

# Introduction

## Le *Recensement* rénové : avancées méthodologiques et apports à la connaissance

Chantal Cases \*

La commission nationale d'évaluation du recensement de la population (Cnerp), mise en place en 2004 à l'initiative du Conseil national de l'information statistique, est un lieu d'observation et de recommandations privilégié depuis la mise en place du recensement rénové<sup>1</sup>. Elle contribue à l'appropriation de ce nouvel outil non seulement par les élus et les personnels territoriaux qui en sont aussi les coproducteurs attentifs, mais encore par les chercheurs et les utilisateurs locaux ou nationaux des données qui constituent également une partie de ses membres. Elle en soutient la qualité, en consacrant notamment un groupe de travail à la constitution d'un de ses outils les plus stratégiques, le répertoire d'immeubles localisés qui conditionne fortement la valeur du sondage réalisé chaque année dans les communes de 10 000 habitants ou plus. La Cnerp prépare aussi les changements futurs, en suivant de près les opérations de dématérialisation de la collecte ; elle a constitué en 2012 l'ossature d'un groupe de travail sur l'évolution des questionnaires. Je suis sûre que ses membres seront, comme je l'ai été en préparant cette introduction, des lecteurs et lectrices passionnés de ce numéro qui fait écho à leur travail et à leurs questionnements. Ce sera aussi le cas, certainement, de tous ceux qui s'intéressent à cet outil central de la statistique publique qu'est le recensement, car le dossier d'articles réunis ici montre toute la portée des changements engagés et des travaux qu'ils permettent.

### Rappel :

Les jugements et opinions exprimés par les auteurs n'engagent qu'eux mêmes, et non les institutions auxquelles ils appartiennent, ni *a fortiori* l'Insee.

\* À la date de rédaction de cette introduction, Chantal Cases était vice-présidente de la Cnerp.

1. « Dix ans d'activité de la Commission nationale d'évaluation du recensement de la population (Cnerp) », *Cnis, Chroniques* n° 12, avril 2015.

La rénovation de la méthode de collecte du recensement de la population est sans conteste une des opérations les plus ambitieuses du système statistique public français depuis de nombreuses années. Il fallait une certaine dose d'audace et d'imagination pour réaliser un recensement grâce à... des enquêtes tournantes par sondage, sans registre de population associé ! Il n'est pas étonnant que des grands noms de la statistique, à l'étranger et en France, aient été les initiateurs de cette réflexion. Jean-Michel Durr et Alain Godinot, qui ont porté ce programme de rénovation au début des années 2000 et jusqu'à son lancement effectif sur le terrain, ont retracé dans l'avant-propos de ce numéro la genèse, l'histoire et les nombreuses facettes de ce projet au long cours, soumis à de fortes contraintes et porteurs de fortes transformations, bien au-delà des statisticiens de l'Insee.

Économie et Statistique était le support le plus naturel pour un dossier consacré au recensement rénové et venant après quelques années d'exercice et de résultats. Un tel numéro spécial avait d'ailleurs été suggéré par le Conseil scientifique de la revue. L'appel à contributions lancé grâce au réseau Quetelet – réseau qui permet aux chercheurs en sciences humaines et sociales d'accéder aux micro-données du recensement mais aussi à de nombreuses enquêtes socio-démographiques et bases de données de la statistique publique et de la recherche, ainsi qu'à des bases de données internationales – a mobilisé un nombre conséquent de contributeurs potentiels et de qualité sur les deux grands axes qu'il avait définis : un axe méthodologique et un axe portant sur les apports de cette nouvelle source à la connaissance. D'où cet épais numéro, riche d'enseignements pour tous ceux qui se préoccupent de questions de population.

Les articles de méthodologie statistique y occupent une place importante, tout au long du dossier. La première partie examine ainsi les conséquences des transformations au cœur du nouveau dispositif, qui ont suscité de nombreuses interrogations lors de son lancement : le passage à un recueil par sondage sur les communes de 10 000 habitants ou plus et à une interrogation exhaustive quinquennale des petites communes, dans l'objectif, entre autres, de produire des résultats annuels et fiables de population légale pour toutes les communes. Les questions qui sont traitées dans les trois articles qui la composent ont beaucoup animé les débats de la Cnerp : la perte de précision due au sondage sur les grandes communes ne risquerait-elle pas de détériorer fortement la qualité des résultats ? Que se passerait-il, en termes d'organisation et en termes de résultats, quand une commune en croissance dépasserait le seuil des 10 000 habitants et devrait passer à la collecte par sondage ? Et comment produire chaque année des données de population légale rendant au mieux compte des évolutions sur le terrain pour les petites communes recensées seulement tous les cinq ans ? Ces trois questions sont successivement traitées par Gwenaëlle Brillhaut et Nathalie Caron, puis Gwenaëlle Brillhaut et Annie Chanut, enfin par Hugues Lerméchin. Le premier article rappelle très clairement et en détail la méthode d'échantillonnage utilisée, ainsi que la méthode d'extrapolation retenue pour produire chaque année des populations légales communales et les résultats statistiques du recensement. Au bout du compte, l'incertitude liée au sondage est estimée à 0,05 % sur le total de la population française de métropole, et des estimations en sont données à divers niveaux de granularité des résultats. Le deuxième article détaille comment se traite en pratique le cas particulier des communes qui franchissent le seuil des 10 000 habitants, à la hausse ou à la baisse. Deux ans successifs de franchissement du seuil sont attendus pour modifier le mode de collecte, exhaustive ou par sondage, afin notamment de pouvoir constituer une base de sondage pour les communes en croissance. Des méthodes transitoires de calcul des populations légales et des résultats statistiques sont également

mis en place. Ces deux premiers articles mettent bien en évidence l'importance de la constitution d'une base de sondage fiable dans les grandes villes, en l'occurrence le répertoire des immeubles localisés, qui fait d'ailleurs l'objet d'une enquête régulière de qualité<sup>2</sup>. Le troisième article détaille un aspect de la mobilisation des sources administratives dans le recensement rénové : l'utilisation des données de la taxe d'habitation pour l'extrapolation des populations légales sur les 40 % des communes de moins de 10 000 habitants pour lesquelles la vague de collecte date de plus de deux ans. Pour cela, il a fallu adapter et compléter la source fiscale, constituée pour connaître l'état du bâti, à la problématique du dénombrement de la population, et s'interroger sur sa qualité au regard de l'utilisation que l'on voulait en faire.

Cette première partie ouvre la voie vers d'autres travaux méthodologiques ; certains prendraient utilement leur place dans de futurs numéros de la revue. Ainsi, les auteures du premier article nous suggèrent que certaines caractéristiques du nouveau recensement – l'annualisation de la collecte et la tenue régulière d'une base d'adresses dans les grandes communes – ont vraisemblablement joué positivement sur la qualité de l'opération ; elles reconnaissent aussi que la précision du recensement ne se limite pas à la mesure des conséquences de l'approche par sondage sur une partie du territoire et qu'il est bien difficile d'en avoir une vue complète. Parmi les autres sources possibles d'erreurs de mesure, en matière de recensement, on pense d'abord aux oublis et aux doubles comptes. À défaut d'enquête post-censitaire, opération rare car coûteuse et délicate à mener, la mobilisation conjointe des données du recensement et de sources administratives, comme ici les données de la taxe d'habitation, ou possiblement d'autres données, est une voie très fructueuse pour explorer ces questions, repérer des écarts éventuels entre sources, leurs origines possibles de part et d'autre, suggérer des méthodes adaptées de correction des données ou d'amélioration de la collecte. Le répertoire statistique des logements en cours à l'Insee y contribuera sans aucun doute. Le travail mené en comparaison avec l'enquête *Famille et Logements* plus loin dans le dossier (article de Pascale Breuil-Génier *et. al*) montre bien tous les apports du travail multi-sources pour évaluer la qualité des informations collectées. À l'heure du *big* et de l'*open data*, la facilitation accrue de la mobilisation de ces sources de données complémentaires, moins diversifiées en informations sur les ménages et les personnes mais rapidement disponibles, est une véritable opportunité. Il est souhaitable que les statisticiens de l'Insee et les chercheurs poursuivent dans ces différentes directions. La transparence, la qualité des données collectées et la richesse des analyses qui en seront tirées ne pourront qu'en bénéficier.

Avec la deuxième partie du dossier, on entre dans le champ de l'exploitation du recensement pour des aspects plus classiques : ici il s'agit d'exemples de son apport aux connaissances en matière d'emploi et de mobilités. À une époque où le débat médiatique sur l'emploi est (trop) largement occupé par la publication mensuelle des chiffres de demandeurs d'emploi inscrits à Pôle Emploi, l'article de Jean-Benoît Massif montre que si le recensement n'est pas l'instrument adapté pour analyser les évolutions conjoncturelles du marché du travail, il est en revanche irremplaçable, même en partie réalisé par sondage, pour les analyses localisées à la commune ou infra-communales des comportements d'activité et d'emploi en fonction de caractéristiques des personnes et de leur ménage. Il prend ici toute sa place, d'autant plus que le concept d'emploi au sens du recensement n'apparaît pas très éloigné de celui du BIT, référence absolue pour les

---

2. Michel Cezard et Olivier Lefevre (2008), « La qualité du recensement », <http://unstats.un.org/unsd/censuskb20/Attachment478.aspx>.

comparaisons internationales. Les élus et décideurs locaux ou les agences d'urbanisme, qui mobilisent régulièrement ces données, n'en seront pas surpris.

Les trois autres articles de cette partie s'intéressent aux mobilités, à des échelles différentes. L'article de Julyan Arbel et Vianney Costemalle est un nouvel exemple de la fertilité de l'utilisation conjointe du recensement et de données administratives, celles des fichiers des titres de séjour en l'occurrence, pour la difficile estimation des flux d'immigration, au-delà de la publication, lors des résultats du recensement, de soldes migratoires apparents obtenus indirectement par différence entre l'évolution de la population mesurée entre deux recensements et le solde naturel déduit de l'état civil. Ils ont choisi pour cela d'utiliser l'inférence bayésienne. Ils proposent ainsi des évaluations des flux de migrants en estimant des taux de présence au recensement et des taux d'obtention de titres de séjour selon le nombre d'années passées sur le territoire. Cet article nous amène à nous questionner sur la difficulté de repérage dans le recensement de la population des nouveaux arrivants, qui ne se considèrent pas toujours à ce moment comme des résidents de long terme, et sur la difficulté qu'ont les immigrants plus anciens à renseigner une date d'arrivée, à une période où, sous l'impulsion européenne, la date de référence de la résidence antérieure qui sert en pratique à compléter cette information sur la date d'arrivée a été modifiée dans le questionnaire. Cet essai, original dans sa méthode, de réconciliation de deux sources régulièrement mobilisées pour estimer les flux de migrations est donc bienvenu. Il est aussi très rassurant sur l'ampleur de la sous-évaluation de la population totale liée à ces difficultés de recensement des migrants récents. Il serait intéressant de prolonger ce type de méthode pour estimer des flux de sortie des migrants, pour lesquels aucune mesure directe n'est disponible. Par ailleurs, l'estimation des entrants sur le territoire progresse également à partir du seul recensement, puisque l'Insee a produit récemment une nouvelle méthodologie en la matière<sup>3</sup> ; de nouvelles recherches à partir des titres de séjour ont été récemment publiées<sup>4</sup>. Il faut se réjouir de ces divers travaux sur un sujet qui est traditionnellement un des points faibles de l'information démographique, en France autant qu'à l'étranger.

Pierre Pistre examine aussi le rôle central des mobilités, mais internes cette fois, et leur contribution à la dynamique démographique des espaces « peu denses » en longue période, en combinant donc les données de l'ancien et du nouveau recensement. Cette perspective d'analyse des territoires fins sur une durée longue est certainement un des plus grands intérêts du recensement. Cette étude envisage également la perspective du traitement des données annualisées pour étudier des évolutions plus conjoncturelles, comme les mobilités dans les petites communes en période de crise : l'auteur constate ainsi un mouvement de réduction de l'installation de nouveaux résidents après 2008-2009 dans les petites et moyennes aires urbaines et les communes isolées. Jean-Paul Hubert, Pierre Pistre et Jean-Loup Madre s'intéressent à d'autres types de données sur ce même espace peu dense. Ils le répartissent en différents types de configurations spatiales selon les densités de population et de voisinage, et y analysent les comportements de recours à l'automobile, afin de chercher à expliquer un paradoxe : la concomitance entre retournement à la baisse des kilométrages parcourus et poursuite de la hausse du taux d'équipement en véhicules motorisés. Le recensement de la population, utilisé en complément

3. Chantal Brutel (2014) « Estimer les flux d'entrées sur le territoire à partir des enquêtes annuelles de recensement », Document de travail Insee n° F1403, juin, et Id. (2014) « Les immigrés récemment arrivés en France - Une immigration de plus en plus européenne », Insee première, n° 1524, novembre.

4. Hippolyte d'Albis et Ekrame Boubtane Id., (2015), « Caractérisation des flux migratoires en France à partir des statistiques de délivrance de titres de séjour (1998-2013) », Population, vol. 70, 2015/3, 477-502.

du Panel Parc Auto de la Sofres –interrogation annuelle d’un échantillon limité et permettant d’estimer les kilométrages parcourus – et des enquêtes nationales transports de 1993-1994 et 2007-2008, est la seule source qui permette d’analyser à un niveau géographique fin, après 2008, les taux d’équipement des ménages en véhicules et les déplacements domicile-travail. Là encore, les auteurs choisissent de produire des séries annuelles à partir des enquêtes annuelles de recensement, cette fois pour les analyser en niveau (moyenne mobile) ou en variations quinquennales pour les communes de moins de 10 000 habitants d’un même groupe de rotation.

Le fait de disposer chaque année d’une enquête de recensement avec des échantillons de grande taille ouvre donc des perspectives d’analyses approfondies, à partir d’exploitations annuelles de ces résultats. La prudence est nécessaire en la matière, car si les échantillons annuels sont équilibrés au niveau régional, cet équilibrage ne porte que sur quelques grandes variables et leur représentativité à des niveaux géographiques plus fins, ou pour des croisements de variables particuliers, n’est pas assurée. Guy Desplanques avait ouvert dès 2008 la voie, ici exploitées par les deux articles sur les espaces peu denses, en comparant certaines données de stock de population (population par sexe et âge) et de flux (migrations par groupes de nationalité) sur plusieurs enquêtes annuelles successives<sup>5</sup>. Il est souhaitable que ces possibilités soient expertisées à l’avenir sur divers champs d’analyse, de manière rigoureuse et pragmatique, afin d’en déterminer les apports et les limites.

La troisième partie du numéro tire les enseignements d’une opération qui, traditionnelle dans le cadre du recensement exhaustif depuis les années 50, devait refaire ses preuves dans celui du recensement rénové : la collecte d’une enquête complémentaire, en l’occurrence l’enquête famille et logements de 2011. Ce type d’enquête apporte, à un moindre coût et sur un échantillon de taille bien plus importante que les habituelles enquêtes statistiques auprès d’échantillons de ménages, des informations complémentaires au recensement et une meilleure compréhension des résultats de celui-ci. Il permet même une appréciation de certains types d’erreurs qu’il comporte, à travers l’exploitation conjointe de leurs questionnaires. À cet égard, l’analyse scrupuleuse concernant l’évaluation par l’une et l’autre source du nombre de couples de même sexe mérite d’être soulignée. Elle montre à la fois la grande qualité des dénombrements (en termes de doubles-comptes) et des informations recueillies au recensement, mais aussi l’incidence relative non négligeable d’erreurs rares sur le repérage de situations elles-mêmes peu fréquentes : couples de même sexe ou enfants en résidence alternée. Techniquement, il fallait que cette enquête réalisée sur une seule année puisse s’adapter à la nouvelle méthodologie du recensement organisée sur cinq ans, et ce premier défi a été relevé avec un plan de sondage adapté. La réussite de cette opération passait aussi, comme celle du recensement, par le soutien et la collaboration avec les communes, au risque sinon non seulement de voir échouer l’enquête, mais de détériorer les résultats du recensement. Au bout du compte, la participation des communes échantillonnées a été massive ; la faisabilité et la non-nocivité de l’enquête sur les opérations principales du recensement ont été largement démontrées et l’expérience pourra donc être reproduite. De plus, la refonte en cours de la feuille de logement bénéficie déjà des enseignements de cette enquête et permettra d’améliorer considérablement la photographie des familles dites « complexes » (recomposées, multi-résidentes...) dans les années à venir.

---

5. Guy Desplanques (2008), « Avantages et incertitudes des enquêtes de recensement en France », *Population-F*, n° 63.

En contrepartie de leur engagement ont été fournies aux communes participantes, plus de 1400, où l'enquête n'a été réalisée que sur quelques centaines de logements tirés au hasard, des estimations de sous-populations d'intérêt construites à partir de l'enquête complémentaire : personnes pacsées, en couple non cohabitant, familles recomposées ou traditionnelles, grands-parents, personnes âgées vivant seules et ayant des enfants résidant à proximité, enfants de moins de 4 ans selon leur mode de garde principal, enfants selon le type de la famille dans laquelle ils vivent, adultes et les enfants occupant régulièrement deux logements, etc. Pascal Ardilly nous explique comment les techniques d'estimation petits domaines ont permis d'évaluer ces informations de grand intérêt pour les élus, sans en cacher les limites. Opération inédite là encore, et estimations audacieuses sans doute, dont le lancement avait d'ailleurs à l'époque été jugé très risqué par le comité de pilotage de l'enquête. Nécessaires, peut-être aussi, à une bonne acceptation par les élus de ces opérations complémentaires et facultatives dont ils n'auraient pu, sans cette opération, qu'être des contributeurs sans en retirer d'information locale utile.

Dans une dernière partie, le recensement rénové français est replacé dans une double perspective européenne. Celle de l'évolution considérable des méthodes au tournant des années 2000, d'abord, avec une perte du monopole des recensements exhaustifs par questionnaire papier, maintenant minoritaires, au profit d'une diversification des méthodes et des supports de collecte. Patrick Festy et Nathalie Le Bouteillec en font un panorama détaillé et fournissent un point de vue comparatif passionnant sur ce qui nous paraît sans doute à tort, d'un point de vue national, comme une opération très normée. Six pays du nord de l'Union européenne procèdent à un recensement virtuel à partir de différents registres ; neuf pays, notamment l'Allemagne, l'Espagne, l'Italie, les Pays-Bas, la Pologne, combinent registres, enquêtes par sondage ou enquêtes exhaustives ; douze pays effectuent encore un recensement classique de la population ; enfin, la France a adopté en 2004 une méthode non conventionnelle, un recensement tournant (*rollingcensus*). On s'aperçoit d'autre part que la forme et les contenus des questionnaires de recensement varient beaucoup entre les pays, le questionnaire français étant nettement le plus court et le plus tourné vers la description de l'activité économique des personnes. Ce travail comparatif avait été particulièrement apprécié au sein du groupe chargé de réfléchir à l'évolution des questionnaires sous l'égide du Cnis<sup>6</sup>. Il avait permis de proposer avec suffisamment de confiance des changements destinés à une meilleure connaissance de certains sujets, comme la santé ou le handicap, à un territoire fin qui est aussi souvent celui de l'action. Ces changements prendront du temps à être mis en place, pour des raisons d'organisation ou de coût. Mais le travail des auteurs de l'article permet de réfléchir aux possibilités d'innovation de manière plus ouverte, en montrant tout le champ des possibles ici et ailleurs. Sans doute, le développement de la collecte par Internet ouvrira-t-il davantage ces possibles dans les années à venir.

Si la diversité des modes de collecte possibles a été reconnue dans le règlement européen de 2008, celui-ci prescrit aussi la transmission d'un certain nombre de variables à des fins de comparaisons entre pays et de conduite des politiques communautaires. Ces variables, regroupées en ligne dans le « Census Hub » européen, permettent de réaliser des études comparatives que l'on espère aussi voir se développer, dans cette revue ou dans d'autres

---

6. J-C. Frécon, C. Cases, F. Clanché et P. Oger (2012) « Évolution du questionnaire du recensement de la population », Rapport Cnis n°130, [http://www.cnis.fr/cms/accueil/publications/Les\\_rapports\\_du\\_Cnis;jsessionid=EC2D2553DF5BEB351B7D48B7C8143973?publication=10166](http://www.cnis.fr/cms/accueil/publications/Les_rapports_du_Cnis;jsessionid=EC2D2553DF5BEB351B7D48B7C8143973?publication=10166).

supports. La transmission de variables comparables nécessite parfois des travaux élaborés de construction de données à partir de collectes nationales dont on a vu toute la diversité de fond et de forme. Ainsi, dans le dernier article du dossier, Pascal Ardilly décrit-il le cheminement méthodologique retenu en France pour fournir une imputation de trois variables importantes : l'indicateur de résidence antérieure, la période d'achèvement de l'immeuble et l'activité au sens du BIT. Pour les deux premières, des changements de questionnaires avaient été réalisés en 2011, modifiant les périodes de référence sans modalité particulière pour assurer un recouvrement, même partiel, pendant une année de collecte donnée. C'est une situation assez fréquente, compte tenu des contraintes de charge de collecte, et qui risque de se reproduire à l'avenir au fil des évolutions, souhaitables par ailleurs, de questionnaires : il faut pouvoir s'ajuster à des besoins qui se transforment tout en produisant des séries de données les moins heurtées possibles au fil du temps. Même pour ces changements relativement simples, l'article reconstruit en détail l'articulation du raisonnement et des choix effectués et en montre de manière très transparente la complexité d'élaboration. La troisième variable nécessitait une reconstruction complète, par modélisation, afin de s'ajuster à un concept un peu différent de celui couvert par les questionnaires. L'auteur nous fait ici toucher du doigt, très concrètement, l'appareillage élaboré qui soutient la difficile « conciliation entre diversité des modalités et harmonie des résultats » mise en avant par les auteurs de l'article précédent comme caractérisant le Census Hub.

Au total, ce triple numéro donne une vision très impressionnante de la technicité et des innovations réalisées à l'occasion de cette opération phare de la statistique publique, autant que de la richesse des analyses qu'elle permet. Comme le soulignent les auteurs de l'avant-propos, le recensement rénové a fait toute sa place dans le paysage ; il a sans doute gagné une nouvelle légitimité. Sa rénovation réussie laisse aussi ouvertes de nombreuses nouvelles possibilités de recherche et d'innovations. Souhaitons qu'elles soient saisies avec autant de maîtrise, de rigueur et d'imagination dans les années à venir. □

