

Paru :
Un nouveau numéro de la revue
« Économie et Statistique »

Le 28 avril 2016 paraîtra la nouvelle publication d'*Économie et Statistique* (n°483-484-485) qui traite des **avancées méthodologiques du Recensement de la population renoué et des apports à la connaissance qui en découlent.**

Au sommaire de ce numéro :

1. Nouvelles méthodologies :

- Le passage à une collecte par sondage : quel impact sur la précision du recensement ?
- Adaptation du recensement à la dynamique des communes qui franchissent le seuil des 10 000 habitants
- Quel est l'apport de la taxe d'habitation à l'extrapolation du nombre de résidences principales au *Recensement de la population* ?

2. Emploi et mobilités via le *RP* :

- De quelle mesure de l'emploi le *Recensement de la population* est-il le nom ? La place du *Recensement de la population* dans le système de suivi de l'emploi
- Estimation des flux d'immigration : réconciliation de deux sources par une approche bayésienne
- Recensements de la population pour l'étude des reprises démographiques et des migrations résidentielles dans l'espace peu dense (1975-2011)
- L'utilisation de l'automobile par les ménages dans les territoires peu denses : analyse croisée par les enquêtes sur la mobilité et le *Recensement de la population*

3. Une opération associée au *RP* : l'enquête *Famille et Logements* :

- Enquête *Famille et Logements* adossée au *Recensement* de 2011 : comment s'adapter à la nouvelle méthodologie des enquêtes annuelles et quels apports ?
- Estimations communales exploitant les données de l'enquête *Famille et Logements* 2011 et du recensement : une opération à hauts risques

4. Convergences européennes :

- Recensements européens : mouvements centrifuges et forces de rappel
- Recensement 2011 et règlement européen : la procédure d'imputation spécifique à trois variables

L'ouvrage est mis à la disposition des journalistes intéressés :

Le 27 avril 2016, à partir de 10h00

Bureau de presse de l'Insee
bureau-de-presse@insee.fr - Tél. : 01 41 17 57 57

Avertissement

Économie et Statistique est une revue scientifique dont la publication est assurée par l'Insee. Les articles qui y paraissent sont publiés sous l'unique responsabilité de leurs auteurs. Ils n'expriment aucune position officielle et n'engagent ni l'Insee, ni les autres institutions auxquelles ces auteurs appartiennent.

Résumés

Le passage à une collecte par sondage : quel impact sur la précision du recensement ?

Gwennaëlle Brilhault et Nathalie Caron

Le recensement rénové de la population mis en place en 2004 est une opération tournante qui a permis d'annualiser la collecte, les dépenses et la publication des résultats, tout en assurant des gains d'efficacité des différents acteurs. Il se caractérise également par le fait que seule une fraction de la population des ménages des communes de 10 000 habitants ou plus est recensée : le sondage s'ajoute donc aux facteurs influençant la qualité d'un recensement traditionnel exhaustif. Cette décision a suscité de nombreuses interrogations autour de la variabilité supplémentaire ainsi introduite pour les estimations basées sur ces nouvelles données du recensement. En particulier, les élus locaux s'inquiétaient de l'impact de cette nouveauté sur l'effectif officiel de la population, lequel détermine le montant des ressources financières allouées aux communes par l'État. L'objectif de cet article est de proposer un historique des travaux chiffrant l'impact du sondage sur la précision des résultats du recensement, et d'en donner un état des lieux actuel. Ces travaux se sont déroulés en deux phases. D'abord, des calculs de précision ont été menés via des simulations basées sur les données du recensement exhaustif de 1999. Ensuite, ces premiers résultats ont été affinés en procédant à des calculs à partir des premières données d'un cycle complet du recensement rénové. Ces travaux méthodologiques démontrent que l'ordre de grandeur de cette imprécision est négligeable et par conséquent que le passage à la collecte par sondage n'a pas dégradé la qualité des résultats du recensement.

Adaptation du recensement à la dynamique des communes qui franchissent le seuil des 10 000 habitants

Gwennaëlle Brilhault et Annie Chanut

Le recensement rénové en métropole et dans les DOM a introduit des méthodes de collecte et de calcul annuel des populations légales différentes selon que les communes se situent ou non au-dessous du seuil des 10 000 habitants. L'importance de ce seuil dans le processus du recensement conduit à un suivi rapproché des communes dont la population l'atteint à la hausse ou le franchit à la baisse. Le décret n° 2003-485 du 5 juin 2003 fixe le délai de prise en compte d'un franchissement de seuil dans les populations légales : après deux constats successifs de population municipale en franchissement de seuil à la hausse (resp. à la baisse), l'Insee a trois ans (resp. cinq ans) pour modifier le mode de collecte d'une commune afin de correspondre à son nouveau statut. Ces délais permettent d'adapter le mode de collecte : préparation d'une collecte exhaustive pour les nouvelles communes de moins de 10 000 habitants et constitution d'une base de sondage (Répertoire des Immeubles Localisés ou RIL) pour les nouvelles communes de 10 000 habitants ou plus. Depuis 2009 (deuxième publication des populations légales basées sur le recensement rénové et donc deuxième constat potentiel d'un franchissement du seuil des 10 000 habitants), une cinquantaine de communes ont été concernées par un changement de leur mode de collecte ; chaque année s'y ajoute une nouvelle vague de communes pour lesquelles on modifie le mode de collecte du recensement. L'étape de constitution d'un RIL pour les nouvelles communes de 10 000 habitants ou plus nécessite un soin particulier car la qualité des résultats du recensement pour ces communes en dépend, au travers de la qualité des échantillons qu'on y tire. Le franchissement de ce seuil a impliqué également la mise au point de méthodes transitoires de calcul des populations légales et des résultats statistiques du recensement rénové, pour les communes franchissant à la hausse ce seuil des 10 000 habitants comme pour celles le franchissant à la baisse.

Quel est l'apport de la Taxe d'Habitation à l'extrapolation du nombre de résidences principales au Recensement de la population ?

Hugues Lermechin

Depuis la rénovation du recensement au début des années 2000, obligation est faite à l'Insee de publier chaque année les populations légales de toutes les communes. Pour les communes de moins de 10 000 habitants, qui sont recensées exhaustivement tous les cinq ans, la difficulté réside dans l'actualisation des populations durant la période intercensitaire. L'utilisation des données de la taxe d'habitation (TH) répond pour partie à ce défi et constitue une innovation notable dans le processus de calcul des populations. La TH permet d'estimer le nombre de résidences principales, approchant ainsi la dynamique de population. Elle contribue à calculer chaque année la population de 40 % des communes de moins de 10 000 habitants, occasionnant par là-même l'estimation d'environ 20 % de la population française.

La TH apporte-t-elle une réelle plus-value à l'estimation des populations communales et de la population totale concernées par son utilisation ? Si le nombre de résidences principales est surestimé, l'erreur occasionnée est faible. En extrapolant le nombre de résidences principales avec la TH durant cinq années successives, l'erreur est ainsi inférieure à 5 % pour près de trois quart des communes et le nombre total de résidences principales n'est surestimé que de 1 %. En réalité, l'extrapolation n'excédant jamais deux années successives, l'erreur est selon toute vraisemblance moindre les années où la TH participe au calcul des populations. Un effet taille apparaît néanmoins : les plus grandes des petites communes sont plus souvent surestimées.

Par ailleurs, les populations légales de sept communes sur dix sont mieux estimées avec la TH, qu'elles ne le seraient par prolongement linéaire des derniers recensements. En outre, la TH restitue finement les ruptures de tendance observées dans l'évolution du parc de logements des communes. C'est à la fois un instrument précis et souple car il s'adapte aux différentes dynamiques d'évolution qui peuvent exister.

De quelle mesure de l'emploi le Recensement de la population est-il le nom ? La place du Recensement de la population dans le système de suivi de l'emploi

Jean-Benoît Massif

Par le passé, en constituant un point d'ancrage décennal des séries officielles d'estimations annuelles d'emploi (EAE), le *Recensement* sous sa forme classique occupait une place centrale dans le système de suivi de l'emploi en France. L'instauration à partir de 2004 d'un nouveau dispositif de recensement à base d'enquêtes partielles annuelles tournantes sur un pas de 5 ans a obligé à repenser cette place. Moins centrale, elle n'en demeure pas moins essentielle. Si le recensement présente désormais l'inconvénient de proposer une mesure de l'emploi lissée sur 5 ans, donc tendancielle, il n'en reste pas moins un complément indispensable aux EAE et à l'enquête *Emploi*.

Moins riche que l'enquête *Emploi* pour analyser les comportements d'activité des personnes et décrire les emplois, le *Recensement* permet en revanche une déclinaison locale très fine de ses résultats, hors de portée de cette enquête. Cette place est légitimée par le fait que la notion d'emploi au sens du recensement est proche de celle du BIT, mesurée par l'enquête *Emploi* et utile aux comparaisons internationales. Si le *Recensement* surestime quelque peu le niveau de l'emploi en étant moins précis que cette enquête dans son questionnement, cela tient à des populations bien particulières, pour lesquelles l'emploi est une situation d'activité *a priori* plus secondaire ou occasionnelle à certaines étapes du cycle de la vie : étudiants, jeunes mères de famille, retraités. Malgré cette légère différence en niveau (de l'ordre de 1,1 %), la tendance de l'emploi mesurée par le recensement est très proche de celle offerte par l'enquête *Emploi*.

Au niveau local, la primauté des sources administratives (EAE) sur le terrain du conjoncturel n'épuise pas l'intérêt des ressources du *Recensement*. Celui-ci reste plus riche pour caractériser les personnes occupant les emplois et leur ménage. Il est aussi moins sujet aux problèmes de localisation fine des emplois, pas encore complètement résolus dans les sources administratives.

Estimation des flux d'immigration : réconciliation de deux sources par une approche bayésienne

Julyan Arbel et Vianney Costemalle

Deux sources au moins permettent d'appréhender les flux migratoires en France. D'une part les titres de séjour délivrés aux migrants majeurs de nationalité étrangère fournissent les dates d'arrivées et les durées de titre de séjour des migrants en situation légale. D'autre part les enquêtes annuelles de recensement de la population qui ont lieu chaque année depuis 2004 donnent une photographie de la population vivant en France au début de l'année, permettant en particulier de recueillir la date d'arrivée en France des personnes nées à l'étranger. En se restreignant à un champ commun aux deux sources, constitué des personnes de nationalité étrangère (hors Espace Économique Européen et Suisse), nées à l'étranger, arrivant en France métropolitaine âgées de 19 ans ou plus, l'article propose un modèle de réconciliation des données en estimant par inférence bayésienne des taux de présence à l'enquête annuelle de recensement et des taux d'obtention de titre de séjour selon le nombre d'années passées sur le territoire. Après avoir identifié parmi les migrants qui obtiennent un titre de séjour ceux qui s'installent au moins un an en France, qui est la condition pour être recensé selon les définitions internationales, les estimations indiquent que tous les migrants entrant une année ne sont pas présents (ou identifiés) au recensement la première année et qu'à partir de la troisième année, la part de ceux qui sont présents au recensement diminue en moyenne de 3 % par an (notamment du fait des sorties du territoire). 70 % de ces nouveaux migrants obtiennent leur titre de séjour l'année de leur arrivée ou l'année suivante. De fortes disparités existent entre nationalités quant à la présence estimée au recensement. De même, les femmes semblent plus souvent recensées que les hommes, et les plus jeunes que les plus âgés. Enfin, l'analyse fine des deux sources permet de comprendre leurs forces et faiblesses respectives en vue d'estimer les flux migratoires.

Recensements de la population pour l'étude des reprises démographiques et des migrations résidentielles dans l'espace peu dense (1975-2011)

Pierre Pistre

Cet article détaille l'analyse des reprises démographiques enregistrées dans les parties les moins denses du territoire français depuis les années 1970. Sur le moyen terme, il met en évidence la progression des croissances de population jusqu'aux années 2000, portées par des soldes migratoires de plus en plus positifs. À la périphérie des agglomérations, notamment les plus peuplées, le développement résidentiel se poursuit par extension. Il est alimenté par des profils individuels bien connus (jeunes actifs, professions intermédiaires, employés, ouvriers, etc.), mais la tendance est aussi à une stratification sociale accrue selon le degré de densité de ces espaces. Hors de la proximité urbaine, les croissances démographiques et migratoires sont aussi devenues majoritaires. L'installation de retraités y joue un rôle moteur depuis les années 1980, et l'originalité de la dernière décennie tient à l'arrivée d'une diversité d'actifs en milieu et fin de carrière professionnelle.

L'approfondissement des analyses pour les décennies 2000 et 2010 révèle par ailleurs un tassement des croissances migratoires après la crise économique de 2008-2009. Il concerne l'ensemble des configurations de l'espace peu dense, en particulier les moins denses, mais il fragilise surtout les reprises dans les communes isolées et les petites et moyennes aires urbaines. Selon les catégories d'espace, l'explication paraît tenir à de moindres migrations entrantes des actifs et/ou des groupes socioprofessionnels moyens et supérieurs.

Les tendances récentes, examinées à partir des données détaillées du recensement et des enquêtes annuelles, demandent à être confirmées par l'exploitation des prochaines données migratoires du recensement, disponibles en 2016 (RP 2013). Pour les communes de moins de 10 000 habitants, le développement de méthodes d'étude annualisées est aussi un enjeu important pour mieux suivre la tendance inégale aux reprises démographiques.

L'utilisation de l'automobile par les ménages dans les territoires peu denses : analyse croisée par les enquêtes sur la mobilité et le Recensement de la population

Jean-Paul Hubert, Pierre Pistre et Jean-Loup Madre

Après des décennies ininterrompues de croissance, les distances parcourues en automobile par les ménages ont plafonné puis diminué, d'abord dans les grandes agglomérations, puis progressivement pendant les années 2000, dans des espaces de moins en moins denses. Or bien des explications données à ce phénomène dans le contexte urbain dense ne s'appliquent pas à des territoires où la mobilité dépend essentiellement de l'automobile. L'objectif de cet article est de mieux comprendre comment l'utilisation de l'automobile par les ménages s'est stabilisée puis a diminué dans les territoires peu denses, en combinant les sources statistiques sur la mobilité (*Parc Auto* et ENT-D) et le recensement de la population.

Cet ensemble de sources fait comprendre plusieurs dynamiques parallèles dans un contexte économique peu favorable. Tandis qu'une partie vieillissante des ménages continue à s'équiper en automobile et à rouler davantage, ceux qui disposent déjà de plusieurs véhicules réduisent les kilométrages parcourus en recentrant leurs activités plus près de chez eux et en reportant une partie de leur mobilité à longue distance sur d'autres modes de transport. En outre, les ménages comprenant un ou plusieurs actifs occupés, généralement multi-équipés, affectent de plus en plus leurs véhicules aux navettes domicile-travail, l'accès à l'emploi en dépendant.

Enquête Famille et logements adossée au recensement de 2011 : comment s'adapter à la nouvelle méthodologie des enquêtes annuelles et quels apports ?

Pascale Breuil-Genier, Guillemette Buisson, Isabelle Robert-Bobée et Loïc Trabut

Des enquêtes spécifiques sur la famille sont adossées au recensement de la population depuis 1954. Le changement de méthode du recensement, avec un passage à des enquêtes par sondage chaque année, a nécessité de revoir la méthodologie des enquêtes associées. L'*enquête sur la famille et les logements* de 2011 est la première enquête couplée au nouveau recensement. Elle a permis de tester la faisabilité d'un tel couplage et de mettre en place le protocole de collecte associé, sans impact négatif sur le recensement. Cette enquête apporte, comme les enquêtes Famille précédentes, des informations complémentaires (dénombrement des personnes pacées ou des familles recomposées par exemple) et une meilleure compréhension des résultats du recensement (comparaison des réponses des mêmes personnes dans les deux sources).

Estimations communales exploitant les données de l'enquête Famille et logements 2011 et du recensement : une opération à hauts risques

Pascal Ardilly

À l'été 2013, l'Insee a diffusé à plus de 1 400 communes des estimations d'effectifs pour diverses sous-populations vivant sur leurs territoires : personnes ayant souscrit un pacs, personnes en couple non cohabitant, grands-parents, personnes âgées vivant seules et ayant des enfants résidant à proximité, etc.

Ces estimations communales se sont appuyées sur la collecte nationale de l'enquête *Famille et logements*, enquête de très grande taille associée à l'enquête annuelle de recensement 2011. Afin de leur assurer une qualité suffisante, on a eu recours à une démarche par modélisation, de type « petits domaines ». La première étape s'appuie sur l'ensemble de l'échantillon de l'enquête. Elle consiste à modéliser les probabilités individuelles d'appartenir à la sous-population d'intérêt. Les variables explicatives utilisées sont des variables binaires disponibles dans le recensement : sexe, groupe d'âge, statut matrimonial, etc. Elles délimitent donc des catégories de population au sein desquelles ces probabilités d'appartenir à la sous-population d'intérêt seront considérées homogènes et indépendantes de la commune de résidence. Une fois évaluées ces probabilités, on obtient les estimations communales en les multipliant par les effectifs communaux des catégories associées (fournis par le recensement). Enfin, on procède à un calage sur l'effectif national de la sous-population d'intérêt issu de l'enquête.

Parce qu'elle fonde les estimations sur des échantillons de grande taille, cette procédure réduit beaucoup la variance d'échantillonnage par rapport à une estimation qui utiliserait seulement les informations de l'enquête au niveau communal. En contrepartie, le recours à un modèle génère un biais car on assimile un comportement communal à un comportement supra communal. L'erreur qui en résulte peut être appréciée à un niveau agrégé et on constate, sur les variables diffusées, qu'elle reste le plus souvent très acceptable.

Recensements européens : mouvements centrifuges et forces de rappel

Patrick Festy et Nathalie Le Bouteillec

En abandonnant la forme traditionnelle du recensement exhaustif périodique au tournant des années 2000, la France et l'Insee participent à un mouvement général en Europe de réforme des modalités de réalisation du recensement de la population, en particulier au cours de la dernière décennie. Certains pays donnent un rôle aux registres de population, d'autres aux sondages, alors que d'autres encore maintiennent une forme classique. S'y ajoute une multiplication des modes de recueil de l'information auprès des individus et des ménages. À cette diversité des modes de collecte au sein de l'espace européen viennent s'ajouter enfin des choix sensiblement différents dans les thèmes abordés au cours des recensements et dans la forme donnée aux interrogations sur ces sujets.

Faut-il donc abandonner l'espoir d'établir une carte cohérente de la population européenne et de ses principales caractéristiques ? La réponse de l'Union européenne et d'Eurostat consiste à laisser les pays libres du choix de leurs procédures, à condition que celles-ci respectent un certain nombre de concepts essentiels, afin que des résultats comparables puissent être produits dans les différents pays (règlement du 9 juillet 2008). Cette conciliation entre diversité des modalités et harmonie des résultats s'inscrit dans la perspective ouverte depuis une cinquantaine d'années par l'élaboration de recommandations par les Nations Unies puis la Conférence des statisticiens européens. Le *Census Hub* inauguré en 2014 en est aujourd'hui la forme aboutie.

Recensement 2011 et règlement européen : la procédure d'imputation spécifique à trois variables

Pascal Ardilly

La modification du questionnaire des enquêtes annuelles de recensement à partir de 2011, a suscité de nombreux travaux méthodologiques complexes, tels que ceux concernant l'imputation massive de trois variables dans le recensement de 2011 : l'indicateur de résidence antérieure, la période d'achèvement de l'immeuble et l'activité au sens du BIT.

Pour les deux premières, le contexte est celui d'un changement de l'information collectée à partir de 2011. À chaque fois, deux méthodes sont proposées. La première est une approche assez simple au moyen d'un modèle, basée sur l'hypothèse d'une distribution uniforme. La seconde, plus compliquée, utilise un calage par commune ou groupe de communes afin de retrouver en moyenne, à partir des données collectées par les enquêtes annuelles en ancienne nomenclature, les structures estimées en utilisant les données collectées par les enquêtes annuelles en nouvelle nomenclature.

L'imputation de l'activité BIT (trois modalités : actif occupé / chômeur / inactif), en métropole et dans les DOM, relève de la construction complète d'une nouvelle variable à partir de variables auxiliaires individuelles corrélées à l'activité et disponibles à la fois dans le recensement et dans l'enquête *Emploi*. Pour cela, malgré une différence dans les modes de collecte qui crée une hétérogénéité que l'on parvient à corriger en partie, on utilise une variable de déclaration d'activité « spontanée » disponible dans les deux sources. Elle permet d'estimer une liaison assez satisfaisante avec l'activité BIT. On calcule dans un premier temps des probabilités de passage d'une modalité d'activité spontanée vers une modalité d'activité BIT. Puis ces probabilités sont ajustées afin de retrouver en moyenne les effectifs des différentes modalités d'activité BIT donnés par l'enquête *Emploi* dans chaque croisement région / sexe. En particulier, il faut que les taux de chômage BIT par région/ sexe après imputation subissent des perturbations minimales.