

# Contribution à la modélisation de l'insertion professionnelle<sup>1</sup>

par Gabriel Tahar\*

**À partir de données portant sur les trajectoires individuelles d'une population de jeunes de bas niveau scolaire, l'auteur propose une analyse qui permet de progresser dans la modélisation statistique de l'insertion professionnelle. Les approches « macro » ou « micro » développées montrent la différence entre définition de niveau statistique et définition au plan individuel.**

C'est en longitudinal qu'il faut analyser l'entrée des jeunes dans la vie active puisqu'elle se caractérise par des trajectoires individuelles diversifiées d'insertion professionnelle. Mais l'on se heurte à la complexité croissante des cheminements de transition entre l'école et ce que l'on appelle couramment « insertion » : on oscille entre description fine peu opératoire et théorie trop caricaturale.

Nous abordons ici deux problèmes méthodologiques majeurs : comment définir « objectivement » l'insertion professionnelle d'une part, comment décrire les trajectoires et en mesurer les différences d'autre part ? Pour comparer les trajectoires individuelles, on a besoin d'une mesure des écarts et l'on a donc introduit une « distance » qui permettra de se situer par rapport à des trajectoires-types mais surtout servira d'indicateur synthétique susceptible d'être ajusté économétriquement. Cette distance se déduira directement des probabilités de transition (éléments

de la matrice de transition entre les états du marché<sup>2</sup> à deux dates  $t$  et  $t'$ ) entre les quatre situations vis-à-vis du marché du travail que nous avons privilégiées :

- 1) l'emploi à durée indéterminée ;
- 2) l'emploi à durée déterminée ;
- 3) les périodes de stage ou de formation ;
- 4) le chômage ou l'inactivité.

Les variables explicatives mises en rapport avec les caractéristiques des trajectoires professionnelles des jeunes sont communément le sexe, l'âge de sortie du système éducatif, le niveau de formation, le nombre d'emplois occupés, etc. Nous y ajouterons des variables de zone ainsi qu'un indicateur de l'adéquation formation-emploi, trop vite évacuée aujourd'hui sous prétexte qu'elle n'est ni automatique ni directe. Nos données concernent une population de jeunes dits « de bas niveau scolaire » qui a l'avantage d'être relativement homogène<sup>3</sup>. Nous proposons ici une analyse en trois étapes de l'insertion professionnelle :

- au niveau agrégé, les trajectoires professionnelles seront considérées comme des processus ponctuels (p.19) ;

\* **Gabriel Tahar**, économiste, directeur de Recherches au CNRS ; détaché en 1998 à l'INRA-Toulouse, équipe ETIQ (Emploi, Territoire, Innovation, Qualité), chercheur associé au LIRHE-CNRS, Université de Toulouse I, CIA-Céreq, travaille actuellement sur le chômage des jeunes et l'analyse des marchés du travail ruraux. Dernière publication : « Démographie d'établissements et processus d'appariement sur les marchés du travail locaux », in *Mélanges en l'honneur de Jean Vincens*, Presses de l'Université des Sciences sociales, Toulouse, 1998 (à paraître).

<sup>1</sup> Ce travail fait suite à une recherche effectuée en collaboration avec Gilles Alloire et financée par le Commissariat général du plan, la Délégation à l'insertion des jeunes ainsi que l'Observatoire régional des emplois et des formations Midi-Pyrénées. Cette recherche entrait dans un programme coordonné par le Céreq (voir Demazière et Dubar, 1994). Je remercie les rapporteurs anonymes de la revue pour la pertinence de leurs remarques.

<sup>2</sup> Un état étant représenté par le vecteur-colonne des fréquences empiriques d'occurrence des quatre situations possibles à un moment donné.

<sup>3</sup> Trop souvent, les études traitent ensemble tous les niveaux de qualification et ne décrivent les effets de ces derniers que sous forme de variables indicatrices, voire simplement d'un nombre d'années d'études.

- au niveau individuel, on définira une date d'insertion correspondant au moment où ne surviennent plus de longues périodes de chômage, les contrats de travail pouvant être à durée déterminée (p.21) ;

- pour traiter de l'hétérogénéité entre individus, on s'intéressera à une mesure des écarts entre les trajectoires individuelles d'insertion en spécifiant une distance fondée sur les probabilités de passage observées entre situations d'emplois stable ou précaire, de formation et de chômage (p.24).

Au préalable, nous aurons posé les termes du débat théorique (§1) et décrit les trajectoires moyennes suivies par les jeunes de notre échantillon (p.17).

## **RECHERCHE D'UNE FORME STRUCTURELLE OU MODÉLISATION EMPIRIQUE ?**

Nous avons à examiner un moment particulier de la vie active, l'insertion, que l'on peut définir comme « le passage de la fin au début » ! Sous la boutade se cache un phénomène fort complexe reliant fin de période scolaire et début de carrière professionnelle. Si un accord peut se faire sans trop de mal sur la date où se termine la formation initiale (encore que l'on observe parfois des prises d'emploi définitif avant cette date, chez les fonctionnaires par exemple), les chercheurs, de l'ultra-empiriste au microéconomiste, sont pour une fois d'accord sur l'impossibilité radicale d'une définition absolue de l'insertion. Il faut néanmoins disposer de repères pour avancer dans l'analyse quantitative et nous proposerons ici une définition de l'insertion se faisant plutôt « en creux », en privilégiant des critères négatifs comme l'absence de chômage récurrent plutôt qu'un critère positif comme le contrat à durée indéterminée (CDI), souvent hors d'atteinte et parfois plus précaire qu'un « bon » CDD (contrat à durée déterminée dans une entreprise solide, prélude à un CDI ultérieur) ! La logique d'une trajectoire n'est pas seulement la stabilisation dans un emploi pérenne.

Analyser l'insertion revient alors en partie à étudier la durée d'un certain parcours initial que l'on qualifie de phase de transition (Allaire et Tahar, 1996). Cette période de transition se justifie « sociologiquement » en passant de la structure des récits personnels à la structure temporelle (Allaire, Beslay, Chartier et Tahar, 1995) qui fait apparaître plus exactement quatre moments : latence et transition proprement dite, correspondant à ce que l'on appelle la phase de transition, puis stabilisation et insertion.

On sait que la modélisation empirique se donne pour but d'ordonner, de simplifier un ensemble complexe de faits afin de mettre en évidence certaines régularités (énoncés descriptifs) à expliquer ultérieurement par la théorie (énoncés explicatifs). La « boîte à outils » de l'économiste nous fournit essentiellement la théorie de la recherche (*job search*), celle de l'appariement (*matching*), celle du signal (*signalling*) et enfin celle de la segmentation (*segmented labor market*).

- La théorie de la recherche utilise le salaire comme paramètre de décision, l'individu décidant ou non de poursuivre sa recherche d'emploi en fonction de la balance entre salaire espéré, salaire offert et coût d'une période supplémentaire de recherche. Mais la concurrence entre jeunes est plus fondée sur un classement que sur une négociation de salaire. Il y a une correspondance « institutionnalisée » entre la rémunération et le poste de travail d'un débutant. Cette approche peut néanmoins s'avérer utile à l'intérieur d'une des phases évoquées plus haut.

- La théorie de l'appariement aborde l'adéquation de l'individu au poste de travail par l'analyse de la confrontation entre emplois vacants et demandeurs d'emplois. Un bon appariement se traduit par un salaire intéressant pour les deux parties et, en conséquence, par une relation d'emploi de longue durée. Cette approche est peu adaptée à l'étude de l'offre de travail des jeunes qui n'est pas un segment spécifique du marché du travail mais se trouve en concurrence avec les travailleurs plus âgés. De plus, on observe souvent dans les parcours initiaux des jeunes ce que l'on appelle des emplois d'attente « faute de mieux » où la durée signifie au contraire l'absence d'opportunité adaptée à la qualification acquise.

- La théorie du signal fournit plus un cadre d'analyse qu'un modèle structurel en rendant compte de l'influence d'une série de caractéristiques individuelles ou d'événements passés sur le futur <sup>4</sup>.

- La prise d'un emploi précaire (ou de plusieurs successivement) par un jeune va le situer sur ce que les théoriciens de la segmentation appellent le marché secondaire : c'est typiquement un choix pouvant avoir une influence sur le futur. La question posée concerne les possibilités de passage plus ou moins facile du marché secondaire au marché primaire où les emplois sont plus attractifs et s'inscrivent dans une carrière. Mais à la différence de l'adulte qui craint à juste titre un enfermement dans le segment secondaire,

<sup>4</sup> Nous utiliserons dans nos estimations économétriques de telles variables explicatives.

l'emploi précaire du jeune peut tout à fait s'inscrire dans une séquence d'acquisition d'une première expérience professionnelle sans forcément donner un signal négatif aux futurs employeurs. La segmentation est alors plus pour eux temporelle que « spatiale ».

La difficulté dans l'utilisation d'un modèle théorique d'offre de travail, quel qu'il soit, n'est pas tant l'hypothèse de rationalité (on peut dire d'une manière très extensive qu'un acte est réputé rationnel si de bonnes raisons le justifient) que le problème du choix d'une fonction d'utilité de l'individu ne se limitant pas au salaire et à la durée de l'emploi. On observe dans les comportements, non seulement une forte hétérogénéité individuelle, ce que l'économètre sait traiter, mais également une hétérogénéité intertemporelle (les préférences évoluant dans le temps en fonction des situations vécues) devant laquelle on est plus démuni.

Le problème de l'insertion est d'examiner des discontinuités et pas seulement de suivre un modèle probabiliste (arrivée des offres et/ou durée de l'emploi) s'appuyant sur un certain déterminisme des trajectoires et des décisions successives fondées sur l'information reçue. On repère néanmoins incontestablement des « signaux » plus ou moins forts, déterminants d'évolutions futures, dont il s'agit d'évaluer l'importance. Avant de s'interroger sur la signification de certaines trajectoires, il est bon de réfléchir sur les cadres structurels dans lesquels se situent ces parcours car il est bien évident, par exemple, que la mobilité dépend du contexte en général et de la demande en particulier ! On en arrive au point à établir avec la théorie du fonctionnement du marché du travail, prenant en compte ce contexte : montée du chômage et transformation profonde des structures d'emploi concernant les jeunes en particulier (emplois dits « aidés » ou emplois du « dispositif »). On sent bien que l'analyse économique de l'insertion ne peut ignorer totalement une certaine sociologie des comportements mettant l'accent sur l'histoire individuelle et l'irréversibilité de certains choix.

Tout cela nous incite fortement à considérer l'insertion professionnelle comme un processus analysant sous contraintes le cheminement dans le temps des individus entre les principaux états qui caractérisent à chaque instant leur situation vis-à-vis du marché du travail.

Le processus d'insertion est situé dans son double rapport à l'emploi et au chômage (Fourcade, in Tanguy, 1986). Pour son rapport à l'emploi, il s'agit de la succession des différents emplois occupés. En ce qui concerne le rapport au chômage, il s'agit alors

des changements d'états vis-à-vis du marché du travail : actif-inactif, actif occupé-chômeur, stagiaire-chômeur, emploi « aidé »-emploi de droit commun, etc.

Le marché du travail est un lieu fictif où se rencontrent une structure d'emploi et des modalités de fonctionnement. Si on le représente formellement par un processus ponctuel, cela correspond à une matrice de transition  $M$  (décrivant le *fonctionnement* du marché sur la période) et un vecteur d'état  $V$  (représentant la *structure d'emploi* à un moment donné). On verra plus loin en détail que cette présentation permet de mieux comprendre les sources de déséquilibre sur le marché du travail en distinguant formellement les flux et les stocks.

## DESCRIPTION DES TRAJECTOIRES MOYENNES

La base de données initiale résulte de l'enquête de cheminement professionnel réalisée par le Céreq dans le cadre de l'Oneva<sup>5</sup> en décembre 1989 auprès d'un échantillon représentatif des jeunes sortis aux niveaux VI, Vbis et V<sup>6</sup> en juin 1986 (300 000 sortants, 10 426 questionnaires exploitables). Elle permet de retracer de manière détaillée mois après mois les parcours de formation et d'activité au cours des 3 années et demie consécutives à la sortie de l'école (Pottier, 1992). S'y ajoute la réinterrogation du même échantillon effectuée sur un nombre restreint de régions en 1993 afin de retracer les parcours sur une période deux fois plus longue et capter ainsi une plus forte proportion d'insertion, quelle qu'en soit sa définition.

La fusion des deux enquêtes ne va pas sans problème. La représentation des trajectoires probabilisées<sup>7</sup> montre une cassure très nette au 55<sup>e</sup> mois pour les CDI (voir le graphique 3, annexe 1), l'évolution des taux de chômage restant à peu près identique. Cette cassure est perceptible sur le graphique 1 précédent pour l'insertion de type 1. Cela ne nuit pas cependant à l'analyse des périodes de transition puisque le graphique résultant de la seule prise en compte des

<sup>5</sup> Observatoire national des entrées dans la vie active.

<sup>6</sup> Niveau VI : sorties de 6<sup>e</sup>, 5<sup>e</sup> et 4<sup>e</sup> ainsi que des formations préprofessionnelles en un an ; niveau Vbis : sorties de 3<sup>e</sup> et des classes du second cycle avant l'année terminale ; niveau V : sorties de l'année terminale des cycles courts professionnels et abandons de la scolarité du cycle long avant la terminale.

<sup>7</sup> C'est-à-dire plus simplement les fréquences relatives mois après mois de l'emploi stable, de l'emploi précaire, de la formation ou du chômage (les situations de chômage étant réunies quelle que soit leur durée en une situation 4).

## Présentation des données utilisées

Pour l'enquête faisant suite à celle de l'ONEVA, six régions aux configurations de systèmes productif et d'emploi contrastées ont été sélectionnées : Alsace, Languedoc-Roussillon, Nord-Pas-de-Calais, Pays de Loire, Midi-Pyrénées et Provence-Alpes-Côte d'Azur. Cette nouvelle enquête a été effectuée en juin 1993 sous la responsabilité du Céreq pour les quatre premières régions, en septembre 1993 sous celle de l'INRA-ESR pour les deux dernières. Réalisée par téléphone, elle ne fournit pas une description des périodes successives aussi détaillée que la précédente mais permet le prolongement du calendrier des activités de 1565 jeunes, même si certains changements d'employeurs ont pu être omis. L'information sur le lieu de travail n'existe, contrairement à l'enquête initiale, que pour le dernier emploi et ce uniquement pour les régions Midi-Pyrénées et PACA. C'est la raison pour laquelle on s'appuiera par la suite sur le lieu du premier emploi, quel qu'il ait pu être, pour éviter au maximum les données manquantes, sachant que le lieu où se situe le premier poste de travail possède un relief particulier qui doit jouer dans les mécanismes d'insertion, telle est du moins l'hypothèse que nous faisons.

C'est ainsi que chaque mois le jeune se trouve dans une des situations répertoriées dans le tableau 1 ci-

Tableau 1  
Description des situations

Description des situations	Code
Emploi stable	1
Emploi précaire	2
Formation ou reprise d'études	3
Chômage ou inactivité de moins de 3 mois	5
Chômage ou inactivité compris entre 3 et 6 mois	6
Chômage ou inactivité de plus longue durée	7
Service national	8

dessus. On a choisi d'éclater les situations de chômage en 3, selon leur durée et, en revanche, de ne pas les distinguer de l'inactivité, distinction très artificielle chez un jeune.

## Remarques

- Les emplois stables regroupent les personnes à leur compte, les contrats à durée indéterminée, les fonctionnaires et agents contractuels de l'Etat. Les emplois précaires regroupent les intérimaires, les aides familiaux, les apprentis et les différents contrats relevant des dispositifs pour l'emploi.

- Une petite difficulté provient du fait que les dates de début et de fin de certaines situations se chevauchent dans le calendrier. Nous avons été amené à effectuer les équivalences suivantes :

- 1,1 ou 1,8 ou 1,3 ou 1,5 ou 1,6 ou 1,7 ou 1,1,8 ou 1,1,8 ou 1,3,8 = 1
- 3,8 = 3
- 3,5 = 5
- 3,6 = 6
- 3,7 = 7

- Dans la suite, le service national entre 2 dates  $t$  et  $t'$  a été « neutralisé » en considérant comme données manquantes les situations entre  $t$  et  $t'$  et en comptabilisant uniquement la transition  $(t-1) \rightarrow (t'+1)$ , et ce à la date  $t'$ .

Finalement le fichier utilisé ici est le produit de ces deux sources. Au cours de la fusion, l'information issue de la base initiale recueillie en 1989 a été privilégiée en cas de contradiction pour une double raison : l'enquête y décrit des événements plus récents et est questionnée en face en face et non par téléphone, ce qui donne des gages de meilleure fiabilité.

non-insérés (voir le graphique 5) ne présente plus de cassure. Le graphique 4 comparé au précédent suggère à ce propos que les enquêtés ont pu traduire en CDI lors de la réinterrogation de 1993 certains des emplois qu'ils percevaient à l'époque comme des CDD.

L'établissement de trajectoires moyennes ne pose pas de problème particulier mais nous allons passer par une étape intermédiaire nécessaire aux calculs de distance ultérieurs. On définit  $S_{it}$  comme la proportion de situations  $i$  par rapport à l'ensemble des situa-

tions au mois  $t$  ( $i$  prenant comme précédemment les valeurs de 1 à 4 selon que l'on se trouve en emploi stable, emploi précaire, formation ou chômage).

La connaissance pour chaque mois de  $q_t = \text{argmax}_i (s_{it})$  permet de définir la trajectoire moyenne comme la suite des situations les plus probables à chaque instant  $t$ . Deux trajectoires moyennes peuvent être ainsi construites selon que l'on arrête ou non les observations individuelles aux dates d'insertion (voir annexe 2). Ces deux trajectoires moyennes s'avèrent bien sûr très différentes :

- après la sortie de la formation initiale (14ème mois), la première, limitée à la phase de transition, présente trois périodes de chômage ou d'inactivité (de respectivement 7, 4 et 9 mois) entre lesquelles s'intercalent trois emplois précaires ou sans doute des emplois aidés dans le cadre du dispositif « jeune » (de respectivement 10, 6 et 1 mois). Elle se termine par du chômage ou de l'inactivité (pendant les 47 mois restant) qui apparaît comme l'état le plus probable (avant l'emploi précaire ou le stage) quand on persiste à n'être pas inséré ;

- après les mêmes 13 mois de formation initiale, la seconde, calculée sur 97 mois quel que soit l'individu, présente également du chômage mais limité à une seule période de 5 mois puis un long emploi précaire (37 mois) et enfin un emploi stable pendant les 42 mois restant.

La trajectoire la plus probable est somme toute assez pauvre et livre peu d'information. De plus, les modifications de situation qu'elle donne sont artificielles : elles ne correspondent pas à des changements réels mais seulement à un changement de rang dans les probabilités d'occurrence des situations. Il est nécessaire d'entrer un peu plus dans le détail des mécanismes sous-jacents.

## ANALYSE DES TRANSITIONS ENTRE ÉTATS

La succession des situations vécues par un individu peut être représentée au plan statistique par un processus ponctuel<sup>8</sup> à espace d'états discrets non homogène par rapport au temps. Le phénomène de persistance du chômage, la propension à subir un épisode de chômage dépendant des chômeurs vécus jusqu'à trois années antérieures (Allaire, Cahuzac et Tahar,

1995), montre le caractère non markovien, du moins à des ordres raisonnables, du processus d'insertion. Il y a donc « dépendance du chemin »<sup>9</sup>.

On part de l'observation des réalisations successives  $s_{it}$  du processus aux différentes dates  $t$  ( $i=1,4$   $t=14,96$ ) représenté par le vecteur-colonne  $V_t$  des probabilités d'appartenance  $s_i$  aux 4 états CDI (1), CDD (2), formation (3) ou chômage (4) à l'instant  $t$ .

On définit alors  $M_t$ , matrice de transition d'ordre 4 composée des éléments  $p_{ij}$  qui sont les probabilités de passer de l'état  $j$  (en  $t$ ) à l'état  $i$  (en  $t+1$ ). La ligne  $i$  est composée des éléments  $p_{i\cdot}$ , les 4 probabilités d'arriver à l'état  $i$  venant d'un des 4 états  $j$  possibles. La colonne  $j$  comprend les 4 éléments  $p_{\cdot j}$ , les 4 probabilités d'arriver à l'un des 4 états  $i$  venant de l'état  $j$ , ce qui implique que la somme des éléments de chaque colonne est égale à 1.

$$M_t = \begin{pmatrix} p_{11t} & p_{12t} & p_{13t} & p_{14t} \\ p_{21t} & p_{22t} & p_{23t} & p_{24t} \\ p_{31t} & p_{32t} & p_{33t} & p_{34t} \\ p_{41t} & p_{42t} & p_{43t} & p_{44t} \end{pmatrix}$$

On passe de l'état  $t$  à l'état  $t+1$  du processus en multipliant  $V_t$  par  $M_t$ . On a donc :

$V_{t+1} = M_t \cdot V_t$  et donc  $V_{t+2} = M_{t+1} \cdot M_t \cdot V_t$  et ainsi de suite d'où

$V_{t'} = M_{t'} \cdot V_t$  avec  $M_{t'} = M_{t'} \cdot M_{t'-1} \dots M_{t+1} \cdot M_t$

$M_{t'}$  dépend de  $t$  et de  $t'$  et non de  $(t-t')$  seulement, ce qui résulte de la non-homogénéité du processus.

On pourrait développer l'analyse en estimant des intensités de transition, limites des probabilités de transition entre  $t$  et  $t+\Delta t$  quand  $\Delta t$  tend vers 0, mais on ne pourrait en tirer des durées moyennes de séjour dans les états du fait de la non homogénéité.

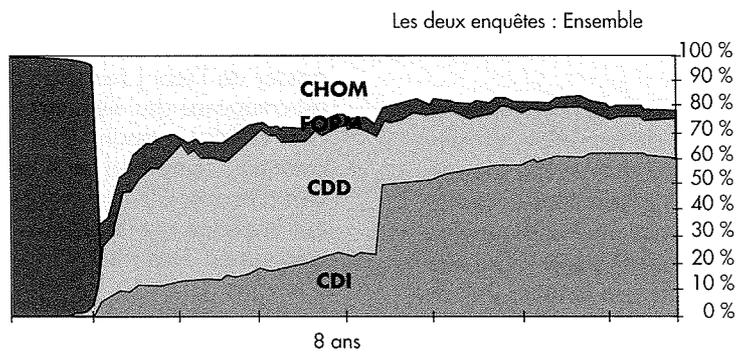
Il est plus intéressant d'utiliser les matrices de transition pour étudier les états stationnaires du processus, où les flux des passages entre les états sont équilibrés de telle sorte que le vecteur d'emploi  $V$  soit invariant.

<sup>8</sup> Un tel processus dénombre ici les occurrences des 4 situations distinctes à aux différentes dates  $t$  envisagées.

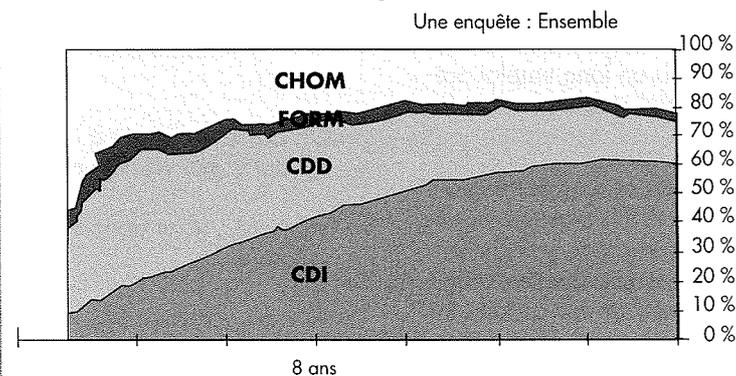
<sup>9</sup> Un processus markovien d'ordre 1 est tel que la situation vécue à l'instant  $t$  ne dépend que de la situation à la date  $(t-1)$ . Ici, toute l'histoire de l'individu aux nombreuses dates qui précèdent, son « chemin parcouru », importe et influence la situation vécue en  $t$ .

**Annexe 1**  
**Fréquences relatives mensuelles des situations**

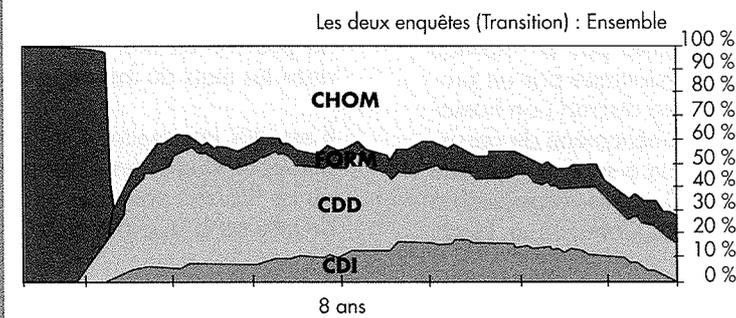
**Graphique 3**  
**1985-1993, ensemble de la population**  
 (Source : enquête 1989 et enquête 1993)



**Graphique 4**  
**1985-1993, ensemble de la population**  
 (Source : enquête 1993)



**Graphique 5**  
**1985-1993, individus en phase de transition**  
 (Source : enquêtes 1989 et 1993)



On a reconnu la définition du vecteur propre associé à la valeur propre 1 de la matrice M (qui est la plus grande des valeurs propres de toute matrice stochastique).

En effet, si l'on doit avoir :  $W = M.W$   
on a  $(M - \lambda.I).W = 0$  avec  $\lambda = 1$

Dans l'ensemble des vecteurs W, définis à une constante multiplicative près, celui qui nous intéresse est alors celui dont la somme des éléments est égale à 1.

À titre d'illustration, nous avons calculé les états stationnaires correspondant aux matrices de transition  $M_1$ ,  $M_6$  et  $M_7$  calculées sur l'ensemble de nos 1 565 individus.

$$W_1 = \begin{pmatrix} 0,276 \\ 0,453 \\ 0,012 \\ 0,259 \end{pmatrix} \quad W_6 = \begin{pmatrix} 0,713 \\ 0,126 \\ 0,017 \\ 0,144 \end{pmatrix} \quad W_7 = \begin{pmatrix} 0,533 \\ 0,157 \\ 0,041 \\ 0,269 \end{pmatrix}$$

On voit ci-dessus que la proportion de CDI « d'équilibre » passe de 27,6 à 71,3 % entre la première et la sixième année. Elle croît donc fortement pour décroître ensuite. On observe parallèlement en septième année une élévation des CDD, des stages et du chômage d'équilibre. Peut-être faut-il voir là un phénomène d'exclusion d'une partie des jeunes qui s'amorcerait une fois que toutes les possibilités d'insertion ont été épuisées ?

Sont présentées en annexe 3 les 7 matrices de transition annuelles ainsi que celle concernant l'ensemble des 83 mois. Les flux observés entre les différentes situations varient fortement dans le temps. On observe, comme pour les vecteurs d'états stationnaires précédents, un changement de tendance entre la sixième et la septième année. La baisse tendancielle des probabilités  $p_{1j}$  d'obtention d'un CDD s'accélère, passant de 0,18 à 0,10 pour le passage CDD-CDI ( $p_{12}$ ), de 0,24 à 0,11 pour les passages chômage-CDI et formation-CDI ( $p_{13}$  et  $p_{14}$ ).

Une définition statistique de l'insertion d'une cohorte peut ainsi correspondre à l'état stationnaire atteint au bout d'un certain temps. Il n'y a donc pas de proportions *a priori* d'individus dans tel ou tel état. Une cohorte sera réputée insérée si, dans une sorte

d'équilibre macroscopique (où les proportions d'individus de catégories différentes restent stables malgré des changements importants de catégories), les passages entre les différentes situations se compensent.

## CHOIX D'UNE DÉFINITION DE L'INSERTION

On comprend que la définition de type macroéconomique qui précède ne peut s'appliquer à l'individu qui vit non pas des probabilités de se trouver dans un état mais une succession de situations. La caractérisation des phases successives de la trajectoire par l'ensemble des états possibles aboutit en général à des séries trop complexes pour être modélisées. De plus, cette succession ne fait pas toujours sens : au-delà du grand nombre de situations successivement occupées, il s'agit de repérer des moments-clés dans le cheminement professionnel des jeunes à l'issue de leur parcours scolaire (Allaire, Beslay, Chartier et Tahar, 1995). Pour rendre compte de la dynamique du processus étudié et de son aspect séquentiel, on a cherché à agréger le processus selon deux phases significatives : une phase de transition dont on cherche à modéliser la durée, compte tenu d'une définition de la date d'insertion à partir de laquelle débute la phase suivante, la phase d'insertion proprement dite (censée se poursuivre au-delà de la date de l'enquête).

La difficulté réside dans le critère de déclenchement de la phase d'insertion. On sait que l'insertion se définit plutôt *ex-post* qu'*ex-ante*. Elle ne dépend pas uniquement de l'obtention d'un emploi à un moment donné mais aussi des situations qui vont suivre. C'est ainsi que pour nous l'insertion est réputée avoir eu lieu dès lors que ne surviennent plus (à l'horizon de l'enquête tout au moins) des événements censés être la preuve de la non insertion. De quels événements s'agit-il ? Il peut être abusif de considérer qu'un court chômage ou une perte d'emploi ruine une insertion. L'insertion peut passer par une succession d'emplois précaires avec mobilité volontaire.

La phase de transition peut être vue comme une phase de recherche d'emploi, effectuée dans un premier temps « à proximité ». Cette proximité est à prendre dans son sens spatial étroit mais également dans un sens professionnel : qualification ou secteur. Mais dans cette acception, l'objectif est l'insertion la plus rapide possible, la mobilité, géographique ou professionnelle, ayant un coût que l'on n'accepte de payer qu'après un certain laps de temps de recherche infructueuse. Il est cependant difficile de séparer la population des mobiles potentiels des immobiles, tout

Tableau 2  
**Les différentes définitions possibles de l'insertion selon les situations acceptées**

Type d'insertion	Emploi stable (code 1)	Emploi précaire (2)	Formation (3)	Chômage -3 m (5)	Chômage > 3 et < 6 (6)
Type 1	oui				
Type 2*	oui			oui	
Type 3*	oui		oui	oui	oui
Type 4*	oui	oui	oui	oui	oui
Type 5**	oui	oui	oui	oui	oui
Type 6***	oui	oui			
Type 7***	oui	oui	oui		
Type 8***	oui		oui	oui	
Type 9***	oui	oui	oui	oui	oui

\* Avec situation 1 obligatoirement au début et à la fin de la période pendant laquelle l'individu est considéré comme inséré.

\*\* Avec début par 1 ou 2 et fin par 1.

\*\*\* Avec fin par 1 ou 2 d'au moins 12 mois.

dépendant des « emplois de réserve » de chacun selon leur localisation et leurs caractéristiques. Sous un angle différent, cette phase de transition peut être vue comme une phase « moratoire » (Galland, 1990) caractérisée par une faible disposition à la mobilité. Des stratégies professionnelles y prennent néanmoins naissance qui pourront changer la disposition envers les différentes formes de mobilité envisageables par la suite.

À partir des calendriers fournis par les données d'enquêtes, on doit définir une procédure qui fixera le changement de phase et donc déterminera la valeur de la durée d'insertion. Plusieurs procédures vont être examinées, qui toléreront dans la phase d'insertion des situations plus ou moins instables. Nous avons défini 7 situations (répertoriées au tableau 1, encadré p.18) que l'on acceptera ou non dans la phase d'insertion, la phase de transition se terminant donc à *partir du moment où n'apparaît plus une situation considérée comme négative*. Pour nous, en particulier, le caractère négatif d'une situation de chômage ou d'inactivité dépend essentiellement de sa durée.

Nous pouvons élaborer différents types d'insertion selon que l'on tolère ou non, partant d'un emploi stable d'abord, stable ou précaire ensuite, certaines situations de chômage, d'inactivité ou de formation. En pratique, on procède en remontant le calendrier à partir de la dernière situation observée. L'insertion est de la sorte définie comme une succession de situations « acceptables ». Selon les différentes variantes testées, on élargit les situations considérées comme acceptables à condition qu'elles soient encadrées par

des périodes d'emploi. Parmi les nombreuses possibilités, nous avons retenu neuf types d'insertions qui sont décrites au tableau 2, en partant de la plus stricte pour aller vers la moins exigeante.

La logique de construction de ces 9 types d'insertion se réfère d'abord à la précarité possible ou non des emplois occupés, ensuite à la durée acceptable des périodes de chômage. Les trois premiers types excluent la possibilité de CDD après l'insertion. À partir du type 4, on accepte que des emplois précaires succèdent à des emplois stables. À partir du type 6, on considère que l'insertion peut démarrer par un emploi précaire mais que l'existence d'un emploi précaire de moins de 12 mois en fin de période d'observation est rédhitoire, faute de « preuve » sur ce qui va suivre. La prise en compte des CDD au même titre que les CDI s'explique par le fait que de plus en plus d'entreprises prennent aujourd'hui « à l'essai » sur contrat à durée déterminée des jeunes qu'elles pourront garder plus longtemps « si tout se passe bien ».

L'hypothèse sous-jacente à ces définitions emboîtées est qu'une période relativement courte (limitée à 3 ou 6 mois) de non-emploi entre deux périodes d'emploi peut correspondre à une période d'ajustement au sein d'une trajectoire de mobilité ascendante, ou tout au moins volontaire, plutôt qu'à une mobilité subie et en quelque sorte un retour « à la case départ ». On aurait pu être plus restrictifs en éliminant les statuts stables courts et l'être un peu moins en autorisant le service national, que l'on a choisi finalement de neutraliser.

À ce stade, nous disposons d'un certain nombre d'ar-

Tableau 3  
**Conséquences des différentes définitions de l'insertion**

Type d'insertion	1	2	3	4	5	6	7	8	9
Proportion d'insérés (en%)	59,2	59,2	59,2	59,2	.....	64,1	66,0	66,8	69,0
Durée moyenne de la phase de transition (en mois)	74,2	74,0	73,7	72,5	67,7	64,2	61,3	60,2	57,4
Nombre moyen d'emplois avant insertion	2,78	2,77	2,75	2,66	2,42	2,25	1,98	1,89	1,67

**Remarques :**

- la durée moyenne de la phase de transition (la « survie » du modèle de durée) est calculée en prenant égale à 97 mois la durée des phases de transition non achevées à la date de l'enquête (durée « censurée à droite » pour le modèle) ;
- le nombre moyen d'emplois, stables ou précaires, sur les 97 mois, sans distinction de phases, est de 3,53 ;
- la durée d'observation de 97 mois retenue par l'enquête intègre la dernière année passée à l'école (avec parfois un emploi concomitant). La durée de la phase de transition après la sortie de l'école s'obtient donc en retranchant 13 mois aux moyennes présentées plus haut.

guments pour définir une notion d'insertion relativement souple qui aurait lieu dès lors que ne surviennent plus des événements interprétables comme une « non insertion », à savoir des interruptions d'activité relativement longues. Avant d'arrêter un choix définitif, il est intéressant de compléter l'analyse par quelques éléments de statistique descriptive.

Les neuf types d'insertion ainsi définis conduisent à des proportions d'insérés, à des durées moyennes de la phase de transition et à des nombres d'emplois avant insertion légèrement différents. Il est à noter que les suites de ces trois paramètres (voir le tableau 3 ci-dessus) sont remarquablement régulières et monotones : ainsi, la proportion d'insérés croît-elle de 59 à 69 %, la survie moyenne (c'est-à-dire le temps moyen mis à s'insérer) décroît-elle de 74 à 57 mois environ et le nombre moyen d'emplois avant insertion passe-t-il de 2,8 à 1,7.

Les interviewes réalisées par ailleurs auprès d'un certain nombre de jeunes de la cohorte (voir Allaire, Beslay, Chartier et Tahar, 1995) nous font pencher en faveur du type 7. Nous allons néanmoins tester les neuf définitions précédentes sur les modèles de durée de la phase de transition associés à chaque type de transition<sup>10</sup>. On se limite aux trois variables explicatives les plus fondamentales, à savoir le sexe, la possession ou non d'un diplôme de formation initiale (caractéristiques individuelles qui différencient bien

notre population) et enfin la présence ou non d'une période de formation pendant la phase de transition<sup>11</sup>. Il ne s'agit pas ici en effet du problème standard du choix du meilleur ensemble de régresseurs pour un modèle mais de celui de la meilleure variable endogène qu'il faut choisir pour cerner le phénomène.

Les résultats économétriques sont présentés annexe 4 (tableau 4). On peut comparer les estimations obtenues au moyen d'un test de significativité globale<sup>12</sup>. Le meilleur ajustement pour les insertions ne considérant que les emplois stables est obtenu avec le type 3, incluant des périodes de chômage inférieures à 6 mois ainsi que des périodes de formation. On améliore l'estimation en prenant en compte les emplois précaires mais là, l'optimum est atteint en ne tenant compte que des chômeurs de moins de 3 mois. La définition 7, en faveur de laquelle des arguments analytiques existaient (comme le souci d'éviter le critère exclusif « CDI ») se trouve confirmée ici : cette définition inclut certains emplois précaires et les chômeurs de moins de 3 mois.

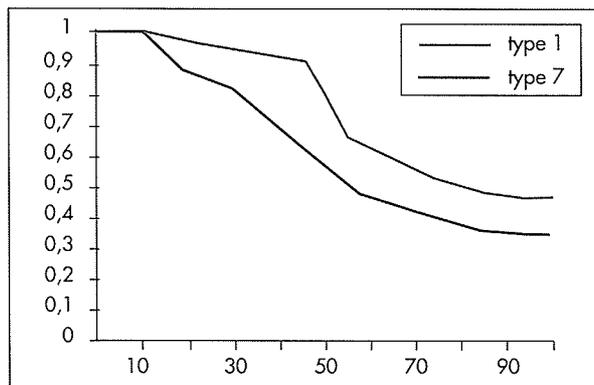
Sur le graphique 1, on constate que la différence entre les insertions de types 1 et 7 va au-delà des durées moyennes et des proportions d'insérés pour affecter le profil de l'évolution du taux d'insertion dans la population de jeunes. À mi-parcours (48ème mois), 30 % des jeunes sont insérés selon la définition 7 alors que moins de 10 % le seraient selon la défini-

<sup>10</sup> On a choisi d'estimer une forme courante de modèle de durée, la spécification paramétrique à distribution Weibull, qui est un modèle à hasard proportionnel dont la variable dépendante est le logarithme du temps passé avant l'insertion.

<sup>11</sup> Ce qui est une façon de différencier ces phases de transition en « contrôlant » l'allongement systématique qui résulte de la présence d'un stage.

<sup>12</sup> Ce qui veut dire que l'on cherche à trouver la définition formelle de l'insertion qui la lie le plus fortement aux caractéristiques individuelles que l'on considère comme les plus déterminantes sur le marché du travail, à savoir le sexe, le diplôme en formation initiale et la formation continue suivie.

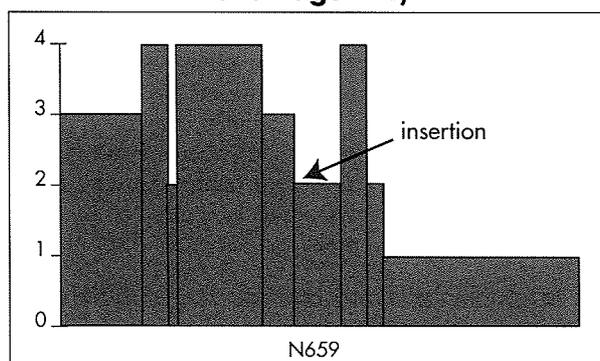
Graphique 1  
Taux de non insérés en fonction  
du temps passé (en mois)



tion la plus stricte. L'écart entre les deux évolutions se réduit ensuite quelque peu pour ne pas dépasser 7 % en fin de période.

L'exemple du graphique 2 ci-dessous (observation n° 659 de l'échantillon) est celui d'un individu inséré au 40ème mois selon notre définition de l'insertion, mais au 56ème mois seulement pour la définition la plus stricte (CDI uniquement).

Graphique 2  
Exemple de trajectoire (emploi stable = 1,  
emploi précaire = 2, formation = 3,  
Chômage = 4)



### MESURE DES ÉCARTS ENTRE TRAJECTOIRES

Quand on classe des chronogrammes (Espinasse, 1993), on considère implicitement que tous les états que peut occuper l'individu à l'instant sont équidistants. Il s'agit ici de fabriquer un indicateur prenant en compte deux idées :

- des situations 1 et 2 peuvent être plus « proches » l'une de l'autre que d'une troisième ;

- une situation 1 peut être plus proche de la situation 2 que cette dernière ne l'est de la situation 1.

Le caractère d'asymétrie n'est pas un défaut puisqu'il est bien normal que la distance entre l'emploi stable et le chômage ( $d_{41}$ ) soit plus grande que celle entre le chômage et l'emploi stable ( $d_{14}$ ) par exemple (voir annexe 3). Inversement mais avec des valeurs différentes, la distance entre le chômage et l'emploi précaire ( $d_{24}$ ) apparaît plus courte qu'entre le chômage et l'emploi stable ( $d_{14}$ ), ce qui ne fait que traduire la difficulté plus grande pour un chômeur d'obtenir un contrat à durée indéterminée. On constate en outre qu'un stage de formation conduit à court terme plus rarement à un emploi stable qu'au chômage mais qu'à long terme la formation continue est « payante » ( $d_{13}$  parfois plus élevée que  $d_{43}$  sur 1 an mais nettement plus faible sur 7).

On va utiliser la matrice de transition déjà définie (voir p.19 indiquer page à l'épreuve) en en déduisant directement une matrice des probabilités conditionnelles au changement d'état qui seront utilisées comme mesures de distance entre les situations.

On a reporté dans l'annexe 1 p. 27 les développements mathématiques sur les caractéristiques de la distance choisie.

A l'instant  $t$ , une situation  $i$ , est d'autant plus « coûteuse » en termes de distance que sa distance élémentaire  $d_{ij}$  à la situation de référence  $j$  est élevée et que la probabilité d'occurrence  $s_{ij}$  de cette situation  $j$  est forte. Par exemple, un chômage pèse plus sur la distance en fin de période qu'en début puisque la probabilité que l'individu moyen s'y trouve lui aussi diminue avec le temps.

Nous avons testé économétriquement la relation entre notre quasi-distance et un ensemble de variables explicatives des caractéristiques des trajectoires que sont communément le sexe, l'âge de sortie du système éducatif, le niveau de formation. Nous y ajoutons, outre des variables de localisation, un indicateur de l'adéquation formation-emploi. Les résultats de l'estimation (voir annexe 4, tableau 4) montrent que la quasi-distance  $D$  s'explique :

- par le sexe, les femmes étant plus éloignées de la référence, quelle qu'elle soit ;
- par l'âge et par la possession d'un diplôme, qui la diminue ;
- par le passage par la formation continue et le nombre d'emplois occupés avant l'insertion, qui l'augmente ;

- par les adéquations formation-emploi aussi bien d'ouvrier que d'employé, qui la diminue ;

- par un premier emploi en zones rurale ou intermédiaire, qui la diminue également.

L'utilisation la plus immédiate de la quasi-distance D réside dans les typologies qu'elle permet de construire, typologies qui ne dépendront plus que du choix de l'ensemble des situations possibles et non des variables explicatives, comme c'est souvent le cas dans les analyses factorielles.

\* \*  
\*

Les trois niveaux d'analyse distincts que l'on a développés situent bien les trois éléments des comparaisons auxquelles on se livre communément en matière d'entrée d'une population de jeunes dans la vie active : le temps mis à s'insérer, le type de carrière prévisible, l'origine des différences individuelles.

Ce travail a essayé de progresser dans la modélisation statistique de l'insertion professionnelle. Il est bien évident que jamais une technique quantitative ne rendra parfaitement compte de la réalité sociale complexe que reflètent les parcours professionnels initiaux des jeunes. Il semble pourtant que nos approches « macro et micro » de l'insertion montrent bien la différence entre définition de niveau statistique et définition au plan individuel.

Après six années, les états stationnaires calculés suggèrent un processus d'insertion terminé pour la cohorte étudiée. Les mécanismes différenciés d'inté-

gration correspondant au processus d'insertion n'agissent plus significativement au cours de la septième année suivant la sortie du système éducatif <sup>13</sup>.

L'insertion d'une cohorte se distingue d'une autre par une insertion globale plus ou moins longue et/ou des proportions d'occupation des différents états différenciées : on pense ici d'une part à l'insertion plus rapide des jeunes qualifiés et aux taux de chômage moins élevés des jeunes diplômés. S'ajoutent des différences interindividuelles que l'on peut quantifier via l'influence de certaines variables explicatives sur la durée d'insertion et les écarts mesurés par rapport à une trajectoire de référence.

La mesure de l'écart entre deux trajectoires par une quasi-distance est féconde au-delà de son intérêt méthodologique car elle permet de synthétiser des différences persistantes au sein de populations relativement homogènes : par exemple les jeunes insérés d'un côté et les jeunes exclus de l'autre.

Les résultats respectifs du modèle de durée standard et du modèle de distance montrent que ce dernier peut constituer un moyen de description alternatif de l'insertion qui ne nécessite pas une définition *a priori* du moment où elle se situe. En effet, notre distance peut se calculer sur l'ensemble des situations connues même si on gagne à préciser une phase de transition.

Plus fondamentalement, cette distance exprime la signification économique des différentes situations envisagées en prenant en compte les changements d'états réellement observés sur le marché du travail pour une cohorte donnée, ce qu'une application directe d'une technique générale d'analyse des données ne permet pas <sup>14</sup>.

Gabriel Tahar  
INRA-Toulouse

<sup>13</sup> Ce n'est pas le moindre intérêt d'avoir poussé l'enquête longitudinale plus loin que les 3-4 ans habituels que d'avoir suggéré cela.

<sup>14</sup> Voir annexes p.27 à 31.

## Bibliographie

Allaire G. et Tahar G. (1996), « L'analyse du différentiel urbain-rural dans l'insertion professionnelle des jeunes de bas niveau scolaire », *Revue d'Économie régionale et urbaine*, n° 2.

Allaire G., Cahuzac E. et Tahar G. (1995), *Persistence du chômage et insertion*, Note LIRHE n° 194, Université de Toulouse I.

Allaire G., Beslay C., Chartier R. et Tahar G. (1995), *Modélisation des trajectoires d'insertion selon quatre phases*, Communication à la 3<sup>e</sup> Journée d'étude du CLERSE « L'insertion professionnelle, nouvelles approches, nouveaux débats », Lille, 7 juin.

Beduwe C. (1991), « Le niveau de diplôme des ouvriers et

des employés, la place des jeunes de niveau V », *Formation Emploi* n° 33, janvier-mars.

Beduwe C., Dauty F. et Espinasse J.-M. (1995), *Trajectoires types d'insertion professionnelle*. Application au cas des bacheliers professionnels, Contribution aux Journées du Longitudinal, Caen, juin, Documents, série séminaires, Céreq, n° 112.

Coeffic N. (1986), « Les jeunes à la sortie de l'école, poids du chômage et risque de déclassement », *Économie et Statistiques*, n° 193-194.

Demazière D. et Dubar C., éd., (1994), *L'insertion professionnelle des jeunes de bas niveau scolaire, trajectoires biographiques et contextes structurels*, Documents, série synthèse, Céreq, n° 91, janvier.

Eckstein Z. et Wolpin K. (1995), « Duration to First Job and the Return to Schooling, Estimates from a Search-matching Model », *Review of Economic Studies*, Avril.

Espinasse J.-M. (1993), *Enquêtes de cheminement, chronogrammes et classification automatique*, Note CEJEE n° 159, Université de Toulouse I.

Galland O. (1990), « Un nouvel âge de la vie », *Revue française de Sociologie*, octobre-décembre.

Geweke J., Marshall R.-C. et Zarkin G.-A. (1986), « Mobility Indices in Continuous Time Markov Chains », *Econometrica*, n° 6, November.

Grelet Y. et Viney X. (1991), « L'insertion professionnelle à l'issue des CAP et des BEP (1980-1988) », *Formation Emploi* n° 33, janvier-mars.

Hecq C., Kestens P., Meulders D., Plasman O. (1994), *Formations et insertion professionnelle, analyse des trajectoires de chômeurs ayant suivi une formation professionnelle*, Communication aux XIV<sup>e</sup> Journées de l'Association d'économie sociale, Lille, septembre.

Huriot J.-M. et Perreur J. (1990), « Distances, espaces et représentations, une revue », *Revue d'Économie régionale et urbaine*, n° 2.

Jones S.R.G. et Riddell W.-C. (1995), The Measurement of Labor Force Dynamics with Longitudinal Data, the Labour Market Activity Survey Filter, *Journal of Labor Economics*, n° 2, April.

Lynch L. (1985), « State Dependency in Youth Unemployment, a Lost Generation? », *Journal of Econometrics*, Annals 1985-1, April.

Lynch L. (1989), « The Youth Labor Market in the Eighties, Determinants of Reemployment Probabilities for Young Men and Women », *Review of Economics and Statistics*, n° 1.

Minni C. et Vergniès J.-F. (1994), « La diversité des facteurs de l'insertion professionnelle », *Économie et Statistique*, n° 277-278.

Narendranathan W. et Elias P. (1993), « Influences of Past History on the Incidence of Youth Unemployment, Empirical Findings for the UK », *Oxford Bulletin of Economics and Statistics*, n° 2.

Penard T. et Sollogoub M. (1995), « Les politiques françaises d'emploi en faveur des jeunes, une évaluation économétrique », *Revue économique*, n° 3.

Pottier F. (1992), « Formes et logiques de mobilité des jeunes à travers l'Observatoire des entrées dans la vie active (EVA) », in Coutrot L., Dubar C. (éd.), *Cheminements professionnels et mobilités sociales*, Paris, La Documentation française.

Rose J. (1984), *En quête d'emploi*, Paris, Économica.

Tanguy L., dir, (1986), *L'introuvable relation formation/emploi*, Paris, La Documentation française.

Taubman P. et Wachter M. (1986), « Segmented Labor Markets », in Ashenfelter O., Layard R. (éd.), *Handbook of Labor Economics*, volume II, North Holland.

Topel R. et Ward M. (1992), Job Mobility and the Careers of Young Men, *Quarterly Journal of Economics*, May.

### Les caractéristiques de la mesure choisie

On va appeler distance entre la situation  $i$  et la situation  $j$  le complément à l'unité de la proportion des passages à la situation  $i$  avec la situation  $j$  comme point de départ, l'ensemble des situations étant construit à partir des 1 565 trajectoires individuelles complètes indépendamment de l'insertion éventuelle (1). La matrice des distances se déduit par conséquent directement de la matrice de transition correspondant aux deux bornes de l'intervalle considéré : elle est constituée des compléments à l'unité des probabilités conditionnelles au changement de situation (c'est-à-dire lorsque  $i \neq j$ ) :

(1)  $d_{ij} = 1 - p_{ji} / (1 - p_{ji})$  (situation  $j$  en  $t$ , situation  $i$  en  $t+1$ ) ;  $d_{ij} = 0$  sinon.

La somme des éléments des colonnes de la matrice des distances n'est pas égale à 1 mais à 2 parce qu'il s'agit des compléments à l'unité et non des probabilités de passage d'un état à un autre (voir annexe 3 les matrices de distance présentées en regard des matrices de transition). Cette définition est relative à la cohorte étudiée et elle ne dépend pas du temps passé dans la situation de départ. On a choisi de prendre les compléments à l'unité des probabilités de passage d'une situation à une autre pour exprimer simplement qu'une probabilité forte traduit une relative proximité et non un éloignement.

On a affaire à une distance un peu particulière puisqu'elle ne satisfait ni l'axiome de symétrie ni l'inégalité triangulaire (2). Le manque d'inégalité triangulaire n'est pas gênant en soi car on n'a pas besoin ici de la cohérence spatiale qui serait requise en cas de représentation dans le plan. On a vu (§5 indiquer page à l'épreuve) que le caractère d'asymétrie est une propriété requise pour l'indicateur que l'on recherche.

On part des probabilités d'appartenance à une situation introduites plus haut (voir §3) qui définissent une trajectoire « de référence », ensemble dans lequel les changements de situations ne sont plus artificiels comme dans les ensembles de situations

les plus probables  $\{qt\}$ . La trajectoire de référence est donc maintenant une matrice  $97 \times 4$   $\{s_{it} ; i = 1, \dots, 4 ; t = 1, \dots, 97\}$  qui est calculée pour tous les individus sur 97 mois, qu'ils s'insèrent ou non pendant la période d'enquête.

On dispose des distances élémentaires  $d_{ij}$  qui vont nous permettre de positionner chaque trajectoire vis-à-vis de la référence. Pour chaque mois  $t$ , la quasi-distance d'un individu  $k$  en situation  $l$  va être la somme des distances entre  $l$  et les 3 autres, pondérées par la probabilité  $s_{jt}$  d'appartenance de la trajectoire de référence à la situation  $j$ . On effectue ensuite une somme pondérée sur l'ensemble des mois constituant la phase de transition, le calcul de la distance s'arrêtant quand l'individu est inséré (au bout de  $t_k$  mois,  $97 \geq t_k$ , pour l'individu  $k$ ) (3). Cela conduit à la définition suivante de la quasi-distance :

$$(2) D_k = \sum_t (\sum_j s_{jt} (d_{jl} | i_{kt})) / t_k$$

$k=1, \dots, 1565$   $t=1, \dots, t_k$   $j=1, \dots, 4$  et  $i_{kt}$  la situation de l'individu  $k$  en  $t$  ( $i=1, \dots, 4$ ).

Cette définition d'une quasi-distance entre une trajectoire donnée et une référence se transpose aisément à une mesure d'écart entre deux trajectoires quelconques.

(1) Cette manière de définir une distance s'inspire de celle utilisée par Hecq et alii (1994).

(2) En effet, on n'a pas nécessairement  $d(i,j) = d(j,i)$  et l'inégalité  $d(i,k) + d(k,j) \geq d(i,j)$  n'est pas vérifiée en général.

Les axiomes de définition d'une distance (ou encore métrique) sont au nombre de 4 : hormis les 2 précédents, il s'agit de la non négativité et de l'identité.  $d(i,j)=0 \Leftrightarrow i=j$  : en effet, on aurait dans le cas contraire  $p_{ji} + p_{ij} = 1$ , ce qui impliquerait que les deux autres probabilités de transition  $p_{ki}$  ( $k \neq i$ ) soient nulles, ce qui n'est jamais observé en pratique (voir annexe 3, où l'on voit que lorsque deux des transitions sont faiblement probables, le calcul donne bien une distance faible pour la transition fortement probable). Formellement néanmoins, il faudrait ajouter à la définition (1) la contrainte  $p_{ij} \neq 0, \forall i, j$ . On parle de quasi-distance quand il n'y a pas symétrie et quand l'inégalité triangulaire n'est pas vérifiée.

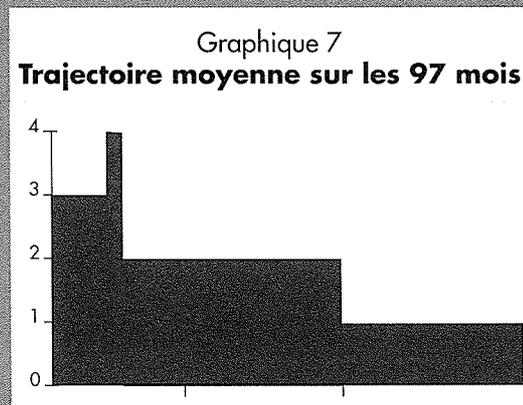
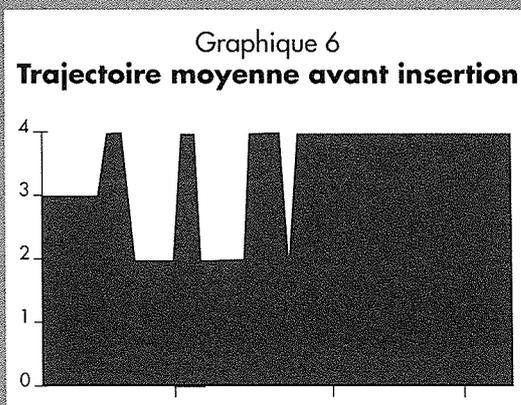
(3) A chaque définition de l'insertion correspond donc un  $t_k$  particulier et donc une mesure. On a choisi ici l'insertion de type 7 déjà retenue (voir §4).

## Annexe 2 Trajectoires moyennes

La trajectoire moyenne peut être établie en ne prenant en compte que les phases de transition dans les trajectoires individuelles ou en considérant au contraire l'ensemble des 97 mois de l'enquête. On trouve ci-dessous la visualisation des trajectoires moyennes correspondant aux deux définitions possibles. La trajectoire moyenne définit sur la phase de transition ne se termine pas par une période

d'emploi mais par du chômage (ou de l'inactivité) puisqu'au fur et à mesure que l'on avance dans le temps, ne subsistent dans l'ensemble considéré que les jeunes non insérés.

Les deux trajectoires moyennes commencent par 13 mois de formation puisque le calendrier de l'enquête démarre une année avant la fin de la formation initiale pour tenir compte des éventuelles prises d'emploi en cours d'études (qui sont malgré tout en nombre très faible).



**Légende :** en abscisse, mois d'observation, de 1 à 97 ; en ordonnées, situation 0 : inséré ou service national, situation 1 : emploi stable, situation 2 : emploi précaire, situation 3 : formation ou stage, situation 4 : chômage ou inactivité.

## Annexe 3 Matrices de transition et de distance

On donne ci-après les matrices de transition  $p(i,j)$  de l'état  $j$  à l'état  $i$  et les tableaux des distances  $d(i,j)$  qui en sont déduites pour les différentes années d'enquête (mois 14 à 25, 26 à 37, 38 à 49, 50 à

61, 62 à 73, 74 à 85, 86 à 96) et pour l'ensemble de la période (mois 14 à 96). On rappelle que les états 1, 2, 3 et 4 correspondent respectivement à un emploi stable, un emploi précaire, une période de formation et une période de chômage. La huitième et dernière matrice, calculée sur l'ensemble de la période, est le produit des 7 précédentes.

(suite p. 29)

Tableau 4

Début : 14 Fin : 25								
Transition					Distance			
1	0.910499	0.142687	0.125236	0.158280	0.000000	0.602074	0.801208	0.758018
2	0.038374	0.641424	0.257170	0.445904	0.571248	0.000000	0.591782	0.318291
3	0.005239	0.028009	0.370018	0.049913	0.941462	0.921889	0.000000	0.923691
4	0.045888	0.187881	0.247576	0.345903	0.487290	0.476036	0.607010	0.000000
Début : 26 Fin : 37								
Transition					Distance			
1	0.930854	0.262230	0.176384	0.274732	0.000000	0.450851	0.779927	0.620684
2	0.033937	0.522479	0.373532	0.414005	0.509194	0.000000	0.533947	0.428394
3	0.002785	0.018346	0.198520	0.035547	0.959717	0.961582	0.000000	0.950922
4	0.032423	0.196945	0.251564	0.275717	0.531089	0.587567	0.686126	0.000000
Début : 38 Fin : 49								
Transition					Distance			
1	0.889297	0.266229	0.227176	0.282520	0.000000	0.485090	0.697015	0.562053
2	0.054979	0.482960	0.356252	0.324122	0.503364	0.000000	0.524865	0.497563
3	0.006387	0.029461	0.250208	0.038458	0.942304	0.943020	0.000000	0.940384
4	0.049337	0.221350	0.166364	0.354900	0.554332	0.571890	0.778120	0.000000
Début : 50 Fin : 61								
Transition					Distance			
1	0.910243	0.284061	0.301895	0.000000	0.000000	0.456199	0.671842	0.532381
2	0.034649	0.477638	0.268077	0.613973	0.613973	0.000000	0.637462	0.584763
3	0.006023	0.032777	0.075629	0.932897	0.932897	0.937253	0.000000	0.882855
4	0.049086	0.205525	0.354398	0.453130	0.453130	0.606548	0.690696	0.000000
Début : 62 Fin : 73								
Transition					Distance			
1	0.915824	0.241024	0.293504	0.242438	0.000000	0.494077	0.640363	0.550211
2	0.031618	0.523595	0.239442	0.270919	0.624385	0.000000	0.706607	0.497369
3	0.004207	0.012340	0.183889	0.025646	0.950025	0.974098	0.000000	0.952420
4	0.048351	0.223041	0.283166	0.460997	0.425590	0.531825	0.653030	0.000000
Début : 74 Fin : 85								
Transition					Distance			
1	0.907959	0.176690	0.235839	0.236499	0.000000	0.628013	0.694550	0.561838
2	0.034439	0.525009	0.221905	0.260447	0.625828	0.000000	0.712597	0.517469
3	0.003181	0.027436	0.227896	0.042806	0.965441	0.942239	0.000000	0.920693
4	0.054421	0.270864	0.314360	0.460248	0.408731	0.429748	0.592852	0.000000
Début : 86 Fin : 96								
Transition					Distance			
1	0.900748	0.099465	0.114193	0.105806	0.000000	0.819572	0.836988	0.728928
2	0.029368	0.448729	0.186806	0.214919	0.704107	0.000000	0.733332	0.449382
3	0.008280	0.047575	0.299482	0.069599	0.916579	0.913699	0.000000	0.821690
4	0.061604	0.404231	0.399520	0.609676	0.379314	0.266729	0.429680	0.000000
Début : 14 Fin : 96								
Transition					Distance			
1	0.702525	0.627659	0.621097	0.629180	0.000000	0.278189	0.359529	0.202164
2	0.101496	0.130439	0.132922	0.129840	0.658807	0.000000	0.862932	0.835356
3	0.023866	0.029701	0.030249	0.029588	0.919771	0.965844	0.000000	0.962480
4	0.172112	0.212201	0.215732	0.211392	0.421422	0.755968	0.777539	0.000000

**Annexe 4**  
**Modèles économétriques**

Tableau 5  
**Estimations des différentes définitions de l'insertion**  
**(modèle de durée paramétrique à distribution Weibull)**

Type d'insertion*	1	2	3	4	5	6	7	8	9
<b>Constante</b>	4,438 (91,2)	4,435 (89,4)	4,433 (88,3)	4,425 (81,8)	4,392 (61,8)	4,245 (64,5)	4,134 (62,5)	4,169 (58,8)	4,094 (54,7)
<b>Sexe</b>	0,1655 (5,49)	0,1695 (5,51)	0,1723 (5,54)	0,1836 (5,46)	0,2312 (5,20)	0,2464 (5,98)	0,2594 (6,33)	0,2632 (5,94)	0,2769 (6,03)
<b>Diplôme</b>	-0,2053 (6,24)	-0,2111 (6,29)	-0,2150 (6,33)	-0,2321 (6,33)	-0,3117 (6,42)	-0,2751 (6,23)	-0,2652 (6,02)	-0,2959 (6,27)	-0,3040 (6,20)
<b>Stage</b>	0,3155 (8,32)	0,3226 (8,35)	0,3258 (8,35)	0,3515 (8,34)	0,4710 (8,42)	0,4410 (8,64)	0,4343 (8,54)	0,4257 (8,01)	0,4334 (7,94)
<b>Stat. de Wald</b> $\sim \chi^2(3)$	149,97	151,61	152,19	149,34	145,22	157,19	159,55	141,01	137,65

t de Student entre parenthèses.

Le test de Wald est asymptotiquement équivalent au test du ratio de vraisemblance mais ne nécessite pas l'estimation du modèle constant.

\* Les types d'insertion sont indiqués dans le tableau 1 p. 22.

**Codage des variables :**

- sexe : 1 pour masculin, 2 pour féminin ;
- possession d'un diplôme : 1 pour oui, 0 pour non ;
- stage (formation continue ou reprise d'études pendant la phase de transition) : 1 pour oui, 0 pour non.

**Annexe 4**  
**Modèles économétriques**

Tableau 6  
**Modèle explicatif (MCO) de la quasi-distance entre une trajectoire individuelle et la trajectoire de référence moyenne (probabilités d'occurrence des différentes situations à chaque instant)**

n = 1565	D
Constante	-0,2437 (1,78)
Sexe	0,07544*** (15,2)
Année de naissance	0,004366* (2,18)
Nb d'emplois avant insertion	-0,01776** (3,24)
Adéquation « ouvrier »	-0,02623*** (3,77)
Adéquation « employé »	-0,06694*** (8,88)
1 <sup>er</sup> emploi en rural profond	-0,02161** (3,06)
1 <sup>er</sup> emploi en rural intermédiaire	-0,02333*** (4,01)
moyenne de D	0,2123
$\sigma$	0,08599
R2 corrigé	0,5607
F (9,1555)	222,8***

t de Student entre parenthèses  
\* : coefficient significatif au seuil de 5 % ;  
\*\* : coefficient significatif au seuil de 1 % ;  
\*\*\* : coefficient significatif au seuil de 0,1 %.

**Signification des variables :**

Toutes les variables sont dichotomiques à l'exception du nombre de périodes d'emploi avant insertion (variable variant entre 0 et 10) et l'année de naissance (variable variant entre 62 et 72).

Hormis les 3 variables déjà décrites au tableau 3, les codages des variables indicatrices utilisées sont les suivants :

- adéquation « ouvrier » : 1 si la spécialité de formation est de type secondaire (groupes 3 à 27 de la nomenclature en 47 groupes de l'Éducation nationale) et le secteur d'activité l'industrie ou le bâtiment ;

- adéquation « employé » : 1 si la spécialité est tertiaire (appartenant au groupe 28 ou au-delà) et l'emploi dans les services ;

- la valeur par défaut de la zone du premier emploi est la zone urbaine, où se situe également la majorité des individus. Le rural profond comprend les communes rurales ou unités urbaines appartenant à une ZPIU de moins de 10 000 habitants. Le rural intermédiaire comprend les communes rurales ou unités urbaines appartenant à une ZPIU comprise entre 10 000 et 50 000 habitants, plus les communes rurales hors unités urbaines appartenant à une ZPIU comprise entre 50 000 et 100 000 habitants.

# L'ORIENTATION

## SCOLAIRE ET PROFESSIONNELLE

revue de  
l'Institut National d'Étude du Travail et d'Orientation Professionnelle

N° 2 - JUIN 1998

### Numéro spécial : L'EXPÉRIENCE SCOLAIRE

Bernadette DUMORA et François DUBET  
L'expérience scolaire : introduction

François DUBET et Danilo MARTUCCELLI  
Sociologie de l'expérience scolaire

Anne BARRERE  
Le travail scolaire : crise des significations et réponse lycéenne

Bernadette DUMORA  
Expérience scolaire et orientation

Lyda LANNEGRAND-WILLEMS  
Adolescence et pairs

Jean-Claude PUJOL et Cécile GONTIER  
L'école et les parents : pratiques et représentations

Joëlle FAVRE-PERROTON  
Racisme et école

Eric DEBARBIEUX, Alix GARNIER et Laurence TICHIT  
Insécurité, incivilités et fracture sociale au collège

Olivier COUSIN  
De l'institution à l'établissement. Le cas des collèges de banlieue

*Abonnement (4 numéros par an) : Tarif 1998*

<b>France :</b> .....	<b>280 FF</b>
<b>Étranger :</b> .....	<b>350 FF</b>
<b>Vente au numéro :</b> .....	<b>95 FF</b>

*Adressez directement commande et paiement à :*  
Régisseur des recettes de l'I.N.E.T.O.P.  
41, rue Gay-Lussac, 75005 Paris  
Tél. : 01 44 10 78 33