

LA PONDÉRATION DE L'ENQUÊTE SANS DOMICILE 2012

Lionel Viglino ¹ & Sylvain Quenum ¹

¹ *Insee, 18 boulevard A. Pinard, 75675 Paris cedex 14,*
lionel.viglino@insee.fr, sylvain.quenum@insee.fr

Sans Domicile est une enquête atypique qui nécessite de répondre à plusieurs défis :

Comment échantillonner et pondérer des individus sans domicile fixe ? Par définition, la population ciblée n'a pas de résidence fixe qu'on puisse sélectionner directement. La solution mise en œuvre est donc un sondage indirect. On échantillonne les lieux que fréquentent régulièrement les personnes ciblées pour dans un deuxième temps y sélectionner les individus à enquêter sur le terrain. Pour réussir à contacter les personnes à interroger par un déplacement sur le terrain, un échantillonnage à 3 degrés est mis en œuvre : tirage d'agglomérations, puis tirage de services \times jours (visites) dans les lieux de fréquentation et enfin tirage sur le terrain des individus à enquêter au sein des usagers venus fréquenter le jour de la visite. Cependant comment s'assurer d'un bon taux de couverture et limiter le biais résultant du fait que des individus pourraient n'avoir aucune chance d'être joints (probabilité de tirage nulle) ? En outre, il est important d'avoir suffisamment de personnes sans-abri parmi les usagers répondants, le stade le plus avancé de l'exclusion. Comment cibler une surreprésentation efficace ? À l'inverse, on peut potentiellement sélectionner plusieurs fois la même personne car elle peut fréquenter régulièrement différents services d'aide. On utilise donc la technique de partage des poids pour donner le bon poids d'échantillonnage à chaque individu enquêté. Mais comment réussir à calculer le bon nombre de liens nécessaire à l'emploi de cette méthode, c'est-à-dire le nombre de fois où l'individu a la possibilité d'être tiré ? Comment gérer la réinterrogation du même individu ?

Comment estimer sur le terrain la proportion de non francophones au sein de chaque service visité ? Une nouveauté de l'enquête en 2012 par rapport à sa première édition en 2001 est que les individus non francophones font partie du champ. Cela conduit à des modes de sélection et de pondération différenciés pour les francophones et les non-francophones. Dès lors, il est important de connaître les nombres - ou du moins, les proportions - de francophones au sein des services d'aide afin de donner les bons poids d'échantillonnages respectifs à ces deux catégories d'individus et de questionnaires. Une seconde phase est donc envisagée dans le tirage de 3ième degré, celui des individus : d'abord en première phase, la sélection d'un certain nombre d'individus - dits contacts - pour répondre à un questionnaire auto-administré (QAA) sans discrimination par rapport à la langue. Ils serviront notamment à estimer la proportion de non francophones ; ensuite en seconde phase, la sélection de personnes francophones parmi les individus de la première phase pour répondre à un questionnaire individuel long (QIL) en face à face. Cette méthode permet en outre d'obtenir des éléments de calage et de traitement de la non réponse pour les QIL parmi l'ensemble des francophones contactés QIL+QAA, étant donné l'absence de source externe de calage.

Comment limiter la dispersion des poids ? Le coeur de cible de l'enquête est avant tout la population francophone. L'idéal serait d'obtenir l'équipondération des individus francophones interrogés pour un QIL. Cependant avant le tirage, l'information sur la proportion de non francophones est jugée de trop mauvaise qualité pour être intégrée comme paramétrage du tirage des visites. A quel niveau rechercher l'équipondération ? Enfin, il faut gérer ex-post de nombreux cas de modifications des caractéristiques des services constatés sur le terrain (fusions, éclatements, rééchantillonnages, fréquentations surestimées ...).