

LA MÉTHODOLOGIE DE COLLECTE DU RECENSEMENT ET DE L'ENQUÊTE NATIONALE AUPRÈS DES MÉNAGES DE 2011 AU CANADA

Patrice MATHIEU, Jean-Pierre MORIN, Jean-François RODRIGUE et Jennifer TAYLOR

Statistique Canada, Direction de la méthodologie

Introduction

Le Recensement de la population est fait à tous les cinq ans au Canada et la collecte du plus récent s'est déroulée de mai à août 2011, avec le 10 mai comme jour de référence. Ce recensement a subi des changements importants par rapport aux recensements précédents. Tout d'abord, une nouvelle méthodologie de collecte a été utilisée. Cette méthodologie, appelée méthodologie par vagues, est constituée d'une série de contacts avec les ménages. Ces contacts avaient pour but de stimuler les ménages à répondre au recensement, en particulier à répondre par Internet.

De plus, le questionnaire long du recensement, qui était auparavant distribué à un logement sur cinq, a été remplacé par une enquête volontaire, l'Enquête nationale auprès des ménages (ENM), dont le questionnaire a été distribué à environ un tiers de la population. Cette enquête recueille des données sur les mêmes sujets que l'ancien questionnaire long du recensement. Une méthodologie de collecte par vagues a aussi été utilisée pour l'ENM, mais adaptée de façon à ce qu'elle s'intègre aussi efficacement que possible dans les opérations du recensement, sans toutefois lui nuire. L'ENM est la plus grande enquête volontaire jamais réalisée par Statistique Canada et elle a dû être mise en œuvre dans un laps de temps relativement court. La planification et la gestion d'une enquête de cette envergure posaient donc des défis importants. De plus, il y avait un risque important de biais causé par une augmentation du taux de non-réponse pour certaines populations et, par conséquent, diverses mesures ont été prises afin de minimiser ce risque.

Ce document décrit les récents changements apportés au Recensement canadien. La méthodologie de collecte du Recensement et de l'ENM de 2011 est décrite, ainsi que quelques mesures prises pour minimiser le risque de biais de l'ENM. Finalement, quelques résultats préliminaires sont présentés.

1. Bref historique du Recensement canadien¹

Le premier recensement du Canada a été mené en 1666 par l'intendant Jean Talon. Les 3 215 habitants de la Nouvelle-France ont ainsi été dénombrés selon l'âge, le sexe, l'état matrimonial et la profession. Devant la nécessité de s'appuyer sur des renseignements fiables pour organiser la colonie et favoriser son développement, Jean Talon a effectué lui-même la majeure partie du recensement en visitant, de porte en porte, les habitants de la colonie.

Au fil des recensements suivants, vu les besoins changeants, on a ajouté des questions sur des sujets tels que le bétail, les récoltes, les bâtiments, les armes à feu, la religion et l'origine ethnique.

Le premier recensement national du Canada après la Confédération a eu lieu en 1871. Selon la Loi constitutionnelle de 1867, un recensement devait être tenu en 1871 et tous les 10 ans par la suite. Le recensement avait alors pour objectif principal de déterminer le nombre de représentants au nouveau

¹ Source: <http://www12.statcan.gc.ca/census-recensement/2006/ref/about-apropos/hist-fra.cfm>

parlement en fonction de la taille de la population de chaque région, mais il couvrait tout de même une variété de sujets et comptait 211 questions portant en outre sur le territoire, les biens immobiliers, les statistiques de l'état civil, la religion, la scolarité, l'administration, le service militaire, la justice, l'agriculture, le commerce, l'industrie et la finance.

À partir de 1906, un recensement quinquennal a eu lieu dans les provinces du Manitoba, de la Saskatchewan et de l'Alberta à la suite de l'entrée de ces provinces dans la confédération canadienne. Dès 1956, le recensement quinquennal était mené à l'échelle du pays. La croissance rapide de la population et de l'agriculture suscitait le besoin de repères tous les cinq ans afin de fournir une base plus précise pour les évaluations annuelles. En 1970, la nouvelle Loi sur la statistique stipule que le recensement doit être effectué à tous les 5 ans. De plus, chaque ménage canadien a une obligation légale de remplir le questionnaire du recensement.

En 1971, l'auto-dénombrement, méthode selon laquelle le répondant remplit lui-même son questionnaire, a été instauré pour la première fois. Les questionnaires étaient livrés par des agents recenseurs puis retournés par la poste dans les grandes régions urbaines, alors qu'ils étaient livrés et recueillis par les agents recenseurs dans les régions rurales. La population dans les régions éloignées était dénombrée comme d'habitude par entrevue. De plus, le Recensement de la population de 1971 comprenait deux types de questionnaire plutôt qu'un seul. Le questionnaire abrégé, distribué aux deux tiers des ménages canadiens, comportait les questions de base sur la population et neuf questions sur le logement. Le questionnaire long, distribué au tiers de la population, comprenait 20 questions additionnelles sur le logement en plus de 30 questions sur la situation socioéconomique de la population. À partir de 1981, le questionnaire long a été distribué à 20% de la population plutôt qu'au tiers.

Au Recensement de 2006, la méthode de livraison des questionnaires par la poste a été introduite dans les plus grandes régions urbaines. Le système postal canadien (Postes Canada) a assuré la livraison des questionnaires du recensement de la population à environ 70% des ménages alors que la majorité des autres ménages s'est fait livrer un questionnaire par un agent recenseur, comme par le passé. De plus, pour la première fois, les ménages avaient l'option de répondre par Internet plutôt que de retourner le questionnaire rempli par la poste.

Au Recensement de 2011, les méthodes de collecte de 2006 ont été conservées mais une méthodologie de collecte par vagues a aussi été mise en place. Suivant cette méthode, plutôt que de recevoir un questionnaire par la poste, une majorité des logements a initialement reçu une lettre invitant les ménages à répondre par Internet ou à composer un numéro de téléphone sans frais pour demander qu'un questionnaire papier leur soit envoyé par la poste. Par ailleurs, en juin 2010, le gouvernement canadien a décidé de rendre volontaire le questionnaire long du recensement pour des raisons de respect de la vie privée. Tous les ménages avaient donc l'obligation de remplir seulement le questionnaire abrégé qui comptait 10 questions. Afin de satisfaire les besoins en données du questionnaire long, Statistique Canada a mis en œuvre l'Enquête nationale auprès des ménages (ENM), une enquête volontaire menée auprès d'environ le tiers de la population. Cette enquête recueille des données sur les mêmes sujets que l'ancien questionnaire long du recensement.

2. Méthodologie de collecte du Recensement de 2011

La méthodologie par vagues est une nouvelle approche de collecte qui a été utilisée pour la première fois pour le Recensement de 2011. Cette méthodologie, mise à l'essai en 2009 (Taylor, 2010) lors d'un Test national du recensement, comprend une série de contacts, ou vagues, avec les ménages. On communique avec les ménages à des moments clés afin de leur rappeler de participer au recensement et de les inciter à remplir le questionnaire. La communication de chaque vague fournit les renseignements nécessaires pour répondre. Cette méthode, qui s'inscrit dans un contexte où chaque ménage canadien est tenu par la loi de répondre aux questions du recensement, est conçue de façon à encourager la réponse par Internet tout en atténuant le risque de diminution de la réponse globale.

Cette nouvelle méthodologie, décrite en détails à la section 2.2, diffère selon le mode de collecte utilisé pour distribuer le matériel du recensement pour une région donnée. Ces modes de collecte sont d'abord décrits à la section 2.1, et la section 2.3 présente quelques autres activités importantes de la collecte du recensement reliées au suivi de la non-réponse.

2.1. Modes de collecte

Les trois modes de collecte utilisés au Recensement canadien sont l'envoi par la poste, le listage-livraison et le dénombrement par entrevue. Afin d'organiser la collecte du recensement le plus efficacement possible, le territoire du Canada est divisé en petites unités géographiques appelées unités de collecte (UC). Il y a environ 46 000 UC au Canada. La taille des UC, en termes de superficie et de nombre de logements, varie beaucoup, mais en moyenne chaque UC contient environ 300 logements. Selon les caractéristiques géographiques et démographiques des UC, un mode de collecte leur est chacune assigné, et le même mode est appliqué à tous les logements d'une même UC.

Cette section décrit les trois modes de collecte, incluant la Figure 1 qui montre leur répartition géographique.

2.1.1. Envoi par la poste

Dans les unités de collecte d'envoi par la poste, on a recours au système postal pour envoyer le matériel du recensement. Ce type d'envoi permet une distribution efficace et coordonnée, sans avoir à recruter et à former un effectif important d'agents recenseurs. Cela est possible grâce au registre des adresses de Statistique Canada, une liste d'adresses qui est mise à jour à partir des résultats des recensements précédents, d'activités ciblées de listage et de diverses sources administratives. Les UC d'envoi par la poste sont typiquement dans les régions urbaines, là où la couverture et la qualité des adresses du registre sont très bonnes. Bien que ce type d'UC couvre maintenant environ 80% des logements au Canada, il représente seulement une très petite partie de la superficie du territoire.

2.1.2. Listage-livraison

Les unités de collecte de listage-livraison sont typiquement des régions rurales. Dans ces régions, le nom de l'occupant doit souvent être utilisé pour effectuer la livraison du courrier par la poste. Or, cette information n'est pas disponible sur le registre des adresses. De plus, on y retrouve aussi des endroits où la couverture du registre est de moins bonne qualité. Ainsi, dans ces UC, des agents recenseurs effectuent le listage des logements et ils font la livraison du matériel du recensement. Environ 18% des logements du Canada se trouvent dans ce type d'UC qui couvre une partie importante de la superficie du territoire.

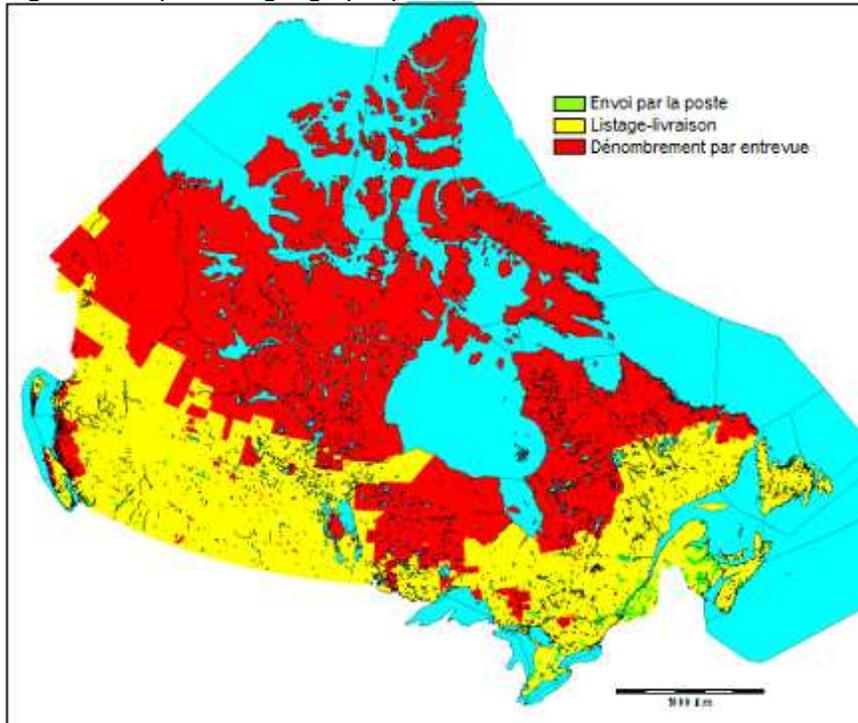
2.1.3. Dénombrement par entrevue

Les unités de collecte de dénombrement par entrevue sont, similairement aux UC de listage-livraison, des UC où le registre des adresses ne permet pas d'effectuer un envoi par la poste. Toutefois, ces UC sont habituellement situées dans des endroits éloignés ou difficile d'accès, ainsi que dans les communautés Autochtones, appelées réserves. Afin de limiter le nombre de déplacements souvent coûteux et logistiquement complexes des agents recenseurs pour faire des suivis à ces endroits, ces derniers ne se limitent pas à faire le listage et la livraison du matériel du recensement, mais remplissent aussi immédiatement le questionnaire avec les répondants. Ce type d'UC couvre un peu plus de la moitié de la superficie du territoire du Canada, mais seulement environ 2% des logements.

Dans certaines régions du nord du Canada, les agents recenseurs commencent le dénombrement par entrevue dès le mois de février plutôt qu'au mois de mai. Ce dénombrement anticipé se fait aux endroits qui sont plus difficile d'accès après la fonte des neiges, ou dans les communautés où une portion significative des habitants part à la chasse ou à la pêche pour de longues périodes au

printemps. Le dénombrement anticipé permet de dénombrer les habitants dans leurs villages avant qu'ils ne soient dispersés et difficiles à rejoindre.

Figure 1 : Répartition géographique des modes de collecte



2.2. Méthodologie par vagues du Recensement

La méthodologie par vagues a été conçue de façon à encourager la réponse par Internet. Il était cependant important de ne pas causer une augmentation significative de la charge de travail lors du suivi des cas de non-réponse sur le terrain, d'où l'importance d'offrir une alternative aux ménages qui ne souhaitent pas remplir leur questionnaire sur Internet. Même si les coûts d'élaboration de l'application Internet sont élevés, cette option comporte de nombreux avantages. Cette méthode de réponse procure des taux de réponse plus élevés à chaque question qu'un questionnaire papier et nécessite donc moins de suivi pour les questions non-répondues (Grondin et Laroche, 2008). Les fonctions intégrées permettent aux répondants de donner des réponses complètes et uniformes grâce à des vérifications et rappels automatiques de non-réponse, aux explications qui leur sont fournies ainsi qu'à l'automatisation de l'enchaînement des questions. L'enregistrement des questionnaires complétés par Internet est aussi beaucoup plus rapide que ceux retournés par la poste, ce qui permet d'avoir un aperçu plus à jour des progrès réalisés dans la collecte et d'utiliser plus efficacement les ressources consacrées au suivi. De plus, aucune opération de saisie des données n'est nécessaire, ce qui permet aux données recueillies d'être disponibles beaucoup plus tôt et de meilleure qualité, tout en réduisant les coûts de traitement.

La méthodologie par vagues s'applique de façon différente selon l'unité de collecte. Trois principaux groupes d'UC ont été définis, et une méthodologie par vague a été développée pour chacun d'eux. En raison de la nature même des UC de dénombrement par entrevue, une méthodologie par vagues n'a toutefois pas été développée pour ces unités. Les sections suivantes ainsi que la Figure 2 donnent un aperçu de la méthodologie par vagues qui a été utilisée pour le Recensement de 2011.

2.2.1. Unités de collecte d'envoi par la poste – Lettre à la vague 1

Premièrement, un ensemble d'unités de collecte d'envoi par la poste a été identifié avec l'aide d'un modèle de régression logistique qui a été construit pour permettre de cibler les ménages les plus

susceptibles de répondre par Internet tout en évitant ceux les moins enclins à répondre au recensement (Côté et al, 2011). Cet ensemble d'UC correspond approximativement à 75% des logements dans les régions d'envoi par la poste (60% de tous les logements au Canada). Dans ce groupe d'UC, la première vague a consisté en l'envoi d'une lettre dans laquelle on demandait aux ménages de remplir le questionnaire en ligne au moyen du code d'accès sécurisé (CAS) qui leur était fourni ou d'utiliser un système automatisé accessible à partir d'un numéro de téléphone sans frais pour demander qu'un questionnaire papier leur soit envoyé par la poste. Les lettres de la vague 1 ont été livrées par le système postal une semaine avant le jour du recensement, soit le 3 mai 2011.

La deuxième vague a consisté en une lettre de rappel envoyée à tous les ménages non-répondants. Cette lettre rappelle aux ménages qu'ils sont tenus par la loi de répondre au recensement. Elle comprend aussi le CAS et le numéro de téléphone sans frais, tout comme la lettre de la première vague. Elle a été livrée aux ménages du 16 au 18 mai, soit à partir du sixième jour suivant le jour du recensement. Les lettres de rappel ont été reçues sur une période de quelques jours en raison du temps nécessaire afin de produire ces lettres pour les non-répondants et du délai variable de livraison requis par le système postal pour effectuer la livraison dans les différentes régions du pays.

À la troisième vague, on a envoyé un questionnaire papier aux ménages non-répondants. Ce questionnaire a été livré aux ménages du 25 mai au 1^{er} juin, soit à partir de quinze jours après le jour du recensement. Les ménages pouvaient toujours répondre en ligne au moyen d'un CAS fourni sur la page couverture du questionnaire. De plus, une lettre accompagnait le questionnaire et elle mentionnait que si le recensement n'a pas été dûment complété au 31 mai 2011, un agent recenseur allait communiquer avec le ménage par téléphone ou en personne pour remplir le questionnaire. La lettre mentionne aussi que si le ménage refuse de répondre aux questions du recensement, le dossier pourrait être déféré au Service des poursuites pénales du Canada qui prendra alors les mesures qui s'imposent en vertu de la Loi sur la statistique.

La quatrième vague, qui a débuté le 1^{er} juin 2011, a été le suivi des cas de non réponse (SCNR) sur le terrain, et incluait aussi un rappel téléphonique automatisé. Le SCNR est décrit en détails à la section 2.3 de ce document.

2.2.2. Unités de collecte d'envoi par la poste – Questionnaire à la vague 1

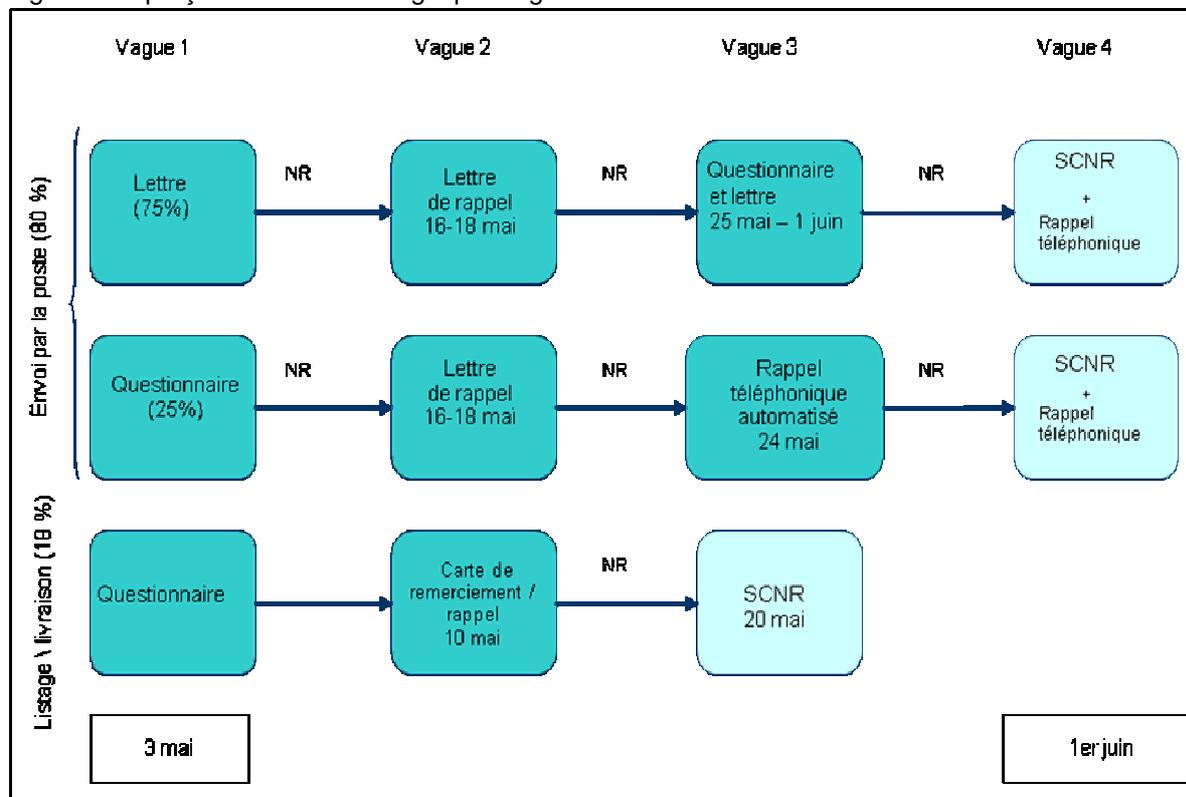
Le deuxième groupe d'unités de collecte où s'applique une variante de la méthodologie par vagues est le groupe constitué des autres UC d'envoi par la poste. Cet ensemble d'UC correspond approximativement à 25% des logements dans les régions d'envoi par la poste (20% de tous les logements au Canada). On considère les ménages dans ces UC comme étant moins susceptibles de bien répondre à la réception de seulement une lettre. Ainsi, la première vague dans ces UC est l'envoi par la poste d'un questionnaire papier. Ce questionnaire fournit un CAS, donnant ainsi l'option au ménage de répondre par Internet. Comme pour le premier groupe, cette vague a eu lieu une semaine avant le jour du recensement, soit le 3 mai. La deuxième vague pour ce groupe était identique à celle du premier groupe. À la troisième vague, les ménages non-répondants de ce groupe pour lesquels un numéro de téléphone était disponible sur la base du recensement ont reçu un rappel téléphonique automatisé le 24 mai. Non seulement cette méthode de communication est relativement peu coûteuse, mais le Test national du recensement de 2009 a aussi montré que ce type de message téléphonique automatisé est presque aussi efficace pour stimuler la réponse que d'envoyer un deuxième questionnaire (Rodrigue et al, 2010). La quatrième vague était identique à celle du premier groupe.

2.2.3. Unités de collecte de listage-livraison

Le troisième groupe d'unités de collecte où s'applique une autre variante de la méthodologie par vagues est le groupe constitué de toutes les UC de listage-livraison. À la première vague, des agents recenseurs ont livré un questionnaire papier à tous les logements dans ces UC autour du 3 mai. Ce questionnaire était assorti également d'un CAS, donnant ainsi l'option au ménage de répondre par Internet. À la deuxième vague, le 10 mai, tous les logements dans ces UC ont reçu par la poste une carte de remerciement/rappel, peu importe qu'ils avaient déjà répondu ou non, car il n'est

généralement pas possible dans ces régions de faire un envoi postal à des adresses précises à partir des adresses disponibles sur le registre des adresses. Finalement, la troisième vague consistait à passer directement au suivi des cas de non réponse sur le terrain à partir du 20 mai. Il y a donc une vague de moins que pour les deux autres groupes d'UC.

Figure 2 : Aperçu de la méthodologie par vagues du Recensement de 2011



Note : NR = Non-répondants

2.3. Suivi des cas de non-réponse

Tel que mentionné à la section précédente, la dernière vague de la méthodologie par vagues correspond au suivi des cas de non-réponse (SCNR), qui a débuté le 1er juin dans les unités de collecte d'envoi par la poste et le 20 mai dans les UC de listage-livraison. À cette étape, les agents recenseurs appellent et visitent les ménages qui n'ont pas répondu. Il y avait aussi un suivi anticipé des non-répondants qui a débuté le 13 mai 2011 dans certaines régions ciblées pour lesquelles les taux d'auto-réponse étaient plus bas ou la collecte était plus difficile lors du dernier recensement. L'anticipation du SCNR vise à obtenir, à la fin des opérations de suivi, des taux de réponse comparables à ceux des régions où la collecte se fait avec moins de difficultés.

Les agents recenseurs avaient à leur disposition des informations pour gérer leur travail dans le Système de gestion du terrain. Ce système accessible via Internet est un système informatique qui a été développé pour le Recensement de 2011 afin de faciliter la cueillette d'information sur les progrès de la collecte. Ceci permettait une meilleure gestion de cette tâche et accélérât la disponibilité des informations sur le progrès et les résultats, tout en facilitant la communication entre tous les échelons du personnel sur le terrain ainsi qu'avec le personnel du bureau central.

Beaucoup de ressources humaines et financières sont nécessaires pour effectuer le SCNR. L'infrastructure de collecte consistait en 37 bureaux locaux de recensement gérant la collecte dans leur région géographique respective et en environ 35 000 employés sur le terrain. Dans chaque bureau local de recensement, il y a aussi une unité de support téléphonique qui était utilisée de façon stratégique en offrant de l'aide aux agents recenseurs en appelant des ménages non-répondants

dans les régions où la collecte progressait plus lentement. Ainsi, les ménages non-répondants avec un numéro de téléphone disponible sur le registre des adresses étaient souvent contactés en premier lieu par cette unité et, si nécessaire, ensuite par les agents recenseurs.

De plus, chaque ménage non-répondant pour lequel un numéro de téléphone était disponible sur la base du recensement a reçu un rappel téléphonique automatisé au début de la période du suivi des cas de non-réponse. Ce message rappelait aux non-répondants leur obligation légale de répondre au recensement. Concurrément, il y avait aussi une campagne publicitaire multimédia dont l'urgence du message s'intensifiait avec la durée des opérations.

2.3.1. Vérification du statut d'occupation des logements

En parallèle avec le SCNR, des opérations sur le terrain ont aussi eu lieu afin d'effectuer la vérification du statut d'occupation des logements. Tout d'abord, il y a le processus de Vérification de l'occupation des logements (VOL). Cette activité, qui commençait le 13 mai 2011, avait pour but d'identifier un nombre significatif de logements inoccupés ou annulés (adresses qui ne sont pas des logements privés) avant le début du SCNR. En identifiant ces logements à une date proche du jour du recensement, la classification devrait être plus précise et même plus facile à effectuer. Il avait été observé lors du Recensement de 2006 que le taux d'erreur dans la classification des logements (occupé, inoccupé ou annulé) augmentait plus la classification était faite loin du jour du recensement. De plus, la VOL permet de réduire la charge de travail du SCNR puisque les logements inoccupés ou annulés identifiés par cette vérification n'ont pas besoin de suivi.

Les régions choisies pour la VOL l'ont été selon certaines caractéristiques reliées à un plus haut taux anticipé d'inoccupation (concentration de nouvelles constructions, de grands immeubles à appartements, de logements sans numéro de téléphone sur la base du recensement, etc.). Le taux de logements inoccupés et annulés dans le quartier, obtenu à la suite du recensement précédent, a aussi été considéré, tout comme certains indicateurs de croissance. Pour être plus efficace, seuls les logements avec les plus hautes probabilités d'être inoccupés ou annulés ont été choisis au lieu d'une unité de collecte au complet. Les agents recenseurs passaient dans les quartiers et classaient chaque logement sur la liste de tâche, si possible en parlant avec un des habitants ou un voisin. Ils encourageaient les gens à compléter leur questionnaire mais ne le faisaient pas avec eux, le but étant d'épurer rapidement les listes de ménages non-répondants pour le SCNR.

Tout de même, des erreurs de classification d'un logement comme étant occupé ou inoccupé se produisent durant le SCNR. Certains logements classifiés comme étant inoccupés sont en fait occupés et certains logements non-répondants sont des logements inoccupés. Ainsi, une autre opération, l'Enquête sur la classification des logements, se déroule après le SCNR. Elle évalue et détermine le statut d'occupation des logements pour lesquels aucun questionnaire complété n'a été reçu (logements inoccupés ou annulés et les cas non résolus). À partir des résultats de l'enquête, on ajuste les comptes du Recensement de la population durant le traitement (Bankier, 2008).

2.3.2. Stratégie de gestion des taux de réponse

Durant la période de collecte, il est essentiel de suivre à tous les jours les progrès réalisés ainsi que l'évolution des dépenses afin d'obtenir des taux de réponse le plus élevé possible et qui sont uniformes à travers le Canada. La stratégie de gestion des taux de réponses mise en place était constituée de quatre étapes, et tenait compte du fait que dans le cas du recensement, contrairement aux enquêtes volontaires, il n'y a pas de critères tels qu'un nombre maximum de contacts à faire avant d'arrêter le suivi auprès d'un ménage. Ces quatre étapes identifiaient les situations nécessitant des changements dans les attentes de qualité pour utiliser les ressources le plus efficacement possible afin d'obtenir des taux de réponse les plus uniformes possible. Ce plan décrivait les actions à prendre pour aider les gestionnaires sur le terrain dans la gestion du SCNR.

La stratégie s'appliquait globalement au niveau des bureaux locaux de recensement, mais la priorisation des ressources dans les bureaux était laissée aux gestionnaires du terrain plus à même de juger de l'utilisation efficace de leur personnel. La gestion des taux de réponse continuait à se faire au niveau de l'unité de collecte, l'unité de travail des agents recenseurs. Aussitôt qu'on jugeait

suffisant le taux de réponse d'une UC, on arrêta la collecte dans cette UC, et le chef d'équipe en charge en était informé. Une fois une UC désactivée pour la collecte, les efforts étaient redirigés dans les autres UC.

Le taux de réponse d'une UC était évalué quotidiennement par rapport à un critère spécifié par la stratégie pour cette unité. De plus, on l'évaluait en tenant compte du taux de réponse des UC avoisinantes. En effet, il y avait un deuxième critère pour l'UC quand le taux de réponse du secteur de recensement² (SR) ou de la subdivision de recensement³ (SDR) satisfaisait un critère à un niveau de géographie plus large. En faisant l'évaluation à deux niveaux, on acceptait un taux de réponse plus bas dans une UC si le taux de collecte était acceptable à un niveau géographique plus agrégé.

La première étape était planifiée pour le mois de juin. La cible était alors un taux de réponse de 98% dans chaque UC, ou un taux de 96% si on avait atteint un taux de réponse de 98% dans le SR ou SDR correspondant. Durant cette période, on suivait attentivement le progrès de chaque bureau local de recensement ainsi que les problèmes potentiels de qualité. Puisque la collecte se déroulait généralement bien, nous avons finalement gardé cette étape jusqu'à la mi-juillet.

La deuxième étape avait pour but la réduction du volume de cas de SCNR quand un endroit prenait du retard dans la collecte par rapport à l'horaire planifié ou si les ressources devenaient insuffisantes. À cette fin, nous arrêtons la collecte dans les UC pour lesquelles l'incidence sur les taux de réponse et la qualité était le plus petit. Ceci permettait au personnel du terrain de concentrer leurs efforts dans les UC avec plus de travail à effectuer ou ayant une plus grande incidence sur le taux de réponse. Nous avons donc désactivé les UC avec le moins de cas non résolus jusqu'à ce qu'une incidence globale jugée raisonnable soit atteinte. Il était prévu de répéter ce type de désactivation d'UC à chaque semaine une fois que cette étape était mise en place. Cette étape fut mise en place le 15 juillet et la désactivation des UC avec la plus petite incidence fut répétée après une semaine.

Lors de la troisième étape, il était prévu que les cibles soient abaissées selon les résultats du Recensement de 2006, mais avec Winsorization pour éviter de cibler des taux trop bas. Ceci permettait de tenir compte des difficultés locales rencontrées lors du recensement précédent. Nous n'avons toutefois pas activé cette étape puisque la collecte progressait suffisamment bien. Cependant, pour aider le personnel du terrain à prioriser leurs efforts durant la dernière semaine de collecte, nous avons accepté à deux reprises toutes les UC ayant moins de quatre cas non résolus.

Une quatrième étape était aussi prévue. Celle-ci aurait été une revue au cas par cas des UC, SR et SDR restants ayant pour but de prendre des décisions permettant d'obtenir des taux de réponse suffisants pour rencontrer les critères minimum de publication. Mais comme pour la troisième étape, étant donné que la collecte allait suffisamment bien, cette étape n'a pas été mise en place.

2.4. Résultats préliminaires

La collecte du recensement s'est terminée le 5 août 2011. Les taux de réponse préliminaires de la collecte du Recensement de 2011 sont présentés dans le Tableau 1. Au niveau national, ce taux de réponse de la collecte est de 98,1%. Les taux de réponse sont près de 98% pour la grande majorité des juridictions et sont dus en partie à l'efficacité de la méthodologie par vagues. Avec cette nouvelle méthodologie, et malgré le fait qu'environ 60% des logements n'ont reçu qu'une lettre à la vague 1, nous avons observé une hausse du taux d'auto-réponse par rapport à 2006.

² Les secteurs de recensement (SR) sont de petites régions géographiques relativement stables qui comptent habituellement entre 2 500 et 8 000 habitants. Ils sont créés au sein de régions métropolitaines de recensement et d'agglomérations de recensement dont le noyau urbain compte 50 000 habitants ou plus d'après le recensement précédent.

³ Terme générique qui désigne les municipalités (telles que définies par les lois provinciales) ou leurs équivalents, soit les réserves indiennes, les établissements indiens, les territoires non organisés et autres régions géographiques définies à cette fin.

Tableau 1 : Taux de réponse de la collecte des logements privés du Recensement de la population de 2011

Canada	98,1%
Terre-Neuve-et-Labrador	98,2%
Île-du-Prince-Édouard	98,3%
Nouvelle-Écosse	98,0%
Nouveau-Brunswick	97,7%
Québec	98,2%
Ontario	98,3%
Manitoba	97,8%
Saskatchewan	97,8%
Alberta	98,0%
Colombie-Britannique	97,6%
Yukon	94,1%
Territoires du Nord-Ouest	96,6%
Nunavut	92,7%

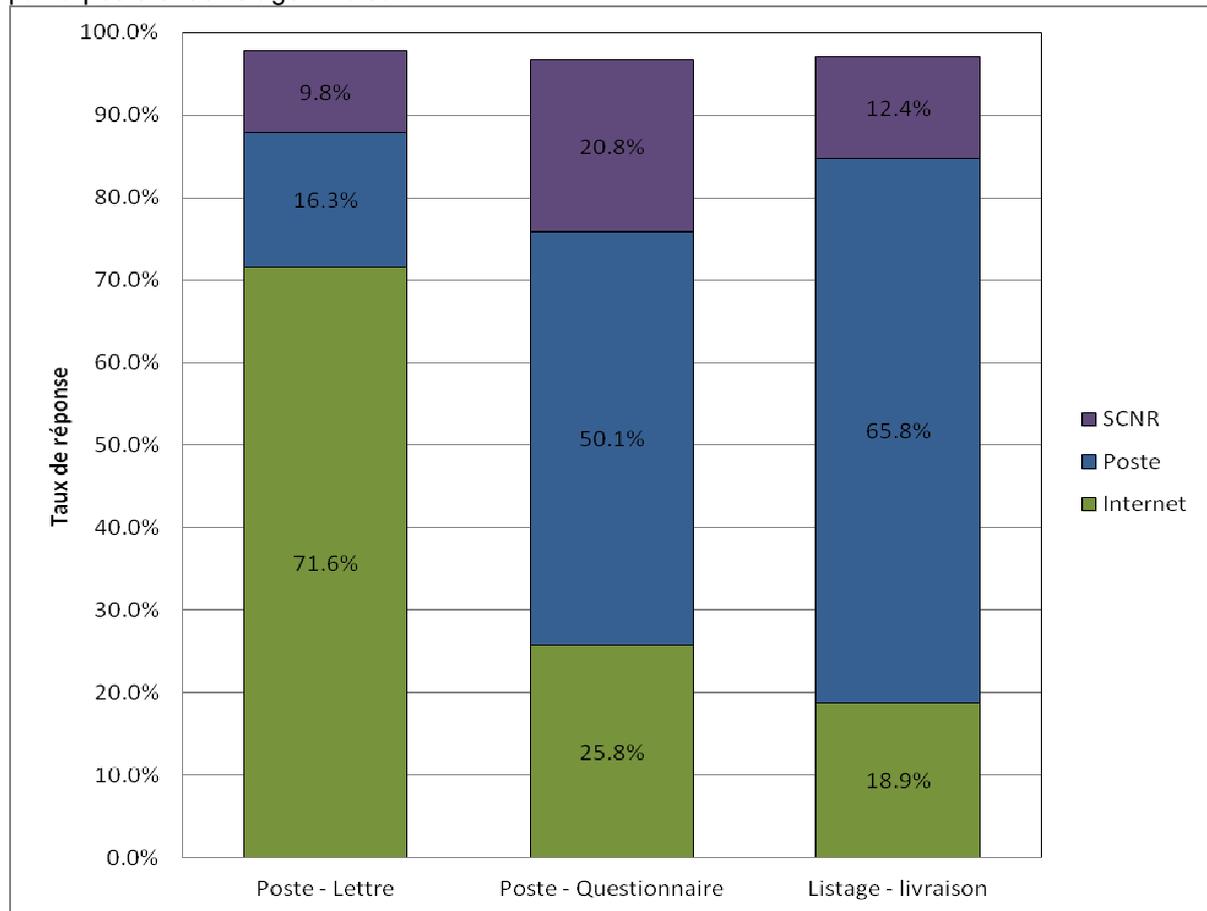
Note : Ces taux provisoires proviennent directement des résultats de la collecte, c'est-à-dire avant le traitement et la vérification de la qualité des données. Ils sont calculés de la façon suivante : le nombre de logements privés qui ont rempli le questionnaire divisé par le nombre de logements privés 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. Source : http://recensement2011.gc.ca/ccr01/ccr01_014-fra.htm

En effet, la proportion de logements qui ont répondu sans l'intervention des agents recenseurs du SCNR a augmentée comparativement au Recensement de 2006. Cette proportion est passée d'environ 79% en 2006 à 85% en 2011. Il semble donc que non seulement la méthodologie par vagues n'a pas augmenté le fardeau de travail des agents recenseurs, mais elle l'a même diminué. Ceci est d'ailleurs surtout le cas pour les logements dans le groupe d'unités de collecte d'envoi par la poste ayant reçu la lettre à la vague 1. Il était prévisible que les taux d'auto-réponse soient plus élevés pour ce groupe comparativement aux autres groupes puisque les UC de ce groupe présentaient de meilleurs taux de réponse en 2006, mais un test effectué au moment de la collecte a démontré que les meilleurs résultats sont aussi attribuables à la méthodologie par vagues tous groupes confondus. La combinaison des trois vagues avant le SCNR s'est avérée efficace pour maximiser l'auto-réponse.

La Figure 3 montre les taux de réponse par Internet, par la poste et par SCNR selon le mode de collecte. On observe que le groupe ayant reçu une lettre par la poste à la vague 1 a atteint le plus haut taux d'auto-réponse ainsi qu'un taux de réponse Internet beaucoup plus élevé que les autres méthodes. Cet aspect sera abordé plus loin dans cette section.

Il est important de mentionner que malgré la proportion importante de réponses par Internet pour le groupe d'UC d'envoi par la poste ayant reçu la lettre à la vague 1, tous les ménages dans ces UC n'avaient pas nécessairement la possibilité ou le désir de répondre par Internet. Afin de maintenir des taux d'auto-réponse élevés, il était donc important que cette lettre soit non seulement combinée avec le questionnaire papier de la vague 3, mais aussi que les ménages de ce groupe qui préféreraient répondre par questionnaire papier puissent demander qu'un questionnaire leur soit envoyé par la poste en utilisant un système automatisé accessible à partir d'un numéro de téléphone sans frais. Environ un million de telles demandes ont été reçues. Plus de la moitié des demandes ont été faites lors des trois premiers jours de la collecte, à partir de la réception de la lettre le 3 mai 2011. Ceci a permis d'envoyer un questionnaire papier plus tôt que la troisième vague, laissant ainsi plus de temps aux ménages pour répondre avant le SCNR. Environ 91% des ménages qui ont fait la demande d'un questionnaire papier ont répondu par eux-mêmes, en moyenne après douze jours. 85% l'ont fait en remplissant ce questionnaire, et 6% en utilisant le code d'accès Internet fourni sur la page couverture du questionnaire.

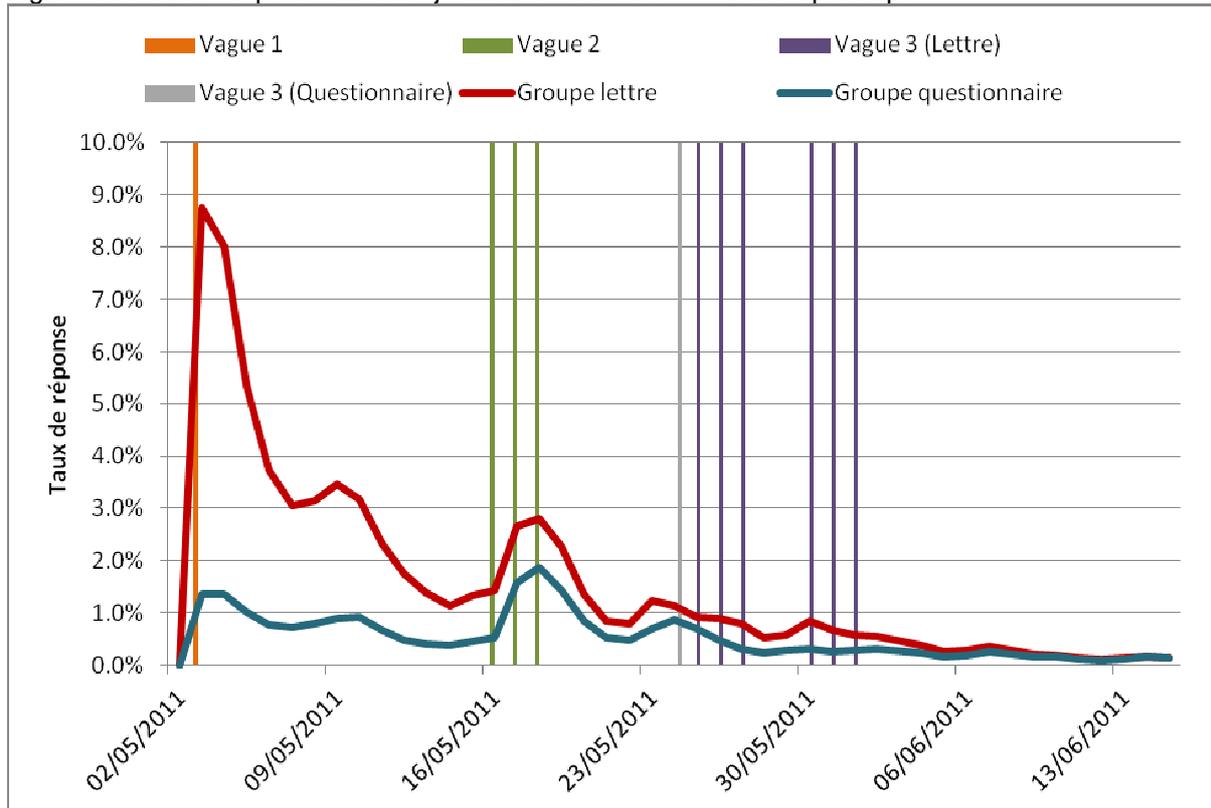
Figure 3 : Taux de réponse à la fin de la collecte selon le mode de réponse pour les régions d'envoi par la poste et de listage-livraison



Une autre mesure de l'efficacité de la méthodologie par vagues est le taux de réponse par Internet. Ce taux a été de 54% pour le Recensement de 2011, une grande augmentation par rapport au 18,3% de 2006 et beaucoup mieux que l'objectif de 45% de 2011. La Figure 4 montre l'évolution des réponses Internet dans les deux types de régions d'envoi par la poste pendant les six premières semaines de la collecte. Les lignes verticales indiquent les dates approximatives de réception des communications des vagues (lettres, questionnaires et rappels téléphoniques automatisés). Nous observons un sommet pour le groupe recevant une lettre immédiatement à sa livraison. Il y a un deuxième sommet pour ce groupe autour du jour du recensement et un troisième à la livraison de la lettre de rappel de la deuxième vague. La lettre de rappel a donc contribué à générer plus de réponses par Internet. Des sommets sont aussi observés pour ce groupe lors de la vague 3. Cependant, ces sommets sont plus étendus et petits, et ce principalement en raison du fait que l'envoi de questionnaires papier lors de cette vague stimule surtout la réponse papier. De plus, le nombre de logements encore non-répondants à cette date est relativement petit. Au total, 71,6% des logements de ce groupe ont répondu par Internet comparativement à 25,8% pour le groupe qui a reçu le questionnaire à la vague 1.

Les sommets pour le groupe recevant un questionnaire à la première vague se présentent environ aux mêmes dates que ceux du groupe lettre mais sont moins grands, surtout à la vague 1 car ces ménages recevaient le questionnaire papier dès le départ. Pour ce groupe, la lettre de rappel a toutefois généré plus de réponses Internet que la vague 1, et aussi plus de réponses par Internet que par questionnaire papier à la vague 2. De plus, le rappel téléphonique de la vague 3 semble avoir eu un petit effet sur la réponse par Internet de ce groupe.

Figure 4 : Taux de réponse Internet journaliers dans les UC d'envoi par la poste



En général, la moitié des réponses par Internet avaient été reçues en date du jour du recensement, environ 75% avant le début du SCNR dans les UC de listage-livraison et environ 90% avant le début du SCNR dans les UC d'envoi par la poste.

3. Méthodologie de collecte de l'ENM de 2011

Tel que mentionné à la section 1, le gouvernement canadien a décidé, au milieu de l'année 2010, que les questions se retrouvant précédemment sur le questionnaire long du recensement seraient recueillies sur une base volontaire pour 2011. Ainsi, tous les logements ont reçu le questionnaire abrégé du Recensement de 2011, alors que le questionnaire long a été remplacé par l'Enquête nationale auprès des ménages (ENM). Cette enquête volontaire recueille donc des données sur les mêmes sujets que l'ancien questionnaire long du recensement, incluant l'origine ethnique, l'éducation, les activités sur le marché du travail, le revenu et le logement. Ces données sont utilisées, entre autres, par tous les niveaux de gouvernement pour planifier, établir et réévaluer des politiques et programmes.

Afin de bénéficier de l'utilisation des ressources et de l'infrastructure du recensement telles que les systèmes de gestion de collecte et les employés déjà sur le terrain, il a été décidé de conduire l'ENM durant la même période que le recensement. Toutefois, cette stratégie avait le potentiel d'augmenter le fardeau des employés et des répondants. La gestion des opérations de collecte de l'ENM devait donc se faire de façon à minimiser toute incidence sur les résultats du recensement puisque ce dernier est prioritaire.

La stratégie de collecte de l'ENM devait aussi tenir compte de la limitation des ressources, des contraintes de temps et de la nature volontaire de l'enquête qui empêche d'être trop insistant et réduit le nombre de contacts possibles avec les non-répondants.

Peu de temps était disponible avant le début des activités de collecte, donc les changements aux systèmes, questionnaires et autres formulaires ont dû être réduits au minimum, dans la mesure du possible, afin de respecter les contraintes opérationnelles déjà en place.

3.1. Plan d'échantillonnage de l'ENM

Un échantillon de 4,5 millions de logements (environ 30% de tous les logements) a été sélectionné pour répondre à l'Enquête nationale auprès des ménages.

Lors des recensements précédents, le questionnaire long était distribué systématiquement à un logement sur cinq, pour une taille d'échantillon d'environ 2,5 millions en 2006. Pour l'ENM de 2011, l'échantillon de 4,5 millions de logements suivait un plan aléatoire systématique stratifié (par province et mode de collecte).

Les logements dans les unités de collecte de dénombrement par entrevue ont tous été inclus dans l'échantillon de l'ENM (environ 200 000 logements). Les logements dans les UC de listage-livraison ont été sélectionnés dans une proportion de un sur trois. Finalement, les logements dans les UC d'envoi par la poste ont été sélectionnés en utilisant des fractions de sondage provinciales qui ont été calculées de façon à i) obtenir la taille d'échantillon totale désirée fixée à 4,5 millions de logements, et, ii) obtenir des fractions de sondage égales pour chaque province (modes de collecte d'envoi par la poste et de listage-livraison combinés). La fraction de sondage moyenne pour les régions d'envoi par la poste était d'environ 29%.

3.2. Méthodologie par vagues de l'ENM

Une méthodologie par vagues a aussi été utilisée pour la collecte de l'ENM, de concert avec celle du recensement. En plus d'être basée sur le mode de collecte de l'unité de collecte, la méthodologie par vagues de l'ENM est aussi basée sur l'état de réponse et le mode de réponse (Internet ou papier) au recensement. Ainsi, parmi les logements sélectionnés pour l'ENM, les trois groupes de logements décrits ci-dessous ont été définis.

Le premier groupe de logements, appelé groupe d'enquête 1, est constitué des répondants au recensement par Internet. Ces répondants incluent non seulement ceux qui ont reçu la lettre à la vague 1 du recensement et qui ont complété leur recensement en ligne, mais aussi ceux qui ont répondu via Internet en utilisant le code d'accès sécurisé imprimé sur le questionnaire reçu par la poste (UC d'envoi par la poste) ou livré à la porte (UC de listage-livraison). Les logements du groupe d'enquête 2 sont les répondants au recensement par questionnaire papier. Finalement, les non-répondants au recensement constituent le groupe d'enquête 3.

La méthodologie par vagues de l'ENM est différente pour chaque groupe d'enquête, et les détails pour chacun d'entre eux sont présentés ci-dessous. Le Tableau 2 en donne aussi un aperçu. Tout comme pour le recensement, en raison de la nature même des UC de dénombrement par entrevue, la méthodologie par vagues n'a pas été appliquée dans ces unités.

3.2.1. Envoi par la poste – Groupe d'enquête 1

Les répondants du recensement par Internet ont été invités à remplir le questionnaire de l'ENM immédiatement après avoir complété le recensement. Après avoir soumis leur questionnaire en ligne (en utilisant le code d'accès sécurisé original), les répondants sélectionnés dans l'échantillon de l'ENM ont été automatiquement dirigés vers un écran de transition les informant qu'ils avaient été sélectionnés pour l'ENM. Cet écran contenait aussi de l'information de base à propos de l'enquête telle qu'une explication de son importance et une mention de sa nature volontaire. Les répondants désirant continuer ont alors été redirigés au questionnaire de l'ENM et l'application passait automatiquement les premières questions qui étaient les mêmes que celles du recensement et qui venaient juste d'être répondues. Ceci constituait la première vague pour le groupe d'enquête 1.

La deuxième vague du groupe d'enquête 1 a été l'envoi d'une lettre de rappel aux ménages non-répondants. La lettre, livrée aux ménages autour du 7 juin, plaidait l'importance de l'enquête.

Aucun autre matériel n'a été envoyé après la vague 2 (i.e. pas de vague 3) et l'étape suivante était le suivi des cas de non-réponse sur le terrain. Le début des suivis s'est fait progressivement par unité de collecte à partir du 8 juin, dépendamment du niveau de réponse au recensement et de la possibilité de redéployer les agents recenseurs pour prêter main forte à la collecte du recensement dans les régions où elle avait du retard.

Les procédures du terrain pour la conversion de la non-réponse de l'ENM étaient similaires à celles du recensement. Les employés, infrastructures et systèmes du recensement ont aussi été utilisés. Lorsque opérationnellement possible, les tentatives de conversion des refus étaient effectuées en personne par des chefs d'équipes et agents recenseurs ayant reçu une formation spécialement à cet effet. La définition d'un refus était la même que celle du recensement, soit un refus clairement exprimé de répondre à l'enquête. Par contre, étant donné la nature volontaire de l'ENM, ce sont plutôt les règles habituelles de refus pour les autres enquêtes volontaires qui s'appliquaient, c'est-à-dire interruption des activités de suivi après deux refus clairs (incluant les refus par téléphone).

De plus, des ressources des centres d'entrevues téléphoniques assistées par ordinateur de Statistique Canada ont été utilisées comme supplément au SCNR. L'utilisation de ces ressources était faite surtout dans les régions où la collecte de l'ENM avait du retard et où le nombre d'employés sur le terrain était insuffisant.

Il est à noter que les logements du groupe d'enquête 1 n'ont jamais reçu de questionnaire papier de l'ENM. Puisque ces logements ont choisi de répondre au recensement par Internet, on faisait l'hypothèse qu'ils utiliseraient le même mode de réponse pour l'ENM s'ils acceptaient d'y répondre. L'envoi d'un questionnaire papier aurait eu peu d'incidence sur le taux de réponse et cela aurait augmenté significativement les coûts de collecte.

3.2.2. Envoi par la poste – Groupe d'enquête 2

Le questionnaire papier du recensement ne faisait aucune mention de l'enquête additionnelle volontaire que le logement pourrait recevoir plus tard. Ainsi, le premier contact avec l'ENM des logements ayant répondu au recensement avec un questionnaire papier au cours du mois de mai était environ 4 semaines après le jour du recensement (autour du 7 juin), alors qu'ils recevaient le questionnaire papier de l'ENM par la poste. Un code d'accès sécurisé pour l'utilisation d'Internet était imprimé sur chaque questionnaire. Ceci constituait la vague 1 pour ce groupe.

Pour la deuxième vague, une lettre de rappel a été livrée aux ménages non-répondants une semaine plus tard, autour du 14 juin. Comme pour la lettre de rappel envoyée au groupe d'enquête 1, cette lettre rappelait aux non-répondants l'importance de répondre à l'enquête volontaire.

Selon les plans initiaux, la vague 3 aurait consisté en l'envoi par la poste d'un deuxième questionnaire papier aux ménages non-répondants. Ce questionnaire, accompagné d'une autre lettre de rappel, devait être livré autour du 6 juillet. Toutefois, en raison de l'incertitude générée par la possibilité d'un arrêt de travail à Postes Canada et de l'opportunité de commencer plus tôt les suivis sur le terrain, cette vague a été annulée.

Le suivi sur le terrain des cas de non-réponse (vague 4) a débuté lorsque le progrès de la collecte du recensement était suffisamment avancé, ce qui s'est produit à partir du 8 juin dans certaines régions. Les mêmes procédures que celles décrites précédemment pour le groupe d'enquête 1 ont été suivies.

3.2.3. Envoi par la poste – Groupe d'enquête 3

La méthodologie par vagues pour les non-répondants au recensement était plus simple que pour les autres groupes d'enquête. Puisque les logements du groupe d'enquête 3 n'avaient pas réagi positivement à aucune des incitations envoyées à répondre au recensement obligatoire, envoyer des lettres, questionnaires ou messages vocaux additionnels pour encourager la réponse à l'enquête

volontaire n'aurait pas eu d'effet significatif. Seul l'envoi du questionnaire de l'ENM pour être reçu autour du 7 juin était prévu, mais cette vague a également été annulée en raison de la possibilité d'un arrêt de travail à Postes Canada afin de donner priorité à la préparation et à la distribution des questionnaires du groupe d'enquête 2 qui s'effectuait en même temps.

Ainsi, le seul contact avec les ménages du groupe d'enquête 3 a été la vague 4, soit le suivi sur le terrain des cas de non-réponse. Celui-ci a commencé le 8 juin, et il se faisait lors de la même visite aux logements que pour le suivi du recensement.

3.2.4. Listage-livraison – Groupes d'enquête 1, 2 et 3

Les logements dans les unités de collecte de listage-livraison ont reçu leur questionnaire du recensement au début du mois de mai. Similairement à la méthodologie par vagues pour les groupes d'enquête des UC d'envoi par la poste, les logements qui ont répondu au recensement par Internet en utilisant le code d'accès sécurisé imprimé sur le questionnaire ont été invités à remplir le questionnaire de l'ENM en ligne immédiatement après avoir terminé le questionnaire du recensement. Aussi, un questionnaire de l'ENM a été livré par les agents recenseurs autour du 7 juin aux logements ayant choisi de répondre au recensement par papier ou n'ayant pas répondu au recensement, tout comme à ceux qui ont refusé l'offre de l'ENM après avoir rempli le questionnaire du recensement par Internet.

Aucun autre matériel n'a été livré aux portes, et le suivi des cas de non-réponse par téléphone ou en personne a commencé aux mêmes dates, selon les groupes d'enquête, que pour les régions d'envoi par la poste. Pour le groupe d'enquête 3, le suivi des cas de non-réponse de l'ENM se faisait aussi lors de la même visite aux logements que pour le recensement.

Tableau 2 : Aperçu de la méthodologie par vagues de l'ENM de 2011

Groupe d'enquête 1 – Répondants au recensement par Internet (UC d'envoi par la poste)	
Vague 1	Offre en ligne immédiatement après avoir complété le questionnaire du recensement
Vague 2	Lettre de rappel reçue autour du 7 juin
Vague 3	Aucune
Vague 4	Suivi des cas de non-réponse par téléphone ou en personne (débutant lorsque le niveau de réponse au recensement était acceptable, à partir du 8 juin dans certaines régions où la collecte du recensement était presque terminée)
Groupe d'enquête 2 – Répondants au recensement par questionnaire papier (UC d'envoi par la poste)	
Vague 1	Questionnaire reçu autour du 7 juin
Vague 2	Lettre de rappel reçue autour du 14 juin
Vague 3	Questionnaire reçu autour du 6 juillet – ANNULÉE
Vague 4	Suivi des cas de non-réponse par téléphone ou en personne (comme pour le groupe d'enquête 1)
Groupe d'enquête 3 – Non-répondants au recensement (UC d'envoi par la poste)	
Vague 1	Questionnaire reçu autour du 7 juin – ANNULÉE
Vague 2	Aucune
Vague 3	Aucune
Vague 4	Suivi des cas de non-réponse par téléphone ou en personne, en même temps que le suivi de la non-réponse du recensement, à partir du 8 juin
UC de listage-livraison	
Vague 1	Offre en ligne immédiatement après avoir complété le questionnaire du recensement (groupe d'enquête 1 seulement)
Vague 2	Questionnaire livré autour du 7 juin à tous les logements sélectionnés (sauf les répondants à l'ENM par Internet)
Vague 3	Aucune
Vague 4	Suivi des cas de non-réponse par téléphone ou en personne à partir du 8 juin (en même temps que le suivi de la non-réponse du recensement pour le groupe d'enquête 3)

3.3. Mesures pour minimiser le biais

3.3.1. Échantillonnage de deuxième phase

Tel qu'indiqué à la section 3.2, un suivi des cas de non-réponse sur le terrain a été effectué pour les logements non-répondants de l'ENM. Les méthodes utilisées pour ce suivi étaient similaires à celles du recensement. De plus, pour les non-répondants au recensement, le suivi des deux enquêtes était effectué lors de la même visite aux logements.

Cependant, tel que prévu selon le plan de l'ENM, un sous-échantillon des non-répondants à l'enquête a été sélectionné après quelques semaines de collecte afin de poursuivre le suivi de non-réponse seulement pour ces cas. Ceci a été fait dans le but de mieux contrôler la charge de travail, de réduire le coût des opérations et, ultimement, de minimiser les risques de biais de non-réponse. Cette approche, proposée initialement par Hansen et Hurwitz (1946), s'avère une méthode efficace à utiliser lorsque le coût de collecte par unité de deuxième phase est élevé ou lorsqu'il y a un besoin d'appliquer des efforts plus intensifs et ciblés sur un plus petit nombre de non-répondants pour le suivi de la non-réponse. Cette méthode a été utilisée avec succès en plusieurs occasions, notamment par l'American Community Survey des États-Unis. Le but pour l'ENM était d'appliquer un effort maximal sur tous les cas retenus pour les suivis plutôt que de diluer les efforts sur tous les cas de non-réponse. Puisque les activités de collecte sont arrêtées pour une importante portion de l'échantillon de l'ENM, le taux de réponse résultant s'en trouve évidemment diminué, mais des taux de réponse pondérés, qui tiennent compte des facteurs de pondération des unités sous-échantillonnées, peuvent être calculés pour donner un meilleur portrait des progrès de la collecte en termes de risque de biais de non-réponse.

Les détails du plan de sous-échantillonnage, tels que la taille et la répartition de l'échantillon, ne pouvaient être complètement déterminés à l'avance. Ils devaient être développés à l'intérieur des limites des contraintes opérationnelles du moment, soit principalement le budget disponible et les capacités limitées des bureaux locaux de recensement qui avaient toujours la collecte du recensement comme priorité. De plus, en raison des modifications apportées au Recensement de 2011 et de la forte couverture médiatique autour de ces changements, il était très difficile de faire des prédictions sur les progrès pendant la collecte du recensement et, surtout, celle de l'ENM. Ainsi, la méthodologie du sous-échantillonnage devait être simple et flexible pour qu'elle puisse être adaptée aux scénarios inattendus qui pouvaient se produire durant la collecte. Il était aussi important de développer un plan de sous-échantillonnage qui ne demandait pas de modifications majeures aux systèmes informatiques en place et qui ne compliquerait pas trop le calcul de la variance. De plus, la stratégie de sous-échantillonnage devait être cohérente avec la méthodologie par vagues décrite précédemment.

La sélection du sous-échantillon a été effectuée le 14 juillet. À cette date, le nombre de logements non-répondants à l'ENM s'élevait à un peu plus de 1,3 millions, pour un taux de réponse d'environ 61%. Selon le nombre d'heures travaillées par les employés du terrain lors des semaines précédentes, et en tenant compte d'une diminution prévue du nombre d'heures travaillées, de la charge de travail restante du recensement et de la date de fin prévue pour la collecte du recensement, il a été déterminé que la charge de travail des employés sur le terrain pouvait être d'environ 450 000 logements d'ici la fin des opérations. Une fraction de sondage de 0,335 a donc été appliquée pour sélectionner un sous-échantillon avec la taille visée. Le sous-échantillon était stratifié et a été sélectionné par échantillonnage systématique avec les logements ordonnés selon leur localisation géographique. L'échantillonnage systématique est relativement facile à mettre en pratique et assure que les endroits où il y a plus de non-réponse incluent plus d'unités dans l'échantillon.

L'échantillonnage de deuxième phase vise non seulement à réduire les risques de biais de non-réponse en permettant un meilleur taux de réponse pour un échantillon représentatif de l'ensemble des non-répondants, mais vise aussi à augmenter le nombre de répondants pour certaines populations cibles. Ainsi, des unités de collecte composées en grande partie de populations d'intérêt pour l'ENM et dont les taux de réponse sont typiquement plus bas ont été identifiées. L'objectif était

ainsi d'augmenter la représentativité des ces populations en augmentant la fraction de sondage (à environ 0,5) dans les UC de ces strates. Les populations cibles sont les suivantes :

- Personnes autochtones vivant hors réserve
- Personnes immigrantes au Canada entre 1996 et 2006 et les résidents non-permanents
- Personnes appartenant à des minorités visibles
- Personnes dans la population active possédant au plus un diplôme d'études secondaire
- Ménages avec un revenu inférieur à 20 000\$

Les UC cibles ont été identifiées en utilisant les réponses au questionnaire long du Recensement de 2006. Ainsi, 6 191 UC ont été sélectionnées pour être sur-échantillonnées. Bien que ces UC cibles ne contiennent que 15,8% de la population canadienne, elles contiennent plus de la moitié des immigrants au Canada (1996-2006) et des personnes appartenant aux minorités visibles, plus du tiers des personnes autochtones hors réserve, 23% de la population vivant dans un ménage avec faible revenu et 17% des personnes avec un bas niveau de scolarité.

Après avoir sur-échantillonné les UC cibles, le reste des logements dans le sous-échantillon a été échantillonné proportionnellement au nombre de non-répondants, avec une fraction de sondage réduite de façon à obtenir la taille de sous-échantillon désirée.

Le sur-échantillonnage des UC cibles devrait contribuer à augmenter le nombre de répondants dans les populations cibles, ce qui devrait réduire la variance due à l'échantillonnage et faciliter l'imputation par donneur pour ces populations. De plus, les UC ciblées ont été traitées différemment lors du processus de gestion des taux de réponse décrit à la section suivante, permettant ainsi de réduire le risque de biais de non-réponse pour les populations cibles.

3.3.2. Stratégie de gestion des taux de réponse

Contrairement aux autres enquêtes volontaires de Statistique Canada où des taux de réponse globaux sont visés et où la collecte est contrôlée en utilisant le nombre de tentatives pour compléter un cas, le contrôle des taux de réponse de l'ENM était plutôt basé sur une méthodologie semblable à la stratégie de gestion de la non-réponse du recensement.

Le but principal de la gestion de la non-réponse du recensement est d'arrêter le dénombrement dans une petite région géographique lorsque le taux de réponse visé est atteint dans cette région. Les ressources supplémentaires peuvent alors être réinvesties dans les régions plus difficiles à dénombrier.

Comme pour le recensement, la collecte de l'ENM était contrôlée quotidiennement. En raison du sous-échantillonnage effectué sur les cas de non-réponse de l'ENM, les taux de réponse pondérés, principalement au niveau de l'unité de collecte, ont été utilisés pour gérer les progrès de la non-réponse. Les progrès de l'ENM variant beaucoup d'une UC à l'autre, les ressources pouvaient être réallouées afin de minimiser autant que possible le risque de biais de non-réponse. Cependant, ceci devait toujours être fait en tenant compte de la priorité accordée au recensement.

Suivant la sélection de l'échantillon de deuxième phase, la première intervention du processus de gestion des taux de réponse a été de fournir des cibles aux employés sur le terrain. Celles-ci étaient le nombre de questionnaires complétés à obtenir par UC qui résulterait en un taux de réponse pondéré de 90%, sans toutefois demander d'obtenir une réponse pour plus de 90% des cas restants d'une UC. Lorsqu'une UC ne contenait pas suffisamment de cas disponibles pour pouvoir atteindre la cible, les cas en surplus étaient assignés à d'autres UC de la même province afin de pouvoir tout de même atteindre un taux de réponse pondéré de 90% au niveau provincial. Certaines UC avaient donc pour cible un taux de réponse pondéré supérieur à 90%. Bien que souvent trop optimistes, les cibles étaient utiles pour l'organisation du travail sur le terrain.

Après quelques semaines, les UC ont été divisées en trois niveaux d'importance selon la présence de populations à risque (UC cibles décrites à la section précédente) et selon l'hétérogénéité des caractéristiques de la population. La mesure d'hétérogénéité des UC a été obtenue à partir d'une méthode mesurant la complexité des combinaisons d'un large éventail de caractéristiques de la population obtenues à partir du questionnaire long de 2006 (Ghafari, 2011). Les UC étaient donc

classées en catégories allant de la plus hétérogène à la plus homogène, et les taux de réponse pondérés visés ont été revus à la baisse en fonction de ce niveau d'importance de l'UC. Les UC ciblées étaient considérées au même niveau que les UC les plus hétérogènes. L'utilisation des niveaux d'importance a permis d'accélérer la désactivation des UC tout en minimisant l'impact sur le biais de non-réponse.

À partir de la mi-août, faisant face à une diminution significative des ressources budgétaires et humaines, il a été décidé de commencer à mettre fin à la collecte dans toutes les UC de plusieurs des bureaux locaux du pays. Ces bureaux ont été choisis en tenant compte des progrès réalisés jusqu'à ce moment en termes de taux de réponse pondéré, et aussi de la distribution de leurs UC dans les différents niveaux d'importance. Ainsi, typiquement, ces bureaux démontraient un progrès satisfaisant, et relativement peu d'UC hétérogènes ou avec populations cibles s'y retrouvaient. Ceci permettait alors de faire une utilisation aussi judicieuse que possible des ressources encore disponibles.

Quelques jours plus tard, la collecte a été arrêtée dans toutes les UC les plus homogènes de quelques grandes villes du pays. Ceci permettait de concentrer les efforts de collecte dans les quartiers les plus hétérogènes de ces villes afin de diminuer le risque de biais de non-réponse.

Finalement, toutes les opérations de collecte ont été arrêtées le 24 août.

3.4. Résultats préliminaires

Tel que mentionné ci-dessus, la collecte de l'ENM s'est terminée le 24 août 2011, soit un peu moins de trois semaines après la fin de la collecte du recensement. Le taux de réponse préliminaire de la collecte de l'ENM s'est élevé à 69,3% pour le Canada, ce qui est loin du taux obtenu lorsque le questionnaire long était obligatoire (environ 95% en 2006), mais du même ordre de grandeur que pour les autres enquêtes ménage menées sur une base volontaire par Statistique Canada. Tel que décrit à la section 3.3, des taux de réponse pondérés ont été calculés et utilisés pour contrôler la collecte dans le but d'avoir un meilleur portrait des progrès de la collecte en termes de biais de non-réponse. Le taux de réponse pondéré de la collecte de l'ENM s'est élevé à 78,5% pour le Canada. Le Tableau 3 contient les taux de réponse de la collecte, non-pondérés et pondérés, par province et territoire.

Tableau 3 : Taux de réponse de la collecte des logements privés de l'ENM de 2011

	Taux de réponse	Taux de réponse pondéré
Canada	69,3%	78,5%
Terre-Neuve-et-Labrador	64,0%	73,7%
Île-du-Prince-Édouard	61,5%	71,1%
Nouvelle-Écosse	65,8%	76,0%
Nouveau-Brunswick	64,5%	75,3%
Québec	72,9%	82,6%
Ontario	68,1%	77,6%
Manitoba	68,8%	76,5%
Saskatchewan	64,5%	74,3%
Alberta	67,5%	76,3%
Colombie-Britannique	70,2%	78,5%
Yukon	61,1%	70,6%
Territoires du Nord-Ouest	83,9%	84,4%
Nunavut	78,1%	78,1%

Note : 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. 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 le retrait des questionnaires non

remplis identifiés comme tels durant le traitement et un ajustement du nombre des logements privés occupés à partir de l'Enquête sur la classification des logements.

Les taux sont calculés de la façon suivante :

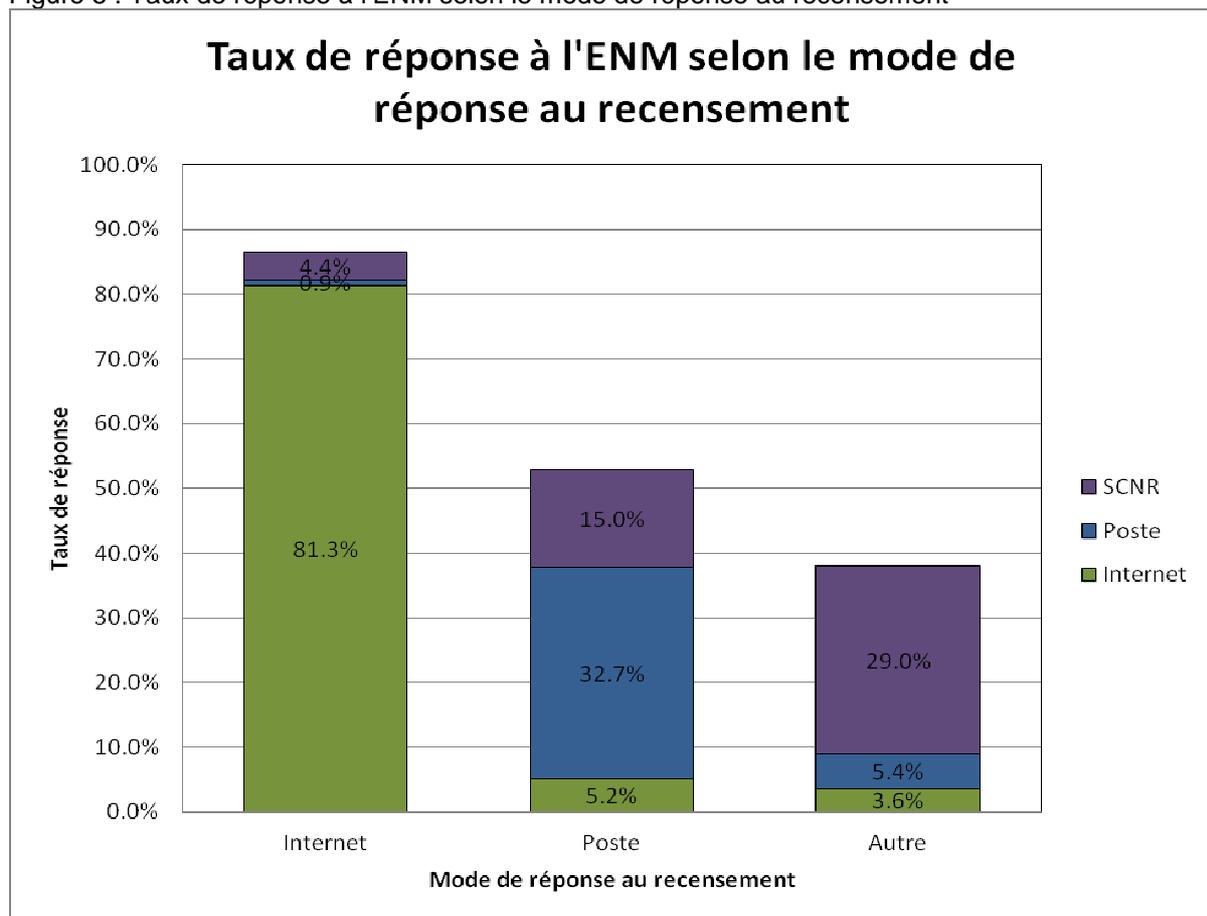
Taux de réponse : le nombre de logements privés choisis qui ont rempli le questionnaire divisé par le nombre de logements privés choisis classés comme occupés par le personnel sur le terrain. Source : http://enm.statcan.gc.ca/nhs-enm_r012-fra.htm

Taux de réponse pondéré : le nombre pondéré de logements privés choisis qui ont rempli le questionnaire divisé par le nombre pondéré de logements privés choisis classés comme occupés par le personnel sur le terrain. Les poids sont dérivés des fractions de sondage initiales et des fractions de sondage pour le sous-échantillon utilisées pour sélectionner des logements non-répondants (à la mi-juillet 2011) pour davantage de suivi. Les taux pondérés sont plus grands que les taux non-pondérés puisqu'ils reflètent le sous-échantillonnage. Aussi, ils reflètent mieux l'effet potentiel de la non-réponse sur les estimations de l'enquête.

Il est aussi intéressant de noter qu'environ 63% des répondants à l'ENM ont rempli leur questionnaire par Internet. Ceci est supérieur à la proportion obtenue pour le recensement, et elle est probablement due en grande partie à la transition facilitée de l'application Internet entre le recensement et l'ENM. Comme on peut le voir dans la Figure 5 qui montre les taux et modes de réponse à l'ENM selon le mode de réponse au recensement, c'est pour les répondants du recensement par Internet que le taux de réponse à l'ENM est le meilleur, et la majeure partie de ces réponses à l'ENM sont aussi par Internet.

Ces résultats donnent un aperçu initial de ce qui a été accompli par l'ENM, mais plusieurs autres aspects seront analysés prochainement afin de mieux évaluer la qualité de l'enquête, notamment l'analyse du biais et les taux de réponse par question.

Figure 5 : Taux de réponse à l'ENM selon le mode de réponse au recensement



4. Conclusion

Plusieurs changements ont été apportés au Recensement de 2011. Tout d'abord, une méthodologie de collecte par vagues a été utilisée. Cette nouvelle méthodologie semble avoir été un succès, tant au niveau de l'auto-réponse que des réponses par Internet.

De plus, le questionnaire long du recensement a été remplacé par l'Enquête nationale auprès des ménages. Ayant bénéficié de son intégration partielle au recensement, les taux de réponses à l'ENM sont dans l'intervalle des taux ordinairement obtenus par les enquêtes volontaires de Statistique Canada mais ils sont tout de même loin de ce qu'ils étaient lorsque ce questionnaire long était obligatoire. Diverses mesures d'atténuation du biais ont été appliquées lors de la collecte, et leur incidence sur la qualité des données sera évaluée subséquemment.

Les changements majeurs apportés au recensement en 2011 ont ouvert la porte à une remise en question plus large du processus de recensement au Canada et une étude évaluant différentes options pour les recensements futurs a été effectuée (Royce, 2011). Les recommandations de cette étude, combinées aux résultats du Recensement de 2011, suggèrent une continuité dans l'approche utilisée au moins pour le prochain recensement. Cependant, la méthodologie de collecte du Recensement et de l'ENM de 2011, qui a été mise en place dans un court délai, sera revue au cours des prochaines années afin d'améliorer l'auto-réponse et la réponse par Internet tout en réduisant les coûts et le fardeau de réponse.

Remerciements

Les auteurs tiennent à remercier Marc Hamel et Normand Laniel pour leurs précieux commentaires.

Bibliographie

- [1] Taylor, J., « Étude de la nouvelle méthode de collecte de recensement dans le cadre du Test du recensement de 2009 », *Recueil du Symposium de 2010 de Statistique Canada*, 2010.
- [2] Grondin, C. et Laroche, D., « Impact of Online Edits and Internet Features in the 2006 Canadian Census », *Working Paper No. 5 of the Work Session on Statistical Data Editing of the Conference of European Statisticians, United Nations Statistical Commission and Economic Commission for Europe, Vienna, Austria*, 2008.
- [3] Côté, A.-M., Laroche, D. et 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 la SSC, Juin 2011, Recueil de la Section des méthodes d'enquête*, 2011.
- [4] Rodrigue, J.-F., Lebrasseur, D., Morin, J.-P. et Taylor, J., « Evaluation of the innovations implemented in the 2009 Canadian Census Test », *Proceedings of the Survey Research Methods Section, American Statistical Association*, 2010.
- [5] Bankier, M., « Canadian Census E&I – Lessons learned from 2006 with plans for 2011 », *United Nations Statistical Commission and Economic Commission for Europe Conference of European Statisticians, Work Session on Statistical Data Editing, Austria*, 2008.
- [6] Hansen, M.H. et Hurwitz, W.N., « The problem of nonresponse in sample surveys » *Jour. Amer. Stat. Assoc.*, 41, 517-529, 1946.
- [7] Ghafari, M., « Homogeneity measure », *document interne, Statistique Canada*, 2011.
- [8] Royce, D., « Rapport préliminaire sur les options méthodologiques pour le Recensement de 2016 », *Statistique Canada*, 2011.