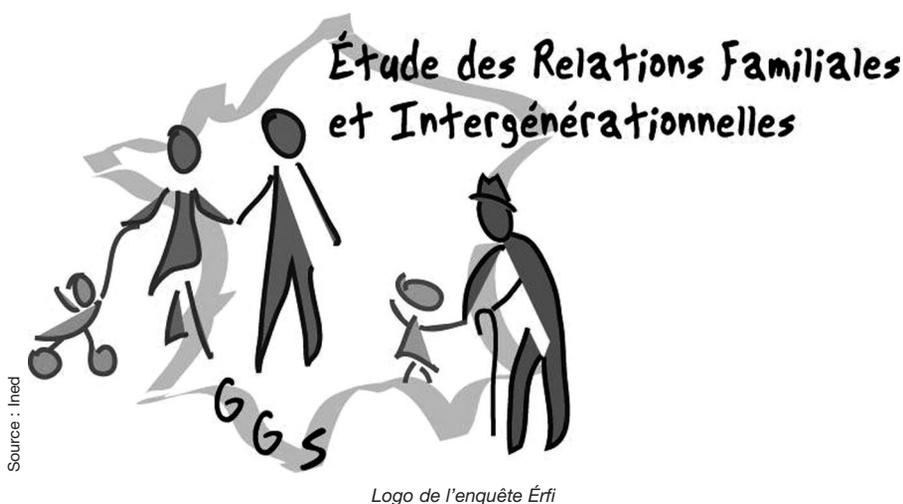


La déperdition dans les enquêtes longitudinales : l'exemple de l'enquête « Étude des relations familiales et intergénérationnelles »

✍ Arnaud Régnier-Loilier*

L'Ined et l'Insee ont réalisé à l'automne 2005 la première vague de l'enquête Érfi (Étude des relations familiales et intergénérationnelles). Les mêmes personnes ont été réinterrogées trois ans après, à l'automne 2008 et une dernière vague est prévue fin 2011. Comme toute enquête répétée dans le temps, l'échantillon initial d'environ 10 000 personnes a subi une érosion relativement importante et seules 6 500 personnes ont répondu au second questionnaire, soit un taux d'attrition de 35 %. Pour prévenir au mieux ce phénomène, une attention particulière doit être portée au suivi de l'échantillon, notamment par le biais d'une gestion rigoureuse des « fiches-adresses » et la recherche systématique d'une « personne-relais ».



Présentation de l'enquête « Étude des relations familiales et intergénérationnelle » (Érfi)

Sur ce modèle d'enquêtes, la Population Activities Unit de la Commission économique des Nations-Unies pour l'Europe (Unece) a engagé à partir de 2000 le *Generations and Gender Programme* (Programme générations et genre, GGP). Il s'agit – entre autres – de promouvoir à l'échelle internationale la réalisation d'une même enquête sur le thème général de la famille (pour plus de détails : Vikat et al., 2005). Trois vagues sont prévues, chacune espacée de trois ans, dans le but de mieux comprendre et de manière beaucoup plus précise qu'au travers d'enquêtes transversales l'ordonnement des événements démographiques et biographiques. L'approche longitudinale de l'enquête *Generations and Gender Survey* (enquête « générations et genre », GGS) ouvre en outre la possibilité d'étudier la corrélation entre, d'une part, l'évolution

Les enquêtes « panel » qui consistent à interroger les mêmes personnes à plusieurs reprises (avec des périodicités variables selon les cas : chaque trimestre, chaque année, tous les trois ans) se sont largement développées en France au cours des dernières années. D'abord très répandues dans les pays anglo-saxons, et souvent associées au champ médical et épidémiologique, les études longitudinales ont trouvé d'autres champs d'application en sciences sociales notamment, dans des thématiques variées. On peut citer, pour la période récente et à titre d'exemples, l'en-

quête « Génération 98 » menées par le Céreq¹ afin de mieux suivre le parcours d'insertion des jeunes sortis de formation initiale (enquête en trois vagues chacune espacée de deux ans), l'enquête « Intentions de fécondité » réalisée par l'Ined² entre 1998 et 2003 (en trois vagues) afin de mieux comprendre les obstacles à la réalisation des intentions de fécondité ou encore l'enquête « Cocon » (cohorte contraceptive entre 2000 et 2003, interrogation annuelle) mise en place par l'Inserm³ et visant à retracer l'évolution des pratiques contraceptives au cours de la vie.

* Institut national des études démographiques (Ined)

1. Centre d'études et de recherche sur les qualifications

2. Institut national d'études démographiques

3. Institut national de la santé et de la recherche médicale

du niveau de revenu, des valeurs et opinions (par exemple) et, d'autre part, les comportements. On ne peut en effet demander de manière rétrospective aux personnes ce qu'elles pensaient ou ce qu'elles gagnaient trois ou cinq ans auparavant : la qualité de l'information serait discutable dans la mesure où les personnes ne peuvent se souvenir avec exactitude de leurs revenus ou de leurs opinions quelques années auparavant (avec, pour les opinions et les valeurs, un risque important de rationalisation *a posteriori*). Enfin, la dimension prospective permet d'inclure un certain nombre de questions d'intentions (de déménager, de se mettre en couple, de se marier, d'avoir un enfant, de changer de situation professionnelle, etc.) et de les confronter aux comportements observés dans les années qui suivent. Il est alors possible de repérer les facteurs et événements qui freinent ou au contraire facilitent la réalisation de ces intentions.

En collaboration avec l'Insee, l'Ined a pris part à ce projet en mettant en place « l'étude des relations familiales et intergénérationnelles » (Érfi), déclinaison française du questionnaire et protocole d'enquête GGS (Unece, 2005 ; Régnier-Loilier, 2006b). Construite sur le même principe qu'une enquête « ménage », la première vague a eu lieu à l'automne 2005 et a permis d'interroger 10 079 personnes âgées de 18 à 79 ans. Il était alors prévu de réinterroger les mêmes personnes trois et six ans plus tard. La seconde vague de l'enquête a eu lieu à l'automne 2008.

L'attrition, un enjeu majeur dans les enquêtes longitudinales

Toutefois, la mise en place de ce type d'enquêtes n'est pas simple.

4. On se limite ici aux « grandes catégories » de sortie d'échantillon. D'autres « fuites » sont possibles : on peut par exemple évoquer des erreurs de saisie qui ne permettent plus de faire le lien entre un répondant et ses coordonnées, le cas de personnes qui ne sont pas disponibles au moment où la collecte a lieu, etc. Nous détaillons ces différentes catégories de sortie plus loin.

Outre le coût qu'elles engendrent, elles posent un certain nombre de problèmes d'ordre méthodologique dont l'un des principaux est l'attrition, c'est-à-dire la perte de la possibilité d'interroger l'ensemble des personnes sur plusieurs vagues. En effet, le suivi d'une même population sur une longue période est difficile, et ce pour plusieurs raisons. En premier lieu, on court le risque à chaque vague qu'un certain nombre de personnes sortent du panel en refusant de participer à un nouveau questionnaire (les questions peuvent leur avoir déplu, les avoir choqués, l'interview peut leur avoir semblé trop longue, etc.). En second lieu, cette population est sujette à un ensemble d'événements qui peut les conduire à déménager : mutation professionnelle, arrivée d'un enfant, séparation, départ en institution. Si certains pays peuvent s'appuyer sur les registres de population pour actualiser l'adresse des personnes, cela n'est pas envisageable partout. En France, il n'est par exemple pas possible de se baser sur les fichiers actualisés de la Sécurité sociale ou des impôts pour tenir à jour le carnet d'adresses. Enfin, comme toute population, le panel est sujet au risque de décès⁴.

Les différentes possibilités de sortie d'échantillon (refus, perte d'adresse ou décès) ne touchent pas de manière aléatoire l'ensemble de la population enquêtée à la première vague : les décès touchent davantage les personnes plus âgées, les pertes d'adresses sont plus fréquentes chez les personnes les plus mobiles (notamment les plus jeunes), les refus émanent davantage de personnes craintives ou ne s'étant pas reconnues dans la thématique de l'enquête. Dans la mesure où la déperdition de l'échantillon est sélective, on court alors le risque que les résultats issus des analyses longitudinales soient biaisés. Si l'attrition observée dans l'enquête « Cocon » (citée précédemment), qui s'est élevée à 1/3 de l'échantillon initial après trois vagues, ne semble pas occasionner de biais significatifs sur les principales variables d'intérêt (Razafindratsima *et al.*, 2004), les conséquences sont en revanche plus

marquées pour l'enquête « Intentions de fécondité », où l'attrition a atteint près de 70 % après cinq ans, invitant à la prudence lors de l'exploitation des données longitudinales (Mazuy *et al.*, 2005).

La maîtrise de l'attrition représente donc un enjeu capital dans la réussite d'une opération de collecte en longitudinal. Il est alors nécessaire de tout mettre en œuvre afin de limiter cette déperdition au maximum entre chaque vague. Par ailleurs, il est indispensable de connaître les caractéristiques des personnes principalement touchées et d'en évaluer les incidences possibles sur les résultats. Cela peut d'abord permettre le calcul d'une variable de redressement pour les exploitations statistiques à partir des données longitudinales. En outre, mieux connaître les facteurs de risques d'attrition peut permettre de mieux anticiper (pour de futures enquêtes) certaines pertes et ainsi mettre en place des stratégies adaptées afin de disposer au final d'un panel qui reste représentatif. On peut par exemple surreprésenter certaines catégories de personnes lors du tirage de l'échantillon initial afin de disposer de suffisamment de répondants lors de la dernière vague ; on peut aussi imaginer un suivi plus attentif de certaines catégories de répondants, etc.

On se donne ici pour objectif de caractériser l'attrition subie par l'échantillon de l'enquête Érfi entre 2005 et 2008, sur la base de critères essentiellement sociodémographiques. En revanche, l'évaluation des biais occasionnés par la déperdition sur les variables d'intérêt de l'enquête ne sera pas évoquée ici.

Différentes raisons de l'attrition entre 2005 et 2008

Au terme du questionnaire de la première vague de l'enquête Érfi (2005), l'enquêteur demandait à chaque répondant : « [...] Nous envisageons de réinterroger dans trois ans les personnes qui ont répondu à ce ques-

Tableau 1 – Détail de l'attrition entre les deux premières vagues de l'enquête Érfi (2005-2008)

	Nombre	%
REpondants à la vague 1 (2005)	10 079	100,0
Refus de participer au Panel	980	9,7
Perdus de vue entre les deux vagues	758	7,5
ECHANTILLON 2008	8 341	82,8
Déchets lors de la seconde vague	1 547	15,3
<i>dont :</i>		
impossible à joindre	221	2,2
entretien impossible à réaliser	190	1,9
refus de participer à la seconde vague	794	7,9
aucun membre du ménage ne correspond à la personne interrogée en 2005	16	0,2
entretien impossible à réaliser afin la fin de la période de collecte	77	0,8
répondant parti sans laisser d'adresse	249	2,5
Hors champ lors de la seconde vague	210	2,1
<i>dont :</i>		
impossible d'accéder à l'adresse	21	0,2
personne décédée	75	0,7
répondant parti en institution	17	0,2
répondant parti à l'étranger	52	0,5
fiche-adresse non traitée	45	0,4
Enquêtes incomplètes	8	0,1
ENQUÊTES RÉALISÉES À LA VAGUE 2 (2008)	6 576	65,2

Source : Ined-Insee, Érfi-GGS2, 2008

tionnaire. Accepteriez-vous que l'Ined reprenne de nouveau contact avec vous dans trois ans, sans que cela représente un engagement de votre part ? ». Si la personne acceptait, l'enquêteur devait alors compléter et lui faire signer une « fiche de suivi » sur laquelle il devait reporter ses coordonnées et, éventuellement, celles d'une ou deux « personnes-relais » (en cas de perte de contact, on peut alors tenter de d'obtenir la nouvelle adresse du répondant *via* ces personnes-relais).

Sur les 10 079 répondants à la première vague de l'enquête Érfi, 11 % ont refusé que l'on garde contact en vue d'une seconde interrogation. Après l'enquête, un courrier de remerciement avait été adressé à chacun, dans lequel on demandait aux personnes n'ayant pas accepté le suivi si « vraiment » elles ne souhaitaient pas participer à la seconde vague. Cette relance a permis de récupérer quelques personnes. Le taux de refus a donc été *in fine* légèrement inférieur à 10 % (9 099 répondants).

Le service des enquêtes de l'Ined était en charge d'assurer le « suivi » du panel entre 2005 et 2008 en mettant

régulièrement à jour les adresses. À cette fin, cinq courriers ont été adressés aux répondants, soit un tous les six mois : lettre de remerciements ; premiers résultats sous forme d'un « 4 pages » (Régnier-Loilier, 2006a) puis seconds résultats sous forme d'un « 8 pages » (Bauer, 2007) afin de les intéresser à l'étude ; carte de vœux pour l'année 2008 ; annonce de la troisième vague de l'enquête. Ces envois devaient permettre de s'assurer que les personnes habitaient toujours la même adresse⁵, et fournissaient l'occasion de leur rappeler de nous transmettre leurs nouvelles coordonnées en cas de déménagement. En outre, un site Internet avait été mis en place avant le début de la collecte pour donner une « visibilité » à l'enquête et favoriser la participation au panel (<http://www-erfi.ined.fr>).

Au terme des trois années de suivi, l'échantillon constitué avant la seconde vague comptait encore 8 341 fiches-adresse (tableau 1). Ainsi, le contact a été perdu avec un peu plus de 750 personnes entre 2005 et 2008. Dans 70 % des cas, ces « perdus de vue » n'avaient pas donné de personnes-relais (contre 45 % pour les personnes encore dans le panel) et

nous ne disposions pas de coordonnées téléphoniques pour 54 % d'entre eux (contre 8 %, respectivement), limitant ainsi les chances de retrouver les coordonnées du répondant⁶. Enfin, lors du terrain à l'automne 2008 (seconde vague), seuls 6 576 répondants ont été interrogés. Le décalage entre la base d'adresses et le nombre d'entretiens réalisés correspond principalement à des refus de participer à la seconde vague et à des personnes sorties du champ de l'étude (décédées, parties à l'étranger) (tableau 1). Le taux d'attrition total s'élève donc, entre les deux premières vagues de l'enquête Érfi, à 35 % (65 % de taux de réponse final).

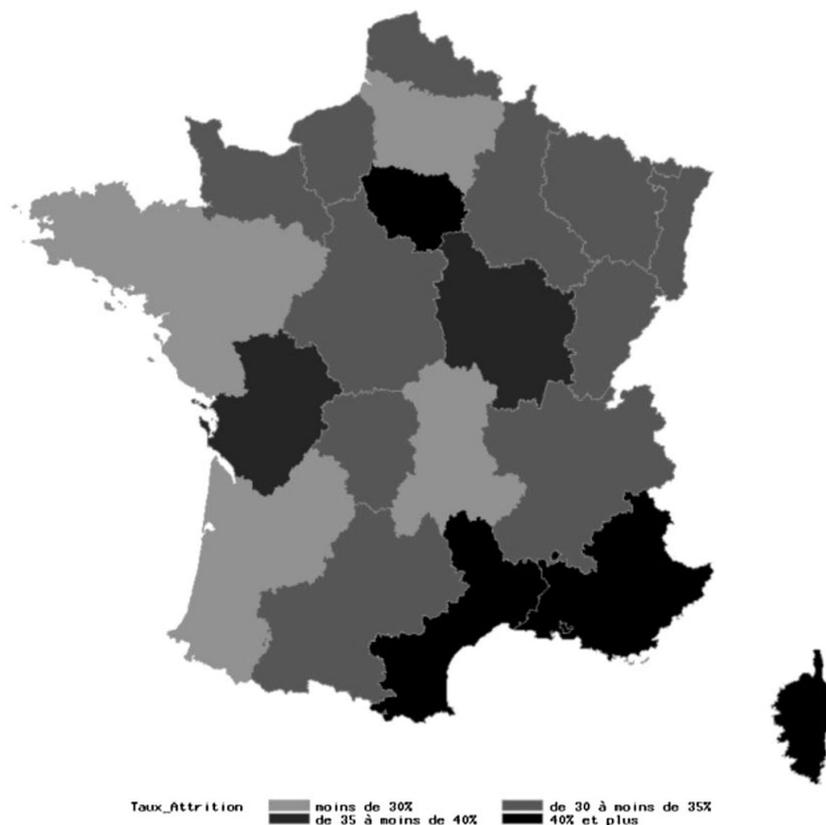
Cette attrition peut être décomposée ainsi :

- pour 10 % de l'échantillon initial, le répondant a refusé de poursuivre l'enquête au terme de la première vague ;
- dans 10 % des cas, soit la personne a été perdue de vue entre 2005 et 2008, soit elle n'habitait plus à l'adresse indiquée lorsque l'enquêteur s'y est rendu à l'automne 2008 ;
- dans 15 % des cas enfin, le répondant a refusé de participer à la seconde vague, était hors champ (décédé, parti en institution, à l'étranger, etc.) ou encore était absent lorsque l'enquêteur s'est présenté à son domicile.

5. En cas de déménagement, le courrier peut en effet être acheminé directement par la Poste à la nouvelle adresse de la personne ou, si la personne n'a pas contracté de procédure de suivi automatique de son courrier, revenir à l'expéditeur avec la mention « NPAI » (n'habite plus à l'adresse indiquée). Nous pouvions alors entamer une procédure de recherche des nouvelles coordonnées du répondant, soit en le contactant directement (s'il avait accepté de laisser son numéro de téléphone), soit par Internet, soit enfin en contactant les personnes-relais lorsque nous en disposions. Si aucune de ces démarches n'avaient été fructueuses, nous renvoyions alors une nouvelle fois un courrier à la dernière adresse connue. Si le courrier revenait de nouveau avec la mention « NPAI », le répondant était considéré comme « perdu » et retiré de la base de suivi.

6. Il est à noter que les « personnes-relais » n'étaient pas « suivies » et leur adresse n'était donc pas actualisée durant l'entre-deux vagues.

Figure 1 – Taux d’attrition totale entre 2005 et 2008, par région



Source : Ined-Insee, enquête Érfi-GGS1-2, 2005-2008

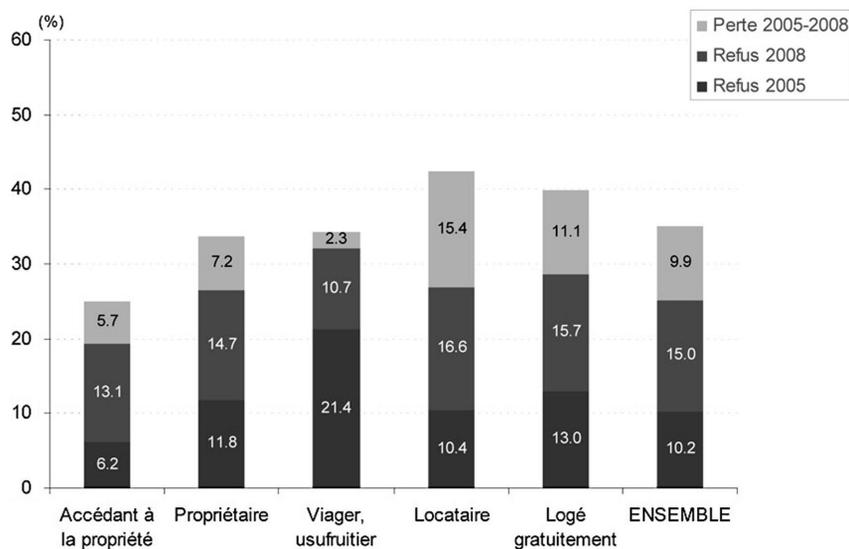
Caractéristiques des personnes perdues entre les deux vagues

L'échantillon de l'enquête Érfi était limité à la France métropolitaine. L'attrition a tout d'abord touché différemment les régions (figure 1), dépassant 40 % en Corse, en Languedoc-Roussillon, en région Provence-Alpes-Côte d'Azur et en Île-de-France. La région parisienne cumule les handicaps avec un accès aux immeubles plus difficile pour les enquêteurs ainsi qu'une forte mobilité professionnelle ; les régions méditerranéennes comptent pour leur part davantage de résidences secondaires et se caractérisent en outre par la plus forte part de pertes d'adresses entre les deux vagues. À l'inverse, quatre régions enregistrent une attrition nettement plus faible, inférieure à 30 % : Pays de la Loire, Picardie, Bretagne et Auvergne.

En lien avec la distribution régionale, l'attrition a été plus faible dans les petites communes avec un taux inférieur à 30 % dans les localités de moins de 5 000 habitants. Ceci s'explique en grande partie par des pertes de contact entre les deux vagues beaucoup plus rares comparativement aux plus importantes communes.

L'attrition est par ailleurs extrêmement élevée auprès des personnes de nationalité étrangère, en particulier chez les maghrébins où elle atteint 65 %. Le taux de refus de participer au panel en 2005 est d'abord plus élevé pour ces populations, de même que les pertes d'adresses. Il est à noter que beaucoup n'avaient pas donné de personnes-relais, ceci expliquant en partie la plus forte proportion de pertes de contact. Mais la plus forte attrition observée chez les étrangers n'a au final que des conséquences limitées dans la mesure où ils ne représentent que 4 % de l'échantillon.

Figure 2 – Taux d’attrition selon le statut d’occupation



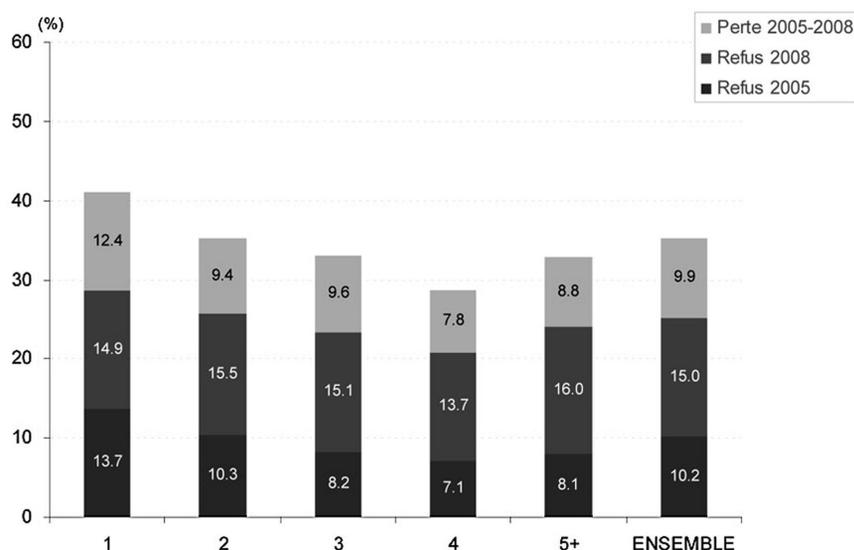
Source : Ined-Insee, enquête Érfi-GGS1-2, 2005-2008

Ces trois cas de figure seront considérés séparément dans la description des caractéristiques des personnes perdues de vue entre les deux vagues dans la mesure où ils

révèlent des comportements bien distincts : le refus de participer à une enquête n'a en effet pas le même sens que le fait de changer de logement.

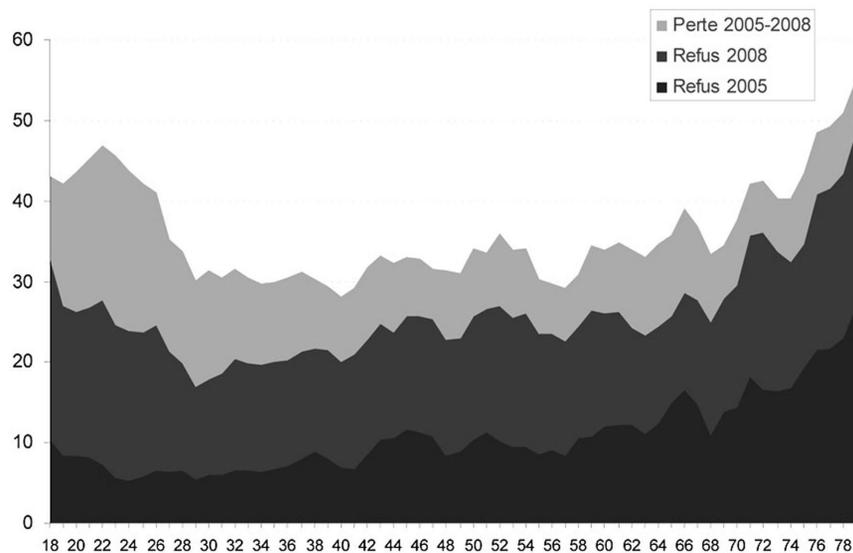
La déperdition apparaît en outre étroitement liée au statut d'occupation du logement (figure 2). Les accédants à la propriété, dont on peut penser qu'ils sont installés dans leur loge-

Figure 3 – Taux d'attrition selon le nombre d'habitants du logement



Source : Ined-Insee, enquête Érfi-GGS1-2, 2005-2008

Figure 4 – Taux d'attrition par âge (moyennes mobiles d'ordre trois)



Source : Ined-Insee, enquête Érfi-GGS1-2, 2005-2008

Note : moyennes mobiles d'ordre trois (moyennes de trois âges successifs, centrées) sauf aux extrêmes (18 = moyenne des valeurs 18 et 19 ans ; 79 = moyenne des valeurs 78 et 79 ans)

ment depuis peu et envisagent moins un déménagement à court terme, enregistrent le plus faible taux d'attrition, en partie dû à des pertes de contact relativement rares entre 2005 et 2008. À l'opposé, les perdus de vue sont beaucoup plus nombreux chez les locataires, conduisant à un taux d'attrition total supérieur à 40 %.

Ces résultats sont finalement peu surprenants, en particulier lorsqu'on les met en perspective avec le risque de mobilité géographique. En effet, 87 % des propriétaires et accédants à la propriété ont déclaré en

2005 qu'ils n'avaient pas l'intention de déménager dans les trois années à venir contre 42 % des locataires ou personnes logées gratuitement. Ce résultat invite donc, pour le suivi des personnes entre 2008 et 2011 (2^e et 3^e vague de l'enquête), à porter une attention toute particulière auprès des personnes ayant l'intention de déménager. On pourrait imaginer, par exemple, un suivi un peu plus régulier pour cette population. La mobilité résidentielle, d'autant plus simple à mettre en pratique qu'elle concerne des ménages de petite taille, a en outre pour conséquence une déper-

dition plus élevée chez les ménages d'une seule personne. Deux effets se conjuguent : plus le ménage compte d'habitants, moins le taux de refus de participer au panel en 2005 et moins le taux de perte d'adresse ont été élevés (figure 3). En lien avec ces résultats, on observe une attrition nettement plus forte auprès des personnes vivant sans conjoint (41 %) qu'auprès des personnes en couple, mariées ou non (32 %).

Ces constats sont en fait à mettre en relation avec la plupart des caractéristiques sociodémographiques des répondants, en particulier leur âge. Si la déperdition diffère peu selon le sexe du répondant (tant en termes de niveaux qu'en termes de raisons), elle est en revanche particulièrement distincte selon son âge (figure 4). Elle est en effet très élevée aux deux extrêmes de la pyramide des âges (en 2005), mais pour des raisons très différentes. Aux jeunes âges, l'acceptation de participer au panel en 2005 est très bonne mais l'attrition tient davantage à des pertes de contact entre les deux vagues. Il s'agit en effet de la population la plus « mobile » : les personnes interrogées alors qu'elles vivaient encore chez leur(s) parent(s) ont pu décohabiter, soit pour aller suivre des études dans une autre ville, soit pour des raisons professionnelles ou encore pour se mettre en couple. En revanche, chez les personnes plus âgées, le taux d'attrition tient en grande partie au refus de participer au panel dès 2005 et également à des sorties d'échantillon « par décès » ou « hors champ » (départ en institutions) plus fréquentes.

L'effet de l'âge se retrouve au travers du niveau de diplôme. Les personnes encore en études en 2005 (donc les plus jeunes) ont connu une déperdition très élevée, en grande partie due à des pertes de contact entre 2005 et 2008, mais aussi à des « déchets », catégorie qui rappelle un englobe un ensemble de situations diverses (aucun habitant du logement ne correspond au répondant de 2005, absence de longue durée, parti à l'étranger, etc.). Les répondants titulaires d'aucun diplôme (mais ayant terminé leurs études) sont

à l'inverse en moyenne plus âgés. On enregistre pour eux un taux d'attrition également très élevé mais que l'on doit avant tout à un taux de refus de participer au panel en 2005 beaucoup plus important que la moyenne.

L'attrition est donc sans surprise beaucoup plus forte parmi les inactifs (les étudiants et retraités y étant très présents). Elle est en revanche plus faible chez les cadres et professions intellectuelles supérieures, chez les professions intermédiaires et chez les agriculteurs, ces derniers étant plus sédentaires – les pertes d'adresses entre 2005 et 2008 sont rares chez eux, inférieures à 5 %.

Vue d'ensemble, « toutes choses égales par ailleurs »

Les résultats descriptifs montrent un effet marqué de l'âge mais aussi des variables qui y sont corrélées. Afin de mesurer l'effet propre de chacune des variables, une modélisation (régression logistique) a été mise en place afin d'estimer la probabilité d'attrition totale entre les deux vagues, « toutes choses égales par ailleurs » (résultats non présentés ici).

Outre les facteurs de risque évoqués précédemment, on tient compte de l'effet de deux variables supplémentaires sur la déperdition : l'état de santé perçu du répondant en 2005 – un certain nombre de refus de participer à la seconde vague parmi des personnes ayant accepté de participer au panel en 2005 sont en effet parvenus au service des enquêtes pour ce motif) ; le fait d'avoir ou

non accepté de signer « l'autorisation exprès » au terme du questionnaire de 2005. On devait en effet obtenir le « consentement exprès » des personnes, au sens de l'article 8 de la loi informatique et libertés, pour enregistrer leurs réponses à deux questions spécifiques qualifiées de « sensibles » par le comité du label⁷ : la religion d'origine ou d'appartenance (catholique, musulman, etc.) et le fait d'être ou non pacsé. Bien que l'on ait rappelé aux enquêtés que les données étaient confidentielles et qu'aucun lien ne pouvait être fait entre les informations nominatives et les réponses au questionnaire, le fait de devoir signer un document sur lequel figurait leur nom et prénom a conduit 6 % des personnes à refuser l'enregistrement de leurs réponses à ces deux questions (alors que dans la très grande majorité des cas, ils avaient répondu à ces questions). Le refus de signer ce document pourrait être lié au refus de participer au panel : il peut s'agir d'un refus « de principe » de signer tout type de documents sur lesquels sont portées des informations nominatives. Enfin, le niveau de diplôme n'a pas été intégré dans la modélisation pour des raisons de colinéarité trop forte avec la catégorie socioprofessionnelle.

La plupart des résultats présentés précédemment sont confirmés, toutes choses égales par ailleurs. Toutefois, l'influence du sexe du répondant se révèle significative, avec une probabilité d'attrition légèrement plus faible chez les femmes. En outre, les personnes vivant seules, pour lesquelles le taux brut d'attrition totale est plus élevé, n'ont pas connu une attrition plus forte comparativement à un adulte vivant en couple sans enfant une fois les autres variables prises en compte (en particulier dès lors que l'on tient compte de l'âge).

Par ailleurs, les deux variables ajoutées dans la modélisation (état de santé perçu en 2005 et signature du consentement exprès) jouent un rôle important sur le risque de déperdition. La probabilité estimée d'attrition atteint 50 % pour les personnes s'estimant en mauvaise santé (contre 31 % en moyenne pour la « personne

de référence »⁸). Et elle atteint 66 % pour les personnes ayant refusé de signer l'autorisation exprès. Sur ce dernier point, on peut penser que la corrélation observée entre l'attrition et l'autorisation écrite d'enregistrement des réponses aux questions « Pacs » et « Religion » est la traduction d'une attitude de principe ou de défiance de la part du répondant vis-à-vis de documents écrits à compléter et à signer. On ne peut pas conclure de ce résultat que la procédure d'autorisation exprès a favorisé l'attrition des personnes ayant refusé de signer le document, et ce d'autant plus que les concepteurs de l'enquête avaient pris soin de demander d'abord au répondant s'il acceptait de participer au panel ; seulement après, on lui rappelait que quelques questions étaient dites « sensibles » et nécessitaient son accord écrit.

Conséquences de l'attrition dans la gestion des cohortes

L'attrition observée entre les deux premières vagues de l'enquête Érfi n'a donc pas été aléatoire mais a davantage touché certaines catégories de répondants. Les personnes seules, hommes, jeunes ou personnes âgées, de nationalité étrangère, résidant en Île-de-France ou dans les régions méditerranéennes, locataires, inactives, en mauvaise santé ont été moins nombreuses à participer à la seconde vague, résultats relativement « conformes » à ceux observés dans d'autres enquêtes longitudinales (Razafindratsima *et al.*, 2004 ; Mazuy *et al.*, 2005 ; Couppié *et al.*, 2007).

S'il n'est pas envisageable d'individualiser le suivi d'une cohorte, on pourrait cependant imaginer de suivre à intervalle plus régulier certaines personnes pour lesquelles le risque de « perte de contact » est plus important. Toutefois, la mise en place d'une telle procédure nécessiterait de pouvoir faire le lien entre les informations nominatives des personnes (adresses) et certaines de leurs réponses au questionnaire. Or ce rapprochement n'est pas permis par la

7. L'information relative au Pacs était en 2005 qualifiée de « sensible » parce que susceptible de décrire « l'orientation sexuelle des personnes ». Pourtant, déjà en 2005, les statistiques montraient que le Pacs n'était pas majoritairement le fait de couples homosexuels. Depuis, la question du Pacs n'est plus sujette à l'autorisation exprès.

8. C'est-à-dire un répondant réunissant l'ensemble des caractéristiques suivantes en 2005 : homme âgé de 30 à 34 ans, résidant en Île-de-France, propriétaire de son logement, ayant accepté de signer l'autorisation exprès au terme de la collecte en 2005, s'estimant en très bonne santé, n'ayant « certainement pas » l'intention de déménager, vivant en couple sans enfant, de nationalité française et employé.

Cnil, afin de garantir la confidentialité des réponses.

Cependant, on a vu que les personnes perdues de vues entre 2005 et 2008 étaient aussi plus souvent des répondants pour lesquels nous ne disposons pas de coordonnées de personnes-relais ou de numéro de téléphone. Ces informations étant liées aux adresses des répondants, on pourrait par contre envisager, entre la seconde et la troisième vague de l'enquête Érfi (prévue en 2011), un suivi plus régulier des personnes pour lesquelles on ne dispose d'aucune information complémentaire. En effet, en cas de déménagement, le courrier est dans bien des cas transféré automatiquement à la nouvelle adresse du répondant, mais la durée de validité de ce transfert est souvent limitée à quelques mois. Ainsi, rapprocher l'envoi de courriers (lesquels rappellent à la personne de nous transmettre sa nouvelle adresse en cas de

déménagement, soit par téléphone, mail ou courrier) augmenterait les chances d'obtenir les nouvelles coordonnées du répondant. Mais il reste difficile d'estimer quelle devrait être la « bonne fréquence » d'envoi pour optimiser le suivi, tout en évitant de « harceler » le répondant.

Il convient en tout état de cause d'insister lors de la formation d'enquêteurs sur le fait qu'il est essentiel de disposer des coordonnées d'au moins une personne-relais, en leur expliquant l'importance fondamentale de cette information sur le suivi. Les conséquences d'une déperdition sélective de l'échantillon peuvent en effet être importantes pour les analyses statistiques. Si le calcul de pondérations ou variables de redressement peut permettre de corriger l'attrition, l'exploitation des données ne reste pertinente et envisageable que dans la mesure où la déformation de l'échantillon a été limitée et que

l'attrition n'affecte pas directement les variables d'intérêt de l'enquête.

Une attrition limitée reste donc la clé de la réussite des enquêtes longitudinales, pour les pays qui, comme la France, ne peuvent s'appuyer sur les registres de population ou fichiers administratifs pour actualiser l'adresse des personnes. L'équipe conceptrice de l'enquête Élife (Étude longitudinale française depuis l'enfance) qui prévoit de suivre, à partir de 2010, 20 000 enfants depuis leur naissance jusqu'à l'âge adulte, devra porter une attention toute particulière au suivi de l'échantillon. Toutes les précautions doivent être prises dès la première vague, afin d'éviter la perte de « fiches-adresse », de risquer qu'il y ait confusion entre plusieurs membres d'un même ménage, et pour disposer d'un maximum d'informations permettant de retrouver les coordonnées du répondant en cas de déménagement. ■

Bibliographie

Bauer D., 2007, « Entre maison, enfant(s) et travail : les diverses formes d'arrangement dans les couples », *Études et résultats*, 570.

Couppie T., Gasquet C., et Lopez A., 2007, « L'enquête Génération 98 à sept ans », in Couppie T., Gasquet C., et Lopez A., (dir.), *Quand la carrière commence... Les sept premières années de vie active de la génération 98*, Céreq, p. 7-11 : <http://www.cereq.fr/carriere/QCC.pdf>

Mazuy M., Razafindratsima N., et de la Rochebrochard E., 2005, « Déperdition dans l'enquête Intentions de fécondité », *Documents de travail*, 129, Ined.

Razafindratsima N., kishimba N., et l'équipe Cocon, 2004, « La déperdition dans la cohorte Cocon entre 2000 et 2002 », *Population*, 59(3-4) : 419-448.

Régnier-Loilier A., 2006a, « À quelle fréquence voit-on ses parents ? », *Population & Sociétés*, 427 : http://www.ined.fr/fichier/t_publication/1211/publi_pdf1_pop_et_soc_francais_427.pdf

Régnier-Loilier A., 2006b, « Présentation, questionnaire et documentation de l'étude des relations familiales et intergénérationnelles », *Documents de travail*, 133, Ined. http://www.ined.fr/fichier/t_publication/1026/publi_pdf1_133.pdf

UNECE, 2005, Generations & Gender Programme. *Survey Instruments*, United Nations, New York and Genova.

Vikat A., Spéder Z., Beets G., Billari F., Désesquelles A., Fokkela T., Hoem J., MacDonald A., Neyer G., Pailhé A., Solaz A., 2007, « Generations and Gender Survey : Towards a Better Understanding of Relationships and Processes in the Life Course », *Demographic Research*, 17 (14) : 389-440.