



INSTITUT
DE LA STATISTIQUE
DU QUÉBEC

www.stat.gouv.qc.ca



Comparaison des
estimations sur le revenu
de l'*Enquête nationale
auprès des ménages*
avec les données fiscales
québécoises

Rapport méthodologique

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
3^e trimestre 2015
ISBN 978-2-550-73243-3 (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.

Août 2015

Cette publication a été réalisée par :

Stéphane Crespo
Agent de recherche socioéconomique

Direction des statistiques
sociodémographiques :

Paul Berthiaume, directeur

Avec la collaboration de :

Sylvie Rheault
Coordonnatrice des conditions de vie
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 Robert Courtemanche de la Direction de la méthodologie et de la qualité pour ses judicieux conseils au cours de cette démarche de validation, Stéphane Ladouceur de la Direction des statistiques sectorielles et du développement durable pour ses conseils sur l'analyse des données de Revenu Québec, et Sophie Desfossés de cette même direction pour les extractions de ces données. Nos remerciements vont enfin au CIQSS de l'Université Laval pour l'accès aux données de l'ENM de Statistique Canada.

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 : 418 961-2406
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

Ce rapport méthodologique ne fournit pas de seuils explicites pour distinguer des estimations de l'ENM selon leur niveau de qualité.

TABLE DES MATIÈRES

7	INTRODUCTION
9	1. PRÉSENTATION DES SOURCES DE DONNÉES
9	1.1 Spécifications des deux sources de données
9	1.1.1 Les données de l' <i>Enquête nationale auprès des ménages</i>
11	1.1.2 Les données de l'Agence du Revenu du Québec
12	1.2 Mesures prises pour améliorer les comparaisons entre les deux sources de données
13	1.3 Limites des comparaisons effectuées
17	2. MÉTHODOLOGIE
17	2.1 Le niveau de concordance : une fonction de la précision et des biais
18	2.2 Indicateurs méthodologiques
19	2.3 Les univers comparés
23	3. RÉSULTATS
23	3.1 Selon l'échelle géographique
25	3.2 Selon le sexe et l'âge
28	3.3 Selon l'indicateur de revenu
30	3.4 Selon le concept de revenu
33	4. RÉSUMÉ ET DISCUSSION

INTRODUCTION

En 2010, le gouvernement fédéral modifie son Programme du Recensement pour 2011¹ de Statistique Canada par l'abandon du caractère obligatoire du questionnaire long, et son remplacement par une nouvelle enquête à participation volontaire, l'*Enquête nationale auprès des ménages* (ENM). Dans le cadre des recensements précédents, le caractère obligatoire du questionnaire long permettait de minimiser le biais de non-réponse en maintenant des taux de réponse supérieurs à 95 %, le biais étant associé aux différences possibles entre les profils des répondants et des non-répondants. Au Québec, le caractère facultatif de la participation à l'enquête s'est traduit par un taux de réponse de 71,9% (Institut de la statistique du Québec, 2014, p. 4²), augmentant ainsi le risque de biais dû à la non-réponse. La pondération effectuée par Statistique Canada permet d'en réduire la portée, mais pas de l'éliminer, surtout pour des sous-univers ou des petits territoires.

Particulièrement pour la donnée sur le revenu, le caractère facultatif de l'ENM, combiné au taux plus élevé de non-réponse pour ces questions, a suscité encore plus de doutes sur la qualité de ces données³. Afin d'éclairer les utilisateurs, une comparaison des données sur le revenu de l'ENM est réalisée avec celles provenant d'un fichier administratif, soit celui des revenus fiscaux détenus par l'Agence de revenu du Québec (RQ).

Le rapport se divise en quatre parties. Les deux sources de données sont décrites dans la première partie. Dans la seconde partie est présentée la méthodologie ainsi que les univers comparés. Les résultats suivent en troisième partie. Les principaux constats terminent le rapport.

Ce rapport offre certains outils de validation aux utilisateurs des données de revenu de l'ENM pour le Québec. Toutefois, ce rapport ne fournit pas de seuils explicites pour distinguer des estimations dites de « bonne » ou de « mauvaise » qualité.

-
1. Le gouvernement fédéral maintient cette décision pour le prochain programme de Recensement (2016).
 2. INSTITUT DE LA STATISTIQUE DU QUÉBEC (2013), *L'Enquête nationale auprès des ménages de Statistique Canada : État des connaissances à l'intention des utilisateurs du Québec*, Québec, Institut de la statistique du Québec, 13 p.
 3. Voir la revue compilée par le Centre interuniversitaire québécois de statistiques sociales (CIQSS) : [En ligne]. [<http://www.ciqss.umontreal.ca/fr/recensement2011.html>].

1. PRÉSENTATION DES SOURCES DE DONNÉES

L'objectif de cette première partie consiste, dans un premier temps, à présenter les principales spécifications des deux sources de données. De celles-ci découlent, dans un deuxième temps, les décisions prises pour améliorer les comparaisons entre les deux sources de données. Dans un troisième temps, on énonce d'emblée certaines limites aux comparaisons effectuées.

1.1 SPÉCIFICATIONS DES DEUX SOURCES DE DONNÉES

Le tableau 1.1 résume les spécificités de ces deux sources de données. Celles-ci sont décrites ci-après.

1.1.1 Les données de l'Enquête nationale auprès des ménages

L'ENM de 2011¹ fournit plusieurs renseignements sur les caractéristiques démographiques, sociales et économiques des personnes vivant au Canada durant une semaine de référence de l'enquête, soit du 1^{er} au 7 mai 2011. Les sujets couverts par l'ENM incluent notamment le revenu des particuliers et des ménages privés².

L'ENM est une enquête transversale par échantillon, dont la participation est volontaire. Un échantillon aléatoire de 4,5 millions de logements, provenant des listes de logements du Recensement de 2011, est sélectionné; celui-ci représente près du tiers des ménages privés au Canada. Au Québec, l'effectif de l'échantillon des particuliers est d'environ 1 636 000 personnes.

La population de référence comprend l'ensemble des personnes de tous âges vivant habituellement au Canada (provinces et territoires). Les catégories suivantes de personnes sont admissibles à l'enquête: les personnes vivant dans les réserves indiennes et autres établissements indiens, les résidents permanents et les résidents non permanents³. Sont exclues les personnes vivant dans des logements collectifs institutionnels (hôpitaux, établissements de soins infirmiers, établissements pénitentiaires) ou dans des logements collectifs non institutionnels (camps de travail, hôtels et motels, résidences pour étudiants), les résidents étrangers, les membres des forces armées d'un autre pays basés au Canada, les résidents visitant temporairement le Canada, les Canadiens vivant à l'étranger et les membres à temps plein des Forces armées canadiennes en poste à l'extérieur du Canada.

Pour la variable du revenu, l'ENM a comme population de référence les particuliers âgés de 15 ans et plus, avec ou sans revenus. Le revenu total se rapporte à l'ensemble des revenus reçus durant l'année 2010: les salaires et traitements, 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, intérêts d'obligations, de dépôts et de certificats d'épargne et autres revenus de placements, les pensions de retraite et rentes, y compris les rentes de REER⁴ et de FERR⁵, les autres revenus privés en espèces⁶, et les transferts gouvernementaux, ces derniers comprenant la pension de la Sécurité de la vieillesse (SV) et du Supplément de revenu garanti (SRG), l'allocation au conjoint et l'allocation au survivant, la prestation du Régime de rentes du Québec (RRQ) ou du Régime de pensions du Canada (RPC), la prestation d'assurance-emploi, la prestation pour enfants, et les autres revenus provenant de sources publiques. Sont exclues les rentrées

1. [En ligne]. [http://www23.statcan.gc.ca/imdb/p2SV_f.pl?Function=getSurvey&SurvId=75585&Instald=75586&SDDS=5178].

2. On y aborde aussi les thèmes suivants: diversité ethnique et immigration, éducation, formation et apprentissage, familles, ménages et logements, langues, peuples autochtones, population et démographie, pensions, dépenses et richesse, société et communauté et travail.

3. Les résidents non permanents comprennent les revendicateurs du statut de réfugié, les détenteurs d'un permis de travail ou d'un permis d'études ainsi que les membres de leur famille vivant avec eux.

4. Régime enregistré d'épargne retraite.

5. Fonds enregistré de revenu de retraite.

6. Les six composantes de revenu que l'on vient d'énumérer forment un concept de revenu nommé le « revenu du marché ».

Tableau 1.1
Grille de spécifications relatives aux sources de données (ENM et RQ)

	ENM	RQ
Descriptions générales	Complément du recensement de 2011 qui vise à donner des renseignements sur les caractéristiques démographiques, sociales et économiques des personnes vivant au Canada	Fichier de données fiscales des particuliers provenant de Revenu Québec, et transmis à l'Institut de la statistique du Québec en vertu d'une entente spéciale
Méthodologies générales	Enquête transversale par échantillon	Recensement
Populations cibles	Ensemble des personnes de tous âges vivant habituellement au Canada dans des ménages privés (personnes en ménage collectif exclues), à l'exclusion, pour le Québec, des réserves indiennes de Wendake, de Kahnawake, d'Akwesasne, de Kanesatake, de Doncaster et de Lac-Rapide	Ensemble des personnes de 15 ans et plus (incluant les personnes vivant dans des ménages collectifs) ayant reçu (ou non) des revenus, à l'exception des personnes vivant dans les unités géographiques suivantes : <ul style="list-style-type: none"> - les communautés autochtones et les territoires non organisés de moins de 40 habitants - les réserves indiennes de Wendake, de Kahnawake, d'Akwesasne, de Kanesatake, de Doncaster et de Lac-Rapide - les communautés religieuses de Notre-Dame-des-Anges et de Saint-Benoît-du-Lac
Définitions du revenu (éléments inclus dans les deux sources)	Revenu de 2010 pour les particuliers de 15 ans et plus (avec ou sans revenus) : <ul style="list-style-type: none"> - salaires et traitements - revenu agricole net - revenu non agricole net provenant de l'exploitation d'une entreprise non constituée en société ou de l'exercice d'une profession - dividendes, intérêts d'obligations, de dépôts et de certificats d'épargne et autres revenus de placements - pensions de retraite et rentes, y compris les rentes du REER et du FERR - autres revenus privés en espèces - transferts gouvernementaux (pension de la Sécurité de la vieillesse et Supplément de revenu garanti, allocation et allocation au survivant, prestations du Régime de rentes du Québec ou du Régime de pensions du Canada, prestations d'assurance-emploi, prestations pour enfants, autre revenu provenant de sources publiques) 	
Définitions du revenu (éléments exclus des deux sources)	<ul style="list-style-type: none"> - rentrées d'argent ponctuelles, comme les gains de loterie, les gains de jeux, les héritages en argent, les règlements monétaires forfaitaires d'assurance - gains en capital - retraits d'un REER - cotisations des employeurs aux régimes de pension agréés, au Régime de pensions du Canada et au Régime de rentes du Québec et à l'assurance-emploi - transferts volontaires entre ménages - loyers imputés - biens et services produits pour le troc - biens produits pour consommation propre 	
Définitions du revenu (éléments exclus selon la source de données)		Crédits d'impôt remboursables fédéraux (prestation canadienne pour enfants, crédit TPS/TVH, prestation fiscale pour le revenu de travail) et prestations provinciales de soutien aux enfants
Mesures de contrôle de qualité des données sur le revenu	<ul style="list-style-type: none"> - utilisation des dossiers fiscaux fédéraux - procédures d'imputation de la non-réponse au revenu en particulier, et de la non-réponse en général 	<ul style="list-style-type: none"> - contrôles internes effectués par le RQ concernant les données de revenu - absence de traitement de la non-réponse totale - détection et correction, par l'ISQ, d'erreurs sur des variables auxiliaires comme l'âge
Caractéristiques de la codification géographique	La classification des MRC et celle des RA ont été reconstituées, car elle n'existaient pas par défaut	148 municipalités ont fait l'objet de 70 regroupements

Compilation : Institut de la statistique du Québec.

d'argent ponctuelles, comme les gains de loterie, les gains de jeux, les héritages en argent, les règlements monétaires forfaitaires d'assurance, les gains en capital et les retraits d'un REER, les cotisations des employeurs aux régimes de pension agréés, au Régime de pensions du Canada, au Régime de rentes du Québec et à l'assurance-emploi, les transferts volontaires entre ménages, les loyers imputés, les biens et services produits pour le troc et les biens produits pour consommation propre.

Des mesures de contrôle de qualité sont réalisées. Notamment, Statistique Canada demande aux répondants la permission d'utiliser leurs dossiers fédéraux de déclarations de revenus. Au Québec, plus des trois quarts des répondants (77,7 %) ont donné leur consentement. Pour les répondants n'ayant pas autorisé l'utilisation de ces dossiers, et n'ayant pas fourni d'information sur le revenu, des procédures d'imputation permettent de pallier la non-réponse sur les variables de revenu.

Finalement, sur le plan de la géographie, l'ENM ne contient pas la classification des *municipalités régionales de comté* (MRC) du Québec, qui consiste en 104 unités géographiques; mais seulement la classification des *Divisions de recensement* (DR) en 98 unités⁷. Cependant, la classification en MRC peut être dérivée en utilisant celle des DR et des *Subdivisions de recensement* (SDR), c'est-à-dire des municipalités. Par ailleurs, l'ENM ne contient pas non plus la classification en *Régions administratives* (RA); celle-ci peut être obtenue en utilisant celle des MRC.

Notons cependant qu'en raison d'une interdiction imposée par Statistique Canada, il n'est pas possible de divulguer⁸ toute estimation de l'ENM se rapportant à des unités géographiques dont la population est inférieure à 5 000 personnes⁹. La présentation des résultats a donc dû être ajustée.

1.1.2 Les données de l'Agence du Revenu du Québec

Les données de RQ sont issues des dossiers de déclaration de revenus à l'échelle provinciale. Ces données sont disponibles à l'Institut de la statistique du Québec en vertu d'une entente spéciale¹⁰. Afin d'assurer un arrimage aux données de l'ENM, le présent rapport exploite des données de l'année 2010, lesquelles ont fait l'objet de sélections et de traitements par l'Institut.

Ces données s'apparentent à un recensement, puisqu'en principe, tous les particuliers ayant produit une déclaration de revenus sont inclus. Le fichier de 2010 contient environ 6 305 000 observations.

La population de référence est plus inclusive que celle de l'ENM, puisqu'elle comprend non seulement les personnes vivant dans des logements privés, mais aussi celles vivant dans des logements collectifs.

Par ailleurs, ces données ne peuvent pas être considérées comme représentatives des personnes âgées de moins de 18 ans, puisque les taux de couverture à cet âge sont faibles ou nuls. Par exemple, entre 15 et 17 ans, ces taux oscillent autour de 10 %.

Enfin, ces données fiscales doivent être considérées comme plus représentatives de la population *avec revenus*, que de la population *en général*, c'est-à-dire *avec* ou *sans* revenus. En effet, toute personne n'ayant pas reçu des revenus en 2010 peut choisir de ne pas produire de déclaration de revenus, quoique le régime fiscal encourage la production d'un tel rapport afin de bénéficier de certains crédits d'impôt (ex. : crédit d'impôt de solidarité). De fait, la population des personnes âgées de 18 ans et plus sans revenus est sous-estimée avec les données fiscales. En 2010, on évalue la population générale des personnes âgées de 18 ans et plus (c.-à-d. avec et sans revenus) à environ 6 406 000 habitants¹¹, alors qu'avec les données de RQ, on obtient 6 197 000 personnes, ce qui représente un déficit de 3,3 %.

7. Les différences entre la classification en DR et la classification en MRC sont les suivantes. La DR Francheville (37) contient les MRC Trois-Rivières et Les Chenaux. La DR Le Saguenay-et-son-Fjord (94) contient les MRC de Saguenay et de Fjord-du-Saguenay, celle de Sept-Rivières-Caniapiscou (97) contient les MRC de Sept-Rivières et de Caniapiscou, celle de Minganie-Le Golfe-du-Saint-Laurent (99) contient les MRC de Minganie et Le Golfe-du-Saint-Laurent, et enfin, celle du Nord-du-Québec (99) contient les MRC de Jamésie, de Kativik et d'Eeyou Istchee. À part ces cas, chacune des autres DR correspond exactement à une MRC.

8. Cette interdiction s'applique aux centres de données de recherche (CDR), dont le Centre interuniversitaire québécois de statistiques sociales (CIQSS) fait partie, et où les données du présent rapport ont été compilées.

9. Il s'agit pour ce critère de l'ensemble de la population des personnes de tous âges.

10. Cette entente permet notamment de développer de nouveaux indicateurs de revenu qui tiennent compte de plusieurs échelles géographiques, dont les municipalités. Notons que cette même entente a permis à l'Institut, après validation auprès des juristes de RQ, d'utiliser ces données pour les comparer à celles de l'ENM.

11. [En ligne]. [http://www.stat.gouv.qc.ca/statistiques/population-demographie/structure/QC_age_et_sexe.xlsx].

1.2 MESURES PRISES POUR AMÉLIORER LES COMPARAISONS ENTRE LES DEUX SOURCES DE DONNÉES

Afin de tenir compte de cette sous-évaluation des personnes sans revenu et des personnes de moins de 18 ans avec les données fiscales, aux fins de comparaison, les données retenues pour l'ENM et pour RQ ne comprennent que la population de 18 ans et plus ayant un revenu¹².

D'autres mesures ont été prises pour améliorer les comparaisons entre les deux sources de données. L'Institut s'est efforcé de rendre compatibles les définitions du revenu. Cette tâche est relativement difficile, notamment pour deux raisons. Premièrement, chaque source de données émane d'un univers fiscal différent (fédéral pour l'ENM¹³ et provincial pour RQ), ce qui rend difficile la reconstitution des composantes de revenu accessibles dans l'ENM (salaires et traitements, revenu agricole net, revenu non agricole net, etc.) à partir des éléments de revenu détaillés qui correspondent aux lignes de la déclaration de revenus. Dans ces circonstances, la prudence est de mise pour toute comparaison ENM-RQ effectuée à partir de chacune de ces grandes composantes (voir section 3.4), plus spécialement les composantes de type résiduel, comme les autres revenus (privés) en espèce¹⁴.

Deuxièmement, les données de RQ à notre disposition n'ont pas permis de comptabiliser tous les éléments entrant dans la définition standard du revenu. Pour commencer, puisque l'impôt fédéral n'est pas mesuré dans ces données, il n'a pas été possible d'effectuer une comparaison de revenu ENM-RQ pour le revenu disponible ou après impôt. Cette comparaison s'est plutôt concentrée sur le concept de revenu total. De plus, tous les éléments entrant dans la définition standard du revenu total ne sont pas comptabilisés dans ces données de RQ. Ainsi, les crédits d'impôt remboursables fédéraux – composés principalement du crédit d'impôt pour TPS/TVH, de la prestation fiscale pour le revenu de travail et de la prestation fiscale canadienne pour enfants (PFCE) – ainsi que

les prestations du programme provincial de soutien aux enfants sont exclus. Comme le niveau de précision de la classification du revenu de l'ENM ne permettait pas de retrancher un à un ces éléments, nous avons exclu de cette enquête au moins les montants de la PFCE et du programme provincial de soutien aux enfants en retranchant les montants de la sous-composante « prestations pour enfants ». Cette sous-composante comprend bien sûr la PFCE et les prestations pour enfants offertes par les provinces, mais elle comprend aussi la Prestation universelle pour la garde d'enfants (PUGE), un élément qui demeure pourtant comptabilisé dans les données de RQ. Donc, afin ne pas nuire aux comparaisons, les montants de la PUGE ont aussi été retirés des données de RQ.

Il est important de signaler que le concept final de revenu total utilisé dans le présent rapport a une valeur strictement analytique, celle de favoriser dans la meilleure mesure possible les comparaisons sur le revenu entre les deux sources de données. Comme les montants mesurés par ce concept alternatif sont nécessairement sous-évalués en raison des exclusions susmentionnées, aucune estimation de l'ENM publiée dans le présent rapport (voir le tableau 2.1 présenté plus loin) ne peut remplacer les estimations officielles tirées de cette enquête.

Les données ont fait l'objet de contrôles internes par RQ. Aussi, l'Institut a procédé à une détection et à une correction d'erreurs sur des variables auxiliaires aux données sur le revenu comme l'âge. Signalons que l'âge est mesuré au 31 décembre 2010. Pour assurer la compatibilité dans les mesures, cette variable a été mesurée à la même date dans l'ENM, en utilisant la date exacte de naissance.

12. Dans les données de l'ENM comme de RQ, on considère qu'une personne est avec revenus lorsqu'elle est dans l'une ou l'autre des deux situations suivantes : 1) son revenu total est non nul ; 2) son revenu total est nul, mais au moins deux composantes du revenu parmi les suivantes sont non nulles : a) les salaires et traitements ; b) le revenu agricole net ; c) le revenu non agricole net ; d) les dividendes, intérêts d'obligations, de dépôts et de certificats d'épargne et les autres revenus de placements ; e) les pensions de retraite et rentes ; f) les autres revenus privés en espèces ; g) les transferts gouvernementaux.
13. En effet, les données de l'ENM sur le revenu sont issues de l'univers de la fiscalité fédérale, d'autant plus que, comme mentionné plus haut, presque quatre répondants sur cinq ont accepté que leurs informations sur le revenu soient puisées à leurs déclarations de revenus.
14. Dans l'ENM, cette catégorie revêt un aspect résiduel important, de telle sorte qu'il est difficile, même en utilisant les définitions détaillées du dictionnaire de l'ENM ([En ligne]. <http://www12.statcan.gc.ca/nhs-enm/2011/ref/dict/index-fra.cfm>), de parvenir à identifier exhaustivement les lignes de la déclaration de revenus de RQ qui recouvrent les revenus en faisant partie.

Aussi, dans le but de satisfaire des règles de confidentialité relatives aux données de RQ, les personnes habitant dans les municipalités suivantes¹⁵ ont été exclues des données de RQ : les communautés autochtones et les territoires non organisés de moins de 40 habitants, les réserves indiennes de Wendake, de Kahnawake, d'Akwesasne, de Kanasatake, de Doncaster et de Lac-Rapide, et les communautés religieuses ayant un statut de municipalité que sont Notre-Dame-des-Anges et Saint-Benoît-du-Lac. Ces personnes ont aussi été exclues des données de l'ENM¹⁶ afin d'assurer la compatibilité des univers. Cette exclusion constitue une cause supplémentaire¹⁷ expliquant que les estimations de l'ENM traitées dans ce rapport diffèrent des estimations standards du revenu total des personnes de 18 ans et plus avec revenus.

Au niveau de la géographie de l'ENM, et toujours afin d'assurer une compatibilité avec les données de RQ, la région de résidence en 2010 a été utilisée, et non celle de la période de référence de l'enquête (2011)¹⁸.

Enfin, signalons une particularité de la codification géographique des municipalités utilisée dans le présent rapport. Un certain nombre de municipalités, de petit effectif et situées en milieu rural pour la plupart, ont fait l'objet de regroupements. En 2010, 148 municipalités ont fait l'objet de tels regroupements, pour atteindre 70 au total¹⁹. Ces regroupements se justifient par le fait que les municipalités impliquées présentent (avant les regroupements) des variations trop importantes dans leur taux de couverture, c'est-à-dire dans le pourcentage de la population des personnes âgées de 15 ans et plus ayant fait une déclaration de revenus. L'absence de ces regroupements aurait donc fait craindre des estimations de revenu biaisées pour notre propre population de référence dans les données de RQ (les 18 ans et plus avec revenus), ce qui aurait pu nuire aux comparaisons ENM-RQ. En conséquence, toute comparaison effectuée aux échelles géographiques des municipalités en tient compte. Au total, après les exclusions susmentionnées et après ces regroupements, 1 087 municipalités sont disponibles pour l'année 2010.

1.3 LIMITES DES COMPARAISONS EFFECTUÉES

Puisque l'objectif de ce rapport est de comparer les données de RQ à celles de l'ENM, il importe de réduire le plus possible les différences conceptuelles entre ces données et celles de RQ.

Comme discuté précédemment, seules sont retenues les personnes de 18 ans et plus ayant un revenu, tant pour les données provenant de l'ENM que de RQ, afin de réduire les biais possibles des données de RQ associés à l'absence de déclaration de revenus pour les personnes dont le revenu est nul et à la sous-couverture des personnes de moins de 18 ans. Toutefois, il n'est pas possible ici d'évaluer la sous-couverture fiscale parmi les personnes à revenu positif, mais faible, par exemple les personnes âgées qui comptent uniquement sur des prestations de la SV et du SRG, ou encore certains jeunes âgés de 18 à 24 ans. La sous-représentation dans les données de RQ de ces personnes paraît envisageable, malgré les incitatifs du régime fiscal pour produire une déclaration (ex. : crédit d'impôt pour la solidarité).

Cela étant dit, la distribution de l'effectif selon l'âge est relativement semblable d'une source de données à l'autre (cf. tableau 1.2). Tout au plus, on constate une légère sous-représentation des personnes de 65 ans et plus dans les données de l'ENM par rapport à celles de RQ. En effet, les personnes de 65 ans et plus représentent 18,4 % des 18 ans et plus avec revenus avec les données de l'ENM, contre 19,8 % pour les données de RQ, un déficit de 1,5 point de pourcentage. Ce déficit s'explique vraisemblablement par le fait que les ménages collectifs, qui sont surtout constitués de personnes de 65 ans et plus, sont exclus de l'ENM, et non par une sous-couverture des personnes âgées dans les données de RQ. Quant aux personnes de 18 à 24 ans, elles constituent 11,0 % des personnes de 18 ans et plus avec revenus dans l'ENM, contre 10,4 % dans les données de RQ, un excédent d'environ un demi-point. Si la sous-représentation de ces personnes dans les données de RQ est plausible, celle-ci n'apparaît pas élevée.

15. Par ailleurs, puisque les estimations de population pour ces municipalités sont moins précises, il est plus difficile d'évaluer le taux de couverture de la population des déclarants dans les données de RQ. Donc, il est plus difficile d'évaluer la qualité de ces données.

16. Notons que les répondants de ces deux dernières municipalités sont, de toute manière, exclus en principe de l'ENM puisque les communautés religieuses constituent des ménages collectifs. Par ailleurs, bien que les personnes vivant dans les réserves indiennes susmentionnées n'aient pas participé à l'ENM, on constate dans l'échantillon la présence de certaines personnes qui habitaient ces réserves en 2010, mais non en 2011 au moment de l'enquête. Ces personnes ont donc été exclues.

17. Comme discuté plus haut, l'autre raison consiste dans le fait que le revenu total ne correspond plus au concept standard.

18. Rappelons que le revenu est mesuré en 2010.

19. Notons que tous ces regroupements ont été effectués entre municipalités voisines, et situées dans les mêmes MRC.

Tableau 1.2

Effectifs des personnes de 18 ans et plus avec revenus selon la source de données (ENM et RQ), le sexe et l'âge, ensemble du Québec, 2010

	Fréquences absolues				Fréquences relatives		
	ENM	RQ	Écart ENM – RQ		ENM	RQ	Écart ENM – RQ
			Relatif ¹	Absolu ²			Absolu ²
	n		%	n	%	point de %	
Ensemble des personnes de 18 ans et plus avec revenus	5 956 210	6 072 700	–1,9	–116 490	100,0	100,0	...
Sexe							
Femmes	2 998 235	3 060 035	–2,0	–61 800	50,3	50,4	–0,1
Hommes	2 957 975	3 012 665	–1,8	–54 690	49,7	49,6	0,1
Âge							
18-24 ans	652 925	633 740	3,0	19 185	11,0	10,4	0,5
25-44 ans	1 955 735	1 984 919	–1,5	–29 184	32,8	32,7	0,1
45-64 ans	2 252 585	2 248 803	0,2	3 782	37,8	37,0	0,8
65 ans et plus	1 094 965	1 205 238	–9,1	–110 273	18,4	19,8	–1,5

 1. L'écart relatif se calcule de la manière suivante : $[(\text{donnée ENM} - \text{donnée RQ}) \div \text{donnée RQ}] \times 100$.

 2. L'écart absolu se calcule de la manière suivante : $[\text{donnée ENM} - \text{donnée RQ}]$.

Source : Enquête nationale auprès des ménages de Statistique Canada et Revenu Québec, 2010, adapté par l'Institut de la statistique du Québec.

Puisque la distribution des effectifs selon l'âge dans l'ENM est calibrée selon les données du Recensement, le caractère relativement semblable des deux distributions suggère que la couverture selon l'âge, dans les données de RQ, est bonne. Les données du tableau 1.2 quant à la distribution selon le sexe conduisent aux mêmes conclusions.

En définitive, si un biais quelconque est présent dans les données de RQ, celui-ci ne devrait pas être important, pour deux raisons : 1) en excluant les personnes ayant eu un revenu nul et les personnes de moins de 18 ans, ces données s'apparentent à un recensement, et donc la couverture de la population doit être satisfaisante ; 2) les déclarations font l'objet de vérifications internes par RQ, ce qui minimise les erreurs de mesure.

Cela dit, signalons que certaines caractéristiques des données fiscales peuvent nuire aux comparaisons entre les deux sources de données. Ainsi, les données de RQ comprennent les personnes vivant en logement collectif, contrairement à l'ENM, ce qui d'ailleurs contribue à expliquer la sous-représentation des 65 ans et plus dans

les données de l'ENM comparativement aux données de RQ. Faute d'informations sur le type de logement, il n'a pas été possible de les exclure de ces dernières données, afin d'assurer la pleine comparabilité. Il devient alors important d'interpréter les résultats sur la concordance ENM-RQ avec prudence, particulièrement chez les personnes âgées.

Aussi, Statistique Canada a mis en garde les utilisateurs de l'ENM sur le fait que ses définitions du revenu – en l'occurrence celles utilisées dans cette enquête – ne correspondent pas *toujours*²⁰ aux concepts utilisés par d'autres organismes. Nous avons décrit dans la section précédente les opérations effectuées pour maximiser la compatibilité des définitions du revenu total entre les deux sources de données. Néanmoins, il faut garder à l'esprit que cet objectif ne peut être parfaitement accompli. Par exemple, rien n'a pu être fait, compte tenu de l'information disponible dans l'ENM, pour retirer de la définition du revenu total les montants du crédit d'impôt pour TPS/TVH et la prestation fiscale pour le revenu de travail, lesquels ne sont pas compris dans les données de RQ.

20. Cet accent mis sur le mot « toujours » est nôtre.

On doit aussi signaler l'absence d'une correspondance exacte dans la localisation géographique des personnes entre les deux sources de données. Par exemple, l'adresse postale communiquée à RQ aux fins de l'impôt risque de ne pas toujours correspondre à l'adresse effective de la personne, du moins l'adresse donnée à l'ENM²¹. Toutefois, le risque de non-correspondance diminue à mesure que l'échelle géographique est plus large, comme c'est le cas des MRC par rapport aux municipalités, et des RA par rapport aux MRC.

Somme toute, en dépit des quelques problèmes de comparabilité que nous venons d'énumérer, les données fiscales constituent une source de données fiables pour effectuer une comparaison avec les estimations de l'ENM. Donc, en présence d'un manque de concordance entre les données imputables à un biais, on considérera, sauf mention contraire, que ce dernier provient plutôt de l'ENM.

21. En effet, lorsque ce n'est pas le contribuable lui-même qui effectue la déclaration de revenus, mais un tiers comme un comptable ou une agence de services comptables, il peut arriver que l'adresse postale soit plutôt celle de ce tiers.

2. MÉTHODOLOGIE

Dans un premier temps sont définies les causes possibles d'écart entre les données de l'ENM et de RQ. Dans un deuxième temps sont décrits les indicateurs utilisés pour mesurer le niveau de concordance entre ces deux sources de données. L'univers des comparaisons est défini dans un troisième temps.

2.1 LE NIVEAU DE CONCORDANCE : UNE FONCTION DE LA PRÉCISION ET DES BIAIS

Deux facteurs sont susceptibles d'influencer le niveau de concordance ou d'écart entre les données de l'ENM et de RQ sur le revenu : la précision et les biais présents dans les données de l'ENM.

La **précision** est associée à la variation échantillonnale d'une enquête. Cette variation échantillonnale est fonction de la taille de l'échantillon, qui elle-même dépend de l'effectif de la population considérée. Or, cet effectif peut varier selon l'échelle géographique ou les univers étudiés (ex. : âge spécifique). Donc, plus l'effectif de la population de référence est petit, plus l'échantillon de l'ENM l'est aussi, et plus l'estimation du revenu est assortie d'une erreur aléatoire. Dans la mesure où les données de RQ correspondent à des données de recensement, les écarts entre les données de RQ et de l'ENM peuvent nous éclairer sur la précision des estimations de l'ENM.

Les **biais** ne découlent pas, quant à eux d'une erreur aléatoire, mais plutôt d'une erreur systématique. Dans le cas de l'ENM, et des enquêtes volontaires de façon générale, l'une des principales préoccupations quant au biais est reliée au taux de non-réponse à l'enquête, et aussi aux taux de réponses partielles à certaines questions. Le principal biais peut provenir de la différence entre les caractéristiques des répondants et des non-répondants. Plus les taux de réponse sont faibles, plus

le risque de biais est élevé, et ce, malgré les procédures comme l'imputation et la correction par la pondération, etc. Aussi, parmi les sources de biais, il faut compter les potentielles erreurs de mesure des caractéristiques qui font l'objet de l'enquête.

Les données de l'ENM sont assorties d'un indicateur de qualité de données qui, calculé à différentes échelles géographiques, permet de chiffrer le potentiel de biais dû à la non-réponse. Cet indicateur se nomme le « taux global de non-réponse » ou TGN. Statistique Canada le définit et en prescrit l'usage comme suit¹ : « *Pour les estimations de l'Enquête nationale auprès des ménages (ENM) de 2011, le taux global de non-réponse (TGN) est utilisé comme indicateur de la qualité des données. Cet indicateur combine la non-réponse totale (ménage) et la non-réponse partielle (question) en un seul taux. La valeur du TGN est présentée aux utilisateurs. Un TGN plus faible indique un risque peu élevé de biais dû à la non-réponse et par conséquent, un risque moins élevé d'imprécisions². Le seuil utilisé pour la suppression des estimations est un TGN de 50 % ou plus.* » Ce seuil est appliqué pour les publications normalisées de Statistique Canada.

Comme il a été explicité précédemment, les biais, s'ils existent, seraient minimes avec les données de RQ. Par contre, le caractère volontaire de l'ENM accroît les risques de biais. Ainsi, les écarts entre les données de RQ et les estimations de l'ENM peuvent illustrer l'importance des biais des estimations de l'ENM.

1. [En ligne]. [<http://www12.statcan.gc.ca/nhs-enm/2011/dp-pd/prof/help-aide/gnr-tgn.cfm?Lang=F>].

2. Le terme utilisé dans la version anglaise de cette citation – *inaccuracy* – nous semble plus approprié que le terme « imprécision » pour décrire la notion de biais sous-tendue par le TGN. En effet, le terme « imprécision » devrait référer davantage à des problèmes d'erreurs dues à l'échantillonnage que d'erreurs dues à des biais. Le terme anglais serait mieux traduit par « inexactitude ».

2.2 INDICATEURS MÉTHODOLOGIQUES

Une conséquence, dans l'ENM, d'une précision insuffisante ou de la présence de biais, est la manifestation d'un écart entre les estimations de cette enquête et les données de RQ. Deux indicateurs sont utilisés pour mesurer un tel écart ou le niveau de concordance. Il s'agit du coefficient de corrélation et de la distribution de l'écart.

Le **coefficient de corrélation** est une mesure³ qui indique jusqu'à quel point le lien qui unit deux séries de nombres est linéaire, donc concordant. Par exemple, imaginons deux séries de données où les résultats sont les mêmes pour chaque municipalité. Dans ce cas, la concordance est parfaite et la corrélation est égale à « 1 ». Plus les écarts s'agrandiront, plus la corrélation diminuera et tendra vers « 0 ».

La corrélation comporte toutefois deux limites. Premièrement, on peut difficilement associer une « valeur » du coefficient de corrélation à un « niveau » de concordance ou de discordance. Par exemple, il est difficile de se représenter ce qu'est un niveau de concordance correspondant à une corrélation de 0,78. On peut en revanche conclure aisément qu'une corrélation de 0,98 correspond à un niveau de concordance supérieur à celui d'une corrélation de 0,78. C'est pourquoi on utilisera cet indicateur non pas en absolu, mais plutôt en termes de tendance.

Deuxièmement, si une concordance parfaite implique une corrélation de 1, le contraire n'est pas nécessairement vrai. En effet, deux séries de données pourraient être parfaitement corrélées, mais ne pas concorder. Reprenons l'exemple précédent des deux séries de données par municipalités. Si l'on ajoute 10 000 \$ à une des deux séries de données, on obtient encore une corrélation parfaite, mais les données ne concordent plus. Cette limite ne permet donc pas d'évaluer si une source de données, en l'occurrence l'ENM, sous-évalue ou surévalue d'une manière relativement systématique le revenu. En d'autres termes, le coefficient de corrélation ne peut détecter certaines formes de biais donnant lieu à des « différences de niveaux » entre les deux sources de données.

Voilà pourquoi un autre indicateur, permettant de cerner la concordance en général, mais aussi ces différences de niveaux en particulier, est également utilisé. Il s'agit de la **distribution de l'écart** entre les données de l'ENM et celles de RQ. Sont calculés selon le cas les écarts absolus (en points de pourcentage) ou les écarts relatifs (en pourcentage)⁴. Les différentes catégories utilisées sont les suivantes :

- 20 et plus ;
- 20 à -10 ;
- 10 à -5 ;
- 5 à -2,5 ;
- 2,5 à 0 ;
- 0 à 2,5 ;
- 2,5 à 5 ;
- 5 à 10 ;
- 10 à 20 ;
- 20 et plus.

Lorsque les écarts sont négatifs, les estimations de l'ENM sont inférieures aux données de RQ, et inversement, les écarts positifs indiquent que les estimations de l'ENM sont supérieures aux données de RQ. Aussi, plus l'écart se situe entre -2,5 et +2,5, plus la concordance est grande et plus la différence de niveaux est faible. À l'inverse, lorsque dans un ensemble d'unités géographiques, les écarts ne sont pas centrés autour de zéro, on peut conclure à une différence de niveaux. Par exemple, si 100 % des écarts relatifs sont répartis entre +5 % et +10 % ou +10 % et +20 %, il y a clairement une différence de niveaux entre les données des deux sources.

3. On utilise ici la corrélation standard de Pearson. Voir le lien suivant pour une définition technique : [En ligne]. [http://en.wikipedia.org/wiki/Pearson_product-moment_correlation_coefficient].

4. On définit ici un écart absolu de la manière suivante : [donnée ENM – donnée RQ]. L'écart relatif se définit en pourcentage comme suit : [(donnée ENM – donnée RQ) ÷ donnée RQ] × 100.

2.3 LES UNIVERS COMPARÉS

Plusieurs univers sont comparés, soit selon l'échelle géographique, certaines populations, des indicateurs et des concepts de revenu.

D'abord, l'**échelle géographique** renvoie à des classifications du territoire québécois. Les quatre échelles suivantes sont considérées, de la plus agrégée à la moins agrégée : l'*ensemble du Québec*, les *régions administratives (RA)*, les *municipalités régionales de comté (MRC)* et les *municipalités*. En 2010, on compte 17 RA et 104 MRC. On compte également 1 165 municipalités après exclusion de celles non considérées dans les données de RQ⁵, et 1 087 après que 148 d'entre elles sont regroupées en 70 (cf. section 1.2).

Pour les échelles géographiques des MRC et des municipalités, on considère aussi les deux axes suivants : un qui ventile les unités géographiques selon des tranches d'*effectif de la population de référence* et un autre qui les ventile selon des tranches de *taux global de non-réponse (TGN)* à l'ENM.

L'effectif de la population de référence d'une unité géographique renvoie au nombre de personnes de cette unité, comme estimé dans l'ENM, qui appartiennent à une population de référence donnée. Quant au TGN, l'Institut dispose de ses valeurs pour toutes les unités géographiques⁶ à toutes les échelles. Celles-ci ont été regroupées en six tranches potentielles :

- 50 % et plus ;
- entre 40 % et 50 % ;
- entre 30 % et 40 % ;
- entre 20 % et 30 % ;
- entre 10 % et 20 % ;
- moins de 10 %.

Les sept **populations de référence** suivantes sont utilisées : l'*ensemble des particuliers âgés de 18 ans et plus avec revenus*, les particuliers de 18 ans et plus avec revenus selon le *sexe* (hommes, femmes) et les particuliers de 18 ans et plus avec revenus selon l'*âge* (18 à 24 ans, 25 à 44 ans, 45 à 64 ans, et 65 ans et plus).

Les quatre **indicateurs de revenu** suivants sont choisis : le *revenu médian*, le *revenu moyen*, la *proportion de la population de référence dont le revenu est inférieur à 10 000 \$*, et la *proportion dont ce revenu est supérieur ou égal à 100 000 \$*. Ces indicateurs ont été choisis étant donné leur caractère standard en statistique descriptive du revenu⁷. La médiane se définit comme la valeur de revenu où la moitié de la population a un revenu inférieur à celui-ci et l'autre moitié un revenu supérieur. La médiane constitue un indicateur de revenu insensible aux valeurs extrêmes⁸. Par contraste, la moyenne est un indicateur de revenu plus sensible aux valeurs extrêmes. Enfin, les proportions de la population dont le revenu est inférieur à 10 000 \$, ou supérieur ou égal à 100 000 \$, constituent des indicateurs complémentaires de revenu. Rappelons que tous ces indicateurs sont calculés dans la population des personnes âgées de 18 ans et plus avec revenus. Une personne est considérée avec revenus si ses revenus sont différents de « 0 ». Autrement dit, lorsqu'une personne est avec revenus, son revenu peut être positif, mais il peut être aussi négatif, c'est-à-dire déficitaire.

Enfin, les neuf **concepts de revenu** suivants sont retenus : le *revenu total*, le *revenu du marché*, les six *composantes du revenu de marché* (cf. section 1.1.1) et les *transferts gouvernementaux*. Le revenu total est la somme des revenus du marché et des transferts gouvernementaux.

Le nombre d'unités de comparaison étant élevé, le nombre potentiel d'analyses croisées l'est aussi. Toute analyse sur la concordance repose en effet sur des statistiques

5. Rappelons qu'il s'agit des suivantes : les communautés autochtones et les territoires non organisés de moins de 40 habitants, les réserves indiennes de Wendake, de Kahnawake, d'Akwesasne, de Kanesatake, de Doncaster et de Lac-Rapide, et les communautés religieuses ayant un statut de municipalité que sont Notre-Dame-des-Anges et Saint-Benoît-du-Lac.

6. Notons que ces TGN des unités géographiques sont calculés en référence à la population totale (c.-à-d. les personnes de tous âges) et non pour une sous-population quelconque.

7. Le revenu médian de la population, le revenu moyen de la population, de même que la distribution de la population selon les tranches de leurs revenus font partie des indicateurs de revenu de la plupart des produits normalisés de Statistique Canada, comme ceux de l'ENM (ex. : produit 99-014-X2011016).

8. Dans le cas du revenu, les valeurs extrêmes sont généralement de très hauts revenus (c.-à-d. des revenus positifs très élevés), et moins souvent de très hauts revenus déficitaires (c.-à-d. des revenus négatifs très élevés).

calculées à partir d'une matrice de données résultant du croisement entre une échelle géographique, une population de référence, un indicateur de revenu et un concept de revenu. Des choix de croisements ont donc été faits. Les tableaux 3.1 à 3.4 de la prochaine partie, sur lesquels se concentre notre analyse, consignent les statistiques de concordance des croisements retenus.

Les principales matrices de données relatives à ces statistiques, quant à elles, sont reproduites dans des fichiers Excel accompagnant le présent rapport. Signalons qu'en raison de l'interdiction, imposée par Statistique Canada, de divulguer depuis les centres de données de recherche (CDR) toute estimation de l'ENM se rapportant à des unités géographiques dont la population est inférieure à 5 000 personnes⁹, et puisque l'objectif du présent rapport n'est pas de fournir systématiquement des estimations, aucune de ces matrices n'en contient. Toutefois, en guise de référence, le tableau 2.1 les fournit à l'échelle de l'ensemble du Québec¹⁰.

Ces matrices contiennent en lignes les unités géographiques à une échelle donnée, et en colonnes les informations suivantes: le code d'unité géographique, le libellé de cette unité, la tranche d'écart dans laquelle l'unité se situe, sa tranche de TGN et sa tranche d'effectif de population de référence. Les fichiers comprennent respectivement les matrices se rapportant à l'échelle géographique (fichier A1), à la population de référence (fichier A2), et à l'indicateur de revenu (fichier A3)¹¹. Les feuilles de chacun de ces fichiers sont nommées par un code (ex. : « A1-1 », « A2-4 », etc.). D'ailleurs, chacune des statistiques de concordance ou d'écart des tableaux 3.1 à 3.3 réfère à un tel code (voir la colonne de droite des tableaux). Ainsi, pour accéder aux données plus détaillées sur lesquelles une statistique a été calculée, il s'agit de repérer son code de référence et d'accéder à la feuille de données correspondante.

9. Il s'agit de la population de tous âges, avec ou sans revenus. Ce seuil fait en sorte que pour environ 85 % des municipalités du Québec, aucune estimation sur le revenu ne peut être divulguée à partir de l'ENM.

10. En parcourant ce tableau, le lecteur constatera d'emblée certains résultats, notamment que les estimations pour les différents concepts de revenu ne sont pas toujours concordantes d'une source de données à l'autre. Ce résultat ainsi que d'autres seront discutés dans la prochaine partie du rapport, en considérant l'ensemble des échelles géographiques, et pas seulement celle de l'ensemble du Québec.

11. Les matrices de données se rapportant à la dimension du concept de revenu ne sont pas reproduites.

Tableau 2.1

Données comparées de l'ENM et de RQ, selon le sexe et l'âge, l'indicateur de revenu, et le concept de revenu, personnes de 18 ans et plus avec revenus, ensemble du Québec, 2010

	Unité	Estimations				Effectif			
		ENM	RQ	Écart		ENM	RQ	Écart	
				Relatif ¹	Absolu ²			Relatif ¹	Absolu ²
		\$ ou %		%	point de %	n		%	n
Volet 1 : Revenu total médian des personnes de 18 ans et plus avec revenus									
	\$	28 265	28 172	0,3	93	5 956 210	6 072 700	-1,9	-116 490
Volet 2 : Revenu total médian des personnes de 18 ans et plus avec revenus, selon le sexe et l'âge									
Sexe									
Femmes	\$	22 994	22 920	0,3	74	2 998 235	3 060 035	-2,0	-61 800
Hommes	\$	34 000	34 080	-0,2	-80	2 957 975	3 012 665	-1,8	-54 690
Âge									
18-24 ans	\$	12 474	12 616	-1,1	-142	652 925	633 740	3,0	19 185
25-44 ans	\$	34 163	33 914	0,7	249	1 955 735	1 984 919	-1,5	-29 184
45-64 ans	\$	34 997	35 445	-1,3	-448	2 252 585	2 248 803	0,2	3 782
65 ans et plus	\$	20 595	20 766	-0,8	-171	1 094 965	1 205 238	-9,1	-110 273
Volet 3 : Revenu total des personnes de 18 ans et plus avec revenus, selon l'indicateur de revenu									
Revenu total moyen	\$	36 680	37 725	-2,8	...	5 956 210	6 072 700	-1,9	-116 490
Proportion dont le revenu total est inférieur à 10 000 \$	%	15,4	14,4	...	0,9	5 956 210	6 072 700	-1,9	-116 490
Proportion dont le revenu total est supérieur ou égal à 100 000 \$	%	3,8	4,2	...	-0,3	5 956 210	6 072 700	-1,9	-116 490
Volet 4 : Revenu médian des personnes de 18 ans et plus avec revenus, selon le concept de revenu									
Revenus du marché	\$	28 309	27 015	4,8	1 294	5 158 740	5 376 688	-4,1	-217 948
Salaires et traitements	\$	31 440	29 493	6,6	1 947	3 921 200	4 097 942	-4,3	-176 742
Revenu agricole net	\$	263	98	168,4	165	44 090	44 904	-1,8	-814
Revenu non agricole net	\$	7 164	7 200	-0,5	-36	483 695	520 643	-7,1	-36 948
Dividendes et revenus de placement	\$	607	657	-7,7	-50	1 708 595	1 926 422	-11,3	-217 827
Pensions de retraite	\$	14 309	10 729	33,4	3 580	893 295	1 182 589	-24,5	-289 294
Autres revenus en espèces	\$	779	2 182	-64,3	-1 403	853 020	393 275	116,9	459 745
Transferts gouvernementaux	\$	6 061	6 190	-2,1	-129	3 949 875	4 289 104	-7,9	-339 229

1. L'écart relatif se calcule de la manière suivante : $[(\text{donnée ENM} - \text{donnée RQ}) \div \text{donnée RQ}] \times 100$. Il est à noter qu'un écart relatif calculé à partir de petits nombres doit être interprété avec prudence, puisqu'il tend à être très élevé.

2. L'écart absolu se calcule de la manière suivante : $[\text{donnée ENM} - \text{donnée RQ}]$.

Source : *Enquête nationale auprès des ménages* de Statistique Canada et Revenu Québec, 2010, adapté par l'Institut de la statistique du Québec.

3. RÉSULTATS

Les résultats se structurent selon les mêmes dimensions de l'analyse, soit l'échelle géographique, la population de référence, l'indicateur de revenu et le concept de revenu.

3.1 SELON L'ÉCHELLE GÉOGRAPHIQUE

Le tableau 3.1 présente les coefficients de corrélation et les distributions de l'écart relatif entre les estimations du revenu total médian des personnes de 18 ans et plus avec revenus, selon les échelles géographiques et, pour l'échelle des MRC et des municipalités, selon les tranches d'effectif et de TGN des unités géographiques.

Plus l'échelle géographique est agrégée, plus la concordance augmente

Dans la population de référence des personnes âgées de 18 ans et plus avec revenus, la corrélation entre les estimations du revenu total médian de l'ENM et de RQ est de 0,97 à l'échelle des RA, de 0,98 à l'échelle des MRC et de 0,82 à l'échelle des municipalités. Donc, que ce soit à l'échelle des RA ou des MRC, la concordance est grande et équivalente. Elle diminue cependant à l'échelle des municipalités.

L'analyse de la distribution de l'écart relatif entre les données donne les mêmes conclusions que l'analyse de la corrélation. Ainsi, plus l'échelle géographique se raffine, plus le pourcentage des unités dont l'écart relatif est proche de zéro diminue, et donc plus la concordance diminue. À l'échelle de l'ensemble du Québec, ce pourcentage atteint 100 %, l'écart relatif étant précisément de +0,3 % (voir tableau 2.1). À l'échelle des RA, le pourcentage des unités géographiques dont l'écart est situé entre -2,5 % et +2,5 % est de 88 % ; à l'échelle des MRC, ce pourcentage s'abaisse à 70 %. Enfin, à l'échelle des municipalités, ce ne sont plus que 23 % des unités qui présentent un écart entre -2,5 % et +2,5 %.

La proportion des municipalités dont les écarts entre les estimations de l'ENM et les données de RQ dépassent 2,5 % est relativement importante. En effet, 34 % des municipalités se retrouvent avec des écarts négatifs ou positifs de 10 % et plus.

Pour une échelle géographique donnée, plus l'effectif des unités géographiques augmente, ou plus le TGN de ces unités diminue, plus la concordance augmente

► Effectif et concordance

L'amélioration de la concordance lorsque l'effectif¹ augmente est plus marquée lorsque l'échelle géographique se raffine. Il y a en effet plus de dispersion dans l'effectif des unités géographiques parmi les municipalités que parmi les MRC. Ainsi, parmi les municipalités, la corrélation des estimations entre les deux sources de données passe de 0,63 parmi les municipalités dont l'effectif est de moins de 250, à 1,00 parmi celles dont il est de 100 000 et plus.

Parmi les MRC, la dispersion de l'effectif est moins élevée puisqu'aucune ne comporte un effectif inférieur à 1 000 personnes. De ce fait, la progression de la concordance est moins prononcée. Par exemple, en excluant le cas des MRC dont la taille de la population de référence se situe entre 1 000 et 4 999 personnes², la corrélation passe de 0,92 parmi les MRC dont l'effectif est entre 5 000 et 24 999 à 1,00 parmi celles dont il est au moins de 100 000.

À la fois pour les municipalités et les MRC, les distributions de l'écart relatif montrent également que plus l'effectif des unités augmente, plus la concordance augmente. À l'échelle des municipalités, le pourcentage des unités dont l'écart se situe entre -2,5 % et +2,5 % augmente de manière systématique en fonction de l'effectif : 5 % (moins de 250) jusqu'à 100 % (100 000 et plus).

1. Dans le reste du texte, lorsque l'on réfère à l'effectif d'une unité géographique, comme une municipalité, une MRC ou une RA, on réfère non pas à l'effectif total de cette unité (c.-à-d. l'ensemble de la population de tous âges), mais à l'effectif de la population de référence considérée (ex : l'ensemble des personnes de 18 ans et plus avec revenus). Afin de ne pas alourdir la formulation, on évitera de faire cette précision systématiquement.
2. Seulement trois MRC ont cette catégorie de taille ; c'est pourquoi nous ne considérerons pas le coefficient de corrélation dans ce cas.

Tableau 3.1

Corrélation et distribution de l'écart relatif entre les données de l'ENM et de RQ, selon l'échelle géographique, l'effectif et le taux global de non-réponse à l'ENM, revenu total médian des personnes âgées de 18 ans et plus avec revenus, Québec, 2010

	Corrélation ²	Écart relatif ¹										Nombre d'unités géographiques	Références du tableau de données ⁴	
		-20,00% et moins	-19,99% à -10,00%	-9,99% à -5,00%	-4,99% à -2,50%	-2,49% à 0%	0,01% à 2,50%	2,51% à 5,00%	5,01% à 10,00%	10,01% à 20,00%	20,01% et plus			n
Volet 1 : Ensemble du Québec	...	0,0	0,0	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	1	A1-1	
Volet 2 : Régions administratives	0,97	0,0	0,0	5,9	0,0	29,4	58,8	5,9	0,0	0,0	0,0	17	A1-2	
Volet 3 : Municipalités régionales de comté	0,98	0,0	1,0	3,8	6,7	28,8	41,3	9,6	8,7	0,0	0,0	104	A1-3	
MRC selon l'effectif des 18 ans et plus avec revenus														
Moins de 250	0		
250 à 499	0		
500 à 999	0		
1 000 à 4 999	0,99	0,0	0,0	0,0	33,3	0,0	0,0	0,0	66,7	0,0	0,0	3		
5 000 à 24 999	0,92	0,0	2,0	8,0	8,0	22,0	36,0	12,0	12,0	0,0	0,0	50		
25 000 à 99 999	0,99	0,0	0,0	0,0	2,6	38,5	46,2	10,3	2,6	0,0	0,0	39		
100 000 et plus	1,00	0,0	0,0	0,0	8,3	33,3	58,3	0,0	0,0	0,0	0,0	12		
MRC selon le taux global de non-réponse (TGN) à l'ENM														
50,00% et plus	...	0,0	0,0	0,0	0,0	0,0	0,0	0,0	100,0	0,0	0,0	2		
40,00% à 49,99%	0,99	0,0	0,0	0,0	28,6	28,6	14,3	0,0	28,6	0,0	0,0	7		
30,00% à 39,99%	0,98	0,0	0,0	4,2	8,3	25,0	33,3	20,8	8,3	0,0	0,0	24		
20,00% à 29,99%	0,99	0,0	0,0	5,7	3,8	28,3	49,1	7,5	5,7	0,0	0,0	53		
10,00% à 19,99%	0,96	0,0	5,6	0,0	5,6	38,9	44,4	5,6	0,0	0,0	0,0	18		
Moins de 10,00%	0		
Volet 4 : Municipalités	0,82	4,6	11,3	12,3	10,0	11,6	11,0	7,0	13,6	10,8	7,7	1 087	A1-4	
Municipalités selon l'effectif des 18 ans et plus avec revenus														
Moins de 250	0,63	16,7	9,4	6,3	6,3	1,0	4,2	5,2	9,4	17,7	24,0	96		
250 à 499	0,63	7,7	16,0	13,8	7,2	6,6	3,3	5,0	6,6	15,5	18,2	181		
500 à 999	0,78	5,9	15,6	15,9	8,1	10,0	5,2	5,9	13,0	14,4	5,9	270		
1 000 à 4 999	0,92	1,0	10,4	12,9	12,1	12,1	13,1	7,4	20,0	7,9	3,0	404		
5 000 à 24 999	0,98	0,0	0,9	7,5	16,0	23,6	28,3	12,3	10,4	0,9	0,0	106		
25 000 à 99 999	0,99	0,0	0,0	0,0	9,5	42,9	33,3	14,3	0,0	0,0	0,0	21		
100 000 et plus	1,00	0,0	0,0	0,0	0,0	33,3	66,7	0,0	0,0	0,0	0,0	9		
Municipalités selon le taux global de non-réponse (TGN) à l'ENM³														
50,00% et plus	0,62	6,9	15,7	12,6	6,3	6,3	4,4	6,3	8,8	15,7	17,0	159		
40,00% à 49,99%	0,81	5,3	12,2	13,8	8,9	11,4	4,9	5,3	16,7	13,4	8,1	246		
30,00% à 39,99%	0,75	4,8	9,6	15,1	12,1	7,7	10,3	5,1	15,4	11,8	8,1	272		
20,00% à 29,99%	0,94	0,5	7,8	9,6	10,6	14,7	18,3	12,8	16,1	6,4	3,2	218		
10,00% à 19,99%	0,95	0,0	6,0	12,0	14,5	26,5	24,1	6,0	4,8	2,4	3,6	83		
Moins de 10,00%	0,92	14,3	7,1	14,3	0,0	0,0	14,3	7,1	7,1	21,4	14,3	14		

 1. L'écart relatif se calcule de la manière suivante : $[(\text{donnée ENM} - \text{donnée RQ}) \div \text{donnée RQ}] \times 100$.

2. Le coefficient de corrélation n'a pas lieu de figurer lorsque le nombre d'unités est égal à 2, puisque dans ce cas il est toujours égal à 1, et lorsque ce nombre est égal à 1, puisque dans ce cas il est non défini.

3. Les municipalités ayant fait l'objet d'un regroupement sont exclues. On compte 992 municipalités après ces exclusions.

4. Le code précédant le tiret correspond au nom du fichier Excel, et le nombre suivant ce tiret correspond au numéro de feuille.

Source : Enquête nationale auprès des ménages de Statistique Canada et Revenu Québec, 2010, adapté par l'Institut de la statistique du Québec.

À l'échelle des MRC, ce pourcentage augmente systématiquement, mais puisque la dispersion de l'effectif est moindre à cette échelle, ce pourcentage ne peut augmenter avec autant d'ampleur qu'à l'échelle des municipalités, du moins à partir d'un effectif de 5 000. Ainsi, ce pourcentage s'élève à 0 % dans le cas des MRC dont l'effectif se situe entre 1 000 et 4 999; il bondit à 58 % pour un effectif entre 5 000 et 24 999; il augmente à 85 % lorsque l'effectif est entre 25 000 et 99 999, puis à 92 % lorsque l'effectif est de 100 000 et plus.

► TGN et concordance

Le tableau 3.1 montre aussi le lien entre le taux global de non-réponse et le niveau de concordance entre les estimations de l'ENM et les données de RQ. De façon générale, plus le TGN diminue, plus la corrélation entre les deux sources de données est élevée. Ce lien est particulièrement observable dans le cas des municipalités. En effet, la corrélation s'élève seulement à 0,62 entre les municipalités dont le TGN est d'au moins 50 %, ce qui est relativement faible. La corrélation augmente à 0,81 pour les municipalités dont le TGN est compris entre 40 % et 50 %, puis atteint 0,92 lorsque le TGN est inférieur à 10 %³. À l'échelle des MRC, la relation entre le TGN et la concordance est faible. En effet, la corrélation ne varie pas de manière sensible : elle est saturée entre 0,96 et 0,99.

Enfin, on décèle aussi une tendance à une augmentation de la concordance à mesure que le TGN diminue avec le calcul des écarts relatifs. À l'échelle des municipalités, seulement 11 % des unités géographiques parmi celles dont le TGN est de 50 % et plus présentent un écart relatif situé entre $-2,5\%$ et $+2,5\%$. Cette faible proportion illustre la faible concordance. Cependant, celle-ci augmente graduellement lorsque le TGN diminue. Ainsi, le pourcentage des unités dont l'écart relatif varie entre

$-2,5\%$ et $+2,5\%$ augmente à 16 % lorsque ces unités ont un TGN entre 40 % et 50 % et atteint 51 % pour celles dont le TGN varie entre 10 % et 20 %.

Par contre, cette relation entre le TGN et la concordance ne s'applique plus lorsque le TGN est inférieur à 10 %. En effet, seulement 14 % des municipalités avec un TGN de ce niveau ont un écart relatif entre $-2,5\%$ et $+2,5\%$ ⁴. Quoique le TGN soit un indicateur satisfaisant de concordance, il ne prédit pas de manière infaillible le niveau de concordance.

À l'échelle des MRC, le TGN est un indicateur relié de manière relativement claire à la concordance, contrairement à ce que l'on a noté pour la corrélation. Le pourcentage d'unités avec un écart situé entre $-2,5\%$ et $+2,5\%$ augmente systématiquement avec la baisse du TGN. Aucune MRC ne se retrouve avec un TGN inférieur à 10 %.

Des analyses additionnelles ont par ailleurs montré que le TGN et la taille comportent des effets sur l'écart relatif qui leur sont propres⁵.

3.2 SELON LE SEXE ET L'ÂGE

Le tableau 3.2 présente les statistiques de concordance ENM-RQ ventilées selon le sexe et l'âge, et également selon les échelles géographiques. Avant d'aborder l'influence de ces deux variables sociodémographiques sur le niveau de concordance, signalons que les données de ce tableau continuent de montrer l'influence de l'échelle géographique sur la concordance. Ainsi, quel que soit le groupe de sexe ou d'âge formant la population de référence, la concordance entre les estimations du revenu total médian des deux sources de données tend à s'abaisser lorsque l'échelle géographique se raffine.

3. Aux fins du calcul de cette corrélation (de 0,92) parmi les municipalités dont le TGN est inférieur à 10 %, celle de Lac-Tremblant-Nord a été exclue, parce qu'elle constitue une valeur extrême fortement influente : en incluant cette municipalité dans le calcul, cette corrélation aurait diminué substantiellement, soit bien en deçà de 0,92.
4. En particulier, 36 % des municipalités, soit 5 sur 14, ont un écart relatif de $+10\%$ et plus. Notons que parmi ces cinq municipalités, on compte Lac-Tremblant-Nord, tandis que les quatre autres sont situées dans le nord du Québec et sont peuplées d'autochtones (Kangiqualujuaq, Salluit, Ivujivik et Tasiujaq). Le brouillage de la relation entre le TGN et la concordance lorsque ce TGN est inférieur à 10 % repose donc sur ces municipalités, du moins en bonne partie. D'ailleurs, contrairement à l'indicateur de corrélation, l'indicateur de l'écart relatif suggère que le brouillage de cette relation ne repose pas seulement sur la municipalité de Lac-Tremblant-Nord (voir la note précédente).
5. Des modèles de régression linéaire avec comme variable dépendante l'écart relatif en valeur absolue et exacte, et comme variables indépendantes l'effectif des unités géographiques et le TGN, tous deux en valeurs exactes, ont été examinés. Par exemple, à l'échelle des municipalités, les résultats suggèrent que le lien entre l'effectif et l'écart, ainsi qu'entre le TGN et cet écart, subsiste lorsque ces deux variables sont simultanément présentes dans le modèle. En effet, les paramètres estimés pour ces deux variables indépendantes demeurent suffisamment éloignés de la valeur nulle. À l'échelle des MRC, ce résultat est observé jusqu'à un certain point.

Tableau 3.2

Corrélation et distribution de l'écart relatif entre les données de l'ENM et de RQ, selon la population de référence et l'échelle géographique, revenu total médian des personnes âgées de 18 ans et plus avec revenus, Québec, 2010

	Corrélation ²	Écart relatif ¹											Nombre d'unités géographiques	Références du tableau de données ³	
		-20,00% et moins	-19,99% à -10,00%	-9,99% à -5,00%	-4,99% à -2,50%	-2,49% à 0%	0,01% à 2,50%	2,51% à 5,00%	5,01% à 10,00%	10,01% à 20,00%	20,01% et plus	points			n
Volet 1 : Ensemble du Québec	...	0,0	0,0	0,0	0,0	0,0	100,0	100,0	0,0	0,0	0,0	0,0	1	A2-1	
Sexe															
Femmes	...	0,0	0,0	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	0,0	1	A2-2	
Hommes	...	0,0	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	0,0	0,0		A2-3	
Âge															
18-24 ans	...	0,0	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	0,0	0,0	17	A2-4	
25-44 ans	...	0,0	0,0	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	0,0		A2-5	
45-64 ans	...	0,0	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	0,0	0,0		A2-6	
65 ans et plus	...	0,0	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	0,0	0,0		A2-7	
Volet 2 : Régions administratives	0,97	0,0	0,0	5,9	0,0	29,4	58,8	5,9	0,0	0,0	0,0	0,0	17	A2-8	
Sexe															
Femmes	0,99	0,0	0,0	0,0	0,0	47,1	47,1	0,0	5,9	0,0	0,0	0,0		17	A2-9
Hommes	0,93	0,0	5,9	0,0	0,0	64,7	29,4	0,0	0,0	0,0	0,0	0,0			A2-10
Âge															
18-24 ans	0,96	0,0	0,0	11,8	5,9	47,1	35,3	0,0	0,0	0,0	0,0	0,0		17	A2-11
25-44 ans	0,98	0,0	0,0	5,9	0,0	23,5	58,8	11,8	0,0	0,0	0,0	0,0			A2-12
45-64 ans	0,99	0,0	0,0	5,9	35,3	47,1	11,8	0,0	0,0	0,0	0,0	0,0	A2-13		
65 ans et plus	0,99	0,0	0,0	0,0	5,9	82,4	11,8	0,0	0,0	0,0	0,0	0,0	A2-14		
Volet 3 : Municipalités régionales de comté	0,98	0,0	1,0	3,8	6,7	28,8	41,3	9,6	8,7	0,0	0,0	0,0	104	A2-15	
Sexe															
Femmes	0,97	0,0	1,0	0,0	9,6	31,7	29,8	16,3	10,6	0,0	1,0	1,0		104	A2-16
Hommes	0,99	1,0	0,0	2,9	15,4	37,5	29,8	7,7	4,8	1,0	0,0	0,0			A2-17
Âge															
18-24 ans	0,79	3,8	6,7	19,2	10,6	17,3	21,2	6,7	7,7	5,8	1,0	1,0		104	A2-18
25-44 ans	0,95	0,0	1,9	1,9	10,6	30,8	29,8	13,5	6,7	3,8	1,0	1,0			A2-19
45-64 ans	0,98	0,0	1,9	14,4	26,9	34,6	11,5	8,7	1,9	0,0	0,0	0,0	A2-20		
65 ans et plus	0,95	0,0	0,0	4,8	14,4	37,5	26,9	7,7	8,7	0,0	0,0	0,0	A2-21		
Volet 4 : Municipalités	0,82	4,6	11,3	12,3	10,0	11,6	11,0	7,0	13,6	10,8	7,7	1 087	1 087	A2-22	
Sexe															
Femmes	0,76	6,2	11,9	10,1	7,8	10,9	9,3	8,1	10,4	12,8	12,5	1 087		1 087	A2-23
Hommes	0,81	8,3	12,1	11,9	9,5	9,8	8,5	7,5	12,9	10,7	8,9	1 087			A2-24
Âge															
18-24 ans	0,28	23,4	12,7	7,6	4,2	5,7	5,6	3,5	7,2	7,6	22,3	1 062		1 087	A2-25
25-44 ans	0,71	10,7	11,6	8,9	7,6	10,4	10,0	7,4	10,9	11,0	11,5	1 085			A2-26
45-64 ans	0,80	11,9	14,1	12,5	8,5	10,5	8,0	5,6	8,5	10,9	9,7	1 087	A2-27		
65 ans et plus	0,51	2,8	7,9	13,8	11,0	12,1	11,6	9,6	11,0	10,6	9,7	1 085	A2-28		

1. L'écart relatif se calcule de la manière suivante : [(donnée ENM – donnée RQ) ÷ donnée RQ] × 100.

2. Le coefficient de corrélation n'a pas lieu de figurer lorsque le nombre d'unités est égal à 1, puisque dans ce cas il est non défini.

3. Le code précédant le tiret correspond au nom du fichier Excel, et le nombre suivant ce tiret correspond au numéro de feuille.

Source : Enquête nationale auprès des ménages de Statistique Canada et Revenu Québec, 2010, adapté par l'Institut de la statistique du Québec.

Comme on peut s'y attendre, à l'échelle du Québec cette concordance est élevée. Ainsi, pour les six groupes concernés, les écarts relatifs de la médiane du revenu total sont toujours compris dans l'intervalle allant de $-2,5\%$ à $+2,5\%$ (tableau 3.2). Plus précisément, les données du tableau 2.1 ont montré que les écarts relatifs exacts sont les suivants: $+0,3\%$ (femmes), $-0,2\%$ (hommes), $-1,1\%$ (18 à 24 ans), $+0,7\%$ (25 à 44 ans), $-1,3\%$ (45 à 64 ans), $-0,8\%$ (65 ans et plus). Par contraste, pour chacun de ces groupes, règle générale, plus l'échelle géographique se raffine, plus les corrélations s'abaissent, et plus le pourcentage des unités géographiques dont l'écart relatif est situé entre $-2,5\%$ et $+2,5\%$ diminue.

Dans la présente section, on se demande plutôt dans quelle mesure une relation est établie entre le niveau de concordance d'une part, et le sexe ou l'âge d'autre part. Autrement dit, lorsqu'on restreint la population de référence seulement aux hommes de 18 ans et plus avec revenus, est-ce que la concordance ENM-RQ est plus élevée ou moins élevée, que lorsqu'on restreint cette population seulement aux femmes de 18 ans et plus avec revenus? Qu'en est-il pour les quatre groupes d'âge?

La concordance ENM-RQ ne varie pas selon le sexe...

Les deux indicateurs utilisés (le coefficient de corrélation et le pourcentage d'unités géographiques dont l'écart relatif se situe dans l'intervalle entre $-2,5\%$ et $+2,5\%$) ne permettent pas de conclure à l'existence d'une relation claire entre le niveau de concordance ENM-RQ et le sexe. Rappelons que plus une corrélation est élevée, ou plus le pourcentage des unités géographiques dont l'écart relatif se situe dans l'intervalle entre $-2,5\%$ et $+2,5\%$ est élevé, alors plus le niveau de concordance est élevé.

Cette relation n'est pas claire pour les deux raisons suivantes. D'abord, pour le coefficient de corrélation, cette relation change de sens quand on compare les RA d'une part aux MRC et aux municipalités d'autre part⁶. Ensuite, il y a absence de convergence entre les résultats des deux indicateurs à l'échelle des RA et des municipalités⁷; c'est seulement à l'échelle des MRC que ces résultats

convergent. Toutefois, les différences constatées selon le sexe dans les niveaux de concordance ne sont pas élevées à cette échelle, à plus forte raison dans le cas des corrélations ($0,97$ pour les femmes contre $0,99$ pour les hommes).

... mais cette concordance est moins élevée chez les personnes âgées de 18 à 24 ans que chez celles âgées de 25 ans et plus

Contrairement au sexe, l'âge semble affecter plus sensiblement la concordance. Du moins à l'échelle des MRC et des municipalités, une opposition se dessine entre les personnes de 18 à 24 ans, pour lesquelles la concordance est relativement faible, et les personnes de 25 ans et plus, pour lesquelles cette concordance est plus forte.

À l'échelle des MRC, la corrélation s'élève à $0,79$ chez les jeunes, alors qu'elle est toujours supérieure ou égale à $0,95$ chez les 25 ans et plus. À l'échelle des municipalités, elle s'élève à seulement $0,28$ chez les jeunes, alors que pour les 25 ans et plus, elle est au moins égale à $0,51$. L'indicateur du pourcentage des écarts situés entre $-2,5\%$ et $+2,5\%$ révèle la même tendance. Ainsi, seulement 39 % des MRC se situent dans ces tranches centrales chez les 18-24 ans, contre respectivement 61 %, 46 % et 64 % respectivement chez les 25-44 ans, les 45-64 ans et les 65 ans et plus, ce qui est toujours plus élevé et dénote donc une meilleure concordance. Dans le même ordre d'idées, seulement 11 % des municipalités se situent dans ces tranches centrales chez les 18-24 ans, contre respectivement 20 %, 19 % et 24 % pour les autres groupes d'âge.

Le lien entre le fait d'être âgé ou non de 18 à 24 ans et le niveau de concordance des données s'interprète à l'égard de la précision. Ainsi, on observe que la part des 18-24 ans dans l'ENM n'est que de 11 %, contre 89 % pour les 25 ans et plus (tableau 1.2); l'effectif de l'ENM est donc relativement faible chez les jeunes, à l'instar de la taille de l'échantillon pour ce sous-groupe, ce qui expliquerait une plus grande imprécision de l'estimation dans l'ENM, avec pour conséquence une plus faible concordance avec les données fiscales. Par ailleurs, de

6. Ainsi, à l'échelle des RA, la concordance est plus élevée pour les femmes que pour les hommes, alors que c'est le contraire pour les deux autres échelles (MRC et municipalités).

7. Ainsi, à l'échelle des RA, alors que la corrélation est plus forte chez les femmes, le pourcentage des écarts situés entre $-2,5\%$ et $+2,5\%$ est le même pour les deux sexes (94 %). Inversement, alors qu'à l'échelle des municipalités, la corrélation est plus élevée chez les hommes, le pourcentage des écarts situés entre $-2,5\%$ et $+2,5\%$ est légèrement plus élevé chez les femmes (20 % c. 18 %).

prime abord, on ne peut associer la faible concordance observée chez les jeunes de 18 à 24 ans à des biais de sous-évaluation de ce groupe, puisque leur répartition dans la population des particuliers avec revenus est relativement similaire entre les données de l'ENM et les données fiscales (11,0% et 10,4% respectivement, voir tableau 1.2). On ne peut toutefois exclure tout biais, dans la mesure où le profil des répondants de 18 à 24 ans de l'ENM pourrait différer du profil des déclarants aux fins de RQ.

Mentionnons qu'à l'échelle des RA, on ne peut conclure à l'existence de ce clivage entre les 18-24 ans et les 25 ans et plus, puisque bien que la corrélation soit plus faible chez les jeunes que chez les plus âgés, le pourcentage des écarts situés entre -2,5% et +2,5% n'est pas systématiquement plus faible (voir tableau 3.2).

Par ailleurs, à l'échelle des municipalités, les corrélations montrent que jusqu'à un certain point, la concordance ENM-RQ serait plus faible chez les personnes de 65 ans et plus que chez celles de 25 à 64 ans. Cette corrélation, en effet, s'élève à 0,51, contre 0,71 pour les 25 à 44 ans et à 0,80 pour les 45 à 64 ans. Cependant, cette tendance ne ressort pas de la comparaison des pourcentages de municipalités dont l'écart relatif se situe entre -2,5% et +2,5%, puisque de manière contradictoire, ce pourcentage est plus élevé pour les 65 ans et plus que pour les deux autres groupes d'âge.

3.3 SELON L'INDICATEUR DE REVENU

Le tableau 3.3 présente les statistiques de concordance ventilées selon les indicateurs de revenu et également selon les échelles géographiques. La population de référence demeure l'ensemble des personnes de 18 ans et plus avec revenus. Pour chaque échelle géographique, les quatre indicateurs suivants sont étudiés : le revenu total médian (variable de référence des sections précédentes), le revenu total moyen, la proportion de la population de référence dont le revenu total est inférieur à 10 000 \$, et la proportion dont ce revenu est supérieur ou égal à 100 000 \$.

Signalons que les statistiques d'écart s'expriment en écart relatif en ce qui concerne les revenus médian et moyen et en écart absolu pour les proportions en deçà de 10 000 \$ ou au-delà de 100 000 \$ (donc des écarts en points de pourcentage).

La concordance est plus élevée avec la moyenne que la médiane à l'échelle des municipalités...

À l'échelle des régions administratives et des MRC, le niveau de corrélation entre les estimations du revenu total moyen et du revenu total médian est du même ordre, quoique celles du revenu moyen soient supérieures. Cependant, à l'échelle des municipalités, le niveau de corrélation est sensiblement plus élevé avec le revenu moyen, soit 0,88 contre 0,82 pour le revenu médian.

... et beaucoup plus élevée à l'échelle des MRC et des municipalités avec la proportion dont le revenu est supérieur ou égal à 100 000 \$ par rapport à la proportion dont celui-ci est inférieur à 10 000 \$

À l'échelle des RA, la corrélation entre les proportions avec revenu supérieur ou égal à 100 000 \$ et celle pour les proportions avec revenu inférieur à 10 000 \$ se situent toutes deux au-delà de 0,95, mais celle pour la proportion avec revenu supérieur ou égal à 100 000 \$ est supérieure (0,99). À l'échelle des MRC, ces corrélations sont respectivement de 0,99 contre 0,87, et à l'échelle des municipalités, de 0,86 contre 0,64. On constate donc que plus l'échelle géographique se raffine, plus la différence en niveaux de concordance entre les indicateurs comparés augmente.

Tableau 3.3

Corrélation et distribution de l'écart relatif ou absolu entre les données de l'ENM et de RQ, selon l'indicateur de revenu et l'échelle géographique, revenu total des personnes âgées de 18 ans et plus avec revenus, Québec, 2010

	Corrélation ³	Écart relatif ¹ ou absolu ²										Nombre d'unités géographiques	Références du tableau de données ⁶
		-20,00 et moins	-19,99 à -10,00	-9,99 à -5,00	-4,99 à -2,50	-2,49 à 0	0,01 à 2,50	2,51 à 5,00	5,01 à 10,00	10,01 à 20,00	20,01 et plus		
Volet 1 : Ensemble du Québec													
Revenu total médian ⁴	...	0,0	0,0	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	1	A3-1
Revenu total moyen ⁴	...	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	0,0	0,0		A3-2
Proportion dont le revenu total est inférieur à 10 000 \$ ⁵	...	0,0	0,0	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0		A3-3
Proportion dont le revenu total est supérieur ou égal à 100 000 \$ ⁵	...	0,0	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	0,0		A3-4
Volet 2 : Régions administratives													
Revenu total médian ⁴	0,97	0,0	0,0	5,9	0,0	29,4	58,8	5,9	0,0	0,0	0,0	17	A3-5
Revenu total moyen ⁴	1,00	0,0	0,0	5,9	47,1	47,1	0,0	0,0	0,0	0,0	0,0		A3-6
Proportion dont le revenu total est inférieur à 10 000 \$ ⁵	0,95	0,0	0,0	0,0	0,0	0,0	94,1	5,9	0,0	0,0	0,0		A3-7
Proportion dont le revenu total est supérieur ou égal à 100 000 \$ ⁵	0,99	0,0	0,0	0,0	0,0	100,0	0,0	0,0	0,0	0,0	0,0		A3-8
Volet 3 : Municipalités régionales de comté													
Revenu total médian ⁴	0,98	0,0	1,0	3,8	6,7	28,8	41,3	9,6	8,7	0,0	0,0	104	A3-9
Revenu total moyen ⁴	0,99	0,0	1,0	4,8	35,6	41,3	11,5	3,8	1,9	0,0	0,0		A3-10
Proportion dont le revenu total est inférieur à 10 000 \$ ⁵	0,87	0,0	0,0	0,0	0,0	20,2	74,0	4,8	1,0	0,0	0,0		A3-11
Proportion dont le revenu total est supérieur ou égal à 100 000 \$ ⁵	0,99	0,0	0,0	0,0	0,0	84,6	15,4	0,0	0,0	0,0	0,0		A3-12
Volet 4 : Municipalités													
Revenu total médian ⁴	0,82	4,6	11,3	12,3	10,0	11,6	11,0	7,0	13,6	10,8	7,7	1 087	A3-13
Revenu total moyen ⁴	0,88	2,0	12,5	19,6	14,4	14,8	9,9	8,1	9,2	5,9	3,6		A3-14
Proportion dont le revenu total est inférieur à 10 000 \$ ⁵	0,64	0,0	1,7	7,2	10,7	21,4	30,7	15,6	9,8	2,4	0,5		A3-15
Proportion dont le revenu total est supérieur ou égal à 100 000 \$ ⁵	0,86	0,0	0,0	0,1	1,9	68,9	24,9	2,9	0,9	0,3	0,1		A3-16

1. L'écart relatif se calcule de la manière suivante : [(donnée ENM - donnée RQ) ÷ donnée RQ] × 100.

2. L'écart absolu se calcule de la manière suivante : [donnée ENM - donnée RQ].

3. Le coefficient de corrélation n'a pas lieu de figurer lorsque le nombre d'unités est égal à 1, puisque dans ce cas il est non défini.

4. On considère pour cet indicateur l'écart relatif.

5. On considère pour cet indicateur l'écart absolu.

6. Le code précédant le tiret correspond au nom du fichier Excel, et le nombre suivant ce tiret correspond au numéro de feuille.

Source : *Enquête nationale auprès des ménages* de Statistique Canada et Revenu Québec, 2010, adapté par l'Institut de la statistique du Québec.

La plus faible corrélation enregistrée concernant la proportion de personnes dont le revenu est inférieur à 10 000 \$ pourrait s'expliquer par la présence d'imprécisions dans la procédure d'imputation des revenus manquants pour les personnes en bas de l'échelle du revenu. 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.

Aux échelles des RA et des MRC, le revenu moyen est légèrement sous-estimé dans l'ENM

Bien que la concordance entre les données apparaisse meilleure en utilisant l'indicateur de revenu moyen que l'indicateur de revenu médian, on constate certaines différences de niveaux impliquant l'indicateur de moyenne. Ainsi, le revenu moyen tend à être sous-estimé dans l'ENM. Cela ressort plus clairement à l'échelle des régions et des MRC. À ces échelles en effet, la distribution de l'écart relatif pour la moyenne est centrée du côté négatif, ce qui indique que l'estimation de l'ENM a tendance à être inférieure à celle de RQ.

Ainsi, à l'échelle des RA, la totalité des régions présentent un écart relatif négatif dans le cas de la moyenne. À l'échelle des MRC, ce pourcentage est 83 %. Mais à l'échelle des municipalités, cette tendance semble s'estomper, car le pourcentage se rapproche de la parité (63 %).

Il reste à savoir pourquoi le revenu moyen est centré du côté négatif de la distribution, indiquant que le revenu moyen est moins élevé dans l'ENM, c'est-à-dire sous-évalué. Cette apparente contradiction s'explique probablement par le fait que l'ENM a sous-évalué les revenus très élevés, voire extrêmes (c.-à-d. dépassant très largement 100 000 \$). En raison de taux de non-réponse élevés, il est possible que les valeurs imputées aient eu tendance à être inférieures aux véritables valeurs. Advenant la confirmation de cette hypothèse, cela expliquerait que l'indicateur de moyenne tend à être sous-évalué avec l'ENM, alors que ce n'est pas le cas avec l'indicateur de médiane. En effet, la moyenne est sensible aux valeurs extrêmes, alors que la médiane ne l'est pas.

3.4 SELON LE CONCEPT DE REVENU

Pour cette dernière analyse, effectuée selon les concepts du revenu, les résultats concernent le revenu médian des personnes de 18 ans et plus avec revenus (relevant de ces concepts). Le tableau 3.4 présente les coefficients de corrélation entre les estimations pour l'ensemble des revenus du marché, pour chacun des six concepts de revenu du marché, et pour l'ensemble des transferts gouvernementaux. Les corrélations pour le revenu total, traitées dans les sections précédentes, sont incluses pour référence dans le tableau.

Signalons que contrairement aux sections précédentes, l'analyse ne porte que sur les corrélations. En effet, pour les concepts spécifiques de revenu (par exemple, les autres revenus en espèces qui sont de nature résiduelle), les montants ont tendance à être plus faibles que pour le revenu total. De ce fait, les estimations du revenu médian des 18 ans et plus avec revenus tendent à être plus proches de 0. La pertinence de l'analyse des écarts relatifs perd de son sens.

L'analyse des coefficients de corrélation 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 0,82 pour le revenu total, 0,80 pour le revenu du marché, contre 0,65 pour les transferts. À l'échelle des MRC et des RA, les mêmes contrastes ressortent, quoiqu'atténués. Ainsi, pour les MRC, la corrélation est de 0,98 pour le revenu total, de 0,99 pour le revenu du marché, contre 0,85 pour les transferts. Pour les RA, les corrélations sont de 0,97 pour le revenu total, 0,99 pour le revenu de marché contre 0,90 pour les transferts gouvernementaux. Soulignons que l'absence des données sur certains crédits d'impôt remboursables fédéraux dans les données de RQ est susceptible de contribuer à cette concordance relativement plus faible pour les transferts.

Tableau 3.4

Corrélation entre les données de l'ENM et de RQ, selon le concept de revenu et l'échelle géographique, revenu médian des personnes âgées de 18 ans et plus avec revenus, Québec, 2010

	corrélation				nombre d'unités géographiques			
	Ensemble du Québec ¹	Régions administratives	Municipalités régionales de comté	Municipalités	Ensemble du Québec	Régions administratives	Municipalités régionales de comté	Municipalités
Revenu total	...	0,97	0,98	0,82	1	17	104	1 087
Revenus du marché	...	0,99	0,99	0,80			104	1 087
Salaires et traitements	...	0,98	0,98	0,73			102	892
Revenu agricole net	...	0,66	0,53	0,24			104	1 068
Revenu non agricole net	...	0,92	0,47	0,08			104	1 084
Dividendes et revenus de placement	...	0,97	0,88	0,43			104	1 082
Pensions de retraite	...	0,98	0,91	0,55			104	1 082
Autres revenus en espèces	...	-0,32	0,10	0,08			104	1 087
Transferts gouvernementaux	...	0,90	0,85	0,65			104	1 087

1. Le coefficient de corrélation n'a pas lieu de figurer lorsque le nombre d'unités est égal à 1, puisque dans ce cas il est non défini.

Source : *Enquête nationale auprès des ménages* de Statistique Canada et Revenu Québec, 2010, adapté par l'Institut de la statistique du Québec.

Par ailleurs, la concordance est meilleure pour les salaires et traitements que pour tous les autres concepts de revenu du marché. Ainsi, à l'échelle des municipalités, la corrélation s'élève à 0,73 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, et les autres revenus privés en espèces), elle ne dépasse pas 0,55, une donnée qui revient aux pensions de retraite. À l'échelle des MRC, la corrélation pour les salaires et traitements est de 0,98, alors que la plus élevée des cinq autres éléments de revenu privé est de 0,91, ce qui revient aussi aux pensions de retraite.

Nous ne pouvons confirmer si les plus faibles niveaux de concordance pour les concepts de revenus du marché autres que les salaires et traitements sont liés à des différences dans la définition de ces concepts détaillés entre les deux sources de données. Cela dit, le fait que, pour ces autres concepts spécifiques, une proportion moins importante de la population des 18 ans et plus ont des revenus que pour les salaires et traitements⁸, contribue à réduire la précision de l'ENM, ayant pour conséquence une plus faible concordance, tel que constaté.

8. Par exemple, seulement une minorité de personnes retire des revenus agricoles, alors qu'une proportion beaucoup plus importante retire des salaires et traitements.

4. RÉSUMÉ ET DISCUSSION

L'objectif du présent rapport technique est de mesurer, sur le thème du revenu, et à différentes échelles géographiques du Québec, le niveau de concordance entre les données de l'ENM et des données puisées à partir des déclarations de revenus de Revenu Québec (RQ). Des efforts particuliers ont été faits pour harmoniser au mieux les univers et les variables contenus dans ces deux sources de données. Dans le doute, certaines nuances ont été faites lors de l'analyse des résultats. Deux indicateurs ont été retenus pour évaluer la concordance entre les deux sources, le niveau de corrélation et les écarts.

Comme attendu, plus l'échelle géographique est agrégée, plus la concordance augmente. Ainsi, les données de l'ENM à l'échelle provinciale et régionale apparaissent fiables, mais cette confiance diminue à l'échelle des MRC, surtout à l'échelle des municipalités. Par ailleurs, pour une échelle géographique donnée, plus les effectifs augmentent ou plus le taux global de non-réponse (TGN) diminue, plus la concordance augmente. Des analyses additionnelles ont suggéré, par ailleurs, que le TGN et la taille comportent des effets sur l'écart relatif qui leur sont en partie propres.

La concordance entre les données fiscales de RQ et les estimations de l'ENM semble plus élevée chez les particuliers de 25 ans et plus. On apprend aussi que la concordance est plus élevée avec la moyenne que la médiane, notamment à l'échelle des municipalités. Par ailleurs, la concordance est beaucoup plus élevée à l'échelle des MRC et des municipalités avec la proportion dont le revenu est supérieur ou égal à 100 000 \$ par rapport à la proportion dont celui-ci est inférieur à 10 000 \$. Enfin, le revenu moyen est légèrement sous-estimé avec l'ENM.

Lorsqu'on décortique l'analyse selon les concepts de revenu, on constate que la concordance est meilleure pour les salaires et traitements que pour tous les autres concepts de revenu du marché (revenu agricole, revenus de placement, pensions de retraite, etc.). Aussi, elle est meilleure pour l'ensemble du revenu du marché que pour les transferts gouvernementaux. Comme pour les analyses précédentes, la concordance pour chacune de ses sources est meilleure lorsque l'échelle géographique est de plus grande taille.

Les résultats confirment les appréhensions relativement à la fiabilité des données de l'ENM pour des échelles géographiques plus fines ou pour des sous-populations. Mais cette analyse a permis de mesurer l'ampleur de ces écarts. Une attention particulière doit donc être apportée aux données par municipalité, pour le groupe d'âge des 18 à 24 ans et les sous-catégories de revenus de marché.

De façon générale, cette recherche met en lumière les pertes possibles à l'égard de la fiabilité de l'ENM à mesure que des croisements de variables sont en cause (ex. : les hommes âgés de 18 à 24 ans) et, à plus forte raison, pour des échelles géographiques fines.

Puisque les données fiscales de RQ ne contiennent aucune caractéristique socioéconomique, il n'a pas été possible de comparer les deux sources de données pour d'autres variables d'intérêt, par exemple, le niveau d'éducation, les professions, le statut d'activité, etc. Toutefois, tout porte à croire que les mêmes constats généraux que ceux décrits dans ce rapport seraient ressortis.

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

En 2010, le gouvernement fédéral modifie le Programme du Recensement de Statistique Canada en remplaçant le questionnaire détaillé et obligatoire par une nouvelle enquête à participation volontaire, l'*Enquête nationale auprès des ménages* (ENM). Ce changement a suscité de nombreuses interrogations. Notamment, plusieurs utilisateurs doutent de la qualité des données sur la variable du revenu, en raison du taux élevé de non-réponse combiné au caractère facultatif de l'ENM.

Afin d'éclairer les utilisateurs de ces données au Québec, ce rapport méthodologique propose une comparaison entre les estimations sur le revenu de l'ENM et des données provenant d'un fichier administratif, soit celui des revenus fiscaux détenus par l'Agence de revenu du Québec (RQ). Des comparaisons sont ainsi effectuées à diverses échelles géographiques : régions administratives, municipalités régionales de comté, et municipalités. Les différences selon l'âge, le sexe, selon divers indicateurs et concepts de revenu sont aussi examinées.