



INSTITUT
DE LA STATISTIQUE
DU QUÉBEC

www.stat.gouv.qc.ca

Enquête nationale auprès des ménages de 2011 : ses portées et limites

Bilan de la consultation de l'Institut
de la statistique du Québec auprès
des ministères et organismes
gouvernementaux du Québec



Pour tout renseignement concernant l'ISQ
et les données statistiques dont il dispose,
s'adresser à :

Institut de la statistique du Québec
200, chemin Sainte-Foy
Québec (Québec)
G1R 5T4
Téléphone : 418 691-2401

ou

Téléphone : 1 800 463-4090
(sans frais d'appel au Canada et aux États-Unis)

Site Web : www.stat.gouv.qc.ca

Ce document n'est disponible qu'en version électronique.

Dépôt légal
Bibliothèque et Archives Canada
Bibliothèque et Archives nationales du Québec
1^{er} trimestre 2015
ISBN 978-2-550-72229-8 (PDF)

© Gouvernement du Québec, Institut de la statistique du Québec, 2015

Toute reproduction autre qu'à des fins de consultation personnelle
est interdite sans l'autorisation du gouvernement du Québec.
www.stat.gouv.qc.ca/droits_auteur.htm.

Janvier 2015

AVANT-PROPOS

En 2010, le gouvernement du Canada a décidé d'apporter des changements majeurs au Programme du Recensement de 2011. En effet, le questionnaire long et obligatoire du recensement a été remplacé par une enquête volontaire auprès du tiers des ménages canadiens, intitulée l'*Enquête nationale auprès des ménages* (ENM). Cette décision a suscité de vives réactions de la part des acteurs socioéconomiques et des chercheurs.

Les premiers résultats de l'ENM ont été diffusés en 2013. À titre de coordonnateur statistique pour le Québec, l'Institut de la statistique du Québec a procédé à une large consultation auprès des ministères et organismes gouvernementaux du Québec (MO) afin de dresser un portrait précis des problèmes rencontrés, tant à l'égard de la qualité des données disponibles que des outils statistiques déployés.

Le présent rapport fait état des résultats de cette consultation. Outre le partage de cette information entre les divers MO, ce diagnostic permet de démontrer, une fois encore, que l'abandon du caractère obligatoire du questionnaire long du recensement est lourd de conséquences, pour notamment l'élaboration et l'évaluation des politiques et des programmes publics. De plus, tant pour les divers MO que pour les chercheurs du Québec et Statistique Canada, ce bilan offre un appui pour la recherche de solutions. Ces stratégies sont d'autant nécessaires que le gouvernement fédéral envisage de reconduire l'ENM en 2016.

Je souhaite remercier tous les intervenants des divers MO rencontrés qui, grâce à leur précieuse collaboration, ont permis la réalisation de ce rapport. Je suis convaincu que les informations présentées dans les pages qui suivent contribueront à alimenter les réflexions visant à permettre aux différents acteurs de la société québécoise de disposer de l'information statistique essentielle à la gouverne et au développement du Québec.

Le directeur général

A handwritten signature in black ink, appearing to read 'SM', written in a cursive style.

Stéphane Mercier

Cette publication a été réalisée par :

Sylvie Rheault
Direction des statistiques sociodémographiques
Marie-Eve Tremblay
Direction de la méthodologie et de la qualité
Anne Baillargeon
Secrétariat et Affaires juridiques

Sous la coordination de :

Sylvie Rheault

Direction générale adjointe
aux statistiques et à l'analyse sociale :

Patricia Caris, directrice générale adjointe

Avec la collaboration de :

Stéphane Crespo, pour l'analyse sur la variable « revenu »
Annie Dupont, pour le soutien à la gestion de l'information
Direction des statistiques sociodémographiques
Danielle Laplante, pour la coordination de l'édition
Esther Frève, pour la révision linguistique
Gabrielle Tardif, pour la mise en page
Direction des communications

Remerciements

Nous remercions Paul Berthiaume et Chantal Girard de la Direction des statistiques sociodémographiques et France Lapointe de la Direction de la méthodologie et de la qualité pour leurs précieux commentaires sur la version préliminaire de ce rapport.

Pour tout renseignement concernant
le contenu de cette publication :

Direction des statistiques sociodémographiques
Institut de la statistique du Québec
200, chemin Sainte-Foy, 3^e étage
Québec (Québec) G1R 5T4

Téléphone : 418 691-2406
ou
1 800 463-4090
(sans frais d'appel au Canada
et aux États-Unis)

Télécopieur : 418 643-4129

Site Web : www.stat.gouv.qc.ca

Avertissement

L'Institut de la statistique du Québec (Institut) n'est pas responsable de la véracité et de la qualité des informations qui lui ont été communiquées par les différents ministères et organismes gouvernementaux du Québec (MO) dans le cadre de la rédaction de ce rapport.

LISTE DES ACRONYMES

AADNC	Affaires autochtones et Développement du Nord Canada	FIPA	Fichier d'inscription des personnes assurées de la Régie de l'assurance maladie du Québec
AD	Aires de diffusion	FMGD	Fichier de microdonnées à grande diffusion
ADS	Analyse différenciée selon les sexes	IBC	Indice de bien-être
AR	Agglomération de recensement	IMSE	Indice de milieu socioéconomique
CANSIM	Canadian socio-economic information management system	LRC	Lois et règlements du Canada
CDR	Centre de données de recherche	LRQ	Lois et règlements du Québec
CLE	Centre local d'emploi	MFR	Mesure du faible revenu
CLSC	Centre local de services communautaires	MO	Ministères et organismes gouvernementaux du Québec
CNP	Classification nationale des professions	MRC	Municipalité régionale de comté
CPE	Classification des programmes d'enseignement	OD	Origine destination
CRE	Conférence régionale des élus	RLRQ	Recueil des lois et des règlements du Québec
CV	Coefficient de variation	RLS	Réseau local de services
EAPA	Enquête auprès des peuples autochtones	RMR	Région métropolitaine de recensement
ECI	Enquête canadienne sur l'incapacité	RQ	Revenu Québec
EDM	Enquête sur les dépenses des ménages	RRQ	Régie des rentes du Québec
EDTR	Enquête sur la dynamique du travail et du revenu	SC	Statistique Canada
EERH	Enquête sur l'emploi, la rémunération et les heures de travail	SCHL	Société canadienne d'hypothèque et de logement
ENM	Enquête nationale auprès des ménages	SCIAN	Système de classification des industries de l'Amérique du Nord
EPA	Enquête sur la population active	SDR	Subdivision de recensement
EREQ	Enquête sur le recrutement et l'emploi au Québec	SFR	Seuil de faible revenu
ESG	Enquête sociale générale	SR	Secteur de recensement
FFT1	Fichier sur la famille T1	TGN	Taux global de non-réponse

TABLE DES MATIÈRES

9	FAITS SAILLANTS
11	INTRODUCTION
13	LES APPRÉHENSIONS ET LES RÉACTIONS EN 2010
13	Un consensus clair contre l'abolition du questionnaire détaillé obligatoire
13	Des craintes manifestes de la part des ministères et organismes gouvernementaux du Québec
14	Les finances publiques
14	La Charte de la langue française
14	Le soutien aux groupes vulnérables
15	Le développement économique régional et d'aménagement du territoire
15	Des interventions fréquentes de l'Institut
17	DÉMARCHE DE CONSULTATION DE L'INSTITUT EN 2014
17	Un bilan en appui à la recherche de solutions
17	La collecte d'information
18	Le traitement
19	À QUOI SERVENT LES DONNÉES PROVENANT DU QUESTIONNAIRE DÉTAILLÉ ?
19	Bien connaître sa population
20	Évaluer nos avancées... ou nos reculs
22	Répondre aux besoins précis
22	Allouer des ressources à l'aide de données probantes
23	QUELLES DONNÉES DU QUESTIONNAIRE COURT DU RECENSEMENT ET DE L'ENM DE 2011 ONT ÉTÉ EXPLOITÉES ?
23	Rappel des dates de diffusion
23	Recensement
23	Enquête nationale auprès des ménages
24	L'outil de départ: les produits normalisés
24	Des commandes nombreuses et variées
25	L'accès aux données dans les centres de données de recherche (CDR), une avenue surtout exploratoire

27	DES OUTILS INSUFFISANTS POUR ÉVALUER LA QUALITÉ DES ESTIMATIONS DE L'ENM
27	Biais
28	Précision
30	Cohérence
30	Comparabilité
32	Règles de confidentialité
32	Soutien obtenu de la part de SC et de l'Institut
34	Expertises
35	DES DÉMARCHES DE VALIDATION EXIGEANTES, MAIS PAS TOUJOURS CONCLUANTES
35	Données du Recensement
36	Données administratives
37	Données provenant d'autres enquêtes
39	En l'absence de sources de données comparables ?
41	DE FORTES INQUIÉTUDES SUR LA QUALITÉ DES DONNÉES SUR LE REVENU
41	Quelques constats des analystes des MO
41	validation avec des données fiscales
42	Selon l'échelle géographique
43	Selon les populations
43	Selon les indicateurs de revenu
44	Selon les concepts du revenu
45	DES DÉCISIONS DIVERGENTES QUANT AUX ESTIMATIONS DE L'ENM POUVANT ÊTRE DIFFUSÉES OU UTILISÉES
45	Difficile de faire fi des pressions
45	Diffuser sans restriction
46	Diffuser avec restriction
46	Des décisions en attente
47	PEU DE PISTES DE SOLUTION POUR LES MO
47	Des ajustements à court terme
47	Utilisation et bonification d'autres sources de données
48	Production de nouvelles enquêtes
48	Amélioration du processus de diffusion des données de l'ENM
49	LES PRINCIPAUX ENJEUX ET DÉFIS QUE REPRÉSENTE L'ENM
49	Un plus lourd fardeau pour les MO
49	Des conséquences négatives pour les politiques et programmes
49	Des impasses pour la prise de décision
50	Des objectifs non atteints

FAITS SAILLANTS

La consultation menée en 2014 par l'Institut de la statistique du Québec auprès des ministères et organismes gouvernementaux du Québec (MO) a permis de dresser un portrait précis des problèmes rencontrés, tant à l'égard des données qui ont pu être utilisées que des outils statistiques disponibles, avec la nouvelle *Enquête nationale auprès des ménages* (ENM) de Statistique Canada. Rappelons que le gouvernement fédéral a modifié son Programme du Recensement, en remplaçant le questionnaire détaillé et obligatoire par une enquête volontaire. En résumé :

- L'ENM est une bonne enquête socioéconomique, mais elle ne répond pas aux objectifs fixés, soit de fournir des données fiables pour certaines sous-populations et pour des unités géographiques fines.
- En raison du peu d'outils méthodologiques mis à leur disposition par Statistique Canada, les MO sont beaucoup laissés à eux-mêmes pour évaluer la qualité des données.
- Les MO doivent effectuer de nombreux travaux de validation, lesquels exigent le déploiement de ressources humaines et financières supplémentaires, ainsi que des expertises spécifiques. Malgré les coûts additionnels que cela implique, les conclusions ne s'avèrent pas toujours concluantes en raison du peu d'outils statistiques mis à leur disposition.
- D'importantes mises en garde ont été effectuées par Statistique Canada sur la comparabilité des données du questionnaire détaillé du Recensement de 2006 et de l'ENM 2011, notamment pour les estimations de taux de faible revenu. Quoique ces mises en garde soient grandement justifiées en raison du passage d'un questionnaire obligatoire à une enquête volontaire, la perte de tendances historiques pose de sérieuses difficultés aux MO.
- Dans 22 % des municipalités du Québec, l'écart relatif du revenu total médian des personnes avec revenus, entre l'ENM et les données fiscales de Revenu Québec, se situe entre plus ou moins 2,5 %. Mais pour 14 % de celles-ci, cet écart se situe entre plus ou moins 20 %, ce qui soulève des questions quant à la qualité de ces données.
- L'élaboration, le suivi et l'évaluation de plusieurs programmes et politiques s'appuient sur des données probantes. Les MO doivent maintenant composer avec des données en moins et, par conséquent, prendre certaines décisions à l'aveugle.
- En somme, l'ENM s'est avérée plus coûteuse à administrer et à traiter que le questionnaire détaillé et obligatoire du Recensement de 2006, et les résultats obtenus sont de moindre qualité.

INTRODUCTION

Le 26 juin 2010, le gouvernement fédéral rend public, par la voie de la Gazette du Canada, le contenu du questionnaire court du Recensement, mais non celui du questionnaire détaillé. En réponse aux interrogations, le gouvernement fédéral précise alors que le questionnaire détaillé et obligatoire du Recensement est aboli et remplacé par une enquête volontaire intitulée l'*Enquête nationale auprès des ménages* (ENM). Rappelons que le Programme du Recensement pour les cycles précédents comprenait un questionnaire court et obligatoire pour toute la population canadienne et un questionnaire détaillé et obligatoire, adressé à un ménage sur cinq. Or, pour 2011, le Programme du Recensement comprend le questionnaire court et obligatoire pour toute la population canadienne et l'ENM. En raison du caractère volontaire de cette enquête, et donc du taux de réponse attendu plus faible qu'un recensement, un ménage sur trois sera sondé au lieu d'un ménage sur cinq.

Le questionnaire court et obligatoire sert essentiellement à obtenir des données populationnelles de base (âge, sexe, état matrimonial), sur la situation du ménage (nombre de membres dans le ménage, type de familles, etc.), ainsi que sur la langue. Le questionnaire détaillé et obligatoire du Recensement, maintenant remplacé par l'ENM, a pour principal objectif de recueillir des données sur les conditions de vie de la population (logement, scolarité, travail, revenu, mobilité et migration, diversité ethnique, etc.). Mais plus encore, d'obtenir des données socio-économiques pour des petites unités géographiques, par exemple, un quartier ou un village.

Le questionnaire court et obligatoire a aussi été modifié entre les recensements de 2006 et 2011. D'abord, avant même l'annonce du changement au Programme du Recensement, on avait prévu collecter des informations pour identifier le nombre de familles recomposées. Par ailleurs, lors du dépôt du contenu du questionnaire court et obligatoire le 26 juin 2010, les questions sur la connaissance des langues officielles et sur la langue parlée à la maison n'étaient pas incluses. Elles ont été ajoutées ultérieurement afin de répondre aux exigences de la Loi sur les langues officielles (L.R.C 1985, ch. 31, 4^e suppl.).

La collecte de données pour le Recensement s'est déroulée en mai 2011, pour un recensement en date du 10 mai, alors que pour l'ENM, la collecte de données a eu lieu de mai à août 2011 pour la même date de référence, soit le 10 mai. La collecte de données de l'ENM chevauchait ainsi une période où un grand nombre de ménages québécois changent d'habitation.

Comme nous le verrons dans ce rapport, les changements apportés au Programme du Recensement en 2011 ne sont pas sans conséquence. Le passage d'une enquête obligatoire à une enquête volontaire n'a, en effet, rien d'anodin. Pour évaluer plus précisément les impacts de ce changement, l'Institut de la statistique du Québec (Institut) a procédé à une large consultation en 2014 auprès des ministères et organismes gouvernementaux (MO). Avec les témoignages et les travaux de ceux et celles qui produisent des données à partir du Programme du Recensement, ce rapport rend compte des principales conséquences de ce changement.

Dans un premier temps, les appréhensions et réactions recueillies par l'Institut en 2010, après l'annonce du gouvernement fédéral, seront résumées. La démarche de consultation de 2014 sera ensuite décrite. Pour évaluer l'importance des changements, nous nous intéressons d'abord aux besoins des MO en termes de données provenant du Programme du Recensement, et nous examinons quelles données ont été exploitées pour 2011. Après cette mise en contexte, nous faisons état des outils statistiques mis à la disposition des utilisateurs de données et des démarches de validation entreprises par les MO et l'Institut pour évaluer la qualité des données de l'ENM. Les décisions prises quant à l'utilisation de ces données ainsi que les pistes de solution envisagées sont aussi abordées. Ce rapport se conclut par la présentation des principaux enjeux et défis liés à l'ENM.

LES APPRÉHENSIONS ET LES RÉACTIONS EN 2010

Dans cette section, nous ferons état des premières réactions qu'a suscitées la décision du gouvernement fédéral d'éliminer le questionnaire détaillé et obligatoire du Recensement et de le remplacer par une enquête volontaire, l'ENM. Ensuite, les propos recueillis par l'Institut auprès des MO en 2010 seront résumés. Finalement, la description des interventions effectuées par l'Institut auprès de son homologue fédéral et des MO termine cette section.

UN CONSENSUS CLAIR CONTRE L'ABOLITION DU QUESTIONNAIRE DÉTAILLÉ OBLIGATOIRE

Rares sont les décisions gouvernementales qui sont décriées par des groupes de la population présentant des intérêts aussi différents. Tant les chercheurs que les municipalités, les syndicats, les représentants d'entreprises privées, les organisations communautaires ou de défense d'intérêt ainsi que les organisations gouvernementales ont manifesté leur désapprobation. Quoique les positions défendues par ces groupes sur la place publique puissent être très divergentes, le consensus est toutefois clair sur le besoin de données fines et de qualité afin de répondre à leur mandat.

Ces oppositions ont trouvé écho à l'Assemblée nationale du Québec, puisque les trois partis qui siégeaient à ce moment ont présenté conjointement le 28 septembre 2010 une motion sans préavis. Elle se lit comme suit :

Pour répondre aux besoins essentiels en termes de données socioéconomiques, que l'Assemblée nationale demande au gouvernement fédéral de revoir sa décision d'éliminer le questionnaire long et obligatoire du Recensement de la population et l'invite à s'assurer que celui-ci continue d'être utilisé.

La motion a été adoptée à l'unanimité par tous les députés, sans abstention.

Avant l'adoption de cette motion, deux lettres ont été transmises au Gouvernement fédéral pour s'opposer à cette décision. La première, en date du 16 juillet 2010,

provenait de M. Raymond Bachand, ministre des Finances du Québec et s'adressait à M. Tony Clément, ministre de l'Industrie et ministre responsable de Statistique Canada (SC). La seconde, en date du 2 août 2010, provenait de M. Claude Béchar, ministre responsable des Affaires intergouvernementales canadiennes et de la Réforme des institutions démocratiques et s'adressait à Mme Josée Verner, ministre des Affaires intergouvernementales.

Hormis les changements qui ont été apportés au questionnaire court et obligatoire du Recensement afin d'y inclure les questions permettant le suivi de l'application de la Loi sur les langues officielles, aucune autre action n'a été prise par le Gouvernement fédéral afin de rectifier la situation. Il est important de mentionner qu'aucune consultation concernant le remplacement du questionnaire détaillé et obligatoire par une enquête volontaire n'avait été faite au préalable auprès des provinces. Soulignons aussi que le 21 juillet 2010, le statisticien en chef de SC, M. Munir A. Sheikh, a remis sa démission afin d'exprimer publiquement son opposition quant à la décision du gouvernement fédéral.

DES CRAINTES MANIFESTES DE LA PART DES MINISTÈRES ET ORGANISMES GOUVERNEMENTAUX DU QUÉBEC

En tant que coordonnateur statistique pour le Québec, l'Institut a mené, à la suite de l'annonce du gouvernement fédéral d'éliminer le questionnaire détaillé et obligatoire du Recensement et de le remplacer par une enquête volontaire, une consultation auprès des MO afin d'identifier et de cerner l'ensemble des impacts attendus des changements apportés au Programme de Recensement de 2011. Dans un premier temps, cette consultation a pris la forme d'ateliers, tenus les 19 et 20 août 2010, auxquels ont participé les représentants d'une cinquantaine de MO. Dans un second temps, l'Institut a réinvité les MO à documenter les impacts appréhendés de ces changements pour leur organisation.

La majorité des problèmes méthodologiques soulevés découlent du caractère volontaire de la nouvelle enquête. Pour les experts et spécialistes rencontrés, cette caractéristique de l'enquête nuira inévitablement à la qualité des données. Par ailleurs, il apparaît que le bruit et la confusion entourant l'opération pourraient être préjudiciables à la qualité de l'exercice. Bien que le questionnaire court du Recensement demeure obligatoire, une certaine confusion pourrait survenir en raison du caractère volontaire de l'ENM, entraînant un plus faible taux de réponse au questionnaire court.

De manière plus spécifique, les quatre conséquences méthodologiques appréhendées sont : 1) la sous-représentation des groupes vulnérables; 2) l'absence ou le manque de fiabilité des données à des niveaux géographiques fins; 3) la perte de comparabilité des données au fil du temps et 4) l'impact sur la qualité des autres enquêtes de SC.

Les impacts sur les grandes missions de l'État québécois appréhendés à ce moment sont nombreux et sont résumés ci-après, en regard de quatre grandes missions de l'État québécois.

Les finances publiques

En lien avec la prévision de revenus, Hydro-Québec anticipe des impacts négatifs sur ses prévisions des ménages privés du Québec et des municipalités régionales de comté ainsi que sur ses prévisions de la demande en électricité et des revenus, en particulier pour le secteur résidentiel. Enfin, la situation rend plus difficile le suivi des programmes destinés à la clientèle à faible revenu et celui sur l'efficacité énergétique.

La Charte de la langue française

Le suivi de la situation linguistique au Québec pourrait devenir plus complexe à réaliser en raison d'une perte de comparabilité historique des données. À la suite des changements majeurs apportés à la méthode de collecte de cette information, la situation linguistique dressée pour 2011 à l'aide des données du prochain Recensement pourra difficilement être comparée à celle des recensements précédents. Les tendances observées seront incertaines, car il sera difficile de distinguer les variations dues aux changements apportés à la méthode de collecte de celles dues à de réels changements survenus dans la situation linguistique au Québec.

Le soutien aux groupes vulnérables

L'application de la Loi sur les services de santé et les services sociaux (L.R.Q., chapitre S-4.2) devient plus incertaine. Cette Loi exige une connaissance des déterminants de la santé et du bien-être ainsi que l'atteinte de niveaux comparables de santé et de bien-être au sein de différents groupes de la population. La fiabilité des données socioéconomiques et sociodémographiques sur les groupes vulnérables et sur l'ensemble de la population à des niveaux géographiques très fins ainsi que la comparabilité historique de celles-ci sont essentielles à l'atteinte des objectifs prescrits par la Loi.

Les programmes d'éducation liés à l'alphabétisation, aux élèves autochtones et aux élèves issus de l'immigration reposent également sur la disponibilité d'informations fiables concernant ces clientèles à des niveaux géographiques très fins. De plus, pour faire le suivi de ces programmes, la comparabilité historique des données est très importante. L'indice du milieu socioéconomique est basé sur des données relatives aux familles, aux ménages, à l'emploi et à la scolarité provenant du questionnaire détaillé du Recensement. Cet indice, qui mesure la défavorisation en milieu scolaire, permet l'identification des écoles québécoises nécessitant des ressources spécialisées supplémentaires pour répondre aux besoins spécifiques de leur clientèle.

La Loi sur le ministère de l'Emploi et de la Solidarité sociale et sur la Commission des partenaires du marché du travail (2001, c. 44, a. 28; 2007, c. 3, a. 30.150) est exigeante en matière de disponibilité d'information spécifique relative au marché du travail. La qualité des services publics reliés à l'emploi et à la politique active du marché du travail pourrait être réduite en raison des changements apportés, l'information qu'ils rendent disponible à la population québécoise étant en grande partie tirée des données du questionnaire détaillé du Recensement.

La connaissance de la situation des groupes en situation de pauvreté est essentielle dans le cadre des mesures de lutte à la pauvreté. Les outils statistiques constituent les pierres d'assise de l'intervention gouvernementale. Pour ce faire, un indicateur de défavorisation matérielle et sociale permet de distribuer les sommes dévolues au Fonds québécois d'initiatives sociales de manière juste et équitable entre toutes les régions du Québec. Ainsi, à long terme, le possible manque de données fiables pourrait jouer au niveau de l'évaluation du *Plan d'action*

gouvernemental pour la solidarité et l'inclusion sociale 2010-2015. Or, la construction de cet indicateur dépend de données provenant du questionnaire détaillé, et ce, pour des niveaux géographiques fins.

La Loi assurant l'exercice des droits des personnes handicapées en vue de leur intégration scolaire, professionnelle et sociale confie à la ministre de l'Emploi et de la Solidarité sociale la responsabilité d'intégrer les personnes handicapées en emploi. Dans le cadre de cette responsabilité, la ministre doit évaluer et mesurer la situation de l'emploi chez les personnes handicapées. Cette évaluation et ce suivi sont réalisés à l'aide d'une enquête postcensitaire, *l'Enquête sur la participation et les limitations d'activité*¹. La fiabilité des données de l'ENM est donc préalable à la réalisation de cette obligation légale.

Le suivi des programmes destinés aux populations ethno-culturelles, le respect des politiques sur l'immigration et l'intégration des immigrants ainsi que l'offre de service répondant adéquatement aux besoins de ces groupes de population nécessitent une bonne connaissance de leurs caractéristiques socioéconomiques. Historiquement, ces derniers sont moins portés à répondre aux enquêtes à participation volontaire, ce qui pourrait engendrer une sous-représentation de ce groupe dans les données. Si tel était le cas, l'information disponible pour l'étude des caractéristiques de ces groupes serait moins fiable, ce qui pourrait entraîner une gestion moins efficiente des programmes et des politiques et nuire à la prise de décision éclairée.

La situation économique des artistes fait l'objet d'un suivi attentif des intervenants en matière culturelle. Les représentants de ces MO appréhendent des difficultés dans le suivi de ce groupe mouvant de la population québécoise.

Le développement économique régional et d'aménagement du territoire

La connaissance des caractéristiques socioéconomiques et sociodémographiques de la population, des ménages et des familles québécoises à l'échelle locale, c'est-à-dire des municipalités régionales de comté (MRC) et de manière plus fine encore, est fondamentale à la prise de décision en matière de développement économique régional et d'aménagement du territoire. Par exemple,

plus de 80 % des municipalités visées par le *Plan d'action gouvernemental à l'intention des municipalités dévitalisées* comptent moins de 1 000 habitants, et près de la moitié, moins de 500 habitants. Il importe notamment de connaître les taux d'emploi, de chômage, de faible revenu et de diplomation des populations de ces municipalités ainsi que le nombre moyen d'enfants par famille. La fiabilité des données de l'ENM de 2011 à des niveaux géographiques fins prend toute son importance.

Les données du questionnaire détaillé et obligatoire du Recensement sont utiles à la prévision de la demande pour la planification des transports urbains. À titre d'exemple, l'étude du navettage entre le domicile et le lieu de travail requiert des données fiables, disponibles selon un découpage géographique très fin.

DES INTERVENTIONS FRÉQUENTES DE L'INSTITUT

En tant que coordonnateur statistique dont la mission première est de produire, d'analyser et de diffuser des informations statistiques officielles, objectives et de qualité, l'Institut est grandement interpellé par ce changement de cap du gouvernement fédéral. Le 15 juillet 2010, le directeur général de l'Institut, M. Stéphane Mercier, écrit au Statisticien en chef de SC pour déplorer cette situation et son impact pour les divers acteurs de la société québécoise. La même journée, l'Institut diffuse un communiqué faisant état de l'impact de l'abandon du questionnaire détaillé et obligatoire du Recensement et son remplacement par une enquête volontaire, notamment sur trois dimensions de la qualité de l'information statistique : la fiabilité, la comparabilité et la cohérence. Il y est aussi stipulé que l'Institut fera les interventions appropriées auprès de SC afin que les impacts méthodologiques pressentis soient atténués et que les besoins de la société québécoise en matière d'information statistique soient comblés le plus adéquatement possible.

Dans la foulée de la consultation menée auprès des MO à la fin de l'été 2010, et dont les principaux constats ont été énoncés précédemment, l'Institut publie, le 16 septembre 2010, un communiqué pour qu'un mécanisme de consultation soit rapidement mis en place afin que des recommandations puissent être faites quant à la formule à retenir pour le Programme de Recensement de 2016.

1. Cette enquête prend le nom de *Enquête canadienne sur l'incapacité* en 2011.

Le 10 décembre 2013, l'Institut réitère son désaccord envers l'abolition du questionnaire détaillé et obligatoire en soulignant que la diffusion des données de l'ENM a fait la preuve de l'inadéquation de cette solution. À ce jour, aucun mécanisme de consultation n'a été mis en place et SC a annoncé le 15 avril 2014 le maintien de la formule adoptée en 2011.

Parallèlement à ces prises de position publiques, les experts de l'Institut sont intervenus à maintes reprises auprès de leurs pairs à SC afin d'obtenir le plus d'outils statistiques possible ou des informations plus précises, par exemple, des mesures de précision ou des données par régions de pondération. Notamment, des représentations ont été effectuées lors de rencontres fédérales-provinciales territoriales qui ont lieu annuellement par thématique. Les rencontres sur les statistiques sociales et le Programme du Recensement ont été particulièrement ciblées. Comme nous le verrons plus loin dans le rapport, les réponses obtenues ne satisfont pas totalement les besoins exprimés par les MO.

Des travaux ont également été menés afin de soutenir le plus adéquatement possible les organisations qui éprouvent des difficultés dans l'analyse de leurs données. Entre autres, le 2 décembre 2013, l'Institut a diffusé

une [note d'information](#) sur l'état des connaissances à l'intention des utilisateurs du Québec. Cette note a été transmise par le directeur général de l'Institut à tous les sous-ministres, sous-ministres adjoints ou associés ainsi qu'aux dirigeants des organismes gouvernementaux le 6 décembre 2013. Soulignons que l'Institut n'est pas maître d'œuvre de cette enquête et n'a pas accès à l'ensemble des données et des métadonnées de l'ENM. Ainsi, l'Institut peut apporter un certain soutien aux utilisateurs, mais ne peut se substituer à SC. Il faut mentionner également que l'Institut n'a pas bénéficié de ressources additionnelles pour réaliser davantage de validation des données de l'ENM.

Enfin, le 3 février 2014, le directeur général de l'Institut a écrit de nouveau aux autorités des MO les invitant à participer à une consultation afin d'évaluer plus précisément les impacts de l'ENM, à partir des constats qu'ils ont faits à la suite de l'exploitation des données. En effet, contrairement à la consultation de 2010 où tous étaient dans l'expectative, la diffusion des données de l'ENM a permis d'évaluer plus précisément ce qu'il en était de cette enquête volontaire. La démarche de consultation entreprise par l'Institut auprès des MO est décrite plus en détail dans la section suivante.

DÉMARCHE DE CONSULTATION DE L'INSTITUT EN 2014

À la suite de la diffusion des résultats de l'ENM, l'Institut a interpellé de nouveau les MO afin de dresser, cette fois-ci, un portrait précis des problèmes rencontrés. Dans les pages qui suivent, nous décrivons plus précisément les objectifs de cette consultation et la démarche utilisée.

UN BILAN EN APPUI À LA RECHERCHE DE SOLUTIONS

Avec la sortie des résultats de l'ENM, que ce soit par la voie des produits normalisés, des compilations spéciales ou du traitement des données à partir du fichier maître dans les centres de données de recherche (CDR), les divers acteurs prennent connaissance des limites de l'enquête et certaines décisions doivent être prises, tant pour la diffusion que pour l'élaboration ou l'évaluation de politiques et de programmes.

Ce diagnostic doit permettre de dresser un portrait précis des problèmes rencontrés, tant à l'égard des données qui ont pu être utilisées que des outils statistiques disponibles, afin de prendre des décisions éclairées. Outre la plus-value que comporte la mise en commun de cette information, ce diagnostic permet de démontrer au gouvernement fédéral, une fois encore, les conséquences de cette décision pour le Québec. De plus, considérant que le gouvernement fédéral a pris la décision de maintenir l'ENM pour 2016, ce bilan sert à détecter les principales lacunes, et donc à identifier les priorités dans la recherche de solutions.

LA COLLECTE D'INFORMATION

Pour amorcer la démarche, une lettre a été expédiée à l'attention des sous-ministres, sous-ministres adjoints et associés ainsi qu'aux dirigeants des organismes gouvernementaux par le directeur général de l'Institut pour les informer de cette démarche et les inviter à communiquer avec nous. Parallèlement, à partir de la liste des MO consultés en 2010 et selon les craintes manifestées, des MO ont été identifiés comme des utilisateurs importants des données de l'ENM. Les personnes ciblées pour la collecte étaient les experts des organisations qui ont eux-mêmes traités et analysés ces données. Il ne s'agissait pas d'obtenir une position officielle des organisations, mais de partager les difficultés rencontrées par les experts de contenu. Cet objectif a été atteint. Les experts de l'Institut qui exploitent les données de l'ENM selon diverses thématiques ont aussi été consultés.

Nous avons conçu et transmis des grilles de collecte aux MO et nous leurs avons demandé de les compléter afin de s'assurer que tous les aspects soient couverts, soit autant leurs besoins, les traitements effectués que les décisions prises. Ces grilles leur permettaient également de se préparer pour les rencontres (téléphoniques ou en vis-à-vis) organisées avec chacun d'eux. Lors des rencontres, les entrevues ont été enregistrées et des résumés des rencontres ont été rédigés, en sus des grilles complétées selon le cas. Par ailleurs, les MO pouvaient nous transmettre les résultats des travaux de validation qu'ils avaient faits, tout comme des copies de leurs ententes pour des commandes spéciales. Dans certains cas, le peu d'impact pour l'organisation ne justifiait pas une rencontre. D'autres ont seulement complété la grille de collecte. À la suite de la première rencontre, des rappels ont été faits auprès des MO et des experts de l'Institut afin de nous transmettre toute autre information nouvelle et pertinente entourant l'exploitation des données de l'ENM.

Malgré l'approche standardisée auprès des MO pour recueillir l'information, les niveaux de détails sont fort divergents d'une organisation à l'autre. Ces divergences reflètent l'apport des données du Programme du Recensement pour les missions de chacune des organisations, mais également les ressources mises à leur disposition pour le traitement des données de l'ENM. Le résumé des démarches de validation présenté plus loin dans ce rapport illustre clairement cette situation.

Par ailleurs, l'Institut a mis à profit les démarches réalisées par la Table de concertation nationale en surveillance (domaine de la santé publique) pour recueillir d'autres informations. En effet, ce groupe a réalisé des ateliers permettant des échanges sur le traitement des données de l'ENM. Aussi, l'Institut a accompagné l'organisation tout au long de l'exercice de rédaction d'un guide de pratiques à l'intention des utilisateurs des données de l'ENM dans le domaine de la santé publique.

Il est important de souligner que ces rencontres ont aussi été l'occasion pour l'Institut de faire profiter les MO des bilans méthodologiques faits par l'Institut et de rappeler certains avis de SC.

Toutes les grandes missions et les thématiques couvertes par le Programme du Recensement ont pu être couvertes par la consultation. La collecte s'est déroulée de février à octobre 2014.

Les organisations suivantes ont été consultées :

- Comité de gestion de la taxe scolaire de Montréal
- Conseil du statut de la femme
- Conseil supérieur de la langue française
- Directeur général des élections
- Emploi Québec
- Ministère de la Famille
- Ministère de la Santé et des Services sociaux
- Ministère de l'Éducation, du Loisir et du Sport
- Ministère de l'Économie, de l'Innovation et des Exportations
- Ministère de l'Emploi et de la Solidarité sociale
- Ministère de l'Immigration, de la Diversité et de l'Inclusion

- Ministère des Affaires municipales et de l'Occupation du territoire
- Ministère du Travail
- Ministère des Transports
- Office des personnes handicapées du Québec
- Office québécois de la langue française
- Régie du logement
- Secrétariat aux affaires autochtones
- Société d'Habitation du Québec
- Table de concertation nationale en surveillance
- Ville de Montréal

Les experts de contenu au sein de l'Institut qui exploitent les données de l'ENM ont aussi été consultés, de même, que certains méthodologistes.

LE TRAITEMENT

L'information recueillie prend différentes formes : présentation PowerPoint, document de travail, publication, grille complétée, résumé de rencontres, messages courriels. Afin de bien couvrir l'ensemble du contenu, toute l'information a été préalablement codée et les résultats ont été déposés dans des fichiers Excel par sous-thématique. Cette liste de codes a été créée en avril 2014, soit après avoir effectué plusieurs rencontres. La participation des trois auteurs de ce rapport aux rencontres a permis la construction d'une grille plus complète.

Afin de ne pas omettre d'informations, nous avons traité les données en s'appuyant sur les méthodes associées à l'analyse de contenu tout en utilisant le logiciel Excel. Outre la diminution de risque d'omission de contenu, cette méthode a facilité la production de ce rapport.

À QUOI SERVENT LES DONNÉES PROVENANT DU QUESTIONNAIRE DÉTAILLÉ ?

Les données du questionnaire détaillé du Recensement de 2006 ont été largement exploitées par les MO. Le principal objectif de cette section consiste à faire ressortir les multiples finalités. Comme la grande majorité des organisations ont exprimé les mêmes besoins en 2011 qu'en 2006, il est plus aisé de les répertorier à partir de l'utilisation qui a été faite des données de 2006, car comme nous le verrons dans les sections suivantes, l'exploitation des données de l'ENM était à peine entamée. Pour quelques organisations, d'autres besoins se sont toutefois ajoutés en 2011. Nous en tiendrons toutefois compte.

BIEN CONNAÎTRE SA POPULATION

Connaître les caractéristiques socioéconomiques de la population du quartier, du village, de la région, ou encore immigrante ou autochtone, est le dénominateur commun de tous les autres objectifs discutés dans ce chapitre. On ne peut, en effet, élaborer des programmes publics ou encore évaluer les besoins d'une population sans la connaître. À titre d'exemple, d'importants changements peuvent survenir au sein d'une ville ou d'un quartier entre les périodes quinquennales. Des résidences privées pour personnes âgées se sont peut-être construites et des entreprises ont pu se développer ou avoir déclaré faillite, etc. Suivre l'impact des caractéristiques démographiques et socioéconomiques de leur population est crucial pour les élus municipaux, mais aussi pour les organismes communautaires, les entreprises, les conférences régionales des élus (CRE), les municipalités régionales de comté (MRC), etc. Créer des profils socioéconomiques par circonscription électorale est aussi très utile aux représentants politiques. Bien représenter sa population exige d'abord la connaissance de celle-ci.

S'ils veulent répondre à leur mission, les intervenants publics et parapublics, qu'ils soient dans le milieu de la santé et des services sociaux, du marché du travail, de la formation, de la protection sociale, du transport, etc.,

doivent avoir une excellente connaissance des caractéristiques des populations qu'ils desservent. Ignorer ces caractéristiques équivaut à demander à une entreprise privée de s'établir sans avoir préalablement effectué une étude de marché.

Les découpages géographiques sur lesquels l'analyse doit porter diffèrent selon les rôles et responsabilités des MO. Pour certaines sous-populations, comme les travailleurs autonomes, l'analyse peut être utile seulement à l'échelle québécoise.

Toutefois, la grande majorité des MO ont besoin d'information à des échelles plus fines (régions, divisions de recensement, secteurs de dénombrement, aires de diffusion, etc.). Les données peuvent être regroupées différemment selon les besoins, par exemple, par territoire de CLSC, de réseaux locaux de services (RLS), de commissions scolaires, de CRE, de centres locaux d'emploi (CLE), de MRC, de tables de quartiers, d'arrondissements, de zones de planification, d'unités de peuplement, etc.

La mission de ces MO exige souvent qu'ils dressent un portrait des populations à l'échelle de leur territoire. Par exemple, la Ville de Montréal, avec d'autres partenaires, a développé un *Atlas des quartiers* pour répondre aux besoins des différents acteurs du développement local. En effet, ce ne sont pas tous les intervenants qui sont à même d'analyser les données. Le rôle de la Ville de Montréal et de ses partenaires consiste donc à les interpréter et à les rendre plus facilement accessibles.

Soulignons que le Vérificateur général du Canada², dans son dernier rapport du printemps 2014, fait mention de l'utilité limitée des données de SC lorsqu'elles ne portent pas sur les régions qui intéressent les utilisateurs. De plus, selon lui, les besoins en données pour de petites régions géographiques et pour des sous-populations situées dans des endroits précis ne seraient pas comblés.

2. BUREAU DU VÉRIFICATEUR GÉNÉRAL DU CANADA. *Rapport du Vérificateur du Canada*. Chapitre 8. Répondre aux besoins en statistiques clés. Statistique Canada. Printemps 2014, p. 13.

Plusieurs indicateurs sont construits à partir des données très fines du Programme du Recensement. Le plus connu et le plus utilisé au Québec est l'indice de défavorisation sociale et matérielle. Cet indicateur est calculé à partir de six variables, dont trois proviennent du questionnaire court et obligatoire du Recensement (proportion de personnes seules, de personnes séparées, divorcées ou veuves et de familles monoparentales) et trois du questionnaire détaillé et obligatoire du Recensement, maintenant devenu l'ENM (revenu moyen, taux d'emploi et proportion de personnes sans diplôme d'études secondaires). L'unité géographique de base pour le calcul de cet indice est l'aire de diffusion. Les ministères de la Santé et des Services sociaux, de l'Emploi et de la Solidarité sociale, l'Agence publique de santé du Canada, l'Institut canadien sur la santé, l'Institut de la statistique du Québec, l'Institut national de santé publique, les agences régionales de santé et de services sociaux, les chercheurs universitaires et d'autres organismes provinciaux (Alberta Health Services, Cancer Care Nova Scotia, Population Data BC, Saskatoon Health region, etc.) utilisent cet indice.

Le milieu de l'éducation compte, par ailleurs, trois indices reliés à la défavorisation. D'abord, deux d'entre eux sont conçus pour l'ensemble du territoire québécois. Il s'agit du taux de faible revenu selon la mesure du seuil de faible revenu (SFR) et de l'indice de milieu socioéconomique (IMSE). Ce dernier, qui est plus largement utilisé, comprend deux variables, soit la sous-scolarisation de la mère et l'inactivité des parents. L'univers est la famille avec enfants de moins de 18 ans. Dans les deux cas, il s'agit de données provenant du questionnaire détaillé et obligatoire du Recensement et maintenant de l'ENM. L'IMSE est disponible depuis 2008 pour 3 568 unités de peuplement réparties au Québec. Le troisième indice concerne plus précisément la région de Montréal. En effet, un indicateur a été développé et une carte de défavorisation a été créée pour le Comité de gestion de la taxe scolaire de Montréal. L'indice est aussi calculé pour les familles avec enfants. On retrouve le revenu familial, la scolarité de la mère, la monoparentalité féminine et l'activité des parents. Environ 480 zones distinctes ont été créées.

Aussi, des indices de performance et de diversité industrielle sont conçus dans le domaine du développement économique. Le milieu autochtone au Québec utilise l'indice de bien-être (IBC) pour les collectivités autochtones, développé par les Affaires autochtones et

Développement du Nord du Canada. Cet indice repose sur quatre variables : le revenu, le travail, le logement et la scolarité.

Soulignons que de nombreuses enquêtes sont réalisées par SC et plusieurs peuvent répondre à certaines des interrogations par rapport au profil socioéconomique, mais aucune d'entre elles n'est conçue pour dresser un portrait à des niveaux géographiques fins. Seul le questionnaire court et obligatoire du Recensement de 2006 permet des analyses aussi fines. La taille de l'échantillon, conjuguée à l'obligation de répondre, assure une bonne qualité. Comme plusieurs intervenants l'ont souligné, ces résultats inspireraient une grande confiance, et même peut-être trop selon certains.

ÉVALUER NOS AVANCÉES... OU NOS RECULS

Les MO mettent en place des politiques et des programmes par l'intermédiaire de lois, de règlements et de plans d'action. Des orientations sont aussi annoncées et des veilles sont mises en place pour en suivre l'application. Par exemple, la politique gouvernementale *Pour que l'égalité de droit devienne une égalité de fait* a été dévoilée le 17 décembre 2006. L'analyse différenciée selon les sexes (ADS) est l'un des instruments de la gouvernance, et elle est présentée comme une approche transversale par la volonté d'intégrer progressivement le principe de l'égalité entre les femmes et les hommes à l'ensemble des décisions gouvernementales ainsi que dans celles des instances locales et régionales. De plus, le Plan d'action 2007-2010 *Pour que l'égalité de droit devienne une égalité de fait* traduit l'engagement du gouvernement à implanter l'ADS au sein de l'appareil gouvernemental. Pour y répondre, le recours aux données détaillées et ventilées selon le sexe est incontournable. Or, les données alternatives pour l'analyse différenciée selon le sexe, provenant tant des sources de données administratives que des enquêtes sociales, n'existeraient pas pour des échelles géographiques relativement fines.

L'objectif d'égalité des chances est aussi la prémisse de plusieurs plans d'action ou politiques à l'égard des populations immigrantes, des minorités visibles, des personnes ayant des incapacités, des populations autochtones et des personnes en situation de faible revenu ou d'exclusion sociale. Les données provenant du questionnaire détaillé et obligatoire servent à évaluer les tendances en matière de conditions de vie de ces

personnes. Comme il s'agit de groupes moins présents au sein de la population et également moins enclins à participer à des enquêtes volontaires, l'objectif d'obtenir un nombre de répondants suffisant et représentatif est crucial, ce à quoi répondait le questionnaire détaillé et obligatoire des recensements passés.

Avec le vieillissement de la population, les données sur les personnes âgées sont de plus en plus nécessaires. Le questionnaire détaillé et obligatoire du Recensement permet de dresser un portrait relativement précis de la situation socioéconomique des personnes âgées vivant en ménages privés, et ce, dans des petites localités. Malheureusement, l'absence de données socioéconomiques sur les personnes en logement collectif offre un portrait tronqué de la réalité des personnes âgées. Sur cet aspect, les données provenant de l'ENM n'offrent aucune solution de plus que le questionnaire détaillé et obligatoire du Recensement.

De nombreux autres enjeux économiques et sociaux, comme le développement des services de garde, le soutien aux familles, la mobilité interprovinciale, l'évolution

du marché du travail et de la formation, l'aménagement du territoire, le respect de la Charte de la langue française, etc. doivent faire l'objet d'un suivi pour en mesurer la portée pour la société québécoise. Ci-après, nous retrouvons une liste de politiques, de plans d'action et de stratégies qui reposent sur les données provenant du questionnaire détaillé et obligatoire. La liste n'est pas exhaustive, mais elle offre un bon aperçu des besoins des MO en matière de données.

Enfin, quoique les MO doivent évaluer régulièrement leurs programmes, des démarches plus systématiques sont demandées dans le cadre de la mise en place d'un processus d'analyse et de révision de programmes, comme annoncé par le ministre des Finances du Québec, lors du dernier discours du budget prononcé à l'Assemblée nationale le 4 juin 2014. Cette révision sera effectuée par la Commission sur la révision permanente des programmes et les travaux de la Commission porteront à la fois sur la pertinence de certains programmes, la réduction ou la révision de leur portée et sur leur évaluation de manière continue.

Tableau 1

Exemples de politiques, de plans d'action et de stratégies nécessitant les données du questionnaire détaillé et obligatoire du recensement ou de l'Enquête nationale auprès des ménages

Politiques sur l'immigration et l'intégration des immigrants ainsi que sur les relations interculturelles
Politiques et stratégies municipales et régionales
Politique nationale sur la ruralité 2007-2014
Politiques d'aide à la population pour sa scolarisation
Politique québécoise en matière d'officialisation linguistique, de terminologie ainsi que de francisation de l'Administration et des entreprises
Politique du Québec en matière de francophonie canadienne
Politiques et stratégies en matière d'assistance sociale
Politique <i>À part entière : pour un véritable exercice du droit à l'égalité</i>
Plan commun de surveillance de l'état de santé de la population et de ses déterminants
Plan d'action régional en santé publique et les différents programmes de santé physique et sociale de l'Agence de santé
Plans d'action locaux en santé publique et tous les programmes et activités qu'ils englobent
Plan d'action gouvernemental pour la solidarité et l'inclusion sociale 2010-2015
Plan d'action visant à favoriser la participation de tous à l'essor du Québec
Plan d'action visant à prévenir la maltraitance envers les personnes âgées
Projet de plan concerté d'occupation des territoires
Plan d'action interministériel en itinérance 2010-2013
Planification stratégique du ministère de la Famille 2008-2012 et 2012-2017
Stratégie minérale du Québec
Stratégie d'action jeunesse 2009-2014
Stratégie pour assurer l'occupation et la vitalité des territoires 2011-2016

RÉPONDRE AUX BESOINS PRÉCIS

Au-delà de la mise en place et de l'évaluation des politiques, les données servent à la gestion quotidienne des programmes ou à l'application des lois. Par exemple, dans l'application de la Loi sur les normes du travail, des travaux ont dû être menés sur les travailleurs du design à partir des données du questionnaire détaillé et obligatoire du Recensement. La situation des travailleurs autonomes et des personnes expérimentées en emploi ont aussi nécessité des analyses particulières, tout comme pour les emplois atypiques.

Par ailleurs, les programmes de formation continue ou la mise en place de programmes de formation professionnelle demandent un suivi constant des réalités du marché du travail, tant du côté du besoin des industries que du potentiel de la main-d'œuvre. Des données précises sur les professions et les industries à une échelle géographique fine sont essentielles pour réaliser ces activités.

Chaque MRC doit soumettre son schéma d'aménagement au ministère responsable. Pour ce faire, elle cherche à s'appuyer sur des données probantes. Par ailleurs, les travaux routiers sont planifiés en tenant compte du navettage de la population, données qui se retrouvent dans le questionnaire détaillé et obligatoire. Les réponses aux requêtes particulières des ministres et des députés, les demandes de citoyens et l'exercice de l'étude des crédits budgétaires doivent s'appuyer également sur des données précises.

Ces quelques exemples montrent que la gestion quotidienne des politiques et des programmes gouvernementaux repose, en partie, sur les données provenant du questionnaire détaillé et obligatoire.

ALLOUER DES RESSOURCES À L'AIDE DE DONNÉES PROBANTES

Comme il a été spécifié précédemment, les données provenant du questionnaire détaillé peuvent faciliter la gestion quotidienne des politiques et des programmes, mais peuvent également servir d'appui à l'allocation des ressources à des individus ou à des organismes. Nous trouvons plusieurs exemples d'allocation fondée sur des données probantes. Mentionnons entre autres les subventions dans les écoles défavorisées, l'ouverture de maternelle 4 ans à temps plein dans les milieux défavorisés, les subventions aux logements, les budgets aux organismes communautaires et ceux du Fonds québécois d'initiative locale, le programme de soutien aux politiques familiales municipales, les subventions aux communautés autochtones, etc.

L'utilisation des données du questionnaire détaillé et obligatoire aux fins d'allocation des ressources repose sur le principe que les données sont représentatives de la réalité. Tant les bailleurs de fonds que les demandeurs doivent s'appuyer sur les mêmes constats une fois que les méthodologies et les indicateurs retenus ont fait l'objet de consensus ou, à tout le moins, ne créent pas de controverses. D'ailleurs, la construction d'indicateurs, tels que ceux présentés précédemment, sert souvent à cette fin. Comme nous le verrons plus tard, ce type d'utilisation met au premier chef l'importance de la reconnaissance commune de la qualité des données utilisées.

QUELLES DONNÉES DU QUESTIONNAIRE COURT DU RECENSEMENT ET DE L'ENM DE 2011 ONT ÉTÉ EXPLOITÉES?

Avant de résumer les démarches réalisées par les MO dans le traitement de données, il est à propos de préciser quels produits ont été utilisés. D'abord, nous ferons un bref rappel des dates de diffusion. Ensuite, nous présenterons les diverses demandes de données selon les produits diffusés par SC ainsi que les constats qui en découlent quant à leur accès.

RAPPEL DES DATES DE DIFFUSION

Recensement

Le 8 février 2012, les premières données provenant du Recensement de la population canadienne sont rendues publiques. Celles-ci concernent les chiffres de population et des logements. Le 29 mai suivant, SC diffuse les données selon l'âge et le sexe, et le 19 septembre il diffuse les données portant sur les familles, les ménages, l'état matrimonial, le type de construction résidentielle et les logements collectifs. Le dernier volet traitant de la langue a été diffusé le 24 octobre 2012.

Les résultats des enquêtes de couverture du Recensement de 2011 ont été diffusés par SC le 26 septembre 2013. Ces enquêtes visent à établir quelle proportion de la population canadienne a été omise au Recensement et quelle part a été dénombrée plus d'une fois. À partir de ces résultats, SC produit les estimations de population basées sur les comptes du Recensement de 2011, rajustés pour tenir compte du sous-dénombrement net du recensement et des réserves indiennes partiellement dénombrées. À noter que les chiffres du recensement ne sont pas modifiés pour tenir compte du sous-dénombrement. La diffusion du 26 septembre 2013 présente des estimations démographiques révisées depuis 2001 et porte uniquement sur la population totale, à l'échelle du Canada, des provinces et territoires. Les estimations de population selon l'âge et le sexe ont été diffusées le 25 novembre 2013, tandis que les estimations infraprovinciales l'ont été le 26 février 2014.

Enquête nationale auprès des ménages

Les trois diffusions de résultats des données de l'ENM ont eu lieu au cours du printemps et de l'été 2013. Les résultats diffusés le 8 mai 2013 portent sur les peuples autochtones, l'immigration et la diversité ethnoculturelle, ceux du 26 juin 2013 concernent le travail, la scolarité, le lieu de travail, les déplacements entre le domicile et le travail, la mobilité et migration et la langue de travail. La troisième diffusion, prévue le 14 août 2013, a été reportée au 11 septembre 2013 et traite du revenu, des gains, du logement et des coûts d'habitation. Des problèmes survenus lors de l'application des routines de validation seraient à l'origine de ce report. Ces sorties ont été accompagnées de guides de référence.

SC a diffusé le 9 octobre 2013 les premiers coefficients de variation (CV) associés aux données de l'ENM à l'échelle des provinces et des territoires. Par ailleurs, comme tous les autres coordonnateurs statistiques du Canada, l'Institut a obtenu le 7 novembre 2013 des CV pour quelques variables des 34 subdivisions québécoises de recensement comptant plus de 30 000 personnes. Ces CV peuvent être obtenus sur demande auprès de l'Institut. À noter qu'aucun autre CV ne sera diffusé avec la parution du rapport technique de l'ENM sur l'échantillonnage et la pondération prévue au début 2015.

L'accès au fichier complet de l'ENM dans les centres de données de recherche a été rendu possible à la fin du printemps 2014. Quant au fichier de microdonnées à grande diffusion (FMGD), qui comprend un échantillon de 2,7 % de la population canadienne, il a été diffusé le 29 juillet 2014. Le fichier hiérarchique FMGD a été pour sa part rendu public le 9 décembre 2014.

L'OUTIL DE DÉPART : LES PRODUITS NORMALISÉS

Les intervenants rencontrés dans le cadre de cette consultation ont consulté, soit la série sur le profil de l'ENM, les tableaux de données ou les perspectives géographiques. Ce constat n'est pas surprenant puisque ces intervenants devaient exploiter directement les données de l'ENM pour qu'ils puissent être rencontrés dans le processus de consultation.

Pour ces derniers, la consultation de ces données est incontournable et sert de référence à toutes exploitations additionnelles, par exemple, pour des commandes spéciales. On a d'ailleurs apprécié le fait que SC ait élargi son éventail de production diffusé gratuitement. De plus, la consultation des données, par l'intermédiaire de l'application Beyond 20/20 est très appréciée en raison de la convivialité de l'outil.

Les nouvelles règles de diffusion pour les produits normalisés de l'ENM qui font en sorte qu'aucune donnée n'est diffusée pour des territoires, dont le taux global de non-réponse (TGN) s'élève à 50 % et plus, ont entraîné, à l'échelle du Québec, la suppression de 183 SDR (correspondant à des municipalités). Ce nombre aurait cependant été beaucoup plus important si le critère de suppression utilisé lors du Recensement de 2006 avait été appliqué (TGN \geq 25 %). Par ailleurs, en raison des règles de confidentialité plus strictes, on retrouve plusieurs cellules mises à 0 dans les tableaux. Comme nous le verrons plus tard, cette situation complique l'analyse des données. Il en est de même pour le traitement croisé des variables.

Quoique ce changement ait été annoncé avant la mise en place de l'ENM, on déplore la perte de données sur le travail non rémunéré parmi les thèmes abordés avec l'ENM. Cette thématique n'est plus traitée, alors qu'elle l'était dans le questionnaire détaillé et obligatoire du Recensement de 2006. Aussi, pour le suivi de la langue, on déplore que le questionnaire de l'ENM ne comprenne aucune question sur la langue d'usage public.

DES COMMANDES NOMBREUSES ET VARIÉES

Pour répondre à des besoins spécifiques, la majorité des MO ont acheté des compilations spéciales à SC. La production de données croisées, ainsi que des données compilées selon des univers géographiques précis sont à l'origine de la plupart de celles-ci. Soulignons que certaines demandes de compilation de données du Recensement ont été faites avec l'objectif de valider la cohérence entre les données démographiques du Recensement et celles de l'ENM.

Au moment de la consultation, certaines commandes de données n'avaient pas encore été effectuées, car on s'interrogeait sur le contenu qui pouvait être obtenu. En effet, lorsque vient le temps de commander des données, il est parfois difficile de prévoir leur niveau de précision. On peut jouer de prudence en augmentant le plus possible la taille des différentes populations à comparer, mais corollairement on peut ne pas obtenir ce qu'on désire. Certains ont joué de prudence en demandant le plus de sous-totaux possible (par exemple, en demandant les données détaillées pour tous les pays de citoyenneté, mais en même temps, des sous-totaux par continent), tandis que d'autres ont fait le choix d'explorer directement le fichier de microdonnées dans les centres de données de recherche (CDR).

La modification dans les classifications des secteurs d'industrie (SCIAN) et des professions (CNP) ainsi que pour les programmes d'enseignement (CPE) entre les années 2006 et 2011 rend plus difficiles les comparaisons dans le temps. Il en est de même pour les changements dans les produits normalisés concernant la mesure du faible revenu. Des compilations spéciales sont alors nécessaires. Certaines comparaisons peuvent se faire à partir des produits normalisés, mais cela demande tout de même une attention particulière afin de pouvoir comparer ce qui peut l'être.

En ce qui a trait au nombre de compilations spéciales, certains MO en ont moins demandé, puisque les diffusions avec les produits normalisés comblent davantage leurs besoins qu'en 2006. D'autres, au contraire, en raison des coûts moins élevés qu'en 2006 pour les compilations spéciales, en ont profité pour en commander davantage. D'ailleurs, ce dernier point concernant la diminution des coûts pour des compilations spéciales, en raison des changements dans la méthode de facturation, a été apprécié par plusieurs MO.

Pour la plupart des MO, les données obtenues répondent à leurs attentes, et les délais sont perçus comme raisonnables. Comme l'a spécifié une organisation, plus la commande est précise, c'est-à-dire qu'elle inclut une liste complète de variables et de catégories requises, plus elle a des chances de répondre aux attentes. Certains ont néanmoins eu des surprises, en raison d'une nouvelle règle de confidentialité qui remplace par 0 l'estimation pour les variables qualitatives lorsque le nombre d'enregistrements est inférieur à 4 dans une cellule³. Avec cette nouvelle règle, il y a plus de valeurs supprimées que dans les estimations provenant du questionnaire détaillé des recensements antérieurs, où cette règle de moins de 4 ne s'appliquait qu'aux estimations quantitatives. De plus, avec l'ENM, en remplaçant les valeurs supprimées par des 0, on ne peut plus distinguer les 0 venant de la méthode de suppression des vrais 0. L'interprétation des données devient très difficile. La règle de l'arrondissement aléatoire pose aussi un défi, mais puisqu'elle a été utilisée dans les recensements précédents, elle n'a suscité aucun commentaire particulier.

Dans le cas d'une commande importante (un consortium d'achat), la commande a été passée tôt, soit avant même que les règles de confidentialité appliquées aux données de l'ENM ne soient connues. Plusieurs éléments ont donc dû être ajustés pour rendre la livraison conforme à ce qui était indiqué dans l'entente. Il semble que la majorité des problèmes que cette situation a entraînés aient été résolus. Toutefois, en raison de nouvelles règles de confidentialité, le produit livré ne répond pas complètement aux attentes. D'ailleurs, d'autres commandes spéciales ont dû être faites à partir des données du questionnaire court et obligatoire pour pallier certaines difficultés.

Les MO prennent de plus en plus conscience de la portée des ententes avec une licence ouverte qui permet un partage plus facile de l'information entre les utilisateurs de ces données. Cette nouvelle disposition est très appréciée, et si elle ne semble pas pour le moment avoir modifié les façons de faire des organisations, il faut s'attendre à des changements de pratique ultérieurement.

L'ACCÈS AUX DONNÉES DANS LES CENTRES DE DONNÉES DE RECHERCHE (CDR), UNE AVENUE SURTOUT EXPLORATOIRE

Auparavant, le recours aux données du fichier maître du Recensement dans les CDR servait surtout dans le cadre de projets de recherche. Pour des données descriptives, l'achat de données par l'intermédiaire de commandes spéciales constitue plutôt la norme. Mais avec l'abolition du questionnaire détaillé et obligatoire, les utilisateurs peuvent plus difficilement apprécier *a priori* la qualité des données qu'ils auront par territoire géographique, du moins sur une base historique. En effet, les nouvelles règles de confidentialité, conjuguées au fait que les régions de pondération⁴ utilisées par SC peuvent différer des univers géographiques étudiées, amènent certains MO à explorer les données des fichiers maîtres.

Deux scénarios ont été observés. Dans certains cas, on avait anticipé que des problèmes surgiraient dans les petites unités géographiques. On voulait vérifier la fiabilité des données en dehors des grands centres urbains pour des unités géographiques particulières avant de demander des compilations spéciales, notamment en examinant les TGN par aire de diffusion. Dans d'autres cas, c'est à la suite de la réception de données de compilations spéciales et la difficulté de les utiliser que des démarches ont été entreprises pour l'accès aux données du fichier maître. Dans ces cas-ci, c'est la cohérence que l'on voulait examiner entre les données démographiques du Recensement de 2011 et celles de l'ENM pour des univers géographiques qui leur sont propres.

Une demande a aussi été faite par l'Institut à des fins de validation pour les données du revenu. L'objectif est de comparer les données provenant des données fiscales de Revenu Québec par municipalités avec les données de l'ENM, et ce, pour les mêmes univers. Cette validation tient aussi compte des divers concepts de revenu. Les résultats de cette démarche sont présentés plus loin dans ce rapport.

-
3. L'estimation de la cellule est supprimée si le nombre d'enregistrements comprenant l'attribut ou la combinaison d'attributs représentée par la cellule (non arrondie et non pondérée) est inférieur à 4.
 4. SC pondère les données de l'ENM afin de réduire ou éliminer les différences avec les comptes de population du Recensement selon les divisions géographiques de pondération.

En ce qui a trait à la mise à jour d'une étude sur la fécondité selon la langue maternelle, on attend d'avoir plus d'indications sur la qualité des données avant d'aller de l'avant. Par ailleurs, une organisation nous a clairement indiqué que l'option d'accéder au fichier maître n'est pas possible, d'une part, en raison des délais importants pour l'accès aux données dans les CDR, et d'autre part, par le manque de ressources internes ayant les habilités nécessaires pour réaliser ces travaux.

Dans le cas des enquêtes postcensitaires, soit l'une sur l'incapacité et l'autre sur les Premières nations, dont les échantillons sont tirés de l'ENM, les demandes d'accès sont plus rares. Parmi les utilisateurs rencontrés, un seul a recours aux données du fichier maître de l'*Enquête canadienne sur l'incapacité* (ECI), l'exploitation des questions 7 et 8 de l'ENM n'étant pas considérée comme une avenue fiable pour mesurer l'incapacité. L'ECI permet de réduire la forte proportion de faux positifs, et évidemment de réaliser une analyse plus spécifique de ce thème. Étant donné les analyses pointues que l'on voulait mener à partir de cette enquête, il apparaissait problématique de s'en remettre aux compilations spéciales. Aucun des MO rencontrés n'a exploité directement les données de l'*Enquête auprès des peuples autochtones* de 2012 (EAPA). Les données publiées par le ministère fédéral des Affaires autochtones et Développement du Nord Canada semblent constituer la principale référence.

Enfin, le FMGD n'a été utilisé par aucun des utilisateurs rencontrés au cours de notre période de consultation, puisque ce fichier a été diffusé tardivement, soit le 29 juillet 2014. Toutefois, selon les principaux besoins exprimés par les MO, les données provenant du FMGD sont peu utiles, en raison de la faible taille d'échantillon de celui-ci. Ce fichier est surtout conçu pour des projets de recherche.

DES OUTILS INSUFFISANTS POUR ÉVALUER LA QUALITÉ DES ESTIMATIONS DE L'ENM

Avant d'utiliser les données du questionnaire court et obligatoire du Recensement ou de l'ENM, les chercheurs doivent se poser de nombreuses questions. SC procède de son côté à plusieurs validations avant la sortie des données et des guides présentant des mises en garde aux utilisateurs. Des présentations ont également eu lieu afin de donner un aperçu du design de l'ENM et de sa qualité. De façon générale, ces guides et présentations sont appréciés. Mais certains MO sont restés sur leur faim. Selon certaines thématiques, les guides sont incomplets (ex. : travail), alors que d'autres sont plus satisfaisants (ex. : scolarité). De plus, la portée pratique de ces documents en regard de l'évaluation de la qualité des données apparaît limitée. On aurait aussi souhaité que des mises en garde spécifiques soient fournies, tant dans les produits normalisés qu'avec la livraison de compilations spéciales. Par exemple, une organisation aurait aimé savoir si les travailleurs autonomes peu scolarisés des MRC rurales ont moins bien répondu au questionnaire d'enquête que les travailleurs autonomes plus scolarisés. En somme, plusieurs MO déplorent d'avoir à prendre des décisions sans avoir tous les outils pour le faire. On reproche d'une certaine façon à SC de faire porter le poids des décisions sur les utilisateurs.

De façon plus pointue, les utilisateurs ont mis en lumière leurs besoins en matière d'outils de mesure de la qualité en lien avec les risques de biais, la précision et la cohérence. Également, le bris de série avec le passage d'une enquête obligatoire à une enquête volontaire, les changements reliés aux règles de confidentialité ainsi que les modifications liées aux classifications ont soulevé plusieurs questions dans la plupart des MO. Ces aspects sont discutés ci-après.

BIAIS

Toute enquête comporte des risques de biais, puisque les non-répondants ont tendance à avoir des caractéristiques différentes des répondants. Avec le questionnaire détaillé et obligatoire des recensements antérieurs, un ménage sur cinq était échantillonné et celui-ci était tenu de répondre au questionnaire. Malgré cette obligation, un certain nombre de ménages ne répondait pas. En 2006, le taux de réponse des ménages s'est élevé à environ 94 %. Or, avec l'ENM de 2011, on obtient plutôt un taux d'environ 72 % pour les données non pondérées et 81 % pour les taux pondérés. Pour obtenir un portrait de la population totale, SC procède à un ajustement des poids des ménages répondants afin qu'ils représentent au mieux les ménages non répondants. Sur ce, certains MO auraient souhaité que SC mette plus d'effort pour l'imputation, plutôt que de s'en remettre à la pondération. Mentionnons que SC a tenté d'utiliser l'imputation pour corriger la non-réponse totale, mais des problèmes de combinaisons improbables trop fréquentes entre les variables l'ont amené à changer de stratégie et à utiliser l'ajustement des poids pour la non-réponse.

Par ailleurs, que ce soit avec le questionnaire détaillé des recensements passés ou avec l'ENM, les répondants peuvent décider ou omettre de répondre à certaines questions. C'est généralement le cas lorsque les questions se trouvent à la fin d'un questionnaire ou lorsque la question est jugée trop sensible par le répondant. On parle alors de non-réponse partielle. Ce sont les questions relatives à la scolarité, aux caractéristiques du logement et au revenu qui sont les plus touchées par ce phénomène dans l'ENM. Le même constat a été observé avec le questionnaire détaillé et obligatoire du Recensement de 2006. Par contre, les taux de non-réponse partielle sont plus élevés avec l'ENM que pour le questionnaire détaillé et obligatoire du Recensement de 2006. Par exemple, pour les questions de logement, ces taux sont en moyenne deux fois plus élevés. Bien que l'imputation réalisée par SC vienne atténuer l'effet de la non-réponse partielle, la proportion de ménages non touchés par aucune forme de non-réponse, totale ou partielle, est considérablement plus faible avec l'ENM que lors des recensements passés. La qualité de l'imputation ne peut égaler la réponse exacte des ménages.

Pour évaluer les risques de biais dus à la non-réponse, SC a calculé un taux global de non-réponse (TGN) par territoire géographique. Ce TGN combine la non-réponse des ménages et la non-réponse aux questions. SC a pris la décision de ne pas diffuser toutes les estimations géographiques dans ses produits normalisés (c.-à-d. grand public) lorsque le TGN est égal ou supérieur à 50 %. Ce critère est beaucoup moins restrictif que pour le questionnaire détaillé de 2006, où la suppression était faite pour un TGN de 25 % et plus. Si cet assouplissement du critère a été apprécié par les utilisateurs de ces données, l'évaluation du niveau de biais jugé « acceptable » leur incombe davantage. La situation est encore plus critique pour les utilisateurs de données de compilations spéciales pour lesquelles les données sont fournies pour des TGN plus grands que 50 %. Par exemple, environ 25 % des municipalités au Québec ont des TGN plus élevés que 50 %, et lorsqu'on descend à des unités géographiques plus petites, cette proportion augmente. Un MO a précisé qu'il aurait souhaité obtenir les TGN comparés de 2006 et 2011 avec les produits normalisés. Un autre souligne que la présence de deux TGN pour les déplacements domicile-travail, soit un pour le lieu de travail et l'autre pour le lieu de résidence, rend l'analyse plus complexe.

Comme l'ont souligné certains utilisateurs, le TGN donne une information sur le risque de biais, mais pas sur le biais lui-même. Par exemple, un TGN de 50 % dans un territoire homogène par rapport à un TGN de même valeur dans un territoire ayant une population diversifiée ne signifie pas la même chose. Le risque de biais devrait être plus élevé dans un territoire ayant une population diversifiée. Tenter de corriger les biais n'est pas un exercice trivial, d'autant plus que l'on risque d'introduire d'autres biais. Si quelques utilisateurs y ont pensé, de telles stratégies n'ont pas été retenues. Également, certains auraient souhaité pouvoir catégoriser les taux pour faciliter la diffusion, comme on le fait pour des coefficients de variation. En raison du caractère subjectif d'une telle catégorisation, la majorité des organisations ont cependant retenu les critères de 25 % (critère de diffusion de SC pour le questionnaire détaillé de 2006) et de 50 % (critère de diffusion de SC pour l'ENM).

Comme il a été spécifié précédemment, les TGN ont été calculés pour des territoires géographiques seulement. Pour les sous-populations, les utilisateurs doivent s'en remettre au taux d'imputation. Les utilisateurs pour des domaines spécifiques d'étude (ex. : une profession ou une industrie en particulier, les autochtones, les immigrants, etc.) se sentent donc encore plus démunis pour évaluer les biais possibles.

Des difficultés liées au calcul des TGN ont aussi été soulevées, car certains résultats étaient difficiles à interpréter. En effet, on pouvait obtenir un TGN pour l'ensemble d'un territoire plus faible que tous les TGN des aires de diffusion qui composent ce territoire, ce qui va à contre-sens. Par exemple, le TGN pour la ville de Cap-Chat est de 25,5 %, alors que les TGN pour les aires de diffusion (AD) qui composent cette ville dépassent tous 40 %. Cette situation serait due aux différences dans les poids utilisés pour le calcul du TGN, qui diffèrent selon le nombre de ménages du sous-échantillon de la deuxième phase d'échantillonnage de l'ENM. Dans l'exemple précédent, ce sont les poids de la 1^{re} phase qui ont été utilisés pour les AD et les poids de la 2^e phase pour Cap-Chat. Ces différences créent de la confusion (d'autant plus que l'on ne sait pas quelle pondération a été utilisée lorsque l'on regarde un TGN pour une petite unité géographique), ce qui rend plus difficiles les analyses sur les risques de biais. Dans une telle situation, l'utilisation des TGN les plus élevés pour évaluer les risques de biais constitue une approche conservatrice.

PRÉCISION

La précision, par ailleurs, rend compte de la variabilité d'une mesure entre les ménages ou les individus et est associée à l'échantillonnage. L'erreur d'échantillonnage est généralement aléatoire et tient compte de la taille de la population et de l'échantillon ainsi que de la complexité du plan de sondage. Différentes mesures de précision, notamment l'erreur type, les intervalles de confiance et le coefficient de variation (CV), dérivés de la variance, sont utilisés. La plus courante est le CV, qui est le rapport entre l'erreur type de l'estimation et l'estimation elle-même.

Avec l'ENM, SC a diffusé des CV à l'échelle des provinces, pour 11 thématiques⁵. SC a aussi remis à l'Institut les CV pour les municipalités de 30 000 personnes et plus, pour ces mêmes thématiques. L'Institut, à titre de coordonnateur statistique pour le Québec, s'est engagé à distribuer ceux-ci à tous ceux lui faisant la demande. À l'échelle provinciale, les CV apparaissent semblables entre 2006 et 2011, mais il y aurait plus de variabilité entre les CV cette fois-ci. À certains endroits, la hausse de la taille de l'échantillon (soit un ménage sur trois au lieu d'un ménage sur cinq) n'a pu compenser la baisse du taux de réponse.

Les besoins en termes de mesures de précision vont bien au-delà de ceux fournis. Or, aucun autre CV ne sera fourni par SC, ni aucun outil permettant de les calculer. En effet, malgré les demandes répétées, SC ne fournira pas dans le rapport technique les facteurs d'ajustement de l'erreur type (comme l'effet de plan de sondage) qui permettraient par le passé de calculer des CV approximatifs. Deux raisons ont été évoquées par SC pour ne pas donner suite à ces demandes. La grande complexité du plan de sondage, notamment avec une 2^e phase d'échantillonnage, rend difficile le calcul de CV ou d'effet de plan. Le calcul de ces mesures nécessiterait un investissement important en temps, et SC n'a pas les ressources disponibles. Les raisons évoquées laissent cependant perplexes plusieurs utilisateurs d'autant plus que l'ENM sera repris en 2016. Des doutes ont été émis sur ces raisons, car certains croient que SC travaille tout de même sur ces questions. Des interrogations ont aussi été formulées en raison de la disponibilité d'un outil et d'une méthode approximative dans le FMGD pour calculer les CV, ainsi que dans l'ECl produite à partir d'un sous-échantillon de l'ENM, alors qu'aucun CV n'est disponible dans le fichier maître de l'ENM (accessible dans les CDR). SC a aussi clairement stipulé qu'aucun taux de réponse brut ne sera fourni au-delà de ceux publiés sur le site Web de SC, soit pour les unités géographiques, provinces et territoires, et pour les subdivisions de recensement. De ce fait, on ne peut estimer le nombre de répondants dans les unités géographiques.

L'Institut a procédé à des calculs d'estimation des effets de plan pour les municipalités de 30 000 personnes et plus, mais en raison de la variabilité observée dans ces effets de plan, vraisemblablement causée par la complexité du plan de sondage et la variabilité dans les taux de réponse, il n'est pas possible de généraliser les résultats à d'autres divisions géographiques ou à d'autres variables. Il faut rappeler que l'Institut n'est pas maître d'œuvre de cette enquête et qu'il n'a donc pas accès à toutes les données et métadonnées liées à l'ENM pour développer une stratégie alternative.

L'absence de mesures sur la précision, notamment les CV, a été déplorée par la grande majorité des MO. Cette situation a pour effet de rendre difficile, voire impossible, l'évaluation de la précision de l'estimation même de façon sommaire, et ce, particulièrement pour des unités géographiques plus fines ou encore des sous-populations. En l'absence de mesures de précision, il est également impossible de mesurer des différences. Le taux de faible revenu est-il plus faible dans une ville que dans une autre? Pour une unité géographique donnée, le taux d'emploi est-il plus élevé ou plus faible chez les personnes de 35-44 ans par rapport à celles de 25-34 ans? Quelle est la profession dominante dans un secteur industriel spécifique? Cette absence de mesures de précision soulève bien d'autres interrogations. Bref, toute conclusion est difficile à tirer.

La notification de SC dans un de ces guides montre bien que les utilisateurs sont laissés à eux-mêmes pour porter un jugement sur la précision de la donnée.

« Étant donné que certains utilisateurs peuvent souhaiter tolérer les plus grosses erreurs avant de pouvoir obtenir des données sur des petits groupes de la population, les suppressions pour des motifs de qualité des données ne sont pas appliquées à la plupart des produits normalisés, même lorsque les estimations sont sujettes à une forte variabilité de l'échantillonnage. Par conséquent, les utilisateurs doivent toujours faire attention aux statistiques basées sur des cellules affichant de faibles estimations⁶ ».

5. Coûts d'habitation (3), immigration et citoyenneté (10), logement (6), minorités visibles (2), mobilité (2), peuples autochtones (2), revenu de la famille (4), revenu des particuliers (4), revenu du ménage (3), scolarité (3) et travail (7). Les chiffres entre parenthèses indiquent le nombre de variables dans chacune des thématiques.

6. STATISTIQUE CANADA (2013), Document analytique, *Guide de référence sur le revenu*, Catalogue N° 99-014-X2011006, 2013, p. 23.

COHÉRENCE

La pondération de l'ENM a été faite par SC de manière à réduire ou à éliminer les différences avec les comptes de population obtenus avec le questionnaire court et obligatoire du Recensement. Cette étape de la pondération, appelée calage, s'est faite à l'échelle d'unités géographiques, appelées régions de pondération. Celles-ci comptent en moyenne 2 300 logements (5 600 personnes). Toutefois, des différences entre les estimations de l'ENM et les comptes provenant du questionnaire court et obligatoire peuvent être observées, même pour des territoires respectant ces tailles. La prudence est donc de mise pour des unités géographiques en deçà de cette taille. SC a émis cette recommandation dans son [Guide de l'utilisateur](#) et suggère donc de comparer les chiffres du Recensement de 2011 et les estimations de l'ENM pour les caractéristiques communes afin de juger de la cohérence. En effet, le manque de concordance permet d'entrevoir un risque de biais pour une unité géographique donnée.

Comme nous le verrons plus tard, plusieurs MO ont fait des analyses de la cohérence des données du Recensement et de l'ENM pour les territoires qui les concernent. Il en est ressorti, entre autres, des différences quelquefois importantes entre les données du Recensement et les estimations de l'ENM lorsque les unités géographiques utilisées ne semblaient pas correspondre aux régions de pondération de SC. Par exemple, à l'Île d'entrée aux îles de la Madeleine, le compte de la population est moins élevé que prévu, mais l'est davantage dans l'autre section de l'Île. Certains MO auraient souhaité obtenir les données démographiques de base par régions de pondération (5 600 habitants et plus), ce qui leur aurait permis de mieux comprendre les différences observées et de redéfinir au besoin leur univers géographique d'analyse pour améliorer la cohérence, et par conséquent la qualité de leurs données. Des demandes ont été transmises pour obtenir des estimations par régions de pondération (ou calage), et également obtenir des tables de correspondance entre ces régions de pondération et les unités géographiques définies par les utilisateurs (ex. : aire de diffusion). SC reconnaît que le traitement statistique au niveau des régions de pondération assure la cohérence entre l'ENM et les chiffres du Recensement de 2011, lorsque les contraintes de calage sont respectées. Cependant, pour des raisons de contrôle de la divulgation, SC ne peut rendre disponibles

ces données. En effet, les régions de pondération sont définies selon des critères spécifiques reliés à la géographie, aux logements et à d'autres informations auxiliaires. Dans certains cas, cela peut permettre la divulgation de données personnelles, entraînant potentiellement une brèche à l'égard de la confidentialité.

Quoique les motifs reliés à la confidentialité ne puissent être contestés, les utilisateurs aimeraient toutefois obtenir des solutions alternatives.

COMPARABILITÉ

Le passage d'une enquête obligatoire à une enquête volontaire introduit ce qu'on appelle communément un bris de série. Dans de telles circonstances, les données de 2011 et de 2006 ne sont plus directement comparables. Trois aspects principaux peuvent limiter cette comparabilité : le risque accru de biais, comme discuté précédemment, des différences entre les populations cibles (logements collectifs non institutionnels inclus dans le Recensement de 2006, mais pas dans l'ENM) et le positionnement et le contenu de certaines questions (identité autochtone, revenus, etc.). À ceux-ci s'ajoutent certains changements de classification. SC recommande la plus grande prudence lors de comparaison, tout spécialement lorsque l'analyse porte sur des échelles géographiques fines. D'ailleurs, sur son site Web, les données du questionnaire détaillé des recensements antérieurs et celles de l'ENM sont sur des pages distinctes.

Comme nous le verrons dans la section suivante, les MO doivent procéder à plusieurs démarches de validation avant de comparer les données de 2006 à celles de 2011. Et une fois que les tendances ont été examinées et comparées à d'autres données disponibles, la réponse n'est pas si simple. Certains utilisateurs se questionnent lorsqu'ils observent une différence, à savoir si celle-ci représente un véritable changement ou si elle est plutôt liée à un biais ou à une erreur d'échantillonnage. Devant l'absence de différence entre les données, les utilisateurs se demandent s'il ne devrait pas y en avoir une. Dit autrement, on peut rejeter des données qui apparaissent *a priori* aberrantes ou accepter une tendance alors qu'il n'y en a pas. Les utilisateurs doivent donc s'en remettre à la méthode du vraisemblable ou du gros bon sens. Des décisions doivent donc être prises par les analystes.

Certains auraient d'ailleurs préféré des consignes plus claires de la part de SC, pour éviter une fois de plus de porter le fardeau de la preuve, alors que pour ceux qui connaissent bien leur domaine, la marge de manœuvre offerte par SC leur convenait davantage.

Comme c'est le cas pour tout changement majeur entre deux enquêtes, le bris de série ne peut être contourné. Avec l'introduction de l'ENM en 2011, et sa reconduite en 2016, on est davantage en mode résignation. La perte en termes d'analyse de tendances est reconnue par tous les utilisateurs, sans pour autant être appréciée.

Les changements de classification conduisent à certaines difficultés, en sus des problèmes de comparabilité liés au passage d'une enquête obligatoire à une enquête volontaire. Ainsi, les données sur l'industrie sont classées selon le Système de classification des industries de l'Amérique du Nord (SCIAN) 2007. La révision est cependant mineure par rapport à celle de 2002. Par contre, il en est autrement pour les données sur les professions qui sont classées en fonction de la Classification nationale des professions (CNP) 2011. En raison des changements considérables, une table de correspondance CNP-S 2006 et CNP-2011 a été produite pour la conversion. Toutefois, selon une organisation qui utilise grandement ces données, certains codes pour la correspondance avec la CNP de 2006 sont manquants. De plus, comme la CNP retenue pour l'*Enquête sur la population active* (EPA) est celle de 2006, et qu'elle ne sera changée qu'en janvier 2015, cela complexifie grandement les analyses. Aussi, pour avoir les deux classifications, il faut avoir deux sorties de tableaux de l'ENM, et il manque des professions si on reconvertit pour la CNP de 2006. Donc, selon les besoins, le choix s'est soit porté sur 2011, avec les bris de série que cela implique, soit 2006. Également, dans le domaine de la scolarité, l'ENM repose sur la Classification des programmes d'enseignement (CEP) de 2011, alors que le Recensement de 2006 s'appuyait sur la CEP de 2001.

Dans le cas des données sur le faible revenu, SC a procédé à des comparaisons entre les taux de faible revenu selon la mesure des seuils de faible revenu (SFR) entre les données du Recensement de 2006 et l'ENM de 2011 et celles de l'*Enquête de la dynamique du travail et du revenu* (EDTR) de 2006 et 2011. La conclusion est

qu'en raison des différences observées et du fait que les estimations ne soient pas directement comparables aux autres enquêtes ménages comme l'EDTR, les estimations des ménages à faible revenu selon la SFR ne sont pas offertes dans les produits normalisés. Pour les obtenir, des compilations spéciales doivent être faites. Même si cette décision peut être comprise, pour les MO qui utilisaient le seuil SFR pour estimer la population à faible revenu depuis de nombreuses années, ils doivent maintenant assumer le coût des compilations spéciales et les délais que cela occasionne.

SC a rédigé un guide de référence sur la langue, et dans celui-ci, il est énoncé que certains changements entre 2006 et 2011 semblent émaner des modifications apportées au positionnement et au contexte des questions linguistiques. Une organisation a demandé un avis plus précis à SC. La personne contactée aurait affirmé que ce changement n'a eu que peu ou pas d'impact sur les réponses. Toutefois, selon cette organisation, ce constat s'appuierait sur des données qui ne leur sont pas facilement accessibles. Il serait donc difficile de contextualiser ces affirmations avec des analyses supplémentaires et ajustées à leurs besoins. Un expert indépendant dans le domaine de la langue a d'ailleurs été consulté par cette organisation, et selon ce dernier, les changements liés à l'emplacement des questions auraient eu des impacts. Soulignons qu'une autre organisation considère qu'elle ne peut se positionner clairement relativement à ces différences, d'où le principe de reprendre les mises en garde de SC pour leur traitement des données. Dans le domaine plus spécifique de la langue du travail, que l'on retrouve uniquement dans l'ENM, les données ne sont disponibles que depuis 2001. Il est alors encore plus difficile d'évaluer les tendances avec le bris de série de 2011.

RÈGLES DE CONFIDENTIALITÉ

Comme nous en avons parlé précédemment, les règles de confidentialité ont été modifiées avec l'ENM. D'abord, une nouvelle règle est appliquée pour les variables qualitatives. Pour un nombre d'enregistrements inférieur à 4, les données d'une cellule sont remplacées par un 0, et non un signe distinctif comme X. Comme il a été précisé précédemment, il est donc impossible de distinguer les vrais 0 des faux 0. Au début des travaux d'analyse, peu de MO avaient pris conscience de cette nouvelle règle, dans la mesure où les premiers exercices s'appliquaient généralement à de grands ensembles. D'autres MO n'ont tout simplement pas pris acte de cette situation. Mais lorsque certains travaux sont devenus plus pointus, il est apparu qu'il n'était pas toujours possible de produire des sous-totaux, des matrices de navettage ou faire de la standardisation directe. L'Institut est intervenu auprès de SC pour connaître les raisons de ce choix, sans toutefois remettre en cause l'objectif du respect de la confidentialité. La réponse fournie par SC indique que si les vrais 0 étaient distingués des faux 0, le risque serait trop élevé de distinguer des valeurs confidentielles. Pour ce faire, il leur aurait fallu masquer davantage de données pour assurer la confidentialité. Les méthodes de pondération pour l'ENM augmentent, en effet, les risques de « bris » de la confidentialité par rapport aux recensements précédents puisque certains individus peuvent se retrouver avec de très grands poids ce qui facilite leur identification dans les tableaux, d'où le resserrement des règles. Toutefois, la question demeure entière sur le choix de la méthode utilisée par SC pour assurer la confidentialité.

Comme porte de sortie, l'Institut a conseillé aux MO d'augmenter la demande pour des sous-totaux lors de compilations spéciales ou encore d'explorer directement le fichier maître de l'ENM dans les CDR. Par ailleurs, certains MO ont vu une opportunité de pouvoir améliorer l'analyse de leurs données en demandant à SC de fournir les données par régions de pondération. Toutefois, comme nous l'avons précisé précédemment dans la discussion sur la cohérence, cette possibilité a été écartée par SC en raison des risques de divulgation accrus par la diffusion des limites de ces divisions géographiques.

SOUTIEN OBTENU DE LA PART DE SC ET DE L'INSTITUT

En sus des questions soulevées ci-haut, les MO demandent à l'occasion d'être accompagnés par des experts en méthodologie de SC pour l'analyse spécifique de données ou pour leur interprétation. Nous avons cherché à obtenir des MO la liste des interventions effectuées auprès de SC, mais nous croyons que plusieurs ont été faites après nos rencontres de consultation. Donc, la synthèse qui suit ne couvre probablement pas toutes les demandes acheminées à SC de la part de MO, ni les réponses subséquentes qui ont été fournies.

Comme il a été précisé au début de cette section, les analystes des MO prennent d'abord possession des informations transmises par SC dans ses divers guides d'accompagnement. Alors que pour certains MO, ces informations sont suffisantes, des questions demeurent sans réponses pour d'autres. Certains ont préalablement approché l'Institut, mais pour des demandes spécifiques, il leur a été suggéré de s'adresser directement à SC qui, étant responsable de cette enquête, est mieux placé pour répondre à ces questions. Une organisation souligne que lors de demande d'avis, SC se montre très prudent, ou pour la citer « ne se mouille pas », et semble retenir beaucoup d'information.

Dans une région moins densément peuplée, on déplore de ne pas avoir de formation personnalisée de la part de SC afin de savoir comment expliquer les écarts entre 2006 et 2011 aux gens de leur milieu. Autrement dit, comment mettre en garde les utilisateurs et leur fournir des explications si eux-mêmes ne comprennent pas.

La plupart des MO attendaient toujours d'avoir terminé leurs travaux d'analyse avant d'acheminer leurs questions à SC. L'un d'entre eux a même précisé qu'il soumettrait son analyse à SC, surtout pour les comparaisons, et leur « plan » est de prendre leurs décisions de publication une fois cette consultation faite. Ce MO ne veut absolument pas porter le fardeau de la preuve de la qualité des données.

Soulignons que dans son rapport de 2014, le Vérificateur général du Canada rend compte des consultations faites auprès d'utilisateur de données de l'ENM. Leur constat rejoint le nôtre. En voici un extrait⁷ :

Statistique Canada a fourni beaucoup de directives aux utilisateurs des données de l'*Enquête nationale auprès des ménages*. Cependant, les utilisateurs de données concernant les petites régions géographiques et sous-populations nous ont fait savoir que, malgré l'utilité des directives générales offertes à ce jour sur les données de cette enquête, ils ne savent pas toujours quelles données relatives à des endroits précis ou à des sous-populations précises sont fiables et comment ces données doivent être interprétées. Par conséquent, certains utilisateurs potentiels pourraient décider de ne pas utiliser ces données, et ceux qui les utilisent pourraient le faire de manière erronée.

Pour l'Institut, outre ses actions posées auprès de SC en tant que coordonnateur statistique (ex. : accès à des CV, calcul des TGN, etc.) et la publication d'une [note d'information](#) à l'intention des utilisateurs du Québec, des MO ont contacté leurs experts pour les soutenir dans l'analyse de leurs données. D'abord, au sein même de l'Institut, les experts en méthodologie sont régulièrement consultés avant la diffusion de données par l'Institut. Soulignons que les conclusions d'une analyse comparative effectuée par des experts de l'Institut sur des données du Recensement de 2006 et l'ENM relativement au lieu de travail et au déplacement entre le domicile et le travail ont été transmises à SC pour information. Également, une étude de validation des données du revenu à l'échelle des municipalités a été effectuée, et les MO ont été informés de cette démarche lors des rencontres de consultation. Les principaux constats de cette analyse sont présentés plus loin dans ce rapport.

Des travaux ont été faits à la demande de certains MO pour estimer des effets de plan. Comme il a été souligné plus tôt, il est impossible de généraliser ces résultats sur de plus petites unités géographiques. Une organisation a, par ailleurs, demandé à l'Institut de prendre en charge le traitement des données de l'*Enquête canadienne sur*

l'incapacité (ECI), enquête dont l'échantillon est tiré de l'ENM. Certains conseils techniques ont aussi été donnés, par exemple, que l'on ne peut recalculer des TGN pour des unités géographiques différentes de celles de SC, car seul SC peut le faire; ou encore que l'on ne peut catégoriser les TGN. Pour une autre organisation, les tableaux et analyses ont été soumis à l'Institut avant la diffusion. Des suggestions de regroupement de catégories pour certaines variables ou d'analyses particulières ont entre autres été faites et des commentaires ont été émis pour que le texte diffusé reflète bien la nature d'une enquête volontaire (estimation, biais, etc.). Des conseils ont aussi été demandés pour la comparabilité des données sur le revenu entre 2006 et 2011. Une organisation privée a également consulté l'Institut avant de procéder à une demande de compilation spéciale. Des suggestions lui ont été proposées afin de mieux cibler sa demande visant l'obtention de données de meilleure qualité.

Des échanges ont aussi eu lieu sur les stratégies statistiques à déployer pour la mise à jour de l'indice de défavorisation sociale et matérielle dans le contexte de l'ENM. Il a d'ailleurs été suggéré de procéder à des analyses à partir du fichier maître de l'ENM pour mieux cibler les unités géographiques à retenir pour ces indices. Les solutions seront aussi partagées avec SC lorsque les travaux seront plus avancés. Par ailleurs, l'Institut a fait des présentations et a participé à des ateliers d'un regroupement d'analystes dans le domaine de la santé publique. Il a également collaboré à la rédaction d'un guide pour leurs utilisateurs.

Enfin, la consultation ayant conduit à ce présent rapport a permis aux MO de poser certaines questions méthodologiques à l'Institut et a offert à l'Institut l'opportunité de fournir les mises en garde appropriées quant à l'analyse et la diffusion des données.

7. BUREAU DU VÉRIFICATEUR GÉNÉRAL DU CANADA. *Rapport du Vérificateur du Canada*. Chapitre 8. Répondre aux besoins en statistiques clés. Statistique Canada. Printemps 2014, p. 17.

EXPERTISES

Pour procéder aux validations et analyses des données de l'ENM, encore faut-il avoir les expertises pour le faire. Pour certains MO, la seule option réside dans l'utilisation des données publiées par SC, et encore là, sans avoir pris la mesure de toutes les mises en garde. Comprendre les tenants et aboutissants d'une enquête n'est pas simple, une certaine formation de base est nécessaire. Par ailleurs, comme le précisait un utilisateur, il faut beaucoup de temps pour porter un regard critique. Or, le temps et les ressources humaines ne sont déjà pas toujours au rendez-vous, même pour fureter sur le site de SC. Les analystes ont également à défendre la pertinence et le temps requis pour effectuer de tels travaux auprès de leurs autorités. Encore là, ce ne sont pas tous les décideurs qui comprennent la complexité d'une enquête, et particulièrement celle de l'ENM.

Celles et ceux qui ont effectué des travaux à partir des données de l'ENM sont unanimes pour affirmer que le temps dévolu à l'analyse de ces données comparativement à celles du questionnaire détaillé du Recensement est décuplé en raison de toutes les étapes de validation à faire. Plusieurs MO doivent s'en tenir au plus urgent et laisser de côté certaines questions. Par exemple, une organisation aurait souhaité une étude de la cohérence des données de comptes des populations entre l'ENM et le Recensement pour toutes les municipalités du Québec, afin de pouvoir fournir ces informations aux municipalités dont le mandat n'est pas de faire de la statistique, mais plutôt des schémas d'aménagement, entre autres choses.

Certains MO ne peuvent faire l'économie de ne pas valider au mieux les données en lien avec les outils et les autres données à leur disposition, car celles-ci sont liées à des processus décisionnels. En l'absence de budgets et de ressources humaines additionnelles, d'autres travaux doivent donc être mis au rancart. En corollaire, plus on doit effectuer des travaux de validation avant d'utiliser les données, plus l'incertitude augmente quant à la qualité de celles-ci pour la prise de décision.

L'institut n'a pas, non plus, bénéficié de ressources additionnelles. Il a été cependant nécessaire de revoir certaines attributions pour soutenir dans la mesure du possible les MO. Également, à la lecture de certaines réponses transmises par les experts de SC lors de demandes spécifiques, par exemple le calcul des CV, il s'avère qu'ils sont également confrontés à des ressources limitées pour effectuer certains travaux. Faut-il rappeler que pour réaliser l'ENM, 22 millions de dollars (22 M\$) ont été ajoutés au budget prévu par rapport à la réalisation du recensement utilisant le questionnaire détaillé obligatoire ?

DES DÉMARCHES DE VALIDATION EXIGEANTES, MAIS PAS TOUJOURS CONCLUANTES

Le passage d'une enquête obligatoire à une enquête volontaire a mis à l'ordre du jour l'importance de l'étape de validation avant toute diffusion de données. Certes, les utilisateurs se sont préoccupés de cette question lors de l'analyse des données des recensements antérieurs, par exemple, en regardant les taux d'imputation, les changements dans les questions ou l'ordre de celles-ci, les changements de classification, etc., mais ces préoccupations sont minimales relativement à ce qui les attendait avec l'ENM. Dans cette section, nous ferons tour à tour état des validations qui ont été faites à partir des données du Recensement de 2011 ou des recensements antérieurs, des données administratives et d'autres types d'enquête. Par ailleurs, les résultats de ces validations seront présentés lorsqu'elles ont été portées à notre attention. Soulignons que plusieurs sources de données différentes peuvent avoir servi pour une seule validation. Nous les traitons cependant séparément pour mieux évaluer les sources de données potentielles.

SC a également procédé à des validations avec plusieurs des enquêtes ou bases de données administratives mentionnées par les MO consultés. Quoique ces validations soient utiles pour les utilisateurs de données de l'ENM, elles n'ont pas été suffisantes pour certaines sous-populations québécoises ou encore pour des univers géographiques plus fins que les RMR ou le Québec.

DONNÉES DU RECENSEMENT

Dès que le TGN franchit la barre des 30 %, une organisation compare systématiquement les estimations de l'ENM et les chiffres de la population du Recensement de 2011. Une autre organisation a comparé les chiffres du Recensement de 2011 avec les données de l'ENM, et ce, pour les divers types de ménage dans chacune des régions administratives du Québec. À la suite de cet exercice, il leur est apparu hasardeux de classer et de comparer les régions entre elles. Des comparaisons ont également été faites dans une autre organisation, plus

particulièrement pour le nombre de familles monoparentales. Selon cette analyse, les données seraient assez près pour l'ensemble du Québec et les régions administratives, mais à l'échelle des SDR, on observerait parfois une sous-estimation assez importante des familles monoparentales. On conclut qu'à un niveau encore plus fin, les risques sont plus importants. Pour une région administrative, une organisation a fait des comparaisons des proportions de familles monoparentales avec enfants de moins de 18 ans dont la chef est une femme dans 481 zones. Selon cette étude, le coefficient de corrélation s'élève à 73 %. Une analyse complémentaire a également été faite dans 50 quartiers. Pour 30 des 50 quartiers, la différence des proportions de l'ENM et du Recensement est plus élevée que de 5 points de pourcentage. L'analyste a toutefois conclu que les résultats sont convergents. Par contre, une autre organisation constate que les comparaisons entre les deux sources de données pour les familles monoparentales dont le chef est de sexe féminin donnent des différences importantes pour de petits volumes de population. Ce MO a d'ailleurs décidé de publier des données en regroupant certaines catégories qui avaient fait l'objet d'analyses séparées avec les recensements antérieurs afin d'avoir de bons volumes de population (ex. : en regroupant les enfants de 0 à 4 ans).

Dans une autre région, les ratios des effectifs de l'ENM par rapport à ceux rapportés du Recensement ont été examinés à l'échelle des secteurs du recensement (SR), et ce, dans l'univers des ménages privés, pour la population totale, la population des personnes seules et pour la population des personnes âgées de 65 ans et plus. Au départ, les TGN ont été considérés comme acceptables, seulement 10 SR sur 80 ont un TGN de plus de 30 % et un seul au-delà de 40 %, mais inférieur à 50 %. En ce qui concerne les ratios, celui touchant les personnes âgées de 65 ans et plus s'est révélé plus problématique. Toutefois, les territoires retenus pour leurs travaux regroupent des SR, et à cette échelle, les ratios ne semblent pas poser de problèmes. Le regroupement de SR permettrait de réduire les erreurs.

Dans un petit territoire de santé d'environ 137 000 personnes, dont les sous-territoires se situaient entre 4 500 et 51 000, il y a eu comparaison entre les chiffres de l'ENM et ceux du Recensement pour plusieurs variables communes. Il y a eu constat d'écart inférieurs à 2 % pour la population totale; des écarts pouvant atteindre 3 % pour les ménages d'une seule personne, 7 % pour la population totale de 65 ans et plus et enfin des écarts pouvant atteindre 9 % pour la population des hommes de 65 ans et plus. Dans ce dernier cas, l'écart se situait à 4 % pour l'ensemble du territoire de santé (137 000). L'ENM était le plus souvent sous-évalué. Leur constat est que l'écart entre les estimations de l'ENM et le recensement varie selon l'indicateur, le territoire et la prévalence du phénomène.

Un analyste a calculé les ratios ENM et Recensement pour quelques variables. Ces ratios ont été mis en parallèle avec les TGN. Aucune corrélation apparente ne ressort pour expliquer les écarts entre les données de l'ENM et le Recensement.

À l'occasion, les données des questionnaires détaillés des recensements antérieurs sont également examinées avec celles de l'ENM pour évaluer les tendances et les ordres de grandeur. Autrement dit, est-ce que l'estimation de l'ENM fait sens considérant les valeurs attendues? Dans le cas d'une analyse concernant la variable scolarité de la mère et les activités des parents pour les quartiers d'une grande ville, l'analyste a conclu que les résultats vont dans le même sens. Une analyse par MRC dans une région administrative a comparé les proportions entre les données du Recensement de 2006 et l'ENM pour les réparations majeures. Les données étaient cohérentes pour la plupart des MRC, sauf deux, dont les données pour plusieurs SDR avaient été supprimées (des produits normalisés). Les différences ne pouvaient s'expliquer par les changements démographiques ou socioéconomiques. L'effet de non-réponse était donc soupçonné. Toujours dans le domaine du logement, et dans une région comptant une population très diversifiée et où les profils démographiques ont beaucoup changé entre 2006 et 2011, on évalue que les raisons des écarts observés peuvent provenir, soit des biais de non-réponse ou des changements démographiques importants. Impossible de conclure.

En comparant les classements des territoires retenus pour des indicateurs de performance, un MO a remarqué un plus grand mouvement dans les rangs entre 2006 et 2011 qu'entre 2001 et 2006 (coefficient de Spearman).

On ne peut conclure que le changement est dû à l'évolution réelle du phénomène observé ou bien au biais de non-réponse. À l'échelle des industries, une organisation conclut que les données entre 2006 et 2011 semblent comparables à l'échelle des divisions de recensement et de la province.

Enfin, une autre étude portant cette fois sur la scolarité à l'échelle du Québec et des régions administratives montre que les changements entre 2006 et 2011 sont cohérents avec les résultats attendus. À noter que cette dernière étude, ainsi que celles sur les réparations majeures, a également examiné d'autres sources de données pour en arriver à ces conclusions.

D'autres poussent l'analyse plus loin en utilisant les données du fichier maître de l'ENM afin de comparer les données de 2006 ayant servi aux calculs de leur indicateur selon des regroupements d'aires de diffusion bien définies, avec les données provenant de l'ENM pour ces mêmes regroupements. L'une des organisations a d'ailleurs comparé leur carte de défavorisation de 2006 avec les TGN de l'ENM. Une corrélation entre la non-réponse et la défavorisation est ressortie. Au moment d'écrire ces lignes, les travaux n'étaient pas terminés.

DONNÉES ADMINISTRATIVES

Soulignons d'emblée que la plupart des bases de données administratives mentionnées par les experts rencontrés ont uniquement pour univers l'ensemble du Québec. Et lorsque celles-ci offrent des données pour des territoires géographiques plus fins, ces derniers peuvent ne pas correspondre aux territoires correspondant aux objectifs d'une autre organisation. Comme le mentionne l'une d'entre elles œuvrant à l'échelle de petits territoires, plusieurs données administratives existent pour dresser un portrait de la vulnérabilité de la population : mortalité, signalement jeunesse, aide sociale, difficultés d'apprentissage et de comportements dans les commissions scolaires. Toutefois, les univers géographiques divergent d'une banque de données à l'autre et correspondent difficilement à la population étudiée.

Les données des migrations internes au Québec sont diffusées par l'Institut, jusqu'à l'échelle des MRC. Ces données s'appuient sur le Fichier d'inscription des personnes assurées de la Régie d'assurance maladie du Québec (FIPA). Ces données sont utilisées par quelques MO pour valider les données sur les changements de

lieu de résidence de l'ENM, surtout pour les migrations entre régions. Toutefois, les données de base du FIPA ne fournissent que des nombres de personnes selon des lieux de résidence, et non les caractéristiques de ces populations. Cette donnée administrative, quoique fiable, ne peut mesurer l'ampleur des biais liés à la caractéristique des répondants et des non-répondants.

Dans le domaine du logement, certaines données administratives ont été explorées et d'autres faisaient partie des options. Dans une région, on a comparé les données du rôle d'évaluation foncière avec celles du questionnaire détaillé de 2006 et de l'ENM. Toutefois, cette donnée comporte ses limites puisqu'un certain pourcentage des logements n'y est pas inscrit (environ 7%). De plus, on ne sait pas si les logements sont occupés ou non.

Comme solutions de validation, d'autres sources de données provenant de la Société canadienne d'hypothèque et de logement (SCHL) ont été mentionnées, comme les stocks de logements, les mises en chantier, les prix des loyers et des maisons, les années de construction, tout comme le registre privé des courtiers immobiliers sur la valeur du logement estimé. Toutefois, à notre connaissance, ces données n'avaient pas encore été exploitées. Une organisation a mentionné d'ailleurs que les sources de données alternatives n'étaient pas nécessairement fiables, notamment pour mesurer des phénomènes de « gentrification ».

Dans une région à faible densité de population, les données concernant le nombre d'immigrants seraient en baisse lorsqu'on compare les données provenant du questionnaire détaillé du Recensement de 2006 et l'ENM de 2011, alors que les données administratives montrent une augmentation. Cette augmentation est aussi perçue sur le terrain par les intervenants de la région. Une organisation a déploré le fait qu'il était difficile d'accéder aux données administratives sur les immigrants.

Certaines données administratives sont disponibles sur la scolarité (fréquentation scolaire, taux de décrochage, nombre de diplômes, etc.). Toutefois, elles le sont par commission scolaire, ce qui ne correspond pas toujours aux délimitations des territoires des programmes des autres MO. Dans le cas d'une grande région, une organisation a comparé les données administratives sur les retards scolaires et la scolarité de la mère. Les résultats sont convergents. Pour la langue maternelle, une organisation a évalué la possibilité d'utiliser la donnée inscrite

dans les bulletins de naissance et de décès. Toutefois, cette variable n'est pas toujours fiable, notamment parce qu'elle est souvent complétée par un tiers.

On a comparé les tendances du questionnaire détaillé du Recensement de 2006 et l'ENM de 2011 sur l'utilisation du transport en commun avec des données administratives sur l'achalandage du transport en commun. Quoique ces deux sources de données ne couvrent pas les mêmes univers, une tendance similaire est observée.

Concernant les origines ethniques, lorsque de grandes différences sont observées entre 2006 et 2011, une organisation effectue des vérifications avec une banque de données sur les immigrants. Par ailleurs, pour la population autochtone, deux registres sont utilisés, soit celui sur les Premières Nations et l'autre sur les nations conventionnées, soit les Inuits, les Cris et les Naskapis. Le premier est sous la responsabilité du gouvernement fédéral (Affaires autochtones et Développement du Nord Canada-AADNC) et l'autre du Gouvernement du Québec (ministère de la Santé et des Services sociaux). Les registres fournissent des données selon l'âge et le sexe. Peu d'analyses ont été produites pour comparer ces données avec l'ENM. La non-participation de plusieurs communautés au Recensement et à l'ENM de 2011, ainsi que la non-publication de données dans les produits normalisés en raison des TGN élevés constituent un frein important aux analyses. En raison du peu de ressources pour effectuer de tels travaux, l'organisation québécoise se tourne beaucoup vers les analyses produites par AADNC.

DONNÉES PROVENANT D'AUTRES ENQUÊTES

D'autres enquêtes ont également été examinées afin de valider les données de l'ENM, ou il est prévu qu'elles le soient. Les principales sources de données mentionnées sont l'EPA, l'EDTR, l'*Enquête des dépenses des ménages* (EDM), l'*Enquête sociale générale* (ESG) et l'*Enquête sur l'emploi, la rémunération et les heures de travail* (EERH). Soulignons que toutes ces enquêtes sont administrées par SC, et les échantillons sont surtout conçus pour des représentativités à l'échelle des provinces. Certaines offrent des données par région administrative ou RMR, mais pas toutes.

La comparaison des données de l'EPA et de l'ENM pour les professions a peu porté ses fruits, les différences entre les regroupements de professions étant trop importantes. Rappelons que l'ENM utilise la classification CNP 2011, alors que l'EPA ne l'utilisera qu'en janvier 2015. Cette différence de classification a été déplorée. Pour les données par industrie, celles de l'EPA seraient discutables. Un analyste souligne qu'il ne serait pas approprié de comparer avec des données considérées moins bonnes que celles de l'ENM. D'ailleurs, SC suggère d'utiliser plutôt l'EERH. Par contre, pour des comparaisons spécifiques dans l'industrie de la culture et des communications, aucun écart aberrant n'a été observé entre l'ENM et l'EPA à l'échelle du Québec. On a tout de même vérifié aussi avec l'EERH et on a constaté que, sauf pour le secteur de l'édition, les tendances sont semblables aux données de l'ENM.

Une comparaison a été faite entre les données de l'EPA et de l'ENM pour comparer les tendances quant au nombre de travailleurs rémunérés et autonomes entre 2006 et 2011 pour 16 industries du Canada et au Québec, ainsi que pour le total de l'économie. Les principales observations sont décrites ci-après.

On constate une augmentation entre le recensement de 2006 et l'ENM du nombre de travailleurs rémunérés au Canada (+ 12,4 %) et au Québec (+ 11,8 %), mais dans une moins grande mesure avec l'EPA : + 5,2 % au Canada et au Québec. Il y a de fortes augmentations de travailleurs rémunérés dans l'industrie des cultures agricoles et de l'élevage (+ 37,7 % au Canada et + 23,5 % au Québec), mais une diminution pour l'agriculture dans l'EPA (-4,8 % au Canada et -6,9 % au Québec). À noter que les secteurs ne représentent peut-être pas exactement les mêmes industries. Dans l'industrie de la foresterie, de la pêche, des mines, de l'extraction de pétrole et de gaz, les données de l'ENM montrent une augmentation de travailleurs rémunérés au Québec (+ 5,8 %), tandis qu'il y a une baisse selon l'EPA (-4,9 %). Dans les industries des services professionnels, scientifiques et techniques, des services d'enseignement et des soins de santé et de l'assistance sociale, la hausse du nombre de travailleurs rémunérés est plus marquée dans l'ENM que dans l'EPA.

En ce qui concerne les travailleurs autonomes, il y a une diminution au Canada (-9,7 %) et au Québec (-5,0 %) avec l'ENM. Selon l'EPA, il y a une augmentation de 9,2 % au Canada et de 13,2 % au Québec. L'industrie de la construction affiche une baisse de 1,9 % au Canada et une hausse de 3,3 % au Québec du nombre de travailleurs autonomes. Les données de l'EPA affichent plutôt une

hausse dans l'industrie de la construction. En 2006, on ne comptait aucun travailleur autonome dans les administrations publiques, alors qu'en 2011, 1 065 personnes travaillent dans l'administration publique fédérale, 1 615 dans celle provinciale et 845 dans celle autochtone ou locale. L'EPA ne nous fournit aucune donnée sur le nombre de travailleurs autonomes dans l'administration publique.

L'EPA a également été utilisé pour valider les données de l'ENM dans le domaine de l'éducation. Notamment, les données régionales sur le niveau de scolarité ont été comparées entre celles de l'EPA et le questionnaire détaillé de 2006. À quelques exceptions près, les données étaient très similaires (entre 1 et 2 points de pourcentage). Les concepts étant les mêmes pour l'ENM et l'EPA en 2011, on a donc conclu que l'EPA est une bonne enquête pour valider les données sur le niveau de scolarité en 2011. Des travaux ont aussi été faits sur le nombre de diplômés universitaires. En utilisant des moyennes sur cinq ans pour tenir compte de la volatilité de l'EPA, on constate que les variations de l'ENM sur le nombre de diplômés universitaires à l'échelle de la province et des régions tiennent la route. En fait, ce seraient les données de l'EPA qui seraient surestimées. D'autres validations ont été faites : 1) Rangements des régions selon la proportion des titulaires d'un grade universitaire. Les observations seraient cohérentes entre 2006 et 2011. Les quelques permutations observées entre les régions sont explicables; 2) Proportions de titulaires de grade universitaire femmes et hommes, par provinces et régions. Les comparaisons sont cohérentes; 3) Évolution des écarts entre les hommes et les femmes. Un peu plus de variabilité à l'échelle des régions, mais sans qu'aucune différence importante soit notée; 4) Les données observées pour l'ENM sont aussi cohérentes avec les données administratives pour les taux d'obtention des diplômes; 5) Les résultats infraprovinciaux sont influencés par la conjoncture et la structure économique, les écarts peuvent donc être liés à ces changements. On conclut que l'évaluation de l'écart entre les sexes, par groupe d'âge et provinces, tient la route. Toutefois, pour les variations infra régionales, il y a trop d'impondérables et elles ne peuvent donc être comparées. Cette analyse soutient que les comparaisons entre le questionnaire détaillé du Recensement de 2006 et l'ENM de 2011 sont possibles à l'échelle du Québec pour la variable de grade universitaire.

Des analyses supplémentaires à partir de l'EPA étaient également prévues dans une autre organisation, une fois celles entre le Recensement et l'ENM complétées à l'égard des comptes de population.

Une tentative a été faite pour comparer les données de revenus de l'ENM avec celles de l'EDTR. Cependant, les données disponibles dans CANSIM ne sont pas suffisamment détaillées pour les croisements retenus avec l'ENM (ex. : revenus par profession pour les RMR de Montréal et Québec). Pour obtenir des données plus précises, il aurait fallu procéder à une demande de compilations spéciales pour l'EDTR, alors que le MO a déjà payé pour une commande spéciale pour l'ENM. Les coûts devenaient trop importants. Une autre organisation a comparé les données sur le faible revenu selon l'EDTR et l'ENM. Des différences sont observées. D'autres analyses étaient à venir. L'extraction de données de l'EDTR sur le logement était aussi envisagée.

L'EDM constituait une option possible pour évaluer les tendances dans le domaine du logement, puisqu'elle est annuelle. Toutefois, depuis 2010, l'EDM a été revue en profondeur, tant sur les aspects méthodologiques que sur son contenu. Le module logement de l'EDM comprend moins de variables qu'antérieurement. Par exemple, le nombre de pièces dans le logement n'est plus collecté. Dans ce contexte, l'ENM devient alors la seule source disposant de données sur la taille du logement. En somme, tant pour l'EDM que pour l'ENM, il n'est plus possible de dégager des tendances historiques.

Pour l'analyse sur les déplacements domicile-travail, les enquêtes « origine destination » (OD) dans des agglomérations importantes sont utilisées. Ces enquêtes réalisées indépendamment d'une région à l'autre ont permis d'accumuler de l'information sur les déplacements de plus des deux tiers de la population québécoise. Quoique très utiles, ces enquêtes ne peuvent à elles seules répondre aux questions sur les déplacements, notamment à l'extérieur des grands centres urbains. Pour valider les déplacements, on se sert également de l'ESG sur l'emploi du temps à l'échelle des RMR.

Des sondages sont réalisés auprès de la clientèle immigrante. Toutefois, comme il est difficile de rejoindre certains groupes, les données obtenues ne présentent pas un portrait fidèle de la population. La validation avec les données de l'ENM s'avère donc difficile. L'ESG a aussi été mise à profit pour des validations dans le domaine de l'éducation. Les données ont confirmé les tendances observées entre le questionnaire détaillé de 2006 et l'ENM de 2011.

EN L'ABSENCE DE SOURCES DE DONNÉES COMPARABLES ?

Pour pouvoir valider les données extraites de l'ENM, encore faut-il avoir des sources de données comparables. Or, particulièrement pour des unités géographiques fines ou pour des variables controversées comme la religion, les sources de données alternatives sont rares, voire inexistantes. Dans les cas où des données administratives sont disponibles, le plus souvent les univers territoriaux ne concordent pas, et peuvent difficilement être conciliés. En effet, les objectifs poursuivis pour la tenue de fichiers administratifs ne cadrent pas nécessairement avec tous les besoins (ex. : commission scolaire c. municipalités ou quartiers). À l'échelle des régions administratives, la situation s'améliore un peu, du moins pour les données administratives. Mais c'est rarement le cas avec des données des enquêtes de SC où les univers les plus fins sont la majorité du temps les RMR. Les défis sont aussi nombreux pour les analyses portant sur des sous-populations (professions, industries, etc.). Les sources de données alternatives sont loin de pouvoir combler des besoins très précis.

Un défi qui nous a été soumis concerne la validation elle-même des sources alternatives. Comme chaque enquête ou fichier administratif comporte sa méthodologie propre, les analystes doivent d'abord s'assurer de bien comprendre les données utilisées. Cette tâche d'appropriation ajoute de la lourdeur au processus de validation des données de l'ENM. Il nous a aussi été souligné qu'auparavant, la tendance était plutôt inverse, on validait les autres sources de données avec les résultats provenant du questionnaire détaillé du recensement. Il faut donc maintenant s'appuyer sur des enquêtes qui peuvent être plus volatiles que l'ENM (plus faible taux de réponse et plus faible échantillon) ou sur des fichiers administratifs qui sont soumis aux aléas des décisions administratives et politiques.

Enfin, de nombreux analystes mettent à profit leurs expertises de la thématique abordée ou leur connaissance du terrain pour évaluer si les données de l'ENM ont du sens. Par exemple, les changements structurels sont généralement lents à percevoir. Une bonne connaissance historique d'un domaine peut faciliter l'interprétation des changements. Cette solution est cependant possible uniquement lorsque les connaissances accumulées par les analystes sont suffisantes. Et au dire même de ces derniers, normalement, la mise à profit de l'expertise va dans l'autre sens, à savoir que les données servent à confirmer ce qu'ils observent sur le terrain.

DE FORTES INQUIÉTUDES SUR LA QUALITÉ DES DONNÉES SUR LE REVENU

L'une des thématiques ayant le plus fait sourciller les analystes rencontrés concerne la donnée sur le revenu. D'ores et déjà, il est connu que les données sur le revenu dans la majorité des enquêtes montrent des taux de non-réponse partielle plus élevés que pour plusieurs autres variables, en raison de la sensibilité de cette variable. Des efforts sont généralement faits pour en réduire les impacts. SC avec l'ENM en a fait de même. Entre autres, on demande aux répondants d'autoriser SC à consulter leurs dossiers fiscaux au lieu de répondre à toutes les questions, augmentant ainsi leur fiabilité. En 2011, 77,7 % des répondants québécois ont donné leur accord, en diminution par rapport au Recensement de 2006 (84,7 %).

Le [Guide de référence sur le revenu](#) produit par SC montre que malgré les efforts, des problèmes persistent. Après avoir validé les données de l'ENM avec celles de l'EDTR et le Fichier sur la famille T1 (FFT1) produit à partir de données fiscales, SC a émis quelques mises en garde importantes sur la comparabilité des données du questionnaire détaillé et obligatoire de 2006 et l'ENM 2011, surtout pour les estimations de faible revenu. SC a même pris la décision de ne pas rendre disponibles dans ses produits normalisés les données sur les taux de faible revenu selon la mesure SFR comme dans le passé. Seuls les taux selon la mesure du faible revenu (MFR) sont disponibles. Les discussions qui ont eu cours dans la presse, par exemple la lettre ouverte de chercheurs de l'Université de Toronto (Globe and Mail, 4 octobre 2013), ont augmenté la méfiance à l'égard de ces données. Or, ces craintes sont encore amplifiées lorsque les données sur le revenu sont exploitées à une échelle géographique relativement fine.

QUELQUES CONSTATS DES ANALYSTES DES MO

Les mises en garde de SC concernant les données sur le revenu ont été entendues. La majorité des MO demeurent prudents lorsqu'ils utilisent ces données. Certains déplorent cependant que pour obtenir des données sur les taux de faible revenu selon la SFR, ils doivent demander des compilations spéciales. Ainsi, en plus d'avoir à composer avec des données moins fiables, ils doivent assumer le coût des compilations spéciales et les délais que cela occasionne.

Peu se sont cependant aventurés à évaluer la qualité des données de revenu. Parmi ceux qui l'ont fait à des échelles géographiques variées, des différences entre les estimations de l'EDTR et de l'ENM ont été constatées, que ce soit pour les données sur le revenu ou le faible revenu. Une organisation a aussi observé que les tendances en ce qui a trait aux revenus des personnes de 65 ans et plus ne s'inscrivent pas dans la continuité. Par contre, à l'échelle de la RMR de Montréal, les données semblent fiables et l'évolution dans le temps est celle attendue. Ce dernier résultat va dans le sens des résultats de l'analyste effectuée par l'Institut, dont les principaux résultats sont présentés ci-après.

VALIDATION AVEC DES DONNÉES FISCALES

Puisque dans son avis, SC semble avoir limité son analyse à l'échelle des provinces et des territoires et des RMR, l'Institut a procédé à des analyses plus poussées pour le Québec, en comparant les données fiscales du Québec (Revenu Québec) avec celles de l'ENM. L'étude concerne le revenu des particuliers de 15 ans et plus avec revenus. De plus, en raison de la meilleure stabilité des estimations du revenu médian que du revenu moyen, les comparaisons portent principalement sur le revenu médian.

Rappelons que la population visée par l'ENM est l'ensemble des personnes de tous âges vivant habituellement au Canada (provinces et territoires) et dans des ménages privés, y compris les personnes vivant dans des réserves indiennes et les résidents permanents et non permanents issus de l'immigration. Quant aux données provenant de Revenu Québec (RQ), elles sont issues des dossiers de toutes les déclarations de revenus des résidents au Québec. Dans le but d'améliorer la comparabilité entre les deux sources de données, la population de référence a été restreinte aux personnes ayant des revenus, et des efforts ont été faits pour le contrôle de la qualité des données de RQ à des niveaux géographiques fins. De plus, seules les variables de revenu de RQ permettant une compatibilité maximale avec celles de l'ENM sont retenues. Malgré ces ajustements, des différences peuvent persister dues au fait notamment que les données de RQ comprennent les personnes vivant en logement collectif, contrairement à l'ENM, et que les définitions dans l'ENM ne correspondent pas nécessairement parfaitement aux définitions utilisées pour les données fiscales (RQ).

L'approche préconisée dans cette analyse est de mettre en perspective deux facteurs fondamentaux : la précision des estimations et la présence potentielle de biais. Deux indicateurs sont utilisés pour mesurer le niveau de concordance entre les deux sources de données : le *coefficient de corrélation* entre les estimations, et la *distribution de l'écart* entre ces estimations. Plus le coefficient de corrélation se rapproche de « 1 », plus le niveau de concordance est élevé. Soulignons toutefois que l'on peut difficilement associer une « valeur » du coefficient de corrélation à un « niveau » de concordance. De plus, s'il s'avérait qu'une source de données sous-évalue ou surévalue le revenu par rapport à l'autre, le coefficient de corrélation ne pourrait pas détecter cette situation. La distribution de l'écart offre des atouts à ces deux égards : elle permet de donner une idée relativement intuitive du niveau de concordance, et permet aussi de détecter toute différence de niveaux entre les estimations des deux sources. En supposant qu'il n'y a pas de telles différences, plus l'écart relatif se rapproche de « 0 », plus le niveau de concordance est élevé. Par exemple, si un pourcentage élevé d'unités de la comparaison présentent un écart relatif d'estimations entre $-2,5\%$ et $+2,5\%$, alors on peut conclure que le niveau de concordance est élevé.

Selon l'échelle géographique

Plus le niveau géographique est agrégé, plus la concordance ENM-RQ augmente, comme l'a montré l'analyse des corrélations et des distributions de l'écart relatif entre les estimations⁸ des deux sources de données. Ainsi, la corrélation est quasi parfaite à l'échelle des RA et des MRC, soit respectivement 0,99 et 0,98, tandis qu'elle s'abaisse à 0,80 à l'échelle des municipalités, dénotant une moins bonne concordance. Par ailleurs, pour les RA, une très grande majorité des unités géographiques (94,1 %) présentent un écart relatif entre les estimations qui est situé entre $-2,5\%$ et $+2,5\%$. Pour les MRC, ce pourcentage diminue à 61,5 %, tandis qu'à l'échelle des municipalités, il n'est plus que de 22,1 %.

De plus, pour un niveau géographique donné, plus l'effectif des unités géographiques augmente, plus la concordance augmente. Ainsi, la corrélation des estimations⁹ entre l'ENM et RQ passe de 0,59 parmi les municipalités dont l'effectif¹⁰ est moins de 250, à 1,00 parmi celles de 100 000 et plus. Entre ces deux tranches extrêmes, la corrélation, c'est-à-dire la concordance, est toujours croissante : 0,60 (250 à 499), 0,74 (500 à 999), 0,90 (1 000 à 4 999), 0,98 (5 000 à 24 999) et 0,99 (25 000 à 99 999). Ces résultats montrent que plus l'effectif augmente, meilleure est la concordance entre les deux sources de données. Dans le même ordre d'idées, le pourcentage des municipalités, dont l'écart relatif ENM-RQ se situe entre $-2,5\%$ et $+2,5\%$, augmente de manière systématique en fonction de cet effectif : 7,7 % (moins de 250), 11,9 % (250 à 499), 18,4 % (500 à 999), 20,5 % (1 000 à 4 999), 49,6 % (5 000 à 24 999), 71,4 % (25 000 à 99 999) et 88,9 % (100 000 et plus). Des conclusions similaires ressortent à l'échelle des MRC et des RA. Cependant, l'amélioration de la concordance lorsque l'effectif augmente, est moins marquée à ces échelles, étant donné que la dispersion de l'effectif y est moins marquée.

Enfin, pour un niveau géographique donné, plus le taux global de non-réponse (TGN) diminue, plus la concordance augmente. La corrélation s'élève seulement à 0,60 entre les municipalités dont le TGN est de 50 % et plus. La corrélation augmente à 0,79 pour les municipalités dont le TGN est compris entre 40 % et 50 % ; elle s'abaisse légèrement pour atteindre 0,74 pour celles affichant un

8. Il s'agit des estimations du revenu total médian des personnes de 15 ans et plus avec revenus.

9. Voir note précédente.

10. Il s'agit de l'effectif non pas de la population totale des municipalités, mais de l'effectif de la population de référence choisie aux fins de la comparaison, soit les personnes de 15 ans et plus avec revenus.

TGN entre 30 % et 40 %. Elle effectue cependant un bond à 0,93 pour celles dont le TGN varie entre 20 % et 30 %, puis augmente à 0,95 avec un TGN entre 10 % et 20 %. Enfin, elle diminue légèrement à 0,89 lorsque le TGN est inférieur à 10 %.

Par ailleurs, en ce qui concerne l'indicateur de l'écart relatif, seulement 15,7 % des unités géographiques parmi celles dont le TGN est de 50 % et plus présentent un écart relatif situé entre -2,5 % et +2,5 %; ce faible pourcentage dénote donc une faible concordance. Le pourcentage des unités dont l'écart relatif varie entre -2,5 % et +2,5 % augmente à 18,7 % lorsque ces unités ont un TGN entre 40 % et 50 %; ce pourcentage reste sensiblement stable (18,0 %) pour les unités dont le TGN varie entre 30 % et 40 %; mais il augmente à 28,9 % pour celles dont le TGN varie entre 20 % et 30 %, puis à 44,6 % pour celles dont il varie entre 10 % et 20 %. Par contre, la relation entre le TGN et la concordance semble se brouiller quand le TGN est inférieur à 10 % : seulement 21,4 % des municipalités avec un TGN à ce niveau, soit trois municipalités sur un total de quatorze, manifestent un écart relatif entre -2,5 % et +2,5 %.

Selon les populations

La concordance entre les deux sources de données est modérément plus élevée chez les hommes que chez les femmes, et beaucoup plus élevée parmi les personnes de 25 ans et plus que chez celles de 15 à 24 ans. Par exemple, à l'échelle des municipalités, la corrélation entre les estimations du revenu total médian est de 0,80 chez les hommes contre 0,73 chez les femmes, tandis qu'elle est seulement de 0,27 chez les 15-24 ans contre 0,69 chez les 25-44 ans, 0,80 chez les 45-64 ans et 0,54 chez les 65 ans et plus. En outre, les différences dans les niveaux de concordance selon la tranche d'âge sont reliées aux différences d'effectifs selon l'âge. En effet, l'effectif des 15-24 ans, et dans une moindre mesure celui des 65 ans et plus, est plus faible que celui des 25-44 ans et des 45-64 ans. De ce fait, l'erreur aléatoire de l'ENM y est en même temps plus élevée. Par ailleurs, on ne peut associer la faible concordance observée chez les jeunes de 15 à 24 ans à des biais de sous-représentation de ce groupe dans les données fiscales, puisqu'avec

ces données, la part que ce groupe représente dans l'ensemble de la population (des 15 ans et plus avec revenus) est presque identique à celle que l'on obtient avec les données de l'ENM.

Ces niveaux de concordance sont atténués par le fait que l'ENM sous-évalue légèrement le revenu des personnes de 15 à 24 ans, et surévalue légèrement celui des femmes et des personnes de 25 à 44 ans. Par exemple, à l'échelle des municipalités, 70,6 % des unités géographiques présentent un écart relatif positif chez les femmes, ce qui dénote une tendance de l'ENM, comparativement aux données de RQ, à surévaluer le revenu total médian des femmes¹¹. De même, 69,0 % des municipalités présentent un écart relatif négatif chez les personnes de 15 à 24 ans, alors que 69,9 % présentent un écart relatif positif chez les 25 à 44 ans, ce qui dénote une tendance de l'ENM, comparativement aux données de RQ, à sous-évaluer le revenu total médian des 15-24 ans et à surévaluer celui des 25-44 ans. Ces différences de niveaux demeurent toutefois difficilement explicables. En effet, ni la taille des effectifs ni les TGN des unités géographiques ne semblent en cause.

Selon les indicateurs de revenu

La concordance entre les données de l'ENM et de RQ est légèrement plus élevée avec la moyenne que la médiane, et plus élevée pour la population dont le revenu est supérieur ou égal à 100 000 \$ que celle dont le revenu est inférieur à 10 000 \$. Par exemple, à l'échelle des municipalités, la corrélation entre les estimations du revenu total moyen est légèrement supérieure à celle des estimations du revenu total médian : 0,89 contre 0,80. Aussi, lorsque les estimations comprennent la population dont le revenu total est supérieur ou égal à 100 000 \$, la corrélation atteint 0,86 à l'échelle des municipalités, contre seulement 0,53 pour la population dont le revenu est inférieur à 10 000 \$. Des imprécisions dans la procédure d'imputation des revenus manquants pour les personnes en bas de l'échelle du revenu expliquent vraisemblablement la faible corrélation. Il n'est pas impossible non plus que la couverture de ces personnes avec les données de RQ soit imprécise, bon nombre d'entre elles étant à la limite du seuil d'imposition.

11. En effet, dans le présent rapport, l'écart relatif est calculé de telle manière qu'un écart négatif implique une sous-estimation de la donnée de l'ENM relativement à celle de RQ, et inversement pour tout écart positif.

Par ailleurs, même si le revenu moyen concorde mieux que le revenu médian du point de vue de la corrélation, le revenu moyen semble sous-estimé avec les données de l'ENM comparativement à celles de RQ. Cette tendance est surtout décelable à l'échelle des MRC et des régions administratives, où respectivement 77,9% et 100% des unités présentent un écart relatif négatif avec cet indicateur. Cette sous-estimation du revenu moyen, qui ne s'applique pas au revenu médian, s'explique vraisemblablement par une tendance de l'ENM à sous-évaluer les valeurs extrêmes du revenu.

Selon les concepts du revenu

Pour l'analyse selon les concepts du revenu, les résultats concernent le revenu médian des personnes de 15 ans et plus avec revenus. L'analyse montre que les estimations du revenu total sont presque aussi concordantes que celles du revenu du marché, et que ces dernières sont plus concordantes que celles des transferts gouvernementaux. Par exemple, à l'échelle des municipalités, la corrélation entre les estimations des deux sources de données est de 0,78 pour le revenu du marché, contre 0,69 pour les transferts. Rappelons qu'elle était de 0,80 pour le revenu total. Soulignons qu'il est possible que l'absence de comptabilisation des crédits d'impôt fédéraux dans les données de RQ contribue à l'explication de cette concordance relativement plus faible pour les transferts.

De plus, la concordance est de loin meilleure pour les salaires et traitements que pour tous les autres concepts de revenu du marché. Ainsi, à l'échelle géographique des municipalités, la corrélation s'élève à 0,72 pour les salaires et traitements, alors que pour chacun des cinq autres concepts (le revenu agricole net, le revenu non agricole net provenant de l'exploitation d'une entreprise non constituée en société ou de l'exercice d'une profession, les dividendes et revenus de placements, les pensions de retraite et rentes, les autres revenus privés en espèces), elle ne dépasse pas 0,55. Nous ne sommes cependant pas en mesure de confirmer si ces plus faibles niveaux de concordance sont liés à des différences mineures dans la définition de ces divers concepts entre les deux sources de données.

En résumé, cette démarche a permis de faire ressortir que les données de l'ENM pour le revenu semblent de meilleure qualité pour les plus grandes unités géographiques, celles dont le taux global de non-réponse est faible, pour le groupe d'âge de 25 à 64 ans, les hommes, pour le revenu moyen, pour les particuliers dont le revenu est plus élevé, pour le revenu total ou pour l'ensemble des revenus du marché.

DES DÉCISIONS DIVERGENTES QUANT AUX ESTIMATIONS DE L'ENM POUVANT ÊTRE DIFFUSÉES OU UTILISÉES

Une fois les données compilées et analysées, les MO doivent prendre des décisions sur ce qui peut être diffusé ou utilisé. Cette étape s'est avérée difficile pour plusieurs analystes qui doivent prendre en compte les besoins de leurs organisations et leurs propres constats sur la qualité des données. Comme il a été précisé précédemment, nombreux sont les analystes qui ne souhaitent pas devoir trancher sur la qualité des données alors qu'ils ne disposent pas de tous les outils statistiques pour le faire.

Dans cette section, nous verrons quels genres de décisions ont été prises par les MO, mais également les raisons ayant motivé ces choix. En premier lieu seront présentés les défis associés aux pressions politiques. Seront ensuite abordées tour à tour les raisons qui ont amené les MO à utiliser les données sans restriction, avec restriction ou avec des mises en garde. La section se termine sur les décisions en suspens.

DIFFICILE DE FAIRE FI DES PRESSIONS

De par leur fonction, les analystes de données dans les MO doivent répondre aux besoins de leur organisation, que ce soit à des fins informationnelles ou décisionnelles. Or, mettre des nuances ou restreindre la diffusion ne rime pas toujours bien avec prises de décision, ou encore avec les attentes des utilisateurs de données. Par exemple, certains ont mentionné que même si des nuances sont apportées ou qu'ils ne diffusent pas eux-mêmes des données, les organisations partenaires peuvent profiter de la large diffusion de SC sur son site Web (produits normalisés gratuits). Ceux qui cherchent des données à tout prix puisent dans ces données gratuites sans tenir compte des avis émis par SC, l'Institut ou par les analystes des MO.

Au moins un MO a décidé de restreindre les données diffusées à la population, mais les analystes sont convaincus que pour des besoins *ad hoc* décisionnels, il sera plus difficile, voire impossible, de refuser de les fournir. Il nous a été souligné que pour répondre aux attentes, une organisation songe à procéder à certaines manipulations des données afin de combler les trous, avec les risques que cela comporte. D'autres coupent la poire en deux en assouplissant leurs critères de diffusion.

Comme dans le passé, les MO diffusaient sans restriction les données des recensements. Les utilisateurs, mais aussi les décideurs, ont de la difficulté à accepter que les règles aient changé. Il est difficile de comprendre que cela prenne plus de temps et de ressources pour valider les données. Les pressions sont particulièrement importantes lorsque les données servent au processus d'allocation de ressources. Remettre en question la qualité d'une donnée ou d'un indice met en avant-plan les désaccords entre les différents acteurs, alors que des consensus sur le choix des indices ou variables avaient pourtant été obtenus lors des travaux antérieurs.

Enfin, s'il est à propos d'émettre des recommandations aux utilisateurs sur le terrain, un intervenant a rappelé que le « paternalisme bon enfant » envers les organisations sur le terrain ne s'avère pas une avenue à retenir.

DIFFUSER SANS RESTRICTION

Pour des raisons différentes, quelques organisations ont choisi de rendre disponibles toutes les données compilées par SC, qu'elles proviennent de compilations spéciales ou non. Dans certains cas, cette décision a été prise, une fois que des analyses ont été faites, dans d'autres cas, on a plutôt choisi de faire confiance aux données fournies par SC. La majorité des organisations reproduisent cependant les mises en garde de SC.

Parmi les raisons évoquées pour diffuser sans restriction et sans analyse, une organisation a mentionné que puisqu'aucune comparaison n'est faite avec 2006, elle ne se sentait pas concernée par les changements de l'ENM. Ainsi, les données détaillées provenant de compilations semi-personnalisées ont toutes été rendues publiques. Une organisation envoie à ses partenaires des avis lorsque des problèmes sont rencontrés dans l'analyse des données provenant de compilations spéciales, mais aucune règle de diffusion n'est suggérée. On tient pour acquis que les personnes sur le terrain sont à même d'évaluer la qualité des données.

D'autres font entièrement confiance aux données publiées par SC, puisque leurs besoins se limitent à des échelles géographiques assez importantes. Des comparaisons avec 2006 sont aussi faites sans restriction. Une raison souvent mentionnée pour diffuser sans restriction tient au fait que les personnes qui ont besoin des données n'ont pas l'expertise pour juger de la qualité des données. On considère que si SC les publie ou les fournit, les données sont sûrement fiables. Enfin, d'autres arguent qu'en l'absence d'autres sources de données, on prend ce qui est disponible.

DIFFUSER AVEC RESTRICTION

Par ailleurs, plusieurs organisations ont choisi de limiter la diffusion des données. Par exemple, une organisation ne publie aucune donnée dont les effectifs estimés sont moins de 100 personnes. Une autre élimine toute publication pour des populations inférieures à 5 000 personnes et fait aussi moins de croisements lorsque la population estimée est moins de 50 000. Aucun croisement de variables dans un territoire de moins de 2 300 logements (ou 5 600 personnes) n'est autorisé dans une autre. Selon les domaines, certains se limitent aux données à l'échelle provinciale ou régionale, surtout lorsque l'analyse nécessite plusieurs croisements ou lorsque le phénomène étudié est rare. Ou encore, on limite les croisements lorsque les données sont à une échelle infrarégionale. Par exemple, des données pour les professions en deçà du CNP à 4 chiffres ne sont pas diffusées à l'échelle régionale. De regroupements de catégories sont aussi appliqués (ex : réparations mineures et majeures). Des décisions de ne pas diffuser les données sur la famille par MRC ont aussi été prises. Pour des données par municipalités, une organisation les diffuse au cas par cas.

L'Institut a également décidé de ne plus diffuser les données à l'échelle des municipalités qui composent chaque région administrative, parmi ses produits normalisés. De plus, certaines données régionales sont diffusées uniquement sous des thématiques particulières, une fois que la démonstration est faite que les données sont fiables.

La majorité des MO rencontrés respectent le seuil du TGN de 50 %. Certains proposent des seuils plus faibles ou conseillent à leurs partenaires régionaux de comparer les données démographiques de l'ENM et du Recensement de 2011 dès que le TGN dépasse 30 %. La majorité considère que l'ENM est une enquête de « Temps 1 » et que l'on ne doit pas comparer avec les recensements antérieurs. Toutefois, même dans les organisations qui partent avec cette prémisse de départ, certaines analyses de tendance sont réalisées lorsque toutes les validations le permettent ou encore lorsque les comparaisons se font sur des indices relatifs.

Soulignons également que certaines organisations, que ce soit à l'échelle du Québec ou à l'échelle régionale, mettent en commun leurs réflexions et leurs travaux.

DES DÉCISIONS EN ATTENTE

Au moment des rencontres, certains MO nous ont mentionné être toujours en réflexion ou faire des démarches de validation. Par exemple, un organisme régional a développé, pour l'ensemble des acteurs, des mégafichiers sur le marché du travail à partir des recensements antérieurs. La mise à jour de ceux-ci est remise en cause. D'autres attendent d'obtenir des réponses claires de SC sur la fiabilité de leurs données avant d'aller de l'avant sur toutes utilisations des données. Enfin, une autre organisation réfléchit sur la façon d'éduquer leurs usagers sur l'utilisation des données. En somme, la diffusion est retardée par rapport aux expériences passées avec le Recensement.

PEU DE PISTES DE SOLUTION POUR LES MO

Après le bilan effectué dans les sections précédentes, il est opportun de se questionner sur les pistes de solution. En effet, il appert clairement que l'ENM ne permet pas de remplir son objectif comme il est stipulé dans le guide de référence de l'ENM, soit de « recueillir des données sociales et économiques sur la population canadienne et de fournir des données à de petits niveaux géographiques [] et pour de petits groupes de population¹² ».

Malgré cette limite importante, plusieurs reconnaissent que l'ENM constitue tout de même une bonne enquête sociale. Mais comme pour toute enquête sociale, il s'agit d'une enquête volontaire avec les défis que cela entraîne, notamment du point de vue de la diffusion et des outils nécessaires pour l'analyse.

Les solutions présentées dans ce chapitre émergent des discussions que nous avons eues avec les organisations rencontrées. Il importe de souligner que celles-ci ne sont pas mutuellement exclusives, plusieurs avenues pouvant être nécessaires pour combler les besoins. Précisons qu'au début des rencontres, la décision du gouvernement fédéral de maintenir l'ENM pour 2016 n'était pas connue, certaines organisations espéraient le retour du questionnaire détaillé du Recensement. Toutefois, le 15 avril 2014, un communiqué émis par SC est venu confirmer les appréhensions en annonçant que la formule adoptée en 2011 (un questionnaire court obligatoire et une enquête détaillée volontaire) serait maintenue pour 2016. Par ailleurs, l'ensemble des modalités pour l'ENM et le Recensement 2016 ne sont pas connues pour le moment.

DES AJUSTEMENTS À COURT TERME

Dans le cas de l'indice de défavorisation sociale et matérielle, des méthodes sont explorées pour calculer des taux ajustés de 2006 aux fins de leur utilisation pour les données de 2011. La standardisation directe n'étant pas possible à utiliser en raison des règles de suppression des données, d'autres méthodes doivent donc être

testées (standardisation indirecte et estimation par le quotient). Les espoirs pour trouver une solution pour 2011 par cette voie s'avèrent possibles, car 95 % des aires de diffusion (AD) de 2011 sont identiques à celles de 2006, et que le caractère relatif de l'indice réduit les risques de mauvaise attribution d'une valeur de l'indice associée à des biais de non-réponse.

Après cet exercice, il est toutefois possible que certains regroupements soient nécessaires. En effet, les AD avec des TGN supérieurs à 50 % ne feront pas partie de la construction de la structure factorielle des indicateurs retenus pour l'indice de défavorisation. Pour répondre aux inquiétudes quant à la comparabilité et au risque de biais dans les AD, l'indice pourrait, après analyse, être réalisé à une échelle géographique plus élevée.

Toutefois, il est clair que cette solution n'est possible que pour l'année 2011. Pour les autres cycles du Recensement, tout doit être repensé. Le concept de défavorisation devra être revu, d'autres sources de données devront être exploitées et un processus de validation devra être instauré pour favoriser l'« acceptabilité sociale » de ce nouvel indicateur.

Comme souligné antérieurement, la possibilité d'utiliser les régions de pondération plutôt que les AD comme base géographique lorsque les AD ont des TGN plus élevés que 50 % s'avère une option intéressante afin d'obtenir des données plus fiables pour 2011 en regroupant des territoires. En raison des règles de confidentialité déterminées par SC, cette option n'a cependant pu être utilisée.

UTILISATION ET BONIFICATION D'AUTRES SOURCES DE DONNÉES

Une organisation souligne la possibilité de partager des données, d'enquêtes ou administratives, entre les divers MO. Cette option semble cependant difficile à réaliser,

12. STATISTIQUE CANADA (2013). *Guide l'utilisateur de l'ENM*. Enquête nationale auprès des ménages, n° 99-001-X2011001 au catalogue, p. 4.

notamment en raison des découpages géographiques différents, mais aussi en raison des contraintes liées à la confidentialité des données. Également, cette option ne permet pas de répondre à tous les besoins puisque les données requises ne sont tout simplement pas disponibles dans d'autres MO.

Dans le domaine de l'éducation, les formulaires d'inscription dans les écoles incluaient déjà la scolarité des parents comme variable. Toutefois, plusieurs personnes ont choisi de ne pas fournir cette information, la réponse étant volontaire. De plus, plusieurs parents se sont plaints quant à la pertinence de cette variable, même si la littérature montre bien le lien entre la réussite scolaire et la scolarité des parents (entre autres variables). Dans ces circonstances, cette variable a dû être retirée du formulaire. Celle-ci pourrait être réintroduite, mais cela impliquerait de longs délais en raison des consultations à mener et des démarches juridiques nécessaires pour obtenir un taux de complétion valable pour cette variable.

Des suggestions ont aussi été faites dans le domaine de l'emploi, comme inclure une variable CNP avec les données fiscales ou encore pour insérer la profession dans les données sur les départs à la retraite. Les données de l'EERH sont intéressantes, mais il n'y pas de données régionales. Des enquêtes québécoises, comme l'*Enquête sur le recrutement et l'emploi au Québec* (EREQ), pourraient aussi être utilisées, voire bonifiées.

D'autres ont exploré la possibilité d'utiliser les données des programmes d'assistance sociale comme mesure de la défavorisation, mais celles-ci constituent une mesure très imparfaite. Comme nous en avons discuté dans le chapitre sur la validation des données, plusieurs sources potentielles ont été examinées, notamment des données administratives (ex. : FIPA, évaluations foncières, etc.). Toutefois, si elles offrent certaines opportunités, elles posent plusieurs défis.

Une organisation de la région de Montréal qui produit un indice de défavorisation a choisi dès l'année 2011 d'utiliser une autre source de données que le Programme du Recensement pour le revenu. Comme son univers se limite aux familles avec enfants de niveau primaire et secondaire, elle a choisi d'utiliser les données provenant du Programme de soutien aux enfants de la Régie des rentes du Québec (RRQ). Soulignons, par ailleurs, que plusieurs organisations verraient d'un bon œil un plus grand recours aux données fiscales du Québec.

PRODUCTION DE NOUVELLES ENQUÊTES

L'option de combler les besoins régionaux ou locaux par des enquêtes particulières a rarement été évoquée comme piste de solutions, et lorsque celle-ci l'a été, les coûts trop importants ont rapidement été mis de l'avant. Par exemple, les autorités locales ou régionales qui ont des besoins de données pour des petites unités géographiques n'ont pas nécessairement les ressources financières pour parrainer une enquête nécessitant des échantillons importants. Rappelons que l'ENM ne s'est pas fait à un coût inférieur au questionnaire détaillé et obligatoire. De ce fait, procéder à une nouvelle enquête pour pallier les lacunes de ce changement revient à accroître les coûts pour obtenir des données auparavant disponibles. De plus, il n'est pas assuré qu'une enquête additionnelle dans un territoire donne de meilleurs résultats si celle-ci demeure de nature volontaire.

Quelques organisations ont soulevé l'hypothèse que l'Institut prenne en charge un recensement à l'échelle du Québec. Mais du même souffle, on reconnaissait qu'une telle option est particulièrement coûteuse et que l'Institut ne peut prendre seul une telle décision. À ces arguments s'ajoute celui de la responsabilité du gouvernement fédéral de mener des recensements de sa population en vertu de la Constitution canadienne.

AMÉLIORATION DU PROCESSUS DE DIFFUSION DES DONNÉES DE L'ENM

Par les critiques énoncées par les MO en ce qui concerne les outils à leur disposition avec l'ENM, on peut également conclure que certaines pistes de solution résident dans l'optimisation des données et métadonnées fournies par SC avec l'ENM. Si on résume, les demandes comprennent :

- la diffusion d'erreur type ou de CV, idéalement de façon concomitante à la sortie des données;
- des indications plus claires sur les biais potentiels;
- des analyses à des niveaux géographiques plus fins;
- un rapport technique produit plus rapidement;
- des regroupements de données qui permettent une plus grande fiabilité (ex. : en fonction des régions de pondération);
- une gestion des règles de confidentialité plus claire pour les utilisateurs.

LES PRINCIPAUX ENJEUX ET DÉFIS QUE REPRÉSENTE L'ENM

Les appréhensions, les constats et les attentes des MO à l'égard de l'ENM ont été présentés au fil des pages de ce rapport. Dans les lignes qui suivent, les enjeux les plus importants sont résumés.

UN PLUS LOURD FARDEAU POUR LES MO

À plusieurs reprises, les analystes rencontrés ont fait ressortir que le peu d'outils statistiques mis à leur disposition les oblige à prendre des décisions souvent à l'aveugle sur la qualité des données, seules leurs expériences et leurs connaissances antérieures étant garantes de la plausibilité des données. Ils sont aussi confrontés aux fortes pressions de la part des responsables de programmes et politiques pour lesquels les données sont essentielles à la prise de décision. Or, ces derniers ne comprennent pas toujours les implications du passage du questionnaire détaillé et obligatoire du Recensement à une enquête volontaire. Les situations qui en découlent, on s'en doute, ne sont pas toujours faciles.

Ce « fardeau de la preuve » s'ajoute à celui qui découle des heures plus nombreuses allouées à la validation des données. Ainsi, en plus de créer de l'insécurité, des heures supplémentaires en ressources humaines doivent être déployées. Ces coûts supplémentaires s'ajoutent donc aux coûts additionnels que SC a dû déboursier pour administrer l'ENM au lieu du questionnaire détaillé et obligatoire. Le seul point positif avancé en regard du fardeau financier est la diffusion plus étendue des données sur le site Web de SC, ainsi que la facturation moindre pour les commandes spéciales.

DES CONSÉQUENCES NÉGATIVES POUR LES POLITIQUES ET PROGRAMMES

Pour des sous-populations bien définies, ainsi que pour des unités géographiques fines, le questionnaire détaillé et obligatoire du Recensement constituait souvent la seule source de données. Or, c'est en regard de ces univers que les questions concernant la qualité des données ont

surtout été évoquées. Comment alors établir des profils de population par quartier? Comment cibler les milieux de garde défavorisés? Comment développer des schémas d'aménagement pour des petites villes? Comment a évolué le travail autonome dans le domaine de l'industrie culturelle? Ainsi, en plus de perdre la comparabilité, les responsables des politiques et des programmes doivent évaluer, élaborer et assurer des suivis, sans avoir la certitude que les données utilisées décrivent adéquatement la situation en 2011.

DES IMPASSES POUR LA PRISE DE DÉCISION

Les recensements antérieurs inspiraient grandement confiance. En effet, ces données permettaient d'objectiver les impressions sur le terrain. Or, les négociations entre les divers acteurs prennent d'abord appui sur des objectifs. Par exemple, si l'objectif est de réduire la pauvreté, on appuiera les décisions sur les faits observés, notamment les taux de faible revenu. Mais qu'arrive-t-il lorsque les acteurs doutent de la donnée sur le taux de faible revenu? Comme l'a précisé un intervenant rencontré, lorsque les acteurs se parlent, il faut que ce soit pour des enjeux et non pour des données.

Cette répercussion de l'ENM n'est pas anodine. En effet, dans certains processus d'allocation des ressources, des mois, voire des années, ont été nécessaires pour que les divers acteurs s'entendent sur les règles d'allocation des ressources et de suivi des indicateurs. Ce faisant, plusieurs MO devront probablement revoir ces règles et indicateurs. Donc, en plus des travaux de validation additionnels, ces derniers devront reprendre le bâton du pèlerin pour arriver à de nouveaux consensus. Cela ne va pas sans coûts. Plusieurs espéraient que l'ENM soit une exception et que le gouvernement fédéral reviendrait à la formule habituelle, à savoir, un Programme du Recensement impliquant un questionnaire court et obligatoire et un questionnaire détaillé et obligatoire pour 2016. Comme ce n'est pas le cas, des pistes de solutions doivent être dégagées.

DES OBJECTIFS NON ATTEINTS

D'emblée, précisons que pour la majorité des points de vue, l'ENM constitue une bonne enquête socioéconomique. Les taux de réponse plus élevés observés par rapport à d'autres enquêtes sociales en constituent une bonne illustration. Toutefois, contrairement à d'autres enquêtes sociales menées par SC, les données de l'ENM n'ont pas été diffusées accompagnées des outils statistiques appropriés.

Malgré l'envergure de cette enquête, son caractère volontaire rend plus hasardeux les conclusions pour des sous-populations ou des pour des unités géographiques fines. Or, l'objectif de cette enquête est précisément de fournir ces données. Pour des données à l'échelle du

Québec, plusieurs autres enquêtes sociales peuvent être mises à profit, mais à mesure que la taille ou la population étudiée diminue, les alternatives diminuent aussi.

Des craintes ont été aussi émises pour l'avenir. Qu'arrivera-t-il aux prochaines enquêtes dont le traitement était associé aux données du questionnaire détaillé et obligatoire du Recensement ?

En somme, comparativement au questionnaire détaillé et obligatoire des recensements antérieurs, l'ENM est plus coûteuse et les résultats obtenus sont de moins bonne qualité. Ainsi, la pertinence de revenir à la formule habituelle demeure.

Des statistiques sur le Québec d'hier et d'aujourd'hui pour le Québec de demain

En 2010, le gouvernement du Canada décide d'apporter des changements majeurs au Programme du Recensement de 2011 en remplaçant le questionnaire long et obligatoire du recensement par l'*Enquête nationale auprès des ménages* (ENM), une enquête volontaire.

Les premiers résultats de l'ENM ont été diffusés en 2013. À titre de coordonnateur statistique pour le Québec, l'Institut de la statistique du Québec a procédé à une large consultation auprès des ministères et organismes québécois afin de dresser un portrait précis des problèmes rencontrés, tant à l'égard des données qui ont été utilisées que des outils statistiques déployés.

Ce rapport, comprenant 9 sections, fait état des résultats de cette consultation. En plus de démontrer, une fois de plus, au gouvernement fédéral que l'abandon du caractère obligatoire du questionnaire long du recensement est lourd de conséquences, les informations qui y sont présentées offrent un appui pour la recherche de solutions et contribuent à alimenter les réflexions afin de permettre au Québec de disposer de l'information statistique essentielle à sa gouverne et à son développement.