

Dossiers administratifs dans le U.S. Census pour la base de sondage des logements collectifs : utilisations possibles et limites¹

Asaph Young Chun et Jessica Gan²

Résumé

Ce document vise à présenter l'utilisation des dossiers administratifs dans le U.S. Census pour les logements collectifs. Le dénombrement des logements collectifs signifie la collecte de données dans des endroits difficiles d'accès, comme les établissements correctionnels, les établissements de soins infirmiers spécialisés et les casernes militaires. Nous examinons les avantages et les contraintes de l'utilisation de diverses sources de dossiers administratifs pour l'élaboration de la base de sondage des logements collectifs, en vue d'améliorer la couverture. Le présent document est un document d'accompagnement de la communication de Chun et Gan (2014) portant sur les utilisations possibles des dossiers administratifs dans le recensement des logements collectifs.

Mots clés : Dossiers administratifs, logements collectifs, base de sondage, recensement

1. Introduction

1.1 Contexte

Les dossiers administratifs (DA) sont des données recueillies principalement aux fins de l'application de programmes administratifs (Federal Committee on Statistical Methodology³, 1980). Les chercheurs ont utilisé les DA pour créer une population ou des bases d'échantillonnage de rechange, en précisant le processus d'estimation au moyen de variables auxiliaires tirées d'une base de données de DA, et en rentabilisant le suivi de la non-réponse⁴. Même si la collecte et l'utilisation de DA remontent aussi loin qu'au temps des pharaons d'Égypte, des progrès importants au chapitre de l'intégration des dossiers administratifs dans les recensements se sont produits au cours des dernières décennies seulement (Nordbotten, 2009; Chun et Scheuren, 2011). De toute évidence, en tirant parti du système de dossiers administratifs centralisé existant et pour assurer un contrôle de l'augmentation des coûts du recensement et maintenir la qualité des données, un recensement fondé sur des dossiers administratifs a remplacé le recensement traditionnel dans plusieurs pays d'Europe, notamment dans les pays nordiques (Wallgren et Wallgren, 2007; Sundgren, 2010).

En principe, le recensement aux États-Unis est tenu pour dénombrer des personnes dans des logements et des logements collectifs (LC) typiques. Le recensement des LC nécessite la collecte de données démographiques dans des endroits où des personnes non apparentées vivent selon des modalités de vie en groupe et dont le propriétaire ou le gestionnaire est une entité fournissant un logement ou des services aux résidents (Williams, De Vos, Russell et Barrett, 2013). Ces LC comprennent des installations comme les résidences collégiales et universitaires, les établissements correctionnels, les casernes militaires, les établissements de soins infirmiers et les établissements de services, comme des gîtes d'urgence et de transition, des soupes populaires et des cantines mobiles offrant des services réguliers. Compte tenu de la nature collective de ces logements, une liste ou des données administratives concernant les personnes vivant dans les LC ont tendance à être maintenues par une entité organisationnelle qui fournit des services de logement. Toutefois, le manque d'ouvrages publiés est évident dans le domaine de

¹ Les points de vue exprimés sont ceux des auteurs et pas nécessairement ceux du U.S. Census Bureau.

² Asaph Young Chun, U.S. Census Bureau, bureau 4K273, Washington, DC 20233. Jessica Gan, Université Rice.

³ Le Statistical Policy Working Paper 6 – Report on Statistical Uses of Administrative Records préparé par le Federal Committee on Statistical Methodology (1980), Washington DC. est le premier rapport américain exhaustif portant sur les diverses façons dont les dossiers administratifs peuvent être utilisés pour compléter les enquêtes et les recensements.

⁴ Voir le document de travail 6 du FCSM (1980) pour trouver des exemples illustrant une gamme variée d'utilisations de dossiers administratifs.

l'application des dossiers administratifs propre aux LC. Dans la conférence fondamentale de 1980 du Federal Committee on Statistics Methodology portant sur l'utilisation de dossiers administratifs à des fins statistiques, aucune communication n'a porté sur les LC. En 2010, le Symposium international de méthodologie de Statistique Canada, qui était consacré à l'interaction entre les données administratives et les recensements et les enquêtes, ne comprenait aucune communication portant sur les défis ou les possibilités liés à l'utilisation de dossiers administratifs dans les logements collectifs ou des cas similaires.

Quelques études récentes ont commencé à porter sur ce sujet, dans le contexte des LC. De façon plus particulière, des études ethnographiques menées pour le Recensement de 2010 ont fait ressortir des défis et des possibilités liés à l'utilisation des données administratives pour la mesure de la couverture du recensement dans les prisons et les pénitenciers, les foyers collectifs pour les femmes, les casernes militaires et les établissements de soins infirmiers spécialisés (Chan, 2012). Le personnel d'évaluation du Census Bureau a fait ressortir les avantages de l'utilisation de données administratives pour l'élaboration de la base de sondage des LC (Williams, Barrett et Williams, 2013) et le recensement de ces logements (Chun et Gan, 2014), dans le cadre du Recensement de 2010. Ces constatations récentes appuient celles de Jonas (2003), selon lequel dans le recensement des LC de 2000, les dossiers administratifs étaient la principale source de données utilisée par près de la moitié des questionnaires destinés aux LC (c.-à-d. 3,7 millions sur 7,5 millions de questionnaires). Jonas a en outre souligné que, dans la plupart des types de LC, les données administratives représentaient la principale façon de produire les rapports individuels de recensement, le questionnaire des LC destiné à chaque personne vivant dans un tel logement – 73 % des établissements de soins infirmiers; 66 % des hôpitaux; 60 % des foyers collectifs; 56 % des établissements correctionnels et 49 % des établissements pour jeunes. Sa recommandation fondée sur des données probantes quant à la mise à l'essai et à l'adaptation de l'utilisation des données administratives pour des LC ciblés appuie l'hypothèse selon laquelle les dossiers administratifs pourraient aider à concevoir un exercice lié aux LC rentable et de grande qualité. Chun, Chan et coll. (2012)⁵ ont exploré une erreur totale et un modèle de coût de LC, afin de déterminer le compromis optimal entre diverses sources d'erreurs de couverture des LC et les composantes de coût connexes. Toutefois, ces efforts de recherche récents ont une portée et des répercussions limitées. Par exemple, une erreur d'enquête totale et un paradigme de coût que Chun, Chan et leurs associés ont appliqué aux LC sont provisoires sans données sur les erreurs et les coûts liés à ces logements. Le National Research Council (2011) a incité le Census Bureau à effectuer des études explorant l'utilisation des DA pour compléter et améliorer la collecte de données traditionnelles sur les LC. Notre document vise à répondre à cette préoccupation scientifique en mettant particulièrement l'accent sur l'établissement de la base de sondage des LC⁶.

Dans le présent document, nous démontrons l'utilité des DA, particulièrement pour l'élaboration d'une base de sondage des LC, en vue d'améliorer la couverture. Nous analysons les résultats du Recensement de 2010, afin de déterminer dans quelle mesure les DA ont servi à la base de sondage des LC et d'examiner les avantages et les limites de diverses sources de DA utilisées pour la validation des unités de LC.

1.2 Aperçu de l'exercice de validation des logements collectifs⁷

La validation des logements collectifs (VLC) de 2010 a servi d'exercice d'établissement de listes de LC, grâce auxquelles tous les LC possibles ont fait l'objet d'un suivi. Le personnel du Census Bureau a déterminé que les autres logements collectifs (ALC) pouvaient être des LC possibles, grâce aux travaux de listage sur le terrain, à des programmes de partenariat, comme le programme Local Update of Census Addresses (LUCA) et la liste existante des LC du Recensement de 2000, notamment. Les responsables des listes de la VLC ont été envoyés dans chacun des ALC, afin de vérifier si une adresse comportait les données géographiques du recensement appropriées, et de déterminer s'il s'agissait ou non d'un LC valide pour le Recensement de 2010. Si le responsable validait l'adresse comme un LC, il déterminait le type de LC à partir des renseignements recueillis dans le cadre de l'interview de VLC et recueillait des renseignements additionnels concernant le LC. Les responsables de l'établissement des listes ont effectué ces travaux au moyen du questionnaire de VLC conçu pour faciliter les interviews sur place.

⁵ Voir Chun, Chan, Barrett, Williams et Belton (2012). « What Counts as Group Quarters? A Glimpse of Total Coverage Errors and Costs Model in Census ». Document présenté aux Joint Statistical Meetings de 2012 dans le cadre de la séance de communications libres organisée par Chun, Chan et Scheuren. San Diego, Californie.

⁶ Le contexte et les détails des constatations clés abordées dans le présent document figurent dans le rapport d'évaluation de validation des logements collectifs du Recensement de 2010 préparé par Williams, Barrett et Williams (2013).

⁷ Des détails de l'exercice de VLC figurent dans Williams, Barrett et Williams (2013).

2. Approche analytique

La principale question de recherche sur laquelle nous nous penchons vise à évaluer si les DA sont utiles pour l'élaboration de la base de sondage des LC et, le cas échéant, à déterminer dans quelle mesure les DA peuvent jouer un rôle pour compléter la base de sondage des LC et à examiner la contribution des diverses sources de DA à la base de sondage des LC.

2.1 Analyse des sources de la base de sondage de VLC de 2010

La base de VLC du Recensement de 2010 comprenait les adresses prédéterminées qui étaient incluses dans la charge de travail initiale, ainsi qu'un certain nombre d'adresses ajoutées pendant la VLC. L'exercice de VLC du Recensement de 2010 a permis de recevoir toutes les adresses déterminées comme étant des ALC pendant l'exercice AC et les adresses déterminées comme des LC potentiels par d'autres sources sélectionnées, y compris les sources figurant ci-après. On a ainsi obtenu une liste de LC potentiels, qui est sujette à une validation individuelle dans l'exercice de VLC. Notre analyse est axée sur la contribution des dossiers administratifs des trois derniers groupes à la base de sondage des LC, par rapport à la contribution des dossiers non administratifs.

- LC du Recensement de 2000 (Précensement de 2010) – adresses de LC du Recensement de 2000 et mises à jour à mi-décennie.
- ALC déterminés pendant l'exercice AC de 2010.
- LC déterminés pendant l'exercice American Community Survey Time of Interview (ACS TOI).
- LC provenant de *dossiers administratifs*, comme le Federal-State Cooperative Program for Population Estimates (FSCPE) – un fichier de déclaration de LC des représentants participant au FSCPE, ainsi que des fichiers opérationnels pour le pays et l'État.
- Adresses de LC fournies par des participants gouvernementaux à l'exercice Local Update of Census Addresses (LUCA), que nous considérons comme des *dossiers administratifs* dans le présent document.
- LC déterminés par des recherches sur Internet effectuées par le National Processing Center (NPC). Le personnel du NPC a fait des recherches sur Internet concernant les emplacements de dénombrement fondé sur les services (DFS) et les foyers collectifs (FC) et a obtenu des renseignements auprès d'élus d'organismes de défense tribaux et nationaux au sujet des personnes sans abri, afin de mettre à jour la liste d'adresses des emplacements DFS (par exemple, gîtes, soupes populaires). Nous considérons cette source comme des *dossiers administratifs*.

2.2 Données : VLC de 2010 et base du Recensement de 2010

Le fichier de la base de VLC de 2010 comprenait des adresses considérées comme des ALC pendant l'AC et des adresses figurant dans le fichier principal d'adresses déterminées comme étant des LC possibles par d'autres sources sélectionnées. Ces adresses ont servi à l'exercice de VLC pour confirmer le statut des ALC comme LC, unités de logement et lieux transitoires, notamment. Le fichier de la base de dénombrement de 2010 a servi à identifier tous les dossiers qui figuraient dans la base de VLC et à déterminer s'ils figuraient ou non aussi dans l'AC. En outre, ce fichier comportait des valeurs pour les variables utilisées pour déterminer le statut de l'ALC, et a aussi servi à déterminer le cheminement de dénombrement de tous les dossiers de la base de sondage initiale de VLC⁸. De nombreuses sources de dossiers administratifs et non administratifs ont fourni à l'occasion la même adresse à la base de VLC. Lorsqu'il y avait chevauchement, nous n'avons pas pu effectuer d'analyse précise selon la contribution des différentes sources. Toutefois, nous avons pu déterminer les sources uniques et les sources multiples pour la même adresse. Cette méthode nous a permis d'isoler l'adresse de source unique et de déterminer les adresses fournies par des sources qui figuraient déjà dans le fichier principal des adresses de 2000 (c.-à-d. Précensement de 2010).

3. Résultats et discussion

3.1 Base de VLC selon la source

Nous avons examiné les résultats de la VLC selon les diverses sources ayant fourni des adresses de LC potentiels à la base de sondage initiale des LC, ainsi que la base finale des LC validés avant le dénombrement des LC (DLC). Cette analyse évalue les répercussions de plusieurs exercices d'établissement de listes.

Le tableau 1 montre le nombre d'adresses fournies par chaque source à la base de VLC. Dans les cas où il y avait

⁸ Les états finaux compris dans le présent document concernent les états de la VLC. Cette étape peut changer plus tard dans les autres activités du recensement.

plusieurs sources pour une adresse, chaque source a été considérée comme ayant fourni l'adresse à la base de VLC. Ainsi, le nombre d'adresses fournies par chaque source correspond à un nombre plus élevé que le nombre d'unités dans la base de VLC, en raison du chevauchement entre les sources multiples des mêmes adresses. Environ 58 % des adresses de la base de VLC figuraient déjà dans la base (c.-à-d. Prérecensement de 2010). Les dossiers administratifs, y compris LUCA, FSCPE et la recherche sur Internet d'emplacements DFS et de foyers collectifs, ont produit au total environ 17 % des adresses de la base de VLC. Les ajouts AC et les ajouts de la VLC ont été à l'origine de 16 % et 20 % des adresses au total, respectivement, de la base de VLC.

Tableau 1.
Nombre d'adresses fournies à la base initiale des LC selon la source – y compris le chevauchement

Source	Nombre	Pourcentage d'unités dans la base de VLC
Unités totales dans la base de VLC	2 551 447 ⁺	100,00
<i>Sources de dossiers administratifs</i>		
LUCA.....	339 206	13,29
FSCPE.....	45 678	1,79
NPC/DFS et FC.....	55 435	2,17
<i>Autres sources</i>		
Prérecensement de 2010	1 467 211	57,51
ACS TOI	4 006	0,16
Ajouts AC	418 779	16,41
Ajouts de la VLC	511 094	20,03
Notes : Les nombres et les pourcentages ne sont pas pondérés. Les nombres et les pourcentages ne correspondent pas au total parce que plusieurs sources ont fourni la même adresse.		
Sources de données : Extrait MTdb de l'univers initial de la VLC de 2010, extrait MTdb du dénombrement de 2010 et fichier ADDUP national de 2010.		

3.2 Base initiale de VLC selon la source pour les logements collectifs valides

Le tableau 2 fournit un aperçu global des résultats des adresses de la VLC déterminées comme étant des logements collectifs valides (LC, LC vacants, UL et LT) et des logements collectifs non valides (en double, supprimés et non résidentiels), selon la source. Dans le tableau 2, lorsqu'il y avait plusieurs sources pour la même adresse, chaque source a été considérée comme ayant fourni l'adresse à la base de VLC. Comme dans le tableau 1, le nombre d'adresses fournies par chaque source est supérieur au nombre d'unités de la base de VLC en raison du chevauchement lorsque plusieurs sources ont fourni la même adresse. Les pourcentages déclarés pour chaque source dans les sections qui suivent sont ceux qui ont été à l'origine de ce chevauchement.

LUCA (dossiers administratifs)

Comme le montre le tableau 2, environ 12 % des adresses fournies par LUCA ont été validées comme des LC durant la VLC. Environ 46 % des adresses de LUCA étaient des doubles, des suppressions et des adresses non résidentielles, environ 39 % étaient des UL et 3 % des LT.

FSCPE et recherche d'emplacements DFS sur Internet par le NPC (dossiers administratifs)

Le FSCPE et la recherche sur Internet d'emplacements DFS et de foyers collectifs faisaient partie de l'exercice de mise à jour des dossiers administratifs. Dans le cadre des programmes géographiques, on a élaboré des procédures pour mettre à jour la base des LC à partir des dossiers administratifs des administrations tribales et locales, des organismes de défense nationaux et d'État et d'autres dossiers administratifs. On a aussi utilisé Internet pour mettre à jour les emplacements DFS et les foyers collectifs. La plupart des adresses fournies par le FSCPE et la recherche sur Internet par le NPC ont été validées comme des LC pendant la VLC. Comme le montre le tableau 2, environ 61 % des 45 678 adresses fournies par le FSCPE et environ 57 % des 55 435 adresses fournies par la recherche sur Internet d'emplacements DFS et de foyers collectifs par le NPC ont été déterminées comme des LC (y compris des LC vacants). Parmi les LC fournis par la recherche sur Internet d'emplacements DFS et de foyers collectifs par le NPC, plus de la moitié (51,1 %) ont été validés comme des foyers collectifs pendant la VLC. Parmi eux, on comptait 14 593 foyers collectifs destinés aux adultes et 1 563 foyers collectifs destinés aux jeunes.

Prérecensement de 2010, ajouts AC de 2010 et ajouts de la VLC de 2010⁹

Le tableau 2 montre aussi que parmi le million et demi d'adresses environ comportant une source Prérecensement de 2010 dans la base de la VLC, environ 9 % ont été validées comme des LC. Parmi les 418 779 ALC ajoutés par suite de l'exercice Adress Canvassing de 2010, 7 % étaient des LC (y compris des LC vacants). Parmi le demi-million d'unités ajoutées par les responsables des listes de la VLC, environ 4 % étaient des LC, y compris des LC vacants.

Base initiale de VLC selon la source unique et les sources multiples

Une analyse supplémentaire abordée par Williams, Barrett et Williams (2013) – non montrée ici en raison du manque d'espace – présente chaque source qui a contribué à la base initiale de VLC, grâce à l'analyse du nombre d'adresses qui étaient fournies, de la façon suivante : source unique; sources multiples dont au moins une Prérecensement de 2010; et sources multiples pour lesquelles toutes les sources ont été ajoutées après les activités de 2010 (c.-à-d. des sources autres que Prérecensement de 2010). Il s'agit de l'approche adoptée pour résoudre le problème de chevauchement lorsque plusieurs sources ont contribué à la même adresse dans la base de VLC.

Pour toutes les sources autres que Prérecensement de 2010, la plupart des adresses fournies figuraient déjà dans la base de sondage. Les pourcentages allaient d'environ 59 % à 73 %. Parmi les 339 206 adresses fournies par LUCA, environ 59 % figuraient déjà dans la base de sondage. Toutefois, LUCA était la seule source pour environ 37 % des adresses ayant contribué à la base de VLC, le pourcentage le plus élevé parmi toutes les sources qui ont été catégorisées comme source unique¹⁰. Plus de 70 % des adresses fournies par les dossiers administratifs du FSCPE (32 584 des 45 678) et la recherche sur Internet d'emplacements DFS et de foyers collectifs (DFS/FC, 40 591 des 55 435) comportaient des sources multiples, dont au moins une était une source Prérecensement de 2010. Les dossiers administratifs du FSCPE et DFS/FC étaient la seule source pour environ 14 % et 15 % des adresses respectivement.

Dans le cas des sources fournissant des adresses de LC, Williams, Barrett et Williams (2013) signalent que le pourcentage de ces adresses figurant déjà dans la liste allait de près de 60 % à plus de 76 %. Les dossiers administratifs, comme le FSCPE et DFS/FC ont fourni les pourcentages les plus élevés (plus de 70 % respectivement) d'adresses figurant déjà dans la base de sondage. LUCA a fourni environ 60 % des adresses de LC qui figuraient déjà dans la base. Même si LUCA comme source unique a fourni davantage d'adresses de logements collectifs valides dans l'ensemble comparativement à ACS TOI, le FSCPE et la recherche sur Internet par le NPC seulement, il a aussi fourni davantage d'adresses non valides (39 %) comme source unique comparativement aux autres sources seulement.

Comparativement aux sources uniques, LUCA comme source unique de dossiers administratifs a aussi fourni davantage d'adresses d'unités de logement (39 % ou 50 877 des 131 672 et de LT, 42 %, ou 4 429 des 10 577 adresses de LT). Plus de 7 % (9 186) des 125 183 adresses ayant LUCA comme source unique dans la base de VLC ont été validées comme des LC, y compris des LC vacants.

⁹ Nous rappelons au lecteur que l'AC a entraîné l'ajout de LC potentiels et que la VLC a entraîné l'ajout de LC, ainsi que l'ajout, selon le plan de sondage, de toutes les unités de logement qui se sont révélées être des logements réguliers.

¹⁰ Nous rappelons au lecteur que les dossiers LUCA de la présente analyse ne représentent pas l'ensemble complet des ALC soumis par les participants au LUCA.

Tableau 2.
Résultats du statut d'adresse initial de la VLC selon la source – y compris les chevauchements

Source des adresses d'unités		LC et LC vacants		Unités de logement		Lieux transitoires		Doubles, suppressions, non résidentiels		
	Nombre*	Pourcentage du total ⁺	Nombre*	Pourcentage du total ⁺	Nombre*	Pourcentage du total ⁺	Nombre*	Pourcentage du total ⁺	Nombre*	Pourcentage du total ⁺
Total	2 551 447	100	205 406	8,05	1 391 477	54,54	80 483	3,15	874 081	34,26
<i>Sources de dossiers administratifs</i>										
LUCA	339 206	100	41 512	12,24	131 672	38,82	10 577	3,12	155,445	45,83
FSCPE	45 678	100	27 792	60,84	5 923	12,97	129	0,28	11,834	25,91
DFS/FC	55 435	100	31 636	57,07	11 697	21,10	529	0,95	11,573	20,88
<i>Autres sources</i>										
Prérecensement de 2010	1 467 211	100	133 222	9,08	747 368	50,94	40 896	2,79	545,725	37,19
ACS TOI	4 006	100	326	8,14	1 402	35,00	403	10,06	1,875	46,80
Ajouts AC	418 779	100	28 768	6,87	98 583	23,54	33 794	8,07	257,634	61,52
Ajouts de la VLC	511 094	100	19 352	3,79	490 051	95,88	1 045	0,20	646	0,13

*Les nombres et les pourcentages ne sont pas pondérés. Il se peut que les nombres ne correspondent pas au total parce que plusieurs sources ont fourni la même adresse. Il se peut que les pourcentages ne totalisent pas 100 en raison de l'arrondissement. Sources de données : Extrait MTdb de l'univers initial de la VLC de 2010, extrait MTdb du Recensement de 2010 et fichier ADDUP national de 2010.

4. Contribution de chaque source à la base postérieure à la VLC

Comme le montre le tableau 4 ci-après, il y avait au total 205 406 LC validés, y compris 44 120 LC vacants au moment de l'interview de la VLC, qui ont été trouvés pendant l'exercice de VLC de 2010. Parmi les différentes sources, celle Prérecensement de 2010 a contribué le plus à la base de sondage de LC, soit environ 65 % des 205 406 LC trouvés pendant l'exercice de VLC. LUCA a aussi fourni un nombre important d'adresses confirmées par la VLC comme des LC, soit plus de 20 % de ces adresses. Le FSCPE et la recherche sur Internet de DFS et de foyers collectifs – les deux sources de dossiers administratifs – ont aussi contribué de façon marquée à la base de sondage de LC validés : 14 % et 15 % respectivement. Les pourcentages comprennent le chevauchement avec d'autres sources et, ainsi, ne totalisent pas 100 %.

Tableau 4.
Nombre d'adresses dans la base postérieure à la VLC, selon la source – y compris le chevauchement

Source des adresses d'unité	Nombre*	Pourcentage du total ⁺	LC		LC vacants	
			Nombre*	Pourcentage du total ⁺	Nombre*	Pourcentage du total ⁺
Total dans la base postérieure à la VLC	205 406	100,00	161 286	100,00	44 120	100,00
<i>Sources de dossiers administratifs</i>						
LUCA	41 512	20,21	33 338	20,67	8 174	18,53
FSCPE	27 792	13,53	25 628	15,89	2 164	4,90
DFS/FC	31 636	15,4	27 999	17,36	3 637	8,24
<i>Autres sources</i>						
Prérecensement de 2010	133 222	64,86	103 604	64,24	29 618	67,13
ACS TOI	326	0,16	189	0,12	137	0,31
Ajouts AC	28 768	14,01	19 212	11,91	9 556	21,66
Ajouts de la VLC	19 352	9,42	18 471	11,45	881	2,00
*Les nombres et les pourcentages ne sont pas pondérés.						
⁺ Les nombres et les pourcentages ne correspondent pas au total parce que plusieurs sources ont fourni la même adresse.						
Sources de données : Extrait MTdb de l'univers initial de la VLC de 2010, extrait MTdb du Recensement de 2010 et fichier ADDUP national de 2010.						

5. Conclusions et programme de recherches futures

On a déterminé que les dossiers administratifs étaient utiles pour améliorer la couverture de la base des LC dans le Recensement de 2010. On a déterminé que les trois principales sources de dossiers administratifs – LUCA, FSCPE, et recherche sur Internet d'emplacements DFS et de foyers collectifs ont contribué considérablement à compléter la base des LC, dans des proportions différentes toutefois. LUCA comme source unique a fourni davantage d'adresses de logements collectifs valides, environ 37 % des adresses qui ont contribué à la base de la VLC, comparativement à ACS Time of Interview, le FSCPE et la recherche sur Internet par le NPC seulement. Toutefois, LUCA a aussi fourni davantage d'adresses de logements collectifs non valides (39 %) comme source unique comparativement aux autres sources seulement. Les dossiers administratifs du FSCPE et de DFS/FC étaient la seule source de la base des LC, à 14 % et 15 % respectivement. Compte tenu des avantages et des limites des dossiers administratifs par rapport à leur utilisation efficace pour la validation des LC et l'élaboration de la base de sondage, nous croyons que les recherches suivantes devraient être envisagées dans un avenir rapproché.

Tout d'abord, même si des types de LC déterminés ont contribué de façon substantielle à la base des LC, environ 22 % (44 120 des 205 406 LC au total) des unités étaient des LC vacants qui avaient un type de LC inconnu. Pour réduire le nombre de LC dont le type est inconnu, nous recommandons de chercher les dossiers administratifs qui pourraient être utilisés pour attribuer des codes de types de LC possibles dans les exercices subséquents liés aux LC. En deuxième lieu, il est recommandé d'effectuer des études pour déterminer les types de LC pour lesquels le taux de validation des LC à partir des dossiers administratifs peut être optimisé et rentabilisé. On encourage le Census Bureau à continuer d'obtenir des dossiers administratifs des gouvernements fédéral, d'État, tribaux et locaux. Les sources combinées de dossiers administratifs (FSCPE et recherche sur Internet d'emplacements DFS et de foyers collectifs) ont contribué à la base de LC dans une proportion de 29 %. Par ailleurs, nous rappelons que la majorité des adresses fournies par les dossiers administratifs (c.-à-d. 61 % des 45 678 adresses fournies par le FSCPE et 57 % des 55 435 adresses fournies par la recherche sur Internet par le NPC) ont été validées comme des LC au cours de l'exercice de VLC de 2010. En troisième et dernier lieu, nous pouvons explorer des façons innovatrices d'utiliser le Registre des entreprises élaboré par le Census Bureau pour les recensements économiques, en vue d'identifier

certaines LC sur une base permanente. Nous devrions aussi explorer une méthode plus efficace et rentable pour tirer parti de la recherche sur Internet qui a été menée en 2010 pour mettre à jour les emplacements DFS et les foyers collectifs, et inclure d'autres types de logements collectifs. Par exemple, les sites Web des collèges et des universités peuvent comporter des renseignements à jour concernant le logement des étudiants et les dortoirs. La recherche sur Internet par le NPC d'emplacements DFS et de foyers collectifs a permis de déterminer 3 900 LC. Ces adresses ne figuraient pas déjà dans le fichier principal des adresses et aucune autre source ne les avait soumises pour les inclure dans la base de VLC de 2010.

Bibliographie

- Chan, A. (2012). 2010 Census Program for Evaluations and Experiments, Center for Survey Measurement. *Investigation of methods to evaluate the coverage of group quarters populations report.*
- Chun, Y., E. Chan, D. Barrett, A. Williams et J. Belton (2012). *What Counts as Group Quarters? A Glimpse of Total Coverage Errors and Costs Model in Census.* Document présenté aux Joint Statistical Meetings de 2012 dans une séance de communications libres organisée par Chun, Chan et Scheuren. San Diego, Californie.
- Chun, Y. et J. Gan (2014). *Analysis of the Sources of Group Quarters Enumeration Data in the 2010 Census.* Document présenté à la séance des méthodes de recherche d'enquête des Joint Statistical Meetings, Boston, Massachusetts.
- Chun, Y. et F. Scheuren (2011). *Pandata Systems to Enhance Survey and Census Systems: Employing Administrative Data to Better Inform Public Policy.* Communication sollicitée présentée aux Joint Statistical Meetings de 2011.
- Jonas, K. (2003). *Group quarters enumeration. Census 2000 Evaluation E.5, Revision I.* Extrait de <https://www.census.gov/pred/www/rpts/E.5%20R.pdf>.
- National Research Council of the National Academies. (2011). *Change and the 2020 Census: Not Whether but How.*
- Nordbotten, S. (2009). *Use of Administrative Data in Official Statistics – Past, Present and Future – With Special Reference to the Nordic Countries.* Official Statistics in Honour of Daniel Thorburn : 205 à 223.
- Russell, D. et D.F. Barrett (2012). *2010 census service-based enumeration operation assessment report.* United States Census Bureau, Decennial Management Division/Decennial Statistical Studies Division, <http://www.census.gov/2010census/pdf/2010%20Census%20Service-Based%20Enumeration%20Assessment%20Report.pdf>.
- Sundgren, B. (2010). *Towards a system of official statistics based on a coherent combination of data sources, including surveys and administrative data.* Document présenté à la conférence annuelle de l'European Survey Research Association, Lausanne, Suisse.
- Wallgren, A. et B. Wallgren (2007). *Register-based statistics: Administrative data for statistical purposes.* John Wiley & Sons, Ltd.
- Williams, J., D. Barrett, A. Williams (2013). *2010 census group quarters validation assessment report.* U.S. Census Bureau, Suitland, MD. [http://www.census.gov/2010census/pdf/2010%20Census%20Group%20Quarters%20Validation%20Operation%20Assessment%20\(Reissue\).pdf](http://www.census.gov/2010census/pdf/2010%20Census%20Group%20Quarters%20Validation%20Operation%20Assessment%20(Reissue).pdf).
- Williams, J., B. De Vos, D. Russell, A. Williams et D. Barrett (2013). *2010 Census group quarters enumeration assessment report.* United States Census Bureau. Extrait de <http://www.census.gov/2010census/pdf/2010%20Census%20Group%20Quarters%20Enumeration%20Assessment%20Report.pdf>.