

UTILISATION DE LA METHODE DE L'ADAPTATION  
DANS L'ESTIMATION DES TAUX DE TRANSITION :  
EXEMPLE DU SYSTEME EDUCATIF

Irina PEAUCELLE

N° 8209

Décembre 1981

UTILISATION DE LA METHODE DE L'ADAPTATION  
DANS L'ESTIMATION DES TAUX DE TRANSITION :  
EXEMPLE DU SYSTEME EDUCATIF

Irina PEAUCELLE

Décembre 1981

Je tiens à remercier C. GOURIEROUX pour la lecture attentive de ce texte, ses remarques et suggestions qui m'ont été très utiles. Les erreurs et imperfections sont bien sûr sous ma responsabilité.

## Sommaire

Introduction	1
I - PROBLEMATIQUE	1
1 - L'estimation des probabilités de transition dans les chaînes de Markov du premier ordre	1
2 - Les principes de l'adaptation	3
3 - Le modèle du Système Educatif et la méthode de l'adaptation appropriée	4
4 - Qualité des données statistiques	9
II - RESULTATS DE L'ANALYSE DE LA STRUCTURE DU SYSTEME EDUCATIF PAR LE MODELE D'ADAPTATION	11
1 - L'analyse rétrospective de la période 1972-1978	11
a) Le degré de précision des estimations avec la méthode de l'adaptation	11
b) Analyse rétrospective de la structure du système éducatif	18
2 - Simulations prospectives	24
a) Prévisions prolongeant les tendances passées	24
b) Problèmes de la recherche d'une structure optimale du système éducatif	28
Conclusion	33
Annexe I	34
Annexe II	35
Annexe III	37
Annexe IV	38
Bibliographie	41

## Introduction

L'objet de ce travail est double : premièrement, on analyse la méthode d'adaptation appliquée à un modèle markovien, deuxièmement, on utilise ces résultats pour la modélisation du système éducatif français.

On a cherché à améliorer la modélisation du système éducatif incorporée dans le modèle macroéconomique AGORA (1978). Le modèle de l'Education cité est un modèle socio-économique fondé sur la théorie des chaînes de Markov <sup>(1)</sup>. La dynamisation de la matrice markovienne est un problème qui intéresse beaucoup d'auteurs. On propose dans ce qui suit d'utiliser la méthode d'adaptation en tant qu'approche pour résoudre ce problème.

La méthode d'adaptation fait partie de la théorie du contrôle optimal qui est utilisée dans de nombreuses disciplines. On expose ci-après une analyse des propriétés de cette méthode en l'appliquant au modèle markovien et dans le contexte spécifique du système éducatif, qui rend son fonctionnement moins souple que dans le cas général.

Avec le modèle rectifié grâce à la méthode d'adaptation, on procède à l'analyse prévisionnelle des flux des élèves dans l'enseignement et des sorties dans la vie active. D'une part, en prolongeant les tendances observées, d'autre part, en analysant les possibilités d'adaptation du système éducatif aux besoins exprimés par le système productif <sup>(2)</sup>. Cette dernière analyse sert de test à la méthode d'adaptation.

## I - PROBLEMATIQUE

### 1 - L'estimation des probabilités de transition dans les chaînes de Markov du premier ordre

Considérer une chaîne de Markov signifie étudier le nombre fini d'états du système : vecteur  $a_1, a_2, \dots, a_r$  en supposant que la probabilité d'être dans l'état  $a_j$  à  $t$  dépend des états où se trouvait le système à des moments antérieurs par l'intermédiaire de l'état  $a_i$  à  $t-1$ .

(1) Ce type de modèle est utilisé dans de nombreux pays pour la planification de l'éducation.

(2) Le problème méthodologique a été posé et une méthode de résolution proposée déjà en 1967 par J. BENARD.

Une probabilité de transition de l'état  $a_i$  à  $t-1$  à l'état  $a_j$  à  $t$  est une valeur donnée  $p_{ij}$ . Les probabilités de transition correspondant au vecteur des états  $a_j$  ( $j=1, \dots, 2$ ) forment une matrice avec les éléments  $p_{ij}$  qui ont les propriétés suivantes :

$$p_{ij} \geq 0 ; \quad \sum_{j=1}^r p_{ij} = 1 ,$$

et la probabilité théorique d'être dans l'état  $a_j$  à  $t$  est égale à :

$$a_{jt} = \sum_{i=1}^r a_{i,t-1} p_{ij} \quad (j=1, \dots, r) \quad (*)$$

Les processus de Markov sont largement employés pour décrire les aspects stochastiques des phénomènes sociaux, spécialement en tant que modèle de prévision. La qualité de ces modèles dépend des estimations de la matrice de transition.

Quand on dispose d'une information complète sur les flux de transition d'un état à l'autre, on est à même de concevoir un modèle de bonne qualité. Mais ces flux sont souvent mal connus. Quand on possède des séries temporelles de données agrégées relatives aux proportions des individus dans chaque état, il est d'usage d'utiliser les méthodes économétriques<sup>(1)</sup> (par exemple la méthode des moindres carrés ordinaires) pour estimer les probabilités de transition. Effectivement, on considère dans ce cas que la fréquence dans l'état  $\hat{a}_{jt}$  se définit ainsi :

$$\hat{a}_{jt} = \sum_i \hat{a}_{i,t-1} p_{ij} + v_{jt} \quad (j=1, \dots, r) \quad (**)$$

où  $v_{jt}$  est une variable des erreurs, dont  $E v_{jt} = 0$ , c'est-à-dire

$$E \hat{a}_{jt} = \sum_i a_{i,t-1} p_{ij}$$

La forme de l'équation de régression linéaire (à partir de(\*\*)) permet d'estimer en se basant uniquement sur les observations de fréquences dans les états  $a_j$  et  $a_{i,t-1}$  ( $i=1, \dots, 2$ ) sans nécessairement avoir d'observations sur le nombre de passages.

(1) Cf. ANDERSON T.W., GOODMAN L.A. (1957), MADANSKY A. (1959), LESTER G. TELSER (1963), MCGUIRE T.W. (1969), T.C. LEE et alii (1970).

La modélisation des processus sociaux avec les chaînes markoviennes a été développée par intégration de différents aspects. Par exemple, par l'incorporation des facteurs exogènes déterminant les probabilités de transition (implicitement supposés indépendants du temps) :

$$P_{ij} = f_{ij}(x_k, \beta_k) \quad (i, j = 1, 2, \dots, r)$$

où  $\beta_k$  sont les paramètres à estimer, qui relient des facteurs exogènes  $x_k$  aux probabilités de transition (R.B. GINSBERG (1972), E. MACRAE (1977)).

Ou encore, dans l'analyse du choix de transition par les personnes vers un état quelconque, on introduit la fréquence d'occupation des états  $a_i$  ( $i=1, 2, \dots, r$ ) par la population entière. Cette fréquence devient le facteur principal de l'attraction individuelle vers différents états et de ce fait détermine la valeur des taux de transition (J. COULISK (1976) (1978)).

Une autre extension concerne le temps de transition qui est considéré par des auteurs soit comme une variable continue (PONTRIAGUINE L.S. et al. (1962), R. GINSBERG (1971)), soit comme une variable à durée de transition mémorisée et variable selon les sous-groupes (BOUDON R. (1973), SCHINNAR A.P. et S. STEWMAN (1978)).

Dans ce qui suit on discutera de l'estimation des probabilités de transition dans les circonstances où on ne dispose pas d'information complète sur les nombres de personnes qui transitent. On utilise les données agrégées relatives aux fréquences d'occupation des états pour estimer la matrice changeant dans le temps. Ce problème est posé en temps fini et avec une mémoire ne dépassant pas les périodes qui se suivent immédiatement. Ainsi, on procédera à l'estimation des probabilités de transition avec la méthode de l'adaptation <sup>(1)</sup>.

## 2 - Les principes de l'adaptation

Les méthodes de l'adaptation servent à l'élaboration des modèles de simulation qui diffèrent des autres modèles de ce genre par leur capacité de refléter les changements dynamiques des conditions extérieures au processus et de rendre compte, d'autre part, de la validité informationnelle de différents éléments des séries temporelles utilisées. Ceci permet de donner des prévisions plus exactes des états suivants du système modulé.

(1) E.C. MACRAE (1977) propose d'utiliser la méthode des moindres carrés généralisés et KELTON C. (1981) (Econometrica, vol. 49, N°2) la programmation quadratique pour l'estimation, à partir des données agrégées, des processus de Markov variant dans le temps.

Les méthodes de l'adaptation comme modèles d'estimation des paramètres se sont développées d'une part à partir du modèle de lissage exponentiel de BROWN R. (1961) (1963). Les modifications et les généralisations de ce modèle (THEIL H., WAGE S. (1964), CHOW W.M. (1965), BOX G.E.P., JENKINS G.M. (1962), TRIGG D.W., LEACH A.G. (1967)) ont donné toute une famille de méthodes d'adaptation avec des caractéristiques diverses, décrite, par exemple, par LOUKACHIN (1979).

D'autre part, pour les systèmes complexes et dans le cas d'une grande incertitude sur la nature du système, la méthode d'adaptation utilisée se rapproche des méthodes d'approximation stochastique dont l'usage multiple est analysé dans KUSHNER H.J., CLARK D.S. (1978) et IOUDIN D.B. (1979).

Le processus de l'adaptation se passe de la façon suivante. Au départ le modèle est dans un état initial  $t_0$  (les valeurs de ses paramètres sont connues). On fait la prévision avec ce modèle (où l'agent économique fait le choix de son comportement). Au moment suivant  $t_1$  on compare les résultats calculés et les valeurs observées de la série à  $t_1$  (où le résultat du choix de l'agent est analysé, à savoir : s'il est compatible avec l'état actualisé de l'environnement). L'erreur de la prévision rétroagit sur l'entrant du système et elle est utilisée, conformément à la logique du modèle, pour permettre le passage à un état correspondant le mieux à la dynamique de la série observée (ou, plus généralement, à l'environnement). Les corrections se font par itérations, jusqu'à ce que la correction soit suffisante pour minimiser l'erreur.

Le mécanisme de l'adaptation est conçu par le chercheur intuitivement, en fonction du système étudié. Etant donné a priori, il est vérifié par la suite d'une façon empirique.

La méthode d'adaptation qu'on décrit ci-dessous est choisie en relation avec le problème étudié qui est le système éducatif.

### 3 - Le modèle du Système Educatif et la méthode de l'adaptation appropriée

Nous étudions l'appareil de formation initiale appelé ci-dessous "Système éducatif". Il est considéré en termes de flux aux différents niveaux.

L'appareil éducatif est repéré en 6 niveaux scolaires : primaire, 1er cycle du secondaire, 2è cycle court, 2è cycle long, 1er cycle de l'enseignement supérieur, 2è et 3è cycles de l'enseignement supérieur et grandes

écoles. Les entrées sont connues d'après des hypothèses démographiques. Elles sont donc considérées comme des variables exogènes.

Les passages d'un niveau éducatif à un autre sont repérés à travers une matrice de passage comme dans les travaux de R. STONE (1971) (1976). Cette matrice détermine aussi les flux de sortie vers la vie active. Etant donné une structure d'enseignement, tous les passages ne sont pas possibles ; ainsi seuls certains coefficients ne sont pas nuls. La valeur des coefficients non nuls peut être estimée (avec le degré de précision différent selon les coefficients <sup>(1)</sup>) à partir des statistiques disponibles.

Dans un travail antérieur, nous avons fait une projection des effectifs scolaires et des sorties à long terme à partir des hypothèses démographiques (fécondité basse et constante) et d'une hypothèse de constance des coefficients de la matrice (AGORA (1978)).

Ici on recherche à dynamiser le modèle éducatif, c'est-à-dire faire en sorte que la matrice de transition ait une valeur changeante par période de quelques années ou par année. Les statistiques disponibles pour évaluer les redoublements, passages, abandons et sorties sont de fiabilité variable et souvent font défaut. Ainsi, les coefficients de transition ne peuvent pas être estimés directement. Ils peuvent à l'inverse être évalués à partir des facteurs observables (flux d'élèves). Les méthodes économétriques (comme elles étaient utilisées par L.G. TELSER (1963) et T.C. LEE et al. (1970)) ne conviennent pas pour ce modèle, car premièrement, on ne dispose pas encore de séries suffisamment longues de flux pour procéder aux estimations des paramètres et, deuxièmement, les coefficients qui seraient les probabilités de passage, resteraient fixes dans les simulations.

Il serait tentant de considérer les coefficients de transition dans le système éducatif comme les fonctions non seulement des flux observables et de la démographie mais aussi de facteurs de natures économiques (comme le niveau de production, des dépenses, de la situation sur le marché du travail et autres). Mais le système éducatif analysé au niveau macro est assez rigide. En effet, seuls 5 coefficients sur les 22 qui ne sont pas nuls dans la matrice de transition résument l'alternative dans le comportement

---

(1) Effectivement, les tableaux d'origine scolaire qui sont vraiment fiables ne couvrent que l'enseignement du second cycle.

des élèves. Ce sont les coefficients de passage du premier cycle du secondaire dans le cycle court ou long du deuxième degré ; le coefficient de sortie après la classe terminale dans la vie active ou ceux de passage vers le supérieur et à l'intérieur du supérieur.

N'ayant pas dégagé dans cette analyse globale les facteurs provoquant les changements de ces coefficients (manque de séries pour tester des hypothèses) on procède dans ce qui suit à l'analyse de variation de tous les coefficients en fonction des flux.

Le modèle de simulation est réduit à deux équations :

$$(1) X_t = p \cdot X_{t-1} + E_t$$

où  $X_t$  - vecteur d'activité du système scolaire hormis le pré-élémentaire (six niveaux : primaire, 1er cycle du secondaire, 2è cycle court, 2è cycle long, 1er cycle du supérieur, 2è et 3è cycles du supérieur)

$p$  - matrice des taux de passage d'un niveau à l'autre

$E_t$  - vecteur des entrées de l'appareil scolaire qui comprend comme première composante les entrants âgés de 6 ans.

$$(2) Y_t = s X_t$$

$Y_t$  - vecteur de sorties par niveau de formation (six niveaux suivants : VI, V bis, V, IV, III et II + I)

$s$  = matrice de taux de sortie de l'appareil scolaire à la fin de la scolarisation par niveau de formation.

Les matrices  $p$  et  $s$  étant complémentaires, la somme des sommes de leurs coefficients de la ligne est égale à l'unité. En effet, tous les enfants entrant sortent un jour ou l'autre du système éducatif (on a négligé les phénomènes de mortalité faible à cet âge).

$$(3) \sum_{j=1}^6 p_{ij} + \sum_{j=1}^6 s_{ij} = 1 \quad i = 1, \dots, 6$$

La matrice d'ensemble a les propriétés de matrice de probabilité de transition de type markovien :

$$(4) p_{ij} \geq 0 \quad s_{ij} \geq 0$$

Les matrices de transition  $p_{ij}$  et  $s_{ij}$  ne restent pas constantes, on utilise la méthode de l'adaptation pour les estimer annuellement. Supposons au départ que, pour chaque nouvelle période, on ne connaisse pas les coefficients structurels du système scolaire (taux de redoublement, de passage, d'abandon et de sortie). Alors ils seront pris égaux à ceux de la période précédente. D'après les flux observés et les flux calculés avec cette hypothèse des coefficients pour la période, il est possible de corriger leurs valeurs. Ainsi, un processus itératif donne une nouvelle matrice des coefficients à chaque période.

On fixe comme objectif à ce processus itératif de minimiser les écarts entre les flux observés et ceux calculés par le modèle.

$$(5) J_t(\phi, \alpha) = |F_t^r - F_t| < \epsilon$$

avec  $F_t^r$  observation des flux :  $X_t^r, Y_t^r$

$F_t$  flux calculés par le modèle :  $X_t, Y_t$

$J_t(\phi, \alpha)$  fonction à minimiser, des facteurs  $\phi$  et des coefficients structurels  $\alpha$  du modèle.

L'algorithme que nous utilisons s'inspire des travaux de TOVSTOUKHA (1974) <sup>(1)</sup>. Le modèle linéaire dynamique avec les équations (1) et (2) peut être résumé, pour les éléments du vecteur  $F_t$  :

$$(6) F_{it} = \sum_{j=1}^{j=N} \alpha_{ij}^t \phi_{ij}^t + u_{it}$$

où  $(\alpha_{ij})$  matrice de coefficients :  $(p_{ij}$  et  $s_{ij})$

$(\phi_{ij})$  matrice de facteurs : vecteurs démographiques ( $E_t$ ), des effectifs ( $X_t$ ), des sorties (diplômés et abandons ( $Y_t$ ))

$u_{it}$  vecteur des variables aléatoires, avec une distribution normale, la moyenne nulle et la dispersion finie.

---

(1) Pour plus de précisions sur les critères de ce choix, voir I. PEAUCELLE (1978).

L'algorithme itératif d'adaptation s'écrit pour le vecteur ligne  $\alpha_i$  de la matrice :

$$(7) \quad \alpha_i^t[n] = \alpha_i^t[n-1] + A_i[n] \nabla_{\alpha_i} Q_i(\phi_i^t[n], \alpha_i^t[n-1])$$

où

$$(8) \quad Q_i(\phi_i^t[n], \alpha_i^t[n-1]) = F_{it}^r - \sum_{j=1}^N \alpha_{ij}^t[n-1] \phi_{ij}^t[n]$$

$i, j = 1, \dots, N$

$n = 1, 2, \dots$  les itérations

$t$  époque

$\nabla_{\alpha} Q$  gradient de la fonction  $Q$

Pour les valeurs initiales  $\alpha_i^1[0]$  de la première période, on prend les coefficients estimés pour l'année de base. Pour une période  $t > 1$ , on prend comme valeurs initiales  $\alpha_i^t[0]$  — les coefficients  $\alpha_i^{t-1}$  obtenus à la période précédente.

La matrice  $A$  ayant les éléments suivants (9) assure la convergence de l'algorithme et satisfait aux conditions décrites précédemment :

$$(9) \quad a_{ij}[n] = \frac{1}{\sqrt{n}(\phi_{ij}^t[n])}$$

Et l'algorithme s'écrit alors

$$(10) \quad \alpha_{ij}^t[n] = \alpha_{ij}^t[n-1] + c_{ij} b_{ij}$$

où :

$$(11) \quad b_{ij} = \left( \frac{(F_{it}^r - \sum_{j=1}^N \alpha_{ij}^t[n-1] \phi_{ij}^t[n])}{\sqrt{n}(\phi_{ij}^t[n])} \right)$$

Etant donné que la matrice  $\alpha$  est markovienne et la somme de ses coefficients égale à 1 dans chaque ligne, l'ajustement de leur valeur au cours de l'itération doit conserver cette contrainte. Une manière de l'introduire dans les formules est de donner une valeur particulière aux coefficients  $c_{ij}$  afin que la somme des écarts d'ajustement des coefficients d'une ligne soit nulle. Ainsi :

$$(12) \quad \begin{cases} c_{ij} = \frac{1}{\sum_j b_{ij}} & \text{quand la correction } b_{ij} \text{ est positive} \\ c_{ij} = \frac{-1}{\sum_j b_{ij}} & \text{quand la correction } b_{ij} \text{ est négative.} \end{cases}$$

On trouvera dans l'annexe II l'algorithme d'adaptation dans les termes du modèle éducatif.

#### 4 - Qualité des données statistiques

Pour estimer la structure dynamique du modèle, on utilise les données globales des flux des élèves, fournies par le Ministère de l'Éducation. Les effectifs par niveau de formation sont connus annuellement d'après les "Statistiques des enseignements" (SEIS). Toutefois, les enseignements supérieurs ne dépendant pas du Ministère de l'Éducation ne sont pas couverts avec beaucoup de précision ; le partage entre le premier cycle et le deuxième plus le troisième cycle du supérieur s'établit approximativement, et le double compte (préparation à la fois de plusieurs diplômes) n'est pas exclu.

Les sorties de la formation initiale du système sont analysées selon la nomenclature des niveaux de formation (Annexe I). Les statistiques disponibles couvrent la période de 1972/73 à 1977/78 <sup>(1)</sup>, ce qui réduit à 6 points le champ de l'analyse rétrospective. Les chiffres indiquant les sorties à un niveau donné regroupent les sortants ayant obtenu le diplôme correspondant au niveau de formation et les effectifs du niveau suivant abandonnant les études avant la fin de la scolarité dans cette formation. Ces derniers seront appelés les abandons. La distinction entre les abandons et les sorties diplômées parmi les sortants par niveau à partir des renseignements du SEIS n'est pas toujours aisée.

---

(1) Source : SEIS, sorties "nettes". Pour les précisions de la nomenclature, les sources et les méthodes des bilans de sortie, voir "Bilan formation-emploi 1973" (1978), et l'article de J. AFFICHARD dans : "L'entrée des jeunes dans la vie active" (1981).

Les flux d'entrée dans le système éducatif sont les moins connus. Le problème ne se pose pas pour les flux annuels d'entrée dans la première année de l'enseignement élémentaire (enfants atteignant l'âge de 6 ans). Pour les autres niveaux, le vecteur des entrées<sup>(1)</sup> correspond aux soldes d'entrée-sortie définitifs couvrant les décès, les flux vers l'enseignement spécialisé (exclu de cette analyse), l'immigration à l'âge scolaire et les entrées tardives dans le système éducatif. Dans l'enseignement supérieur, la reprise des études après une rupture de la scolarité est fréquente. Ces entrées dans le supérieur, qui ne sont pas les flux provenant directement du secondaire, sont mal connues.

Avant de procéder à l'adaptation, on vérifie que la contrainte de cohérence entre les flux soit respectée :

$$(13) \quad X_{t+1} - X_t = E_{t+1} - Y_t$$

Compte tenu des données dont on dispose, qui n'assurent pas cette contrainte, on procède à la correction à la proportionnelle des flux, dont les estimations initiales sont les plus approximatives. Cette correction est équivalente à l'ajustement qu'on doit faire des coefficients de transition dans la matrice markovienne pour assurer l'égalité à l'unité de la somme des coefficients en colonne.

Les corrections portent sur les effectifs du supérieur, les entrées dans le supérieur et tous les niveaux de sorties.

Le coefficient de correction :

$$(14) \quad K_t = \frac{\sum_{i=1}^4 E_{i,t+1} + \sum_{i=1}^6 Y_{i,t} - \sum_{i=1}^4 X_{i,t+1}}{\sum_{i=5}^6 X_{i,t+1} + \sum_{i=1}^6 Y_{i,t} - \sum_{i=5}^6 E_{i,t+1}}$$

Dans une première étape, on a fait cette correction des statistiques. Les chiffres obtenus sont considérés comme étant des observations par rapport auxquelles on fait fonctionner l'algorithme d'adaptation.

-----  
(1) Dans notre analyse.

II - RESULTATS DE L'ANALYSE DE LA STRUCTURE DU SYSTEME EDUCATIF PAR LE MODELE D'ADAPTATION.

1 - L'analyse rétrospective de la période 1972-1978

a) Le degré de précision des estimations avec la méthode de l'adaptation.

On trouvera en Annexe II les équations et l'algorithme de l'adaptation du modèle utilisé (décrits dans le chapitre I, paragraphe 3).

On recherche les valeurs de 22 coefficients des matrices  $p$  et  $s$  que l'analyse du système éducatif suppose non nuls. Dans un premier temps on a estimé, pour toute la période, les coefficients d'après diverses enquêtes statistiques plus ou moins complètes et précises.<sup>(1)</sup> Dans un second temps on a réalisé l'adaptation. D'abord l'adaptation concerne les 22 coefficients  $\alpha_{ij} \neq 1$  ( $p_{ij}$  et  $s_{ij}$ ). Ils sont corrigés à chaque étape de l'opération du produit  $c_{ij} b_{ij}$  (équations (11) et (12)) :

$b_{ij}$  dépend de l'écart entre les valeurs observées et les valeurs résultant des anciennes valeurs des paramètres ( $b_{ij}^1$  pour les effectifs et  $b_{ij}^2$  pour les sorties)

$c_{ij}$  est un facteur de correction pour garder aux matrices leur nature markovienne (respectivement  $c_{ij}^1$  pour les effectifs et  $c_{ij}^2$  pour les sorties)

Dans les estimations, bien évidemment, les coefficients nuls ne sont pas changés. Ceci est obtenu en annulant la valeur des écarts ( $b_{ij}$ ) calculés pour ces coefficients. C'est d'ailleurs, une particularité de l'adaptation utilisée ici : la multiplicité des contraintes sur les coefficients (valeurs nulles, total des lignes constant) limite les variations qui résultent de l'adaptation à chaque itération.

./.

(1) Estimations faites pour le modèle AGORA (1978).

Après on a procédé à l'adaptation qui tient compte de ces contraintes, en limitant l'adaptation à 15 coefficients. Les sept autres résultent d'une résolution directe du système d'équations, c'est-à-dire des contraintes.<sup>(\*)</sup> Ainsi, dans l'algorithme de l'adaptation proprement dit, les valeurs des écarts ( $b_{ij}$ ) sont annulées pour ces coefficients ( $\alpha_{ij}$ ).

Pour juger de la qualité de ces deux méthodes, on compare les résultats des simulations avec l'adaptation de la matrice structurelle du système (après 5 itérations à l'intérieur de chaque période<sup>(1)</sup>) et une simulation faite sur une matrice fixe se rapportant chaque fois aux valeurs observées<sup>(2)</sup>. Premièrement, on compare les erreurs globales moyennes (absolues et relatives) des simulations, c'est-à-dire les sommes des erreurs par niveau de formation et de sortie.

TABLEAU I  
ERREUR ABSOLUE TOTALE (moyenne par année)  
DES SIMULATIONS  
(en milliers de personnes)

$$(15) \quad EA = \frac{1}{m} \sum_{t=1}^m \sum_{j=1}^{12} |P_j^t - A_j^t| \quad \text{où} \quad m = 5 \text{ périodes (de 1972 à 1977)}$$

$P_j^t$  - valeur des effectifs et des sorties calculés  
 $A_j^t$  - valeur des effectifs et des sorties observés.  
 $j=1, \dots, 6$ -niveaux de formation,  
 $j=7, \dots, 12$ -niveaux de sorties

	SIMULATIONS AVEC		
	LA MATRICE FIXE	ADAPTATION DE TOUS LES COEFFICIENTS	CALCUL DE 7 COEFFICIENTS ET ADAPTATIONS DES AUTRES
Effectifs	84,7	46,2	16,4
Sorties	63,6	37,7	29,3

(1) Au-delà de 5 itérations la convergence de l'algorithme de l'adaptation est très ralentie.

(2) Les méthodes de comparaison des modèles sont décrites, par exemple, par Henri THEIL (1970) et A. ERMILOV (1979).

(\*) Ceci pour éliminer la redondance existante dans les équations (1)-(4).

On aperçoit l'avantage des modèles adaptatifs quant à la diminution des erreurs dans les estimations par rapport à la simulation avec la matrice fixe. La méthode faisant varier tous les coefficients non nuls de la matrice apparaît moins efficace que celle où on calcule a priori quelques coefficients de transitions.

Dans les deux premières colonnes du tableau l'erreur la plus grande atteint les estimations des effectifs, par contre, ce sont les sorties qui sont les plus biaisées dans le dernier cas. Ces erreurs absolues, en milliers de personnes, ont naturellement une importance différente selon la valeur des flux sur lesquels elles portent.

TABLEAU 2  
ERREUR RELATIVE MOYENNE

$$ER = \frac{EA}{\frac{1}{m} \sum_{t=1}^m \sum_j A_j^t}$$

où EA = erreur absolue (TABLEAU I)

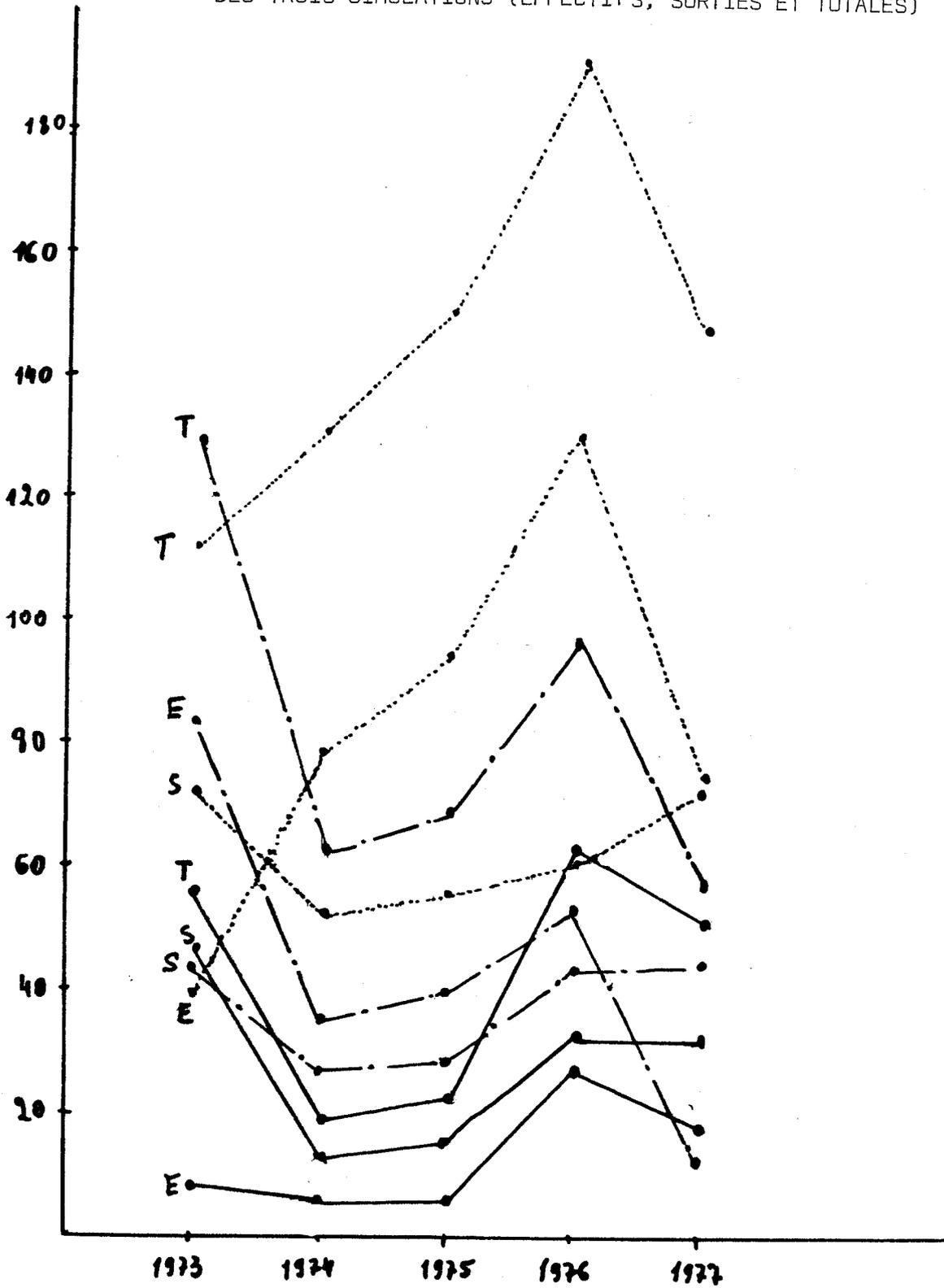
$A_j^t$  = valeurs observées (respectivement : effectifs, sorties et leurs sommes)

	SIMULATION AVEC			%
	MATRICE FIXE	ADAPTATION SUR TOUS LES COEFFICIENTS	CALCUL DE 7 COEFFICIENTS ET ADAPTATION SUR LES AUTRES	
effectives	0,89	0,49	0,17	
sorties	10,02	5,94	4,61	
totales	1,47	0,83	0,45	

En valeur relative les erreurs portent surtout sur les estimations des sorties, car elles cumulent, d'une part, les imprécisions sur les coefficients de structure (taux d'abandons et de sorties diplômées), d'autre part, les erreurs non corrigées sur les effectifs à la période précédente dont ils sont fonctions (voir l'équation (2)). La méthode de l'adaptation améliore sensiblement les résultats dans les deux cas de simulation (colonne 2 et 3).

De plus, on présente les valeurs des erreurs globales des simulations année par année dans les graphiques 1 et 2.

ERREUR ABSOLUE EN MILLIERS DE PERSONNES  
DES TROIS SIMULATIONS (EFFECTIFS, SORTIES ET TOTALES)



— 7 coefficients calculés et adaptation  
- · - adaptation de tous les coefficients  
····· simulation avec la matrice fixe

Pour les variantes de simulation avec adaptation, deux périodes apparaissent spécialement "difficiles" pour atteindre le but de l'adaptation qui est la diminution maximale de l'erreur entre les valeurs calculées et observées. Ce sont les années 1973 et 1976. Les estimations pour l'année 1973 sont fortement marquées par les coefficients de l'année de base qui est 1972 avec toute l'incertitude qu'on a sur ces coefficients. En 5 itérations pour 1973 l'erreur absolue reste importante. Pour la simulation avec les coefficients fixes c'est, au contraire, la moindre erreur car la structure du système a justement le moins changé depuis l'année précédente et les coefficients fixes donnent de relativement bonnes prévisions.

A partir de cette année 1973, l'approximation est meilleure pendant deux périodes et la détérioration a lieu en 1976 où l'on enregistre les erreurs les plus importantes dans les trois variantes. Cette chute de la qualité sera discutée dans le paragraphe suivant, où on examine la situation plus en détail.

Le graphique 2 retrace les erreurs relatives aux effectifs de l'année  $t$ . L'erreur relative maximale sur la période de 1972-1977 est relevée dans la simulation avec la matrice fixe. Elle dépasse en 1976 2 % des effectifs totaux. Dans le second modèle adaptatif elle est, pour la même année, 0,55 % des effectifs. Tandis que dans la variante de l'adaptation de tous les coefficients non nuls de la matrice (premier modèle), l'erreur maximale est en 1973 égale à 1,4 % des effectifs de l'année <sup>(1)</sup>.

Etant donné les résultats, on a poursuivi l'analyse avec le meilleur modèle. La discussion qui suit concernant la période rétrospective, utilise donc le modèle de l'adaptation avec le calcul a priori de 7 coefficients.

---

(1) Rappelons que la sensibilité des premiers résultats de l'adaptation aux paramètres de la période de base ( $t_0$ ) est très forte.

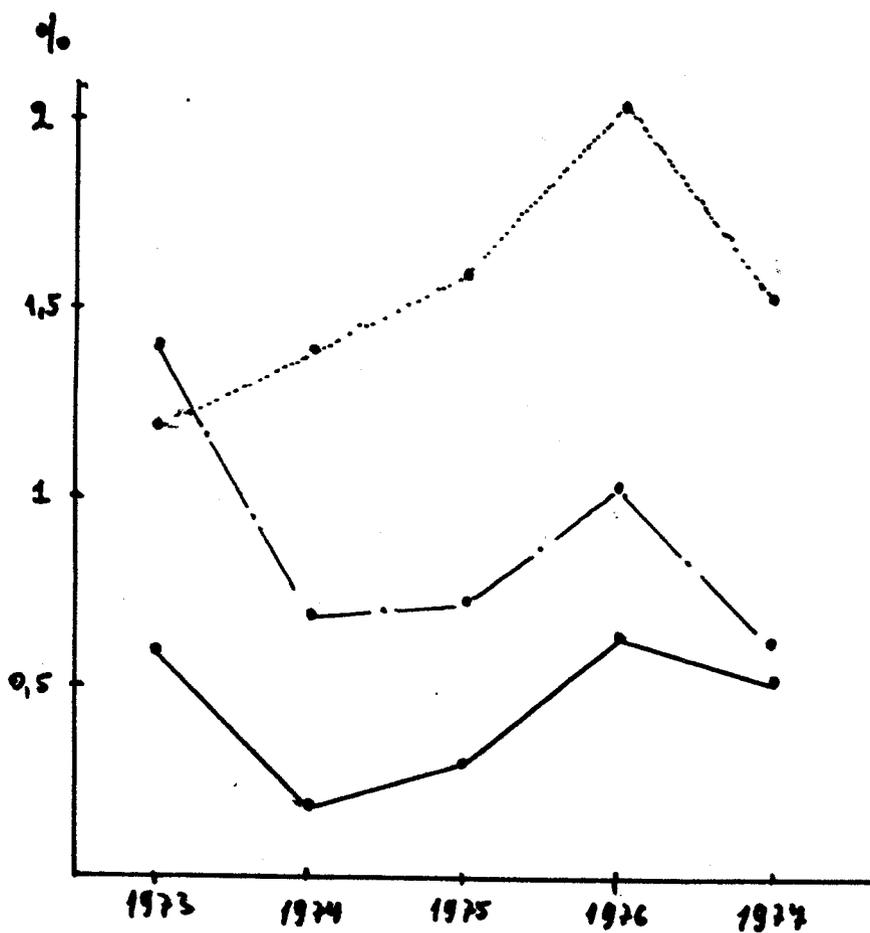
GRAPHIQUE 2

ERREURS RELATIVES TOTALES (SOMME DES ERREURS SUR LES EFFECTIFS ET LES SORTIES) DE L'ANNEE t PAR RAPPORT AUX EFFECTIFS DE L'ANNEE t.

$$ER_t = \frac{\sum_{j=1}^1 |P_j^t - A_j^t|}{\sum_{j=1}^6 A_j^t}$$

j = 1, ..., 6 niveaux de formation

j = 7, ..., 12 niveaux de sorties



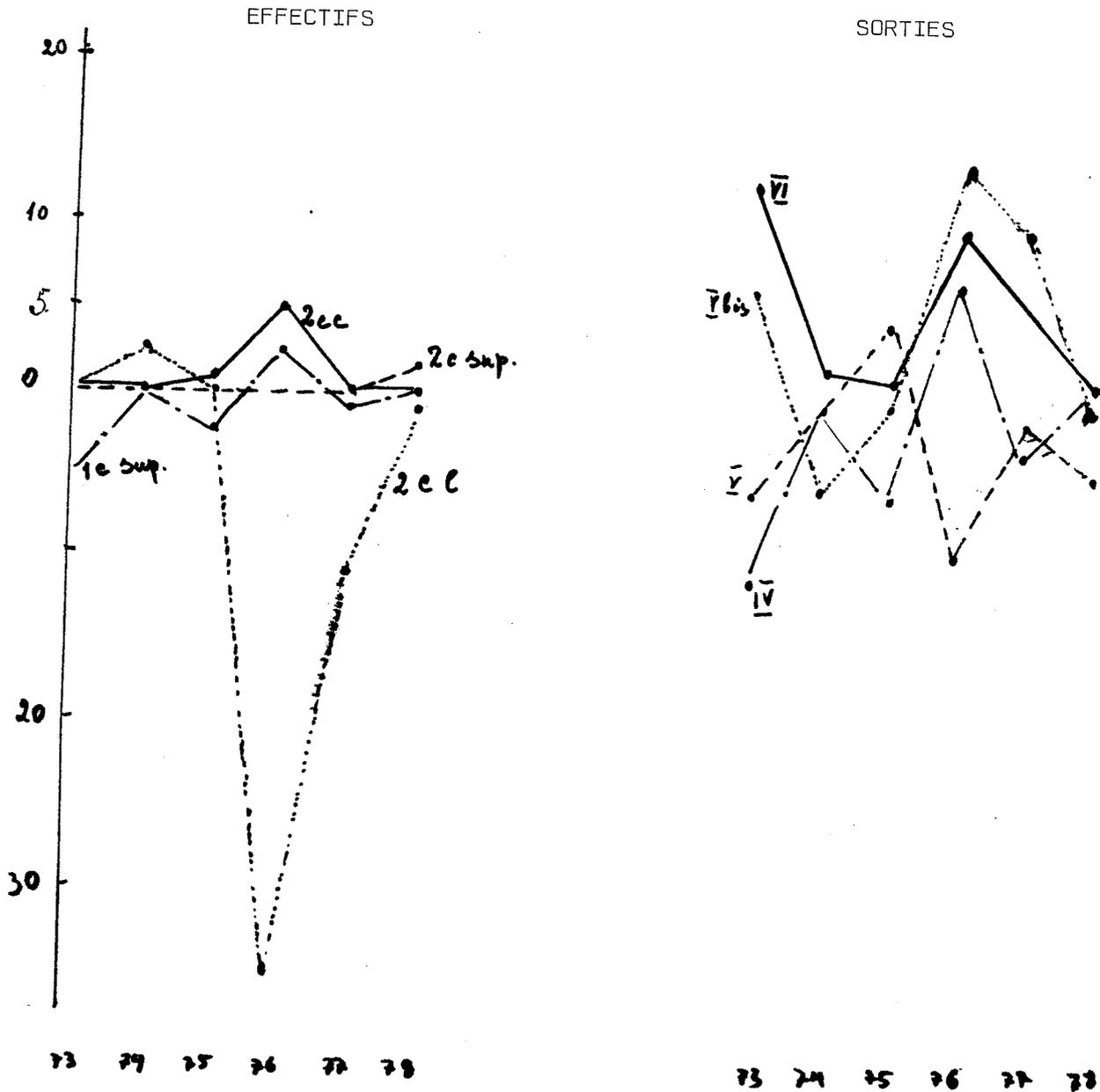
- 7 coefficients calculés et adaptation
- · - · adaptation de tous les coefficients
- · · · · simulation avec la matrice fixe.

./.

Selon les niveaux d'études et les niveaux de sortie, la qualité d'approximation n'est pas la même (graphique 3).

GRAPHIQUE 3

ECARTS DES SIMULATIONS PAR NIVEAUX (EN MILLIERS DE PERSONNES)



Les écarts sont faibles dans la chronique des effectifs, sauf la surestimation de 37 milles en 1976 dans le second cycle long. Cette dernière ne présente pourtant que 3,8% des effectifs concernés. Les autres écarts sont inférieurs 1 % en variation relative. L'incertitude est plus grande du côté des flux de sortie <sup>(1)</sup>. L'erreur selon les années et les niveaux atteint jusqu'à 14 % des flux (en 1976 pour le niveau Vbis de sortie). Cette approximation relativement faible vient de ce que les contraintes gênent l'efficacité de la formule d'adaptation. Celle-ci calcule des écarts pour tous les coefficients et voudrait les changer tous alors qu'on la contraint à n'en changer que quelques uns.

A l'intérieur de chaque période tous les coefficients non nuls convergent asymptotiquement vers leurs valeurs optimales, quand n augmente (n = nombre d'itération) (Annexe III : approximation des coefficients du modèle (année 1976) ).

b) Analyse rétrospective de la structure du système éducatif.

Degré de variance de la structure

Les résultats d'estimations des coefficients de transition avec la méthode de l'adaptation sont données dans l'Annexe IV (TABLEAU A1). Le graphique (p. suivante) montre l'évolution de la valeur de chaque coefficient entre 1972 et 1976.

Pour chaque coefficients le taux moyen sur la période rétrospective

de variance ( $\text{var} = \frac{\sum_i^2}{N} - \frac{(\sum a_i)^2}{N^2}$ ) et l'écart-type ont été calculés. On donne

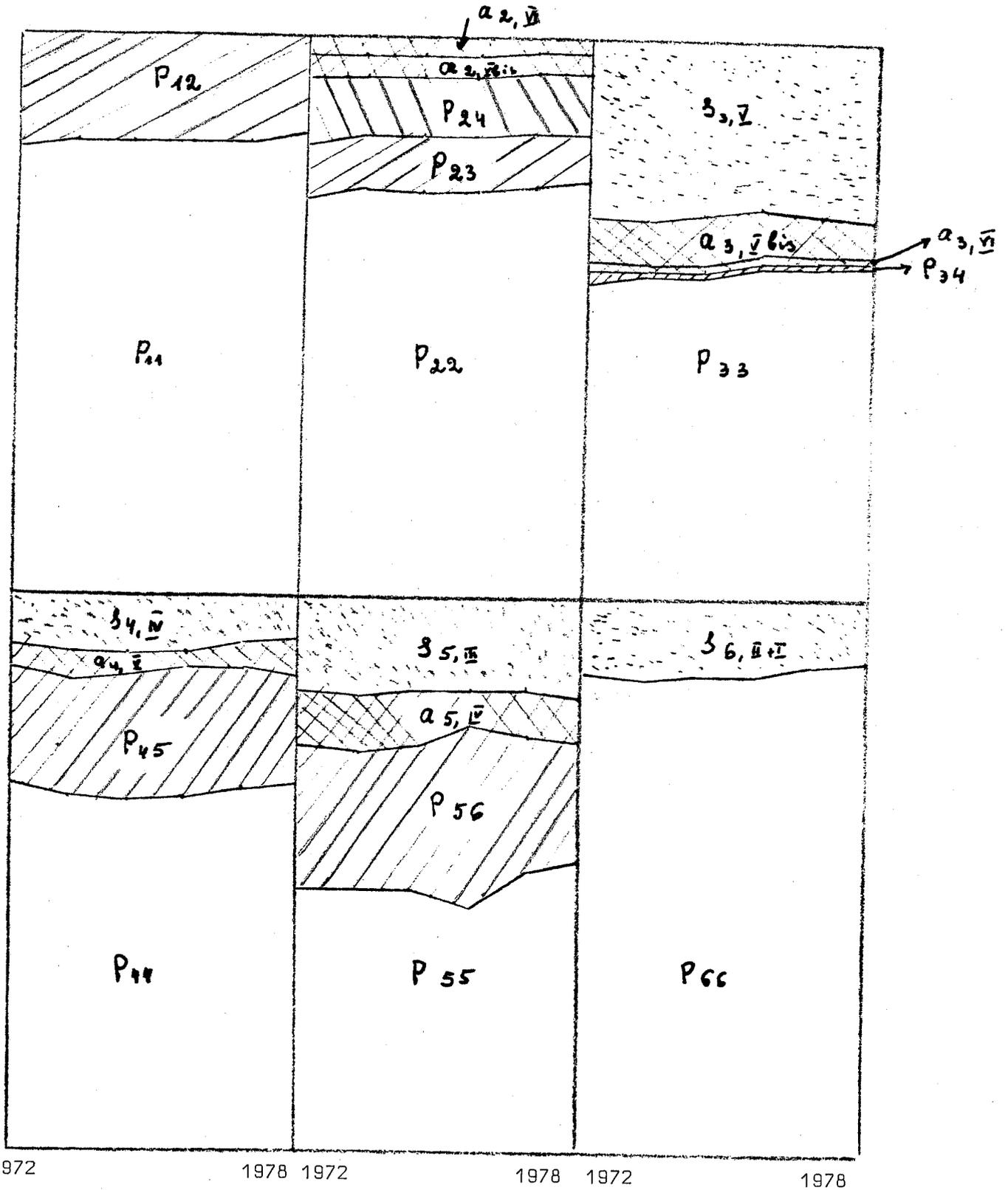
le tableau des résultats en Annexe IV, TABLEAU A2.

---

(1) Les erreurs portant sur les sorties des niveaux III et II+I étant nulles n'apparaissent pas sur le graphique 3.

Evolution des valeurs des coefficients structurels  
du modèle éducatif entre 1972 et 1978

□ % d'effectifs restant dans le cycle    ▨ % de passage    ▩ % d'abandons    ▤ % de sort: diplômées



Les coefficients de la diagonale

Les coefficients de la diagonale de la matrice de passage ( $p_{ij}$ ,  $i = j$ ) représentent les proportions des effectifs qui restent au même niveau à l'époque suivante. Leur valeur est proportionnelle à la durée du cycle de formation au taux d'entrée et au taux de redoublement et inversement proportionnelle au taux d'abandons.

Parmi ces facteurs la durée du cycle est le facteur majeur. C'est pourquoi les taux plus grands sont dans le deuxième et troisième cycle du supérieur et dans le primaire, où la durée moyenne est égale (primaire) ou supérieure à 5 ans. Dans le système parfait, sans redoublement et abandon, à flux d'entrée constant, les coefficients de la diagonale seraient égaux aux taux théoriques.

TABLEAU 3

	$P_{11}$	$P_{22}$	$P_{33}$	$P_{44}$	$P_{55}$	$P_{66}$
Taux théorique	0,8	0,75	0,5-0,67	0,75	0,5	0,75-0
taux moyen sur la période	0,811	0,725	0,59	0,65	0,48	0,86

On remarque que dans le premier cycle du secondaire ( $p_{22}$ ) et le deuxième cycle long ( $p_{44}$ ) le taux moyen est très inférieur au taux théorique. Cela signifie que la durée moyenne d'étude est inférieure à la durée théorique et que les abandons avant la fin de la scolarité dans ces cycles sont très nombreux.

Par contre, dans le primaire (où les abandons sont absents à cause de la scolarité obligatoire pour les enfants qui ont l'âge d'être dans l'enseignement élémentaire) le taux moyen est supérieur au taux théorique et ceci à cause des redoublements assez fréquents.

La variance des coefficients de la diagonale (sauf  $p_{11}$ ) est relativement importante par rapport aux autres coefficients de la matrice. Elle est particulièrement forte ( $p_{55}$ ) dans le 1er cycle du supérieur, vu que le comportement des étudiants quant à la prolongation de leurs études change d'une période à l'autre.

Les coefficients de passage d'un niveau à l'autre ( $p_{ij}$ ,  $i \neq j$ )

Les autres coefficients de la matrice de passage qui mesurent les proportions de transition vers un autre état qui n'est pas le même qu'à la période précédente, reflètent les possibilités du choix des élèves et les capacités d'accueil. En plus, la valeur des taux est fonction de la durée moyenne du cycle d'où sortent les élèves et des proportions de ceux d'entre eux qui entrent dans la vie active.

Le taux de passage dans le premier cycle du secondaire a la valeur complémentaire au taux des effectifs qui restent dans le primaire ( $p_{11} + p_{12} = 1$ ).

Les passages dans le cycle long après le premier cycle du secondaire sont plus fréquents que les passages dans le cycle court (les coefficients moyens sont égaux respectivement à 11,1 % et 9,8 % des effectifs).

La variance des coefficients de passage est faible. L'exception présente la variance du coefficient de passage entre les cycles dans le supérieur ( $p_{56}$ ) qui, de même que le coefficient du taux de poursuite dans le premier cycle du supérieur, reflète le comportement des étudiants en liaison avec la conjoncture des diplômés sur le marché du travail.

Les taux de sortie du système avec ou sans diplômes

A part les coefficients concernant l'enseignement supérieur (abandons du 1er cycle et sorties du second et troisième) les autres varient faiblement.

En moyenne 6,6 % d'effectifs du 1er cycle du secondaire quittent annuellement l'école, dont 3,8 % sans aucune qualification (niveau VI). La majeure partie des sortants du second cycle court ont le diplôme de sortie correspondant à la fin de leurs études (niveau V) et éventuellement ils sortent avec le niveau V bis. Par contre, le pourcentage de sortie sans qualification est très faible (16,6 %, 7,2 % et 1 % respectivement). Il y a deux fois plus de sorties diplômées du second cycle long que d'abandons (respectivement 9,1 % et 4,1 % des effectifs du cycle). Le diptyque des sorties du 1er cycle du supérieur est analogique (sorties diplômées représentent 16,6 % et sans diplôme ( $a_{5,IV}$ ) 9,3 %).

#### Dynamique de la structure

Le Graphique 4 et le Tableau A2 attestent une forte stabilité de la structure du système éducatif durant la période.

Une légère tendance au prolongement de la scolarité peut être accusée dans l'enseignement élémentaire et le premier cycle du secondaire. Dans le dernier cas ceci est dû à la diminution de la proportion d'élèves quittant définitivement le système mais aussi à l'arrivée dans le cycle des entrants de naissance des années 1964-65. Pourtant les taux d'abandon restent forts pour ce cycle (comme on l'a vu dans le Tableau 3).

Au-delà du premier cycle du secondaire, les élèves dépassent l'âge de la scolarité obligatoire et ils ont plus de "liberté" pour choisir entre la poursuite des études et le débouché professionnel. Les déformations des coefficients de transition à partir du second cycle font penser aux changements du comportement des jeunes face à la crise de 1975.

Le fléchissement marqué du taux d'élèves qui restent dans le second cycle court entre 1974 et 1975 doit être expliqué par le taux de redoublement accru. La durée d'étude dans ce cycle n'est pas déterminée, car les formations en 2 et en 3 ans y sont regroupées. Entre 1973 et 1977

les formations en 2 ans sont passées de 33 % à 39 % du total des effectifs (dépendants du Ministère de l'Education <sup>(1)</sup>) ce qui correspond à la diminution de la durée moyenne institutionnelle (sans redoublement) qui a pu entraîner l'accélération de la "rotation". Mais ceci ne s'est pas produit.

Les sorties avant la fin des études diminuent sensiblement, surtout au niveau V bis ( $a_{3,Vbis}$ ), ce qui correspond aux observations faites par J. AFFICHARD (1981, p.11). La dynamique des sorties en niveau V, avec des ruptures en 1974 et 1976 ne dégage pas de tendance perceptible.

Le pourcentage des élèves qui restent dans le second cycle long ( $p_{44}$ ) a eu, entre 1973 et 1978, deux tendances d'évolutions opposées. Jusqu'à 1975 une tendance à la baisse (-2,4 % par an) et une remontée (de +2,2 % par an) depuis.

Confronté à ce phénomène, on pense que les élèves abandonnaient ( $a_{4,V}$ ) ou sortaient ( $s_{4,IV}$ ) du système éducatif sans craintes pour l'avenir professionnel. Mais à partir de 1975, sous "la menace du chômage" ils préfèrent rester, en redoublant éventuellement, à l'école. D'ailleurs, les taux d'abandon et surtout de sortie diminuent plus que la proportion de passage vers le supérieur, qui reste très important ; plus de 21 % d'élèves effectuent ce passage annuellement.

Pour l'enseignement supérieur, l'évolution des coefficients témoigne de la prolongation volontaire de la scolarité post-obligatoire. La durée moyenne d'étude (exprimée par  $p_{55}$  et  $p_{66}$ ) augmente et les taux de sortie baissent. Cette tendance ne s'est pas produite en continuité, entre 1975 et 76 le système d'enseignement supérieur a été déstabilisé (la moins bonne précision des simulations pour 1976 le montre également, § 1). Le comportement des élèves et des jeunes en général, ayant le baccalauréat et/ou le diplôme équivalent à bac +2 ans, a changé brutalement.

-----  
(1) Bilan Formation - Emploi 1977 (1981) p.20-21.

Une perturbation dans l'adaptation du modèle avec ses multiples contraintes en découle, sans permettre l'explication simple de la baisse instantanée du taux d'étudiants dans le 1er cycle du supérieur et l'augmentation de 10 % (instantanée aussi) du taux de passage dans le cycle supérieur long.

## 2 - Simulations prospectives

Deux types de prévisions s'appuyant sur la méthode de l'adaptation sont possibles. Le premier utilise les chroniques des paramètres du modèle estimés sur la période antérieure, l'autre permet l'optimisation du flux dans le système selon les critères désignés.

### a) Prévisions prolongeant les tendances passées

Il y a plusieurs manières d'utiliser les résultats d'estimations en dynamique des paramètres du modèle sur la période rétrospective. D'une part, on peut faire les simulations avec la structure du système, décrite par la matrice de transition estimée pour la dernière année qui précède la projection, ou en se référant aux moyennes de chaque coefficient. D'autre part, on peut procéder en introduisant un scénario d'évolution sur la base des tendances observées<sup>(1)</sup>. Cette dernière variante est certainement la plus fructueuse. Pourtant, le champ des observations rétrospectives étant trop court dans notre cas, les tendances ne se sont pas prononcées distinctement. En effet, il a été constaté (voir II.1.b) une certaine prolongation de la scolarité, surtout dans les enseignements post-obligatoires, mais aussi une déformation en 1975-76 comme réaction vive à la crise économique. Ainsi, l'incertitude empêche un scénario cohérent, et on a écarté cette voie d'analyse prévisionnelle.

A titre d'exemple on propose les prévisions faites avec une matrice composée des moyennes des coefficients de passage et de sortie sur la période de 1972-1978. Ces coefficients moyens figurent dans le Tableau A2 de l'Annexe IV. Avec la matrice composée de ces coefficients, l'extrapolation est

---

(1) Si on disposait des séries longues d'évaluation des coefficients on pourrait aussi estimer les fonctions tendanciennes économétriques avec les schémas autorécursifs pour chacun d'entre eux.

faite jusqu'en 1985. On suppose que tous les enfants à l'âge de 6 ans entrent dans le système éducatif et on néglige la mortalité et les migrations des enfants. En plus, les entrées à d'autres niveaux que le primaire sont exclues par définition.

Les estimations des flux de sortie par niveau et des effectifs dans chaque filière sont présentées dans le Tableau 4.

Tableau 4 : Estimations des flux de sortie par niveau de formation et des effectifs du système éducatif : 1980-1985

Hypothèse coefficients moyens prolongés constants/en milliers

Sortie par niveau	1980	1981	1982	1983	1984	1985
I + II	86	89	92	95	97	99
III	69	69	70	70	71	71
IV	131	132	133	133	133	133
V	280	281	281	281	281	281
V bis	142	142	142	141	141	141
VI	128	128	128	128	127	127
Effectifs						
Primaire	4580	4577	4566	4535	4457	4348
1 c. sd	3175	3169	3164	3158	3148	3126
2 c. c.	759	759	759	759	758	756
2 c. l.	1020	1023	1024	1024	1023	1022
1 c. sup.	421	425	427	428	429	429
2+3 c. sup.	675	696	715	732	747	761
Total des entrées <sup>(1)</sup> (variable exogène du modèle)	862	854	832	779	734	727
Total des sorties	836	841	846	848	850	852

(1) Source : Statistiques des enseignements, Ministère de l'Éducation.

Il est évident que le modèle (et les estimations le prouvent) avec la matrice de transition constante ne peut refléter que les modifications de flux intervenues à la suite de variation de la structure démographique. Dans notre cas ceci se manifeste par la baisse des entrées en primaire à partir de 1983 qui provoque en conséquence une diminution des effectifs dans ce cycle. De même, la haute natalité par rapport à la période actuelle dans les années 1960 se traduit sur le nombre d'étudiants et de sorties de l'enseignement supérieur.

Les prévisions de sorties présentées dans le Tableau 4 diffèrent sensiblement de celles faites par le Ministère de l'Education, Tableau 5. Les divergences principales et leurs causes sont les suivantes :

1) Dans l'analyse rétrospective on a cherché à adapter la structure du modèle aux observations de la formation initiale en excluant l'apprentissage <sup>(1)</sup>. Il s'ensuit une divergence des proportions de répartition entre les niveaux de sorties VI, V bis et V des deux prévisions, car celle du Ministère prend en compte l'apprentissage (dans le Tableau 5 qui présente des données d'ensemble). Ceci est un écart de signification du cadre statistique.

2) La plus grande similitude est remarquée au niveau IV (secondaire + supérieur), quant à la tendance de stabilité des flux et aux chiffres de sorties eux-mêmes.

La discordance entre la répartition entre les niveaux III et II+I tient à l'imprécision existant dans le partage des effectifs entre le cycle court et long supérieur et à l'appréciation assez floue des abandons universitaires.

3) Le modèle utilisé ici distingue les niveaux d'éducation de façon agrégée, "étouffant" en quelque sorte les tendances divergentes qui sont justement mises en évidence dans l'analyse du Ministère, qui suit les évolutions des non-réinscriptions (interprétées comme des sorties) par classe dans le secondaire et par discipline universitaire dans le supérieur.

-----  
(1) Selon les données présentées par J. AFFICHARD (1981) p. 10.

Tableau 5 : Flux de sorties "nettes" du système éducatif  
(apprentissage inclus) par niveau de formation  
1980 à 1985

Niveau \ Année	1980	1981	1982	1983	1984	1985
Niveaux I et II	74,8	74,2	75,2	76,1	77,1	78,1
Niveau III	75,5	77,2	79,1	80,8	82,0	82,8
Niveau IV sup.	56,5	59,0	62,2	64,6	67,1	68,8
TOTAL supérieur	206,8	210,4	216,5	221,5	226,2	229,7
Niveau IV sec.	64,6	67,3	68,7	69,7	69,8	70,4
Niveau V	384,7	400,7	402,8	404,2	411,6	417,0
Niveaux Vbis et VI	96,0	88,9	83,1	75,7	69,0	63,3
TOTAL secondaire	545,3	556,9	554,6	549,6	550,4	550,7
TOTAL GENERAL	752,1	767,3	771,1	771,1	776,6	780,4

Source : Formation-emploi, VIII<sup>e</sup> Plan (1980).

4) L'avantage des prévisions faites à partir de notre modèle (équations (1) et (2)) est de prendre en considération à côté des flux des sorties, les effectifs scolaires, ce qui permet de suivre tous les élèves scolarisés jusqu'à la sortie, quel que soit le niveau et quelle que soit l'appartenance administrative de l'établissement d'enseignement <sup>(1)</sup>. Les évolutions des effectifs et de sorties sont étudiées séparément dans les études du Ministère. Les estimations des effectifs dans la période rétrospective concordent ; en revanche, les sorties obtenues par la modélisation avec les coefficients constants et les prévisions de sorties données par le Ministère divergent.

(1) C'est-à-dire autres que les Ministères de l'Education, de l'Agriculture et de la Santé.

b) Problème de la recherche d'une structure optimale du système éducatif

Ce problème peut être posé dans un contexte de planification centralisée de l'éducation. Mais dans ce travail, on l'examine uniquement comme un moyen de tester la pertinence de la méthode d'adaptation pour la résolution des problèmes de l'optimisation à long terme.

On se donne les flux optimaux à des périodes précises et on adapte le système jusqu'à un régime de fonctionnement permettant d'atteindre cet objectif.

Pour les flux dits "optimaux" dans notre analyse, on prend les estimations de la demande de travail adressée aux jeunes entrant dans la vie active faites dans le modèle AGORA (1978). Ces estimations de la demande de travail découlent d'un scénario "optimiste" concourant à améliorer tant l'emploi que les équilibres financiers.

La méthode d'adaptation permet le passage de l'état correspondant à l'année 1978, à un autre le plus proche possible de l'état désiré pour 1984. Comme si on cherchait à minimiser le chômage des jeunes par la restructuration du système éducatif. Le scénario "optimiste" du modèle AGORA prévoit la demande suivante de travail des jeunes :

Tableau 6 : Demande par niveau de formation à l'entrée dans la vie active (en milliers) en 1984

VI	Vbis	V	IV	III	II+I
121	133	244	108	47	46

Tableau 7 : Flux de sorties observés en 1978 (en milliers)

VI	Vbis	V	IV	III	II+I
131	121	285	109	77	71

Le problème dynamique à résoudre consiste à trouver une trajectoire équilibrée qui corresponde à la fonction objectif : minimisation de la variation des variables calculées des flux de sortie autour des valeurs des flux désirées. On cherche les coefficients structurels et les valeurs des variables (les effectifs et les sorties) leur correspondant qui satisferaient aux contraintes du modèle et à la fonction objectif (équations 5-12 de la première partie de ce texte).

La solution d'un tel problème d'optimisation donne :

- 1) la trajectoire d'évolution des variables du modèle (les flux d'effectifs dans ce cas), et
- 2) la trajectoire du changement des coefficients de la matrice.

On impose un passage progressif de la trajectoire qui s'est dessinée dans la période rétrospective à la trajectoire optimale sous l'hypothèse de régularité de transformation de la structure de flux de sortie. Ce passage se fait sur la période de 6 ans (1978-1984).

Tableau 8 : Ecart absolu entre les flux de sorties recherchés et les flux calculés par la méthode d'adaptation (le nombre d'itérations est égal à 5)

Années Périodes	1978 1	1979 2	1980 3	1981 4	1982 5	1983 6	1984 7	Dans le cas des estimations sans adaptation en 1984
Niveaux de sortie								
VI	8	4	3	3	1	2	0	6
Vbis	10	6	2	1	0	12	2	8
V	0	7	5	8	10	31	10	37
IV	16	15	18	10	10	11	9	25
III	0	1	0	0	0	0	0	30
II+I	0	0	0	0	0	0	0	53
Somme des écarts absolus	34	33	28	22	21	56	21	159

Les écarts présentés dans le Tableau 8 montrent le degré d'approximation des estimations à la trajectoire optimale. On voit que l'algorithme de l'adaptation, décrit dans le premier chapitre, est suffisamment efficace pour réagir aux changements radicaux imposés au système. La somme des écarts diminue au fur et à mesure bien que le régime imposé devienne de plus en plus contraignant (cette somme des écarts représente de 2 à 3 % de la somme des variables). Une exception se remarque dans la 6ème période : les perturbations sont dues à la diminution en 1983 de l'arrivée d'enfants à l'âge scolaire dans le système scolaire (voir Tableau 4 p. 28).

La dernière colonne (Tableau 6) présente les écarts des estimations de sortie dans le scénario de la prolongation du fonctionnement du système éducatif selon les tendances observées dans le passé, toujours rapportés aux flux "désirés".

Comment se modifie la structure du système ?

Dans le cadre des flux démographiques d'entrée sans diminution (jusqu'à 1983) et sous la contrainte de diminution totale des flux de sortie, les coefficients de la diagonale ( $t_{ii}$ ) croissent et ceux de sortie et d'abandon diminuent (graphique 5).

Les plus importantes modifications de la répartition des effectifs ont lieu dans l'enseignement supérieur où les restrictions sur les sorties sont les plus fortes. Les changements des coefficients dans le primaire et le secondaire apparaissent être moins grands. Mais il faut se rappeler que la sensibilité de la réaction des variables aux changements des coefficients est plus forte pour ces filières que pour le supérieur, car ils manient beaucoup plus d'effectifs.

./.

GRAPHIQUE 4

L'évolution de la structure du système éducatif  
dans le modèle d'optimisation des flux

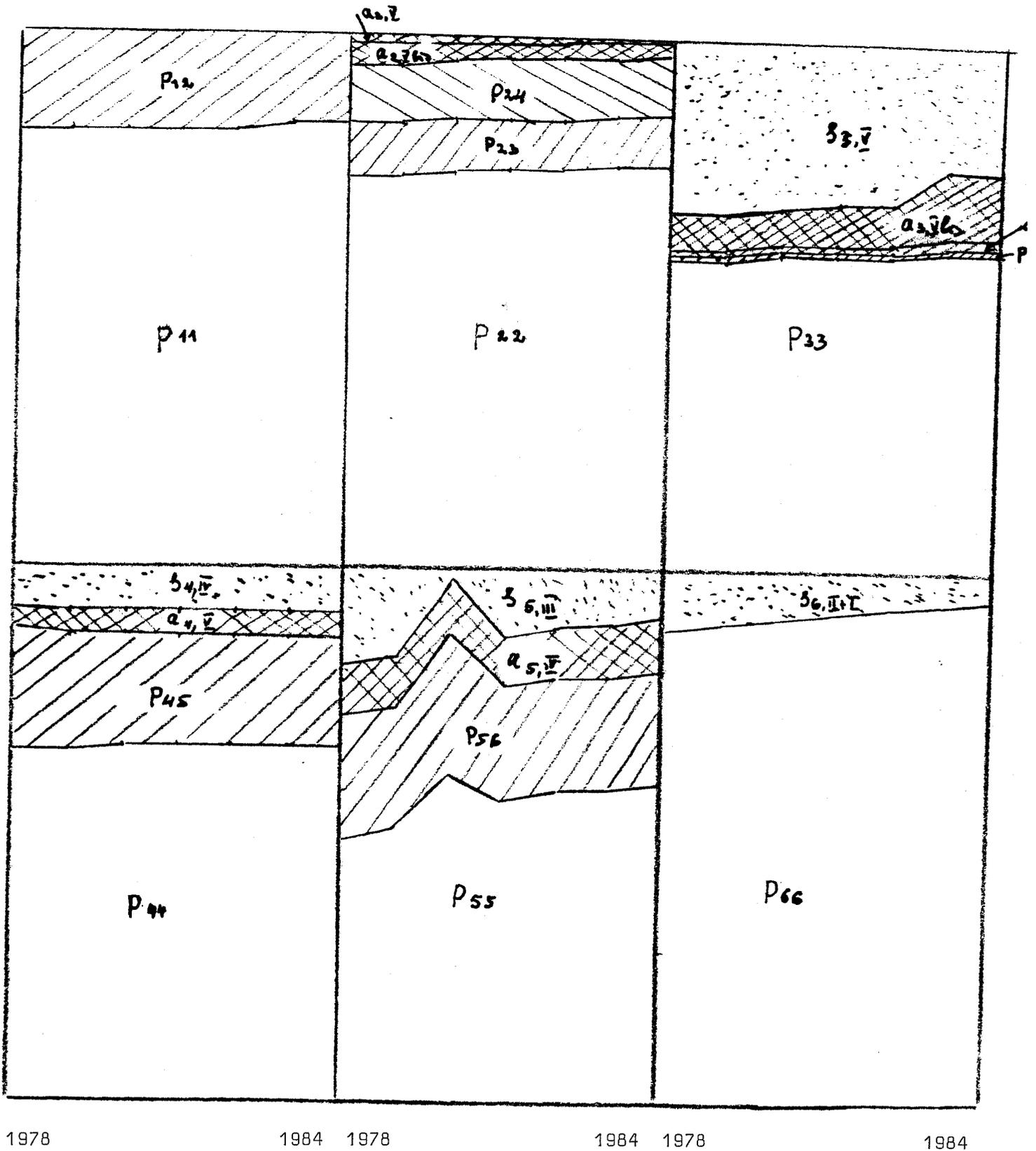


Tableau 9 : Taux de croissance annuel moyen des effectifs et des sorties dans le scénario d'optimisation des flux de sortie par niveau de formation (en %)

	Primaire	1er c.sd	2è c.c.†	2è c.long	1er c.s.	2+3è c.s.
Effectifs	-0,6	0,38	0,85	0,62	2,8	5,9
Sorties	VI	Vbis	V	IV	III	II+I
	+1,3	+1,6	-2,5	-0,15	-7,9	-7,0

Pour assurer en 1984 la structure de flux de sortie demandée (Tableau 6) à la structure d'entrée exogène prévoyant le passage progressif au taux d'évolution porté dans le Tableau 9, les variables (effectifs scolarisés) doivent avoir la trajectoire d'évolution indiquée dans le Tableau 9 et les coefficients de la matrice doivent suivre les trajectoires inscrites sur le Graphique 4.

### Conclusion

La méthode de l'adaptation est utilisée pour estimer les probabilités de transition dans le système éducatif. Il apparaît que cette méthode est assez efficace pour traiter les problèmes au niveau des comportements généraux. Les manières de dynamiser les transitions dans les chaînes markoviennes des autres auteurs (discutés dans la première partie de ce texte) s'appliquent généralement au comportement individuel. Mais les capacités de la méthode de l'adaptation se trouvent réduites dans le traitement du système éducatif où à ce niveau de globalité les possibilités de transition sont limitées aux cursus principaux.

La série de données des flux de sortie à l'heure actuelle est très courte (6 points). La reconstitution de la structure du système à travers l'estimation des proportions des flux entre les niveaux éducatifs et les flux de sortie n'est d'ailleurs possible dans ces conditions qu'avec la méthode de l'adaptation. Pourtant, les changements observés durant la période 1972-1978 ne permettent pas de dégager de scénario d'évolution prospectif. C'est pourquoi on se limite à faire une projection avec une matrice constante, dont les coefficients sont estimés sur la période rétrospective.

L'approche de la modélisation dynamique dans la projection est exposée à la fin à titre d'exemple pour illustrer la possibilité du traitement des effets en retour de l'offre de travail <sup>(1)</sup> sur le fonctionnement du système scolaire (impact sur les taux de transition).

Cette analyse permet d'avancer une réponse "chiffrée" à la question que posent les planificateurs <sup>(2)</sup> aujourd'hui : à quel niveau éducatif et à quelle échelle faut-il organiser les formations permanentes des jeunes pour retarder leur sortie dans la vie active ?

---

(1) Une étude récente de P. PETIT (1981) précise les fonctions d'appel aux jeunes, exprimé par le système productif, en améliorant le traitement modélisé de la relation formation-emploi.

(2) Un des objectifs que se donne le Plan intérimaire de 1982-1983 pour diminuer le chômage.

Annexe I

Nomenclature des niveaux de formation

La version actuelle de cette nomenclature, dite nomenclature interministérielle ou nomenclature des niveaux du Plan, a été fixée par les travaux de la Commission statistique nationale de la formation professionnelle et de la promotion sociale, approuvée par le Groupe permanent du Comité interministériel de la formation continue (décision du 21 mars 1969).

Niveau VI. - Sorties du premier cycle du second degré (6è, 5è, 4è, 3è pratique) et des formations pré-professionnelles en un an (CEP : certificat d'éducation professionnelle, CPPN : classe pré-professionnelle de niveau, CPA : classe préparatoire à l'apprentissage).

Niveau V bis. - Sorties de 3è et des classes du second cycle court avant l'année terminale.

Niveau V. - Sorties de l'année terminale des cycles courts professionnels et abandons de la scolarité du second cycle long avant la classe terminale.

Niveau IV. - Sorties des classes terminales du second cycle long et abandons des scolarisations post-baccalauréat avant d'atteindre le niveau III.

Niveau III. - Sorties avec un diplôme de niveau bac + 2 ans, dans certains cas bac + 3 ans (DUT : diplôme universitaire de technologie ; BTS : brevet de technicien supérieur ; instituteurs ; DEUG : diplôme d'études universitaires générales ; écoles de la Santé, etc.).

Niveau II + I. - Sorties avec un diplôme de second ou troisième cycle universitaire, ou un diplôme de grande école.

Annexe II

Equations d'adaptation du système éducatif modélisé dans AGORA.

Terminologie

- $XEAH^t[n]$  vecteur des effectifs estimé à l'itération n de l'algorithme à t
- $XEAH^t$  vecteur des effectifs observé à t
- $XEAH_i^t$  élément du vecteur  $XEAH^t$   $i = 1, 2, \dots, 6$
- $XEAH_k^t[n]$  } élément du vecteur  $XEAH^t[n]$   $k = 1, 2, \dots, 6$
- $XEAH_j^t[n]$  }
- $TPH^t[n]$  matrice de passage estimée à l'itération n
- $TPH_{ij}^t[n]$  élément de la matrice  $TPH^t[n]$
- $TNH^t[n]$  matrice de sortie du système scolaire estimée à l'itération n
- $TNH_{ij}^t[n]$  élément de la matrice  $TNH^t[n]$
- $ENTH^t$  vecteur des entrées observé
- $ENTH_i^t$  élément du vecteur  $ENTH^t$   $i = 1, 2, \dots, 6$
- $XEOH^t$  vecteur des sorties du système scolaire observé
- $XEOH^t[n]$  vecteur des sorties estimé à l'itération n
- $XEOH_i^t[n]$  élément du vecteur  $XEOH^t[n]$   $i = 1, 2, \dots, 6$
- $b_{ij}^1, b_{ij}^2$  élément de correction du coefficient du modèle
- $c_{ij}^1, c_{ij}^2$  coefficient de normalisation pour garder la somme des coefficients du modèle égale à 1

Le modèle à l'itération n

$$XEAH^t[n] = TPH^t[n-1] XEAH^t[n-1] + ENTH^t$$

$$XEOH^t[n] = TNH^t[n-1] XEAH^t[n-1]$$

./.

Annexe II (suite)

L'adaptation

$$\text{TPH}_{ij}^t[n] = \text{TPH}_{ij}^t[n-1] - b_{ij}^{1t}[n] c_{ij}^{1t}[n]$$

$$b_{ij}^{1t}[n] = \frac{(\text{XEAH}_i^t - \sum_k \text{TPH}_{ik}^t[n-1] \text{XEAH}_k[n] - \text{ENTH}_i)}{\sqrt{n} \text{XEAH}_j^t[n]}$$

$$\text{TNH}_{ij}^{2t}[n] = \text{TNH}_{ij}^{2t}[n-1] - b_{ij}^{2t}[n] c_{ij}^{2t}[n]$$

$$b_{ij}^{2t}[n] = \frac{(\text{XEOH}_i^t - \sum_k \text{TNH}_{ik}^t[n-1] \text{XEAH}_k[n])}{\sqrt{n} \text{XEAH}_j^t[n]}$$

$$b_i^{1t+}[n] = \sum_{j/b_{ij}^{1t} > 0} b_{ij}^{1t}[n]$$

$$b_i^{1t-}[n] = \sum_{j/b_{ij}^{1t} < 0} b_{ij}^{1t}[n]$$

$$b_i^{2t+}[n] = \sum_{j/b_{ij}^{2t} > 0} b_{ij}^{2t}[n]$$

$$b_i^{2t-}[n] = \sum_{j/b_{ij}^{2t} < 0} b_{ij}^{2t}[n]$$

$$c_{ij}^{1t}[n] = \frac{1}{b_i^{1t+}[n] + b_i^{2t+}[n]}$$

si  $b_{ij}^{1t}[n] > 0$

$$c_{ij}^{1t}[n] = \frac{-1}{b_i^{1t-}[n] + b_i^{2t-}[n]}$$

si  $b_{ij}^{1t}[n] < 0$

$$c_{ij}^{2t}[n] = \frac{1}{b_i^{1t+}[n] + b_i^{2t+}[n]}$$

si  $b_{ij}^{2t}[n] > 0$

$$c_{ij}^{2t}[n] = \frac{-1}{b_i^{1t-}[n] + b_i^{2t-}[n]}$$

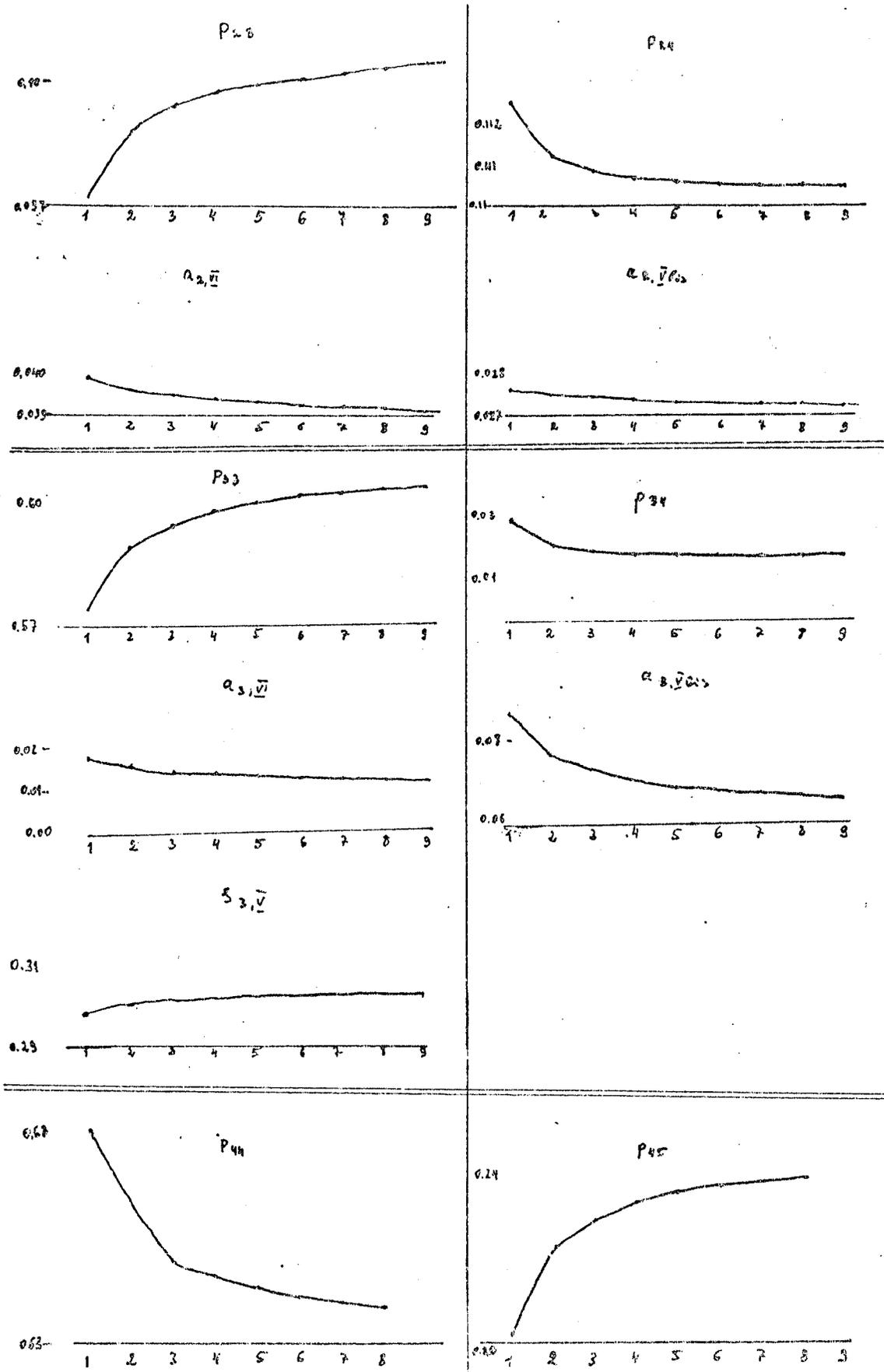
si  $b_{ij}^{2t}[n] < 0$

$$\left. \begin{array}{l} \text{si } b_i^{1t+}[n] + b_i^{2t+}[n] < \epsilon \\ \text{ou } b_i^{1t-}[n] + b_i^{2t-}[n] < \epsilon \end{array} \right\}$$

arrêt de l'algorithme.

Annexe III

Approximation des coefficients  
du modèle (Année 1976)  $n=1, \dots, 9$  itérations



Annexe IV

Tableau A1 - Evolution de la structure du système éducatif  
(coefficients de transition dans le système  $p_{ij}$  et  
vers la sortie dans la vie active  $s_{ij}$ )

Coefficients	Primaire	1er c.sec.	2è c.court	2è c.long	1er c.sup.	2 et 3è c.s
<u>1973</u>	0.801978	0	0	0	0	0
$p_{ij}$	0.198022	0.709035	0	0	0	0
	0	0.102951	0.574562	0	0	0
	0	0.117554	0.021296	0.658942	0	0
	0	0	0	0.213426	0.467665	0
	0	0	0	0	0.258485	0.856843
	0	0.0416088	0.0141406	0	0	0
$s_{ij}$	0	0.0288510	0.0732145	0	0	0
	0	0	0.316786	0.0358304	0	0
	0	0	0	0.0918009	0.103662	0
	0	0	0	0	0.170187	0
	0	0	0	0	0	0.143157
	0	0	0	0	0	0
<u>1974</u>	0.812605	0	0	0	0	0
$p_{ij}$	0.187395	0.726104	0	0	0	0
	0	0.098031	0.583189	0	0	0
	0	0.109207	0.009273	0.640842	0	0
	0	0	0	0.213546	0.469249	0
	0	0	0	0	0.254515	0.856465
	0	0.038949	0.010365	0	0	0
$s_{ij}$	0	0.027708	0.075344	0	0	0
	0	0	0.321827	0.047875	0	0
	0	0	0	0.097736	0.105245	0
	0	0	0	0	0.170991	0
	0	0	0	0	0	0.143535
	0	0	0	0	0	0
<u>1975</u>	0.809528	0	0	0	0	0
$p_{ij}$	0.190472	0.719445	0	0	0	0
	0	0.100171	0.577236	0	0	0
	0	0.111414	0.009117	0.628491	0	0
	0	0	0	0.222673	0.469518	0
	0	0	0	0	0.258833	0.861585
	0	0.039847	0.0103113	0	0	0
$s_{ij}$	0	0.029123	0.088798	0	0	0
	0	0	0.314534	0.045957	0	0
	0	0	0	0.102877	0.105305	0
	0	0	0	0	0.166344	0
	0	0	0	0	0	0.138415
	0	0	0	0	0	0

---

<u>1976</u>	0.807859	0	0	0	0	0
	0.192141	0.72273	0	0	0	0
$P_{ij}$	0	0.100111	0.605365	0	0	0
	0	0.11131	0.009529	0.636934	0	0
	0	0	0	0.228585	0.440203	0
	0	0	0	0	0.331183	0.858734
	0	0.038208	0.0089298	0	0	0
	0	0.027640	0.0735548	0	0	0
$s_{ij}$	0	0	0.302619	0.037642	0	0
	0	0	0	0.096837	0.074836	0
	0	0	0	0	0.153778	0
	0	0	0	0	0	0.141265
<u>1977</u>	0.809061	0	0	0	0	0
	0.190939	0.73194	0	0	0	0
$P_{ij}$	0	0.096328	0.599713	0	0	0
	0	0.109714	0.010503	0.655916	0	0
	0	0	0	0.222719	0.498854	0
	0	0	0	0	0.252465	0.878799
	0	0.035922	0.00856	0	0	0
	0	0.026095	0.06738	0	0	0
$s_{ij}$	0	0	0.312844	0.03826	0	0
	0	0	0	0.08411	0.08368	0
	0	0	0	0	0.164999	0
	0	0	0	0	0	0.121201
<u>1978</u>	0.824629	0	0	0	0	0
	0.175371	0.738823	0	0	0	0
$P_{ij}$	0	0.094029	0.603887	0	0	0
	0	0.107712	0.010808	0.671203	0	0
	0	0	0	0.215709	0.518983	0
	0	0	0	0	0.222042	0.886957
	0	0.034518	0.007032	0	0	0
	0	0.024919	0.055758	0	0	0
$s_{ij}$	0	0	0.321513	0.0391	0	0
	0	0	0	0.074986	0.0868296	0
	0	0	0	0	0.172145	0
	0	0	0	0	0	0.113043

Tableau A2 - Taux moyen, variance et écart-type des coefficients de la matrice de transition

	Coefficients de "rotation" dans le système					
	P <sub>11</sub>	P <sub>22</sub>	P <sub>33</sub>	P <sub>44</sub>	P <sub>55</sub>	P <sub>66</sub>
Taux moyen	0.811	0.725	0.59	0.649	0.477	0.866
Variance	0.00004	0.00009	0.00016	0.00021	0.00064	0.00014
Ecart-type	0.00736	0.0104	0.0141	0.016	0.0277	0.01314

	Coefficients de passage vers les niveaux suivants					
	P <sub>12</sub>	P <sub>23</sub>	P <sub>24</sub>	P <sub>34</sub>	P <sub>45</sub>	P <sub>56</sub>
Taux moyen	0.189	0.098	0.111	0.0115	0.219	0.263
Variance	0.000049	0.0000086	0.0000088	0.0000186	0.0000322	0.001114
Ecart-type	0.0077	0.0032	0.00325	0.0047	0.0062	0.036

	Coefficients d'abandons du système scolaire ; n = VI, V, Vbis, IV, III, II+I = sortie a <sub>i,n</sub> i = 1, ... 6 = niveaux de formation					
	a <sub>2,VI</sub>	a <sub>2,Vbis</sub>	a <sub>3,VI</sub>	a <sub>3,Vbis</sub>	a <sub>4,V</sub>	a <sub>5,IV</sub>
Taux moyen	0.0381	0.0275	0.0097	0.072	0.041	0.093
Variance	0.0000063	0.0000023	0.000005	0.000097	0.0000201	0.00014
Ecart-type	0.00275	0.00164	0.00246	0.0108	0.0049	0.013

	Coefficients de sorties diplômées; s <sub>i,n</sub>			
	s <sub>3,V</sub>	s <sub>4,IV</sub>	s <sub>5,III</sub>	s <sub>6,II+I</sub>
Taux moyen	0.315	0.0915	0.166	0.133
Variance	0.0000397	0.0000889	0.000037	0.00014
Ecart-type	0.0069	0.01033	0.0066	0.0129

Bibliographie

AGORA (1978)

Le modèle AGORA : contribution à l'analyse des relations entre productions marchandes et productions non marchandes. Rapport ronéoté CEPREMAP, 384 p.

AFFICHARD J. (1981)

Quels emplois après l'école : la valeur des titres scolaires depuis 1973, Economie et Statistique n° 134

ANDERSON T.W., GOODMAN L.A. (1957)

Statistical Inference About Markov Chains,  
The Annals of Mathematical Statistics, 28

BENARD J. (1967)

General optimization model for the Economy and Education,  
OCDE, Mathematical models in educations planning, Paris

BILAN FORMATION-EMPLOI 1973 (1978)

BILAN FORMATION-EMPLOI 1977 (1981)

Les Collections de l'INSEE, D.59 et D.78

BOUDON R. (1973)

Mathematical structures of social mobility,  
San Francisco, Jossey Bass

BOX G.E.P. et JENKINS G.M. (1962)

Some statistical aspects of adaptative optimization and control,  
Journal of the Royal Stat. Sec., ser. B, Vol. 24 n° 2

BROWN R.G., MEYER R.F. (1961)

The fundamental theorem of exponential smoothing  
Oper. Res. vol. 9 n. 5

BROWN R.G. (1963)

Smoothing forecasting and prediction of discrete time series, N.y.

CHOW W.M. (1965)

Adaptative control of the exponential smoothing constant,  
J. of Indust. Engineering, vol. 16, N 5

CONLISK J. (1976)

Interactive Markov Chains,  
Journ. of Math. Sociology, Vol. 4

- CONLISK J. (1978)  
A stability theorem for an interactive Markov chain,  
J. of Math. Sociology, Vol. 6
- GINSBERG R.B. (1971)  
Semi-Markov processes and mobility,  
Journ. of Math. Sociology, Vol. 1
- GINSBERG R.B. (1972)  
Incorporating causal structure and exogenous information with  
probabilistic models : with special reference to choice, gravity,  
migration, and Markov chains,  
Jour. of Math. Sociology, Vol. 2
- L'ENTREE DES JEUNES DANS LA VIE ACTIVE (1981)  
Econom. et Statistique, Numéro Spécial, N 138
- ERMILOV A.P. (1979)  
Analyse de l'exactitude des prévisions macroéconomiques,  
dans : Modélisation économétrique (1979) Novossibirsk 1979 (en russe)
- IOUDIN D.B. (1979)  
Les problèmes et méthodes de la programmation stochastique, Moscou  
(en russe)
- KUSHNER H.J., CLARK D.S. (1978)  
Stochastic Approximation Method for Constrained and Unconstrained  
Systems, Springer, Verlag
- LEE T.C., JUDGE G.G., ZELLNER A. (1970)  
Estimating the parameters of the Markov probability model from  
aggregate time series data,  
North Holland Publishing Company, Ams. London
- LOUKACHIN I.P. (1979)  
Les méthodes d'adaptation pour les prévisions à court terme,  
Moscou, "Statistica" (en russe)
- MACRAE E.C. (1977)  
Estimation of Time-Varying Markov Processes with Aggregate Data,  
Econometrica, 45, N 1
- MADANSKY A. (1959)  
Least Squares Estimation in Finite Markov Processes,  
Psychometrica, 24

- MC GUIRE T.W. (1969)  
More on Least Squares Estimation of the Transition Matrix in a Stationary First-Order Markov Process from Sample Proportion Data, Psychometrica, 34
- PEAUCELLE I., PETIT P. (1978)  
Simulation à long terme de la relation formation-emploi, Colloque franco-soviétique "Méthodologie de la planification à moyen et long terme", Paris, Septembre
- PEAUCELLE I. (1979)  
La méthode d'adaptation dans les modèles économiques. Analyse de quelques approches,  
Note Ronéotée Université de Paris I (Rapport CORDES)
- PETIT P. (1981)  
Insertion des jeunes et évolutions sectorielles de l'emploi : une approche macroéconomique d'un marché du travail segmenté
- PONTRYAGIN L.S., BOLTYANSKII V.G., GAMKRELIDZE R.V., MISHCHENKO F.F. (1962)  
The mathematical theory of optimal processes (translation from the russian), Interscience publishers
- ROSENQUIST G. (1980)  
Combining micro and macro data in estimation of the transition probabilities of a finite Markov chain,  
Contributed paper at the Fourth World Congress of the Econometric Society, Aix-en-Provence
- SCHINNAR A.P., STEWMAN S. (1978)  
A class of Markov models of social mobility with duration memory patterns, J. of Math. Sociology, vol. 6
- SPILERMAN S. (1972)  
The analysis of mobility processes by the introduction of independent variables into a Markov Chain, American Sociological Review, 37.277-294
- STONE R. (1971)  
The fundamental matrix of the active sequence,  
Fifth International Conference on Input-output Techniques, Geneve, January

STONE R. (1976)

Le système de statistiques sociales et démographiques (ONU)  
Ec. et St. N 75, Février

TACHIBANAKI T. (1980)

Education, occupation and earnings : A recursive approach for France,  
European Economic Review, Vol. 13, N 1

TELSER Lester G. (1963)

Least-Squares Estimates of Transition Probabilities  
Christ and others in Measurement in Economics,  
Stanford Univ. Press

THEIL H., WAGE S. (1964)

Some observations on adaptative forecasting,  
Management Science, Vol. 10, N 2

THEIL H. (1970)

Programmation économique appliquée (en russe)

TOVSTAIKHA T.I. (1974)

Analyse des algorithmes discrets quasi-optimaux de l'identification  
(en russe), Automatique et Télémechanique, N 4, Moscou

TRIGG D.W., LEACH A.G. (1967)

Exponential smoothing with an adaptative response rate,  
Oper. Res. Quart. Vol. 18, N 1

TURNOVSKY S.J.

Macroeconomic analysis and stabilization policy