et aux phénomènes structurels qui modulent et limitent les tendances à l'ajustement, au mouvement. Les deux premiers facteurs créent la tendance à la mobilité, les derniers la bloquent et perpétuent ainsi les déséquilibres que la mobilité permettrait de réduire".

β) Les facteurs explicatifs de l'orientation spatiale de la mobilité

L'étude est là encore, menée en terme relatif : la variable expliquée retenue est la part des travailleurs qui, dans un département, ne travaillent pas dans la branche où il auraient été occupés, si la croissance avait été la même dans toutes les branches du département. Cette étude est complétée pour chaque branche par l'introduction de la variable "écart départemental du pourcentage de travailleurs du secondaire travaillant dans la tranche B entre 1962 et 1968". Les variables explicatives départementales introduites étaient de 8 types : indicateurs de transport et télécommunications, de proximité de marché, de niveau d'éducation, de composition socio-professionnelle de la population, de marché du travail, d'attraction du département et la cas échéant, des indicateurs spécifiques aux branches.

Les diverses régressions multiples tentées semblent montrer que le phénomène est en fait très complexe : certains facteurs jouent un rôle différent selon les périodes. Par ailleurs, la différenciation dans l'espace des structures régionales paraît jouer un rôle moins important dans le déplacement des branches que la différenciation des structures internes aux branches.

Il nous semble au terme de cette étude que le passage de la description à la prévision est assez délicat bien que possible à partir des équations fournies, au prix de l'introduction d'un grand nombre d'hypothèses simplificatrices.

Examinons maintenant la démarche que nous avons retenue.

SECTION 2 - RECHERCHE D'UN SOUS-MODELE DE DETERMINATION DES EMPLOIS INDUSTRIELS

Examinons successivement la démarche retenue et l'analyse des facteurs de localisation.

2.1 - Présentation de la démarche retenue

Comme nous l'avons indiqué au chapitre 5⁽¹⁾ l'analyse de système menée conduit à un clivage entre emplois fondamentaux et emplois induits. Cette notion sera affinée dans un premier paragraphe, puis nous verrons les principes d'analyse que nous avons retenus.

2.11 - Stratification des emplois industriels

Les emplois industriels sont considérés dans leur ensemble comme fondamentaux parce que les biens produits sont consommables a priori n'importe où ; c'est à dire que, au problème près du coût du transport, la vente des biens produits ne dépend pas de la zone dans laquelle la production s'effectue. Toutefois, comme on l'a souligné au chapitre 4, la localisation du marché potentiel a une incidence sur la compétitivité des établissements, en particulier par le truchement des différences de coût de transport.

A l'occasion d'investissements nouveaux ou de remplacement, les responsables d'entreprises peuvent être amenés à reconsidérer pour diverses raisons, une localisation qui traditionnellement s'opérait dans une zone donnée. Ces raisons peuvent être de deux types : certaines sont liées à la nature de l'emploi, d'autres à la région.

Les facteurs liés à la nature de l'emploi sont à rechercher principalement dans la définition des processus productifs. Ceux-ci par la taille des investissements qu'ils impliquent (pour atteindre des économies d'échelle nécessaires pour rendre l'entreprise compétitive) peuvent entraîner que les variations du potentiel productif (pris au sens large) ne s'opèrent que par "sauts" importants.

(1) page 133

En outre la localisation des nouvelles unités productives peut réquerir la réunion d'un certain nombre de conditions permissives d'implantation telle que la disponibilité d'eau, la présence d'infrastructures de transport, ... Il en résulte que seul un certain nombre de localisations est techniquement possible. L'implantation du centre de décision joue alors un rôle fondamental dans le choix final de la zone où l'investissement sera effectué.

Les facteurs liés à la région sont de plusieurs sortes : il y a tout d'abord des effets d'entraînement qu'exercent certains établissements, moyens ou gros, sur des petites entreprises dont le marché reste essentiellement local et dont la vocation est la production de biens non finaux. Ensuite, on trouve des établissements qui produisent des biens finaux et correspondent à une version industrialisée de petites manufactures. Dans un cas comme dans l'autre la création de ces établissements relève d'une décision locale et est en grande partie fonction des capacités locales de financement.

Il ressort de ces considérations qu'il serait souhaitable de tenir compte de la notion d'étendue du marché et de celle du rattachement à la zone du centre de décision pour parfaire les études menées jusqu'ici, ces deux critères étant d'ailleurs en grande partie liés. Pour cerner correctement la notion d'étendue du marché il faudrait disposer de tableaux inter-régionaux d'échanges inter-industriels, ce qui, nous l'avons déjà dit, est actuellement exclu. Nous nous bornerons donc à essayer de saisir statistiquement la notion de centre local de décision en espérant que l'introduction de cette stratification révélera un certain nombre de régularités utilisables pour la prévision. On distinguera donc *les emplois autonomes* c'est à dire rattachés à des établissements pour lesquels la décision s'effectue localement, des *emplois non autonomes*.

Le centre de décision est au siège social de l'entreprise à laquelle appartient l'établissement, on prendra comme critère de l'autonomie de la décision le fait que l'établissement se trouve dans le même département que l'entreprise. Dans la mesure où la population active est rattachée non pas à son domicile mais à son lieu de travail, un certain nombre d'anomalies imputables aux migrations alternantes devraient disparaître. Ne devraient subsister que celles qui sont dues aux agglomérations interdépartementales. Par ailleurs, pour tenir compte de la notion d'étendue du marché il semble utile de prendre comme proxis, la notion de taille des établissements. C'est pourquoi la stratification que nous retiendrons pour cette étude comporte deux groupes : la population des emplois non autonomes rattachés à des établissements ayant plus de 50 salariés (seuil arbitrairement choisi) et les autres emplois. Cette partition sera bien entendu employée simultanément avec celle des secteurs d'activités INSEE.

Une exploitation du fichier INSEE des établissements de 1961 à 1969, intégrant des renseignements tirés du fichier des entreprises doit nous permettre d'obtenir les renseignements demandés. Nous devons ajouter

qu'une étude de type exploratoire a été menée par D MALKIN⁽¹⁾ sur le fichier d'aide de la DATAR avec la prise en compte des facteurs création ou extension et d'une partition de l'espace en fonction des taux de prime. Les résultats encourageant de cette étude incitaient le SESAME à poursuivre dans cette voie mais en partant d'une population non biaisée et en raisonnant sur des variations globales (et non plus sur les seuls flux de créations et d'extensions).

Le traitement informatique des deux fichiers n'a pas été mené à son terme par l'organisme qui en était chargé. Indiquons malgré tout les traitements que nous comptons effectuer sur ces données.

2.12 - Principes d'analyse retenus

Notre démarche par rapport à celle décrite dans la section 1 est de type global. Toutefois, la disposition de séries temporelles et la stratification adoptée permet deux choses qui n'étaient pas possibles jusqu'ici : une analyse dynamique sur une courte période (3 ans) et une étude des variantes annuelles d'emplois non autonomes.

2.121 - Analyse dynamique sur une courte période

L'inconvénient de l'utilisation des périodes intercensitaires est que les indicateurs statiques sont supposés stables tout au long de la période, ce qui est une hypothèse pour le moins contestable. Il est possible ici de raisonner sur deux périodes assez courtes et centrées sur les recensements de 1962 et 1968. L'hypothèse de la stabilité des indicateurs statiques (le plus souvent tirés des recensements) devient alors acceptable.

L'analyse portera sur la variation des emplois par strates, telles qu'elles sont définies ci-dessus, et les variables explicatives correspondent au corps d'hypothèses implicitement posé par les différentes études de (2) localisation portées à notre connaissance et plus particulièrement celles du CETEM et celles décrites au chapitre 4. L'outil d'analyse utilisé sera celui de la régression optimale (que nous avons utilisé tout au long des études.

2.122 - Une étude des variations annuelles des emplois non autonomes

Il semble intéressant de compléter ci-possible l'approche précédente par une vision dynamique à

(1) Etude sur les facteurs de localisation des emplois industriels (DATAR - SESAME décembre 1971)

(2) Il s'agit des modèles d'implantation industrielle du BASS, du AAM et du modèle de Hamilton décrits au chapitre 4, page 97 et suivantes. On peut signaler en outre l'intéressante étude de JP. JAEGER "Approche de la localisation industrielle en France à partir d'une analyse factorielle des structures régionales" (La Revue juridique et économique du Sud-Ouest - 1971 n° 3 page 364 et suivantes) mais qui est menée au niveau des régions de programme.

long terme pour chercher à expliquer des renversements de tendance qui pourraient se produire dans la distribution spatiale de ces variations (après avoir éliminé les causes d'évolutions sectorielles valables pour l'ensemble du pays). Ceci présente un double intérêt, d'une part cette étude peut permettre de mettre en évidence l'apparition progressive de seuils de saturation, et d'autre part, permettre de nuancer l'étude précédente qui relève de ce qu'on pourrait appeler l'analyse dynamique comparative. L'étude porterait sur la liaison entre des trends et des stocks, avec des comparaisons entre branches.

2.2 - Analyse des facteurs de localisation

(Pour mémoire)

CHAPITRE 14 - RESULTATS GLOBAUX - CONCLUSIONS

On dispose maintenant des résultats des différents sous-modèles, il faut tout d'abord obtenir une vision d'ensemble, c'est-à-dire calculer l'état potentiel global puis rechercher une typologie des ajustements qui permettraient de comprendre le passage de cet état potentiel à la réalité observée en 1968. Dans une dernière section on s'efforcera de dresser le bilan de ces travaux.

SECTION 1 - RESULTATS d'ENSEMBLE

Pour obtenir l'état potentiel quelques options complémentaires sont à prendre, que l'on examinera avant de faire les premiers commentaires d'ensemble de ces résultats.

1.1 - Options complémentaires

Certaines options concernent la connaissance de données non déterminées par le modèle et les autres concernent le problème de cohérence globale des divers sous-modèles

- a) Options sur la connaissance a priori de certaines données
- α) Certaines données sont considérées comme connues sans erreurs à titre de simplification.

Il s'agit tout d'abord de l'ensemble des emplois du secondaire qui restent en dehors du champ de détermination de l'emploi potentiel comme on a pu le voir au chapitre 13.

Les effectifs des militaires de carrière, qui ont été exclus de la définition du tertiaire, sont supposés déterminés et répartis spatialement sans erreur par les services de prévision du Ministère de la Défense Nationale, d'ailleurs la tâche de ces services est singulièrement facilitée par le caractère contraignant des ordres de mutation et la possibilité de bloquer tout recrutement pour diminuer les effectifs (par le seul jeu de retraites).

Enfin, les effectifs de la CAE pêche sont de par la nature de l'activité exercée, géographiquement localisée. Il ne saurait être question de rechercher un sous-modèle "pêche" sur l'ensemble des départements (ce qui est de règle pour tous les sous-modèles recherchés), en outre, il s'agit là d'un domaine où le rôle de la concurrence internationale est très fort, la prévision ne peut donc relever pour ce secteur que d'un modèle très spécifique.

- β) En ce qui concerne la définition des effectifs de tertiaire on a multiplié le taux de tertiaire potentiel par la population potentielle et non celle observée, ce qui semble davantage dans la logique de définition d'un état potentiel.
- b) Options relatives à la cohérence globale des divers sous modèles

Lorsque l'on compare dans un sous-modèle le cumul des valeurs départementales calculées avec la valeur observée correspondante (= France entière - Région Parisienne et Corse) l'écart est comme on a pu le constater à diverses reprises, assez important.

Or, il est vraisemblable que la détermination des effectifs nationaux d'une catégorie donnée est meilleure par un modèle national qui, lui, peut tenir compte de données supplémentaires qui n'existent qu'à ce niveau, telles que les complémentarités intersectorielles des évolutions de productivité et de données en valeur (en particulier de la demande finale intérieure et étrangère). En ce qui concerne les migrations intérieures le problème est un peu différent car se superpose un problème d'identité (somme des immigrants intérieurs = somme des émigrants intérieurs) à celui de la détermination du montant de ces migrants intérieurs ; le problème de l'identité est a priori insoluble car il n'y a pas de raisons particulières de privilégier l'aspect émigration plus que celui d'immigration. On n'est pas concerné dans le cas présent par ce problème d'identité puisque la Région Parisienne a été exclue du champ de l'étude pour des raisons de modifications de découpages géographiques. En ce qui concerne le volume de migrants, les travaux de COURGEAU sur les migrations multiples laissent penser qu'il est possible avec une erreur relative faible de le calculer pour une période et un découpage donné.

(1) de l'ordre de 10 % seulement si l'on avait considéré comme stable les diverses probabilités calculées sur la période 54 - 62

Il semble donc préférable de prendre la valeur calculée par un modèle national. Dès lors la cohérence entre modèle spatiaux et nationaux n'est plus assurée. Pour résoudre ce problème on peut par exemple cadrer à la proportionnelle les résultats des modèles spatialisés sur les modèles nationaux. Le procédé est admissible si l'écart n'est pas imputable à quelques départements isolés pour lesquels des facteurs explicatifs importants auraient été ignorés.

C'est cette façon de procéder qui a été retenue finalement et le cadrage des résultats (avec et sans coefficients de spécificité de régression) sur la valeur observée correspondante (Province) a été effectué pour chacun des sous-modèles (mais les résultats et graphiques que l'on va étudier ont été établis aussi dans ces facteurs correctifs de cadrage).

1.2 - Examen d'ensemble des résultats

Les écarts relatifs entre valeurs observées et valeurs calculées⁽¹⁾ ne sont que rarement supérieurs à 5 % ce qui laisse présager que des ajustements sont possibles sans remise en cause fondamentale des modèles. L'approche qui utilise les corrections pour les coefficients de spécificité ne donne pas de résultats meilleurs que celle qui ne prend que les estimations ponctuelles. Enfin on peut noter que l'écart entre offre et demande potentiels n'excède pas non plus 5 % sauf dans 2 départements.

Un certain nombre de régularités observées résultent d'écarts en cascade : c'est ainsi qu'une sous-estimation de la population totale entraîne une sous estimation de demande d'emploi (à supposer les taux d'activités potentiels et observés soient proches) ainsi qu'une sous-estimation de l'offre d'emploi tertiaire.

Pour ces raisons d'interdépendance, il est sans intérêt d'établir un tableau synoptique d'ensemble sans réflexions complémentaires sur les liaisons entre variables, réflexions qui mènent à l'établissement de typologie.

(1) Voir tableaux des pages A 19 et A 20

SECTION 2 - RECHERCHE d'une TYPOLOGIE d'AJUSTEMENT

La recherche de similitudes de "comportements" des divers départements n'est guère aisée. Deux démarches sont possibles pour résoudre ce délicat problème.

2.1 - L'analyse systématique des écarts

On peut tout d'abord procéder à une analyse fine des écarts et effectuer une classification des divers départements par "famille" d'écart. Mais le nombre de familles possibles croît géométriquement avec le nombre de sous-modèles à prendre en compte. Ici par exemple nous avons les résultats de 7 sous-modèles, si l'on se contente de 2 classes d'écarts (selon le signe) on obtient 2^7 familles possibles soit 128 familles, c'est-à-dire qu'il y aurait plus de familles que de départements à prendre en compte. Or il est difficile de se contenter de deux classes car un certain nombre de valeurs calculées sont très proches de la valeur observée correspondante. L'introduction d'une troisième classe mène à prendre en compte 2187 "familles" et il est vraisemblable que l'on obtiendrait pour chaque famille retenue 1 ou 2 départements seulement. Que dire alors de ce type de démarche dans un modèle définitif comprenant plusieurs dizaines de sous-modèles ne s'articulant pas forcément par des systèmes de Cramer de causalité.

Pour compléter malgré tout la description de ce type de démarche, il faut ajouter qu'il est nécessaire de trouver des indicateurs qui permettent de dissocier chaque famille en début de période afin d'être en mesure de déclencher le mécanisme d'ajustement approprié dans une utilisation prévisionnelle du modèle d'ensemble. En outre il faut que chaque mécanisme d'ajustement puisse faire l'objet d'une quantification correcte sur le passé et que le modèle théorique sous-jacent présente un caractère plausible.

Bref cette démarche n'a rien d'évident.

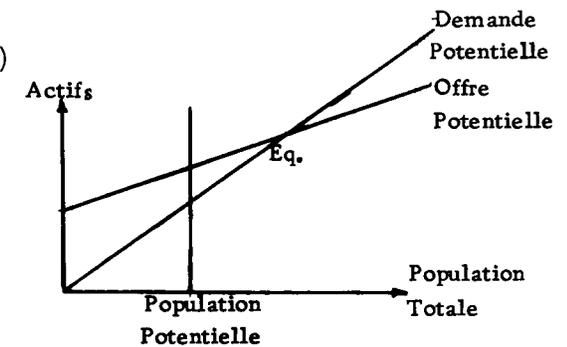
2.2 - L'appel à un modèle de référence

Il est possible aussi de raisonner à un niveau agrégé en faisant appel à un modèle de référence pour faire apparaître des similitudes. Le choix de ce modèle est quelque peu arbitraire et obère la démarche suivie et les résultats obtenus. On présentera l'un (1) de ces sous-modèles possibles (qui a finalement servi de référence).

L'état potentiel est défini par trois grandeurs :

- Population potentielle
- Demande d'emploi = (Population Potentielle). (Taux d'activité Potentiel)
- Offre d'emploi = (emploi exogène) + (BTP Potentiel) + (Agriculteurs Potentiels) + (Taux de tertiaire Potentiel). (Population Potentielle)

Si l'on part d'un système de référence "population totale - population active", l'état potentiel se traduit par deux droites où l'emploi est fonction de la population totale (2) et une droite de population potentielle (qui peut, bien entendu, se situer à droite ou à gauche du point d'intersection des droites d'emploi). Il est bien évident que cette représentation n'a pas grand sens lorsque l'on est proche de l'origine.



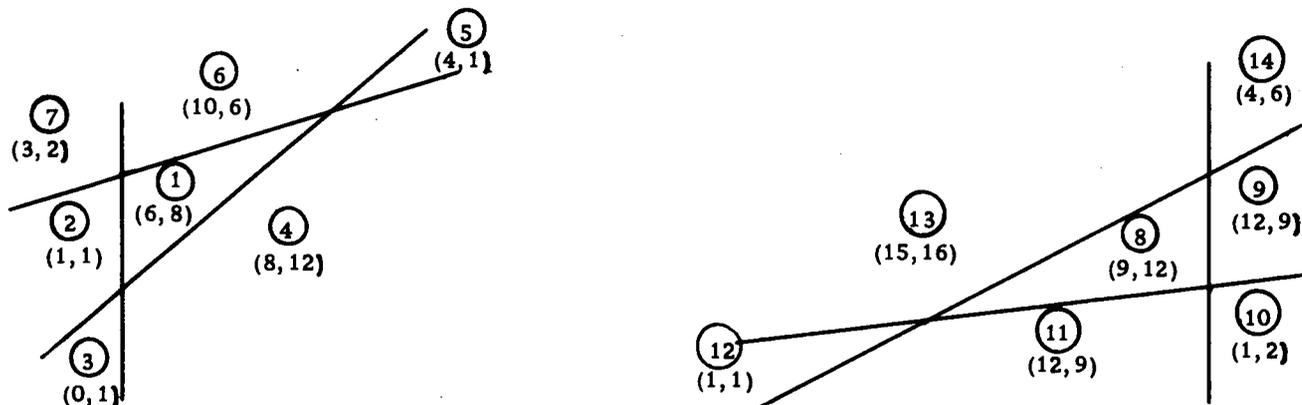
On peut interpréter l'intersection des droites de demande et d'offre comme un point d'équilibre puisque, à ce volume de population totale l'offre et la demande d'emploi coïncident. Par ailleurs le couple observé de population totale - population active peut se situer n'importe où sur ce graphique. C'est en raisonnant sur les différents cas de figure possibles que l'on définira les typologies possibles d'ajustement.

- (1) Elaboré conjointement avec MM. HORPS et MALKIN. Cette démarche est jugée comme l'une des moins mauvaise possible par leurs promoteurs sans pour autant leur donner satisfaction.
- (2) Ceci n'est possible que parce que l'on a retenu le sous-modèle tertiaire en taux bien que les résultats de l'approche en effectifs semblaient aussi bons. Si cette dernière approche avait été retenue il aurait fallu choisir un autre modèle de préférence

Pour rechercher quelques situations type assez facilement explicables, deux remarques préalables sont nécessaires :

- le volume d'emplois observé correspond à des emplois effectivement occupés, tandis que les taux d'activité calculés intègrent le chômage. On peut donc logiquement espérer que le point observé se situe en-dessous de la droite de demande potentielle d'emploi et qu'une partie du désajustement est imputable au chômage. Si le point se situe au-dessus il faudra chercher à expliquer les causes de cette superactivité ; celle-ci peut s'expliquer par une tension importante sur le marché de l'emploi c'est-à-dire que l'offre potentielle est supérieure à la demande, c'est-à-dire que le point observé doit alors se situer entre les droites de demande et d'offre
- si le point d'équilibre correspond à une certaine stabilité on peut espérer que la population observée se situe entre la population potentielle et celle d'équilibre mais pas en deçà si la population potentielle est inférieure à la population d'équilibre, ni au delà si les positions respectives de ces populations sont inverses. En effet cela implique que les mécanismes d'ajustement tendent à faire converger l'état potentiel vers une situation d'équilibre.

Pour opérer une classification des départements il faut définir différentes zones, en s'aidant des schémas suivants où les nombres cerclés repèrent chaque zone et où le couple de nombres entre parenthèses indique le nombre de départements pour lesquels le point observé se situe dans la dite zone, le premier pour l'approche sans utilisation de coefficients de spécificité des régressions et le second dans le cadre de cette approche avec utilisation de ces coefficients.



On trouvera dans le tableau synoptique de la page suivante la classification des départements dans ces diverses zones.

Les considérations précédentes conduisent à juger la situation d'un département comme satisfaisante si le point observé se trouve dans les zones 1, 4, 5, 8 et 9 ce qui ne se produit que dans un peu plus du tiers des cas. Mais il faut tout de suite ajouter que dans la majorité des cas restants, les points observés sont très proches des triangles de référence (- de 1% d'écart), ce qui est assez bon.

Toutefois on peut dire que la précision avec laquelle ces diverses données sont connues rend un peu vaine la recherche d'une typologie rigoureuse basée sur ces seules observations. En outre il ne semble pas, à l'analyse des résultats que les tensions révélées par l'état potentiel puissent être considérées comme le reflet de tensions s'étant effectivement produites sur la période 1962 - 1968 et de ce point de vue la démarche tentée ne permet pas de faire jouer à l'état potentiel un rôle de révélateur de tension. Enfin une réflexion sur la quantification des mécanismes d'ajustement devrait permettre pour les cas jugés satisfaisants de faire coïncider le potentiel au réel par des modifications de pentes, d'abscisses et d'ordonnées à l'origine ; mais la hiérarchie de ces trois types de modification résulte de décisions arbitraires, ou suppose de repartir d'une analyse fine des écarts, ce qui ramène à la démarche typologique initiale.

Pour toutes ces raisons, on ne poussera pas davantage une analyse qui ne pourrait mener qu'à des résultats sans grand intérêt parce que non objectivement fondés et c'est donc un constat d'échec de la confrontation des états réels et potentiels que l'on est obligé de faire.

CLASSIFICATION des DEPARTEMENTS SUIVANT la POSITION des POINTS REELS
par RAPPORT au SCHEMA de REFERENCE

Zone	Sans coefficient de spécificité de régression	Avec coefficient de spécificité de régression
1	2/70/43/80/28/48a/	79/65/43/80/5a/60/28a/48a/
2	39/	14/
3		23/
4	7a/9/11a/12a/36/40/71/66/	84/4/7/11/12/38a/85a/74/84a/32/35/66/
5	87/24/23/3/	26/
6	45/22/1/89/58/19/18/51/27/60/	72/1/37/19/51/27/
7	61/10/41/	76/21/
8	65/67a/35/31a/44a/47a/57a/54a/42a/	30/71/70a/59/33/62a/8a/44a/47a/57a/54a/42a/
9	6/26/84/4a/64a/68a/69a/38a/83a/34/59a/33/	87/6/24/64a/68a/69a/90/67a/25a/
10	46/	46/40/
11	13/29/56/85/90/81/74/32/62/8/25/53/	9/13/29/56/83/81/36/31/53/
12	5/	40/
13	76, 72, 79, 52, 17, 14, 55, 86, 82, 73, 21, 16, 15, 50, 49,	52, 17, 55, 45, 89, 88, 82, 63, 61, 58, 18, 16, 50, 49, 41, 39,
14	37/88/63/30/	22/86/73/15/10/2/

a) Emploi observé compris entre l'offre et la demande potentielle

SECTION 3 - CONCLUSIONS

Les résultats obtenus sont finalement assez décevants compte tenu de l'importance des efforts entrepris. Pourtant cette étude a bénéficié de plusieurs facteurs favorables : tout d'abord pour cerner progressivement les difficultés on s'est limité à des systèmes de Cramer de causalité et on a admis qu'un certain nombre de données étaient connues sans erreur ; la prise en compte de données observées a éliminé toute cause de désajustement supplémentaire. D'autre part le programme de régression optimale employé a permis d'utiliser au mieux les techniques de régressions multiples et de pouvoir effectuer sans trop de difficultés des choix corrects (sur le plan statistique) sur n'importe quel ensemble d'indicateurs. Enfin la base de données constituée et la bibliothèque de programme mise au point facilitaient grandement les opérations d'interprétation et de mise au point.

On peut donc dresser un bilan presque définitif de ce type d'approche à la fois sur les techniques économétriques et sur la modélisation entreprises.

3.1 - Remarques sur l'application des techniques économétriques utilisées

Différentes techniques ont été utilisées pour le choix du découpage et la détermination des lois d'évolution.

3.11 - Techniques relatives au choix du découpage

On a pu remarquer que ces techniques laissaient une part assez grande à l'arbitraire dans la mesure où les critères retenus le sont (mais on ne voit guère comment faire autrement) et où, dans les techniques de type probabiliste, l'appréciation du risque de seconde espèce est très subjective.

...

Mais ceci est indépendant de l'adéquation du ou des indicateurs retenus pour le problème. Or dans le cas présent le choix de l'indicateur est très délicat d'une part parce que l'aspect spatial de la définition du marché de l'emploi fait intervenir simultanément des données statiques et des données dynamiques (par le biais de mécanismes jouant au cours du temps) et d'autre part parce que dans la démarche heuristique suivie il ne pouvait être fait allusion à un ensemble de mécanismes théoriques pré-existants. (mais même s'il en avait été ainsi le problème de la pondération entre ces divers mécanismes serait resté entier). On peut ajouter que de toute façon la démarche est contrainte en partie par les statistiques disponibles.

Il semble donc que ce problème de la pertinence d'un découpage spatial à la notion de marché de l'emploi (ou mieux encore celui de la recherche d'un découpage spatial optimal) nécessite pour avoir une solution définitive qu'un certain nombre d'investigations complémentaires soient entreprises mais qui ne seront possibles qu'au prix d'un effort théorique complémentaire et de l'élaboration de données nouvelles.

3.12 - Remarques sur les techniques de préhension des régularités.

Une première remarque s'impose : la dispersion des résultats autour des hyperplans de régression est toujours très grande, que l'approche soit faite en taux ou en effectifs. Ceci s'explique par une distribution très étendue des valeurs départementales dont la courbe de répartition est fortement dissymétrique à gauche dans la plupart des cas. L'obtention de meilleures lois d'évolution passe donc par une stratification des départements (à supposer que le cadre départemental ne soit pas à remettre en question).

Cette stratification souhaitable sur le plan statistique remet en cause l'adéquation du découpage spatial adopté puisque le marché de l'emploi est supposé avoir des lois communes pour toutes les zones du découpage (dans le cas contraire il ne serait pas pertinent). La remise en cause de ce postulat jusqu'ici admis pose alors le problème de la justification théorique de tel ou tel seuil de stratification, celui de la perméabilité des groupes retenus entre eux et enfin celui de la réduction du nombre de degrés de liberté dans le test des hypothèses des modèles relatifs à chacun des groupes.

Par ailleurs, la dispersion observée n'est pas stable dans le temps, ce qui conduit à n'avoir aucune amélioration (sauf dans un cas) lorsque l'on utilise la technique des coefficients de spécificité. Ceci

...

peut être dû au fait que la loi d'évolution de ces coefficients que l'on a arbitrairement retenue (stabilité de ces coefficients) n'est pas valable et qu'il faille lui en substituer une autre (stabilité des écarts...) mais aussi peut être au fait que l'hyperplan de référence (estimé dans la détermination du modèle) n'est plus correct lors du calcul de l'état potentiel, en effet la validité de la loi d'évolution est subordonnée à celle de l'hyperplan.

La logique complète de la démarche probabiliste (laquelle repose on l'a vu au chapitre 7 sur des hypothèses complémentaires) conduit à utiliser la notion d'intervalle de confiance, mais l'écart-type résiduel est tel qu'il est hors de question de raisonner autrement qu'en estimation ponctuelle. Ce rejet est imputable en fait à la dispersion des valeurs départementales dénoncée ci-dessus. Dans ces conditions et sans rentrer dans des considérations théoriques complexes concernant l'intervalle de confiance des offres et demandes d'emploi(1) on peut se demander s'il est intéressant de rechercher une typologie des mécanismes d'ajustement alors qu'il est vraisemblable que les ellipsoïdes de confiance des offres et demandes globales pour un département englobent vraisemblablement le point observé et les données potentielles, et très certainement dans de nombreux cas, des données calculées par une simple règle de trois!

3.2 - Remarque sur la modélisation

3.21 - Critiques internes

Il y a tout d'abord une certaine contradiction à travailler sur une période intercensitaire, imposée par la disponibilité des données, et l'hypothèse d'indépendance des décisions sous-jacentes au calcul d'offre et de demande potentielle. L'état potentiel obtenu est très certainement dénaturé par rapport à ce qu'il devrait être et peut être aurait-il mieux valu, compte tenu des statistiques existantes, abandonner ce type d'approche. On peut alors, pour justifier la poursuite de ce travail exploratoire, arguer du fait que si les états calculés dans chaque sous-modèle ne revêtent pas la nature d'état potentiel et ne conservent que celle d'intermédiaire de calcul la comparaison du "produit de ces diverses régularités" avec la réalité n'en conserve pas moins un intérêt certain.

(1) En effet compte tenu de l'interdépendance des variables explicatives des divers sous modèles, cet intervalle de confiance ne correspond pas à la somme des intervalles de confiance des valeurs constitutives de l'agrégat.

On peut remarquer ensuite qu'un certain nombre de régularités observées sur le passé intègrent implicitement un ensemble d'actions politiques. L'utilisation, en mode prévisionnel de ces régularités revient à reconduire ces actions politiques ce qui explique sans doute un certain nombre de distorsions.

Enfin, on peut rappeler que l'absence de référence explicite à des modèles théoriques, nécessitée par l'approche concrète effectuée, reste une faiblesse de cette démarche. On peut toutefois rappeler que chaque fois que l'on a cherché à faire "coller" la réalité à un modèle théorique c'est toujours avec beaucoup de mauvaise grâce que celle-ci s'y est prêtée et les tests classiques du type χ^2 ne sont guère fameux. On peut ajouter que jusqu'à présent les auteurs de modèles n'ont que rarement cherché à faire cet exercice et lorsque ces modèles ont eu une utilisation prévisionnelle, il est encore plus rare que la comparaison du prévu et de l'observé ait été esquissée. Ce comportement est préjudiciable à l'avancement de la science économique qui ne peut progresser qu'en profitant de son expérience acquise, c'est-à-dire surtout de ses erreurs de jeunesse. On peut craindre à ce propos que la science économique ne soit pas prête de sortir de sa période d'adolescence. Cette observation n'implique d'ailleurs en rien une vision mécaniste des choses car fort heureusement les impondérables continueront à donner du sel à la prévision ; ce que nous reprochons ici c'est la grande pudeur dont font preuve les spécialistes de la prévision dans la recherche a posteriori de la cause de ces impondérables (ou dans la diffusion de ces éléments ce qui revient au même).

3.22 - Critiques externes

Le processus de modélisation suivi est tributaire du découpage retenu (qui on l'a vu dépend des techniques et des données) et d'un corps d'observations. Ce corps d'observations est nécessaire pour alimenter la modélisation en hypothèses car l'approche déductive n'est pas suffisante. Il semblerait qu'en ce domaine les efforts de clarification que nous avons entrepris soient insuffisants et qu'il faille accumuler encore un certain nombre d'observations avant de pouvoir progresser davantage. Mais ces observations, qui visent à relier le phénomène étudié à un (ou plusieurs) autre phénomène, se heurte assez vite à une insuffisance de données. Un effort sérieux est sans doute à opérer à la fois dans la constitution de données et dans la poursuite descriptive, mais cet effort ne pourra être fructueux qu'au prix d'une coordination étroite. On peut espérer alors que dans une quinzaine d'années de nouvelles études permettront de mettre au point un modèle prévisionnel satisfaisant en matière d'emploi localisé.

3.3 - Vers un dépassement

Il nous semble en fin de compte que notre effort n'aura pas été inutile car il aura permis de mettre en évidence un certain nombre de problèmes méthodologiques, et de souligner certaines carences statistiques, indépendamment de l'apport de quelques études descriptives originales. Sur le plan de la démarche l'analyse de système tentée n'a pas produit tous ses effets, mais ceci est très certainement imputable aux problèmes et carences que l'on vient de signaler.

L'outil de réflexion et de formalisation qu'est l'analyse de système est très certainement appelé à un brillant avenir en analyse régionale durant ces prochaines années. Notre espoir est que celle sur laquelle notre effort s'est appuyé sera, une fois qu'un certain nombre de blocages aura disparu, reprise et approfondie par d'autres chercheurs car nous restons convaincus que le problème de la prévision spatialisée des tensions de l'emploi sera, ces prochaines années, un sujet de préoccupation majeur pour le politique et conséquemment un axe de préoccupation privilégié de recherche pour la science économique.

AMÉNAGEMENT DU TERRITOIRE

LA DOCUMENTATION FRANÇAISE

29-31, QUAI VOLTAIRE 75340 PARIS CEDEX 07

AM. 1140-1

schéma
général
d'aménagement
de la France

emploi et espace

49

éléments pour un modèle
de prévision

②

TRAVAUX ET RECHERCHES DE PROSPECTIVE

schéma
général
d'aménagement
de la France

**emploi
et espace**

DECEMBRE 1974

éléments pour un modèle
de prévision

2

EMPLOI ET ESPACE

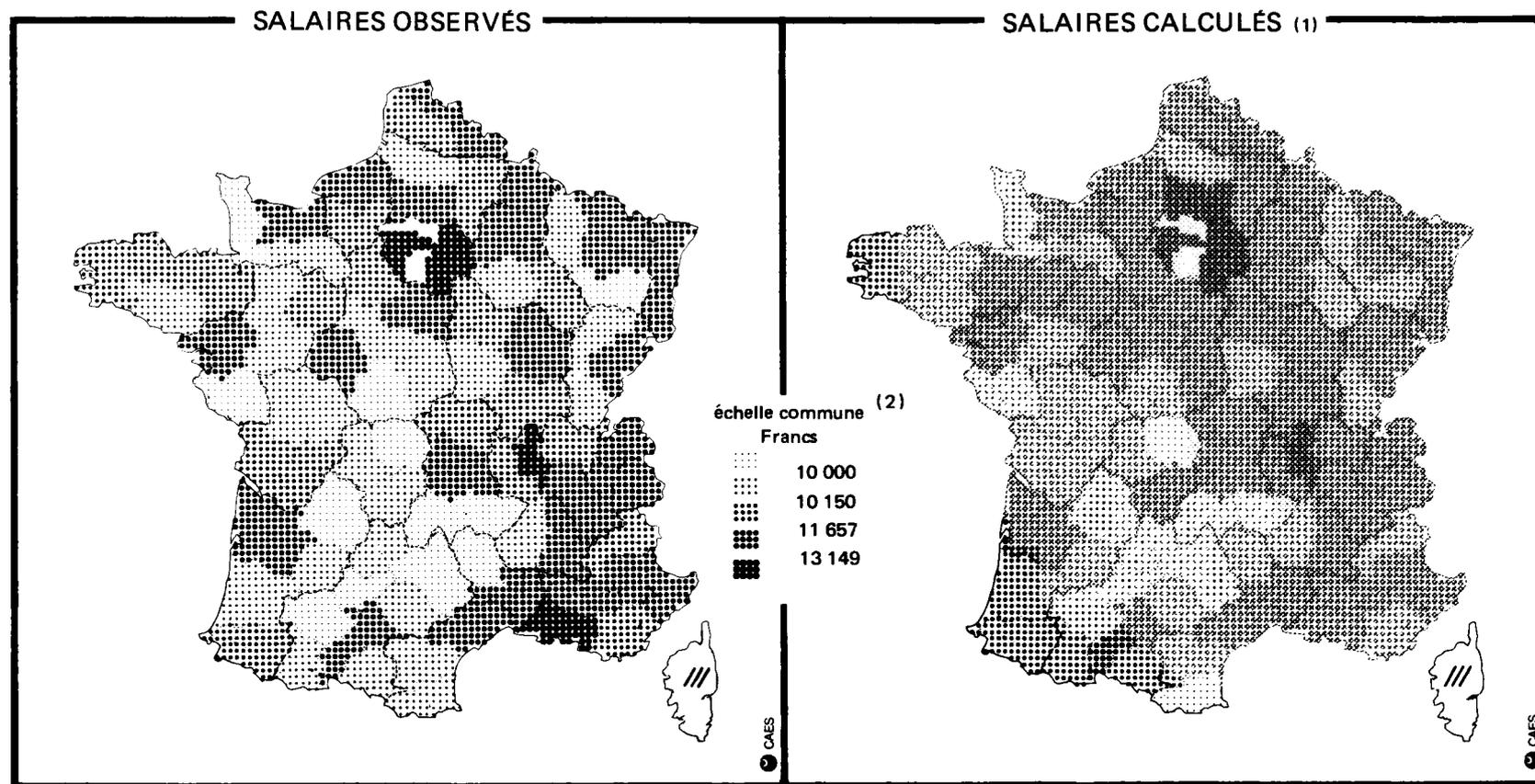
Fascicule 2

Sommaire

- A 1 - Disparités départementales des salaires annuels moyens en 1968
- A 2 - Tableau des soldes migratoires interdépartementaux 62 - 68
- A 3 - Migrants interdépartementaux 62 - 68, nés pendant la période intercensitaire
- A 4 - Taux d'activité en mars 1968 des migrants interdépartementaux 62-68
- A 5 - Comparaison des effectifs d'immigrants interdépartementaux 54 - 62 observés et de ceux calculés par le sous-modèle retenu
- A 6 - Comparaison des effectifs d'émigrants interdépartementaux 54 - 62 observés et de ceux calculés par le sous-modèle retenu
- A 7 - Immigrants extérieurs 62 - 68 (rapatriés d'Algérie exclus)
- A 8 - Comparaison des immigrants extérieurs (hors rapatriés d'Algérie) 54 - 62 observés et de ceux calculés par le sous-modèle retenu
- A 9 - Comparaison des taux d'activité observés en 1968 et de ceux calculés par le sous-modèle retenu
- A 10 - Etude de l'évolution de la part des effectifs du B.T.P. dans la population active départementale comparativement à l'évolution "France Entière"
- A 11 - Comparaison des effectifs du B.T.P. observés en 1962 et de ceux calculés par le sous-modèle retenu
- A 12 - Comparaison des effectifs du B.T.P. observés en 1968 et de ceux prévus par le sous-modèle retenu
- A 13 - Comparaison de l'évolution des effectifs d'agriculteurs départementaux comparativement à l'évolution "France Entière"
- A 14 - Comparaison des effectifs d'agriculteurs observés en 1962 et de ceux calculés par le sous-modèle retenu

- A 15 - Comparaison de l'évolution des taux de tertiaire (par rapport à la population totale) départementaux et nationaux
- A 16 - Catégories responsables d'une "accélération" ou d'un "resserrement" relatif du taux de tertiaire
- A 17 - Comparaison des effectifs observés du tertiaire en 1962 et de ceux calculés à l'aide du sous-modèle retenu
- A 18 - Comparaison des effectifs du tertiaire observés en 1968 et de ceux prévus dans le sous-modèle retenu : écarts relatifs des taux par rapport à la population totale
- A 19 - Etat potentiel sans et avec corrections par coefficients de spécificité de regression

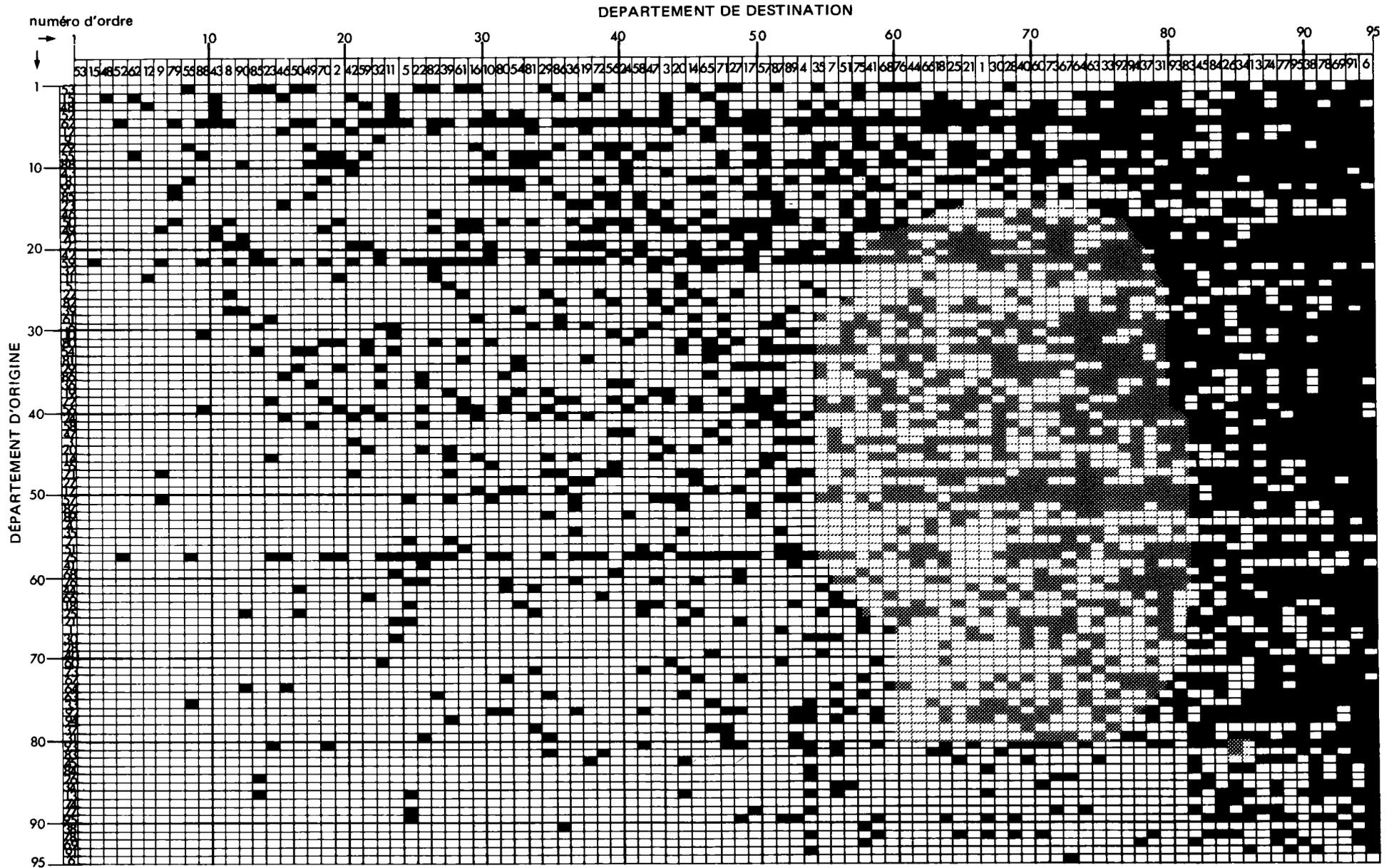
DISPARITÉS DÉPARTEMENTALES DES SALAIRES ANNUELS MOYENS EN 1968



La comparaison de ces deux cartes permet de constater qu'une partie importante des disparités spatiales de salaires est imputable aux différences spatiales de structure de population active. On trouvera page 28, un commentaire plus détaillé de ce phénomène.

(1) Moyenne des salaires par sexe et catégorie socio-professionnelle observée dans chaque département, pondérée par l'importance relative de la catégorie d'actifs considérée dans la population active française (exprimée en personne - année - travail). Si l'on remplace le système de pondération «France entière» par celui observé dans le département, on obtient alors, le salaire annuel moyen départemental observé.

(2) 13 149 est le salaire annuel moyen pour l'ensemble des salariés français, tandis que 11 657 correspond au salaire annuel moyen des salariés qui résident en dehors de la Région Parisienne.

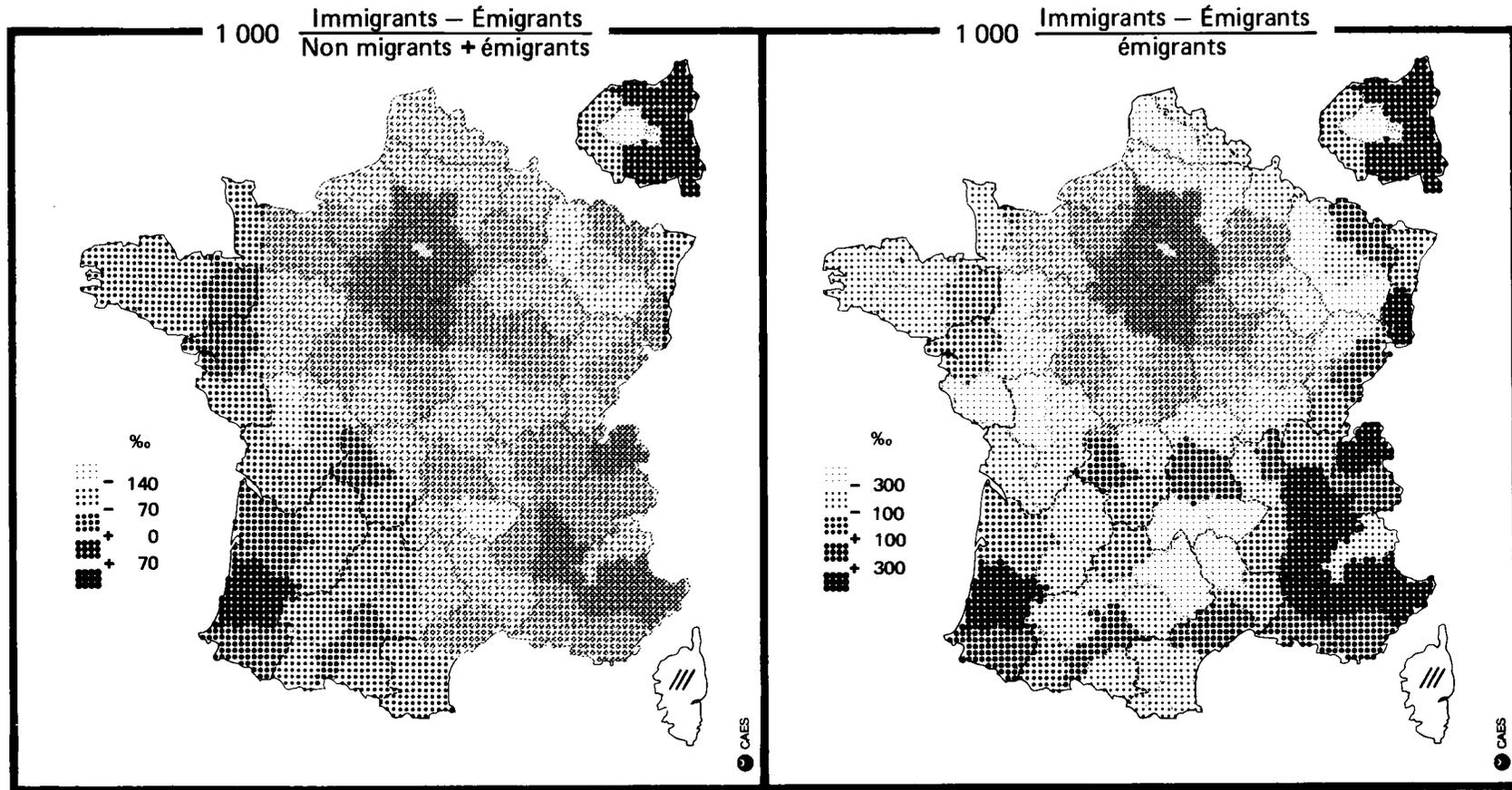


CAES

case noire: le département repéré en ligne bénéficie dans ces échanges migratoires avec le département repéré en colonne d'un solde migratoire significativement positivement.

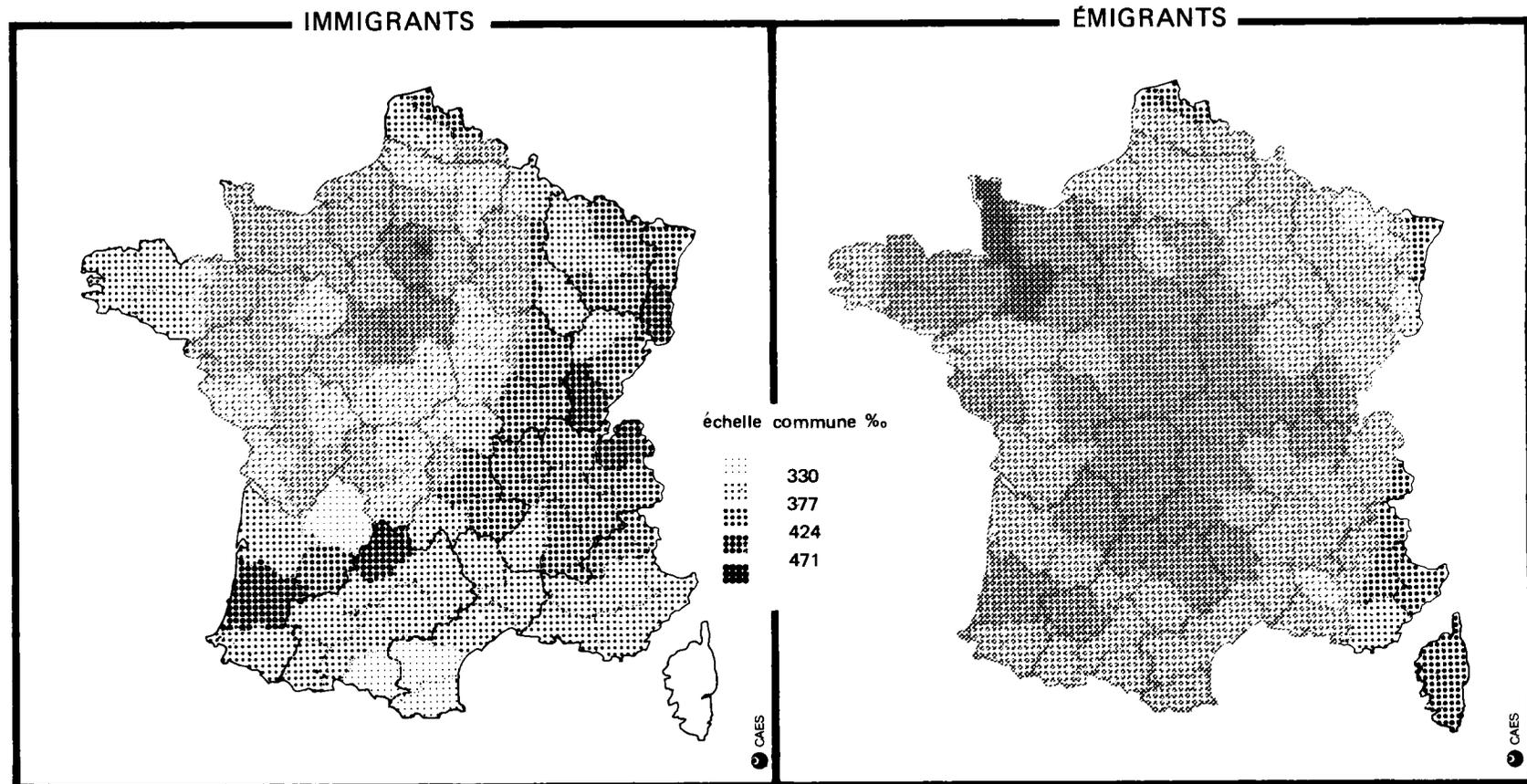
Le tableau met en évidence l'existence d'une certaine transivité dans les filières migratoires puisque, après classement, dans plus de 80% des cas, un département a des échanges bénéficiaires avec tous ceux de rang inférieur, et déficitaires avec ceux de rang supérieur. On trouvera page 189 un commentaire détaillé de ce phénomène.

MIGRANTS INTERDÉPARTEMENTAUX 62-68, NÉS PENDANT LA PÉRIODE INTERCENSITAIRE



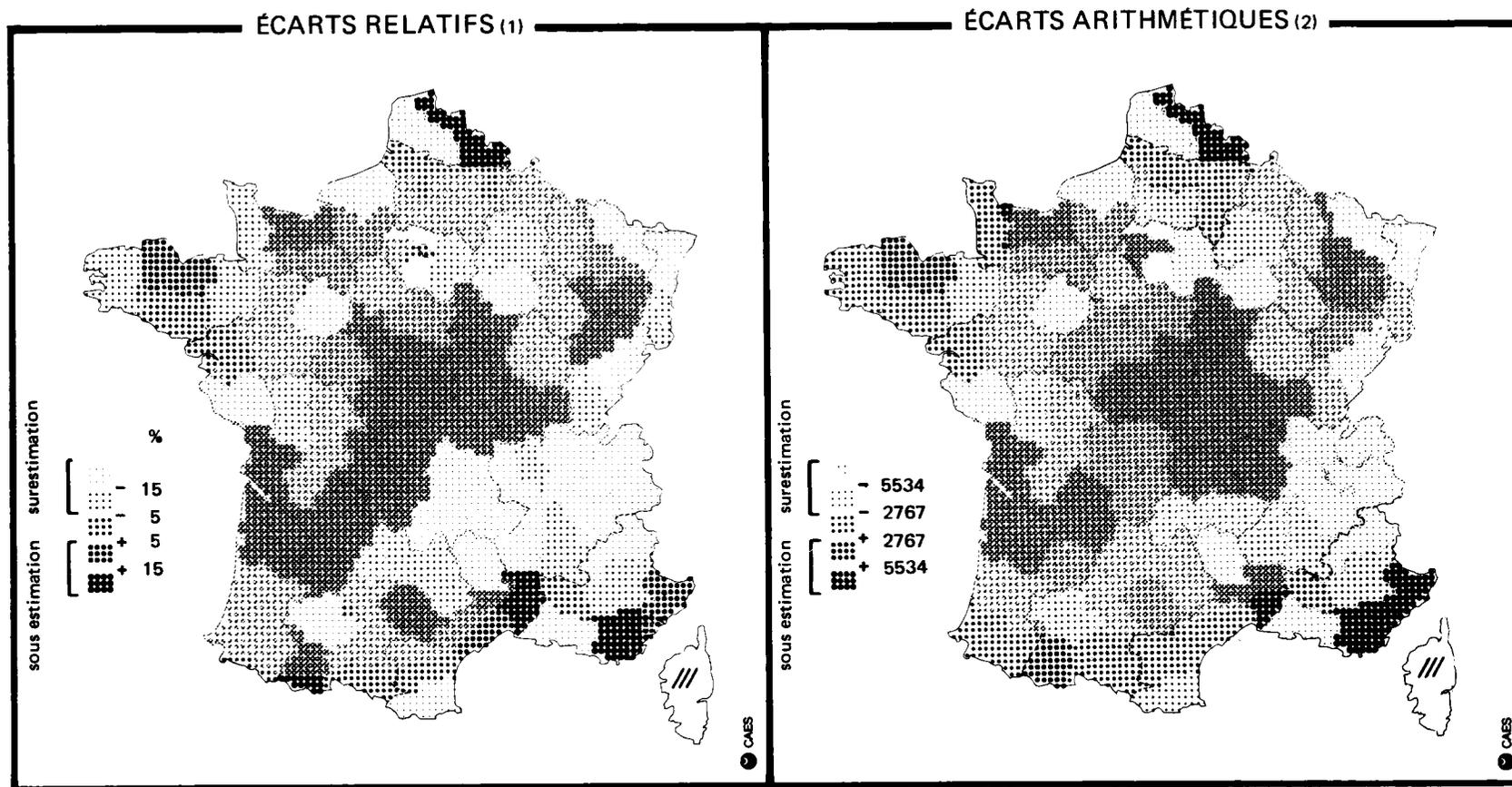
Le premier ratio mesure l'apport net d'enfants imputable aux échanges migratoires, comparativement à la situation que l'on aurait eu en l'absence de mouvements migratoires ; tandis que le second ratio indique la valeur de l'apport net pour 1 000 départements. Le commentaire de ces cartes se trouve page 198 et suivantes.

TAUX D'ACTIVITÉ EN MARS 1968 DES MIGRANTS INTERDÉPARTEMENTAUX 62-68



On peut constater à l'aide de ces deux cartes, l'existence d'une profonde dissymétrie entre les caractéristiques d'activité de l'émigration et de l'immigration ; en effet les départements présentent peu de différence à l'émission, mais les migrants actifs se dirigent préférentiellement sur quelques zones. On trouvera page 201, un commentaire détaillé de ces différences d'activité et de structures par catégorie socio-professionnelle des migrants interdépartementaux. On peut toutefois ajouter que le taux d'activité moyen des migrants interdépartementaux est de 420‰ alors que celui de la population non migratoire est de 409‰.

COMPARAISON DES EFFECTIFS D'IMMIGRANTS INTERDÉPARTEMENTAUX 54-62 OBSERVÉS ET DE CEUX CALCULÉS PAR LE SOUS-MODELE RETENU

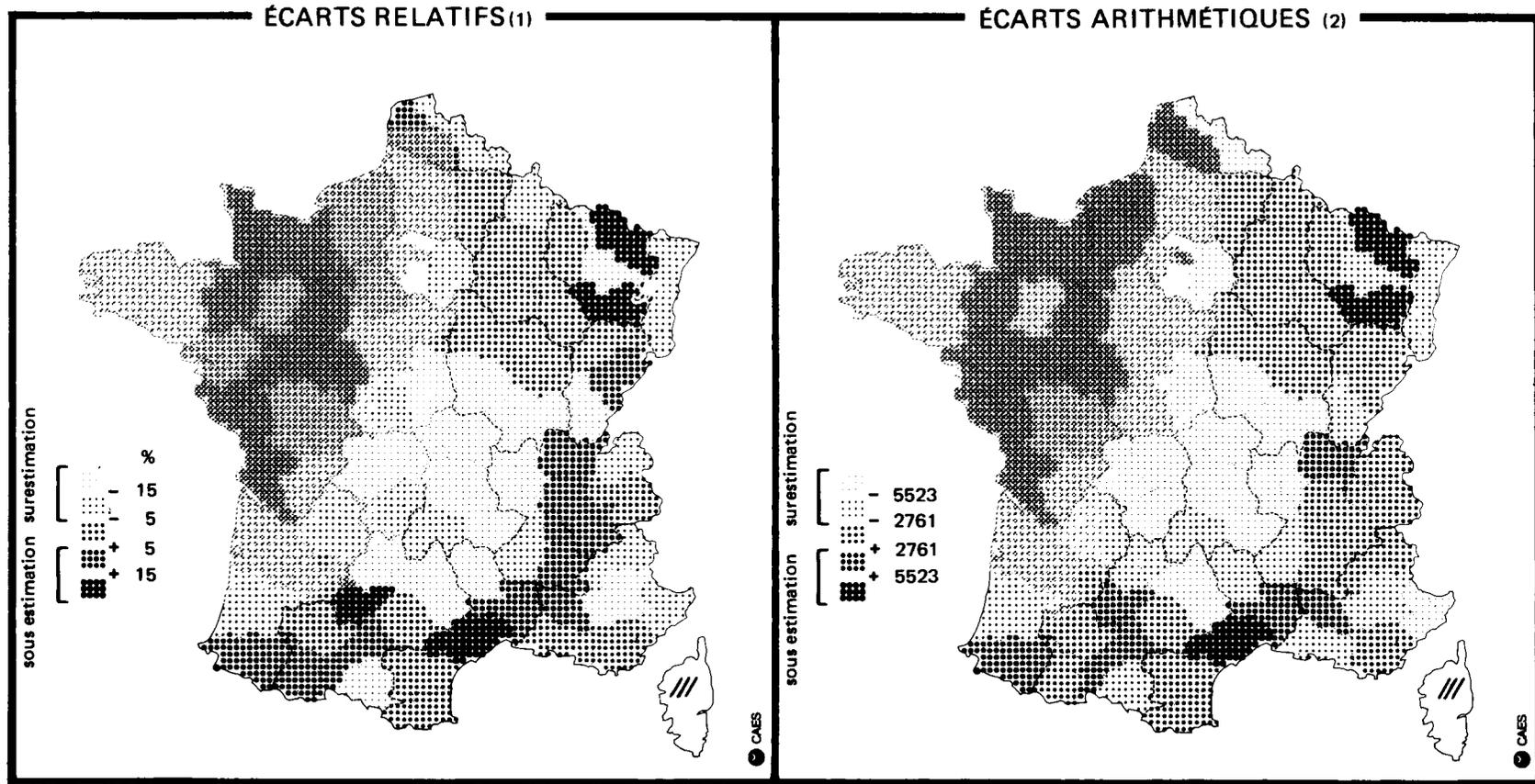


On trouvera page 213 et suivantes, les raisons du choix du sous-modèle proposé, ainsi que les commentaires des résultats.

(1) = $100 \frac{\text{observé} - \text{calculé}}{\text{observé}}$

(2) = $\text{observé} - \text{calculé}$

COMPARAISON DES EFFECTIFS D'ÉMIGRANTS INTERDÉPARTEMENTAUX 54-62 OBSERVÉS ET DE CEUX CALCULÉS PAR LE SOUS-MODELE RETENU

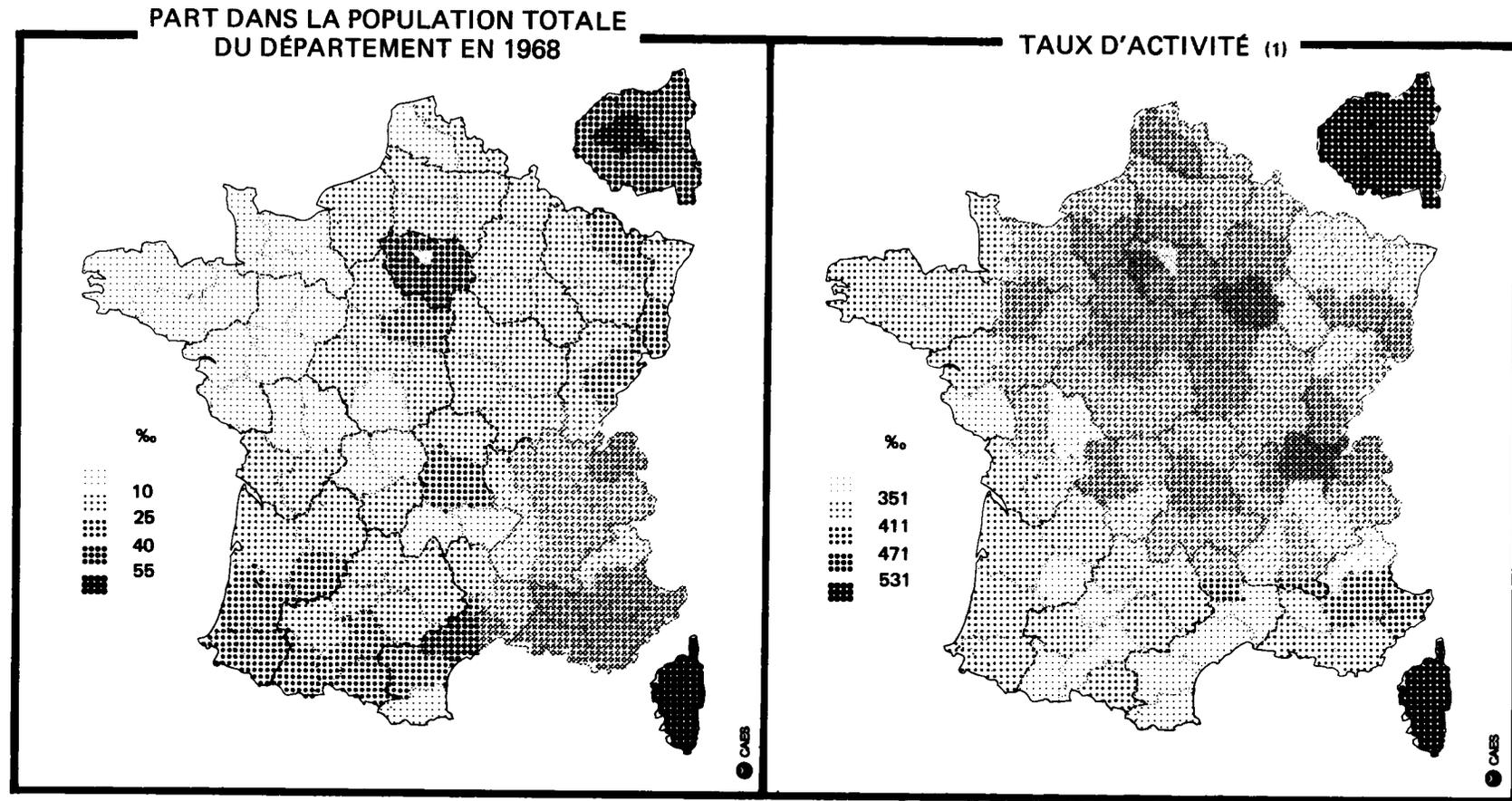


On trouvera page 218 et suivantes, les raisons du choix du sous-modèle proposé ainsi que les commentaires des résultats.

(1) = $100 \frac{\text{observé-calculé}}{\text{observé}}$

(2) = $\text{observé} - \text{calculé}$

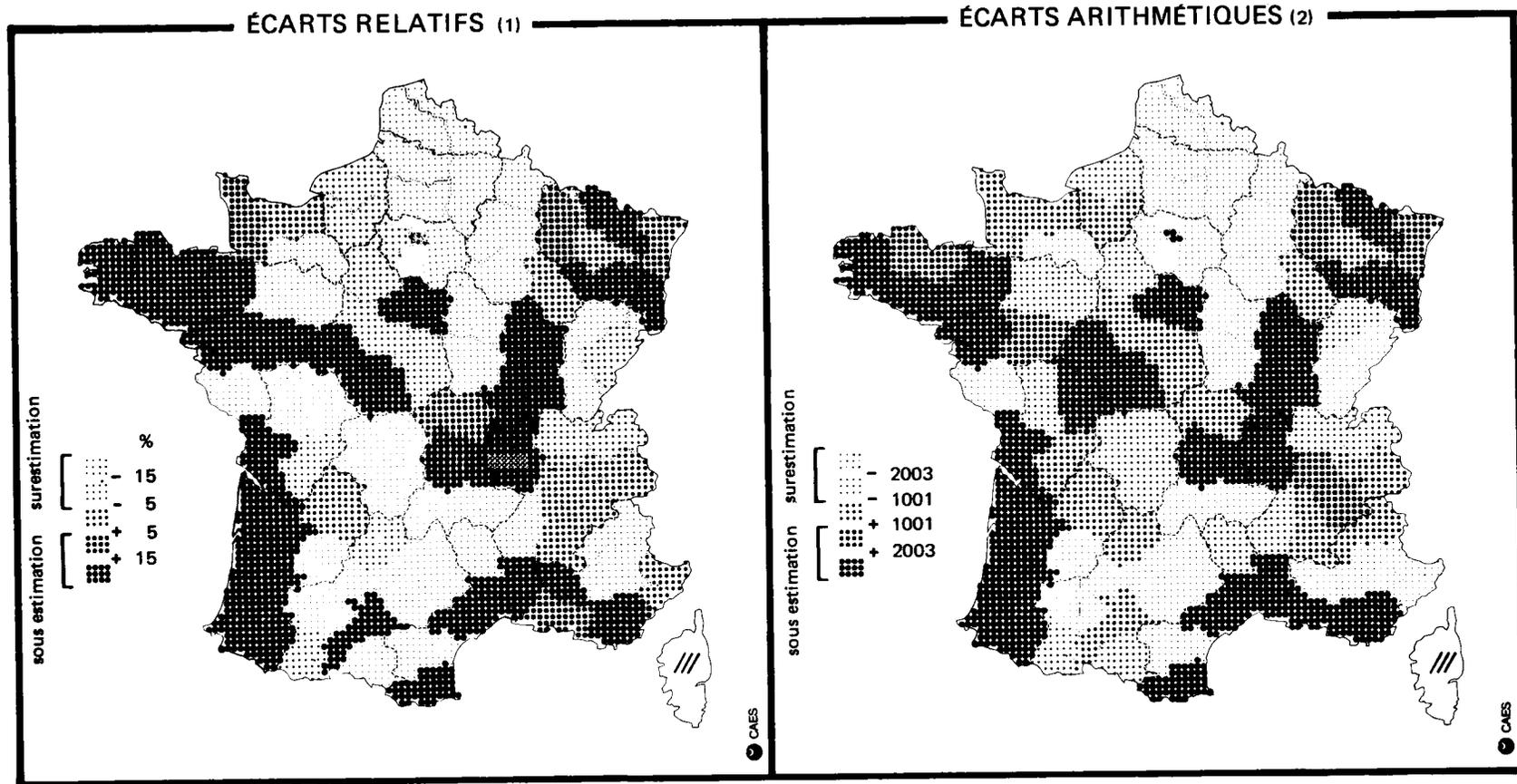
IMMIGRANTS EXTÉRIEURS 62-68 (RAPATRIÉS D'ALGÉRIE EXCLUS)



Ces deux cartes illustrent deux des principales caractéristiques de ces immigrants extérieurs : leur polarisation spatiale et leur activité supérieure à l'activité moyenne du reste de la population. L'étude détaillée des caractéristiques de ces immigrants débute page 237.

(1) 411‰ est le taux moyen d'activité de l'ensemble des résidents, sur le territoire en 1968 et 471‰ le taux moyen des immigrants extérieurs.

COMPARAISON DES IMMIGRANTS EXTÉRIEURS (HORS RAPATRIÉS D'ALGÉRIE) 54-62 OBSERVÉS ET DE CEUX CALCULÉS PAR LE SOUS-MODELE RETENU

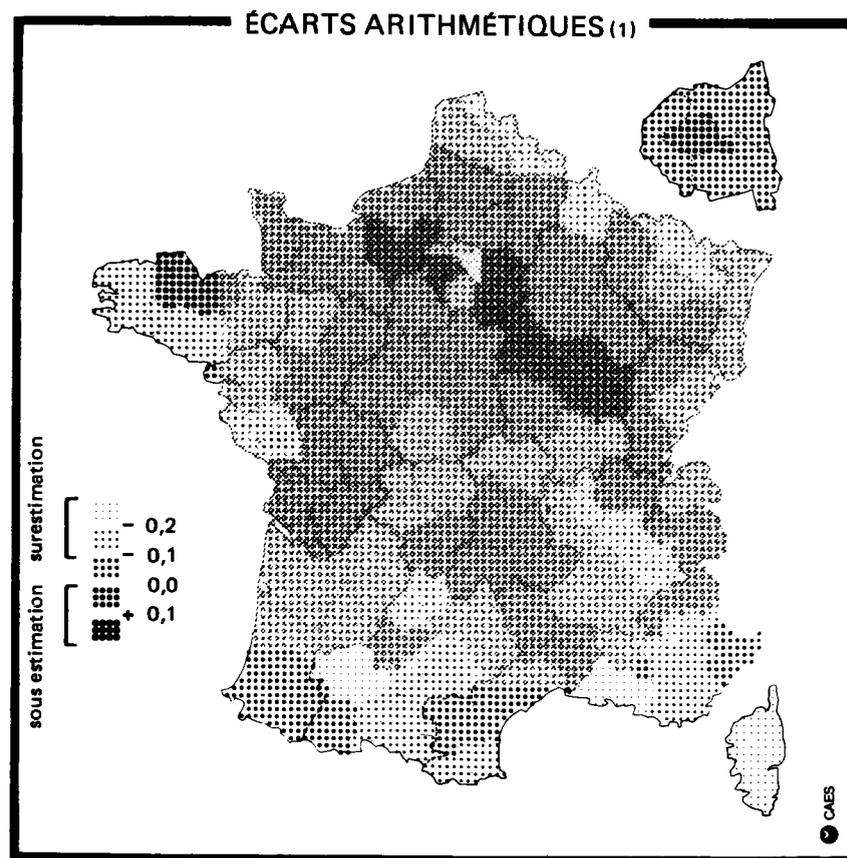


Les raisons du choix du sous-modèle retenu, ainsi que les commentaires des résultats se trouvent exposés page 343 et suivantes.

(1) = $100 \frac{\text{observé} - \text{calculé}}{\text{observé}}$

(2) = $\text{observé} - \text{calculé}$

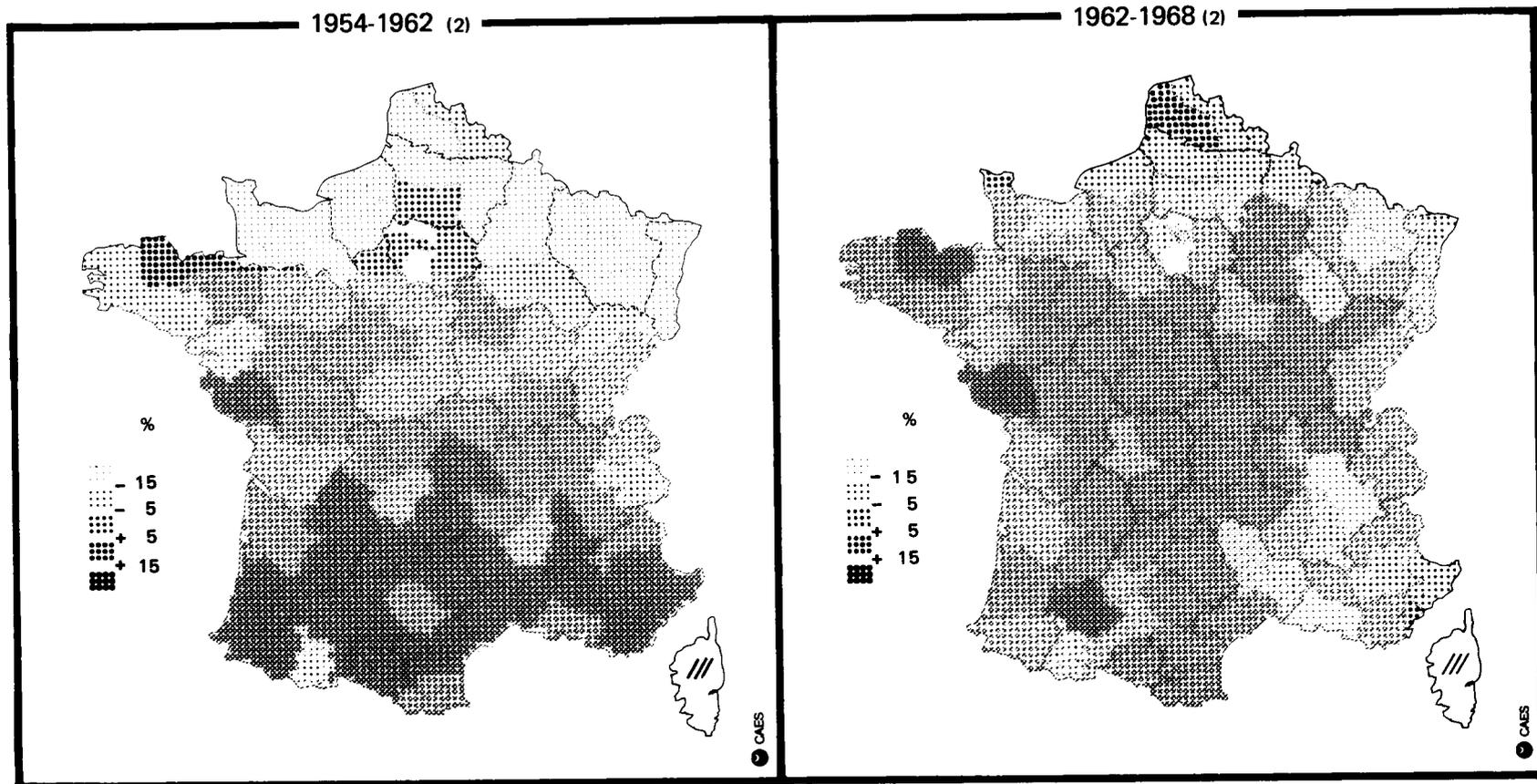
COMPARAISON DES TAUX D'ACTIVITÉ OBSERVÉS EN 1968 ET DE CEUX CALCULÉS PAR LE SOUS-MODELE RETENU



Le taux d'activité moyen pour l'ensemble de la population ne résidant pas en Région Parisienne était de 40,1% en 1968. On trouvera page 252 et suivantes, l'analyse de ces résultats.

(1) Taux observé - taux calculé

ÉTUDE DE L'ÉVOLUTION DE LA PART DES EFFECTIFS DU BTP DANS LA POPULATION ACTIVE DÉPARTEMENTALE COMPARATIVEMENT A L'ÉVOLUTION « FRANCE ENTIÈRE » (1)

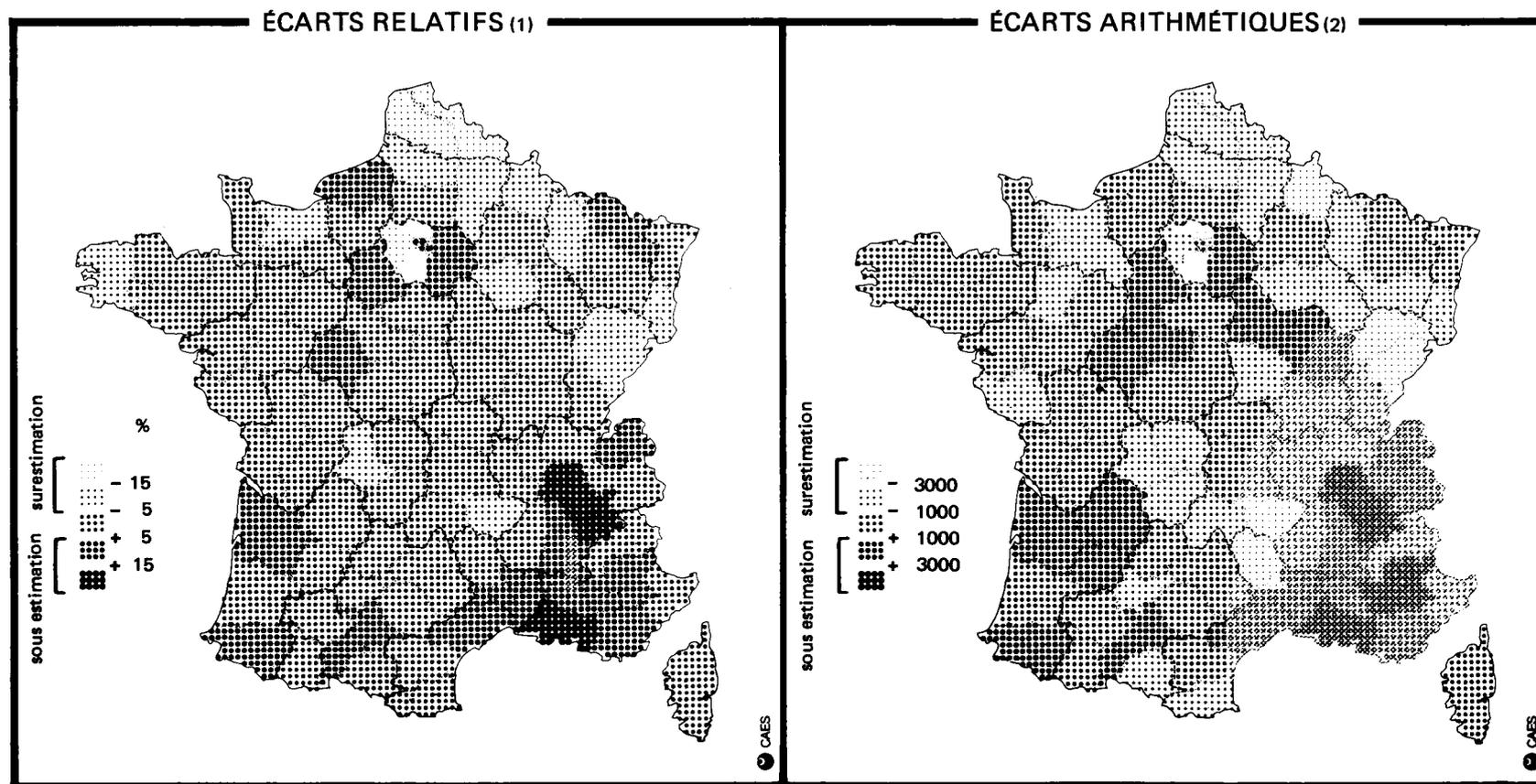


Il semble au vu de ces deux cartes qu'il n'y ait guère de continuité prononcée dans l'évolution du phénomène. La même analyse a été aussi conduite en terme d'effectifs. Ces évolutions sont analysées en détail à partir de la page 263.

(1) La valeur observée est celle du taux constaté en fin de période, la valeur calculée résulte de l'application à la valeur départementale observée en début de période du taux considéré, du coefficient de croissance (= rapport de la valeur de fin de période, à celle de début de période) observé pour l'ensemble de la France.

(2) $= 100 \times (\text{observé} - \text{calculé}) / \text{observé}$

COMPARAISON DES EFFECTIFS DU BTP OBSERVÉS EN 1962 ET DE CEUX CALCULÉS PAR LE SOUS-MODELE RETENU

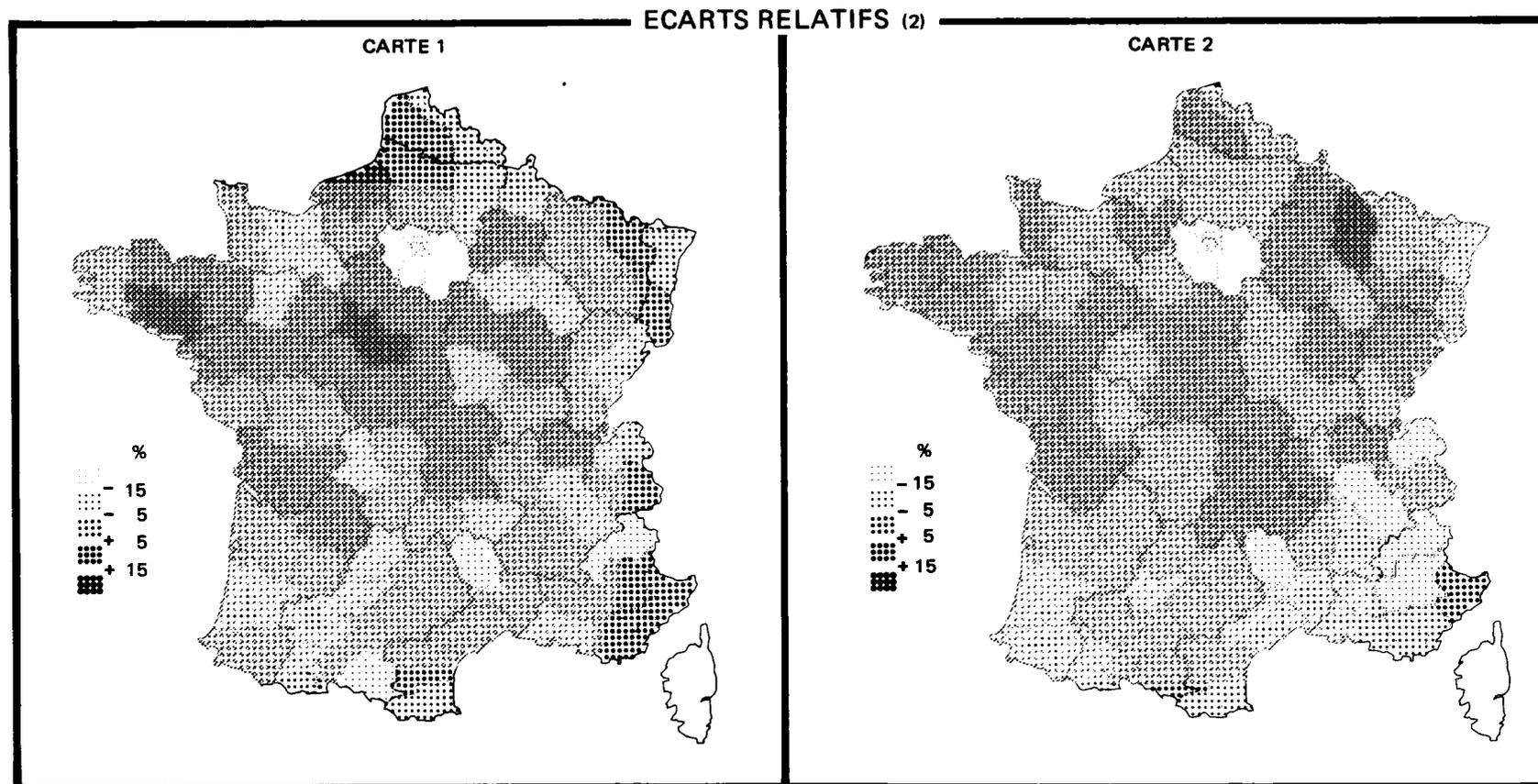


On trouvera page 277 et suivantes, les raisons du choix du sous-modèle retenu, ainsi que les commentaires des résultats.

(1) $100 \frac{\text{observé} - \text{calculé}}{\text{observé}}$

(2) $\text{observé} - \text{calculé}$

COMPARAISON ⁽¹⁾ DES EFFECTIFS DU BTP OBSERVÉS EN 1968 ET DE CEUX PRÉVUS PAR LE SOUS-MODELE RETENU

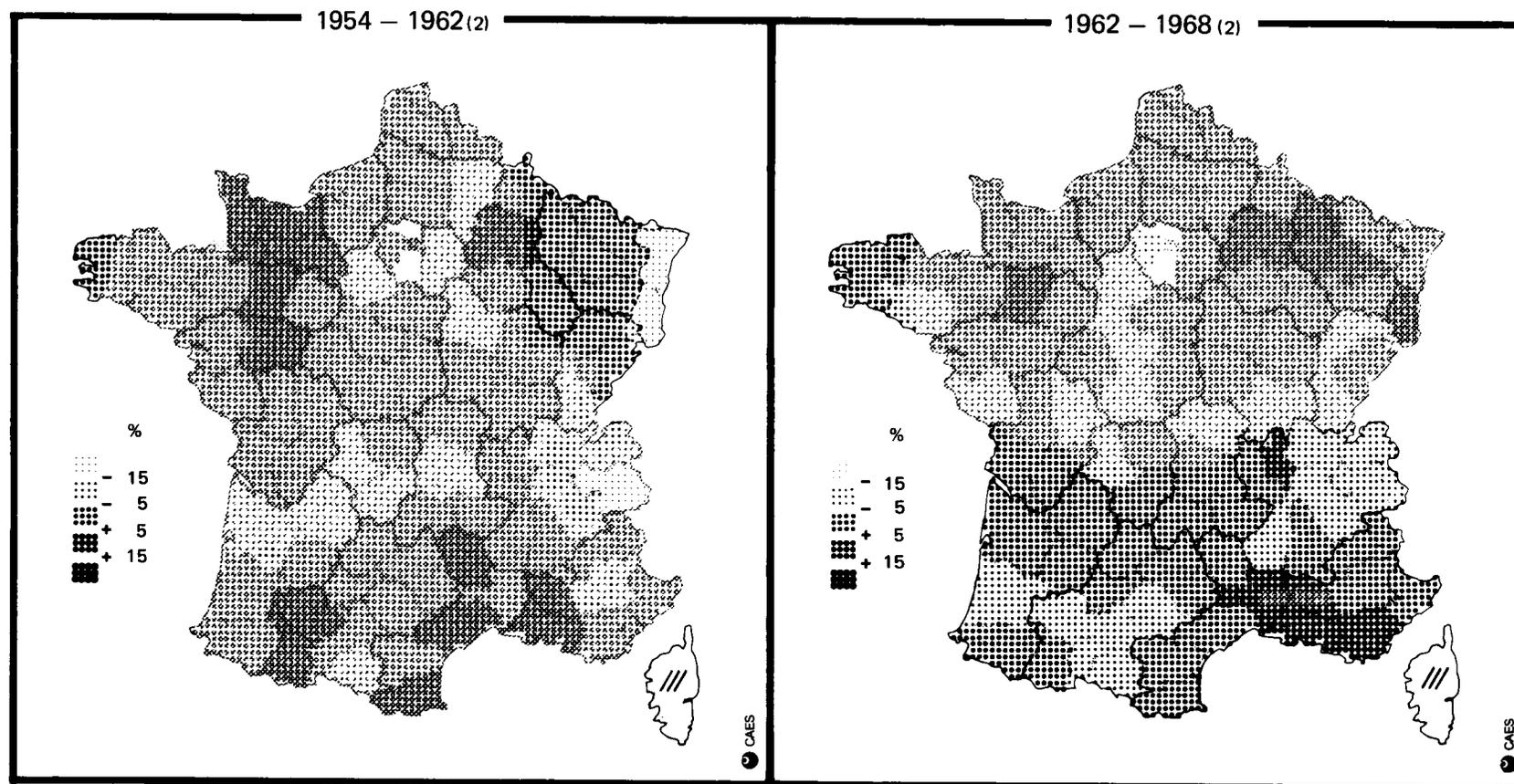


Le commentaire de ces résultats se trouve page 279.

(1) Les prévisions départementales ont été cadrées sur la valeur prévue en 1968 pour l'ensemble du territoire. Par ailleurs, on peut indiquer que les données départementales de la carte 1 correspondent aux estimations ponctuelles fournies par la régression multiple, et corrigées par le cadrage ; tandis que celles de la carte 2 intègrent une correction supplémentaire et qui consiste à utiliser les coefficients de spécificité de régression départementaux calculés lors du « calibrage » du modèle.

(2) $\pm 100 (\text{observé} - \text{calculé}) / \text{observé}$.

COMPARAISON DE L'ÉVOLUTION DES EFFECTIFS D'AGRICULTEURS DÉPARTEMENTAUX COMPARATIVEMENT
A L'ÉVOLUTION « FRANCE ENTIÈRE » (1)

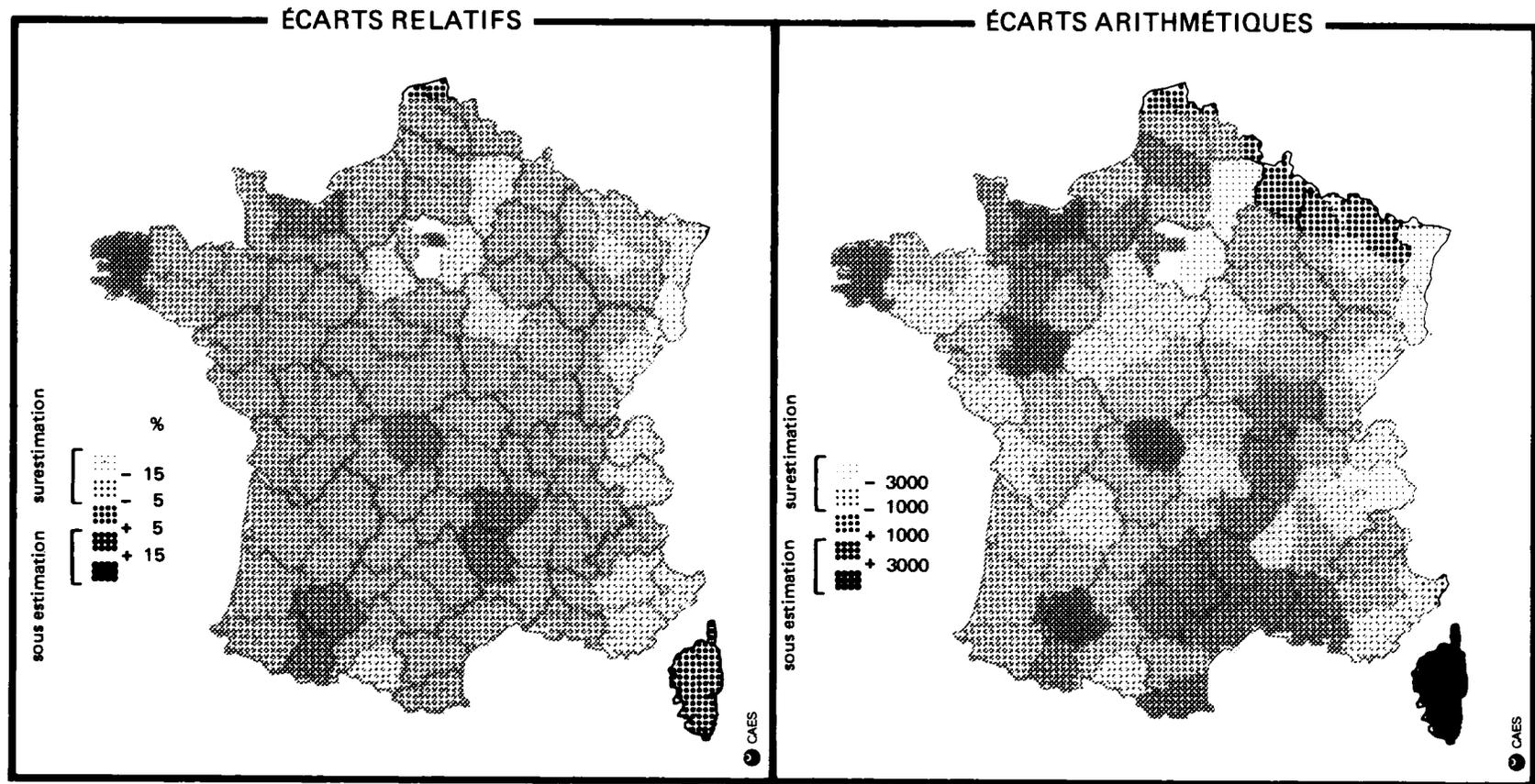


L'analyse de ces cartes se trouve aux pages 291 et suivantes.

(1) La valeur observée est celle de fin de période, la valeur calculée résulte de l'application à la valeur départementale observée en début de période du taux considéré du coefficient de croissance (= rapport de la valeur de fin de période, à celle de début de période) observé pour l'ensemble de la France.

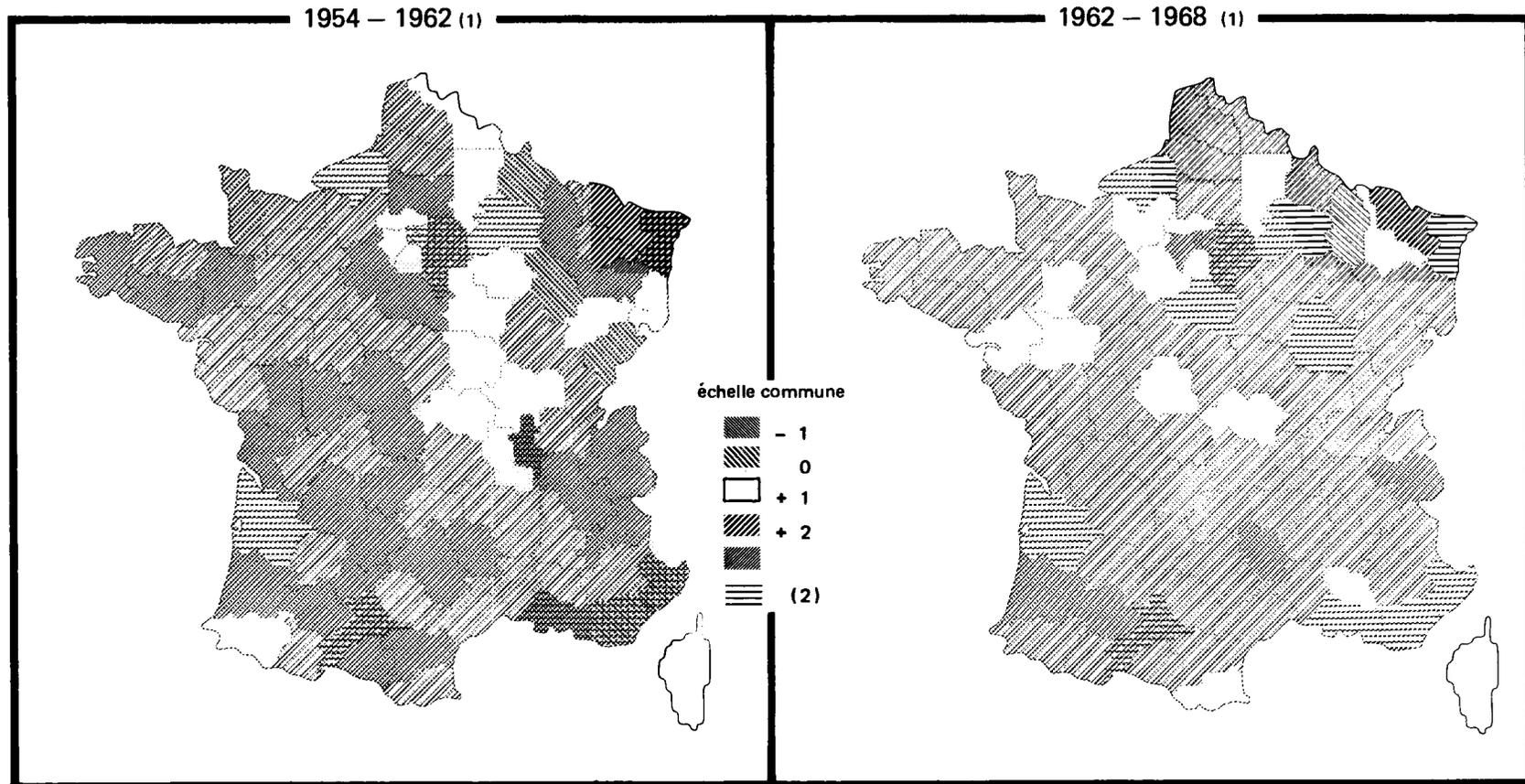
(2) = $100 \text{ (observé-calculé) / observé}$.

COMPARAISON DES EFFECTIFS D'AGRICULTEURS OBSERVÉS EN 1962 ET DE CEUX CALCULÉS PAR LE SOUS-MODELE RETENU



On trouvera page 303 et suivantes, les raisons du choix du sous-modèle retenu, ainsi que les commentaires des résultats.

COMPARAISON DE L'ÉVOLUTION DES TAUX DE TERTIAIRE (PAR RAPPORT A LA POPULATION TOTALE)
DÉPARTEMENTAUX ET NATIONAUX

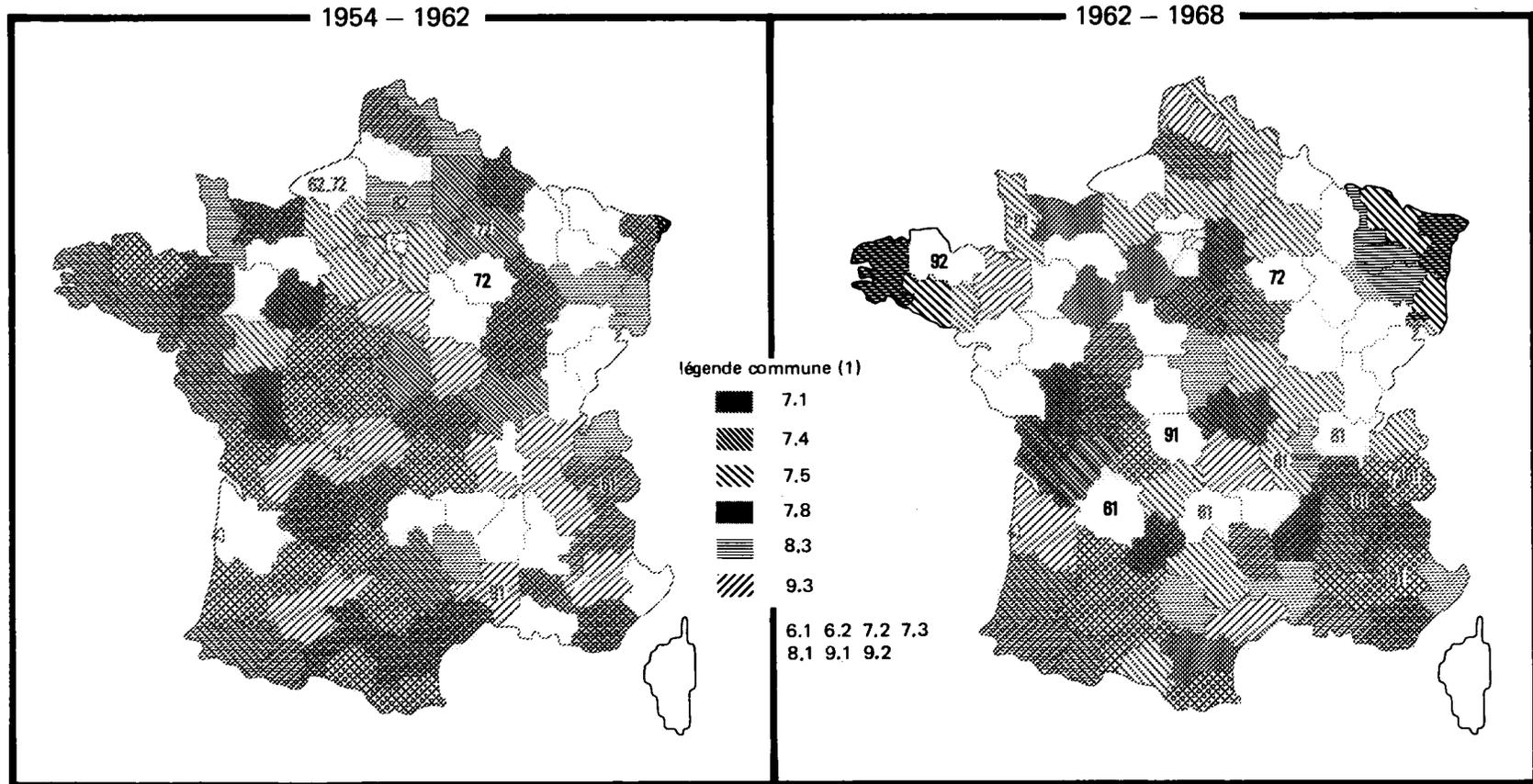


On peut apprécier à l'aide de ces cartes l'évolution des disparités départementales en tenant compte du fait que sur la première période, le taux de croissance du taux du tertiaire national était de 13%, tandis que sur la seconde période, il était de 15% et que seuls les départements dont la délimitation est en traits pleins avaient en début de période un taux de tertiaire supérieur à la moyenne de l'ensemble du pays. Ces cartes sont commentées page 315 et suivantes.

(1) Rapport du taux d'évolution départementale au taux d'évolution nationale.

(2) Département ayant en début de période un taux de tertiaire supérieur à la moyenne de la France Entière

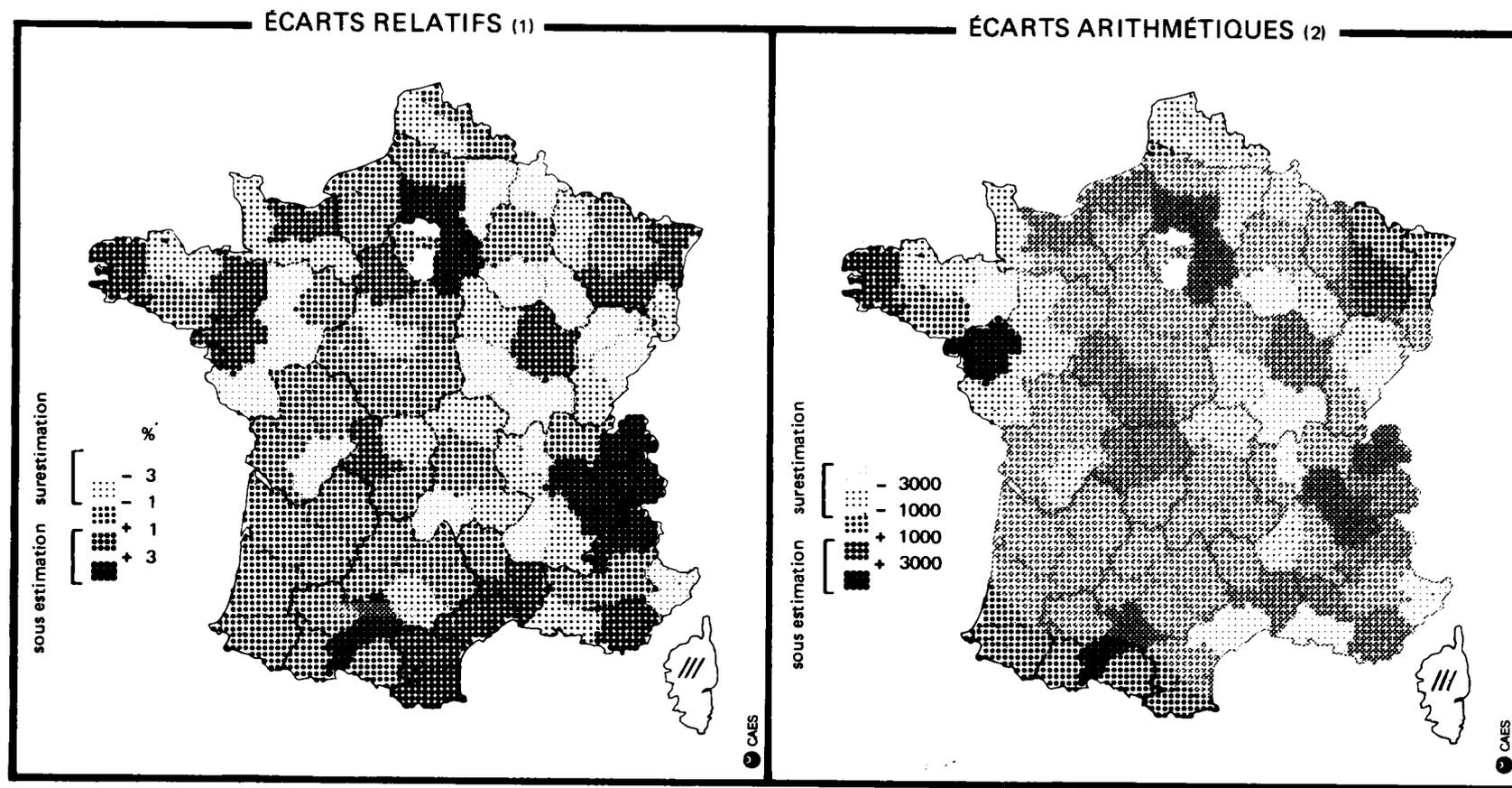
CATÉGORIES RESPONSABLES D'UNE « ACCÉLÉRATION » OU D'UN « RESSERREMENT » RELATIF DU TAUX DE TERTIAIRE DE TERTIAIRE



Ces deux cartes illustrent une étude typologique détaillée de l'évolution départementale du tertiaire, qui débute page 324.

- (1) 7.1 Commerces agricoles et alimentaires de gros 7.4 Autres commerces de gros 7.5 Autres commerces de détail 7.8 Banques et assurances
8.3 Services aux particuliers 9.3 Administrations publiques
6.1 Transports terrestres 6.2 Transports fluviaux, maritimes, aériens 7.2 Commerces agricoles et alimentaires de détail 7.3 Hôtellerie, débits de boissons
8.2 Services domestiques 8.1 Services aux entreprises 9.1 Eau, gaz, électricité 9.2 Transmissions et radios.

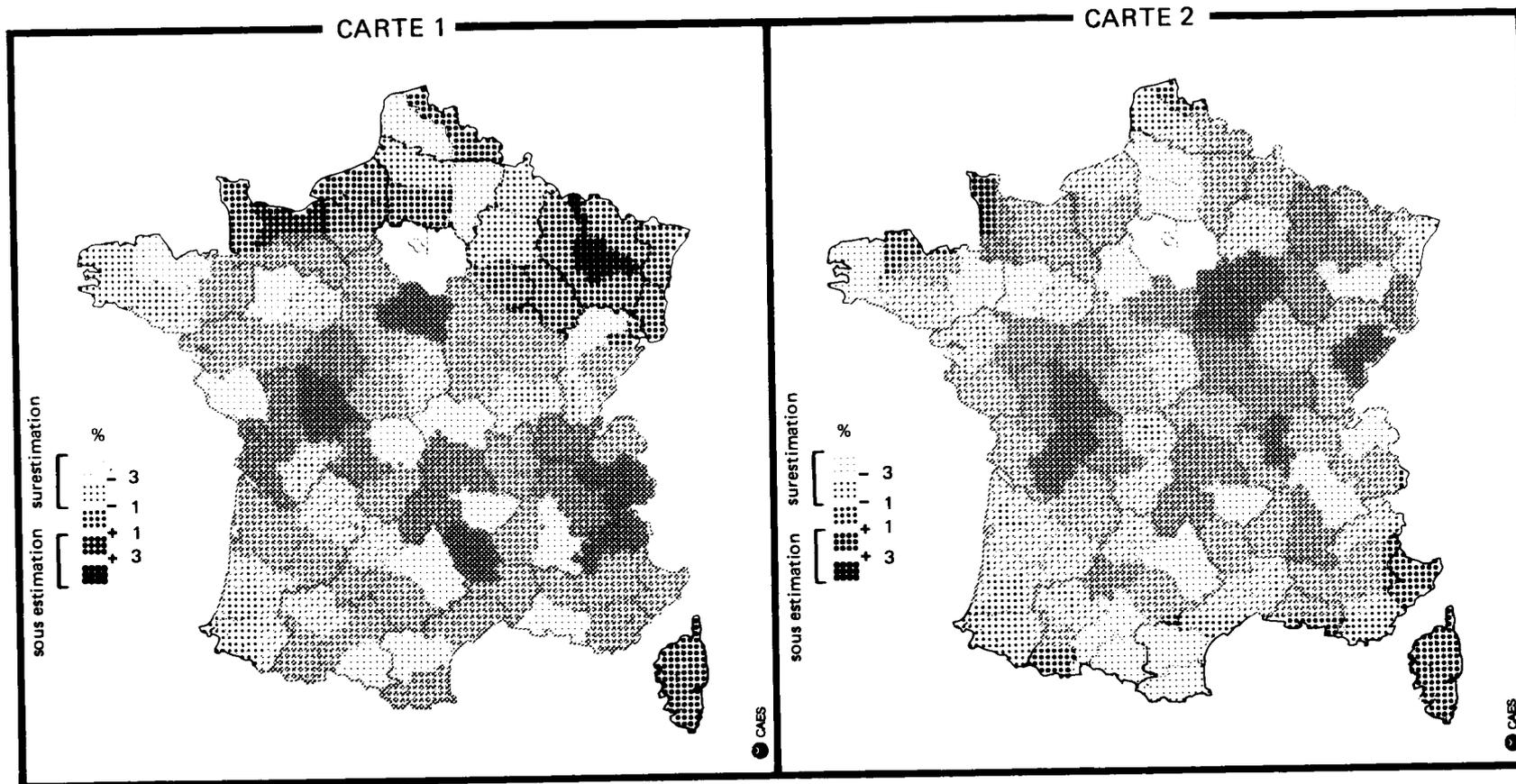
COMPARAISON DES EFFECTIFS OBSERVÉS DU TERTIAIRE EN 1962 ET DE CEUX CALCULÉS A L'AIDE DU SOUS-MODELE RETENU



On trouvera page 351 et suivantes, les raisons du choix du sous-modèle retenu, ainsi que les commentaires des résultats.

- (1) $100 \frac{\text{observé-calculé}}{\text{observé}}$; l'échelle retenue ici est différente des précédentes « illustrations » de modèle, du fait de la « qualité » des résultats.
 (2) $(\text{observé} - \text{calculé})$

COMPARAISON⁽¹⁾ DES EFFECTIFS DU TERTIAIRE OBSERVÉS EN 1968 ET DE CEUX PRÉVUS DANS LE SOUS-MODELE RETENU : ÉCARTS RELATIFS⁽²⁾ DES TAUX PAR RAPPORT A LA POPULATION TOTALE.



Les commentaires de ces résultats se trouvent page 352 et suivantes.

(1) Voir note page A 12

(2) = 100 (observé - calculé) / observé

ÉTAT POTENTIEL SANS CORRECTION PAR COEFFICIENTS
DE SPÉCIFICITÉ DE RÉGRESSION

ÉTAT POTENTIEL AVEC CORRECTIONS PAR COEFFICIENTS
DE SPÉCIFICITÉ DE RÉGRESSION

	OBSERVÉ		ÉQUILIBRE		ÉTAT POTENTIEL		
	Pop. Tot.	Pop. act.	Pop. Tot.	Pop. act.	Pop. Tot.	demande	offre
1 AIN.....	23924	150056	336510	147933	316339	138366	144288
2 AISNE.....	52635	195660	513637	188154	511133	188411	188672
3 ALLIER.....	34633	145492	368064	144741	371500	157150	142675
4 BASSES ALPES.....	10481	46116	103051	39137	106115	38634	38940
5 HAUTES ALPES.....	9179	36624	39776	39100	40730	35427	37611
6 ARDENNES.....	27247	272497	246074	246074	246074	246074	246074
7 ARDECHE.....	25493	94224	254926	94224	244231	94514	97309
8 ARDENNES.....	30934	110564	107059	114421	108845	114421	108845
9 ARTOIS.....	13484	49788	132794	49788	13667	51430	50088
10 AUVERGNE.....	27032	114768	285340	113944	278469	114774	113814
11 AUBE.....	27827	94000	281470	94000	27462	94000	94000
12 AVALOIS.....	24157	106716	280844	106716	27411	106694	107683
13 BASSES NORMANDIES.....	147027	517130	1289694	486800	1475962	558397	572084
14 CALVADOS.....	51964	210274	553737	210274	533655	212718	210095
15 CANTAL.....	15312	60503	140607	60503	14701	63708	61822
16 CHARENTE.....	33102	143188	405400	122474	333674	128470	125684
17 CHARENTE MARITIME.....	42620	177448	468474	184443	527477	190373	177435
18 CHER.....	30400	120696	303496	118155	311122	120017	114818
19 CORSE.....	23766	98172	232006	98172	231473	98172	94266
20 COTE D'OR.....	42119	168008	426000	144444	417371	163333	164447
21 COTES DU NORD.....	50610	203400	499734	199911	530075	200080	199889
22 CREUSE.....	15488	60406	159760	60406	157372	60406	60406
23 DORDOGNE.....	31470	150134	371741	149337	344817	148462	148462
24 DROME.....	42626	171958	398121	141114	425456	172395	164579
25 DUOIS.....	34249	133952	332931	132040	37427	132054	131115
26 DURANDE.....	38334	155986	395570	140044	379334	153404	157783
27 EURE ET LOIRE.....	30221	125924	318110	119466	342236	117537	123028
28 FINISTERE.....	7414	297440	711115	297440	711115	297440	301811
29 GARONNE.....	47854	160388	490474	163444	491141	164296	163966
30 HAUTE GARONNE.....	65071	259270	666844	258704	697704	270494	264886
31 GERS.....	18154	71912	147456	78493	140575	75322	74032
32 GIRONDE.....	100940	391074	10172012	391074	10172012	391074	391074
33 HERAULT.....	59140	197948	407754	204811	575643	195245	200790
34 ILLE ET VILAINE.....	65272	279720	650444	240441	640000	275905	278461
35 INDE.....	24716	100172	246620	100172	246620	100172	100172
36 INDE ET LOIRE.....	42787	175912	439820	173425	434454	171979	172737
37 ISERE.....	74634	304400	774731	323411	728221	314579	314579
38 JURA.....	23955	95164	231491	93144	235977	94029	93748
39 LANDES.....	27734	111936	471482	114400	275704	116163	114948
40 LOIR ET CHER.....	24790	111360	266460	109349	273174	121166	110467
41 LOIRE.....	22924	22924	22924	22924	22924	22924	22924
42 HAUTE LOIRE.....	25834	84400	815130	46450	40377	42265	45501
43 LOIRE ATLANTIQUE.....	46113	334408	833010	324770	874233	341444	331740
44 LOIRET.....	42623	160788	415411	171191	419617	172984	171882
45 LORRAINE.....	15120	57746	145400	57746	145400	57746	57746
46 LOT.....	62400	240400	114400	240400	114400	240400	240400
47 LOT ET GARONNE.....	7726	30180	88110	33424	74731	28422	31990
48 LOZERE.....	54671	239560	577401	234390	602651	244736	237796
49 MACHES.....	43194	191808	491400	190319	453394	191190	190088
50 MAINE.....	41480	197488	400005	197488	400005	197488	197488
51 HAUTE MAINE.....	21430	81868	213800	81868	213800	81868	81868
52 MAYENNE.....	26274	117420	261179	118497	262724	121263	119477
53 MEURTHE ET MOSELLE.....	70841	261090	657994	244992	727443	271439	265671
54 MEUSE.....	29951	77574	198434	71723	204961	79077	78808
55 MOSELLE.....	34097	131808	315048	131808	315048	131808	131808
56 NIEVRE.....	97131	338884	890391	327761	975135	358167	338083
57 NORD.....	247705	918000	242411	84420	247462	90742	89072
58 ORNE.....	241790	877488	425488	845941	243522	912489	871774
59 OISE.....	34097	121830	322754	121830	322754	121830	121830
60 PAS DE CALAIS.....	24852	128420	294493	128143	241904	122713	126443
61 PAYS DE LA LOIRE.....	139716	459620	1351423	451341	1409977	470932	458445
62 PUY DE DOME.....	34774	226400	531166	214442	456394	227407	219363
63 HAUTES PYRENEES.....	52873	194954	496400	194954	496400	194954	194954
64 HAUTES PYRENEES.....	22873	87474	228433	90982	228433	90982	90982
65 PYR. ORIENTALES.....	24154	96340	205496	104302	278795	98442	101434
66 PYR. ORIENTALES.....	62737	327708	803437	320570	622767	321113	323309
67 HAUT RHIN.....	54032	233400	338921	215494	543416	234756	227600
68 HAUT RHIN.....	15414	59434	15414	59434	15414	59434	59434
70 HAUTE SAONE.....	41814	81100	205400	78481	214252	40587	79471
71 SAONE ET LOIRE.....	50316	219952	541474	218010	542649	226594	220094
72 SARTHE.....	46184	194190	466661	194244	464233	188315	191217
73 SAVOIE.....	24602	118780	280430	118780	280430	118780	118780
74 HAUTE SAVOIE.....	37854	161252	401019	17271	361429	15814	165444
75 SEINE.....	—	—	—	—	—	—	—
76 SEINE MARITIME.....	111394	405928	1125220	405975	1117714	406973	406973
77 SEINE ET MARNE.....	—	—	—	—	—	—	—
78 YVELINES.....	—	—	—	—	—	—	—
79 CANTON DE SEVRES.....	32646	130216	330414	128782	320740	128782	128743
80 SEVRES.....	51211	194836	524411	197240	506443	190675	194700
81 TARN.....	33201	124888	318954	123437	334204	123337	123593
82 TARN ET GARONNE.....	14397	70144	177440	67251	146474	70400	68415
83 VAUCLUSE.....	34097	131808	315048	131808	315048	131808	131808
84 VAUCLUSE.....	35397	132148	351799	133463	349317	132740	132466
85 VENDEE.....	42125	164660	412064	165786	409124	164100	164949
86 VIENNE.....	44620	146874	327761	125441	339174	125766	125446
87 HAUTE VIENNE.....	14359	54400	14359	54400	14359	54400	54400
88 YONNE.....	48200	198610	383400	157247	349400	189402	189402
89 YONNE.....	24334	108316	287054	104431	285246	109224	106414
90 TERR. DE BELFORT.....	11845	47264	111484	45048	114444	46998	45419
91 ESSONNE.....	—	—	—	—	—	—	—
92 HAUTS DE SEINE.....	—	—	—	—	—	—	—
93 SEINE SAINT DENIS.....	—	—	—	—	—	—	—
94 VAL DE MARNE.....	—	—	—	—	—	—	—
95 VAL D'OISE.....	—	—	—	—	—	—	—
96 HAUTS DE SEINE.....	402474	1562824	3914790	1535105	4084744	1576379	1582491
97 HAUTS DE SEINE.....	44901	172190	426461	140774	447441	173408	171023
98 ECART-TYPE.....	35907	137510	334001	132146	360990	142870	134397

	OBSERVÉ		ÉQUILIBRE		ÉTAT POTENTIEL		
	Pop. Tot.	Pop. act.	Pop. Tot.	Pop. act.	Pop. Tot.	demande	offre
1 AIN.....	339240	150056	434450	146703	327854	143797	144576
2 AISNE.....	526350	195660	539416	197291	537511	195443	195682
3 ALLIER.....	346330	145492	368064	142713	372744	146211	144014
4 BASSES ALPES.....	104810	46116	103051	39790	107113	38422	37616
5 HAUTES ALPES.....	91790	36624	39776	36432	40731	36453	37414
6 ARDENNES.....	272470	272497	246074	246074	246074	246074	246074
7 ARDECHE.....	254930	94224	263020	102024	245442	94086	92946
8 ARDENNES.....	309340	110564	302930	111452	312749	115154	112766
9 ARTOIS.....	134840	49788	139000	52015	136764	50564	51569
10 AUVERGNE.....	270320	114768	276940	114444	276593	118281	118471
11 AUBE.....	278270	94000	273790	94000	273790	94000	94000
12 AVALOIS.....	241570	106716	272964	105414	272426	105395	105401
13 BASSES NORMANDIES.....	1470270	517130	1302330	491404	1300406	565293	551461
14 CALVADOS.....	519640	210274	527469	206210	536360	211408	209717
15 CANTAL.....	153120	60503	140607	60503	140607	60503	60503
16 CHARENTE.....	331020	143188	405400	174000	331020	129460	129400
17 CHARENTE MARITIME.....	426200	177448	471044	169907	441785	177459	172893
18 CHER.....	304000	120696	306011	119424	294912	117441	117090
19 CORSE.....	237660	98172	233094	94939	232327	90448	93474
20 COTE D'OR.....	421190	168008	426000	161405	420667	163487	162547
21 COTES DU NORD.....	506100	203400	500446	200195	495314	197455	199426
22 CREUSE.....	154880	60406	159674	60406	157413	63413	64432
23 DORDOGNE.....	314700	150134	371741	149337	344817	148462	148462
24 DROME.....	426260	171958	398121	141114	425456	172143	174174
25 DUOIS.....	342490	133952	332931	132040	327400	129455	130121
26 DURANDE.....	383340	155986	395570	140044	384122	154394	157999
27 EURE ET LOIRE.....	302210	125924	318110	119466	311933	119151	122945
28 FINISTERE.....	74140	297440	711115	297440	711115	297440	297440
29 GARONNE.....	478540	160388	490474	163444	484436	165133	154789
30 HAUTE GARONNE.....	650710	259270	666844	258704	697704	270494	264886
31 GERS.....	181540	71912	147456	78493	140575	75322	74032
32 GIRONDE.....	1009400	391074	10172012	391074	10172012	391074	391074
33 HERAULT.....	591400	197948	407754	204811	575643	195245	200790
34 ILLE ET VILAINE.....	652720	279720	650444	240441	640000	275905	278461
35 INDE.....	247160	100172	246620	100172	246620	100172	100172
36 INDE ET LOIRE.....	427						

Impression.
LA DOCUMENTATION FRANÇAISE
PARIS

AMÉNAGEMENT DU TERRITOIRE

LA DOCUMENTATION FRANÇAISE

29-31, QUAI VOLTAIRE 75340 PARIS CEDEX 07