
REDRESSER UNE ENQUÊTE LONGITUDINALE : LE PANEL POLITIQUE DE LA VILLE

Marie SALA(*), Guillaume CHAUVET(**)

(*) CGET, Sous-direction de l'observation et des analyses statistiques

(**) Univ Rennes, ENSAI, CNRS, IRMAR - UMR 6625

marie.sala@cget.gouv.fr

guillaume.chauvet@ensai.fr

Mots-clés. Estimations transversales, panel, partage des poids, redressement.

Résumé

Le panel politique de la ville est une enquête longitudinale collectée annuellement, de 2011 à 2014, auprès des ménages des zones urbaines sensibles (Zus) de France métropolitaine. Près de 2000 ménages et 3000 individus de 16 ans ou plus, obtenus selon un plan de sondage à plusieurs degrés, étaient enquêtés chaque année en face à face et suivis d'une année sur l'autre, y compris en cas de déménagement – en Zus ou hors Zus. Afin de compenser l'attrition et de représenter les nouveaux arrivants dans le champ de l'enquête, un échantillon complémentaire de logements en Zus était tiré et enquêté chaque année.

Deux types de redressement ont été effectués : un redressement longitudinal pour suivre les trajectoires des individus, et un redressement transversal par vague, afin de disposer d'une « photographie » des zones urbaines sensibles pour chaque année de collecte. Les redressements transversaux effectués à partir de la deuxième vague d'enquête posent le plus de défis méthodologiques, car ils requièrent d'utiliser la méthode généralisée de partage des poids. En effet, à partir de la deuxième vague, la population d'intérêt est approchée par deux sous-échantillons : le premier sous-échantillon est composé des ménages comprenant des individus panels de la vague précédente, ménages dont la composition a pu changer entre-temps mais qui sont enquêtés intégralement. Le second sous-échantillon se compose de ménages enquêtés pour la première fois et qui auraient pu, ou non, être enquêtés la fois précédente selon qu'ils résidaient ou non en Zus un an auparavant.

Au sein du premier sous-échantillon de ménages issus des individus panel, après correction de l'attrition, la méthode du partage des poids a été appliquée à partir de la deuxième vague pour mettre en commun les individus panel – pour lesquels on dispose d'un poids issu de la première vague d'enquête, redressé de l'attrition – et les corésidents, pour l'instant dépourvus de poids. Pour le deuxième sous-échantillon, les traitements sont plus classiques : on dispose de poids d'échantillonnage et il s'agit principalement de corriger la non-réponse totale. Enfin, on applique la méthode de l'estimation composite pour mettre en commun les deux sous-échantillons.

Au-delà des aspects théoriques, les redressements transversaux des différentes vagues de l'enquête ont soulevé un certain nombre de questions pratiques du fait de l'étroitesse du champ de l'enquête, basé sur un zonage infra-communal.

Abstract

The french “Zus” are deprived urban neighbourhoods set on a sub-municipal level and targeted by specific public policies. A longitudinal survey, called PPV, was collected annually from 2011 to 2014 amongst households living in such areas. Individuals were followed from one year to another, even if the composition of their household had changed in the meantime. In that case, all members of the new household are surveyed. An additional sample was surveyed each year to compensate for cohort attrition. Besides the longitudinal estimation, a cross-sectional estimation has been performed, which required the use of the generalized weight share method. From the second wave of the survey, the individuals could indeed be sampled by two means : by the longitudinal sample or the additional one. This article presents the theoretical and practical issues of the cross-sectional estimation.

Introduction

Si les enquêtes longitudinales sont menées avant tout pour suivre des trajectoires dans le temps, elles peuvent aussi se montrer utiles en coupe transversale. La pondération transversale de ce genre d'enquête s'appuie sur la méthode généralisée du partage des poids [4]. Le panel politique de la ville permet d'appliquer cette méthode dans un contexte original, puisque le champ de l'enquête est réduit aux seules zones urbaines sensibles de France métropolitaine. La méthode du partage des poids exige d'identifier les personnes qui entrent ou qui sortent du champ de l'enquête entre deux interrogations. Dans les enquêtes en population générale, il faut alors identifier les personnes ayant immigré ou émigré du territoire, ainsi que les nouveaux-nés et les personnes décédées. Pour le panel politique de la ville, il s'agit de déterminer si les individus interrogés résidaient au sein d'une zone urbaine sensible un an auparavant, ce qui s'avère délicat dans la pratique.

1 Présentation du panel politique de la ville

1.1 Contexte et objectifs de l'enquête

Le panel politique de la ville a été mis en place en 2011 à la demande du conseil d'orientation de l'Observatoire national des Zus (Onzus) et du Secrétariat général du comité interministériel des villes (SG-CIV). Le panel a pour objectif d'améliorer l'observation et l'évaluation des effets de la politique de la ville sur les individus qui en bénéficient, au-delà des effets de cette politique sur les quartiers.

En effet, un débat récurrent dans l'histoire de la politique de la ville oppose une vision des quartiers prioritaires comme des quartiers « nasses » à une autre où ces quartiers sont vus comme des « sas » [1]. Dans le premier cas, on suppose que les habitants de ces quartiers se renouvellent très peu au cours du temps : le fait que ces quartiers concentrent toujours des populations précarisées malgré l'intervention de la politique de la ville serait la preuve de son inefficacité. À l'opposé, d'autres voient les quartiers de la politique de la ville comme des « sas » où la mobilité résidentielle serait élevée : ces quartiers accueilleraient des populations en grande difficulté, qui, lorsque leur situation s'améliore, déménagent ailleurs. Le fait que les indicateurs socio-économiques qui portent sur ces quartiers restent dégradés ne permettrait donc pas de juger de l'efficacité de la politique de la ville, puisque les personnes qui ont bénéficié de cette politique n'y résident

potentiellement plus.

D'autres besoins viennent justifier la mise en place du panel politique de la ville. D'une part, les effets du plan national de rénovation urbaine mis en place au début des années 2000 devaient être évalués. D'autre part, malgré les nombreuses enquêtes thématiques de la statistique publique exploitables à l'échelle des zones urbaines sensibles, les analyses de l'Onzus se heurtaient au manque de croisement possible entre les sujets (logement, sécurité, emploi...).

La volonté d'évaluer les effets à long terme de la politique de la ville sur ses bénéficiaires justifie la dimension longitudinale de l'enquête : le protocole du panel politique de la ville prévoit de suivre les individus enquêtés - tous habitant une zone urbaine sensible (Zus) lors de leur première interrogation - pendant quatre ans, y compris en cas de déménagement en dehors d'une Zus.

Afin de permettre le croisement de différentes thématiques, le questionnaire de l'enquête couvre des thèmes variés. Dans la mesure du possible, les questions posées sont identiques à celles figurant dans les grandes enquêtes de la statistique publique, afin de disposer d'un point de comparaison avec l'ensemble du territoire national et de pouvoir situer les résultats à l'échelle des zones urbaines sensibles dans leur contexte.

1.2 Plan de sondage et pondérations initiales

Le champ de l'enquête, pour la première vague réalisée en 2011, est limité aux zones urbaines sensibles de France métropolitaine situées intégralement dans des communes de plus de 10 000 habitants, soit 558 Zus sur les 717 de France métropolitaine. Cette restriction était due au fait que l'information sur la localisation des adresses en Zus ne figurait systématiquement que pour les communes dotées d'un répertoire d'immeubles localisés, donc les communes de plus de 10 000 habitants.

L'échantillon est sélectionné selon un plan de sondage à trois degrés. Un échantillon de zones urbaines sensibles est d'abord sélectionné, puis un échantillon de logements est sélectionné dans ces quartiers. Une unité de vie est ensuite tirée dans chaque logement, et tous les individus de 3 ans et plus sont enquêtés.

1.2.1 Sélection de l'échantillon de quartiers

Au premier degré, la population U_1 de quartiers est partitionnée en $H = 4$ strates définies selon le degré d'avancement du programme de rénovation urbaine. Un échantillon S_1 de $n_1 = 40$ quartiers est sélectionné, avec une allocation par strate proportionnelle au nombre total de résidences principales estimé avec le recensement de la population. L'échantillon est obtenu par sondage stratifié, avec au sein de chaque strate tirage à probabilités proportionnelles au nombre de résidences principales. On note π_i la probabilité de sélection du quartier i . Le même échantillon de quartiers sera utilisé pour toutes les vagues d'enquête.

1.2.2 Sélection de l'échantillon de logements

Au second degré, et dans chaque quartier sélectionné, on tire un échantillon de logements à partir d'une base de sondage constituée de 5 enquêtes annuelles de recensement (EAR) : pour la première vague d'enquête, il s'agit des EAR 2007 à 2011. Ce tirage vise à obtenir *in fine* des probabilités de sélection égales pour le second degré d'échantillonnage. En pratique, ces probabilités ne sont pas égales car la probabilité d'être sélectionné dans une EAR diffère selon les logements. On note S_i^t l'échantillon de logements sélectionné dans le quartier i lors de la vague 1, et $\pi_{j|i}^t$ la probabilité de sélection d'un logement j dans le quartier i . Le suffixe t est relatif à la vague d'enquête.

Le poids de sondage d'un logement j en vague 1 vaut donc :

$$d_j^t = \underbrace{\frac{1}{\pi_i}}_{\text{poids du quartier}} \times \underbrace{\frac{1}{\pi_{j|i}^t}}_{\text{poids du logement dans le quartier}} \quad \text{pour tout logement } j \in i. \quad (1)$$

1.2.3 Sélection de l'unité de vie enquêtée

Dans chacun des logements enquêtés, une unité de vie (UV) était sélectionnée afin d'être enquêtée. On considère qu'une sous-partie des habitants d'un logement forment une unité de vie lorsqu'ils font budget commun : dans la plupart des logements, l'ensemble des individus du logement forment une même unité de vie ; néanmoins, certains logements abritent plusieurs unités de vie. Dans le protocole de l'enquête, aucune méthode de tirage d'UV « kish » n'a été prévue. On considérera donc que la sélection est aléatoire, à probabilités proportionnelles au nombre d'adultes de 16 ans ou plus dans l'unité de vie. On note S_j^t l'échantillon d'unités de vie sélectionnées dans le logement j lors de la première vague d'enquête, et

$$\pi_{k|ij}^t = \frac{N_k^{t,16+}}{N_j^{t,16+}} \equiv \frac{\text{Nb d'adultes de 16 ans et + de l'UV } k}{\text{Nb d'adultes de 16 ans et + du logement } j} \quad (2)$$

la probabilité de sélection d'une UV k dans le logement j du quartier i . En pratique, $\pi_{k|ij}^t$ vaut 1 pour la plupart des UV, sauf dans le cas de 61 logements dans lesquels coexistent plusieurs UV.

1.2.4 Sélection de l'échantillon d'individus

Dans chaque unité de vie, tous les individus de 2 ans et plus sont enquêtés. On note $S_{k,ind}^t$ l'échantillon d'individus sélectionnés dans l'unité de vie k . Comme les individus de l'UV sont enquêtés exhaustivement (les enfants de moins de trois ans mis à part), le poids de sondage d'un individu est celui de son unité de vie, soit, pour un individu l de l'UV k du logement j et du quartier i :

$$d_{l|ij}^t = \frac{1}{\pi_{k|ij}^t}. \quad (3)$$

Le poids de sondage non conditionnel d'un individu l vaut :

$$d_l^t = \underbrace{d_j^t}_{\text{poids du logement}} \times \underbrace{d_{l|ij}^t}_{\text{poids de l'individu dans le logement}} \quad \text{pour tout individu } l \in k \in j \in i. \quad (4)$$

1.3 Traitement de la non-réponse totale

On rencontre un problème de non-réponse totale au niveau des logements et des individus. Pour les logements, la non-réponse totale est traitée par la méthode des groupes homogènes de réponse (GHR). On note $S_{r,j}^t$ l'échantillon de logements répondants. Le poids du logement j corrigé de la non-réponse totale est

$$d_{rj}^t = \frac{d_j^t}{\hat{p}_j^t}, \quad (5)$$

avec \hat{p}_j^t la probabilité de réponse estimée du logement j .

On rencontre également un problème de non-réponse totale pour les individus, également traité selon la méthode des GHR. On note $S_{r,ind}^t$ l'échantillon d'individus répondants. Le poids de l'individu l corrigé de la non-réponse totale est

$$d_{rl}^t = \frac{d_{rj}^t}{\hat{p}_l^t} \text{ pour tout individu } l \in j, \quad (6)$$

avec \hat{p}_l^t la probabilité de réponse estimée de l'individu l .

Les poids obtenus sont finalement calés à l'aide d'informations auxiliaires issues des enquêtes annuelles de recensement à l'échelle des Zus. Ces informations sont connues au niveau logement et au niveau individuel. On obtient au niveau logement les poids calés w_j^t , et au niveau individuel les poids calés

$$w_l^t = \underbrace{w_j^t}_{\text{poids logement calé}} \times \underbrace{\frac{d_{l|ij}^t}{\hat{p}_l^t}}_{\text{poids individu redressé de la non-réponse}} \text{ pour tout individu } l \in k \in j \in i. \quad (7)$$

1.4 Échantillons complémentaires à partir de la deuxième vague d'enquête

À partir de la deuxième vague d'enquête, deux sous-échantillons sont enquêtés. D'une part, les individus issus de la première vague sont suivis, et leurs unités de vie sélectionnées sont enquêtées en vague 2. D'autre part, un échantillon complémentaire d'individus est sélectionné et enquêté en vague 2, en première interrogation, afin de compenser l'attrition. En effet, le panel a été conçu pour être exploitable de manière transversale : pour chaque vague, la taille de l'échantillon devait être suffisante pour être exploitée de manière indépendante.

Comme pour l'échantillon initial, les logements des échantillons complémentaires étaient sélectionnés à partir des cinq dernières EAR disponibles.

Deux types de redressement ont été mis en place : des pondérations longitudinales ont été calculées pour étudier les trajectoires individuelles, et des pondérations transversales qui permettent de donner une photographie des zones urbaines sensibles une année donnée. C'est ce deuxième type de redressement que nous allons présenter ici, car il présente le plus de défis méthodologiques : à partir de la deuxième vague d'enquête, la plupart des individus enquêtés pouvaient l'être par chacun des deux sous-échantillons, ce qui requiert d'utiliser la méthode généralisée du partage des poids.

2 Le suivi des individus panel : mise en place de la méthode généralisée du partage des poids

À partir de la deuxième vague, le premier sous-échantillon est composé de tous les ménages qui comprennent au moins un individu panel : ce sont les individus et non les ménages qui sont suivis d'une vague à l'autre, mais l'ensemble du ménage d'un individu panel est interrogé, y compris s'il comprend des individus qui ne faisaient pas partie de l'enquête l'année précédente. Ces individus, qui sont alors enquêtés pour la première fois bien qu'ils fassent partie d'un ménage avec un individu panel, sont appelés corésidents ou cohabitants.

Le redressement de ce sous-échantillon s'est effectué en deux temps. On corrige tout d'abord de l'attrition individuelle observée entre les deux vagues. On utilise ensuite la méthode de partage

des poids [3] afin de passer d'une pondération individuelle issue de la vague précédente à une pondération au niveau unité de vie pour la vague redressée.

2.1 Traitement de l'attrition

Dans cette partie, nous prenons l'exemple du redressement de la vague 2, mais le raisonnement est semblable pour le redressement des vagues 3 et 4 : dans ce cas, le premier sous-échantillon est composé des ménages issus des individus enquêtés lors de la vague précédente.

On note S_r^t l'échantillon qui réunit les individus dont les unités de vie étaient répondantes en vague 1. S_r^t peut contenir des individus qui n'avaient pas initialement répondu en vague 1, bien que leur unité de vie ait été répondante. Le poids de départ d'un individu $l \in S_r^t$ vaut :

$$d_l^{t,t+1} = \underbrace{d_{rj}^t}_{\text{poids logement redressé NR}} \times \underbrace{d_{l|ij}^t}_{\text{poids individu dans le logement}} \quad \text{pour tout individu } l \in k \in j \in i, \quad (8)$$

où le poids d_{rj}^t du logement j redressé de la non-réponse est donné en (5) et le poids de sondage $d_{l|ij}^t$ de l'individu l dans le logement j est donné en (3).

Les individus de S_r^t font l'objet d'un suivi afin de pouvoir identifier l'unité de vie à laquelle ils appartiennent lors de la vague 2. On note $S_r^{t,t+1} \subset S_r^t$ le sous-échantillon des individus de S_r^t qui sont retrouvés en vague 2. Le traitement de l'attrition s'effectue par groupes homogènes de réponse. On note $\hat{p}_l^{t,t+1}$ la probabilité estimée pour un individu l de S_r^t d'appartenir à une unité de vie répondante en vague 2. Le poids de l'individu l corrigé de l'attrition vaut :

$$d_{rl}^{t,t+1} = \underbrace{d_l^{t,t+1}}_{\text{poids départ individu}} \times \underbrace{\frac{1}{\hat{p}_l^{t,t+1}}}_{\text{correction attrition}} \quad (9)$$

2.2 Partage des poids

Après correction de la non-réponse, seuls les individus interrogés lors de la première vague de l'enquête disposent d'un poids, et les cohabitants en sont toujours dépourvus. L'application de la méthode généralisée du partage des poids permet de passer d'un poids individuel pour les seuls individus panels à un poids au niveau unité de vie pour toutes les unités de vie du premier sous-échantillon.

Pour une unité de vie k contenant au moins un individu $l \in S_r^{t,t+1}$, le poids de l'unité de vie est donné par la somme des poids des individus de $S_r^{t,t+1}$ qu'elle contient, rapporté au nombre total N_k^t d'individus de k qui étaient dans le champ de l'enquête lors de la vague précédente d'enquête. Le poids correspondant pour une unité de vie k est donc :

$$d_{rk}^{t,t+1} = \frac{\sum_{l \in k \cap S_r^{t,t+1}} d_{rl}^{t,t+1}}{N_k^t} \quad (10)$$

À titre d'exemple, le ménage 1 de la figure 1 était composé de trois individus lors de la première vague, mais entre les deux vagues, l'un d'entre eux a décohabité et n'a pas été retrouvé. Le poids associé au nouveau ménage en vague 2 sera égal à la somme des poids des deux individus retrouvés corrigés de l'attrition, divisé par le nombre de lien avec la base de sondage, à savoir 2. Le ménage 3 s'est agrandi entre les deux vagues, et le nouvel arrivant résidait en zone urbaine sensible lors de la vague 1, bien qu'il ne fasse pas partie de l'échantillon initial. Seuls les individus 5 et 6 du nouveau ménage formé disposent de poids après correction de l'attrition. À l'issue du partage des poids, le logement associé a un poids égal à la somme des poids des individus 5 et 6, divisé par

3, étant donné que l'individu 7 était dans le champ de l'enquête lors de la vague précédente. En revanche, le ménage 4 qui s'est également agrandi entre les deux vagues hérite du poids de l'individu 8 inchangé, puisque le corésident n'était pas dans le champ de l'enquête lors de la vague 1.

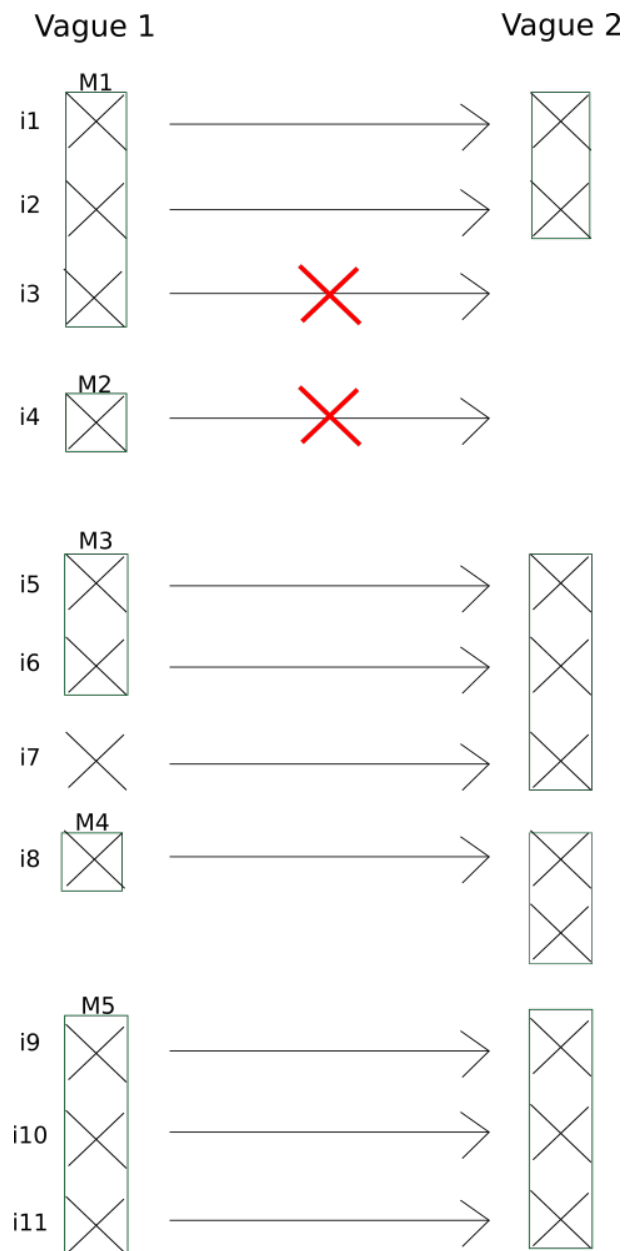


FIGURE 1 – Suivi des individus entre les vagues 1 et 2

Sur le même modèle que ce qui a été effectué lors de la première vague d'enquête, on peut obtenir un poids $d_{rj}^{t,t+1}$ au niveau logement en multipliant le poids de l'unité de vie par le nombre d'adultes de 16 ans ou plus de l'unité de vie, rapporté au nombre total d'adultes de 16 ans ou plus du logement, le tout en considérant le logement en vague 2. On obtient :

$$d_{rj}^{t,t+1} = \underbrace{d_{rk}^{t,t+1}}_{\text{poids de l'UV}} \times \underbrace{\frac{N_k^{t+1,16+}}{N_j^{t+1,16+}}}_{\text{correction du poids de l'UV dans le logement}} \quad (11)$$

2.3 En pratique

70 % des individus enquêtés lors de la première vague ont été retrouvés et enquêtés un an plus tard. Ce taux d'attrition était peu ou prou celui attendu, étant donné que PPV n'a pas le caractère d'enquête obligatoire. Pour le redressement transversal de la deuxième vague, il a aussi fallu écarter de la base les ménages ayant déménagé hors d'une zone urbaine sensible entre 2011 et 2012 : la base transversale de la vague 2 doit être représentative des seules zones urbaines sensibles en 2012. Au final, sur les 3 000 individus interrogés lors de la première vague, 2 200 sont enquêtés en deuxième vague.

Le traitement de la centaine de corésidents a posé le plus de problèmes pratiques, du fait de la nécessité de déterminer leur appartenance ou non au champ de l'enquête un an auparavant. Il s'agit donc, pour chacun des corésidents, de déterminer s'ils habitaient une zone urbaine sensible au printemps 2011. L'application de la méthode généralisée du partage des poids n'avait pas été anticipée dans le questionnaire de l'enquête, et on ne connaît pas précisément l'adresse des corésidents un an auparavant. Pour déterminer s'ils étaient alors dans le champ de l'enquête de la vague 1, on utilise des proxys. Notamment, une question permet de déterminer si l'individu habitait le même logement qu'un an auparavant, un logement différent dans le même quartier ou un logement dans un autre quartier. À partir de la troisième vague d'enquête, on demandait explicitement aux enquêtés leur adresse un an auparavant, ce qui a permis d'améliorer la qualité du redressement.

3 Le traitement des nouveaux ménages et la mise en commun des deux sous-échantillons

Les échantillons complémentaires sélectionnés à partir de la deuxième vague d'enquête permettent de couvrir l'ensemble du champ en vague 2, c'est à dire les ménages des zones urbaines sensibles. La base de sondage utilisée pour le tirage de l'échantillon complémentaire de la vague 2 est constituée des EAR 2008 à 2012. Le redressement de ce sous-échantillon suit globalement les mêmes étapes que celui de l'échantillon initial en vague une : calcul des poids de sondage et traitement de la non-réponse totale.

La mise en commun des deux sous-échantillons pose davantage de défi, puisque les pondérations ont été établies de telle manière que chacun d'entre eux représente l'intégralité de la population enquêtée.

3.1 Traitement de la non-réponse totale

Soit S_i^{t+1} l'échantillon de logements sélectionnés dans le quartier i lors de la vague 2. On note :

- d_j^{t+1} le poids de sondage correspondant au logement j ,
- S_j^{t+1} l'échantillon d'unités de vie sélectionnées dans le logement j dans la vague 2,
- S_l^{t+1} l'échantillon d'individus sélectionnés dans l'unité de vie,
- d_l^{t+1} le poids de sondage correspondant à l'individu l .

De la même manière que pour l'échantillon initial de la première vague d'enquête, on rencontre un problème de non-réponse totale au niveau des logements et des individus. La non-réponse totale est ici aussi traitée à l'aide de groupes homogènes de réponse. Pour les logements, on note \hat{p}_j^{t+1} la probabilité de réponse estimée du logement j . Le poids du logement j corrigé de la non-réponse totale est :

$$d_{rj}^{t+1} = \frac{d_j^{t+1}}{\hat{p}_j^{t+1}} \quad (12)$$

3.2 Estimation composite

La mise en commun des deux sous-échantillons se fait en tenant compte du fait qu'un logement contenant au moins un individu qui était dans le champ de l'enquête lors de la première vague peut être atteint par les deux sous-échantillons, alors qu'un logement qui ne contient que des individus qui entrent dans le champ de l'enquête en 2012 ne peut être atteint que par l'intermédiaire de l'échantillon complémentaire.

Si le logement j ne contient aucun individu qui était dans le champ de l'enquête en 2011, il représente les logements qui peuvent seulement être échantillonnés par le sous-échantillon de vague 2. On laisse donc son poids inchangé lors de la mise en commun des deux sous-échantillons :

$$d_{cj}^{t+1} = d_{rj}^{t+1} \quad (13)$$

où d_{rj}^{t+1} est donné en (12).

Si le logement j contient au moins un individu qui était dans le champ de l'enquête en 2011, il représente les logements qui peuvent être échantillonnés *via* les deux sous-échantillons. On lui attribue le poids :

$$d_{cj}^{t+1} = \begin{cases} \theta_j d_{rj}^{t,t+1} & \text{si } j \text{ est dans l'échantillon issu des individus de la vague 1} \\ (1 - \theta_j) d_{rj}^{t+1} & \text{si } j \text{ est dans l'échantillon complémentaire} \end{cases} \quad (14)$$

Le cas où un logement est sélectionné dans les deux sous-échantillons ne s'est pas produit en pratique, et n'a donc pas été traité.

Les pondérations θ_j et $1 - \theta_j$ correspondent au poids de chaque sous-échantillon : elles sont calculées en fonction du nombre de répondants par sous-échantillon et par quartier. Pour un logement j situé dans le quartier i , on prend donc :

$$d_{cj}^{t+1} = \begin{cases} \frac{n_{ri}^t}{n_{ri}^t + n_{ri}^{t+1}} d_{rj}^{t,t+1} & \text{si } j \text{ est dans le sous-échantillon issu de la vague 1,} \\ \frac{n_{ri}^{t+1}}{n_{ri}^t + n_{ri}^{t+1}} d_{rj}^{t+1} & \text{si } j \text{ est dans le sous-échantillon complémentaire,} \end{cases} \quad (15)$$

avec :

- n_{ri}^t le nombre de logements répondants dans le quartier i dans le sous-échantillon issu de la vague 1,
- n_{ri}^{t+1} le nombre de logements répondants dans le quartier i dans le sous-échantillon complémentaire.

Une méthode simple pour réaliser cette estimation composite aurait été de diviser tous les poids des logements qui comprennent au moins un individu qui était dans le champ de l'enquête un an auparavant par deux, deux étant le nombre de liens entre le logement et la base de sondage. Néanmoins, cette méthode aurait conduit à des poids beaucoup plus dispersés. En effet, le sous-échantillon complémentaire existe avant tout pour compenser l'attrition des ménages précédemment enquêtés, et le prestataire chargé de l'enquête avait pour instruction de réaliser un nombre fixe d'entretiens par quartier, tous sous-échantillons confondus. Dans les quartiers où l'attrition était faible, les logements du sous-échantillon complémentaire, peu nombreux, avaient un poids élevé, et diviser les poids de tous les logements qui auraient pu être enquêtés deux fois par deux aurait conduit à leur accorder beaucoup d'importance au regard des logements de l'autre sous-échantillon. En outre, les pondérations retenues visent à limiter la variance de l'estimateur composite [2].

3.3 Passage au niveau individuel et calage sur marges

Chaque individu l d'un logement j va recevoir comme poids de départ le poids du logement après mise en commun des deux sous-échantillons, soit :

$$d_{cl}^{t,t+1} = d_{cj}^{t,t+1} \quad (16)$$

Ce poids vaut aussi pour les corésidents qui partagent la même unité de vie qu'un individu panel issu de la vague 1.

Pour les individus, on note \hat{p}_l^{t+1} la probabilité de réponse estimée de l'individu l en vague 2. Le poids de l'individu corrigé de la non-réponse totale est :

$$d_{rl}^{t+1} = \underbrace{d_{cl}^{t,t+1}}_{\text{poids départ individu}} \times \underbrace{\frac{1}{\hat{p}_l^{t+1}}}_{\text{correction non-réponse}} \quad (17)$$

Enfin, comme pour l'estimation transversale de la première vague, les poids sont calés sur des marges connues à la fois sur les logements et sur les individus. On obtient au niveau logement les poids calés w_j^{t+1} , et au niveau individuel les poids calés w_l^{t+1} .

3.4 En pratique

La principale difficulté posée par la mise en commun des deux sous-échantillons était encore une fois de définir si les individus enquêtés pour la première fois lors de la deuxième vague d'enquête étaient dans le champ de l'enquête un an auparavant, autrement dit s'ils habitaient alors dans une zone urbaine sensible. Cette opération a dû être effectuée sur l'ensemble des individus du deuxième sous-échantillon, soit 2 090 individus. Là encore, le questionnaire de la vague 2 ne comportait pas de question sur l'adresse précise de résidence des enquêtés en 2011. Un certain nombre d'entre eux ont été retrouvé en vague 3 et ont accepté de donner leur adresse de résidence en 2011. Pour les autres individus, la probabilité de résider en Zus en 2011 a été estimée à partir des questionnaires remplis de la vague 3.

Conclusion

La mise en place de la méthode généralisée du partage de poids pour le redressement transversal du panel politique de la ville a été compliquée par la difficulté de déterminer le lieu précis de résidence antérieure des individus interrogés : leur adresse antérieure précise (et non leur seule commune antérieure de résidence) devait être renseignée, ce qui a pu paraître intrusif et a conduit à une non-réponse élevée pour cette question. L'estimation longitudinale ne demandait pas ce niveau de précision, puisque l'échantillon n'était alors pas représentatif de l'ensemble de la population des zones urbaines sensibles, mais uniquement de la population qui résidait en Zus en 2011.

En 2014, soit la dernière année de collecte du panel politique de la ville, une nouvelle géographie prioritaire a été définie et les quartiers prioritaires de la politique de la ville ont remplacé les zones urbaines sensibles. Malgré cela, les résultats tirés de PPV restent pertinents, car il existe une continuité entre les deux géographies prioritaires : notamment, toutes les Zus enquêtées par le panel sont restées dans la géographie prioritaire après la réforme, même si les contours ont pu varier.

Références

- [1] Laurent DAVEZIES et Philippe ESTÈBE. « Le sas ou la nasse » (2010). Étude réalisée pour l'Anru.
- [2] Jean-Claude DEVILLE et Pierre LAVALLÉE. « Sondage indirect : les fondements de la méthode généralisée du partage des poids ». *Techniques d'enquêtes* 32.2 (2006), p. 185.
- [3] Pierre LAVALLÉE. « Le sondage indirect, ou la méthode généralisée du partage des poids » (2002).
- [4] Pierre LAVALLÉE. « Pondération transversale des enquêtes longitudinales menées auprès des individus et des ménages à l'aide de la méthode du partage des poids ». *Techniques d'enquêtes* 21 (1995), p. 27-35.