

Reconstruire des trajectoires de mobilité résidentielle

Éléments d'une analyse biographique des données de l'EDP

Daniel
Courgeau,
Éva Lelièvre
et Odile
Wolber *

Dans un pays sans registre de population, l'échantillon démographique permanent (EDP) constitue un suivi géographique unique des individus, bien qu'il ne donne le lieu de résidence qu'à certaines dates (recensement, mariage, naissance d'enfants). Afin d'exploiter cette source, il est toutefois nécessaire de reconstituer les itinéraires individuels et d'en évaluer les limites.

Une analyse de ces trajectoires permet de préciser les grands traits de la mobilité géographique des Français. Des principaux résultats se dégagent le fort enracinement des indépendants, encore plus marqué lorsqu'ils sont fils d'indépendants, une mobilité résidentielle augmentant avec le niveau d'éducation, et une faible durée de séjour pour les cadres ou les étudiants dont le père est lui-même cadre.

* Daniel Courgeau et Éva Lelièvre appartiennent à l'Ined, Odile Wolber est statisticienne à l'IHESI.

Les noms et dates entre parenthèses renvoient à la bibliographie en fin d'article.

En appariant les données issues des bulletins individuels des recensements et des bulletins statistiques de l'état civil depuis 1968, l'échantillon démographique permanent de l'Insee (appelé EDP par la suite) est une source biographique exceptionnelle tant par sa taille que par les informations qu'il comporte. L'EDP, en effet, procure un suivi longitudinal individuel d'informations concernant la vie familiale, la carrière professionnelle et la mobilité géographique (Courgeau et Lelièvre, 1989). Ainsi, pour les individus ayant eu 15 ans en 1968, l'analyse du parcours familial est aisée, du moins tant qu'on ne fait pas intervenir les périodes de cohabitation : le mariage, les naissances

successives d'enfants y sont enregistrés et datés au fur et à mesure qu'ils se produisent, avec un certain nombre de caractéristiques des deux conjoints et de leurs enfants.

Une biographie fragmentaire

En revanche, si l'on s'intéresse à la vie migratoire ou professionnelle, on ne dispose plus d'une datation aussi précise des événements (déménagement, changement d'emploi, promotion, etc.), car ce fichier ne saisit la résidence et la profession des individus qu'aux dates des recensements et des événements

familiaux (cf. schéma 1). Dans ce cas, les dates observées encadrent l'événement migratoire ou professionnel que l'on veut étudier, ce qui revient à dire que l'on a une *biographie fragmentaire* (Courgeau et Najim, 1995). Les méthodes usuelles d'analyse des biographies, appelées aussi "analyse de durée", ne sont dès lors plus directement utilisables.

Cependant, comme cette biographie fragmentaire fournit une information non négligeable sur la mobilité géographique et professionnelle, elle peut être utilisée pour donner une estimation de ces mobilités et les relier à la vie familiale des individus. Mais cela n'est possible que si un certain nombre d'hypothèses et de conditions sont correctement vérifiées (cf. schéma 1).

Deux hypothèses nécessaires

Afin d'exploiter des données biographiques fragmentaires, telles que les données migratoires de l'EDP, deux hypothèses principales sont faites.

En premier lieu, la densité des points d'observation doit être suffisante pour supposer raisonnablement qu'un seul changement d'état ait pu intervenir entre deux observations consécutives. Si par exemple, plusieurs changements de département se produisent entre deux dates d'observation, aucune information ne sera disponible sur les départements intermédiaires où l'individu a vécu, et aucune migration ne sera même enregistrée s'il effectue un retour vers son département initial. L'hypothèse est donc faite qu'un seul, au maximum, des événements étudiés (migration par exemple) peut se produire entre deux observations. Pouvoir localiser les individus à la fois aux recensements et aux divers événements familiaux accroît la densité d'observation dans le temps, ce qui réduit au minimum le risque d'omission.

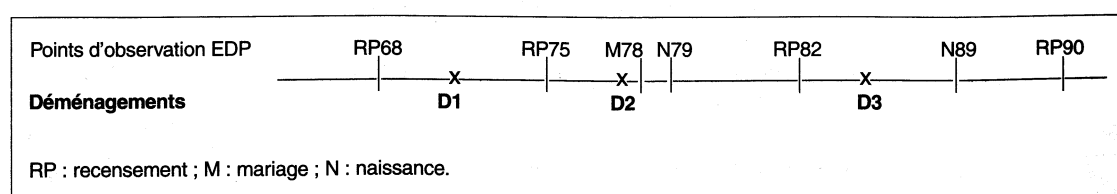
En second lieu, il est nécessaire, pour conduire une analyse biographique, de faire l'hypothèse que le plan de sondage, ici le plan d'observation, est *non informatif* (Hoem, 1985), c'est-à-dire que l'observation des trajectoires individuelles se fait indépendamment de leur déroulement, à des moments aléatoires. Cette seconde hypothèse est parfaitement vérifiée dans le cas des recensements, mais pas pour les données d'état civil collectées à l'occasion des événements familiaux. Il n'y a *a priori* aucune raison pour que les migrations spatiales ou la mobilité professionnelle soient indépendantes des événements familiaux. Ainsi, un couple installé dans un studio va, à la suite de la naissance d'enfants, déménager vers un logement de plus grande taille. De même, la naissance des enfants peut induire un changement d'activité pour les parents. Dans ces cas, l'observation des dates de naissances successives risque de n'être plus indépendante des migrations ou des changements professionnels survenus dans l'intervalle (1). Le risque de biais est alors d'autant plus important que l'on travaille sur des migrations à courte distance très liées à l'accroissement de la taille des familles.

Une vérification expérimentale

Afin d'éprouver ces hypothèses dans le cas de l'utilisation de l'EDP, un travail expérimental préparatoire a été conduit en tronquant artificiellement des données complètes. Cela était possible en utilisant les données d'une enquête rétrospective recueillant tous les événements familiaux, migratoires et professionnels (enquête *Triple Biographie* ou *3B* de l'Ined). En rendant ces biographies artificiellement fragmentaires, il est possible de vérifier les biais ainsi introduits (voir pour plus de détails Courgeau et Najim, 1995).

1. On trouvera une formulation probabiliste de ces problèmes dans Courgeau et Najim (1995), pp. 166-167.

Schéma 1
Exemple de trajectoire reconstituée à partir des données de l'EDP



La densité des points d'observation est faible si l'on travaille sur les seules données de l'état civil, les résultats s'améliorent sur les seules données des recensements et deviennent optimaux lorsque l'on utilise à la fois l'état civil et les recensements. Les biais sont également réduits pour les changements de logement ou de commune, mais deviennent importants pour les migrations à plus longue distance, entre les départements par exemple. On arrive ainsi à une sous-estimation de la mobilité interdépartementale, du fait en particulier de l'importance des retours vers un département d'origine. Ce sont là des risques inhérents à cette méthode d'observation, que seule une observation plus détaillée permet de mettre en évidence.

L'utilisation des données de l'enquête 3B permet de vérifier également l'importance des biais introduits par les dépendances entre vie familiale et vie migratoire, c'est-à-dire entre le plan d'observation et le phénomène étudié. L'effet de ces dépendances sur l'estimation des migrations est cependant réduit, d'autant plus que l'on utilise simultanément l'information fournie par les recensements. Elles entraînent essentiellement un changement de calendrier, mais affectent peu l'intensité du phénomène étudié.

Les données de l'EDP se prêtent donc à une analyse de durée adaptée aux données fragmentaires grâce aux modèles développés à l'*Ined*. Néanmoins, les résultats d'une analyse de la mobilité des individus dépendent fortement de la qualité des informations disponibles sur leur localisation.

Restaurer la cohérence des trajectoires

Pour que l'information fournie par l'EDP puisse être utilisée pour une telle estimation, il faut également s'assurer de sa qualité. Cette vérification s'avère indispensable, car les renseignements recueillis lors des recensements et par les officiers d'état civil peuvent être incorrectement déclarés ou omis par les intéressés : certaines personnes ou même certains ménages échappent au recensement (Coeffic, 1993) ; certains individus, bien que recensés, n'ont pas été appariés dans l'EDP parce que les éléments permettant de les intégrer dans le fichier (jours ou mois de naissance, par exemple) font défaut ou sont mal déclarés.

Une telle vérification a fait l'objet d'un travail important qui sera publié par ailleurs (Cour-

geau, Lelièvre et Wolber, 1998). Sans entrer dans les détails des analyses, il est utile de présenter les principaux résultats obtenus. Cette vérification a en effet conduit à éliminer dans les trajectoires individuelles des informations de localisation trop incohérentes et à en redresser d'autres. Cette étape était indispensable, les incohérences dans les trajectoires brutes reconstituées interdisant leur analyse.

Analyser la mobilité géographique des individus de l'EDP oblige à reconstituer leur trajectoire résidentielle et à s'assurer de leur vraisemblance. Une typologie des diverses incohérences rencontrées dans ces trajectoires a pu être établie à partir d'une étude approfondie des divers lieux occupés tout au long de la vie des individus. De cet inventaire découle une série d'hypothèses *les plus vraisemblables selon les éléments disponibles*, qui constituent les règles de " correction " appliquées pour restaurer la cohérence des trajectoires. Cependant, comme les éléments d'information sont succincts, plusieurs règles de décisions alternatives sont vraisemblables et plusieurs fichiers résultant de règles de décisions différentes ont été constitués. Toutefois ici le seul fichier utilisé est celui où les redressements effectués sur les données sont les moindres même si la comparaison des résultats d'une même analyse obtenus à partir des différents fichiers reste très intéressante, mais elle n'est pas encore finalisée.

Un exemple d'incohérence : la déclaration du lieu de résidence au 1^{er} janvier de l'année du recensement précédent

L'utilisation de la déclaration de résidence au 1^{er} janvier de l'année du recensement antérieur s'avère problématique. Un premier type de situation concerne les " faux absents " : en effet, cette question permet de dénombrer les individus qui déclarent un lieu de résidence en France au 1^{er} janvier de l'année du recensement antérieur alors que l'on n'a pas de bulletin pour eux à ce recensement, soit qu'ils n'aient pas été saisis au recensement précédent, soit que le bulletin individuel n'ait pas été récupéré par l'EDP, soit qu'ils se soient trompés dans la datation de leur précédent séjour. Les proportions de faux absents pour les individus de l'EDP nés entre 1945 et 1960 aux années des recensements de 1968, 1975 et 1982 (2) sont les plus élevées

2. Pour ces années, on dispose de la localisation des individus au recensement (en mars) que l'on confronte à leur déclaration (faite au recensement suivant) de résidence au 1^{er} janvier.

avec un maximum d'environ 14 % autour de 22-23 ans (cf. graphique I).

D'une façon générale, les incohérences sont maximum pour les localisations de l'année 1982 car les données de localisation géographique de l'EDP sont redressées (3) pour le recensement de 1990. Ainsi, aux problèmes de déclarations s'ajoutent les différences dues au redressement et la chance est moindre de retrouver les individus en 1990. On remarquera que les déclarations des femmes sont plus souvent concordantes que celles des hommes, et cela à tous les recensements.

Des invraisemblances existent aussi entre le lieu déclaré pour la résidence au 1^{er} janvier du

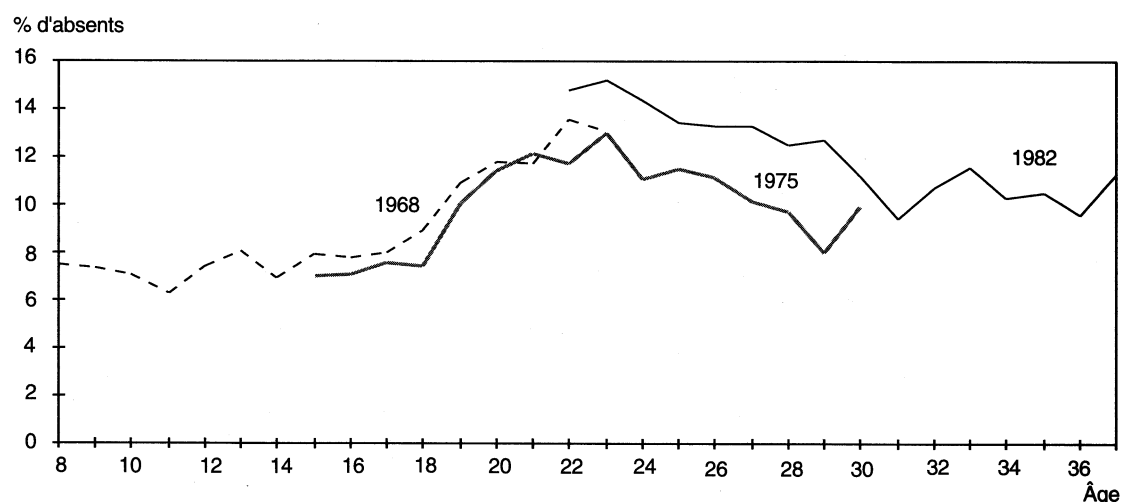
recensement précédent et celui recueilli en mars au moment de ce recensement. En dépit de la mauvaise fiabilité de cette déclaration, il est néanmoins indispensable d'utiliser cette information. Dans le fichier utilisé ici, elle est conservée quand le lieu déclaré coïncide avec une des localisations encadrantes, sinon l'information est déplacée dans le temps en fonction de critères de cohérence élaborés à l'aide des autres lieux de la trajectoire. La complexité des nombreux cas rencontrés ne permet pas de les détailler ici.

Au mariage, on dispose de trois lieux : la déclaration de la résidence antérieure de l'individu, le lieu du mariage et la déclaration de la résidence conjugale après le mariage. En exami-

3. Les données brutes du recensement de 1990 n'ont pas été appariées au reste de l'échantillon de l'EDP contrairement aux données des recensements de 1968, 1975 et 1982.

Graphique I

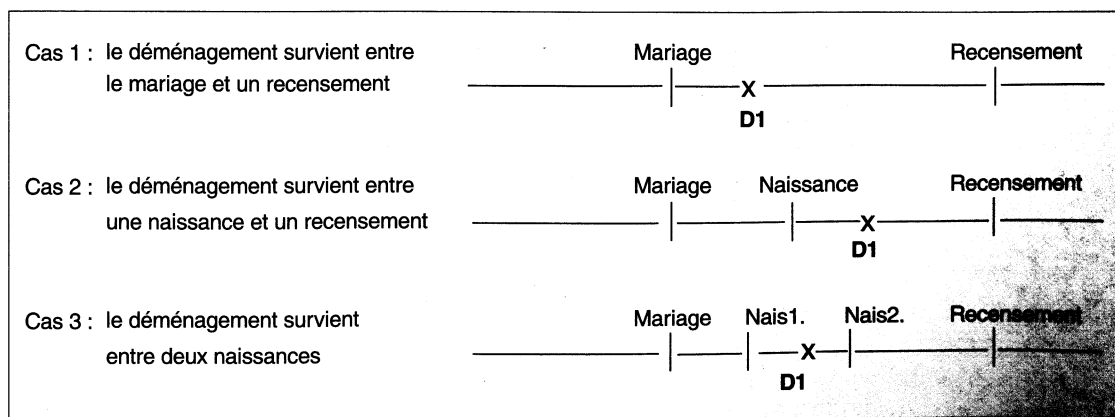
Proportion d'individus absents à un recensement alors qu'ils déclarent au recensement suivant qu'ils résidaient en France cette année-là



Lecture : une partie des personnes qui ont déclaré (au recensement de 1975) qu'elles résidaient en France au 1^{er} janvier 1968, ne figurent pas au recensement de 1968. La proportion culmine à 13,6 % pour celles qui avaient 22 ans en 1968.
Champ : individus nés de 1945 à 1960.
Source : EDP.

Schéma 2

Trajectoires sélectionnées à partir des données de l'EDP



nant ces déclarations et celles qui les entourent on s'aperçoit que souvent le lieu de la cérémonie ne correspond pas à un lieu de résidence, que le lieu déclaré comme antérieur peut correspondre à la résidence de l'enfance, etc. Un ou deux de ces lieux doivent alors être supprimés (Courgeau, Lelièvre et Wolber, 1998).

Une méthode d'estimation des durées de séjour non observées

Le principe d'une telle estimation n'est rappelée ici que dans le cas le plus simple, celle-ci ayant été présentée de façon plus détaillée dans Courgeau et Najim (1995).

Le mariage est pris ici comme origine du processus puisque l'état civil fournit une date précise et un lieu d'installation. La date de la première migration après le mariage n'est pas connue, mais peut être encadrée par deux points observés au recensement ou à l'état civil. Lors de la première observation, l'individu réside au lieu déclaré juste après le mariage (4) ; lors de la seconde, il réside ailleurs (cf. schéma 2). Une migration s'est donc produite entre ces deux dates sans qu'elle puisse être située avec plus de précision.

4. Le mariage peut être ce premier point d'observation quand la date suivante renseignée correspond à un changement de résidence.

Encadré

UNE APPROCHE STATISTIQUE DES BIOGRAPHIES FRAGMENTAIRES

Plaçons-nous dans le cas simple où le début d'observation est connu avec certitude (date de mariage), mais où la date de migration T , ne peut qu'être située entre deux dates t_1 et t_2 . Bien entendu, la seconde date peut ne pas être connue si l'individu est toujours dans le même lieu lors de la dernière observation.

Nous supposons ici que la variable aléatoire T est discrète (annuelle par exemple). Nous cherchons alors à mesurer les probabilités $m_h = P(T = t_h)$ pour que la migration se produise l'année t_h ($1 \leq h \leq r$, où r est la durée maximum observée). Introduisons des paramètres α_h^l qui pour l'individu, l , sont égaux à l'unité lorsque $t_1^l \leq t_h \leq t_2^l$, nuls dans le cas contraire. La vraisemblance des observations peut alors s'écrire :

$$L = \prod_{l=1}^n \left(\sum_{h=1}^r \alpha_h^l \cdot m_h \right) \quad (1)$$

où N est le nombre d'individus observés. Les valeurs de m_h qui rendent cette vraisemblance maximum sont obtenues par une procédure d'ajustement itératif (De Gruttola et Lagakos, 1987). Pour ce, introduisons une variable aléatoire égale à 1 si la vraie valeur observée de T est égale à t_h pour le l ème individu et égale à 0 sinon. Alors l'espérance mathématique de cette variable conditionnée par α^l est égale à :

$$\mu_h^l = \frac{\alpha_h^l \cdot m_h}{\sum_h \alpha_h^l \cdot m_h} \quad (2)$$

Il en résulte l'estimation suivante de m_h :

$$\hat{m}_h = \frac{\sum_l \mu_h^l}{N} \quad (3)$$

Pour obtenir cette estimation, partons de valeurs initiales arbitraires de m_h , puis calculons les valeurs de μ_h^l correspondantes à l'aide de (2) qui permettent de recalculer une nouvelle valeur de \hat{m}_h à l'aide de (3). On répète ces itérations jusqu'à ce que les écarts entre les valeurs successives de m deviennent négligeables. On arrive ainsi à un maximum local ou à un point-selle de la vraisemblance. Le calcul de l'opposée de la matrice des dérivées secondes de $\log L$, permet à la fois de vérifier que l'on n'est pas sur un point-selle (si les valeurs propres de cette matrice sont toutes positives) et fournit grâce à son inverse une estimation de la matrice des variances et covariances de \hat{m} . Cette méthode peut se généraliser au cas où à la fois le début du séjour et sa fin sont encadrés par deux dates chacun, et au cas où des caractéristiques peuvent jouer sur cette durée de séjour à l'aide d'un modèle à risques proportionnels (Kim *et al.*, 1993). Pour vérifier qu'un tel modèle s'adapte bien à une caractéristique binaire par exemple, on peut éclater les observations en deux sous-populations, l'une ayant cette caractéristique, l'autre ne l'ayant pas. On estime alors les probabilités de migrer dans ces deux groupes m_h^1 et m_h^0 , qui peuvent fournir des estimations des quotients cumulés de migration jusqu'à l'année h : H_h^1, H_h^0 . En portant les logarithmes de ces quotients cumulés en fonction de la durée, on obtient deux courbes à peu près parallèles si le modèle à risques proportionnel est vérifié (Courgeau et Lelièvre, 1989, pp. 121-122).

Pour estimer les probabilités de migrer à chaque durée de séjour, on procède par itérations successives pour construire une vraisemblance dans laquelle ces probabilités, inconnues au départ, sont retenues si la durée se situe entre les deux dates encadrantes pour un individu donné, et annulées dans le cas contraire. La recherche des probabilités rendant cette vraisemblance maximale est obtenue grâce à une procédure d'ajustement itératif (De Gruttola et Lagakos, 1989) qui fournit également la matrice des variances et covariances de ces diverses estimations (cf. encadré).

Cette méthode peut être généralisée à l'ensemble des migrations suivant le mariage pour lesquelles à la fois le début du n^{ème} séjour et sa fin sont inconnus et encadrés par deux dates d'observations chacun (cf. schéma 2).

L'effet des caractéristiques individuelles disponibles sur ces probabilités de migrer est ensuite estimé (effet du niveau d'éducation, de la profession et de l'âge au mariage sur la mobilité des individus). Mettre en évidence ces effets suppose qu'ils agissent de façon proportionnelle sur les quotients de migration : il est possible de vérifier la validité d'un modèle à risques proportionnels en scindant la population soumise au risque en sous-populations selon les catégories de variables retenues (cf. encadré). Comme précédemment, il reste alors à calculer une vraisemblance pour l'ensemble des individus, dont le maximum sera obtenu pour les valeurs des probabilités que l'on cherche à estimer (Kim *et al.*, 1993).

Une application à la mobilité des personnes mariées nées entre 1953 et 1956

Étudier les changements de commune après le mariage met en évidence la mobilité au niveau le plus fin. Cette mobilité est liée à l'amélioration des conditions de logement tout au long de la vie. Pour cela un échantillon d'individus mariés après 1968 (date de prise en compte des bulletins d'état civil dans l'EDP) est suivi pendant une durée suffisante après leur mariage, en l'occurrence les générations nées entre 1953 et 1956 (5).

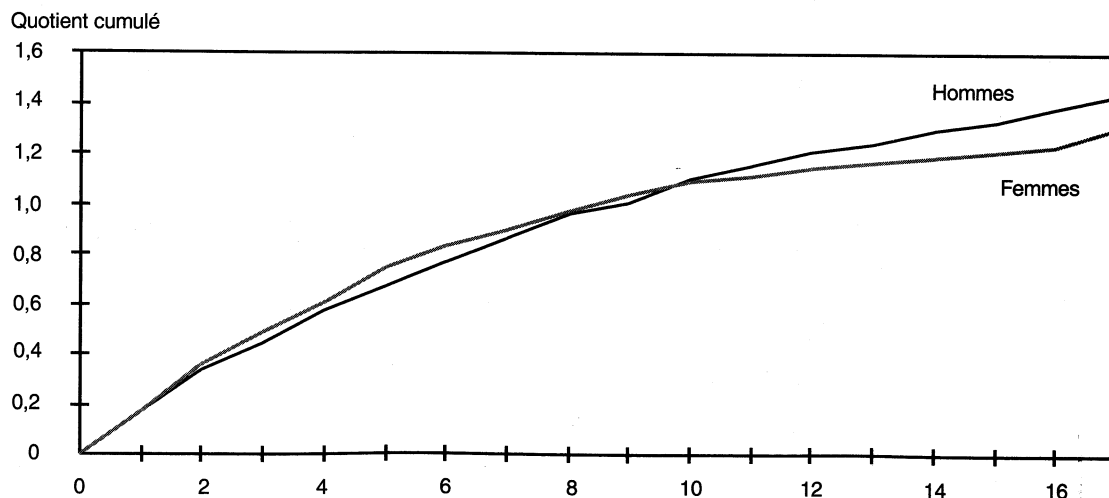
À partir des 16 217 individus de l'EDP nés en 1953-1956 en France (à l'exclusion de la Corse), 12 060 personnes mariées sont retenues pour lesquelles on dispose au moins d'un point d'observation après le mariage. Disposer d'un lieu (qui peut d'ailleurs être identique à celui du mariage) à une date ultérieure est en effet indispensable pour pouvoir mener l'estimation (6). Rappelons que dans l'EDP le nombre de dates figurant pour un individu est variable : plus nombreuses si l'individu était présent aux recensements de 1968, 1975, 1982, 1990 et s'il multiplie les actes d'état civil (mariages, naissances, etc.) ; réduit à une date pour un individu

5. L'examen détaillé de la qualité de l'EDP amène à ne pas prendre en compte les personnes nées à l'étranger, du fait du grand nombre de mariages célébrés hors de France à une date inconnue, et les personnes nées en Corse, pour lesquelles la qualité des recensements laisse à désirer, du moins dans le passé (Lery et Lenain, 1972).

6. Dans l'échantillon retenu, 30,8 % des individus n'ont pas changé de résidence depuis leur mariage.

Graphique II

Mobilité selon la durée de mariage des générations 1953-1956



Champ : individus nés de 1953 à 1956, mariés après 1968 et repérés dans l'EDP au moins 1 fois après leur mariage.
Source : trajectoires reconstituées à partir des données de l'EDP.

repéré à un seul recensement et n'ayant pas été enregistré par l'état civil (personne célibataire, par exemple).

Les premières estimations obtenues sont celles des quotients instantanés (ou risques) de migration après le mariage pour les hommes et les femmes nés entre 1953 et 1956 (cf. graphique II). Les courbes obtenues pour les hommes et les femmes ne sont pas significativement différentes, et comme on pouvait s'y attendre, la propension à migrer diminue avec la durée de mariage (Bonvalet et Lelièvre, 1991).

Ces estimations peuvent être un peu affinées en utilisant les caractéristiques individuelles disponibles dans l'EDP, on s'attend en effet à ce que la mobilité des individus varie selon leur âge, leur origine sociale, leur niveau d'éducation et leur profession (Courgeau, 1985). Malheureusement, l'EDP fournit une gamme

restreinte de variables pour ce type d'étude. Les variables de profession sont mal saisies dans les bulletins d'état civil, et les modèles utilisés ne peuvent prendre en compte que des variables fixes mesurées au même moment de la trajectoire des individus. Les déclarations concernant la profession, faites au recensement le plus proche, sont alors utilisées pour caractériser la mobilité après le mariage, un déménagement qui pour certains intervient l'année de leur mariage, et pour d'autres 10 ans après. La déclaration du niveau d'éducation faite au recensement le plus proche du mariage est également utilisée, mais on constate, grâce aux données de l'EDP, le manque de fiabilité de cette mesure qui varie d'un recensement à l'autre, alors que l'individu a terminé ses études. Quatre variables par sexe (âge au mariage, profession, niveau d'éducation et profession du père) sont retenues pour préciser la nature de la mobilité observée (cf. tableau 1). Les regroupements ont été motivés au vu de la distribution

Tableau 1
Composition socio-démographique de l'échantillon retenu

En %

Hommes		Femmes	
Variables retenues		Variables retenues	
Âge au mariage		Âge au mariage	
Moins de 20 ans	8,0	Moins de 19 ans	19,1
De 21 à 22 ans	25,5	20 ans	15,4
23 ans	15,0	21 ans	15,8
De 24 à 25 ans	21,2	De 22 à 23 ans	21,5
De 26 à 29 ans	20,4	De 24 à 25 ans	12,2
Plus de 30 ans	9,9	Plus de 26 ans	16,1
Profession		Profession	
Indépendants	5,0	Cadres, professions intellectuelles supérieures, employés du secteur public	9,0
Cadres, professions intellectuelles supérieures, employés du secteur public	8,8	Professions intermédiaires	9,0
Professions intermédiaires	11,3	Employées du secteur privé	34,0
Employés du secteur privé	7,2	Ouvrières	15,4
Ouvriers qualifiés	21,5	Étudiantes	12,1
Ouvriers non qualifiés	21,5	Inactives	20,6
Étudiants et inactifs	24,6		
Niveau d'éducation		Niveau d'éducation	
Sans diplôme	14,6	Sans diplôme	15,1
CEP	12,4	CEP	18,0
BEPC	7,8	BEPC	11,3
BEP, CAP	36,5	BEP, CAP	25,8
Baccalauréat	13,1	Baccalauréat	15,4
Diplôme universitaire	15,7	Diplôme universitaire	14,4
Profession du père		Profession du père	
Indépendants	19,4	Indépendants	19,4
Cadres et professions intellectuelles supérieures	6,0	Cadres et professions intellectuelles supérieures	5,4
Professions intermédiaires	9,8	Professions intermédiaires	9,7
Employés	9,9	Employés	10,6
Ouvriers qualifiés	19,3	Ouvriers qualifiés	20,6
Ouvriers non qualifié	15,1	Ouvriers non qualifié	16,9
Retraités	8,8	Retraités	6,5
Non renseignés	11,7	Non renseignés	10,9

Champ : individus nés de 1953 à 1956, mariés après 1968 et repérés dans l'EDP au moins 1 fois après leur mariage.
Source : EDP.

initiale et de la mobilité attendue des diverses sous-populations.

Les distributions observées sont tout à fait cohérentes par rapport à celles que donnent les enquêtes *Formation et qualification professionnelle* (FQP) de l'Insee. Les résultats des analyses qui suivent sont donc généralisables à ces générations. Néanmoins, du fait de la sélection des individus mariés, les femmes sont, en moyenne, moins diplômées que l'ensemble des femmes de ces générations et les hommes un peu plus diplômés. En gras, figure pour chaque variable la catégorie de référence : pour l'âge au mariage il s'agit de la valeur modale, et pour

les autres variables nominales de la modalité la plus fréquente, excepté pour la profession des hommes pour laquelle on a souhaité conserver une cohérence avec la profession du père. En ce qui concerne l'âge au mariage et la profession des personnes, la décomposition adoptée reflète les différences qui existent entre hommes et femmes. Ainsi les indépendantes ne sont pas assez nombreuses pour constituer une classe distincte et il est intéressant d'étudier séparément le comportement migratoire des inactives et des étudiantes. En revanche, pour le niveau d'éducation et la profession du père, les modalités retenues sont les mêmes pour les hommes et pour les femmes.

Tableau 2
La modélisation de la mobilité des hommes
A - Les facteurs de la mobilité des hommes

Variables retenues (1)	Variables séparées	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Âge au mariage					
Moins de 20 ans	1,150**		1,301**		
De 21 à 22 ans	1,063				
23 ans	1,000				
De 24 à 25 ans	1,045				
De 26 à 29 ans	0,998				
Plus de 30 ans	1,059				
Profession					
Étudiants et inactifs	1,252**	1,172**	1,183**	1,130**	
Cadres, professions intellectuelles supérieures, employés du secteur public	1,618**	1,430**	1,481**	1,551**	
Professions intermédiaires	1,276**	1,111*	1,134**	1,235**	
Employés du secteur privé	1,142*			1,112	
Ouvriers qualifiés	1,000				
Ouvriers non qualifiés	0,920*			0,928	
Indépendants	0,531**	0,572**	0,530**	0,577**	
Niveau d'éducation					
Sans diplôme	0,904*				0,899**
CEP	0,905*				0,906*
BEPC	1,064				
BEP, CAP	1,000				
Baccalauréat	1,190**	1,165**	1,189**		1,166**
Diplôme universitaire	1,499**	1,312**	1,367**		1,418**
Profession du père					
Indépendants	0,777**	0,860**		0,872**	0,771**
Cadres et professions intellectuelles supérieures	1,346**	1,100		1,216**	1,112
Professions intermédiaires	1,210**	1,132**		1,159**	1,131**
Employés	0,972				
Ouvriers qualifiés	1,000				
Ouvriers non qualifiés	0,894**	0,987***		0,985***	0,939
Retraités	1,007				

1. Pour des raisons liées aux contraintes imposées par le logiciel d'estimation, du fait de la perte de robustesse des calculs, les modèles 1 à 4 ne peuvent inclure qu'un nombre limité de variables.

* Significatif au seuil de 5 %.

** Significatif au seuil de 1 %.

Lecture : le risque de migrer d'un diplômé universitaire est 1 fois et demie (x 1,499) celle d'un homme ayant un BEP ou un CAP.

Champ : hommes nés de 1953 à 1956, mariés après 1968 et repérés dans l'EDP au moins 1 fois après leur mariage.

Source : trajectoires reconstituées à partir des données de l'EDP.

La mobilité des hommes est marquée par la profession

L'analyse des variables prises séparément montre une grande mobilité des plus jeunes encore étudiants. Le changement matrimonial pour ceux-ci correspond moins à une stabilisation de leur régime de mobilité dans la mesure où les choix professionnels ne sont pas encore effectués. Pour ceux dont la carrière professionnelle est déjà entamée, l'enracinement très fort des indépendants (artisans-commerçants, agriculteurs) (7) est conditionné par la nature de leur emploi et une sédentarisation significative, mais reste bien moindre que celui des ouvriers non qualifiés, que l'on attribue plutôt à la précarité de leur situation économique. La propension des salariés à migrer s'élève avec le milieu social. De même, les moins diplômés apparaissent moins mobiles que les autres. On peut enfin se demander si c'est une forte hérédité sociale qui conduit à un effet presque identique de la profession du père et de celle de l'enquêté sur sa mobilité géographique, et si l'un de ces effets disparaîtrait si on les considérait conjointement.

Dans les modèles 1 à 4 (cf. encadré) faisant intervenir simultanément en groupes différents les variables disponibles, l'ensemble des modalités significatives au seuil de 1 % sont maintenues et les enfants d'ouvriers non qualifiés se détachent. En particulier, quand on tient compte simultanément de la profession de l'individu

et de celle de son père, les deux effets restent chacun significatifs. Ainsi, dans le modèle 3, le fait d'être indépendant réduit la mobilité d'environ 40 % (en la multipliant par 0,577) et pour un indépendant fils d'indépendant cette tendance est encore plus affirmée puisque la réduction est de 50 % ; en revanche, si son père occupait une profession intermédiaire, sa mobilité n'est que 67 % de celle d'un ouvrier fils d'ouvrier ou d'employé (cf. tableau 2-A).

La mobilité estimée pour certains profils particuliers montre des comportements très différenciés suivant les groupes retenus (cf. tableau 2-B). Chez les indépendants, les fils d'indépendants sont les plus enracinés, de même que pour les ouvriers non qualifiés, les moins mobiles sont ceux dont la profession du père est également ouvrier non qualifié. À l'inverse, les étudiants les plus diplômés fils de cadre et les cadres sont les plus mobiles.

Les étudiantes et les inactives sont les plus mobiles

Si aucun effet significatif de l'âge des femmes au mariage n'apparaît sur leur mobilité après celui-ci, la mobilité des étudiantes après leur mariage est très forte. Peu installées, elle sont les plus susceptibles de réemménager rapidement après leur mariage dans une résidence sans doute plus propice à la constitution de leur famille. La possession d'un diplôme et l'exercice d'une profession supérieure favorisent également la mobilité de celles qui en sont munies comparé à la plus grande sédentarité des ouvrières et des non-diplômées plus captives de leur environnement.

En prenant en compte la profession de l'individu au mariage et son niveau d'éducation (modèle 1), les liaisons qui existent entre ces deux caractéristiques ressortent. Les comportements résidentiels des femmes peu et très diplômées s'opposent, avec une confirmation de l'enracinement des premières, les plus mobiles apparaissant être les étudiantes et les inactives. Le modèle 3 confirme ces oppositions (cf. tableau 3).

Parmi les étudiantes, les plus mobiles sont celles dont la famille est en mesure de les aider matériellement. Leur première résidence conjugale ne correspond pas à une installation à long terme car elles débutent leur vie de femme mariée tout en étant engagées dans des études longues. À l'inverse, la sédentarité, que l'on

7. Les effectifs réduits de ces deux groupes ne permettent pas de les distinguer dans l'analyse.

Tableau 2 (suite)

B - La mobilité de quelques profils d'hommes

Modèle avec interactions	Effets
Étudiants, père cadre, professions intellectuelles supérieures, professions intermédiaires ayant le baccalauréat ou plus	1,699**
Cadres, professions intellectuelles supérieures, employés du secteur public	1,618**
Professions intermédiaires	1,276**
Autres étudiants	1,190**
Employés du secteur privé	1,142*
Ouvriers qualifiés	1,000
Autres ouvriers	0,950
Ouvriers non qualifiés, père ouvrier non qualifié, sans diplôme ou CEP	0,750**
Autres indépendants	0,563**
Indépendants, père indépendant, BEP/CAP	0,421**

* Significatif au seuil de 5 %.

** Significatif au seuil de 1 %.

Lecture : le risque de migrer d'un employé du secteur privé est légèrement supérieur à celui d'un ouvrier qualifié mais le double d'un indépendant.

Champ : hommes nés de 1953 à 1956, mariés après 1968 et repérés dans l'EDP au moins 1 fois après leur mariage.

Source : trajectoires reconstituées à partir des données de l'EDP.

peut supposer contrainte, des ouvrières les moins diplômées se confirme. Les femmes occupant une profession intermédiaire (souvent dans l'enseignement et la santé) ont également une mobilité plus importante. Les inactives, qui ne forment pas un groupe homogène, ont la mobilité de leur époux. Celles dont le niveau d'éducation est moyen ou élevé et mariées à des hommes très mobiles favorisent même la mobilité du ménage puisque les problèmes de relocalisations professionnelles ne se cumulent pas.

Vers une meilleure fiabilité des informations géographiques

Le suivi longitudinal des trajectoires résidentielles individuelles à partir des données de

l'EDP se révèle malgré tout difficile à conduire sur un tel fichier, en dépit de sa vocation. Les données requièrent une préparation importante avant de se prêter à l'analyse et celle-ci ne peut ensuite être conduite qu'avec des méthodes de durée spécialement adaptées pour les parcours fragmentaires. De plus, le choix de caractéristiques individuelles pour une telle analyse est réduit, ce qui restreint fortement la portée des résultats obtenus.

Les problèmes de repérage de la localisation des individus, inhérents au fichier de données, ne sont en aucun cas engendrés par l'appariement des différentes sources, mais résultent plutôt de la nature des données (recensement, état civil) et de la qualité très inégale des informations sur la localisation géographique.

Tableau 3
Les facteurs de mobilité des femmes

Variables retenues (1)	Variables séparées	Modèle 1	Modèle 2	Modèle 3
Âge au mariage				
Moins de 19 ans	1,071			
20 ans	0,969			
21 ans	1,000			
De 22 à 23 ans	0,984			
De 24 à 25 ans	0,961			
Plus de 26 ans	0,947			
Profession				
Cadres, professions intellectuelles supérieures, employées du secteur public	1,177**	1,139**	1,169**	
Professions intermédiaires	1,235**	1,101	1,214**	
Employées du secteur privé	1,000			
Ouvrières	0,881**	0,968	0,889**	
Étudiantes	1,402**	1,303**	1,377**	
Inactives	1,216**	1,262**	1,218**	
Niveaud'éducation				
Sans diplôme	0,841**	0,817**		0,828**
CEP	0,856**	0,838**		0,843**
BEPC	1,067			
BEP, CAP	1,000			
Baccalauréat	1,118**	1,064		1,088*
Diplôme universitaire	1,318**	1,182**		1,270**
Profession du père				
Indépendants	0,948			
Cadres et professions intellectuelles supérieures	1,272**		1,166**	1,089
Professions intermédiaires	1,113*		1,080	1,047
Employés	1,012			
Ouvriers qualifiés	1,000			
Ouvriers non qualifiés	0,971			
Retraités	0,973			
Non renseignés	1,060			
1. Pour des raisons liées aux contraintes imposées par le logiciel d'estimation, du fait de la perte de robustesse des calculs, les modèles 1 à 3 ne peuvent inclure qu'un nombre limité de variables.				

* Significatif au seuil de 5 %.

** Significatif au seuil de 1 %.

Lecture : le risque de migrer d'une ouvrière est moindre (x 0,881) que celle d'une employée du secteur privé.

Champ : femmes nées de 1953 à 1956, mariées après 1968 et repérées dans l'EDP au moins 1 fois après leur mariage.

Source : trajectoires reconstituées à partir des données de l'EDP.

Au contraire, l'appariement que propose l'EDP, en confrontant les divers lieux de résidence, déclarés et recueillis, permet de con-

clure sur la fiabilité des informations géographiques contenues dans les recensements et les bulletins d'état civil. □

BIBLIOGRAPHIE

Bonvalet C. et Lelièvre E. (1991), " Nuptialité et mobilité " in *La nuptialité : évolution récente en France et dans les pays développés*, T. Hibert et L. Roussel (Eds), INED/PUF, Paris, pp. 237-253.

Coeffic N. (1993), " L'enquête post-censitaire de 1990. Une mesure de l'exhaustivité du recensement ", *Population*, vol. 48, pp. 1655-1682.

Courgeau D. (1985), " Changements de logement, changements de département et cycle de vie ", *L'Espace Géographique*, n°4, pp. 289-306.

Courgeau D. et Lelièvre É. (1989), *Analyse démographique des biographies*, Éditions de l'Ined, PUF, 268 p.

Courgeau D., Lelièvre É. et Wolber O. (1998), *Localisation des individus au cours de leur vie : tests pour l'utilisation de l'Échantillon démographique permanent*, à paraître, Coll. Insee Méthodes.

Courgeau D. et Najim J. (1995), " Analyse de biographies fragmentaires ", *Population*, vol. 50, pp. 149-168.

De Gruttola V.G. et Lagakos S.W. (1989), " Analysis of Doubly-Censored Survival Data with Application to AIDS ", *Biometrics*, vol. 45, pp. 1-11.

Hoem J. (1985), " Weighting Misclassification and other Issues in the Analysis of Survey Samples of Life Histories ", *Longitudinal Analysis of LabourMarket Data*, Heckman & Singer eds., Cambridge University Press, pp. 249-293.

Kim M.Y., De Gruttola V.G. et Lagakos S.W. (1993), " Analysing Doubly-Censored Data with Covariates, with Application to AIDS ", *Biometrics*, vol. 49, pp. 13-22.

Lery A. et Lenain J. (1972), " La population de la Corse : 200 000 ou 300 000 habitants ", *Économie et Statistique*, n° 32, pp. 53-59.