

NICOLE COËFFIC

Le recensement de 1990 : l'enquête de mesure d'exhaustivité

Journal de la société statistique de Paris, tome 134, n° 3 (1993), p. 3-20

http://www.numdam.org/item?id=JSFS_1993__134_3_3_0

© Société de statistique de Paris, 1993, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>



I

COMMUNICATION

LE RECENSEMENT DE 1990 : L'ENQUETE DE MESURE D'EXHAUSTIVITE

par Nicole COEFFIC

Institut National de la Statistique et des Études Économiques (INSEE)

Immédiatement après le recensement de la population de 1990, l'INSEE a réalisé une enquête ayant pour objectif d'apprécier la qualité du dénombrement sur le territoire métropolitain (Corse exclue)¹.

Cette enquête post-censitaire n'était qu'un élément d'un plus large dispositif mis en place en 1990 par l'INSEE pour l'amélioration et le contrôle de la qualité, non seulement de la collecte, mais aussi de la saisie et du chiffrage des documents du recensement. Un numéro d'INSEE-Méthodes, consacré à « Diverses questions méthodologiques concernant le recensement de la population de 1990 », paraîtra en 1994.

Pour mettre au point l'enquête post-censitaire de 1990, nous nous sommes inspirés de l'expérience d'une première enquête que l'INSEE avait menée après le recensement de 1962 [Vangrevelinghe G., 1963]. Que l'INSEE n'ait pas réalisé d'enquête post-censitaire entre 1962 et 1990 a sans doute plusieurs explications. La première est que la réalisation de chaque recensement de la population mobilise une très grande partie du personnel des Directions régionales de l'INSEE et, qu'en raison de la faiblesse des moyens disponibles, il nous est généralement difficile de programmer de nouveaux travaux pendant cette période. La seconde est qu'une telle étude représente une tâche très délicate, dont on peut craindre qu'elle ne reconduise une petite partie des erreurs du recensement. Néanmoins, même si les résultats obtenus comportent une part d'incertitude, ils donnent des ordres de grandeur et des enseignements que nous jugeons intéressants.

Tout d'abord, nous rappellerons les grandes lignes de l'organisation du recensement et les sources possibles des erreurs de dénombrement qui en découlent. Puis, nous présenterons la méthodologie de l'enquête et quelques résultats.

1. Communication faite en 1991 à la Société ainsi qu'à la 49^e session de l'Institut International de Statistique à Florence en août 1993.

1. Les grandes lignes de l'organisation du recensement

Le recensement français s'effectue par collecte sur le terrain [Léry, 1990]. En 1990, cette collecte s'est déroulée en deux temps :

- En février 1990, 1 400 agents recenseurs ont été chargés de recenser sur le terrain environ 2 000 000 de personnes vivant dans des « communautés », c'est-à-dire les casernes, les internats des établissements d'enseignement, les établissements pénitentiaires, les foyers de travailleurs ou d'étudiants, les maisons de retraite...
- En mars-avril 1990, 110 000 agents recenseurs ont eu à dénombrer tous les logements d'habitation (environ 26 000 000) : résidences principales, mais aussi logements vacants et résidences secondaires, et à recenser les occupants des résidences principales (un peu plus de 54 500 000 personnes). Le 1^{er} jour du recensement (5 mars) a été consacré au recensement des habitations mobiles et des personnes sans domicile fixe.

Les agents recenseurs avaient la tâche de remettre les imprimés de collecte au responsable de chaque communauté ou à l'un des occupants de chaque résidence principale, puis de les récupérer une fois qu'ils étaient remplis.

Pour permettre un ratissage systématique des communautés et des logements, le territoire avait été découpé à l'avance en 320 000 districts de recensement. Dans les zones à forte densité de population, un district correspondait généralement à un pâté de maison, voire à une portion de pâté de maison si celui-ci était trop gros (plus de 600 habitants). Dans les zones rurales, les communes de moins de 500 habitants constituaient un seul district, les communes de taille supérieure étaient découpées en plusieurs districts, correspondant à des portions de territoire bien délimitées. Chaque agent recenseur avait la responsabilité de recenser un ou plusieurs districts.

Le découpage du territoire français a été préparé pendant l'année ayant précédé le recensement. Il a consisté en une mise à jour des plans déjà établis pour le précédent recensement de 1982.

Tout le travail de préparation et d'exécution de la collecte du recensement a été effectué par l'INSEE en étroite association avec les maires et le personnel des mairies des 36 500 communes françaises.

Les questionnaires du recensement, collectés par les agents recenseurs, transitaient par les mairies, avant d'être acheminés dans les 22 Directions régionales de l'INSEE.

Une innovation du recensement de 1990, destinée à améliorer la qualité de la collecte, a consisté à faire *un contrôle d'exhaustivité à l'aide du fichier de la « taxe d'habitation »* (impôt payé pour chaque logement). Dans la plupart des communes de plus de 20 000 habitants, des agents ont été chargés de vérifier que, sur chaque district, les nombres de logements, classés par catégorie, étaient proches de ceux figurant sur les listings édités à partir de ce fichier. En cas d'incohérences significatives, les agents recenseurs sont retournés sur le terrain pour corriger et compléter le recensement.

2. Sources possibles des erreurs de dénombrement

Les sources possibles des erreurs de dénombrement sont les suivantes :

- *Des logements ou des communautés peuvent être omis.* Si le logement omis est une résidence principale ou une communauté, leurs occupants ont toute chance d'être omis au recensement, sauf si, par exception, ils sont recensés ailleurs.
- *Des logements peuvent être classés dans une catégorie erronée.* Ainsi *des résidences principales peuvent être considérées à tort* comme résidences secondaires ou logements vacants, *donc recensées sans occupants.* Ce type d'erreur se produit principalement pour les ménages absents pendant la période de recensement, ceux dont les membres effectuent des déplacements fréquents et ont des horaires particuliers, ceux qui partagent leur temps entre deux résidences, ... Les personnes qui résident dans ces résidences principales mal classées risquent alors d'être omises au recensement, sauf si elles sont recensées ailleurs (par exemple, par erreur, dans leur résidence secondaire).

A l'inverse, des résidences secondaires ou des logements vacants peuvent être *classés à tort résidences principales.* Les personnes qui y sont recensées sont alors susceptibles d'être comptées deux fois au recensement, si elles sont aussi recensées dans leur résidence principale. Ces classements erronés peuvent provenir d'erreurs de collecte, mais aussi résulter de fraudes. Il arrive en effet que des mairies cherchent par ce biais à « gonfler » le nombre des habitants de leur commune.

- Une mairie qui voudrait *frauder* peut aussi créer des logements avec une adresse fictive et des occupants fictifs.
- Certaines personnes *refusent* de se faire recenser. Certes, ces refus n'entraînent pas systématiquement des omissions au recensement. En effet, si *tout le ménage refuse*, normalement, l'agent recenseur doit tout de même le recenser. Il a la consigne de se renseigner auprès des voisins, et d'ouvrir le nombre de bulletins individuels, qui convient et portant, pour le moins, la mention « refus ». Mais cette situation peut être source d'erreurs : l'agent recenseur peut oublier d'appliquer cette consigne ; ou obtenir des voisins des renseignements erronés sur le nombre de bulletins individuels à créer.

Par contre, si, *à l'intérieur d'un ménage, une personne veut échapper au recensement*, alors que les autres membres du ménage remplissent un bulletin individuel, l'erreur de recensement est quasiment certaine, car il y a peu de chances que l'agent recenseur puisse repérer ce type d'omission et la corriger.

- Des personnes peuvent aussi être omises isolément dans leur résidence principale, si la personne qui remplit les documents du recensement *oublie de les déclarer.* Ces omissions involontaires concernent assez souvent les personnes ne faisant pas partie du noyau familial (salariés logés chez leur employeur, étudiants logés chez un particulier, personnes âgées hébergées chez leurs enfants, ...).

- A l'inverse, dans une résidence principale recensée, des personnes peuvent être *comptées en trop*, si elles résident la plus grande partie de l'année ailleurs. Elles risquent alors d'être comptées deux fois. Par exemple, un étudiant, qui vit habituellement dans une chambre en ville, peut, par erreur, être déclaré par ses parents comme habitant encore avec eux.

De même, les personnes *vivant dans une communauté et ayant conservé des liens avec un logement familial* peuvent être recensés à la fois dans cette communauté et dans ce logement. Aussi, pour les élèves internes et les militaires du contingent, la méthode de recensement a-t-elle été conçue pour limiter le risque de double compte : le bulletin individuel collecté en février dans la caserne ou l'internat a été, chaque fois que possible, réintégré dans la résidence personnelle de l'intéressé pendant le recensement des logements (au moment de la réintégration, on a pu alors supprimer les éventuels doubles). Mais cette procédure de réintégration n'a pas été prévue pour les membres des autres communautés.

Enfin, le *déménagement* pendant la période de recensement est source d'erreurs. La période de recensement des logements s'est étalée sur plusieurs semaines. Quand un ménage, ou une personne dans un ménage, *change pendant ce laps de temps de résidence principale*, il risque d'être omis, ou au contraire d'être compté deux fois.

Ce même type de problème se pose pour les personnes *ayant quitté un logement pour vivre en communauté ou vice-versa*, puisque le recensement des communautés a précédé d'au moins un mois celui des logements. On savait à l'avance que séparer le recensement des communautés de celui des logements risquait d'occasionner quelques erreurs de dénombrement, mais une telle organisation permettait, à l'inverse, des gains de qualité substantiels. En effet, les agents chargés de recenser les communautés ont pu recevoir une formation spécialisée et approfondie, et concentrer tous leurs efforts au recensement des seules communautés sur un court intervalle de temps.

3. La méthode de l'enquête post-censitaire de 1990

Les principaux points de la méthode de l'enquête post-censitaire étaient les suivants :

- des enquêteurs de qualité devaient dénombrer tous les logements situés dans un échantillon de zones géographiques. Pour chaque logement, un questionnaire approfondi permettait d'en connaître les occupants, même occasionnels, de noter comment ces occupants partageaient leur temps entre ce logement et d'éventuels autres lieux d'habitation, d'obtenir l'adresse précise de ces autres lieux de résidence ;
- puis le personnel de bureau effectuait un 1^{er} appariement avec le recensement, qui consistait à rechercher tous les documents du recensement remplis pour les logements situés sur les zones-échantillon, afin de déterminer si les personnes enquêtées avaient été ou non recensées à l'adresse du logement enquêté ;

LE RECENSEMENT DE 1990

- en cas d'incohérence inexplicable entre l'enquête et le recensement, les enquêteurs devaient retourner sur le terrain pour obtenir des explications et corriger l'enquête si elle comportait des erreurs ;
- un 2^e appariement était ensuite effectué par le personnel de bureau. Il s'agissait de regarder dans les documents du recensement si les enquêtés avaient été ou non recensés à l'une de leurs éventuelles autres adresses (autres que celle du logement enquêté).

En rapprochant ensuite les résultats du 1^{er} et du 2^e appariement, on pouvait donc déterminer, pour chaque enquêté, s'il avait été recensé 0 fois, 1 fois (à l'adresse du logement enquêté ou à une autre adresse) ou 2 fois.

3.1. Le calendrier de la collecte

Le calendrier de la collecte était soumis à deux contraintes. L'enquête devait se dérouler le plus près possible de la date de fin du recensement. En effet, des mouvements pouvaient se produire entre la date du recensement et celle de l'enquête : déménagements, décès et naissances. Ceux-ci rendaient le travail de comparaison entre le recensement et l'enquête plus délicat. On avait donc intérêt à ce qu'ils soient les moins nombreux possibles. A l'inverse, pour envoyer les enquêteurs sur le terrain, il fallait attendre que l'agent recenseur ait terminé son travail sur le (ou les) district(s) correspondant à l'aire à enquêter. Il fallait en effet éviter que le recensement puisse être modifié du fait du passage de l'enquêteur, puisque le but de l'enquête n'était pas d'améliorer le recensement, mais d'en mesurer les limites.

Le démarrage de la collecte de l'enquête a donc été échelonné en fonction des dates de retour des documents du recensement. Finalement, la majorité des interviews ont eu lieu entre la mi-avril et la fin juin.

3.2. Echantillon de la collecte

L'enquête a été réalisée auprès d'un échantillon aréolaire. Cet échantillon n'a pas été constitué pour la circonstance, il avait été tiré en 1983 sur la base du recensement de 1982 pour les enquêtes sur l'Emploi. L'échantillon de chaque enquête Emploi étant renouvelé par tiers chaque année, on avait tiré une provision d'échantillons assez large permettant un renouvellement sur une dizaine d'années. On a donc pu utiliser pour l'enquête post-censitaire un échantillon n'ayant jamais servi.

Les aires ont été tirées en 1983 à l'intérieur de strates définies par le croisement des régions et de 10 catégories d'agglomérations classées en fonction du nombre d'habitants. Pour éviter le découpage de tout le territoire national, on a procédé dans chacune des strates précédemment définies à un tirage à plusieurs degrés. Le dernier degré du tirage a consisté à découper l'unité tirée au sort en zones d'environ 200 logements. Ce travail de découpage en zones a été fait à l'aide de cartes géographiques. Ensuite, on tirait au sort une zone-échantillon et un plan très détaillé de cette dernière était établi par un enquêteur à partir de relevés de terrain. La zone-échantillon était alors découpée en aires d'une quarantaine de logements.

Le découpage devait obéir aux impératifs suivants :

- chaque zone, comme chaque aire, devait avoir des limites repérables sur le terrain ;
- les zones définies devaient être aussi semblables que possible les unes aux autres, ce qui a conduit à découper des zones ayant une hétérogénéité interne aussi grande que possible : dans le cas d'une commune rurale par exemple, toute zone devait comprendre une partie du bourg, une partie d'habitat dispersé et une partie de terrains non bâtis. On devait, dans la mesure du possible, s'efforcer de respecter cette même règle pour le découpage en aires, ce qui pouvait conduire à définir des aires constituées de 2 morceaux séparés : une partie d'habitat dense, une partie des écarts.

L'échantillon de l'enquête post-censitaire était constitué d'environ 650 aires, comportant donc toutes au moment du tirage une quarantaine de logements. Au total, le nombre de logements de l'échantillon était d'à peu près 30 000 logements, ce qui correspond à un taux de sondage d'environ 1/900.

Lors de l'enquête post-censitaire, l'enquêteur devait repérer de manière exhaustive tous les logements se trouvant à l'intérieur des limites des aires-échantillon, y compris les logements neufs construits depuis la date du tirage de l'échantillon.

L'INSEE a choisi d'utiliser un échantillon aréolaire, parce que c'était la seule solution permettant de repérer à l'enquête les logements oubliés au recensement, grâce à un « ratissage » systématique des logements situés sur les aires-échantillon, effectué par des enquêteurs de qualité.

Certes, ce type d'échantillon a un inconvénient. Certaines erreurs au recensement peuvent dépendre de circonstances locales (capacités de l'agent recenseur, structure de l'habitat) et ne pas être réparties uniformément sur tout le territoire. L'échantillon aréolaire enregistre donc des « effets de grappe », qui diminuent un peu la précision des estimations. Les aires tirées étaient de petite taille de manière à minimiser cet inconvénient.

Du fait de son échantillon aréolaire, l'enquête mesure assez mal le volume des fraudes pouvant être commises par certaines mairies et consistant en ajouts de logements qui n'existent pas ou de personnes dans des logements déjà recensés. Ces fraudes sont en effet étroitement localisées. De plus, les logements qui sont créés frauduleusement avec une adresse fictive ont, pour certains, une probabilité nulle de figurer dans une aire d'un échantillon aréolaire, si l'adresse comporte un nom de rue qui n'existe pas. Heureusement, en 1990, les risques de fraude ont été extrêmement limités en raison d'un accroissement des opérations de contrôle et de suivi pendant la collecte du recensement.

3.3. Champ de la collecte

Seuls les logements étaient dans le champ de l'enquête : autrement dit, les « communautés », telles que casernes, internats, foyers de travailleurs, ... n'étaient pas enquêtées. Cependant, une partie des personnes logeant dans des communautés

étaient prises en compte dans l'enquête, dès lors qu'elles conservaient un lien avec un ménage « ordinaire » (c'est-à-dire l'ensemble des occupants d'un logement) et étaient enquêtées dans ce ménage. Les personnes sans domicile fixe et les personnes vivant habituellement dans une habitation mobile étaient aussi hors champ de l'enquête.

3.4. Exécution de l'enquête sur le terrain et questionnaires de l'enquête

Une des principales difficultés de l'enquête était d'éviter que certaines erreurs de recensement (omissions de logements ou de personnes) ne se reproduisent lors de la collecte. La méthode d'enquête a donc été conçue pour tenter d'échapper le plus possible à cet écueil.

Tout d'abord, l'enquêteur devait bien repérer sur le terrain les limites de l'aire-échantillon qu'il avait à enquêter, et, à l'intérieur de ces limites, faire un inventaire complet des immeubles et des logements d'habitation et même des locaux non habités, mais pouvant servir à l'habitation.

Ensuite, l'enquêteur ne devait pas se contenter d'interviewer les ménages déclarant que le logement enquêté était leur résidence principale, mais il devait interviewer tous les ménages qu'il lui était possible de joindre dans l'aire, et n'arrêter toute tentative de prise de contact que s'il avait l'assurance (par les voisins généralement) que le logement était occupé moins de 3 mois par an. On voulait, par cette méthode, couvrir un champ un peu plus large que les résidences déclarées principales par leurs occupants, et ne pas risquer de manquer les résidences que les enquêtés (ou leurs voisins) déclarent secondaires, alors qu'en réalité elles sont occupées par les ménages plus longtemps sur l'année que leur(s) autre(s) résidence(s). En effet, on a vu qu'un certain nombre d'omissions de personnes au recensement peuvent provenir de ce que certaines résidences principales sont classées à tort secondaires. Il fallait éviter que les enquêteurs, à leur tour, ne reproduisent le même type d'erreur.

Quand l'enquête était acceptée par un ménage, les données recueillies étaient les suivantes :

- liste des personnes résidant même très occasionnellement dans le logement ou y ayant résidé depuis le 1^{er} janvier ;
- puis, pour chaque personne,
 - date de son arrivée éventuelle dans le logement ou de son départ (en cas d'emménagement, déménagement, décès, naissance après le 1^{er} janvier) ; durée d'occupation du logement depuis un an ;
 - liste des « autres résidences », dont elle dispose en permanence (par exemple, résidence secondaire, ...), ou dans lesquelles elle a logé, même occasionnellement, depuis le 1^{er} janvier (par exemple, logement de ses parents, établissement collectif : caserne, internat, foyer, ...) ; adresse précise de ces « autres résidences » et durée d'occupation depuis un an ;
- enfin, toujours pour chaque personne, renseignements sur son état civil (sexe, âge, nationalité, ...) et sa situation professionnelle.

LE RECENSEMENT DE 1990

Sur ce questionnaire , on peut faire plusieurs remarques :

- Il permettait de *connaître les mouvements récents de personnes dans le logement* (emménagements, déménagements,...) susceptibles d'expliquer les divergences entre la liste des personnes enquêtées et celle des personnes recensées. Au lieu de prendre comme période de référence, pour l'observation de ces mouvements, l'intervalle séparant la date du recensement et celle de l'enquête, on a préféré faire débuter cette période au 1^{er} janvier, pour deux raisons : on ne voulait pas focaliser l'interview sur la date du recensement, pour éviter que les enquêtés ne cherchent à être cohérents avec ce qu'ils avaient déclaré lors du recensement. Par ailleurs, couvrir une période plus large que celle commençant à la date du recensement nous paraissait permettre d'éviter des oublis de déclaration de certains mouvements.
- Le questionnaire était conçu de manière à *enquêter la population la plus large possible*. Au stade de l'enquête sur le terrain, on voulait éviter avant tout d'omettre des personnes. Ainsi, dans le questionnaire, étaient rattachées au logement enquêté un nombre très large de personnes, notamment toutes celles qui y résidaient occasionnellement. En effet, bien que résidant épisodiquement dans un logement enquêté, certaines catégories de personnes pouvaient n'avoir aucune chance d'être prises en compte par l'enquête ailleurs : celles qui effectuent des déplacements longs et fréquents ; celles qui vivent habituellement dans une communauté ; les personnes, sans résidence personnelle, qui sont hébergées tantôt chez les uns , tantôt chez les autres, ... Aussi, après que l'enquêté avait donné spontanément la liste des occupants du logement, l'enquêteur lui demandait de citer toutes les personnes ayant habité le logement depuis le 1^{er} janvier, puis toutes les personnes absentes et logeant temporairement ailleurs, « par exemple, jeune vivant en caserne, ... ; personne à l'hôpital ou logée ailleurs pour son travail ». Enfin, il posait une question sur la présence éventuelle de locataires ou de sous-locataires, de salariés logés, d'enfants en nourrice. Ces catégories risquaient en effet d'être souvent spontanément omises à l'enquête (comme au recensement).
- On demandait aux enquêtés *la liste de leurs « autres résidences »*, pour pouvoir vérifier ultérieurement s'ils y avaient été ou non recensés. *Là encore, on a essayé de « ratisser large »*. L'objectif était de ne pas omettre « d'autre résidence », et on a considéré qu'il n'était pas gênant de prendre en compte, à l'inverse, des « autres résidences » en trop, c'est-à-dire des résidences où les personnes avaient très peu de chances d'avoir été recensées. Par exemple, les personnes hospitalisées fin janvier 1990 pour un séjour de longue durée (en principe d'au moins 3 mois) devaient être recensées dans l'hôpital. Pour ne pas manquer des personnes qui auraient été recensées dans un hôpital, on a considéré qu'un hôpital était une « autre résidence », dès lors que l'enquêté y avait séjourné fin janvier, pour un séjour d'au moins un mois, c'est-à-dire d'une durée inférieure à celle requise normalement pour être recensé pendant une hospitalisation.

3.5. 1^{er} appariement manuel et retours sur le terrain

Une fois la collecte terminée, dans chacune des Directions régionales de l'INSEE, le personnel de bureau chargé de l'enquête devait procéder à un appariement manuel des questionnaires de l'enquête et des documents remplis au recensement, pour tous les logements situés sur les aires de l'enquête. Il s'agissait de regarder si, sur le territoire correspondant à l'aire-échantillon, on avait bien repéré les mêmes logements lors de l'enquête et lors du recensement. En même temps, on devait appairer les personnes.

L'appariement des logements n'était pas toujours simple. Par exemple, dans les zones rurales, les adresses sont souvent très peu formalisées et étaient très imprécises sur les documents du recensement, en particulier dans les hameaux. Pour essayer de rapprocher les logements enquêtés de ceux qui avaient été recensés, il a fallu parfois retourner sur le terrain.

Mais l'appariement des personnes était encore plus délicat. En effet, le champ de l'enquête était plus large que celui du recensement, puisqu'on enquêtait des personnes dont le logement enquêté n'était pas la résidence principale ; par ailleurs, des mouvements de personnes pouvaient s'être produits entre la date du recensement et celle de l'enquête. Certaines divergences entre le recensement et l'enquête étaient donc « normales ». Ainsi, il était normal qu'une personne enquêtée n'ait pas été recensée à l'adresse du logement enquêté, si elle y avait emménagé après la date du recensement ; si le logement enquêté n'était pas sa résidence principale ; si, à la date du recensement, elle n'était pas encore née, etc. Par contre, chaque fois qu'une divergence entre le recensement et l'enquête était, au vu des renseignements collectés lors de l'enquête, inexplicable, il était demandé aux enquêteurs de retourner sur le terrain pour obtenir des compléments d'information. Ces retours auprès des enquêtés permettaient non seulement de compléter les questionnaires, mais aussi de corriger d'éventuelles erreurs commises lors de l'enquête. Il fallait en effet absolument éviter que des erreurs de l'enquête puissent être imputées à tort au recensement.

La difficulté de ce 1^{er} appariement était donc de bien isoler les cas de non-appariements nécessitant un « retour terrain ».

En outre, lors de ce 1^{er} appariement, les agents devaient vérifier que les personnes enquêtées n'avaient pas omis de déclarer une « autre résidence » : par exemple, ils devaient s'assurer que les personnes ayant emménagé récemment dans le logement enquêté avaient bien déclaré l'adresse de leur résidence antérieure (car elles pouvaient y avoir été recensées). Si cette information était manquante, les enquêteurs devaient contacter de nouveau le ménage, par visite ou téléphone, pour obtenir le renseignement.

3.6. 2^e appariement manuel

Rappelons que, sur le questionnaire de l'enquête, figurait, pour chaque personne enquêtée, la liste de toutes leurs « autres résidences », autres que celles correspondant au logement enquêté. Un 2^e appariement (complètement séparé du 1^{er}) a consisté à

rechercher les documents du recensement correspondant à toutes ces « autres résidences », et à noter si les enquêtés y avaient été ou non recensés.

Contrairement aux logements enquêtés qui étaient regroupés de manière aréolaire, les « autres résidences » étaient dispersées sur tout le territoire national. On comprend que, dans ces conditions, la qualité de ce 2^e appariement reposait en grande partie sur celle des adresses des « autres résidences » données par les enquêtés.

En cas d'impossibilité à apparier « l'autre résidence » avec un logement ou une communauté recensée, on ne concluait au non-recensement *certain* de la personne dans son « autre résidence » que si l'adresse fournie avait pu être validée par consultation du Minitel. Dans le cas contraire, si l'adresse n'avait pu être validée ou si elle était manquante, on était incapable de déterminer avec certitude si l'enquêté avait été ou non recensé dans son « autre résidence ». C'est pourquoi, on le verra plus loin dans la présentation des résultats, parmi les omissions au recensement comptées par l'enquête, on isole celles qui ne sont que « probables » ou « possibles » en sus de celles sont « caractérisées ». Une marge d'incertitude similaire affecte le nombre des personnes comptées une fois : elle est constituée de personnes recensées « peut-être ou probablement » une fois.

3.7. Exploitation

Lors de la collecte, on devait enquêter toutes les personnes logeant ou ayant logé depuis le 1^{er} janvier 1990, même très occasionnellement, dans les logements de l'échantillon. Si, au stade de l'exploitation, on avait gardé dans le champ des tableaux tous les enquêtés affectés de leur coefficient de pondération initial, on aurait sur-représenté les personnes ayant logé dans plusieurs logements entre le 1^{er} janvier et la date de l'enquête, puisque, en théorie, celles-ci étaient représentées par l'enquête plusieurs fois.

Il s'agissait donc d'éviter ce biais. Pour ce faire, on a choisi la méthode suivante.

On a déterminé quelle était, pour chaque enquêté, *sa résidence principale à la date de l'enquête*. Si cette résidence principale correspondait au logement-échantillon, celui où la personne avait été enquêtée, on gardait la personne dans le champ de l'exploitation. Quand cette résidence principale correspondait à une « autre résidence » que le logement enquêté, la personne était exclue du champ de l'exploitation.

Pour déterminer quelle était la résidence principale des personnes à la date de l'enquête, on ne se fondait pas sur la déclaration spontanée des enquêtés. Si on l'avait fait, on risquait de reproduire certaines des erreurs du recensement. Au contraire, la détermination de la résidence principale à la date de l'enquête a été effectuée a posteriori par le personnel de bureau selon des règles précises. Lorsque la personne enquêtée avait plusieurs résidences, la règle générale était de classer en résidence principale celle où elle logeait la plus grande partie de l'année (les durées d'occupation de ses différents logements figuraient sur le questionnaire). Par contre, pour un certain nombre de cas particuliers – par exemple, celui d'un salarié logé par son employeur, ayant par ailleurs un domicile familial –, la règle établie ne tenait pas

compte de la durée d'occupation, mais elle désignait clairement le logement qui devait être classé conventionnellement en résidence principale (dans l'exemple cité, c'était le logement chez l'employeur). Autre cas particulier, celui des personnes résidant habituellement dans une communauté, mais enquêtées dans leur logement familial : on considérerait que ce logement familial était leur résidence principale.

A priori, il aurait pu paraître préférable de coller au recensement en gardant dans le champ de l'exploitation les personnes dont le logement enquêté était la résidence principale à la date du recensement (plutôt qu'à celle de l'enquête). A l'examen, on s'aperçoit que la mise en pratique de cette méthode pose problème. En effet, pour les ménages ayant changé de résidence principale entre les deux dates, il aurait été difficile de remplir un questionnaire à l'adresse du logement qu'ils occupaient à la date du recensement, puisque, précisément, à la date de l'enquête, ils n'y étaient plus présents. Or, certaines données, indispensables à l'enquête, ne sont pas de celles qu'on peut obtenir facilement auprès des voisins. Par exemple, on a besoin de connaître, pour chacun des membres du ménage enquêté, l'adresse de leurs éventuelles « autres résidences » et notamment celle de leur nouvelle résidence principale, afin de vérifier qu'ils n'y ont pas été recensés.

La méthode retenue a l'avantage que, pour les ménages ayant changé de résidence principale entre la date du recensement et celle de l'enquête, les renseignements sur leur situation actuelle et leur situation à la date du recensement peuvent être facilement collectés à leur nouvelle adresse. Elle présente le petit inconvénient, incontournable, que les personnes ayant quitté le territoire français métropolitain ou décédées entre la date du recensement et celle de l'enquête sont hors champ de la collecte. Comme ces populations sont peu nombreuses, cette exclusion n'a sans doute pas un effet significatif sur les résultats.

3.8. Traitement des non-réponses

Bien évidemment, dans cette enquête, il fallait essayer de limiter au maximum les non-réponses, car on pouvait supposer a priori une corrélation forte entre la non-réponse du ménage à l'enquête et un recensement incorrect de ses membres. Aussi avait-on prévu diverses formes de rattrapages pour essayer d'obtenir des renseignements auprès des ménages non-répondants. A un ménage qui refusait l'enquête, les Directions régionales de l'INSEE commençaient par adresser, comme d'habitude, une lettre de rappel ; si le ménage persistait dans son refus, on lui adressait un questionnaire postal succinct. Pour un ménage absent, la procédure de rattrapage était la suivante. L'enquêteur essayait d'obtenir auprès des voisins l'adresse d'une éventuelle autre résidence, et des tentatives d'interview par téléphone devaient être effectuées à cette deuxième résidence. En réalité, ce mode de rattrapage a été très peu utilisé. Par contre, plus systématiquement, lors de la phase des « retours terrain », les enquêteurs ont tenté de contacter à nouveau les ménages précédemment absents. Puis, en dernier recours, les Directions régionales leur ont envoyé un questionnaire postal.

Finalement, les rattrapages ont permis de récupérer l'interview de 2,7 % des résidences principales, représentant environ la moitié des non-réponses initiales. Il

subsiste néanmoins un taux d'échec (rapporté aux seules résidences principales) de 2,5 %.

Si le taux de non-réponses est faible par rapport à une autre enquête, il apparaît élevé si on le compare au taux d'erreur de recensement mesuré par l'enquête, en raison de la corrélation possible entre non-réponse à l'enquête et non-recensement. Généralement, on redresse les non-répondants en les supposant semblables aux répondants de mêmes caractéristiques. Dans cette enquête, il nous a paru préférable de les assimiler plutôt aux ménages qui, n'ayant pas répondu lors de la phase principale de l'enquête, avaient finalement fait l'objet d'un rattrapage (par questionnaire postal, par téléphone ou lors des retours terrain). En effet, ces ménages « non-répondants finalement répondants », difficiles à joindre lors de la phase principale de l'enquête, ou parfois récalcitrants pour répondre, ont été sans doute recensés dans des conditions plus proches de celles des non-répondants que de celles des ménages ayant répondu à l'enquête sans difficulté.

Pour effectuer ce redressement, on a réparti ménages non-répondants et « ménages non-répondants finalement répondants » en 4 classes. La 1^{re} était constituée des ménages dont le logement enquêté (qui était leur résidence principale à la date de l'enquête) avait été, au recensement, soit oublié, soit classé dans une autre catégorie (résidence secondaire, logement vacant ou occasionnel). Les autres ménages, dont le logement enquêté avait été recensé comme résidence principale, étaient répartis en 3 classes, selon la taille d'agglomération (variable corrélée avec le risque d'erreur au recensement comme, on le verra, les résultats le montrent). Dans chaque classe ainsi définie, on a corrigé le poids des « non-répondants finalement répondants » de manière à compenser les ménages défailants.

Finalement, un tel redressement a augmenté le taux global d'omission et de double compte au recensement, respectivement, de 0,1 point.

4. Quelques résultats

4.1. Résultats globaux

L'enquête post-censitaire permet de conclure que le taux d'omission de la population au recensement de 1990 est sans doute proche de 2 % ; ces omissions sont, d'une certaine façon, compensées partiellement par des doubles comptes, qui représentent environ 1 % de la population.

Les omissions sont classées en trois catégories, selon leur plus ou moins grand degré de certitude.

- Les omissions caractérisées correspondent à des oublis de recensement repérés sans ambiguïté par l'enquête : les personnes ont été omises dans le logement enquêté et dans leurs éventuelles « autres résidences ».

LE RECENSEMENT DE 1990

- Les *omissions probables ou possibles* correspondent à des enquêtés non recensés dans le logement enquêté (qui était leur résidence principale à la date de l'enquête) et ayant déclaré une deuxième adresse qui n'a pas pu être appariée avec le recensement :
 - dans le cas des *omissions probables*, l'enquêté a donné l'adresse précise de son « autre résidence », mais cette adresse est un peu incertaine, car elle n'a pas été trouvée sur Minitel, c'est-à-dire dans le fichier des abonnés au téléphone. Or cet échec ne permet pas de préjuger de la validité de l'adresse, car beaucoup « d'autres résidences » sont des résidences secondaires et celles-ci, souvent, ne sont pas équipées du téléphone ;
 - dans le cas des *omissions possibles*, la deuxième résidence n'a pas pu être appariée parce qu'on n'en connaissait pas l'adresse (celle-ci était manquante ou visiblement fausse). On sait que la personne a été omise dans le logement enquêté (qui était sa résidence principale à la date de l'enquête), mais on ne sait pas si elle n'a pas été recensée dans cette deuxième résidence.

*Tableau 1. Erreurs de recensement des personnes en 1990
(en % exprimé par rapport à la population totale)*

Omissions caractérisées	1,6
Omissions probables	0,3
Omissions possibles	0,2
Recensées une seule fois	93,7
Recensées probablement ou peut-être une fois (doubles comptes possibles)	3,4
Doubles comptes caractérisés	0,7
Doubles comptes probables	0,1
Population totale	100,0

Parmi les personnes recensées une fois, on isole :

- les personnes dont on sait *sans ambiguïté* qu'elles ont été recensées une fois ;
- les *personnes probablement ou peut-être recensées une fois* : cette catégorie est constituée de personnes recensées au moins une fois, mais ayant déclaré une « autre résidence » qui n'a pas pu être appariée avec le recensement (ou bien l'adresse de cette deuxième résidence était manquante, incomplète ou fausse ; ou l'adresse était précise, mais elle n'a pas été trouvée au Minitel, et on ne sait pas si le non-appariement avec le recensement n'est pas due à une erreur d'adresse). Une petite partie de ces personnes « probablement ou peut-être recensées une fois » ont sans doute été en réalité comptées deux fois au recensement.

Parmi les **doubles comptes**, on isole deux catégories :

- les *doubles comptes caractérisés* : on a retrouvé pour les personnes deux bulletins individuels remplis au recensement, chacun dans deux logements différents.
- les *doubles comptes probables*. Il s'agit d'élèves internes ou des militaires du contingent pour lesquels on a trouvé un 1^{er} bulletin individuel dans un logement d'habitation, qui figurent sur la liste des noms des membres recensés de l'internat ou de la caserne, mais pour lesquels le 2^e bulletin individuel, rempli dans la caserne ou l'internat, est introuvable. Ce 2^e bulletin peut avoir été compté dans un autre logement non déclaré à l'enquête, parmi les « réintégrables non réintégrés » ...

4.2. Origine des omissions

En sus de cette estimation globale de l'erreur de couverture, l'enquête donne des éléments sur l'origine des oublis de recensement des personnes.

Ainsi le risque d'omission est plus important pour les personnes **rattachées à plusieurs lieux d'habitation** (y compris celles qui ont déménagé pendant la période de recensement). Le pourcentage d'omissions (caractérisées et probables) est particulièrement élevé pour les personnes qui ont changé de résidence principale pendant le recensement (15,8 %), il est encore fort pour celles qui sont rattachées à une communauté (5,7 %), un peu moins pour celles qui ont un point d'attache dans plusieurs logements (2,5 %). Il est sensiblement plus faible pour les personnes stables, qui n'ont qu'un seul logement (1,3 %). Toutefois, ces dernières étant largement majoritaires dans l'ensemble de la population, leur part parmi les omissions l'est aussi : 57 %. Les personnes rattachées à une communauté ne représentent que 9 % des omissions, celles rattachées à plusieurs logements, 17 %, celles ayant changé de résidence principale pendant le recensement, 17 % également.

Sur 100 personnes omises, 60 correspondent à des **ménages omis dans leur totalité** : 11 à des ménages ayant déménagé, 25 à des ménages non-mobiles dont la résidence principale a été oubliée au recensement, 24 à des ménages non-mobiles dont la résidence principale a été classée à tort logement vacant ou résidence secondaire.

4.3. Résultats en fonction de quelques critères

L'enquête permet aussi de montrer que la fréquence des erreurs *varie en fonction de divers critères socio-démographiques ou géographiques*.

Ainsi, le taux d'omission varie en fonction de l'âge . Les nouveaux-nés, ceux nés en 1990 avant la date du recensement, sont assez souvent omis, mais ils sont peu nombreux, et les erreurs qui les affectent pèsent peu sur l'ensemble. Les omissions sont aussi particulièrement fréquentes pour la tranche d'âge des 20-29 ans. Le pourcentage d'omis diminue ensuite légèrement en même temps qu'augmente l'âge, pour remonter un peu au-delà de 70 ans. Quant aux doubles comptes, ils sont très fréquents

LE RECENSEMENT DE 1990

entre 20 et 24 ans, plus rares pour les très jeunes (moins de 10 ans) ou les personnes « d'âge mûr » (après 35 ans).

Tableau 2. Erreurs de recensement des personnes selon quelques critères (% exprimé par rapport à l'ensemble de chaque catégorie de population)

	% d'omissions caractérisées et probables	% de doubles comptes caractérisés et probables
ENSEMBLE	1,9	0,8
20-24 ans	4,7	2,5
25-29 ans	3,3	0,9
Personne vivant seule	3,5	1,4
Français	1,6	0,8
Étrangers	3,6	0,5
Chômeurs	4,5	0,6
Étudiants (18 ans et plus)	5,1	3,6
Communes rurales	1,1	0,7
Agglomération parisienne	3,2	0,7

C'est donc pour les « jeunes adultes » de 20 à 30 ans et surtout ceux de 20 à 25 ans qu'on observe le plus d'erreurs. Ce n'est pas étonnant. A cette étape de la vie, les liens avec le domicile familial commencent à se distendre, la mobilité est plus forte et le rattachement à plusieurs résidences plus fréquent.

A presque tous les âges, les taux d'omission sont un peu plus importants **pour les hommes que pour les femmes**.

La fréquence des omissions est deux fois plus élevée pour les **personnes vivant seules** que pour l'ensemble. Elles occupent plus souvent de petits logements, voire des logements « marginaux » (chambre d'hôtel, pièce unique, chambre d'étudiants), qui sont donc plus difficiles à repérer pour les agents recenseurs. Elles sont sans doute assez souvent absentes de leur domicile, et l'agent recenseur a pu croire celui-ci vacant. Elles sont aussi nettement plus souvent comptées en double (1,4 %) qu'en moyenne.

Les **étrangers** ont été environ deux fois plus souvent oubliés au recensement que les français. Ce résultat est toutefois partiel, puisque les étrangers qui vivent dans des communautés, dans des foyers de travailleurs par exemple, et ne sont pas rattachés à un logement, sont exclus du champ de l'enquête.

LE RECENSEMENT DE 1990

Les taux d'erreurs varient aussi en fonction de **la situation professionnelle** : pour les actifs ayant un emploi, les inactives de moins de 60 ans, les taux d'omission et de double compte sont identiques aux taux moyens. Les inactifs de 60 ans et plus, et les retraités sont un peu moins souvent oubliés qu'en moyenne (c'est d'ailleurs surtout vrai des plus jeunes de la catégorie). Par contre, les chômeurs échappent plus fréquemment au recensement (4,5 %), ainsi que les étudiants de plus de 18 ans (5,1 %). Parmi ces derniers, on compte aussi beaucoup de doubles comptes (3,6 %).

Enfin, on observe que le taux d'omission des personnes s'accroît régulièrement avec la **taille de l'agglomération**. Il est de 1,1 % dans les communes rurales et progresse jusqu'à 3,2 % pour l'agglomération parisienne.

En réalité, en l'absence des contrôles d'exhaustivité effectués pendant la collecte du recensement dans les communes de plus de 20 000 habitants (par **rapprochement avec le fichier de la « taxe d'habitation »**, comme on l'a expliqué plus haut), la hiérarchie des taux d'omission par taille d'agglomération aurait été plus accentuée. L'enquête permet en effet de mettre en évidence que ce contrôle a permis de réduire de manière sensible les taux d'omission de résidences principales dans les zones urbaines. On s'attendait à ce que, comme en 1962, ces taux progressent continûment en même temps que l'accroissement de la taille de l'agglomération. Or, en 1990, ils marquent une baisse nette quand on passe le seuil des petites unités urbaines de plus de 20 000 habitants, pour remonter ensuite au fur et à mesure qu'augmente le nombre d'habitants de l'agglomération.

Conclusion

Pour conclure, on retiendra que les résultats de l'enquête doivent être lus avec quelque précaution. En sus des marges d'incertitude sur le nombre d'omissions ou de doubles comptes qui résultent d'un échec du 2^e appariement, les erreurs aléatoires peuvent être importantes quand les effectifs sur lesquels portent les résultats sont faibles. Enfin, malgré toutes les précautions prises, on ne peut pas être sûr que certaines erreurs de recensement ne se sont pas reproduites lors de la collecte de l'enquête (omissions de personnes, omissions de logements).

BIBLIOGRAPHIE

- BARNES R. "Post Censal Surveys in Great Britain", *Population Trends* 57 (Autumn 1989), HMSO, London
- Evaluation of the 1981 Census : Post Enumeration Survey (1983), *OPCS Monitor CEN 83/4*, OPCS, London
- Erreur de couverture dans le recensement : section spéciale, dans *Techniques d'enquête*, volume 14, numéro 2, décembre 1988, Statistique Canada
- FAY R.E., J.S. PASSEL, J.G. ROBINSON (1988), *The Coverage of the Population in the 1980 Census*, in 1980 Census of Population and Housing. Evaluation and Research Reports, February 1988, United States Bureau of the Census
- LÉRY A. (1990), "le 32^e recensement de la population : la France s'est comptée en 1990", *Courrier des Statistiques*, n° 56, INSEE
- VANGREVELINGHE G. (1963), Recensement démographique de 1962. Étude de l'exhaustivité du dénombrement, *Etudes statistiques*, octobre-décembre 1963, INSEE

Résumé

En France, à l'occasion du recensement démographique de 1990, l'Institut National de la Statistique et des Etudes Economiques a réalisé, auprès de 30 000 logements, une enquête consacrée à la mesure et à l'étude des erreurs de dénombrement.

La méthode de l'enquête repose sur les principes suivants :

- l'échantillon choisi était aréolaire, et les enquêteurs devaient dénombrer de manière exhaustive tous les logements situés sur les aires de l'échantillon. Par ce moyen, on pouvait récupérer dans l'enquête les logements ayant été omis au recensement.
- pour chaque logement, un questionnaire approfondi permettait d'en connaître tous les occupants, même occasionnels, de noter comment ces occupants partageaient leur temps entre ce logement et d'éventuels autres lieux d'habitation, d'obtenir l'adresse précise de ces autres lieux de résidence.
- les agents des directions régionales de l'INSEE devaient ensuite rechercher dans les documents du recensement si les personnes enquêtées avaient été recensées à l'une de leurs adresses déclarées à l'enquête. Ce travail permettait donc de déterminer, pour chaque enquêté, s'il avait été recensé 0 fois, 1 fois ou 2 fois.

L'enquête donne tout d'abord une estimation de l'erreur globale de couverture. Ainsi, d'après l'enquête, le taux d'omission des personnes au recensement de 1990 est proche de 2 %. Ces omissions sont compensées, d'une certaine façon, par des doubles comptes, qui concernent environ 1 % de la population.

L'enquête permet également de donner des informations sur les sources d'erreurs. Par exemple, les omissions de personnes proviennent dans un peu plus de la moitié des cas d'une erreur de recensement de leur résidence principale : celle-ci a été omise ou classée à tort logement vacant ou résidence secondaire par l'agent recenseur. Les personnes qui vivent seules, qui déménagent pendant la période de recensement, qui

ont plusieurs lieux d'habitation sont souvent omises ou comptées deux fois. Le risque d'omission des personnes s'accroît en même temps qu'augmente la taille d'agglomération.

Summary

Following the French population census of 1990, the National Institute for Statistics and Economic Studies carried out a survey of 30 000 housings for the purpose of measuring and studying counting errors.

The method of the survey relied on the following principles :

- the chosen sample was clustered and surveyors had to make an exhaustive count of all housings in the selected areas. The survey could this way retrieve housings which slipped through the census ;
- a thorough questionnaire made it possible to know all, even occasionnal, occupants of every housing, note how much time they spend in their possibly different dwelling places, and get the exact address of these locations ;
- clerks in the regional offices then searched through the census forms to find out whether the persons recorded by the survey were listed in the census at one of the addresses given in the survey. It could thus be established for each person queried in the survey whether he had not been counted in the census, counted once, or twice.

The survey first provided an estimation of the global coverage error. Thus, according to the survey, the rate of omission of persons in the 1990 census was close to 2 %. These omissions were in a way compensated by double counts, which involved slightly under 1 % of the population.

The survey also supplied information on the sources of errors. For instance, somewhat more than half the omissions of people derived from a census error as to their main place of residence : the latter was omitted or erroneously classified by the census enumerators as vacant or as secondary residence. Persons living alone, persons who moved during the census period or those who lived in several places were often omitted or counted twice. The larger the agglomeration, the greater was the risk of omission of persons.