

**Conseil économique et social**

Distr. générale  
13 mars 2012  
Français  
Original: anglais

---

**Commission économique pour l'Europe****Conférence des statisticiens européens****Groupe d'experts des recensements de la population  
et des habitations****Quatorzième réunion**

Genève, 24 et 25 mai 2012

Point 5 de l'ordre du jour provisoire

**Collecte de données par Internet****Collecte de données par Internet dans le cadre  
du recensement de la population canadienne****Note de Statistique Canada***Résumé*

À l'occasion du recensement de 2011, le Canada a procédé à plusieurs changements d'ordre méthodologique. Tout en s'inspirant des succès du recensement de 2006, lorsque ont été introduites la méthode de l'envoi de questionnaires par la poste grâce à un registre des adresses et la possibilité de répondre par Internet, de nombreux changements ont été apportés en 2011 pour obtenir un taux plus élevé de réponses par Internet, tout en réduisant les risques de non-réponse. Une méthodologie par vagues a constitué le principal élément de l'approche suivie. Cette méthodologie est fondée sur l'envoi répété de courrier, des contacts téléphoniques et des rappels en personne pour inciter les non-répondants à répondre au recensement, de préférence par Internet. Cette méthode a été testée avec succès en 2009, avant d'être utilisée en 2011.

**I. Le recensement au Canada**

1. Le présent document a été élaboré par Marc Hamel, Statistique Canada.
2. Le Canada effectue un recensement de la population et des habitations tous les cinq ans. À l'occasion du plus récent, qui a eu lieu en mai 2011, il a procédé à plusieurs changements d'ordre méthodologique. Tout en s'inspirant des succès du recensement de 2006, lorsque ont été introduites la méthode de l'envoi de questionnaires par la poste grâce

à un registre des adresses et la possibilité de répondre par Internet, de nombreux changements ont été apportés en 2011 pour obtenir un taux plus élevé de réponses par Internet, tout en réduisant les risques de non-réponse. Une méthodologie par vagues a constitué le principal élément de l'approche suivie. Cette méthodologie est fondée sur l'envoi répété de courrier, des contacts téléphoniques et des rappels en personne pour inciter les non-répondants à répondre au recensement, de préférence par Internet. Cette méthode a été testée avec succès en 2009, avant d'être utilisée en 2011.

## **II. Les modes de collecte**

3. Le Canada a recours à différentes méthodologies de collecte pour ses recensements. Elles sont brièvement décrites dans Rodrigue *et al.* (2012). La principale est l'envoi du matériel de recensement par la poste grâce aux informations figurant dans un registre d'adresses. Cette méthodologie a été utilisée pour environ 79 % des logements en 2011. Pour les autres, un questionnaire a été remis aux habitations par un agent recenseur dans le cadre d'une opération de listage-livraison (19 %), ou la collecte des renseignements a été faite directement en porte à porte par un agent recenseur (2 %). La collecte en porte à porte est surtout utilisée pour les communautés des Premières Nations et les régions reculées et septentrionales. La poste ne pouvant distribuer du courrier sans adresse à toutes les habitations, on ne peut pas recourir plus à la méthode de l'envoi du matériel. Le registre des adresses de Statistique Canada n'indique pas les noms des personnes qui habitent à telle ou telle adresse. L'adresse étant le point de départ pour le dénombrement de la population au Canada, il est indispensable d'envoyer des lettres ou des questionnaires spécifiques à des adresses spécifiques pour déterminer qui a répondu et garantir ainsi un dénombrement complet. La méthode du listage-livraison est employée principalement dans les zones rurales où le courrier n'est pas distribué directement à l'habitant mais à une boîte postale, ce qui exige de connaître exactement le nom du destinataire à la bonne adresse.

## **III. L'approche par vagues**

4. L'approche par vagues suivie dans le cadre du recensement de 2011 est en grande partie fondée sur la théorie de Dilman (2007). Cette approche n'a pas été appliquée de la même façon selon le mode de collecte retenu. Il y a eu deux approches différentes dans les régions d'envoi par la poste, et une seule pour les régions de listage-livraison. Le choix de l'approche dans les régions concernées par l'envoi par la poste a été fonction de la propension de la population à se dénombrier et du taux d'utilisation d'Internet lors du recensement de 2006.

## **IV. Approches dans les régions d'envoi par la poste**

5. Pour environ 75 % des habitations dans les régions d'envoi par la poste, une seule lettre a été envoyée lors de la première vague. Ce groupe a été identifié comme le groupe 1. Les lettres ont toutes été distribuées par la poste le 3 mai. La lettre demandait aux ménages de remplir le questionnaire en ligne au moyen du code d'accès sécurisé personnel qui leur était fourni ou d'appeler un numéro de téléphone sans frais pour demander qu'un questionnaire papier leur soit envoyé par la poste (appels adressés au système automatisé de demande de questionnaires). Le numéro de téléphone sans frais était aussi fourni au cas où les ménages aient des questions ou des problèmes. Cette approche qui n'offrait pas la simple possibilité de répondre à un questionnaire papier a été utilisée dans les régions où il était le plus probable que les ménages soient enclins à répondre au recensement. Pour le groupe 2 correspondant à 25 % des logements dans les régions d'envoi par la poste, un

questionnaire papier a été envoyé, accompagné aussi d'un code d'accès sécurisé personnel pour répondre en ligne et du numéro de la ligne téléphonique du recensement.

6. La deuxième vague a débuté le 10 mai avec l'envoi d'une lettre de rappel à tous les ménages non répondants de la première vague. La lettre, similaire à celle de la première vague, indiquait aux ménages qu'il n'était pas trop tard pour répondre. Elle incluait une nouvelle fois le code d'accès sécurisé personnel de sorte qu'il était inutile d'avoir la lettre initiale ou le questionnaire papier pour répondre en ligne. Les deux groupes ont reçu la même lettre pour la deuxième vague. Les lettres ont été imprimées en utilisant un format tout-en-un, où la lettre se replie dans ce qui est aussi l'enveloppe (un seul morceau de papier), ce qui a permis de produire plus de 6 millions de lettres de rappel en l'espace de six jours. Cette façon de procéder permettait aussi de s'assurer de la correspondance entre chaque adresse et code d'accès sécurisé, aspect essentiel pour bien comptabiliser chaque retour à la bonne adresse.

7. La troisième vague a été lancée le 18 mai, huit jours après le jour du recensement. Pour le groupe 1, un questionnaire papier a été envoyé à tous les ménages qui n'avaient pas répondu et n'avaient pas commandé de questionnaire via le système automatisé de demande de questionnaires. Un nouveau code d'accès leur était aussi fourni pour Internet. La lettre accompagnant l'envoi rappelait plus explicitement aux ménages qu'ils étaient tenus par la loi de répondre au recensement. Là encore, on a mis en place un processus de correspondance pour que l'adresse et le code d'accès sécurisé soient associés à un logement précis.

8. Pour le groupe 2 qui avait déjà reçu un questionnaire papier lors de la vague 1, un message téléphonique automatisé a été adressé aux ménages non répondants. Ce message, envoyé à la mi-journée pour être laissé sur le répondeur téléphonique, rappelait de compléter le questionnaire soit en ligne soit par la poste. On disposait de numéros de téléphone pour environ 60 % des logements, numéros principalement fournis par des sources publiques.

9. La quatrième vague, qui a débuté le 1<sup>er</sup> juin, consistait pour les agents recenseurs à contacter les non-répondants par téléphone ou en personne. Si un contact personnel ne pouvait être établi lors du suivi initial, un message était laissé sur le répondeur téléphonique ou un «avis de passage» laissé à la porte pour indiquer l'objet de la visite et demander aux ménages d'appeler la ligne téléphonique du recensement ou de contacter le bureau local de recensement le plus proche pour se faire aider à compléter le questionnaire. Les activités de suivi des cas de non-réponse se sont poursuivies jusqu'à ce que les objectifs de taux de réponse aient été atteints dans toutes les régions.

## **V. Traitement dans les régions de listage-livraison**

10. Dans les régions de listage-livraison, des questionnaires papier ont été distribués à tous les ménages par les agents recenseurs de Statistique Canada durant la première vague, qui s'est déroulée du 2 au 10 mai. Le questionnaire était accompagné d'un code d'accès sécurisé personnel pour répondre en ligne.

11. À la deuxième vague, tous les logements ont reçu par la poste une carte de rappel car il n'était pas possible d'envoyer uniquement ce rappel aux ménages non répondants.

12. Aucun traitement n'étant prévu dans le cadre de la troisième vague pour ces régions, le suivi des cas de non-réponse a débuté le 20 mai, dix jours après le jour du recensement. Le passage rapide aux activités de suivi était conforme à l'approche suivie lors du recensement de 2006 et permettait de faire en sorte que les agents recenseurs utilisés pour la distribution des questionnaires soient pour la plupart encore disponibles pour le suivi.

## VI. Système automatisé de demande de questionnaires

13. Le système automatisé de demande de questionnaires est un système automatisé accessible à partir d'un numéro de téléphone qui a été mis en place pour permettre aux ménages qui n'avaient reçu qu'une seule lettre dans le cadre des deux premières vagues de commander un questionnaire papier s'ils le préféraient. Après avoir appelé le numéro figurant sur toutes les lettres, les répondants étaient invités à saisir leur code d'accès sécurisé en utilisant les touches de leur téléphone. Le système adressait alors un message à la poste canadienne lui demandant d'envoyer un questionnaire papier au répondant. Les ménages qui ont fait la demande d'un questionnaire papier par ce biais ont répondu en moyenne en douze jours. Les répondants qui ne disposaient pas d'un téléphone avec des touches étaient invités à appeler la ligne téléphonique «spéciale recensement» pour obtenir de l'aide.

## VII. L'approche par vagues mise à l'essai

14. Les approches utilisées en 2011 étaient fondées sur les résultats des tests réalisés en 2009. Le test de recensement sur un échantillon de 100 000 logements, qui utilisait l'approche de base (quatre vagues), s'est révélé probant (Taylor 2010). En effet, le taux de réponse Internet pour le test effectué sur une base volontaire s'est élevé à 77,8 %, soit un taux nettement supérieur aux résultats du test de 2004. Les taux d'autoréponse (61,5 %) ont été aussi plus élevés qu'escompté (on tablait sur 47,5 %).

15. Quatre variantes de l'approche ont aussi été testées sur des échantillons de 5 000 logements chacun, et les résultats ont été comparés à ceux d'un groupe témoin (Taylor 2010). Deux combinaisons différentes de vagues ont été testées, chacune avec deux lettres différentes pendant la première vague. La première combinaison (deux échantillons distincts) incluait une lettre (première vague), une lettre de rappel (deuxième vague) et un questionnaire papier (troisième vague). La seconde combinaison remplaçait les deuxième et troisième vagues par l'envoi d'un questionnaire dans le cadre d'une deuxième vague qui a eu lieu plus tard que la deuxième vague de la première combinaison. Aucun suivi sur le terrain n'a été réalisé auprès de l'un quelconque des échantillons. Le test a montré que l'approche comprenant trois envois par la poste (première combinaison) promettait d'obtenir un taux de réponse Internet élevé tout en réduisant au minimum les risques de non-réponse. Elle a donc été retenue pour 2011.

16. Lors du test, on a aussi comparé les résultats obtenus lorsqu'on rédigeait de deux façons différentes la lettre envoyée pendant la première vague. Une lettre invitait les destinataires à répondre le jour du recensement (le 10 mai) tandis que l'autre les invitait à répondre sous dix jours. Les deux lettres ont été utilisées dans le cadre de chaque combinaison, une lettre par échantillon. Le test n'a pas révélé de différence importante selon le type de lettre. C'est pourquoi, la lettre proposant de répondre sous dix jours a été choisie, l'idée étant que les répondants auraient plus de jours pour répondre sur Internet et que l'on éviterait ainsi une surcharge de connexion le même jour, ce qui réduirait les risques de problème technique. Le Canada utilisant une approche *de jure* pour son recensement, on a estimé que le fait d'inviter les destinataires à répondre sous dix jours ne poserait pas de problèmes de précision sur le dénombrement de la population. Lors des précédents recensements, les questionnaires étaient toujours distribués avant le jour du recensement et une proportion importante de ménages les remplissait avant ce jour.

## VIII. Gestion des taux de réponse

17. Une stratégie a été mise en place pour gérer le processus de collecte afin d'atteindre les taux de réponse cibles et d'obtenir des taux les plus uniformes possible par région. Il s'agissait aussi d'utiliser au mieux les ressources humaines et financières tout en réalisant les objectifs en matière de réponse. La stratégie impliquait de disposer quasiment en temps réel de renseignements sur l'état d'avancement de la collecte. Elle faisait aussi appel à un modèle dynamique pour indiquer la fin prévue de la collecte par région en se fondant sur un certain nombre de paramètres – taux d'autoréponse, productivité des agents recenseurs dans le cadre du suivi des cas de non-réponse, et le nombre d'heures quotidiennes travaillées par agent recenseur et par zone. Ces renseignements étaient analysés plusieurs fois par semaine durant la collecte afin de déterminer si les différentes étapes de la stratégie devaient être mises en place.

18. La stratégie, constituée de quatre étapes, a été appliquée individuellement à plus de 45 000 unités de collecte dans le pays. Ces unités correspondent aux unités géographiques découpées pour la planification des opérations de collecte. La première étape prévoyait d'arrêter le suivi des cas de non-réponse dans toutes les unités de collecte du pays si le taux de réponse enregistré atteignait 98 %. On pensait qu'avec un tel taux, on disposerait d'informations de qualité aux fins d'une diffusion. Cela ne signifiait pas que le taux de réponse plafonnerait à 98 % dans la mesure où l'on continuerait vraisemblablement de recevoir des questionnaires remplis sur Internet ou par courrier pour ces zones.

19. Dans le cadre de la deuxième étape, il était prévu d'arrêter la collecte avant d'atteindre un taux de réponse de 98 % dans les unités de collecte très restreintes (juste quelques logements) car cela n'aurait guère d'incidence sur les résultats. La troisième étape consistait à stopper la collecte lorsqu'on enregistrerait les taux de réponse obtenus en 2006, et la quatrième à adopter des mesures qui permettraient d'obtenir le taux de réponse le plus élevé (mais avant d'atteindre les 98 %) ou les résultats de 2006. Au fur et à mesure que les taux de réponse cibles étaient atteints dans les unités de collecte ou que la collecte était arrêtée, les ressources étaient réallouées aux unités de collecte voisines où le travail n'était pas fini. Les étapes 2, 3 et 4 ne devaient être activées que si l'on estimait que la collecte dans une zone ne serait pas achevée avant la date finale prévue, ou que le budget prévu pour la collecte serait entièrement dépensé. Pour 2011, les troisième et quatrième étapes n'ont pas été mises en place.

## IX. Résultats

20. Les différentes stratégies et méthodes de collecte mises en place ont permis d'atteindre les objectifs de taux de réponse fixés pour le recensement de 2011. Le tableau ci-après indique les taux de réponse de la collecte, les taux de réponse par Internet et les taux d'autoréponse au niveau national et selon les provinces.

Tableau 1  
(En pourcentage)

	<i>Réponse</i>	<i>Internet</i>	<i>Autoréponse</i>
Canada	98,1	54,4	84,1
Terre-Neuve-et-Labrador	98,2	37,6	84,9
Île-du-Prince-Édouard	98,3	36,4	85,8
Nouvelle-Écosse	98,0	42,2	85,4
Nouveau-Brunswick	97,7	48,2	85,3

	<i>Réponse</i>	<i>Internet</i>	<i>Autoréponse</i>
Québec	98,2	52,3	84,9
Ontario	98,3	57,0	84,7
Manitoba	97,8	47,4	83,1
Saskatchewan	97,8	41,8	82,7
Alberta	98,0	57,7	82,1
Colombie-Britannique	97,6	60,5	83,3
Yukon	94,1	35,5	58,4
Territoires du Nord-Ouest	96,6	9,1	26,2
Nunavut	92,7	0,0	0,4

*Notes:* Ces taux provisoires proviennent directement des résultats de la collecte, c'est-à-dire avant le traitement et la vérification de la qualité des données. Ils sont calculés de la façon suivante: le nombre de logements privés qui ont rempli le questionnaire divisé par le nombre de logements privés classés comme occupés par le personnel sur le terrain. Les taux de réponse subséquents à la collecte seront produits une fois le traitement et la vérification des données effectués. Entre autres améliorations, ces taux de réponses finaux comprendront les logements collectifs et un ajustement du nombre de logements privés occupés à partir de l'Enquête sur la classification des logements.

21. L'approche par vagues a permis non seulement d'obtenir un taux de réponse élevé par Internet mais aussi une hausse du taux d'autoréponse. L'envoi d'une lettre à une grande partie des ménages uniquement lors des vagues 1 et 2 a poussé les répondants à utiliser massivement Internet comme mode de réponse. Les rappels fréquents semblent aussi avoir incité les personnes à répondre alors qu'elles auraient attendu la visite d'un agent recenseur pour le faire.

22. Après la collecte, l'évaluation qualitative des messages, y compris des lettres et de la publicité payée, a montré que les Canadiens interrogés étaient en général au courant du recensement et avaient compris qu'ils devaient répondre. Il est difficile de savoir si la controverse autour du remplacement du formulaire long par l'Enquête nationale auprès des ménages (ENM) dans le cadre du recensement de 2011 et la couverture médiatique de cette affaire ont contribué à mieux informer la population.

23. Lorsqu'on examine les modes de réponse avec l'approche par vagues, on constate une hausse du taux de réponse lorsque de nouveaux rappels ont été envoyés aux non-répondants (Rodrigue *et al.*, 2010, graphique non inclus). Les vagues ayant pu se chevaucher dans certaines régions (avec moins de jours pour réagir au rappel), il sera difficile d'évaluer l'impact réel de chaque vague. On peut néanmoins conclure que l'envoi d'une lettre à une partie des ménages a eu l'effet désiré (tableau 2). Le groupe 1 a surtout répondu par Internet (71,6 %). Pour le groupe 2, le taux de réponse par courrier (50,1 %) montre que les répondants utiliseront différents modes de réponse lorsque plusieurs possibilités sont à leur disposition.

24. Ces résultats sont aussi logiques en ce sens que le groupe 1 comprenait des régions ayant enregistré des taux élevés d'autoréponse en 2006. D'une manière générale, les résultats ont dépassé les attentes avec un taux de réponse global de 98 %, un taux de réponse cible par Internet de 40 % et un taux d'autoréponse cible de 80 %.

Tableau 2  
(En pourcentage)

Mode de collecte	Mode de réponse				Non-réponse	Total
	Courrier	Internet	Ligne d'assistance téléphonique	Suivi des cas de non-réponse		
Groupe 1 – lettre	16,1	71,6	0,7	9,1	2,3	100
Groupe 2 – questionnaire	50,1	25,8	0,8	20,0	3,4	100

## X. Les avantages de l'utilisation d'Internet comme mode de réponse

25. Le fait de répondre par Internet présente un certain nombre d'avantages. Les réponses reçues par ce biais sont en général d'une qualité nettement supérieure dans la mesure où l'application guide les répondants grâce à toutes les questions pertinentes. Même si un répondant n'a pas besoin de répondre à toutes les questions avant de passer à la suivante, le taux de non-réponse tend à être plus bas. Le tableau 3 permet de comparer les taux de non-réponse entre le recensement de 2006 et celui de 2011. La réduction du taux de non-réponse en 2011 est entièrement due à la forte proportion de réponses reçues par Internet.

Tableau 3

Numéro question 2011	Questions	Taux de non-réponse 2006	Taux de non-réponse 2011	Écart (2011-2006)
Recensement				
2	Sexe	1,4	0,8	-0,6
3	Date de naissance	0,7	0,5	-0,2
4	Situation matrimoniale	2,9	2,5	-0,4
5	Partenariat enregistré	6,7	4,4	-2,3
6	Lien avec la personne 1	1,7	1,4	-0,3
7	Langues officielles	n/a	1,6	n/a
8A	Langue parlée à la maison	n/a	1,8	n/a
8B	Autres langues	n/a	2,0	n/a
9	Langue maternelle	2,0	1,9	-0,1
	Accord divulgation dans quatre-vingt-douze ans	8,1	5,5	-2,6

26. Internet simplifie aussi le traitement des réponses renvoyées. La saisie des données se fait du côté du répondant, et les données entrées sont toujours dans un format exploitable par la machine. Internet permettant un traitement plus propre, il facilite aussi les opérations de codage et d'édition si nécessaire.

27. Internet est aussi plus écologique car il permet de réduire l'utilisation massive de questionnaires papier. L'option Internet peut aussi contribuer à réduire les coûts d'impression, d'envoi et de retour de courrier et, dans certains cas, de traitement, à condition que l'initiative revête une certaine envergure. Le Canada étudie actuellement à partir de quel moment l'option Internet devient intéressante.

## **XI. Enseignements tirés**

28. Les approches utilisées pour accroître le taux de réponse par Internet dans le cadre du recensement canadien se sont révélées fructueuses. Un certain nombre d'enseignements importants doivent encore être tirés pour apporter des améliorations aux prochains recensements. Il a fallu beaucoup plus de temps que prévu pour élaborer les rappels et les envoyer à divers endroits du pays, ce qui signifie que l'on n'a pas tiré pleinement parti des effets d'une vague avant de lancer la suivante. Statistique Canada examinera les moyens d'appliquer de façon plus rigoureuse l'approche à l'avenir, en particulier les moyens de mettre beaucoup moins de temps pour la production et l'envoi des rappels dans le cadre de chaque vague.

29. L'utilisation de cette approche exige aussi que l'on dispose d'un fichier d'adresses de très grande qualité. On disposait d'un petit nombre d'adresses en double pour certains logements qui ne correspondaient pas en fait à des immeubles collectifs. Dans ces cas, les répondants complétaient en général l'un des questionnaires et ignoraient l'autre, ce qui signifie qu'ils continuaient à recevoir des rappels pour le questionnaire non rempli. Ces problèmes n'ont pas été réglés avant la quatrième vague, lorsqu'un agent recenseur a pu vérifier qu'il n'existait pas de deuxième logement. Des mesures seront prises pour améliorer la qualité du fichier d'adresses, et mieux résoudre ce type de situations durant la collecte.

30. Après le recensement, l'évaluation du matériel de communication a montré que de nombreux Canadiens ne se servaient pas beaucoup des services traditionnels de la poste, ce qui fait que l'approche par vagues telle que mise en œuvre en 2011 ne leur était pas d'une grande utilité. Ces personnes règlent la plupart de leurs factures en ligne, et n'ont pas besoin de relever souvent leur boîte à lettres. Si l'on ne sait pas combien de personnes cela représente, il est probable que leur nombre augmente à l'avenir. Statistique Canada étudiera différents moyens d'adresser à l'avenir des invitations à participer de différentes façons au recensement, par exemple en invitant les répondants à se préenregistrer pour remplir le questionnaire en ligne, ou utiliser le service de poste électronique proposé par la poste canadienne pour contacter certains foyers. Avec le service de poste électronique, les Canadiens peuvent obtenir une adresse électronique auprès de la poste, et se faire envoyer leurs factures et tout autre courrier similaire via cette adresse.

## **XII. Conclusion**

31. L'utilisation de l'approche par vagues dans le cadre du recensement de la population de 2011 a donné d'excellents résultats. Il sera relativement simple d'apporter d'autres améliorations lors des prochains recensements pour augmenter encore plus le taux de réponse par Internet, ce qui permettra d'optimiser les processus connexes tels que le traitement des données et l'assurance qualité.

## **Bibliographie**

Dillman, D. A., «Mail and Internet Surveys: The tailored Design Method, second edition», Wiley 2007.

Taylor, J., «2009 Census Test: analysis of the supplementary sample», document interne, 2010.



Taylor, J., «Étude de la nouvelle méthode de collecte de recensement dans le cadre du test du recensement de 2009», Recueil du colloque 2010 de Statistique Canada, 2010.

Rodrigue, J.-F., Mathieu, P., Morin, J.-P., Taylor, J., «La méthodologie de collecte du Recensement et de l'Enquête nationale auprès des ménages de 2011 au Canada», communication présentée lors des Journées de méthodologie statistique, Paris, France, 24-26 janvier, 2012.

Côté, A.-M., Laroche, D., Wang, Q., «Identification des unités de collecte qui ont reçu la lettre destinée à promouvoir la réponse par Internet au recensement de 2011», Assemblée annuelle de 2011.

---