

Les méthodes de correction du sous-dénombrement. Le cas du recensement des États-Unis

CORRECTING FOR AN UNDERCOUNT. THE CASE OF THE U.S. CENSUS

LOS MÉTODOS DE CORRECCIÓN DEL « SUBRECuento » : EL CASO DEL CENSO DE LOS ESTADOS UNIDOS

Barbara A. Bailar

Volume 18, Number 2, Fall 1989

Sous-dénombrement et estimation de population

URI: <https://id.erudit.org/iderudit/010016ar>

DOI: <https://doi.org/10.7202/010016ar>

[See table of contents](#)

Publisher(s)

Association des démographes du Québec

ISSN

0380-1721 (print)

1705-1495 (digital)

[Explore this journal](#)

Cite this article

Bailar, B. A. (1989). Les méthodes de correction du sous-dénombrement. Le cas du recensement des États-Unis. *Cahiers québécois de démographie*, 18(2), 257–264. <https://doi.org/10.7202/010016ar>

Article abstract

The United States has had an undercount of the population in every census it has undertaken. The undercount is particularly serious for minorities, for men, and for young persons. Since 1940, research has been conducted on estimating the size of the undercount. Though the methods were sufficient for producing a national estimate, they were lacking at providing estimates for subgroups or regions. Methods tested in the last decade appeared very promising, but a political decision was made not to adjust the census. That decision has now been repealed and the 1990 U.S. Census may be the first to reflect an adjustment for the undercount.

## **Les méthodes de correction du sous-dénombrement. Le cas du recensement des États-Unis**

Barbara A. BAILAR\*

Tout recensement vise à dénombrer la population pour qu'elle soit dûment représentée dans les divers organes législatifs. Il est évident qu'on ne parviendra jamais à dénombrer tout le monde. Mais cela ne justifie pas que le recensement des États-Unis aboutisse régulièrement à un sous-dénombrement plus élevé de certains groupes minoritaires, situation qui a entraîné des problèmes sérieux ces dernières années.

Le problème du sous-dénombrement n'est pas nouveau. Déjà, à propos du tout premier recensement des États-Unis, en 1790, George Washington déplorait qu'on n'ait dénombré que 3,9 millions d'habitants alors qu'il avait la certitude que le pays en comptait au moins 4 millions. Les documents historiques montrent que des commentaires de même nature reviennent au long des deux siècles durant lesquels l'administration a tenu des recensements réguliers aux États-Unis.

Pourtant, jusqu'aux années 1970, la question du sous-dénombrement n'a suscité qu'un intérêt théorique, même si le recensement sert traditionnellement, en vertu de la constitution américaine, à déterminer la répartition des membres de la Chambre des représentants (Carlucci, 1980).

En 1970, le Congrès a institué des programmes d'allocation de fonds fédéraux aux États, municipalités et autres entités infranationales (Fellegi, 1980). Les montants octroyés en vertu

---

\* Directrice du Conseil exécutif de l'American Statistical Association, Alexandria (Virginia, USA).

de ces programmes sont attribués annuellement sur la base des chiffres du recensement le plus récent. Ces transferts, qui ont atteint à un moment donné les 80 milliards de dollars US, sont actuellement de l'ordre de 32 milliards de dollars US.

Quand les gouvernements locaux ont constaté que le sous-dénombrement de la population entraînait une réduction des sommes qui leur étaient allouées en vertu des programmes fédéraux, ils ont fait pression sur le *U.S. Bureau of the Census* pour qu'il améliore la précision des recensements.

Le *Census Bureau* était conscient du problème depuis longtemps. Il a tenté d'y remédier en multipliant les règles qui entourent la collecte des données de population (Bailar, 1983). Il en est résulté un accroissement de la complexité du recensement. Le *Bureau* a également organisé avec soin la publicité du recensement de 1980, faisant appel à des personnalités connues dans tout le pays. En dépit de ces efforts, le sous-dénombrement a persisté. La situation s'est même aggravée, car les nouvelles méthodes ont produit une surénumération de la population de plusieurs régions, plus particulièrement des régions en pleine croissance. Il est devenu évident que l'application de méthodes de plus en plus complexes ne viendrait pas à bout du sous-dénombrement.

Les poursuites judiciaires intentées contre le *Census Bureau* par 37 gouvernements locaux ont porté le problème devant l'opinion publique. Les plaignants voulaient que l'on ajuste les données du recensement de 1980 pour corriger le sous-dénombrement. L'une de ces causes présente un intérêt particulier : celle de la ville et de l'État de New York. À l'issue du premier procès, en 1980, un appel a été logé devant une cour de juridiction supérieure, qui a ordonné la tenue d'un nouveau procès. Celui-ci a eu lieu en 1984, mais la décision n'a été rendue qu'à la fin de l'année 1987.

La ville et l'État de New York soutenaient qu'il existe des méthodes statistiques de correction des recensements (Eriksen et Kadane, 1985) et ils les ont présentées. Le *Census Bureau* a allégué que la correction ne pouvait être faite en l'absence de données d'enquête suffisamment sûres. À l'appui de cet argument, il a présenté douze estimations jugées également valides par ses statisticiens, assurant qu'il procéderait volontiers aux corrections s'il savait comment les faire. Le juge John Sprizzo a tranché en sa faveur. Mais la portée du jugement est claire : la décision ne s'applique qu'au recensement de 1980.

Puisque sa stratégie de défense reposait entièrement sur la nécessité de méthodes statistiques et de contrôle plus adéquates, en 1984, le *Census Bureau* a constitué une équipe de recherche sur le sous-dénombrement en vue du recensement de 1990. L'équipe avait le mandat de passer en revue et de comparer les méthodes de mesure de la sous-énumération, de formuler des recommandations pour 1990, et de tester et d'évaluer la méthode retenue. Elle a étudié trois méthodes.

La première, fondée sur l'analyse démographique, a été mise au point par Ansley Coale, à Princeton University, dans les années 1950. Elle repose sur l'utilisation de registres démographiques indépendants du recensement, tel celui de l'état civil. Essentiellement, si on veut connaître par exemple l'effectif des Noirs de sexe masculin du groupe d'âge 20-24 ans en 1990, on remonte aux déclarations de naissance d'enfants noirs de sexe masculin faites au cours de la période 1965 à 1970. Les enregistrements de décès permettent de déterminer le nombre d'entre eux qui sont morts avant 1990, et les dossiers de l'immigration, les adjonctions au groupe survenues depuis 1965. Enfin, puisque aucun registre des émigrants n'est tenu, on fixe, en recourant au maximum d'informations disponibles, le chiffre approximatif des pertes par émigration. Ainsi, on obtient l'estimation du nombre de Noirs de sexe masculin de 20-24 ans en ajoutant les immigrants à l'effectif de la cohorte initiale des nouveau-nés de la période 1965-1969 et en retranchant les décès et les départs qui ont marqué l'évolution du groupe.

Cette méthode a été utilisée aux États-Unis pour estimer la population aux années de recensement 1950, 1960, 1970 et 1980. Elle a produit les résultats les plus fiables qu'il ait été possible d'obtenir. Comparés aux données des recensements, ces résultats ont révélé un important sous-dénombrement, spécialement des groupes minoritaires et de certains groupes d'âge. En 1980, la sous-énumération des Noirs atteignait encore 6 % et dépassait 15 % pour les 25 à 35 ans de sexe masculin. De façon générale, les omissions sont plus fréquentes pour les hommes que pour les femmes et touchent surtout les jeunes adultes.

Après le recensement de 1970, la correction des données de recensement a monopolisé l'attention. Les estimations obtenues par la méthode de l'analyse démographique ont été promues au rang de norme. Elles comportaient pourtant des défauts considérables. Disponibles uniquement au plan

national, elles ne fournissent aucune donnée relative aux régions, aux États ou aux villes. Les migrations internes entre États et entre villes ne sont pas connues, de sorte qu'on ne peut produire d'estimations fiables au niveau infranational. Il n'est pas davantage possible d'en fournir selon le revenu, le niveau d'instruction ou quelque autre caractéristique jugée pertinente. De plus, ces estimations ne prennent pas en compte l'immigration illégale. Or l'évaluation de ces mouvements aux États-Unis conduit à un large éventail de résultats, spécialement en 1970 et en 1980. Enfin, cette méthode fondée sur les composantes démographiques manque de robustesse en ce qui a trait à la mesure de la part du sous-dénombrement dans l'erreur de couverture totale. En 1980, les méthodes de recensement utilisées ont produit un nombre important de doubles comptes, particulièrement dans les banlieues. Les données provenant de la méthode démographique ont été comparées à celles du recensement, mais comme ce dernier comprenait trois millions de doubles comptes, la mesure du sous-dénombrement est trop faible d'autant.

La deuxième méthode examinée est celle de la contre-vérification du recensement, qu'utilise Statistique Canada. Les États-Unis l'avaient testée à l'occasion du recensement de 1960, mais l'avaient trouvée impraticable. L'équipe de recherche sur le sous-dénombrement, après avoir multiplié les tests en vue du recensement de 1990, a démontré que cette méthode demeurerait inapplicable aux États-Unis. La contre-vérification est fondée sur la sélection d'un échantillon représentant chaque segment de la population constitutive de l'univers du recensement. Il s'ensuit que l'échantillon est tiré des sous-populations suivantes : les personnes énumérées au recensement précédent, celles qui auraient dû l'être et ne l'ont pas été, les enfants nés depuis lors et les immigrants arrivés après l'énumération. Les registres des naissances relèvent des États, qui en contrôlent l'accès. Or, certains États refusent de mettre leurs fichiers à la disposition du *Census Bureau* ou en restreignent l'accès pour certaines données, par exemple pour celles qui concernent les enfants adoptés. Mais, plus fondamentalement, la faiblesse de cette méthode réside dans la difficulté de retracer des individus dont les coordonnées connues peuvent être vieilles de dix ans.

Peut-être la méthode est-elle applicable au Canada parce que les recensements s'y font tous les cinq ans plutôt que tous les dix ans et que les adresses y sont, dans l'ensemble, plus à jour.

La troisième méthode évaluée par l'équipe de recherche est celle qui consiste à contrôler les données du recensement par une enquête postcensitaire. Le *Bureau* teste cette méthode depuis 1940. Chaque fois, elle donne de plus faibles estimations de la couverture que la méthode démographique. En 1980, c'est en rapprochant l'enquête sur la population active et le recensement que le *Bureau* a obtenu l'évaluation du sous-dénombrement. Pour diverses raisons, l'utilisation d'une enquête existante plutôt que d'une enquête ad hoc a été jugée inadéquate. L'équipe a donc proposé une enquête indépendante, fondée sur un échantillon stratifié, et l'a testée de façon extensive en 1985, 1986 et 1987. Les résultats ont été impressionnants. À Los Angeles-Est, l'enquête postcensitaire a montré que le pré-test du recensement avait omis 9 % de la population totale et 11 % du groupe d'origine espagnole. À Meridien (Mississippi), région rurale à forte population noire, elle a permis de calculer une sous-énumération de l'ordre de 6 %. La plupart des problèmes qu'avait entraînés l'utilisation de l'enquête sur la population active, en 1980, se trouvaient résolus. La mesure du sous-dénombrement des migrants demeure problématique, mais leur prise en compte par le recensement l'est également.

Un problème de nature statistique demeure : comment distribuer le sous-dénombrement mesuré au plan national entre les États, les régions, les comtés et les îlots de recensement ? L'enquête postcensitaire se révèle à cet égard supérieure aux autres systèmes de mesure de la sous-énumération.

Avant la sélection de l'échantillon de l'enquête postcensitaire, le pays tout entier est découpé en strates. Environ cent strates sont définies en fonction de caractéristiques liées à la sous-énumération. Par exemple, une strate pourrait être constituée d'îlots de recensement de la ville de New York comprenant plus de 50 % d'habitants de race noire, une autre d'îlots des États de l'Ohio, de l'Indiana et du Missouri dont les résidents sont principalement des banlieusards blancs.

Comme le *Census Bureau* a déjà partagé en îlots tout le territoire des États-Unis, y compris les zones rurales, cette opération n'est plus à faire. Il reste à classer chaque îlot dans la strate à laquelle il appartient par ses caractéristiques. Après quoi, par strate, on procède au tirage des îlots de manière à réduire le plus possible l'erreur échantillonnale. Puis on dresse la liste des îlots retenus, et des enquêteurs collectent, par

logement, les renseignements voulus sur le nombre et les caractéristiques des résidants. Sur la base des résultats de cette enquête, on évalue la couverture du recensement (sur- ou sous-énumération) pour chacune des cent strates. Vu l'homogénéité des strates, l'évaluation de la couverture peut être étendue, dans chacune, à tous les îlots non compris dans l'échantillon. Cette démarche a été appliquée avec succès à Los Angeles.

Le *Census Bureau* a révisé soigneusement le programme de l'enquête postcensitaire et a jugé que des gains sensibles avaient été réalisés par rapport à l'expérience de 1980 (Wolter, 1986). Il a pris la décision de collecter les données et d'estimer le sous-dénombrement, mais il a reporté à l'automne 1990 la décision de corriger ou non le recensement. De cette manière, le *Bureau* lui-même et les autres organismes, groupes ou personnes intéressés devaient avoir le loisir d'examiner les résultats du recensement et de l'enquête postcensitaire avant qu'on ne statue sur la pertinence de la correction.

Malheureusement, une décision politique, stipulant qu'on ne corrigerait pas le recensement de 1990, est venue trancher la question prématurément. Une poursuite judiciaire a été intentée en novembre 1988, et la cause devait être entendue en juillet 1989. Le jour où le procès devait débiter, un règlement hors-cour a été obtenu. Le Département du Commerce, dont dépend le *Census Bureau*, a accepté de revenir sur sa décision de ne pas ajuster les résultats du recensement de 1990. Toutes les publications du recensement de 1990 contiendront donc une note spécifiant que les chiffres publiés sont sujets à ajustement, et que l'ajustement sera effectué, le cas échéant, avant le 15 juillet 1991.

En outre, le Secrétaire au Commerce a accepté de constituer un groupe de huit personnes chargé de revoir toutes les données, de manière à pouvoir prendre sa décision en connaissance de cause. La moitié des membres de cette commission a été sélectionnée à partir d'une liste de sept noms proposée par les plaignants. Ainsi, à l'heure actuelle, trois groupes suivent très attentivement tout le processus de recensement. Le comité consultatif de l'*American Statistical Association* fera — au nom du groupe — des recommandations au directeur du *Census Bureau*, qui recevra également les recommandations d'un groupe de travail de la *National Academy of Science*. Enfin, les huit membres du groupe de travail feront, à titre individuel et à titre collectif — s'ils s'entendent sur une position commune —, des recommandations qu'ils adresseront directement au

Secrétaire au Commerce (et non au directeur du *Census Bureau*). Il se peut évidemment qu'il y ait accord de toutes les parties sur la voie à suivre. Mais il se peut aussi que la cause se retrouve prochainement devant la cour...

La conclusion principale que l'on peut dégager de ces péripéties est sans doute qu'il a fallu attendre que des millions de dollars (de transferts) soient en jeu pour que le sous-dénombrement cesse d'être une préoccupation purement académique et que le recensement soit soumis à un examen critique rigoureux. Mais il faut peut-être aussi admettre que le résultat d'un dénombrement et la meilleure estimation de la population sont deux choses différentes.

### RÉFÉRENCES BIBLIOGRAPHIQUES

- BAILAR, Barbara A., 1983. « Counting versus Estimation in a Census—A Difficult Decision ». In *1983 Proceedings of the Social Statistics Section*, Washington (DC), American Statistical Association, 42-49.
- CARLUCCI, Carl P., 1980. « The Impact of Adjustment to the 1980 Census on Congressional and Legislative Apportionment ». In *Proceedings of the 1980 Conference on Census Undercount*, Washington (DC), Bureau of the Census, 145-152.
- ERICKSEN, Eugene P. et Joseph B. KADANE, 1985. « Estimating the Population in a Census Year: 1980 and Beyond ». *Journal of the American Statistical Association*, 80, 98-113.
- FELLEGI, Ivan, 1980. « Should the Census Be Adjusted for Allocation Purposes—Equity Considerations ». In *Proceedings of the 1980 Conference on Census Undercount*, Washington (DC), Bureau of the Census, 193-203.
- WOLTER, Kirk M., 1986. « Some Coverage Error Models for Census Data ». *Journal of the American Statistical Association*, 81, 338-346.



**RÉSUMÉ — SUMMARY — RESUMEN****BAILAR Barbara A. — LES MÉTHODES DE CORRECTION DU SOUS-DÉNOMBREMENT. LE CAS DU RECENSEMENT DES ÉTATS-UNIS**

*Les États-Unis ont connu un sous-dénombrement à chacun de leurs recensements. Ce sous-dénombrement est particulièrement considérable pour les groupes minoritaires, les hommes et les jeunes. Depuis 1940, on a tenté de quantifier le phénomène. Les méthodes utilisées étaient sans doute suffisantes à l'échelle nationale, mais n'étaient pas appropriées dans le cas des sous-groupes ou des régions. Les méthodes mises au point au cours de la dernière décennie sont plus prometteuses, mais on a pris, à l'échelon politique, la décision de ne pas ajuster les données du recensement. Le renversement tout récent de cette décision pourrait faire en sorte que le recensement de 1990 soit le premier à introduire un ajustement pour le sous-dénombrement.*

**BAILAR Barbara A. — CORRECTING FOR AN UNDERCOUNT. THE CASE OF THE U.S. CENSUS**

*The United States has had an undercount of the population in every census it has undertaken. The undercount is particularly serious for minorities, for men, and for young persons. Since 1940, research has been conducted on estimating the size of the undercount. Though the methods were sufficient for producing a national estimate, they were lacking at providing estimates for subgroups or regions. Methods tested in the last decade appeared very promising, but a political decision was made not to adjust the census. That decision has now been repealed and the 1990 U.S. Census may be the first to reflect an adjustment for the undercount.*

**BAILAR Barbara A. — LOS MÉTODOS DE CORRECCIÓN DEL "SUBRECUESTO": EL CASO DEL CENSO DE LOS ESTADOS UNIDOS.**

*Los Estados Unidos han conocido un "subrecuento" a cada uno de sus censos. Este subrecuento es particularmente considerable para los grupos minoritarios, los hombres y los jóvenes. Desde 1940, se ha intentado cuantificar el fenómeno. Los métodos utilizados eran sin duda suficientes, a escala nacional, pero no eran apropiados en el caso de los subgrupos o las regiones. Los métodos puestos a punto en el curso del último decenio son más prometedores, pero se ha tomado, a nivel político, la decisión de no ajustar los datos del censo. El rechazo reciente de esta decisión podría hacer que el censo de 1990 sea el primero a introducir un ajuste para el subrecuento.*