

PANEL D'INDIVIDUS VERSUS PANEL DE LOGEMENTS OU : QUE PEUT-ON DIRE DE LA QUALITÉ DU PANEL EUROPÉEN ?

P. BREUIL-GENIER, N. LEGENDRE et H. VALDELIEVRE

Insee, Division Revenus et Patrimoine

Introduction : panel d'individus versus panel de logements

Le panel européen mis en place à la demande d'Eurostat en 1994 est un *panel d'individus* : toutes les personnes considérées comme répondantes lors de la première vague (automne 1994) ont été de nouveau sollicitées pour les vagues annuelles suivantes à condition de ne pas avoir été non-répondantes pendant deux vagues de suite¹. La collecte de la vague 7 s'est déroulée à l'automne 2000, mais seules les vagues 1 à 4 (1994 à 1997) seront exploitées ici. Les « individus panels » sont suivis même s'ils déménagent, sauf s'ils sortent du champ de l'enquête (en déménageant en institution ou à l'étranger) ou bien sûr s'ils décèdent (dans ces deux cas, on qualifiera ces individus de « hors champ » dans la suite).

Un panel d'individus diffère donc dans son principe d'un panel de logements. En effet, dans des panels de logements - comme les enquêtes Emploi ou les Enquêtes Permanentes Conditions de Vie (PCV) - une partie des logements enquêtés une année sont de nouveau enquêtés l'année suivante (un tiers dans les enquêtes Emploi, la moitié dans les enquêtes PCV). Dans le cas de l'enquête Emploi, il n'est donc possible de reconstituer un historique des trajectoires d'activité sur trois ans que pour les personnes qui n'ont pas déménagé pendant trois ans. Or on peut penser que les personnes n'ayant pas déménagé ne sont pas représentatives de la population française. C'est par exemple le constat que fait Magnac (1997)² qui exploite trois années de l'enquête Emploi pour étudier l'influence des stages sur l'insertion professionnelle des jeunes de 18 à 29 ans : ne retenir pour l'étude que les jeunes qui ont été enquêtés trois années de suite conduit à travailler sur un échantillon de personnes plus jeunes, venant de familles où les parents ont un niveau d'éducation plus faible mais divorcent moins³. Le biais potentiel lié à l'utilisation d'un sous-

¹ A chaque vague d'enquête sont également interrogées toutes les personnes qui appartiennent au ménage d'un individu panel.

² T. Magnac (1997) : « Les stages et l'insertion professionnelle des jeunes : une évaluation statistique », *Economie et Statistique*, n°304-305, pp. 75-94.

³ Magnac (1997) note par ailleurs que ce biais de sélection ne serait pas un problème pour l'analyse économétrique si les caractéristiques inobservables qui expliquent la sélection n'influençaient pas le

échantillon de personnes n'ayant jamais déménagé est par ailleurs d'autant plus susceptible d'être important que les populations étudiées sont mobiles ou dans des situations précaires. Ainsi, dans l'étude de Lagarenne et Legendre (2000)⁴, 20 % des travailleurs pauvres présents dans l'enquête Revenus Fiscaux de 1996 (basée sur le tiers médian de l'enquête Emploi de 1997) ne répondent pas à l'enquête Emploi 1998. Cette proportion monte à 32% pour les travailleurs pauvres qui ont connu l'année précédente un parcours incluant des passages vers l'inactivité, alors qu'elle n'est que de 9 % pour les personnes qui ont exercé une activité d'indépendant pendant les douze derniers mois.

Si se restreindre aux personnes n'ayant jamais déménagé pour réaliser des études longitudinales est susceptible de générer des biais, ce que l'on illustrera dans le cas du panel européen (Ib), le recours à un panel d'individus ne résout pas en pratique tous les problèmes. Ainsi, il est difficile de suivre les individus qui déménagent (Ia), et ceux que l'on retrouve ne sont pas parfaitement représentatifs de l'ensemble (Ib). Toutefois, l'utilisation d'une pondération adaptée pour l'exploitation longitudinale d'un panel de logements peut limiter ce dernier biais (Ic). Si la pondération peut également permettre de réduire le biais lié à l'absence de suivi en cas de déménagement dans le cadre d'un panel de logement, elle conduit à multiplier très fortement le poids de certains individus (Id).

Au final, le recours à un panel d'individus reste sans doute de toute façon préférable, dès lors que le phénomène étudié n'est pas indépendant de l'existence d'un déménagement, ou que le nombre de vagues d'enquêtes à exploiter est important ou la population à étudier est mobile⁵. Mais quel est le prix à payer pour disposer d'un panel d'individus ? Financièrement d'abord, le coût moyen par questionnaire rempli est plus élevé (IIa), d'où des échantillons souvent plus réduits. La précision des phénomènes que l'on peut étudier à partir du panel reste-t-elle alors suffisante, ou, en d'autres termes, celui-ci continue-t-il à être suffisamment représentatif de la population de l'année courante au fur et à mesure que le nombre de vagues augmente (IIc) ? On verra que sur l'exemple des taux de transition entre emploi, chômage et inactivité, l'éventuelle perte de représentativité de la vague 4 du panel par rapport à l'enquête Emploi semble limitée.

phénomène étudié (en l'occurrence l'insertion) mais précise que cela n'est probablement pas le cas (par exemple, si un jeune qui trouve un emploi stable quitte ses parents, il n'est plus observé).

⁴ C. Lagarenne, N. Legendre (2000) : « Les travailleurs pauvres en France : facteurs individuels et familiaux », *Economie et Statistique*, n°335.

⁵ Nous ne discutons pas dans cette étude de l'alternative consistant à n'effectuer qu'une enquête ponctuelle comportant un important calendrier rétrospectif. Ce type d'enquête a l'avantage de ne pas poser de problème d'attrition, mais a l'inconvénient d'être sensible aux problèmes de mémoire (ce qui interdit sans doute les questions sur certains thèmes, comme les revenus).

1. Peut-on (et faut-il ?) suivre les individus qui déménagent ?

1.1 Plus de 8 individus sur 10 sont encore répondants après un déménagement

Parmi les individus panels de 17 ans et plus (encadré 1) qui remplissent un questionnaire individuel lors d'une vague, environ 90 % en remplissent encore un la vague suivante (tableau 2). Bien que le taux de réponse soit plus faible parmi ceux qui ont déménagé (annexe), il reste élevé (entre 75 et 84 % selon les vagues). Au total, 82 % des personnes interrogées en première vague ont également répondu aux deux suivantes. Ce taux de réponse sur trois vagues sur les données françaises est proche de ceux observés en Belgique, en Grèce, en Allemagne ou au Luxembourg. Seuls le Portugal et l'Italie arrivent à des taux de réponse nettement supérieurs (90 %), tandis qu'en Grande-Bretagne et en Irlande les taux de réponse sont plus faibles (65 %) ⁶.

Si l'on avait renoncé à suivre les individus qui déménagent, la part d'individus de la première vague pour lesquels on aurait eu des réponses pour les trois vagues suivantes aurait été de 58 % (et non de 73 %, tableau 1), soit une diminution supplémentaire de 20 % de la taille de l'échantillon. Et cette perte d'échantillon s'accroît avec le nombre de vagues ⁷.

Tableau 1 - Incidence du déménagement sur le pourcentage d'individus réinterrogés

	%	
	Absence de suivi en cas de déménagement	Suivi en cas de déménagement
Vague 2	82	88
Vagues 2 et 3	71	82
Vagues 2, 3 et 4	58	73

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panels de 17 ans ou plus ayant rempli un questionnaire individuel à la vague 1 – Données non pondérées

Lecture : parmi les individus interrogés à la vague 1, 73% répondent aux trois vagues suivantes, mais seuls 58% répondent aux trois vagues suivantes et ne déménagent pas entre les vagues 1 et 4

⁶ Eurostat (2000) : Income Poverty and Social exclusion in the member states of the European Union.

⁷ Elle deviendrait particulièrement forte dans un panel « long ». En effet, tous âges confondus, une personne sur deux a déménagé entre les recensements de 1990 et de 1999 (Baccaïni B. : « Les migrations en France entre 1990 et 1999 », *Insee Première* n°758, février 2001).

Tableau 2 - Taux de déménagement et taux de réponse en fonction de l'existence d'un déménagement

	v1->v2		v2-> v3		v3-> v4		v1-> v4		
	réparti- tion	taux de réponse	réparti- tion	taux de réponse	réparti- tion	taux de réponse	réparti- tion	taux de réponse v4	taux de réponse v2, v3,v4
n'ont pas déménagé	91	90	90	94	89	91	75	80	77
ont déménagé	8	79	9	84	10	75	21	71	71
<i>dont :</i> <i>ont déménagé avec leur ménage</i>	6	81	7	86	7	77	15	73	70
<i>dont :</i> <i>ont déménagé en se séparant de leur ménage</i>	2	73	2	79	2	70	6	68	63
hors champ (dont décès)	1	0	1	0	1	0	3	2	0
total /moyenne	100	88	100	93	100	88	100	84	73
effectif	14334	12677	12992	12018	12533	11067	14334	10785	10409

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panels de 17 ans ou plus ayant rempli un questionnaire individuel à la première des deux vagues considérées - Données non pondérées

Lecture : parmi les individus de 17 ans ou plus ayant rempli un questionnaire individuel en vague 1, 91 % n'ont pas déménagé en V2, et parmi eux, 90 % ont répondu au questionnaire V2.

Dans plus de la moitié des cas, les non-réponses faisant suite à un déménagement seraient dues au fait que le nouveau ménage n'a pas été retrouvé (soit l'enquêteur n'a pas obtenu la nouvelle adresse, soit le ménage n'a pas été retrouvé à cette adresse). Lorsque la personne est retrouvée à sa nouvelle adresse, sa probabilité de réponse semble comparable à celle des enquêtés n'ayant pas déménagé (tableau 3).

Au sein de ceux qui déménagent, le taux de non-réponse varie encore en fonction du « type » de déménagement. Il est en effet plus important chez les personnes qui quittent leur ménage (enfants qui partent, conjoints qui se séparent) que chez celles qui déménagent avec l'ensemble de leur ménage. Les taux de non-réponse sont particulièrement élevés (plus d'un tiers) en cas de séparation de conjoints ou de départ d'une personne autre qu'un conjoint ou un enfant. En effet, c'est dans ces cas que retrouver l'adresse de la personne qui est partie est le plus difficile (tableau 3).

Tableau 3 - Fréquence et causes de la non-réponse en fonction du type de déménagement

	répartition	taux de non-réponse	dont : non réponse liée à une adresse non retrouvée	dont : non réponse liée à un refus, une absence de contact ou une raison inconnue
déménagement sans éclatement	76	18	10	8
déménagement avec éclatement	24	26	14	12
séparation du conjoint	19	36	21	15
départ d'un enfant	73	22	11	10
départ d'une autre personne	8	44	22	21
total déménagement	100	20	11	9

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : Individus panels de 17 ans et plus ayant rempli un questionnaire individuel à la vague précédente et qui ont déménagé entre les deux vagues tout en restant dans le champ de l'enquête. Moyennes non pondérées sur les vagues 2, 3 et 4.

1.2 Ce sont les jeunes et les moins insérés sur le marché du travail qui déménagent le plus

Parmi les individus de 17 ans et plus répondants en vague 1 qui sont restés dans le champ de l'enquête lors des trois vagues suivantes, 22 % ont déménagé entre les vagues 1 et 4, parmi lesquels 68 % sont répondants les 3 vagues suivantes. Ne pas réinterroger ces individus après leur déménagement pose deux problèmes : en premier lieu, cela réduit fortement la taille de l'échantillon exploitable pour des études longitudinales - donc la précision des estimations - mais surtout, cela risque d'introduire un biais dans les résultats, si les individus que l'on ne réinterroge pas sont différents de ceux que l'on réinterroge (et en particulier si ces différences portent sur des caractéristiques inobservables). Or, comme nous allons le voir, les individus qui sont les plus susceptibles de déménager sont les jeunes et les personnes les moins insérées sur le marché du travail⁸, c'est-à-dire deux populations pour lesquelles l'intérêt d'études longitudinales est particulièrement important. Mais nous verrons aussi que ce sont également ces populations que l'on a le plus difficulté à suivre en cas de déménagement.

Près d'un jeune de 17-25 sur deux déménage en trois ans

La probabilité de déménagement varie fortement en fonction de l'âge. Elle est maximale pour les 21-25 ans, qui sont 55 % à déménager au cours des trois années étudiées, et elle reste très élevée pour les 26-30 ans (43 %) et les 17-20 ans (36 %). Après 26 ans, le taux de déménagement en trois ans décroît avec l'âge, jusqu'à un taux de 7 % chez les personnes de plus de 61 ans (tableau 4) .

⁸ Voir aussi les résultats de Gobillon L. (2000) : « Emploi, logement et mobilité géographique », miméo Insee, sur les migrations inter-urbaines.

Tableau 4a : Ventilation des individus panels selon l'existence d'un déménagement et la réponse aux vagues 2 à 4, en fonction des caractéristiques des individus en début de période.

Caractéristiques individuelles	Ensemble			Pas de déménagement			Déménagement			Part	
	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Non retrouvés		Autres non réponses
Niveau d'études atteint											
Aucun diplôme	100	72	28	91	67	24	9	5	2	2	26
CAP, BEP	100	74	27	78	60	18	22	14	5	4	20
BEPC	100	75	25	78	60	17	23	15	4	4	26
Baccalauréat	100	75	25	71	55	16	29	21	3	6	7
Baccalauréat plus 2 années	100	79	21	66	55	11	34	24	4	5	12
Supérieur à bac plus 2	100	78	22	64	51	13	36	27	4	5	10
Total des individus	100	75	25	78	60	18	22	15	3	4	100
Occupation											
Actif occupé	100	78	22	76	61	15	24	17	3	4	51
Chômeur	100	64	36	63	43	21	37	21	10	6	8
Étudiant	100	72	28	60	45	15	41	27	7	7	9
Retraité	100	73	27	93	69	24	7	4	1	2	22
Autre inactif	100	75	26	84	64	20	16	11	3	3	11
Total des individus	100	75	25	78	60	18	22	15	3	4	100
Âge											
17-20 ans	100	71	29	64	48	16	36	23	6	7	7
21-25 ans	100	69	31	45	33	13	55	37	10	9	9
26-30 ans	100	76	24	57	44	13	43	32	7	5	10
31-40 ans	100	79	21	75	62	13	25	17	4	4	19
41-50 ans	100	78	22	89	70	18	11	7	2	2	19
51-60 ans	100	75	25	89	67	21	12	8	1	2	13
Plus de 61 ans	100	72	28	93	67	25	7	4	1	2	23
Total des individus	100	75	25	78	60	18	22	15	3	4	100

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panel de 17 ans ou plus en 1994 ayant rempli un questionnaire individuel en première vague. Individus hors champ en vagues 2, 3 ou 4 exclus. Caractéristiques de 1994. Données pondérées (pondérations de la vague 1)

Tableau 4a (suite) : Ventilation des individus panels selon l'existence d'un déménagement et la réponse aux vagues 2 à 4, en fonction des caractéristiques des individus en début de période.

Caractéristiques individuelles	Ensemble			Pas de déménagement			Déménagement			Part	
	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Non retrouvés		Autres non réponses
Statut des actifs											
Salarié	100	78	22	75	60	15	25	18	3	4	88
Indépendant	100	76	24	84	65	19	16	11	3	3	13
Total des actifs	100	78	22	76	61	15	24	17	3	4	100
PCS des actifs											
Agriculteur	100	82	18	90	74	16	11	9	0	2	4
Artisan	100	71	29	80	60	20	20	11	5	4	7
Cadre supérieur	100	80	20	73	60	13	27	20	2	4	13
Profession intermédiaire	100	82	18	74	62	12	27	21	2	4	22
Employé	100	77	23	75	59	16	25	18	4	4	29
Ouvrier	100	76	24	76	61	16	24	16	4	4	25
Total des actifs	100	78	22	76	61	15	24	17	3	4	100
Type de contrat de travail des salariés											
CDD	100	73	27	61	47	14	39	26	6	7	10
CDI	100	79	21	76	62	15	24	18	3	4	90
Total des salariés	100	78	22	74	60	14	26	18	3	4	100

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panel de 17 ans ou plus en 1994 ayant rempli un questionnaire individuel en première vague. Individus hors champ en vagues 2, 3 ou 4 exclus. Caractéristiques de 1994. Données pondérées (pondérations de la vague 1)

Tableau 4b : Ventilation des individus panéls selon l'existence d'un déménagement et la réponse aux vagues 2 à 4, en fonction des caractéristiques de leur ménage en début de période.

Caractéristiques ménage	Ensemble			Pas de déménagement			Déménagement			Part
	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Non retrouvés	
Type de ménage										
Personne seule	100	72	28	74	55	19	26	16	4	5
Couple sans enfant	100	73	27	81	60	21	19	13	3	3
Couple avec enfant	100	78	22	77	62	15	23	16	3	4
Famille monoparentale	100	70	30	68	51	18	32	20	5	7
Autre type de ménage	100	73	27	80	60	20	20	12	4	4
Total des ménages	100	75	25	78	60	18	22	15	3	4
ZEAT										
Ile de France	100	71	30	75	55	20	25	15	6	4
Bassin parisien	100	76	24	80	62	18	20	14	3	3
Nord	100	74	26	80	60	20	20	14	2	4
Est	100	76	24	80	61	20	20	15	2	3
Ouest	100	78	22	80	63	16	21	14	3	3
Sud-ouest	100	77	23	80	63	17	20	14	2	4
Centre-est	100	77	24	76	59	17	24	17	2	4
Méditerranée	100	73	27	75	59	16	26	15	7	4
Total des ménages	100	75	25	78	60	18	22	15	3	4
Strate										
Commune rurale	100	78	22	86	68	17	14	10	2	3
Unité urbaine < 20000 hab	100	76	24	78	60	17	22	16	3	4
20000 <= u.u. <= 100000 hab	100	72	28	76	56	21	24	17	3	5
u.u. de plus de 100000 hab	100	74	26	73	56	17	27	18	5	4
Commune de l'unité urbaine de Paris	100	72	28	75	56	19	25	15	6	4
Total des ménages	100	75	25	78	60	18	22	15	3	4

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panéls de 17 ans ou plus en 1994 ayant rempli un questionnaire individuel en première vague. Individus hors champ en vagues 2, 3 ou 4 exclus. Caractéristiques de 1994. Données pondérées (pondérations de la vague 1)

Tableau 4b (suite) : Ventilation des individus panels selon l'existence d'un déménagement et la réponse aux vagues 2 à 4, en fonction des caractéristiques de leur ménage en début de période.

Caractéristiques ménage	Ensemble			Pas de déménagement			Déménagement			Part
	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Non retrouvés	
Statut d'occupation du logement										
Locataire(yc, sous locataire et logé gratuitement)	100	73	28	61	46	15	39	27	6	39
Propriétaire	100	76	24	89	69	20	11	7	2	61
Total des ménages	100	75	25	78	60	18	22	15	3	100
Propriétaire (pour locataires)										
Famille	100	75	25	77	58	18	23	17	3	14
Propriétaire privé	100	72	28	50	38	12	50	34	9	46
HLM ou organisme public	100	73	27	70	52	18	30	21	4	41
Total des ménages de locataires	100	72	28	62	46	16	38	26	6	100
Revenu mensuel / uc (quartiles)										
Revenu ≤ 4400 F	100	70	30	76	55	22	24	15	5	25
4400 F < revenu ≤ 6500 F	100	74	26	80	61	19	21	14	3	25
6500 F < revenu ≤ 9330 F	100	79	21	78	64	14	22	15	3	25
9330 F < revenu	100	76	24	78	61	17	22	16	3	25
Total des ménages	100	75	25	78	60	18	22	15	3	100

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panel de 17 ans ou plus en 1994 ayant rempli un questionnaire individuel en première vague. Individus hors champ en vagues 2, 3 ou 4 exclus. Caractéristiques de 1994. Données pondérées (pondérations de la vague 1)

Tableau 4c : Ventilation des individus panels selon l'existence d'un déménagement et la réponse aux vagues 2 à 4, en fonction de leur trajectoire sur le marché de l'emploi entre nov.1993 et oct.1994.

	Ensemble			Pas de déménagement			Déménagement				Part	
	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Non retrouvés	Autres non réponses		
Trajectoire d'emploi entre nov. 93 et oct.94 (12 mois)												
Emploi tous les mois	100	79	21	78	63	15	22	16	3	3	58	
tous les mois à temps complet	100	79	21	76	62	15	24	18	3	4	44	
au moins un mois en temps partiel	100	82	18	79	66	13	21	16	3	3	7	
au moins un mois en indépendant	100	76	25	84	66	18	16	10	3	4	8	
Emploi et non emploi	100	69	31	61	44	16	39	25	8	7	13	
emploi puis chômage	100	62	38	59	43	16	41	20	12	9	2	
chômage puis emploi	100	71	29	63	47	16	37	24	6	8	2	
alternances d'emploi et de chômage	100	68	32	59	42	17	41	26	10	5	3	
activité et inactivité (+ chômage)	100	72	29	61	45	16	39	26	6	6	6	
Jamais d'emploi	100	73	28	71	54	17	29	18	6	5	29	
Ensemble (âge actif)	100	76	24	74	58	16	26	18	4	4	100	

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panel de 17 ans à 60 ans répondants en 1994. Individus hors champ en vagues 2, 3 ou 4 exclus. Données pondérées (pondérations de la vague 1)

Bien que les 17-25 ans aient un fort taux de non-réponse en cas de déménagement (environ un tiers de non-répondants), suivre ceux qui déménagent permet d'obtenir des réponses sur quatre vagues pour 70 % de cette classe d'âge, alors que n'interroger que ceux qui n'ont pas déménagé ne permettrait d'obtenir cette information longitudinale que pour 48 % des 17-20 ans et 33 % des 21-25 ans (tableau 4), soit, en moyenne sur ces tranches d'âges, un échantillon deux fois plus faible qu'en cas de suivi lors de déménagements. L'intérêt (tout du moins en termes de taille d'échantillon) de suivre les individus qui déménagent reste important au sein de la classe d'âge immédiatement supérieure (26-30 ans), puisque, sur 76 % d'individus répondant aux quatre vagues, seuls 44 % n'ont pas déménagé. En revanche, au-delà de 30 ans, la proportion d'enquêtés qui pourraient être interrogés avec succès quatre vagues de suite en l'absence de suivi en cas de déménagement dépasse 60 %.

Mener des analyses de comportement sur la seule population des jeunes n'ayant pas déménagé implique de généraliser à l'ensemble des jeunes des comportements observés sur la moitié ou un tiers d'entre eux. Cela n'est bien sûr problématique que si les comportements des jeunes qui déménagent diffèrent de ceux qui ne déménagent pas. Or, pour les jeunes en particulier, on peut craindre l'existence de liens directs et indirects entre changements sur le marché du travail et changements de domicile (encadré 3). En effet, chez les moins de 25 ans, le déménagement correspond près de quatre fois sur dix au départ du foyer parental (tableau 5). Et les trajectoires des jeunes quittant leurs parents diffèrent sans doute fortement de celles des jeunes restant chez leurs parents, notamment en matière d'accès à l'emploi. Plusieurs études récentes ont abordé les liens entre les différentes étapes des jeunes vers l'autonomie. Ainsi, la moitié des jeunes des générations 1968 à 1971 ont quitté leurs parents avant 21,1 ans. Un quart les a quittés avant 19,2 ans, et un quart après 23,8 ans. L'âge médian de l'accès au premier emploi (21,2 ans) est très proche de l'âge médian de premier départ de chez les parents. L'accès à un emploi stable et la mise en couple sont eux plus tardifs (respectivement 23,3 ans et 23,4). Et l'âge d'accès au premier logement indépendant (non payé par les parents ou mis gratuitement à disposition par la famille) est intermédiaire (22,5 ans). Le nombre d'années médian entre premier emploi de plus de 6 mois et accès à un logement autonome est proche de zéro) (Galland, 2001)⁹.

⁹ O. Galland (2001) : « Entrer dans la vie adulte : des étapes toujours plus tardives mais resserrées », *Economie et Statistique*, n°337-338. Cette étude exploite les calendriers rétrospectifs des enquêtes Jeunes de 1992 et 1997.

Tableau 5 - Type de déménagement chez les 25 ans et moins

	répartition
déménagement sans éclatement	55
déménagement avec éclatement	45
séparation du conjoint	3
départ d'un enfant	39
départ d'une autre personne	3
total déménagement	100

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : Individus panels de 17 à 25 ans ayant rempli un questionnaire individuel à la vague précédente et qui ont déménagé entre les deux vagues. Moyennes non pondérées sur les vagues 2 à 4.

Pour essayer de mettre en évidence des liens directs entre changements de logement et de situation sur le marché du travail, une première manière de faire consiste à comparer la situation d'emploi des jeunes qui déménagent avant et après leur déménagement. En fait, les déménagements s'accompagnant d'une modification quasi-simultanée de situation sur le marché de l'emploi restent peu fréquents (encadré 3). Ainsi, seuls 5 % des déménagements des 17-20 ans correspondent à une transition des études à l'emploi dans les mois entourant le déménagement. Dans 64 % des cas, les 17-20 ans sont en études deux mois avant comme deux mois après le déménagement, et dans 14 % ils sont en emploi avant comme après. Chez les 21-25 ans, la situation la plus fréquente est l'emploi avant et après le déménagement (deux tiers des déménagements). Dans cette tranche d'âge, seuls 7 % des déménagements s'accompagnent d'accès à l'emploi.

Ces résultats suggèrent que la plupart des déménagements ne seraient pas liés à l'accès à l'emploi, mais pourraient par exemple correspondre à des déménagements avec la famille (ce que montrait déjà l'analyse par type de déménagement). Toutefois, ces résultats ne tiennent compte que des liens *directs et quasi simultanés* entre accès à l'emploi et déménagement. Or, un étudiant peut commencer à travailler tout en résidant chez ses parents, et ne les quitter que 6 mois après, une fois son autonomie financière assurée. Lorsqu'il déménagera, il sera considéré ici comme passant de l'emploi à l'emploi, alors que l'on pourrait considérer que son déménagement comme son accès à l'emploi s'inscrivent dans un processus d'accès à l'autonomie (lien direct mais non simultané entre accès à l'emploi et déménagement) (cf. dossier « jeunes : l'âge des indépendances » dans *Economie et Statistique* 337-338, 2001). Et par ailleurs, ces résultats ne tiennent pas compte de liens *indirects* entre emploi et logement (encadré 3).

Pour tenir compte des liens indirects ou à plus long terme entre accès à l'emploi et déménagement, on peut comparer les taux d'emploi à trois ans (i.e. en septembre 1997) des jeunes initialement en études en septembre 1994 (tableau 6). La proportion de jeunes en emploi est beaucoup plus élevée chez ceux qui ont déménagé que chez les autres. L'erreur commise sur le taux d'accès à l'emploi en n'étudiant que les jeunes n'ayant pas déménagé serait de 6 points pour les 17-20 ans

(19 % au lieu de 25 %), et de 19 points pour les 21-25 ans (35 % au lieu de 54 %). Et, selon la tranche d'âge considérée, ceux qui ont déménagé représentent la moitié ou plus des deux tiers de ceux qui ont un emploi ! Ainsi, étudier les trajectoires d'accès à l'emploi des jeunes sur trois ou quatre ans sans disposer d'information sur ceux qui ont déménagé conduit au minimum à très fortement sous-estimer les taux d'accès à l'emploi, et, au pire, à des estimations de comportements dont on n'est pas assuré qu'elles puissent être généralisées à l'ensemble de la population.

Tableau 6 - Taux d'emploi à trois ans (septembre 1997) en fonction de la situation en septembre 1994 et de l'existence d'un déménagement

Situation en sept. 1994	taux d'emploi à trois ans (septembre 1997)			Part de ceux qui ont déménagé parmi ceux qui ont un emploi
	Ensemble	Personne n'ayant pas déménagé entre les vagues 1 (1994) et 4 (1997)	Personne ayant déménagé entre les vagues 1 (1994) et 4 (1997)	
17-20 ans en étude	25	19 [après redressement 19]	36	49
21-25 ans en étude	54	34 [après redressement 35]	72	71
Chômeurs (tous âges)	49	44 [après redressement 47]	60	42

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panel de 17 ans ou plus ayant rempli un questionnaire individuel aux vagues 1 et 4. Résultats pondérés à l'aide d'une pondération longitudinale redressant de la non-réponse en vague 4, sauf le résultat « après redressement » qui est, lui, pondéré à l'aide d'une pondération redressant également de l'absence de suivi en cas de déménagement (cf. infra).

Lecture : Parmi les 17-20 ans en études en septembre 1994, 25 % occupent un emploi en septembre 1997. Cette proportion est de 19 % parmi ceux qui n'ont pas déménagé, et de 36 % parmi ceux qui ont déménagé, si bien que ces derniers représentent 49 % de ceux qui ont un emploi en septembre 1997

Plus d'un chômeur sur trois déménage en trois ans, et 26% d'entre eux ne sont pas retrouvés

Les personnes qui semblent les moins insérées sont également particulièrement mobiles, et souvent assez difficiles à ré-interroger en cas de déménagement. Comme dans le cas des jeunes, le lien entre déménagement et changement de situation sur le marché du travail peut être direct (par exemple, déménagement vers le lieu où l'on a trouvé un emploi) ou indirect (les moins insérés sur le marché de l'emploi sont aussi ceux dont le logement est le moins stable) .

Ainsi, parmi les chômeurs d'octobre 1994, 37 % vont déménager dans les trois ans (contre 22 % pour l'ensemble de la population de 17 ans ou plus). A peine plus de la moitié de ces chômeurs ayant déménagé seront réinterrogés. En effet, les chômeurs constituent la catégorie dont il est le plus difficile de retrouver l'adresse en cas de déménagement (celle-ci n'est pas retrouvée une fois sur quatre). Les enquêtés d'âge actif (17-60 ans en première vague) qui déménagent ont des trajectoires d'activité initiales plus instables ou moins insérées que les enquêtés qui ne déménagent pas. Ainsi, le taux de déménagement en 3 ans (i.e. entre les vagues 1 et 4) est de 39 % parmi les personnes ayant connu l'emploi et le non-emploi lors de la première année, contre 26 % dans la population d'âge actif dans son ensemble.

Au sein des salariés actifs occupés, 39 % de ceux qui sont en CDD en début de période vont déménager dans les trois ans, contre 24 % pour les actifs employés en CDI. De manière générale, le taux de déménagement est plus élevé chez les salariés (25 %) que chez les indépendants (16 %). On retrouve ce résultat lorsque l'on analyse les taux de déménagement par catégorie socio-professionnelle : les agriculteurs sont beaucoup plus sédentaires que les autres actifs (11% seulement déménagent), les artisans également mais dans une moindre mesure (20%), tandis que les salariés ont des taux de déménagement plutôt supérieurs à la moyenne, variant de 24 % pour les ouvriers à 27 % pour les cadres.

Même si les écarts de taux de déménagement et de taux de réponse en fonction de la situation de la personne sur le marché de l'emploi sont moins marqués que les écarts en fonction de l'âge, ils restent forts¹⁰. De ce fait, les risques de biais liés à l'absence de suivi des personnes qui déménagent, moins importants que ceux liés à l'âge, restent notables. Ainsi, parmi les chômeurs de septembre 1994, 44 % de ceux qui n'ont pas déménagé ont un emploi trois ans plus tard, contre 49 % de l'ensemble. Si l'on cherche à estimer le taux d'emploi à trois ans de l'ensemble des chômeurs à partir des seuls chômeurs n'ayant pas déménagé (en les repondérant pour les rendre représentatifs de l'ensemble), on aboutit à 47 %, soit une valeur peu différente de la valeur estimée sur l'ensemble des chômeurs (49 %) (tableau 6). Mais il n'en reste pas moins que 42 % des chômeurs qui ont un emploi en fin de période ont déménagé pendant les trois années étudiées.

¹⁰ Des premiers résultats de régressions logistiques confirment que l'âge est effectivement un facteur explicatif de premier ordre, mais que, à âge donné, la situation sur le marché de l'emploi pourrait avoir un impact moins important que le laissent suggérer les statistiques descriptives : ainsi, si les personnes les moins insérées sur le marché de l'emploi déménagent plus, c'est avant tout parce elles appartiennent aux âges où l'on franchit les différentes étapes menant à l'indépendance.

Ainsi, les résultats ci-dessus suggèrent

1. que les *statistiques descriptives* que l'on peut tirer d'échantillons de personnes n'ayant pas déménagé peuvent ne pas être trop éloignées des statistiques que l'on obtient sur des panels d'individus, si l'une des trois conditions suivantes est satisfaite :
 - soit les comportements de ceux qui ont déménagé et des autres sont proches,
 - soit les taux de déménagement pour la population concernée ne sont pas trop importants (population peu mobile, ou période d'observation courte¹¹),
 - soit le redressement par post-stratification permet de rapprocher ces deux estimations, ce qui n'est possible que si les différences entre ceux qui déménagent et les autres s'expliquent bien en fonction de caractéristiques observables prises en compte dans le calcul de la pondération ;
2. que, même si, grâce à l'usage d'une pondération adaptée, les statistiques descriptives sont peu faussées, se limiter aux personnes n'ayant pas déménagé *réduit fortement l'échantillon* d'étude quand les taux de déménagement sont élevés, et réduit encore plus fortement l'échantillon des personnes qui connaissent l'événement étudié si celui-ci est plus fréquent parmi ceux qui déménagent. Cela peut rendre délicate la généralisation à l'ensemble de la population de comportements observés sur ceux qui ne déménagent pas, et ce, d'autant plus que ceux qui déménagent diffèrent des autres, en particulier si ces différences portent sur des *caractéristiques non observables* (et donc non susceptibles d'être redressées). Cette étude traite peu des problèmes posés par les biais liés à des caractéristiques inobservables. Toutefois, si on considère une variable future comme inobservable (ou comme révélatrice d'une variable inobservable à la date courante), l'exemple des taux d'emploi à trois ans met en évidence des différences importantes parmi ceux qui déménagent et les autres¹².

Une autre manière d'essayer d'apprécier l'impact potentiel de variables difficilement observables est d'étudier le taux de déménagement futur en réponse à des questions d'opinion. Un déménagement futur est d'autant plus probable que la personne est peu satisfaite de son logement : le taux de déménagement dans les

¹¹ La période à laquelle il est fait référence ici est la période pendant laquelle la personne ne « doit » pas déménager. Cette période correspond donc à l'intervalle de temps entre la première et la dernière vague d'enquête, et non à la longueur du calendrier d'activité exploité.

¹² Malheureusement, dans cet exemple, seul l'âge est contrôlé, et non les autres caractéristiques observables. D'autres analyses plus complètes des biais introduits par des caractéristiques non observables peuvent être envisagées, mais n'ont pas été menées ici faute de temps. Par exemple, on pourrait comparer dans une équation de salaire les résidus obtenus pour les individus qui vont déménager et pour les autres.

trois ans dépasse 50 % pour ceux qui ne sont « pas satisfaits du tout » de leur logement en début de période, et décroît jusqu'à 16 % pour les personnes qui sont « très satisfaites » de leur logement. Une analyse rapide « toutes choses égales par ailleurs » confirme que cet effet reste très significatif à âge, sexe, occupation et catégorie sociale donnés. Les autres questions d'opinion ont un effet plus modéré sur la probabilité de déménagement, et cet effet est souvent plus corrélé aux déterminants socio-économiques usuels (donc, peut être plus facilement corrigé). Ainsi, un tiers des personnes se déclarant « pas satisfaites du tout de leur travail ou situation principale » vont déménager dans les trois ans, contre environ un cinquième de celles qui s'en déclarent « assez satisfaites », « satisfaites » ou « très satisfaites ». Toutefois, à âge et situation sur le marché de l'emploi donnés, la satisfaction vis-à-vis de l'emploi n'apparaît plus liée de manière significative au déménagement. Il en est de même, pour les actifs occupés, de la satisfaction vis-à-vis de la distance à leur lieu de travail.

Le déménagement est plus fréquent aux bas niveaux de vie, dans les familles monoparentales et chez les locataires.

Compte tenu de l'importance du taux de déménagement et des taux de réponse pour déterminer les risques de biais et de diminution de la taille d'échantillon, il est important d'avoir une idée des variations de ces indicateurs en fonction des caractéristiques des personnes et de leurs ménages (tableaux 4). L'objet ici n'est pas de fournir un modèle explicatif de la probabilité de déménagement en fonction de caractéristiques individuelles. En effet, le déménagement dépend vraisemblablement largement d'événements familiaux (départ de chez les parents, mise en couple, arrivée d'un enfant, divorce...) qui ne sont pas évoqués dans ce texte. Le but est plutôt de fournir des indications sur le type de populations que l'on suit moins bien dans des panels de logements, afin de permettre au lecteur souhaitant réaliser une étude sur ce type de données d'apprécier les risques de biais pour les populations qui l'intéressent. Nous nous limitons donc ici au commentaire de statistiques descriptives (s'appuyant sur les caractéristiques des personnes et des ménages en début de période, c'est-à-dire en 1994).

Ce sont les individus appartenant à des ménages dont le revenu mensuel du ménage par unité de consommation se situe dans le 1er quartile (soit moins de 4400 F par mois) qui ont la plus forte probabilité de déménager dans les trois ans (24 %). Il semblerait que la probabilité de déménagement, ainsi que la difficulté à retrouver les personnes, soient encore plus importantes pour le premier décile de revenu. Les locataires sont également plus mobiles, et ils le sont d'autant plus que le propriétaire est un propriétaire privé (50 % de taux de déménagement). A l'opposé quand le propriétaire est un membre de la famille le taux de déménagement est nettement plus faible (23 %). Enfin, ce taux est de 30 % lorsque que le propriétaire est un organisme public.

Les personnes appartenant à des familles monoparentales sont celles qui ont la probabilité la plus forte de déménager, soit 32 %, contre 26 % pour les personnes seules célibataires. Ces déménagements correspondent dans environ la moitié des cas

au départ d'un enfant, et dans l'autre moitié, à un déménagement de l'ensemble de la famille monoparentale. A l'inverse, ce sont les couples sans enfant qui sont les plus stables (taux de déménagement de 19 %).

C'est dans les unités urbaines de plus de 100 000 habitants ainsi que dans l'unité urbaine de Paris que l'on observe la plus forte proportion de déménagements (respectivement 27 % et 25 %). C'est également sur Paris que la part des non retrouvés est la plus forte (ils représentent un quart des cas de déménagement). A l'opposé, les communes rurales sont celles qui connaissent le moins de déménagements (14 %). Si on regarde les individus ayant déménagé suivant la Zone d'Etudes et d'Aménagement du Territoire (ZEAT) de résidence initiale, la Méditerranée, l'Île de France puis la région Rhône-Alpes ont des taux de déménagement supérieur aux autres régions. C'est également dans les ZEAT à plus fort taux de déménagement qu'il est le plus difficile de retrouver les personnes qui déménagent (un quart d'entre elles ne sont pas retrouvées en Île-de-France ou en Méditerranée, contre un septième environ dans l'ensemble de la France).

En revanche, la nationalité ainsi que le sexe de l'individu n'ont pas d'effet significatif sur la probabilité de déménager.

1.3. Dans quelle mesure la pondération permet-elle de compenser la non-représentativité des individus suivis ?

Si suivre les personnes en cas de déménagement limite les risques de biais, cela ne résout pas tous les problèmes de « représentativité » de la population suivie par rapport à la population dans le champ. En effet, les personnes qui répondent à une vague donnée ne sont pas « représentatives » des personnes interrogées à la vague précédente, car l'attrition d'une vague sur l'autre est sélective : parmi les personnes répondant à une vague, les personnes les moins bien insérées sont plus susceptibles d'être non-répondantes à la vague suivante. Ainsi, 79 % des personnes interrogées lors de la première vague qui ont été en emploi toute l'année précédant la collecte répondent également aux trois vagues suivantes, tandis qu'elles ne sont que 62 % parmi celles qui ont connu l'emploi puis le chômage l'année précédant la première collecte (tableau 7c). Dans cette partie, nous nous intéressons à l'importance des biais liés à cette sélectivité de l'attrition, et examinons dans quelle mesure l'utilisation d'une pondération (longitudinale) adaptée permet de les réduire¹³.

¹³ Notons que le calage (longitudinal) dans un panel se distingue d'un calage transversal. En gros, le principe du calage longitudinal est de rendre les répondants de la vague n+i représentatifs des répondants de la vague n. Ce calage peut donc s'appuyer sur les informations recueillies dans la vague n, et éventuellement sur des informations extérieures (structure de la population en n+i). A l'inverse, pour une enquête transversale, le calage ne peut s'appuyer que sur les quelques informations contenues dans la base de sondage (par exemple, le type de logement) et sur des informations extérieures (par exemple, la distribution par âge et sexe de la population à l'enquête Emploi). Les possibilités de redressement sont donc plus limitées dans une opération transversale.

Sont particulièrement sur-représentés parmi les personnes interrogées en première vague qui ne seront pas réinterrogées aux trois vagues suivantes (tableaux 7) :

- les sans-diplômes (29 % de ceux qui ne seront pas réinterrogés, contre 25 % de ceux qui seront réinterrogés),
- les chômeurs (11% contre 6%), et, dans une moindre mesure, les retraités (23 % contre 21 %),
- les 21-25 ans (12 % contre 9 %) et les plus de 61 ans (26 % contre 22 %),
- parmi les actifs, les artisans (10 % contre 7 %), ouvriers (27 % contre 24 %) et les employés (31 % contre 29 %),
- parmi les salariés, ceux qui sont en Contrat à Durée Déterminée (13 % contre 10 %),
- les personnes seules (16 % contre 13 %) et les couples sans enfant (28 % contre 26 %),
- les habitants d'Ile de France (23 % contre 18 %),
- les locataires (42 % contre 38 %),
- les 25 % les plus pauvres (30 % contre 23 %),
- les personnes ayant alterné entre emploi et non-emploi au cours des douze derniers mois (17 % contre 12 %), et celles n'ayant pas eu d'emploi (33 % contre 27 %).

Le biais sur les statistiques descriptives (non pondérées) dû à l'attrition dépend, avant pondération, de deux facteurs : de la part des personnes réinterrogées et de l'importance des différences entre les personnes réinterrogées et non-réinterrogées. A condition d'essayer de suivre les personnes qui déménagent et de s'intéresser à des populations pas trop spécifiques, les personnes non réinterrogées restent en général minoritaires (elles représentent 24 % de l'ensemble des enquêtés d'âge actif dans le cas d'un suivi sur 4 vagues). De plus, les différences (observables) entre personnes réinterrogées et personnes non réinterrogées restent encore modérées dans le cas d'un suivi sur quatre ans (cf. ci dessous). Au total, ces différences sont suffisamment peu marquées en population générale pour que la sélectivité de l'attrition ne se traduise pas par des biais importants sur les statistiques descriptives. Ainsi, sur l'exemple de la trajectoire sur 12 mois avant la première collecte des personnes d'âge actif, 61 % de celles qui seront réinterrogées ont été en emploi tous les mois, contre 51 % de celles que ne seront pas réinterrogées. Comme les secondes représentent 24 % du champ et les premières 76 %, la proportion de personnes en

emploi sur douze mois sur l'ensemble de la population est de 58 %¹⁴, valeur peu éloignée du 61 % mesuré sur les seules personnes réinterrogées.

Mais bien sûr, le biais lié à l'attrition pourrait devenir important dans le cadre d'une étude centrée sur une sous-population mal suivie (par exemple les chômeurs) pour des variables dont la moyenne serait très différente sur l'échantillon suivi et l'échantillon non suivi. Par ailleurs, en général, les caractéristiques associées à une plus forte probabilité de déménagement sont également celles qui se traduisent par une moindre probabilité de réinterrogation en l'absence de déménagement. L'absence de suivi en cas de déménagement accentue donc les biais inhérents à l'utilisation d'un panel.

Le recours à une pondération adaptée (pondération longitudinale, encadré 3) permet de corriger en partie les biais liés à l'attrition, tout du moins pour la part de ces biais qui s'explique en fonction de caractéristiques observées dans les vagues d'enquêtes précédentes¹⁵. Ainsi, l'estimation de la proportion de personnes toujours en emploi l'année précédant la première vague effectuée à partir des répondants aux quatre vagues (60 %) se rapproche de la valeur obtenue sur les répondants de première vague (61 %) (tableau 8). Ce résultat n'est pas mécanique, dans la mesure où la procédure de redressement longitudinal n'utilise pas d'information sur la situation des individus par rapport à l'emploi pour le calage.

¹⁴ soit $76\% \cdot 61\% + 24\% \cdot 51\%$.

¹⁵ ou éventuellement uniquement dans la vague courante si l'on effectue un calage sur des marges connues à l'aide d'autres enquêtes.

**Tableau 7 : Représentation de différents sous-échantillons
7a : en fonction de caractéristiques observables des personnes en 1994.**

Caractéristiques individuelles	Ensemble			Pas de déménagement			Déménagement				Taux de déménagement	
	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Non retrouvés	Autres non réponses		Part des non-retrouvés dans les non réponses
Niveau d'études atteint												
Aucun diplôme	26	25	29	30	29	35	11	10	14	15	48	9
CAP, BEP	20	20	21	20	20	20	20	18	27	21	55	22
BEPC	26	26	25	26	26	25	26	27	27	25	51	23
Baccalauréat	7	7	6	6	6	6	9	9	6	10	37	29
Baccalauréat plus 2 années	12	12	10	10	11	7	18	19	15	17	45	34
Supérieur à bac plus 2	10	10	8	8	8	7	16	18	12	12	47	36
Total des individus	100	100	100	100	100	100	100	100	100	100	49	22
Occupation												
Actif occupé	51	53	45	49	51	43	56	59	48	52	47	34
Chômeur	8	6	11	6	5	9	13	11	21	12	62	37
Étudiant	9	9	10	7	7	7	16	16	17	17	50	41
Retraité	22	21	23	26	25	29	7	6	6	11	32	7
Autre inactif	11	11	12	12	12	13	8	8	9	8	50	16
Total des individus	100	100	100	100	100	100	100	100	100	100	49	22
Âge												
17-20 ans	7	7	8	6	6	6	12	11	13	14	47	36
21-25 ans	9	9	12	5	5	7	23	23	26	22	53	55
26-30 ans	10	10	9	7	7	7	19	21	19	13	59	43
31-40 ans	19	20	16	19	20	14	22	22	21	20	50	25
41-50 ans	19	20	17	21	22	19	10	9	10	11	45	11
51-60 ans	13	13	13	15	14	15	7	7	5	8	38	12
Plus de 61 ans	23	22	26	27	26	32	8	7	6	12	32	7
Total des individus	100	100	100	100	100	100	100	100	100	100	49	22

Tableau 7a (suite) : en fonction de caractéristiques observables des personnes en 1994.

Caractéristiques individuelles	Ensemble		Pas de déménagement		Déménagement			Part des non-retrouvés dans les non réponses	Taux de déménagement
	Ensemble	Répondants V2, V3 et V4	Ensemble	Répondants V2, V3 et V4	Ensemble	Répondants V2, V3 et V4	Non retrouvés		
Statut des actifs									
Salarié	88	86	86	87	92	92	90	89	47
Indépendant	13	14	14	13	9	8	10	11	45
Total des actifs	100	100	100	100	100	100	100	100	46
PCS des actifs									
Agriculteur	4	4	5	5	2	2	0	2	0
Artisan	7	7	8	7	6	5	11	8	53
Cadre supérieur	13	13	12	12	14	15	9	13	38
Profession intermédiaire	22	23	22	23	24	27	16	22	38
Employé	29	29	29	28	30	30	33	27	52
Ouvrier	25	24	25	25	25	23	31	28	49
Total des actifs	100	100	100	100	100	100	100	100	46
Type de contrat de travail des salariés									
CDD	10	10	9	8	16	15	18	20	44
CDI	90	90	92	92	84	85	82	80	47
Total des salariés	100	100	100	100	100	100	100	100	46

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panel de 17 ans ou plus ayant rempli un questionnaire individuel en 1994. Individus hors champ en vagues 2, 3 ou 4 exclus. Caractéristiques de 1994.

7b : en fonction de caractéristiques observables des ménages en 1994.

Caractéristiques des ménages	Ensemble		Pas de déménagement		Déménagement			Part des non-retrouvés dans les non réponses	Taux de déménagement
	Ensemble	Répondants V2, V3 et V4	Ensemble	Répondants V2, V3 et V4	Ensemble	Répondants V2, V3 et V4	Non retrouvés		
Type de ménage									
Personne seule	14	13	13	13	15	15	16	19	45
Couple sans enfant	26	26	28	27	31	23	23	19	55
Couple avec enfant	42	44	42	44	36	43	46	40	47
Famille monoparentale	4	4	3	3	4	5	5	7	41
Autre type de ménage	14	14	15	14	15	13	12	14	51
Total	100	100	100	100	100	100	100	100	49
ZEAT									
Ile de France	19	18	19	18	22	22	20	20	60
Bassin parisien	17	17	18	18	17	15	16	14	48
Nord	7	7	7	7	8	6	6	8	30
Est	9	9	9	9	10	8	9	6	42
Ouest	13	14	14	14	12	13	13	11	12
Sud-ouest	10	11	11	11	10	9	10	11	30
Centre-est	12	12	12	12	11	13	14	7	14
Méditerranée	13	12	12	12	11	15	12	24	15
Total	100	100	100	100	100	100	100	100	49
Strate									
Commune rurale	27	28	29	30	26	17	18	18	42
Unité urbaine < 20000 hab	16	17	16	16	16	17	13	18	40
20000 < u.u. <= 100000 hab	13	13	13	12	15	14	15	10	35
u.u. de plus de 100000 hab	27	27	26	26	26	33	35	30	53
Commune de l'unité urbaine de Paris	17	16	16	16	17	19	17	17	61
Total	100	100	100	100	100	100	100	100	49

Tableau 7b (suite) : en fonction de caractéristiques observables des ménages en 1994.

Caractéristiques des ménages	Ensemble			Pas de déménagement			Déménagement				Part des non-retrouvés dans les non réponses	Taux de démenagement
	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Non retrouvés	Autres non réponses		
Statut d'occupation du logement												
Locataire, sous locataire et logé gratuitement)	39	38	42	30	30	33	69	70	70	64	51	39
Propriétaire	61	63	58	70	71	67	31	30	30	36	44	11
Total	100	100	100	100	100	100	100	100	100	100	49	22
Propriétaire (pour locataires)												
Famille	14	14	12	17	17	16	8	8	6	10	39	23
Propriétaire privé	46	46	47	37	38	37	60	60	67	55	56	50
HLM ou organisme public	41	41	40	46	46	47	32	32	27	36	43	30
Total	100	100	100	100	100	100	100	100	100	100	51	38
Revenu mensuel / uc (quartiles)												
Revenu ≤ 4400 F	25	23	30	24	23	30	27	25	37	24	59	24
4400 F < revenu ≤ 6500 F	25	25	25	25	25	26	23	23	23	25	46	21
6500 F < revenu ≤ 9330 F	25	27	21	25	27	20	25	26	22	24	46	22
9330 F < revenu	25	26	24	25	25	24	25	27	18	27	39	22
Total	100	100	100	100	100	100	100	100	100	100	49	22

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panel de 17 ans ou plus ayant rempli un questionnaire individuel en 1994, Individus hors champ en vagues 2, 3 ou 4 exclus. Caractéristiques de 1994.

Tableau 7c : en fonction des trajectoires d'emploi sur les douze derniers mois avant la première collecte.

Trajectoire d'emploi entre nov. 93 et oct. 94 (12 mois)	Ensemble			Pas de déménagement			Déménagement				Taux de déménagement	
	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Autres	Ensemble	Répondants V2, V3 et V4	Non retrouvés	Autres non réponses		Part des non-retrouvés dans les non-réponses
Emploi tous les mois	58	61	51	61	63	55	50	53	37	48	43	22
tous les mois à temps complet	44	46	38	46	47	41	39	43	27	37	43	24
au moins un mois en temps partiel	7	7	5	7	8	5	6	6	4	5	47	21
au moins un mois en indépendant	8	8	8	9	9	9	5	4	5	6	46	16
Emploi et non emploi	13	12	17	11	10	13	19	18	25	21	54	39
emploi puis chômage	2	2	3	2	2	2	3	2	6	5	57	41
chômage puis emploi	2	2	3	2	2	2	3	3	3	4	42	37
alternances d'emploi et de chômage	3	3	4	2	2	3	5	4	7	4	65	41
activité et inactivité (+ chômage)	6	5	7	5	5	6	9	9	9	9	50	39
Jamais d'emploi	29	27	33	28	27	31	31	29	39	31	56	29
Ensemble (âge actif)	100	100	100	100	100	100	100	100	100	100	50	26

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panel de 17 ans à 60 ans répondants en 1994. Individus hors champ en vagues 2, 3 ou 4 exclus. Données pondérées (pondérations de la vague 1)

Tableau 8 : Biais lié à la sélectivité de l'attrition et effet correcteur du redressement – exemple de la proportion d'enquêtés ayant toujours été en emploi et de la proportion d'enquêtés ayant connu l'emploi et le non-emploi.

8a : Proportion d'enquêtés ayant toujours été en emploi entre novembre 93 et octobre 94.

Echantillon	Type de pondération	Proportion / Biais	Ensemble	Pas de déménagement	Déménagement entre V1 et V4
Répondants de 1994	Transversale 1994	Proportion	58,4%	61,4%	49,5%
Répondants de 1994 à 1997	Transversale 1994	Proportion	60,8%	63,2%	52,9%
		Biais	2,4%	1,8%	3,4%
Répondants de 1994 et 1997	Longitudinale	Proportion	59,9%	63,0%	51,0%
		Biais	1,5%	1,6%	1,5%
Répondants de 1994 et 1997 n'ayant pas déménagé	Longitudinale	Proportion	59,3%	62,1%	50,8%
		Biais	0,9%	0,7%	1,3%
		Proportion	60,5%		
		Biais	2,1%		

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panel de 17 ans à 60 ans répondants en 1994. Individus ayant rempli un questionnaire individuel en vague 1 et qui ne sont pas hors champ aux vagues 2, 3 et 4. Données pondérées.

8b : Proportion d'enquêtés ayant connu l'emploi et le non-emploi entre novembre 93 et octobre 94.

Echantillon	Type de pondération	Proportion / Biais	Ensemble	Pas de déménagement	Déménagement entre V1 et V4
Répondants de 1994	Transversale 1994	Proportion	13,0%	10,7%	19,4%
Répondants de 1994 à 1997	Transversale 1994	Proportion	11,7%	9,9%	17,8%
		Biais	-1,3%	-0,8%	-1,6%
Répondants de 1994 et 1997	Longitudinale	Proportion	12,2%	10,0%	18,7%
		Biais	-0,8%	-0,7%	-0,7%
Répondants de 1994 et 1997 n'ayant pas déménagé	Longitudinale	Proportion	12,4%	10,3%	18,7%
		Biais	-0,6%	-0,4%	-0,7%
Répondants de 1994 et 1997 n'ayant pas déménagé	Longitudinale	Proportion	11,4%		
		Biais	-1,6%		

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panel de 17 ans à 60 ans répondants en 1994. Individus ayant rempli un questionnaire individuel en vague et qui ne sont pas hors champ aux vagues 2, 3 et 4. Données pondérées.

1.4 La pondération permettrait-elle de compenser l'absence de suivi des individus qui déménagent ?

Sur quatre vagues, les biais liés à l'attrition semblent modérés, et peuvent être réduits grâce à une pondération adaptée. Bien que les biais liés à l'absence de suivi en cas de déménagement soient, eux, plus importants, on peut se demander s'il ne serait pas également possible de les corriger par re-pondération. C'est l'exercice auquel nous nous livrons maintenant. Pour ce, les individus qui ont déménagé entre les vagues 1 à 4 sont considérés comme non-répondants, ce qui porte le taux de non-réponse à 38 % (non-réponse « au sein du même logement »), contre 22 % si ceux qui déménagent ne sont pas considérés comme non répondants¹⁶. Considérer ceux qui ont déménagé comme non-répondants divise donc le taux de réponse par 1,2. Les poids des « répondants au sein du même logement » de la vague 4 doivent donc en moyenne être multipliés par 1,2 par rapport aux poids obtenus lorsque les personnes qui déménagent sont suivies, et par 1,6 par rapport aux poids de première vague. En fait, les taux de « non-réponse au sein du même logement » ne sont pas homogènes dans la population. Selon les caractéristiques des personnes, ils s'échelonnent entre 15 et 82 %. Parmi les 55 catégories homogènes vis-à-vis de la non-réponse au sein du même logement qui ont été définies (encadré 3), quatre ont un taux de non-réponse au sein du même logement égal ou supérieur à 80 %, et 20 ont un taux de non-réponse qui dépasse 50 %. Trois des quatre catégories dont le taux de non réponse est le plus élevé correspondent à des individus de 25 ans à 35 ans. Comme les taux de non-réponse au sein du même logement sont très dispersés, le facteur correctif appliqué aux pondérations est très variable d'une catégorie à l'autre : les poids initiaux sont multipliés au moins par 1,2 et au plus par 5,6 (par 1,09 au moins et 2,2 lorsque ceux qui déménagent sont suivis). Les poids obtenus sont donc beaucoup plus dispersés que lorsque ceux qui déménagent sont suivis (tableau 9) : ils varient dans un rapport de 1 à 22 (contre un rapport de 1 à 9 en cas de suivi lors des déménagements). Un seul individu peut alors fortement peser dans les résultats, ce qui peut poser des problèmes s'il est atypique (ou si les informations le concernant sont affectées d'erreurs de mesure).

¹⁶ Taux de non-réponse calculés sur les enquêtés de première vague de 17 ans et plus qui ne sont pas hors champ lors des vagues suivantes.

Tableau 9 : Caractéristiques des poids longitudinaux en vague 4 des individus panel

	Pondération valable pour l'ensemble des répondants	Pondération valable pour les répondants n'ayant pas déménagé
Moyenne	3 961	4 945
Ecart type	1 326	2 243
D1	2 824	3 152
Q1	3 176	3 564
Médiane	3 693	4 359
Q3	4 402	5 526
D9	5 226	7 363
C95	5 931	9 020
C99	9 892	13 348
Minimum	2 195	2 265
Maximum	19 485	49 200

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panel de 17 ans et plus en 1997 ayant rempli un questionnaire individuel en vague 4.

Lecture : dans le cas où les individus sont suivis lorsqu'ils déménagent, un individu panel remplissant un questionnaire individuel en vague 4 représente en moyenne 3 961 personnes dans la réalité. En l'absence de suivi en cas de déménagement, un individu panel « répondant au sein du même logement » en vague 4 représente 4 945 personnes dans la réalité, soit 1,2 fois plus.

C'est pour les 21-30 ans que l'absence de suivi en cas de déménagement aurait l'impact le plus fort sur les poids. En effet, dans ce groupe d'âge, l'absence de suivi en cas de déménagement conduirait en moyenne à presque doubler les poids par rapport aux poids calculés en cas de suivi. Pour les 17-20 ans, l'incidence du suivi sur les poids resterait assez importante (poids 1,5 fois plus élevés que les poids calculés en cas de suivi). Elle serait en revanche relativement faible pour les plus de 30 ans (multiplication par 1,1 à 1,3).

S'il est en théorie possible de redresser les données pour corriger de l'absence de suivi des personnes qui déménagent (ainsi que de l'attrition), cela ne permet évidemment pas de compenser les pertes d'effectifs dans certaines catégories : les poids calculés sont très dispersés, et peuvent être assez élevés. Par ailleurs, une telle pondération, basée sur le redressement suivant les caractéristiques socio-démographiques usuelles, ne redresse qu'imparfaitement les biais. Ainsi, le taux de personnes qui ont été continûment en emploi l'année précédant la première vague est estimée après redressement à 60,5 % (à partir de l'échantillon des répondants au sein du ménage), soit une valeur plus proche de la valeur non-redressée (60,8 %) que de la vraie valeur (58,4 %) (tableau 8). L'estimation de la proportion de personnes qui ont connu l'emploi et le non-emploi souffre des mêmes défauts (estimation après redressement de 11,4 %, plus proche de la valeur non redressée (11,7 %) que de la vraie valeur (13 %)). Le redressement, s'il permet de corriger des biais s'expliquant bien en fonction des caractéristiques observables, et notamment les « effets de structure », atteint ses limites lorsqu'il s'agit de tenir compte de biais qui ne sont pas fortement corrélés aux caractéristiques socio-économiques usuelles.

2. Quel est le prix à payer pour suivre les personnes qui déménagent ?

2.1 Un coût par interview réalisé plus élevé

Un panel d'individus est plus coûteux qu'un panel de logements. En effet, en premier lieu, un effort supplémentaire doit être fait pour suivre les individus, et éventuellement pour assurer le transfert de leur dossier d'enquête à une direction régionale de l'Insee et / ou à un autre enquêteur. L'interrogation des personnes retrouvées peut également s'avérer plus coûteuse. C'est le cas par exemple si elles ont déménagé dans une zone hors échantillon maître de l'Insee (d'où des coûts de transports plus élevés pour les enquêteurs ainsi que des pertes de temps et des coûts de repérage des nouvelles adresses). Par ailleurs, il est plus fréquent que les personnes qui déménagent ne soient pas enquêtées (soit parce qu'elles ne sont pas retrouvées, soit parce qu'elles sont en fait hors champ, décédées...), d'où un coût moyen par interview réalisé plus élevé. Enfin, les panels d'individus imposent en général une charge plus lourde aux enquêtés, car ceux-ci sont en moyenne suivis plus longtemps que dans un panel de logements (parce que le nombre de vagues est plus élevé ou que les individus ne sortent pas de l'enquête par déménagement). Cette charge implique des efforts particuliers pour s'assurer la coopération des enquêtés (cadeaux, courriers, diffusion de résultats...). Elle pourrait également se traduire par des refus supplémentaires (entraînant des coûts supérieurs par interview réalisée), sauf si le fait que le même enquêteur retrouve le même enquêté d'une année sur l'autre facilite la coopération de ce dernier ou si le fait que certaines variables (état civil) ne soient pas redemandées à chaque fois permet d'alléger notablement la charge d'enquête (Caron et Ravalet, 2000¹⁷).

A budget donné, l'échantillon d'un panel d'individus sera donc *a priori* de taille plus réduite que celui d'un panel de logements. 7 344 ménages ont répondu à la première vague du panel européen, et des questionnaires individuels ont été remplis pour environ 14 000 individus de 17 ans et plus. Est-il alors possible d'estimer valablement sur la base d'un échantillon de cette taille la fréquence d'événements relativement rares, comme les transitions sur le marché de l'emploi ? C'est la question à laquelle nous nous intéressons maintenant.

¹⁷ Caron, N. et Ph. Ravalet (2000) : « Estimation dans les enquêtes répétées : application à l'enquête Emploi en continu », Insee, document de travail de l'Unité Méthodes Statistiques n°0005. Toutefois, en pratique, l'état civil est vérifié à chaque enquête, ce qui limite le gain en temps de questionnement.

2.2 Construction des calendriers d'activité à partir du panel, et qualité des données

Toutefois, avant de pouvoir comparer les résultats tirés du calendrier d'activité du panel et ceux que l'on observe sur les enquêtes Emploi (panel de logements de taille plus importante), il convient de traiter certains problèmes spécifiques au calendrier du panel, comme les incohérences que l'on observe sur les recouvrements de calendrier entre les différentes vagues ou les cas où plusieurs occupations peuvent être déclarées pour un même mois (encadré 4).

Si l'impact de ces problèmes est faible si l'on s'intéresse à une situation un mois donné sur le marché du travail, ils peuvent être plus importants si l'on s'intéresse aux transitions. En effet, les transitions restent peu fréquentes un mois donné : à peine plus de 1 % par mois (cf. ci-dessous), ce qui représente dans le panel moins de 200 observations par mois. Ainsi, un problème de déclaration, de définition ou de qualité des données qui conduirait à générer 20 transitions supplémentaires par mois augmenterait de 10 % la probabilité de transition.

Compte tenu de l'importance potentielle de ces problèmes pour l'analyse des transitions, il nous a semblé important de commencer notre travail par une évaluation de la qualité des données, et éventuellement une correction de ces dernières. L'impact des corrections menées dépend des variables étudiées. Ainsi :

- L'apurement a conduit à modifier moins de 1 % des mois renseignés, mais a surtout conduit à renseigner environ 3,5 % des mois en plus (notamment les mois pour lesquels seule une activité « non principale » (activité secondaire, maladie, chômage partiel, formation...) était renseignée, ou les mois correspondant à des trous de calendrier). Au total, 10 % des individus sont concernés.
- Le choix des règles de définition de la situation principale du mois introduit un léger biais en faveur de l'activité, puisqu'une transition vers le chômage inférieure à un mois n'est pas prise en compte. Si l'on avait choisi de privilégier le chômage sur l'emploi puis sur l'inactivité, seuls 0,3 % des mois auraient été modifiés, mais 5,4 % des enquêtés de 17 ans et plus auraient été concernés. Seul un mois aurait été modifié pour la moitié de ces derniers (2,8 %). Pour 2,3 % des individus (soit 373 individus), ce choix aurait entraîné une modification du nombre de transitions (celui-ci aurait varié d'une unité dans 0,8 % des cas, de deux unités dans 1,1 % des cas). La séquence des transitions (suite des activités exercées) est également modifiée dans ce cas. Toutefois, cela n'implique une modification de la trajectoire sur 48 mois (selon la typologie du tableau 10) que dans 0,8 % des cas (soit 80 individus parmi ceux qui sont présents sur les quatre vagues) : si l'on privilégie le chômage sur l'emploi, plus de la moitié de ces individus ne sont plus continûment en emploi sur 48 mois.

Tableau 10 : Incidence de la règle de définition de l'activité principale et de l'apurement sur la définition des trajectoires sur 48 mois

Règle de priorité pour déterminer l'activité principale :	OUI		NON
	Emploi domine chômage qui domine inactivité	Chômage domine emploi qui domine inactivité	
Prise en compte des corrections ponctuelles :			
Emploi tous les mois	52,8	52,3	57
tous les mois à temps complet	38,6	38,2	42
au moins un mois en temps partiel	7,2	7,1	8
au moins un mois en indépendant	7,1	7,0	8
Emploi et non emploi	30,5	30,8	31
emploi puis chômage	1,8	2,0	2
chômage puis emploi	1,1	1,2	1
alternances d'emploi et de chômage	10,1	10,2	11
activité et inactivité (+ chômage)	17,5	17,3	17
Jamais d'emploi	16,6	16,9	11
Trajectoire à trous	0,0	0,0	1
Total	100	100	100

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee
 Champ : individus panel de 17 à 60 ans ayant rempli un questionnaire individuel en 1994, 1995, 1996 et 1997. Données pondérées.

Au total, l'absence de hiérarchisation entre les différentes occupations possibles (les situations d'emploi étant placées sur le même plan que la maladie, la formation, ou le chômage partiel) comme la variété des choix offerts - mais pas toujours compris - se sont traduits par des calendriers parfois difficilement exploitables sans travail supplémentaire, car ne permettant pas de déterminer facilement si, un mois donné, la personne a été active occupée, chômeuse ou inactive. Ces problèmes devraient en partie disparaître à partir de la vague 5, grâce à la généralisation de la Collecte Assistée Par Informatique (qui réduira par ailleurs les erreurs de saisie). En effet, dans le questionnaire informatique, l'enquêté doit décrire les différentes périodes consécutives qu'il a connues sur le marché de l'emploi. Ces différentes périodes ne portent que sur les principales situations proposées sur les questionnaires papier (et excluent notamment la maladie, les autres formations...). Ce nouveau mode d'interrogation assure qu'aucun mois ne puisse être vide (toute période doit débiter à la fin de la précédente) et permet de déterminer sans ambiguïté quand l'individu a été actif occupé, chômeur ou inactif¹⁸. Au sein de chacune de ces périodes, l'enquêté doit ensuite préciser le nombre de jours de maladie, de chômage partiel ou de formation qu'il a connus, ce qui permet de récupérer une information proche de celle recueillie avec les questionnaires papier. Enfin, lors de l'enquête, une incohérence entre les déclarations d'une vague et celles de la vague précédente peut être repérée informatiquement dès la collecte, ce qui donne à l'enquêteur la possibilité de la faire corriger à l'enquêté. Avec le passage à CAPI et l'expérience de plus en plus grande des enquêteurs, la qualité des calendriers d'activité du panel devrait donc notablement s'améliorer à partir de la vague 5.

¹⁸ Il permet par ailleurs de connaître exactement la durée en jours de chaque période et d'observer toutes les transitions, même s'il y en a plusieurs dans le mois.

2.3 L'estimation des taux de transition avec un échantillon plus faible reste-t-elle satisfaisante (comparaison des taux de transition entre enquêtes emploi et panel européen) ?

Pour « valider » les transitions observées sur le panel, on a souhaité les rapprocher de celles observées sur les calendriers d'activité de l'enquête Emploi. Cette validation n'est bien sûr que partielle. En effet, les calendriers de l'enquête Emploi sont, comme ceux du panel, déclaratifs (la codification d'une activité repose largement sur la perception de la signification des différentes rubriques par l'enquêté). Toutefois, elle permet de s'assurer que la taille limitée de l'échantillon du panel, les incohérences observées lors des recouvrements ou encore les différents choix méthodologiques qui ont été faits (traitements des mois à plusieurs occupations, des occupations « non principales ») ne conduisent pas à des résultats très différents de ceux que l'on tirerait de l'analyse des calendriers des Enquêtes Emploi¹⁹.

Pour tenir compte du fait que les modalités des calendriers d'activité sont différentes dans l'enquête Emploi et dans le panel - et que, plus les rubriques sont détaillées, plus le « bruit » est important (encadré 4) - on a choisi de se limiter à trois catégories dans la suite de ce travail : emploi, chômage et inactivité.

Les transitions les plus fréquentes sont les transitions entre emploi et chômage (0,4 % des individus un mois donné, soit une cinquantaine de cas par mois²⁰) (tableau 11). En moyenne sur l'ensemble de la période étudiée (mars 1993 à septembre 1997), les taux de transition observés sur les enquêtes Emploi et le panel européen sont proches, sauf pour les transitions entre emploi et inactivité, pour lesquelles les taux de transitions sont supérieurs de 0,1 % dans le panel.

¹⁹ En effet, dans l'enquête Emploi, seule une occupation peut être déclarée un mois donné, et le nombre de types d'occupations est plus restreint. Enfin, nous n'exploitons ici les calendriers d'activité année par année, si bien qu'aucun problème de recouvrement de calendrier ne se pose.

²⁰ Ces probabilités de transition sont non conditionnelles à la situation de l'individu le mois précédent. Autrement dit, on rapporte le nombre de transitions emploi-chômage observées au nombre d'individus présents, et non au nombre d'individus en emploi.

Tableau 11 : Comparaison des taux de transition sur le marché du travail entre le panel et les enquêtes Emploi 1993-1997

Enquête	Période d'observation	%							Total
		Emploi à chômage	Emploi à inactivité	Chômage à emploi	Chômage à inactivité	Inactivité à emploi	Inactivité à chômage		
Panel européen	1/93 à 11/97	0,40	0,19	0,43	0,05	0,20	0,10	1,36	
	1/93 à 12/96	0,42	0,20	0,44	0,05	0,21	0,10	1,42	
	3/93 à 9/97	0,42	0,20	0,44	0,05	0,21	0,10	1,42	
Enquêtes Emploi	3/93 à 9/97	0,46	0,14	0,46	0,03	0,15	0,09	1,33	

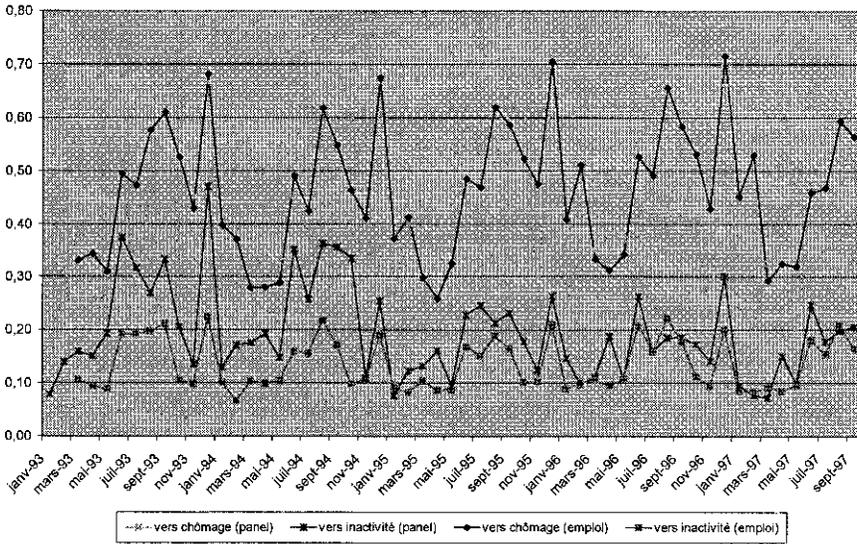
Source : Insee, *panel européen, vagues 1994, 1995, 1996, 1997; enquêtes Emploi 1994, 1995, 1996, 1997, 1998*

Champ : individus de 17 ans ou plus lors de l'enquête

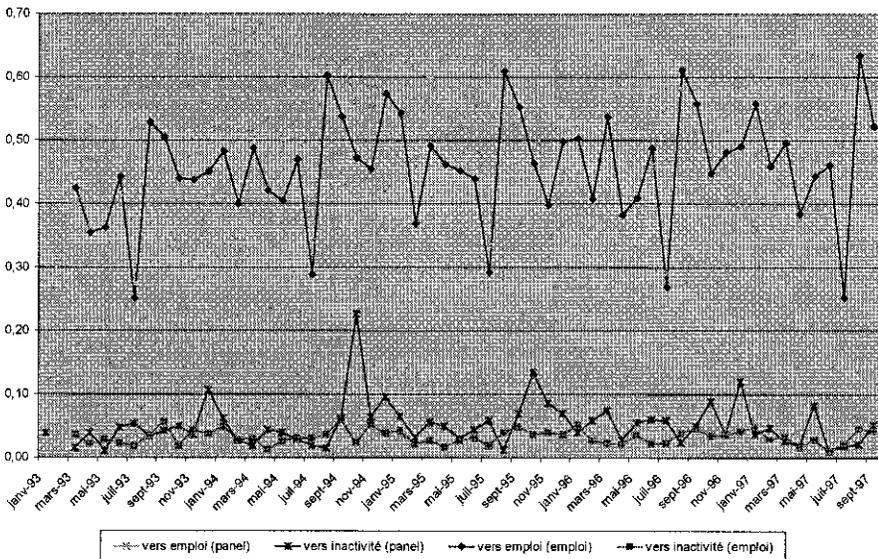
Lecture : En moyenne sur la durée d'observation du panel (janvier 93 à septembre 97), la probabilité d'observer une transition emploi/chômage un mois donné pour un individu est de 0,40 %. Sur la période correspondant aux calendriers de l'enquête Emploi, elle est de 0,42 %.

Graphiques 1 et 2 - Comparaison des données des trois premières vagues du panel européen et des Enquêtes Emploi

1. Transitions depuis l'emploi



2. Transitions depuis le chômage



Source : Insee, panel européen, vagues 1994 à 1997; enquêtes Emploi 1994 à 98.

Champ : individus de 17 ans ou plus lors de la vague d'enquête (pondérés de manière à être représentatifs de la population qui avait 17 ans ou plus)

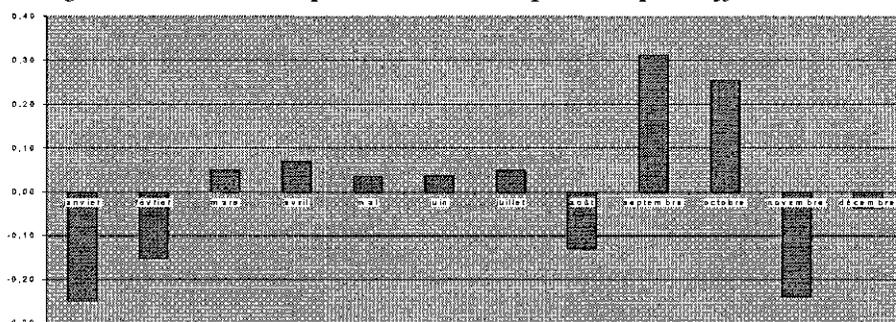
Toutefois, si les transitions entre emploi et inactivité sont plus fréquentes sur le panel européen que sur les enquêtes Emploi, cet effet est surtout marqué sur la période correspondant à la première vague du panel, à cause probablement d'un problème de questionnement (encadré 4). Les différences, plus mineures, qui subsistent au-delà de la première vague pourraient être dues au fait que l'importance des calendriers d'emploi est différente dans les deux enquêtes. Alors que le calendrier d'activité, très complet, est au coeur du panel, il est en fin de questionnaire de l'Enquête Emploi (et ne porte que sur l'activité principale)²¹. On retrouverait alors un phénomène analogue à ce que l'on observe sur les revenus par exemple : plus les questionnements sont détaillés, plus ils permettent de récupérer de l'information²². Ici, l'accent mis sur l'importance du calendrier dans le panel pourrait conduire à recueillir un nombre de transition supérieur dans cette enquête. La différence entre le panel et les enquêtes Emploi pourrait également être en partie liée aux thèmes qu'elles traitent. En effet, on note que si les taux de transition entre emploi et inactivité sont supérieurs dans le panel, les taux de transition entre emploi et chômage sont eux inférieurs à ceux de l'enquête Emploi : sensibilisés dans cette dernière enquête par les questions portant sur le chômage, les enquêtés pourraient-ils se déclarer plus facilement au chômage qu'en inactivité ?

Par ailleurs, l'écart entre le taux de transition dans le panel et dans l'enquête Emploi dépend fortement du mois calendaire, probablement en raison des effets de mémoire : les transitions apparaissent plus fréquentes dans l'enquête Emploi que dans le panel entre novembre et mars (mois de collecte de l'enquête Emploi), tandis qu'elles sont plus importantes entre mars et octobre dans le panel (graphique 4). Fait exception à cette règle le mois d'août, avec un nombre supérieur de transitions dans l'enquête Emploi (pour une raison inexplicée). Enfin, les valeurs observées lors des recouvrements de calendrier dans le panel sont particulièrement élevées, et il est difficile de faire la part exacte sur ces deux mois de l'effet de mémoire et des incohérences lors des recouvrements entre vagues du panel (encadré 4).

²¹ On peut se demander si, notamment, le panel permettrait de mieux saisir les alternances entre emploi ponctuel et inactivité des personnes principalement inactives. Ceci sera à examiner sur les données.

²² Accardo J. et M. Fall (1996) : « La mesure des revenus dans les enquêtes ménages et fiabilité des indicateurs d'inégalités dérivés », *Document de travail n°9602 de la Direction des Statistiques Démographiques et Sociales, Insee.*

Graphique 4 - Comparaison des taux de transition mensuels entre emploi, chômage et inactivité dans le panel et dans les enquêtes Emploi : effet du mois



Source : Insee, panel européen, vagues 1994 à 1997; enquêtes Emploi 94 à 98

Champ : individus de 17 ans ou plus lors de l'enquête

Lecture : Sur la période 1994-1996, les taux de transition entre emploi, chômage et inactivité observés sur le panel sont en moyenne supérieurs de 0,4 % à ceux observés sur les Enquêtes Emploi. A cet écart moyen s'ajoutent des écarts supplémentaires selon le mois considéré : -0,25 % en janvier, -0,15 % en février...

Enfin, le taux de transition dans le panel diminue de plus en plus (relativement au taux de transition dans l'enquête Emploi) au fur et à mesure que le nombre de vagues augmente (tableau 12 et graphiques 1 et 2). Cette baisse de la fréquence des transitions dans le panel peut être interprétée comme le signe d'une perte de représentativité du panel au fur et à mesure que le nombre de vagues augmente. Cette baisse reste toutefois encore modérée en vague 4 : l'écart absolu entre taux de transition est de 0,1 % (pour un taux de transition total de 1,4 % environ). Il est très légèrement plus important si on estime les taux de transition dans le panel à partir des seuls individus répondants lors des trois vagues. Il faudrait pouvoir disposer d'un nombre supérieur de vagues pour voir dans quelle mesure la baisse de représentativité pourrait être gênante²³.

²³ L'étude devrait alors sans doute se baser plutôt sur des probabilités de transition conditionnelles.

Tableaux 12 - Différence entre les taux de transition sur le marché du travail estimés sur le panel et les enquêtes Emploi 1993-1997

12a. Ensemble des individus de 17 ans et plus

	1993 (mars à décembre)	1994	1995	1996	1997 (janvier à septembre)	moyenne
emploi à chômage	-0,05	0,01	-0,01	-0,08	-0,08	-0,04
emploi à inactivité	0,11	0,11	0,04	0,03	0,02	0,06
total emploi vers non emploi	0,06	0,11	0,04	-0,05	-0,06	0,02
chômage à emploi	-0,06	0,01	0,03	-0,04	-0,04	-0,02
inactivité à emploi	0,10	0,10	0,05	0,03	0,00	0,06
total non-emploi vers emploi	0,04	0,11	0,08	-0,01	-0,04	0,04
chômage à inactivité	0,01	0,02	0,03	0,02	0,00	0,02
inactivité à chômage	-0,02	0,03	0,01	0,00	-0,01	0,01
total	0,10	0,27	0,15	-0,04	-0,10	0,09

Champ : individus de 17 ans ou plus lors de l'enquête

12b. Individus de 17 ans et plus présents lors des quatre vagues du panel

	1993 (mars à décembre)	1994	1995	1996	(janvier à septembre)	moyenne
emploi à chômage	-0,06	-0,02	-0,03	-0,08	-0,09	-0,05
emploi à inactivité	0,12	0,11	0,06	0,04	0,01	0,07
total emploi vers non em	0,06	0,09	0,03	-0,04	-0,08	0,01
chômage à emploi	-0,07	0,00	0,01	-0,04	-0,04	-0,03
inactivité à emploi	0,11	0,10	0,06	0,03	0,00	0,06
total non-emploi vers em	0,03	0,10	0,07	0,00	-0,05	0,03
chômage à inactivité	0,02	0,03	0,03	0,03	0,00	0,02
inactivité à chômage	-0,01	0,03	0,01	0,01	0,00	0,01
total	0,09	0,25	0,13	-0,01	-0,12	0,08

Source : Insee, panel européen, vagues 1994, 1995, 1996, 1997; enquêtes Emploi 1994, 1995, 1996, 1997, 1998

Champ : individus de 17 ans ou plus lors de l'enquête

Lecture : En moyenne de mars 93 à septembre 97, la probabilité d'observer une transition emploi/chômage un mois donné est inférieure de 0,04 points dans le panel si on l'estime à partir de l'ensemble des individus, et de 0,05 si on l'estime à partir des seuls individus présents lors des trois vagues.

Au final, la validation interne des calendriers d'activité du panel comme la comparaison avec les enquêtes Emploi indiquent que les transitions entre emploi et chômage que l'on définit à partir du panel sont de bonne qualité, et en particulier que le risque de transitions fictives (liées à la complexité du calendrier) est limité. Et si l'attrition biaise sans doute légèrement les taux de transition, ce biais reste en vague 4 plus limité que celui lié à l'effet de mémoire.

Conclusion

Les enquêtes répétées visent souvent plusieurs objectifs : le statisticien souhaite à la fois disposer de données en coupe fiables, pouvoir mesurer avec précision les évolutions entre deux vagues d'enquêtes successives, et disposer de données individuelles longitudinales. Malheureusement, ces objectifs sont difficilement compatibles (Caron et Ravalet, 2000) :

- La reconstitution de micro-données longitudinales portant sur une période relativement longue (4 ans ici) n'est possible que grâce à un panel d'individus. En effet, bien que le taux de déménagement soit faible chaque année (de l'ordre de 10 %), le taux de déménagement cumulé sur plusieurs années devient lui beaucoup plus élevé, et il s'ajoute de plus à l'attrition : ne pas suivre les personnes qui déménagent revient à se priver d'une part de plus en plus importante de l'échantillon, et nuit fortement à la représentativité de ce dernier. Les statistiques descriptives peuvent en être fortement modifiées (exemple des taux d'emploi à trois ans des jeunes étudiants), non seulement parce que le déménagement peut être directement lié au phénomène étudié (les jeunes qui trouvent un emploi quittent leurs parents), mais surtout, parce qu'il peut également y être lié plus indirectement en raison de l'hétérogénéité individuelle (les personnes dont la situation sur le marché du travail est la plus précaire sont également celles dont la situation de logement est la plus instable, notamment parce qu'il s'agit fréquemment de jeunes en phase d'accès à l'indépendance).
- Pour pouvoir estimer les évolutions d'une grandeur avec précision, un panel d'individus semble également préférable, dans la mesure où il permet de s'assurer que les variations observées d'une période à l'autre ne sont pas (ou peu) liées à des changements d'échantillon.
- A l'inverse, la qualité des estimations en coupe est meilleure dans des enquêtes transversales ou des panels de logements, qui ne sont pas sensibles à l'attrition²⁴.

Si le panel de logements semble un bon compromis pour disposer à la fois d'estimations en coupe de qualité et de mesures d'évolutions précises (enquête Emploi, enquêtes Permanentes Conditions de Vie), il ne permet pas de suivre les personnes sur longue période. Ainsi, choisir entre un panel d'individus et un panel de logements revient dans une large mesure à arbitrer entre représentativité en coupe et possibilité de suivi sur longue période. Ce travail montre qu'il est possible d'arbitrer

²⁴ Il faut toutefois signaler que le recours à un panel de logements, donc à une enquête répétée, peut avoir une incidence sur la qualité de l'information transversale. En effet, les ménages peuvent réagir différemment lors de passages ultérieurs de l'enquêteur (soit qu'ils comprennent mieux l'enquête, soit au contraire qu'ils essaient de limiter le temps de questionnement). Et l'attitude de ce dernier peut également varier (passation plus rapide des questions sur les points qui n'ont pas changé...). Ainsi, il semblerait que les trois tiers de l'enquête Emploi, même redressés indépendamment, ne conduisent pas à des résultats transversaux identiques.

en faveur d'un panel d'individus (sur 4 ans) sans devoir sacrifier de manière importante la qualité des estimations en coupe. En effet, la sélectivité de l'attrition sur 4 ans est suffisamment modérée pour que la structure de l'échantillon de répondants aux quatre vagues reste assez proche de la structure des répondants de départ. Par ailleurs, inclure des individus non-panels dans l'échantillon transversal permet de redresser une partie de ces biais. Et l'usage de pondérations longitudinales permet également de réduire les écarts observés entre estimations menées sur un échantillon cylindré et échantillon complet²⁵.

Au total, les taux de transition entre emploi, chômage et inactivité estimés à partir du panel restent proches de ceux de l'enquête Emploi. Mais bien sûr, il serait intéressant de refaire ce type d'étude avec un panel de plus longue durée...

²⁵ Tout du moins, ceux de ces écarts qui sont liés à des caractéristiques observables.

Encadré 1 - Champ de l'étude

Dans cette étude, l'unité statistique retenue sera l'individu panel de 17 ans et plus :

- En premier lieu, l'étude sera basée sur les *individus* et non sur les *ménages*. En effet, il est délicat de raisonner au niveau du ménage dès lors que l'on souhaite effectuer des analyses longitudinales, car les contours d'un ménage ne sont pas stables dans le temps (par exemple en cas de décès, divorce, décohabitation, mariage, naissance...). Notamment, le ménage peut être « éclaté » si seuls certains de ses membres déménagent.
 - Le choix d'une limite d'âge inférieure à 17 ans est à la fois dicté par les données (les questionnaires individuels du panel, qui contiennent notamment les calendriers d'activité, ne sont posés qu'aux 17 ans et plus) et par les thèmes qui vont servir d'exemple (principalement, l'analyse des trajectoires d'emploi). Pour l'analyse des trajectoires sur le marché de l'emploi, une limite d'âge supérieure sera également imposée (60 ans en première vague).
 - Enfin, dans la première partie, la plupart des résultats portent sur les seuls individus ayant rempli un questionnaire individuel en première vague. Parmi ces 14 334 individus, 10 409 (73 %) ont également répondu aux trois vagues suivantes. En revanche, dans la seconde partie, les taux de transition mensuels entre emploi, chômage et activité sont - sauf mention du contraire - estimés sur l'ensemble des individus de 17 ans et plus à la vague considérée (16 007 personnes), qu'ils aient ou non été présents depuis la première enquête (il suffit qu'ils soient présents lors des mois étudiés).
-

Encadré 2 - Liens entre déménagement et modification de situation sur le marché de l'emploi

Compte tenu de l'importance des thématiques liées à la reprise ou la perte d'emploi (en particulier dans les études portant sur des panels), il est intéressant de voir dans quelle mesure déménagement et modification de la situation sur le marché de l'emploi peuvent être liés²⁶. Un tel lien peut être direct (par exemple, une personne déménage pour aller exercer un emploi qu'elle a trouvé loin de son domicile, ou elle déménage à la suite d'une perte d'emploi vers un logement moins cher) ou indirect (les personnes les moins insérées sur le marché de l'emploi auraient une résidence moins stable).

Des liens directs ou des changements simultanés peu importants...

Intéressons-nous dans un premier temps aux liens « directs » entre changements sur le marché de l'emploi et déménagement. Une première manière de le voir si de tels liens existent est de comparer le statut d'occupation des personnes qui déménagent avant et après leur déménagement (tableau E.1). Que ce soit chez les jeunes ou sur l'ensemble de la population, les changements de situations coïncidant avec un déménagement semblent rares. Plus précisément, dans 87 % des cas, les personnes qui sont en emploi, au chômage, en études ou en inactivité (autre que les études) deux mois avant un déménagement sont encore dans la même situation deux mois après²⁷. Et il n'est pas possible de savoir à travers le calendrier d'activité si ces personnes ont changé d'emploi (ou d'établissement d'études).

Ainsi, si l'on prend l'exemple des 26-50 ans, ceux-ci sont 9 fois sur 10 dans la même situation deux mois avant et deux mois après le déménagement (tableau E.1). Dans 4 % des cas, les déménagements les concernant s'accompagnent d'un retour à l'emploi, et dans 5 % des cas, d'une perte d'emploi. Ce dernier chiffre est un peu plus élevé pour les femmes (6 %) que pour les hommes (4 %). L'étude de Dormont et Dufour-Kippelen (2000) sur les jeunes peu diplômés²⁸ confirme qu'il y a des interactions significatives entre durées d'accès au CDI et durées de cohabitation avec

²⁶ Les analyses de cet encadré ne sont menées que pour les seules personnes qui répondent, pour lesquelles il est possible d'exploiter la date de déménagement, le calendrier d'activité et le motif de déménagement. Plus précisément, l'échantillon est restreint aux personnes ayant répondu aux quatre vagues.

²⁷ Comparer les situations deux mois avant et deux mois après permet de s'affranchir d'éventuels « bruits » lors du mois du déménagement (absence d'activité durant ce mois pour permettre le déménagement). Une variante « 3 mois avant - 3 mois après » a été effectuée, et donne des résultats analogues.

²⁸ Dormont B. et Dufour-Kippelen S. (2000) : « Insertion professionnelle et autonomie résidentielle : le cas des jeunes peu diplômés », *Economie et Statistique*, n°337-338.

les parents, mais que ces interactions restent modestes en comparaison avec l'effet d'autres variables explicatives.

Tableau E.1 – Situations sur le marché de l'emploi deux mois avant et deux mois après le déménagement

	ensemble	hommes	femmes
de 17 à 20 ans			
études ---->études	64	ns	ns
emploi ----> emploi	14	ns	ns
sans emploi (hors études) -->sans emploi	11	ns	ns
dont initialement au chômage	8		
dont initialement inactif (hors études)	3		
études --> emploi	5	ns	ns
autres situations (emploi--> pas emploi, sans emploi-->emploi)	6	ns	ns
dont initialement inactif (hors études)	3		
ensemble	100		
de 21 à 25 ans			
emploi ----> emploi	67	70	61
études ---->études	9	9	7
études ----> emploi	4	5	6
emploi-----> sans emploi	5	5	5
sans emploi (hors études)--> emploi	3	6	5
sans emploi (hors études)-->sans emploi	12	4	17
dont initialement au chômage	8	3	13
dont initialement inactif (hors études)	4	1	4
ensemble	100	100	100
de 26 à 50 ans			
emploi ----> emploi	74	86	64
sans emploi (hors études) -->emploi	4	4	4
emploi ---->sans emploi	5	3	6
sans emploi (hors études)--> sans emploi	16	6	25
dont initialement au chômage	7	4	9
dont initialement inactif (hors études)	9	2	16
autres situations (études---->études)	1	1	1
ensemble	100	100	100
de 51 à 60 ans			
inactivité ----->inactivité	41	ns	ns
emploi ----> emploi	33	ns	ns
emploi---->sans emploi	14	ns	ns
sans emploi (hors études)-->sans emploi	11	ns	ns
dont initialement au chômage	1	ns	ns
dont initialement inactif (hors études)	10		
ensemble	110		
plus de 61 ans			
inactivité ----->inactivité	99	ns	ns
autres situations (emploi--->emploi, emploi---> sans emploi)	1	ns	ns
ensemble	100		

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : ensemble des déménagements des individus panel de 17 ans ou plus ayant rempli un questionnaire individuel en 1994, 1995, 1996 et 1997. Individus hors champ en vagues 2, 3 ou 4 exclus.

Une deuxième façon d'étudier les liens directs entre déménagement et changement de situation sur le marché de l'emploi est de demander aux personnes qui déménagent la raison de leur déménagement. Seuls 6 % des déménagements sont liés à une reprise d'emploi (tableau E.2). 15 % supplémentaires sont justifiés par une « autre raison liée à l'emploi », comme un rapprochement du lieu de travail. Et les 80 % restants relèvent d'autres motivations (raisons liées au logement...).

Tableau E.2 - Principal motif du déménagement entre deux vagues successives en fonction de la trajectoire entre ces deux vagues

	vous ou une autre personne du ménage avez trouvé un emploi ici	pour une autre raison liée à l'emploi (rapprochement du lieu de travail)	pour des raisons liées au logement	pour d'autres raisons	ensemble des principaux motifs exprimés
toujours en emploi	4	15	53	28	100
alternance emploi / non emploi	13	19	30	39	100
jamais en emploi	3	12	35	50	100
ensemble	6	15	44	36	100

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panels de 17 à 60 ans ayant rempli un questionnaire individuel en 1994 et aux deux vagues consécutives étudiées. Pourcentages sur les seules réponses exprimées (non-réponse d'environ 5 %). Moyennes sur les principaux motifs de déménagements exprimés dans les vagues 2, 3 et 4

Ces résultats, comme les précédents, invitent à relativiser l'importance des liens directs entre déménagement et situation sur le marché de l'emploi : dans la plupart des cas, le changement de logement ne serait pas concomitant à une modification de la situation sur le marché de l'emploi.

... mais des comportements fortement différenciés entre ceux qui déménagent et les autres.

Les analyses ci-dessus ne portent que sur des changements quasi-simultanés dans les domaines du logement et de l'emploi. Or, le déménagement peut être assez éloigné dans le temps de la reprise d'emploi, par exemple dans le cas d'une personne qui retrouve un emploi et ne décide de déménager que lorsqu'elle est sûre que son nouvel emploi sera durable. Par ailleurs, ces analyses ne tiennent pas (ou peu) compte des liens indirects entre emploi et logement liés à l'hétérogénéité individuelle (les personnes les moins bien insérées sur le marché de l'emploi sont sans doute plus susceptibles d'avoir une situation de logement instable).

Pour juger de l'ensemble de ces effets, il est possible de comparer les situations vis-à-vis de l'emploi des personnes qui déménagent et des autres. Ainsi, le taux d'emploi à trois ans de jeunes élèves ou étudiants est beaucoup plus élevé parmi ceux qui ont déménagé (cf. texte). Plus généralement, les trajectoires sur le marché de l'emploi de ceux qui ont déménagé sont beaucoup moins stables que celles de ceux qui n'ont pas changé de logement : 38 % de ceux qui ont déménagé entre l'automne 1994 et l'automne 1997 ont connu à la fois l'emploi et l'absence d'emploi durant cette

période, contre 20 % de ceux qui n'ont pas changé de logement. Autrement dit, 39 % de ceux qui ont connu à la fois l'emploi et le non-emploi ont déménagé. Le lien indirect entre emploi et logement joue sans doute un rôle important dans ces résultats, car les taux de déménagement entre 1994 et 1997 semblent varier à peu près autant en fonction de la trajectoire entre 1994 et 1997 (tableau E.3) qu'en fonction de la trajectoire antérieure (tableau 7c).

Tableau E.3 - Trajectoire entre novembre 1994 et octobre 1997 (vagues 2 à 4) selon l'existence d'un déménagement entre les vagues 1 et 4

trajectoire entre nov. 94 et oct 97 (36 mois)	ensemble	pas de déménagement	déménagement entre v1 et v4	taux de déménagement
Emploi tous les mois	54%	57%	46%	22%
tous les mois à temps	39%	40%	37%	24%
au moins un mois en temps	8%	8%	6%	20%
au moins un mois en	7%	9%	3%	11%
Emploi et non emploi	25%	20%	38%	39%
emploi puis	2%	2%	2%	23%
chômage puis	1%	1%	2%	44%
alternances d'emploi et de	8%	7%	14%	42%
activité et inactivité (+)	13%	10%	20%	40%
Jamais d'emploi	21%	23%	15%	19%
Total	100%	100%	100%	26%

Source : panel européen, vagues 1 à 4 (1994 à 1997), Insee

Champ : individus panel de 17 à 60 ans répondants en 1994, 1995, 1996 et 1997. Données pondérées.

Au total, se restreindre à un échantillon de personnes n'ayant pas déménagé pour étudier les trajectoires sur le marché de l'emploi conduit à se limiter à un échantillon de personnes dont les trajectoires d'emploi sont plus stables ou plus tournées vers l'emploi. Cela ne s'explique pas tant par le fait qu'elles n'ont pas connu de déménagement associé à un changement d'emploi, que par le fait que leurs caractéristiques diffèrent (et que l'effet de certaines de ces caractéristiques, inobservables, ne peut être corrigé par redressement). Les statistiques descriptives tirées d'échantillons de personnes n'ayant pas déménagé sont donc biaisées vers une plus grande stabilité ou une meilleure insertion sur le marché du travail. Mais surtout, une étude des changements sur le marché de l'emploi menée sur les seules personnes n'ayant pas déménagé conduit à ne pas tenir compte d'une part très importante des personnes qui ont connu l'événement étudié.

Encadré 3 Les poids longitudinaux des individus panel adultes dans le panel européen

Le principe général de calcul des pondérations longitudinales

Dans le panel, le poids d'un individu change d'une vague à l'autre. En effet, à chaque vague, les individus panel répondants doivent être représentatifs de la population de l'année, et leurs poids doivent donc être recalés pour tenir compte de l'attrition d'une vague sur l'autre. Ainsi si sur 100 individus répondants à la vague n (supposés identiques), seuls 50 sont répondants à la vague n+1, le poids (longitudinal) moyen de chaque répondant de n+1 sera le double de son poids en n. En fait, les individus ne sont pas identiques, si bien que le coefficient multiplicatif appliqué à chacun dépend de ses caractéristiques. Plus précisément, la méthode de calcul des poids est la suivante. Un modèle logit dichotomique permet dans un premier temps de repérer les variables qui permettent de mieux prédire la non-réponse. Sur la base de ces résultats, les individus sont regroupés en fonction de leurs caractéristiques pour définir des catégories homogènes par rapport à la non-réponse (chaque catégorie devant être d'effectif suffisant). Enfin, on calcule pour chaque catégorie le taux de non-réponse d'une vague à la suivante, ou plus généralement, à la vague n+i. Les pondérations de la vague n+i (notées $BASE_{n+i}$) s'obtiennent alors à partir des pondérations de la vague n par la formule : $BASE_{n+i} = BASE_n / (1 - \text{taux de non-réponse})$.

Les facteurs explicatifs de la non-réponse d'une vague sur l'autre.

Les taux de non-réponse varient en fonction des caractéristiques des enquêtés. Dans une optique de calcul de pondérations longitudinales, il est possible d'étudier les variations des taux de réponse à la vague n+i (sachant que la personne est répondante en vague 1 ou en vague n) en fonction de toutes les informations recueillies aux vagues précédentes. L'analyse des taux de non-réponse est donc plus riche que dans une enquête transversale, puisque l'on n'est pas obligé de se restreindre à l'analyse de l'impact des caractéristiques décrites dans la base de sondage.

Sur l'ensemble de la population, le principal facteur explicatif de la non-réponse d'une vague sur l'autre est l'appartenance à un ménage dont certains membres sont partis entre les deux vagues considérées. Viennent ensuite : la zone urbaine de résidence (l'agglomération parisienne a un taux de non-réponse plus élevé) et le type de contrat de travail de la personne de référence (les taux de réponse sont plus élevés quand la personne de référence est en CDI qu'en CDD), puis sa catégorie socio-professionnelle (ce sont les ménages agriculteurs qui répondent le plus, et les ménages d'ouvriers et d'artisans ou commerçants qui répondent le moins) et le niveau des charges de logement (des charges trop élevées entraînent une probabilité de non-réponse accrue à la vague suivante). Selon les vagues et les modèles, d'autres

variables peuvent avoir une influence, comme le type de ménage (les couples avec enfants ont des taux de non-réponse d'une vague sur l'autre plus faibles) ou le diplôme de la personne de référence (l'attrition est plus faible pour les diplômés de l'enseignement supérieur)²⁹.

Dans le cas où les personnes qui déménagent ne sont pas suivies, sont considérées comme « non-répondantes au sein du même logement » à la fois les personnes qui ne répondent plus, mais également toutes celles qui ont déménagé. Les facteurs explicatifs de cette non-réponse au sein du même logement ne sont plus les mêmes que précédemment : notamment, le fait d'appartenir à un ménage dont certains membres sont partis n'a plus d'influence, alors que la variable d'âge devient très fortement significative. Finalement, parmi les 6 variables retenues pour expliquer la non-réponse au sein du même logement, la plus discriminante est l'âge de la personne. Viennent ensuite le statut d'occupation du logement et la catégorie socio-professionnelle de la personne de référence, et enfin le type de contrat de travail de la personne de référence, la zone urbaine de résidence et le niveau des charges de logement. Par regroupement de classes d'individus obtenues par croisement de ces modalités, 55 catégories homogènes vis-à-vis de la non-réponse au sein du même logement sont définies.

²⁹ Cf. document de travail à venir sur la pondération de la vague 4 du panel.

Encadré 4 : qualité du calendrier d'activité du panel européen

Des calendriers particulièrement complexes et détaillés

Dans sa version française, le calendrier d'activité du panel européen est particulièrement complexe et détaillé. Il permet de connaître mois par mois depuis janvier 1993, les situations professionnelles des individus âgés de 17 ans et plus l'année de l'enquête. 17 situations différentes sont proposées : 9 correspondent au croisement entre le type d'emploi (Contrat à Durée Indéterminée, Contrat à Durée Déterminée ou indépendant) et la durée du travail (temps complet, temps partiel de plus ou de moins de 15 heures par semaine), 4 correspondent aux principales formes d'inactivité (retraite ou préretraite, études initiales, autre) ou au service national. Enfin, les 4 dernières modalités concernent l'existence d'activités secondaires, de formations (non initiales), de chômage partiel et enfin, d'absences pour cause de maladie. Dans le questionnaire français du panel européen³⁰, deux situations peuvent être recueillies pour un mois donné.

A l'inverse, seule la situation principale du mois est demandée dans le calendrier rétrospectif de l'enquête Emploi, et celle-ci n'est codée qu'en 8 modalités (1. A son compte ou aidait un membre de sa famille dans son travail ; 2. Salarié pour une durée non limitée ; 3. Contrat à durée déterminée, mission d'intérim, apprentissage, travail saisonnier ; 4. "Stages de la formation professionnelle", ou contrat d'aide à l'emploi ; 5. Au chômage ; 6. Elève, étudiant, stagiaire non rémunéré ; 7. Militaire du contingent ; 8. Autre situation : retraité, préretraité, retiré des affaires, femme au foyer, autre).

Qualité des calendriers d'activité du panel au sein de chaque vague..

Dans un premier temps, des contrôles de qualité et de cohérence ont été menés séparément sur les fichiers d'activité correspondant aux trois premières vagues du panel. Plus de 500 lignes des calendriers d'activité³¹ ont été corrigées sur la première vague (soit 2,6 % des lignes), près de 300 en deuxième vague (1,4 %) et plus de 150 en troisième vague (0,8 %). La recherche d'erreurs ayant été faite sensiblement de la même manière sur les trois vagues, il est intéressant de constater à travers ces chiffres que la qualité du panel s'est améliorée avec le temps (ce que l'on peut, au moins pour partie, imputer à la formation des enquêteurs) .

Voici une liste non exhaustive des problèmes détectés sur les lignes des calendriers d'activité :

³⁰ Ce n'est pas le cas dans la version européenne du questionnaire.

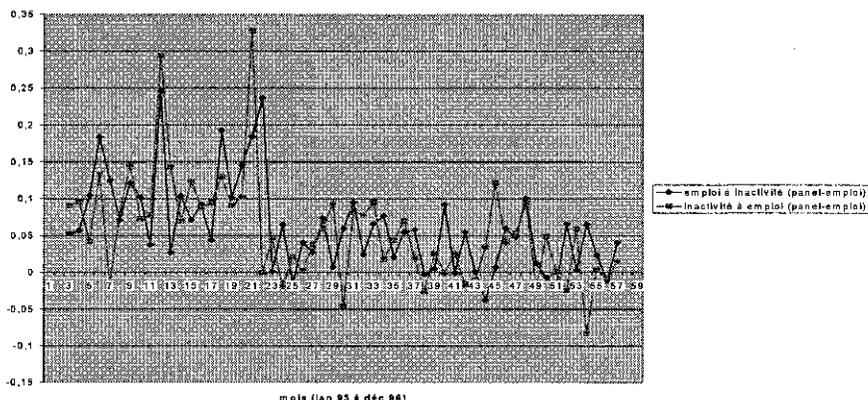
³¹ Il existe une ligne d'activité par type d'activité (occupation) déclarée au cours d'une vague par individu.

- Les lignes d'activités pour lesquelles aucun mois n'était rempli ont été renseignées dans la mesure du possible à l'aide des informations des calendriers de revenu (ou des questionnaires individuels). Cela concerne moins de 0,5 % des lignes en V1 (dont la moitié a été corrigée), et cette proportion diminue nettement en V2 et en V3.
- Les individus ne déclarant pour une vague donnée que des mois de maladie se sont vu ajouter une seconde ligne d'activité (voire une troisième), en fonction de leurs déclarations au questionnaire individuel (occupation, dates de fin d'activité...) et aux autres questionnaires (calendrier d'activité des autres vagues, fiches revenus...). En effet, toutes les lignes 'maladie' ne concernaient pas des travailleurs en absence temporaire (congés maladie), mais elles étaient parfois remplies par des retraités ou d'autres inactifs. Ces corrections portent sur 150 personnes en V1, mais sur moins d'une quinzaine en V2 et moins de 10 en V3.
- De nombreuses erreurs de saisie sur les codes activité ont également été détectées et corrigées, comme la confusion à la saisie entre les codes '11' et '17' ou '10' et '16'.
- De manière générale, l'activité « chômage partiel » est apparue peu fiable. Alors que cette modalité visait à récupérer de l'information sur les individus auxquelles leur entreprise imposait une activité réduite (du fait d'une perte d'activité de nature économique), certains enquêtés se sont déclarés en chômage partiel car à temps partiel et souhaitant travailler plus, ou encore car ils avaient quitté leur emploi au milieu de mois et avaient donc été partiellement chômeurs un mois donné. Enfin, certains chômeurs ont sans doute été classés en chômage partiel parce que la ligne d'activité « chômage partiel » figurait sur la première page du questionnaire (et non la ligne chômage, reléguée sur la suivante).
- Une analyse systématique a été menée pour détecter les actifs occupés qui auraient connu un ou deux mois d'inactivité pendant l'été. En effet, certaines personnes se sont déclarées inactives pendant leurs congés (parfois non sans raison, comme pour les assistances maternelles ou les surveillants dans les lycées). Lorsqu'il n'y avait pas de déclaration de changement d'activité, on a alors considéré que la personne était en fait occupée. Au total, les déclarations d'une dizaine d'actifs ont été corrigées en vague 1. Les ruptures de calendrier lors des mois d'été étaient en revanche beaucoup plus fréquentes pour les étudiants (qui ne se déclaraient plus en études lors des mois d'été, et étaient donc classés parmi les autres inactifs). Les corrections pour les étudiants ont été effectuées au cas par cas.

Comme l'indiquent certains résultats ci-dessus, la qualité de l'information dans le panel semble s'être améliorée avec le nombre de vagues. Certains problèmes n'apparaissent en fait qu'en vague 1, du fait d'un questionnement légèrement différent (ou d'enquêteurs moins rompus aux subtilités de l'enquête). Ainsi, une

comparaison entre les taux de transition entre emploi et inactivité dans le panel et les enquêtes Emploi indique que ces taux sont surestimés dans le panel pendant les 21 premiers mois (correspondant au calendrier d'activité de la vague 1) (graphique E.1). En effet, en première vague, la ligne « autre inactif » n'existait pas dans les calendriers d'activité. Cette ligne devait être reconstituée en négatif (absence d'autre activité), avec les risques que cela comporte.

Graphique E.1 - Transitions entre emploi et inactivité : comparaison des taux de transition du panel et de l'enquête Emploi



Source : Insee, panel européen, vagues 1994 à 1997; enquêtes Emploi 94 à 98

Champ : Individus de 17 ans ou plus lors de l'enquête

Lecture : Sur la période mars 1993 (mois 3)- septembre 1994 (mois 21), les taux de transition entre emploi et inactivité observés sur le panel sont en moyenne supérieurs de 0,1% à ceux observés sur les Enquêtes Emploi.

... et incohérences détectées lors des recouvrements entre vagues

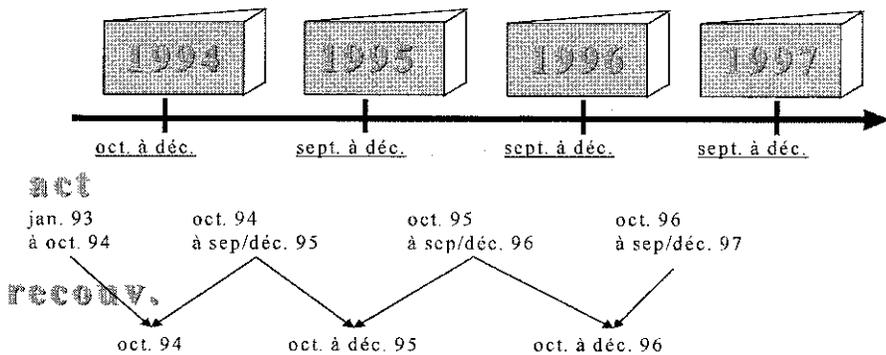
On détecte également des incohérences sur les calendriers d'activité lorsque ceux-ci se recouvrent, comme au mois d'octobre 1994 entre les deux premières vagues, ou aux mois d'octobre à décembre 1995 (resp. 1996) entre les vagues 2 et 3 (resp. 3 et 4) (schéma E.1). Chaque mois, de 1 à 4 % d'individus changent d'occupation « principale » d'après les trois premières vagues du panel (graphique E.2). Bien sûr, ces changements sont de nature très différente. Pour certains, il s'agit simplement d'un passage du temps partiel au temps plein, et pour d'autres, de la perte d'emploi ou du passage à la retraite. Mais plus les changements sont observés à un niveau fin, plus les risques d'erreurs (transitions fictives ou mal datées) sont importants :

- En vague 1 par exemple, la première ligne du calendrier d'activité (correspondant au CDI temps plein) a souvent été renseignée pour des individus qui se déclaraient en CDD ou à temps partiel lorsqu'on les interrogeait sur leur situation du moment dans le questionnaire individuel. Heureusement, la qualité du remplissage des calendriers d'activité s'est améliorée avec le nombre de vagues.

- Certains changements « mineurs » sont moins bien repérés par ces calendriers (par exemple, les transitions temps partiel / temps plein ou inactivité / chômage)...

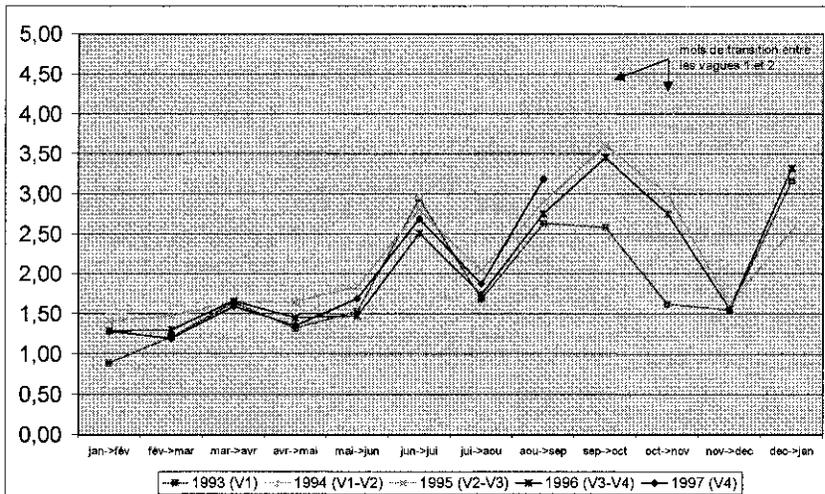
Ces phénomènes expliquent les forts taux de transition observés entre les vagues 1 et 2 (mois de septembre à novembre 1994). Alors que les courbes correspondant aux différentes années sont proches sur les autres mois, la période automnale se caractérise par des écarts plus importants (graphique E.2). On constate notamment que les taux de transition observés entre septembre et novembre 1994 (passage du calendrier vague 1 au calendrier vague 2) sont plus importants que ceux observés à la même époque en 1993 (relevés sur le calendrier de la seule première vague). L'accroissement du taux de transition en automne se réduit lors des deuxième et troisième vagues (puis lors des troisième et quatrième vagues) (graphique E.2), ce qui renforce l'idée que la qualité du panel s'améliore notablement avec le nombre de vagues³².

Schéma 1 - Les calendriers d'activité des quatre premières vagues du panel : mois sur lesquels portent les enquêtes et périodes de recouvrement



³² Si l'on regroupe les Contrats à Durée Indéterminée (CDI) (resp. les contrats à durée déterminée (CDD), les activités indépendantes) de plus de 15 heures par semaine avec les CDI (resp. les CDD, les activités indépendantes) de moins de 15 heures par semaine (et compte tenu des apurements préalablement effectués), seules 179 lignes activités n'ont pas la même valeur pour le mois d'octobre 1994 en vagues 1 et 2, et 148 pour les mois d'octobre à décembre en vagues 2 et 3. Ces problèmes de recouvrement ont été traités de manière à assurer la plus grande continuité dans les calendriers. Une attention particulière a été portée aux individus ne déclarant pas d'activité lors du mois d'enquête.

Graphique E.2 - Taux mensuels de changements d'occupation principale d'après le panel européen en fonction du mois et de l'année (nomenclature du calendrier d'activité en 19 postes).



Source : Insee, panel européen, vagues 1994, 1995, 1996, 1997

Champ : individus de 17 ans ou plus lors de la vague d'enquête (pondérés de manière à être représentatifs de la population qui avait 17 ans ou plus)

Résumer toutes les situations déclarées pour un mois donné en une occupation principale

Dans le calendrier d'activité du panel, plusieurs occupations différentes peuvent être déclarées un même mois. Ainsi, un individu peut être le même mois en CDD temps plein et en maladie (occupations simultanées), ou bien en CDD temps plein et au chômage s'il a été licencié au cours du mois (occupations consécutives). Le traitement des cas d'occupations multiples peut bien sûr avoir une influence sur les transitions que l'on repère. Pour cette étude, les différents types d'occupations présents un mois donné ont été synthétisés en une occupation dite principale. Pour ce, l'activité a été privilégiée sur le chômage, lui-même privilégié sur l'inactivité. Ce choix implique que de courts passages par le chômage (inférieurs à un mois) ne sont pas pris en compte. Implicitement, on ne s'intéresse donc qu'aux périodes de chômage ayant au moins une durée de l'ordre du mois. Au terme de cette étape, les occupations obtenues pour certains mois restaient difficiles à classer entre activité, inactivité et chômage. C'était le cas par exemple lorsque l'occupation principale du mois était une activité secondaire (!) ou une formation complémentaire (qui peut avoir lieu dans le cadre d'un emploi ou non). C'était également le cas pour des occupations telles la maladie ou le chômage partiel, qui - compte tenu des déclarations observées - ne semblaient pas uniquement concerner des actifs. Un peu plus de 1 % des mois d'activité étaient concernés. Le choix qui a été fait est de

prolonger les occupations déclarées le mois précédant et le mois suivant la maladie, le chômage partiel, la formation complémentaire ou l'activité secondaire. Lors de ce traitement, entre un cinquième et un quart des cas de chômage partiel ou de maladie ont ainsi été classés en chômage ou inactivité, ainsi que la moitié des cas de formations complémentaires, et les deux tiers des activités secondaires. Au total, 7 % des calendriers individuels ont été modifiés à cette étape.

Annexe : taux de déménagement et d'éclatement de ménages.

Une sous-estimation négligeable des taux de déménagement

Dans le panel européen, plusieurs variables contiennent de l'information sur l'existence d'un déménagement ou d'un « éclatement de ménage » (situations dans lesquels certains membres présents dans le ménage à l'enquête précédente le quittent). Toutefois, ces variables ne sont pas toutes renseignées systématiquement. Certaines ne le sont notamment que lorsque que le ménage accepte de répondre au questionnaire. Pour la présente étude, toute l'information disponible a été utilisée pour repérer les cas de déménagement ou d'éclatement de ménage, mais les taux de déménagement et d'éclatement restent malgré cela encore légèrement sous-estimés. En effet, si on peut supposer que l'on arrive à détecter presque tous les cas où tous les membres du ménage quittent le logement, en revanche, quand le ménage est non-répondant (et que tous ses membres n'ont pas déménagé), il n'est pas possible de savoir si certains membres du ménage l'ont tout de même quitté. On est amené à supposer qu'aucun membre du ménage n'est parti, ce qui conduit à des taux de déménagement (ou d'éclatement de ménage) légèrement sous-estimés. Ce biais est toutefois négligeable. En effet, la non-réponse est de l'ordre de 10 %, et le taux d'éclatement de ménage est de quelques pourcents (et est sans doute plus faible au sein des ménages qui déménagent), soit un biais inférieur à 0,5 %³³.

Réconcilier des informations multiples mais souvent incohérentes

Déménagement du ménage

Lorsque le ménage est non-répondant, on utilise les variables TRESFA, remplies par l'enquêteur, qui décrivent les résultats des essais de contact de ce dernier avec le ménage, ainsi que la variable TCHAD (qui décrit l'existence d'un changement d'adresse depuis la vague précédente - téléphonique³⁴ ou en vis-à-vis - et est donc plus difficilement exploitable). Lorsque le ménage est répondant, on peut de plus utiliser la question qui lui est directement posée de savoir s'il a déménagé (LOGID). Les codes des communes de résidence lors des différentes vagues d'enquêtes sont aussi utilisés pour trancher les cas litigieux. Toutefois, toutes ces informations ne sont pas toujours cohérentes entre elles (cf. notamment les travaux de L. Gobillon).

³³ En revanche, le fait que l'on ne puisse pas savoir si une personne a quitté le ménage si ce dernier ne répond pas interdit de calculer de manière valable des taux de réponse pour les individus restant dans le ménage, selon que d'autres membres du ménage sont partis ou non (on aboutirait par exemple en vague 1 à un taux de réponse en vague 2 anormalement élevé de 96 % pour les individus qui n'ont pas déménagé à l'inverse d'autres membres de leur ménage : ce taux élevé est simplement dû au fait qu'on ne sait en général quand un individu a quitté le ménage que si un autre membre au moins du ménage accepte de répondre).

³⁴ Des vagues téléphoniques ont été intercalées entre les premières enquêtes en vis-à-vis réalisées chaque année à l'automne. Elles avaient pour principal objet de repérer les principales modifications affectant le ménage, mais n'ont pas été exploitées.

Notamment, dans environ la moitié des cas, le code de la commune change alors que le ménage déclare ne pas avoir changé de logement. Il s'agit souvent de problème de codage d'arrondissement. Dans tous les cas, l'information donnée par le ménage a été privilégiée, en particulier quand elle s'accompagne de réponses à des questions sur la date et les motifs du déménagement.

Déménagement de l'individu

Pour les ménages répondants, les personnes ayant quitté leur ménage sont identifiées avec la variable TNOC (mouvement des personnes), et bien sûr par comparaison des présents dans le ménage aux différentes vagues ou examen du numéro de ménage.

Au total, les informations relatives aux mouvements des individus ou des ménages ou aux difficultés rencontrées pour interroger le ménage ou certains de ses membres sont multiples, souvent incohérentes ou incomplètes. La détermination des personnes ayant déménagé repose donc sur des règles de décision, dont la modification pourrait faire légèrement varier les statistiques sur la non-réponse et ses causes.
