



## INTÉGRATION DE L'EDTR ET DE L'ENQUÊTE SUR LES FINANCES DES CONSOMMATEURS (EFC)

L'Enquête sur les finances des consommateurs (EFC), qui est effectuée tous les ans en avril comme supplément à l'Enquête sur la population active (EPA), a de tout temps été la source des estimations annuelles du revenu de Statistique Canada, y compris les statistiques sur le faible revenu. L'enquête sert à recueillir des données sur le revenu selon la source, pour l'année civile précédente, et utilise des données sur le travail et des données démographiques de l'EPA pour élargir la portée analytique.

En avril 1997, on recueillera des données pour la dernière fois dans le cadre de l'EFC. L'échantillon de l'EDTR sera utilisé en remplacement, à partir de l'année de référence 1997, pour la production d'estimations annuelles ou transversales du revenu, en plus de la production de données longitudinales sur le revenu.

Deux raisons ont justifié cette intégration. Tout d'abord, il existe

### Note de la rédaction

Le présent numéro marque non seulement le début d'une nouvelle année de parution de La Dynamique, mais aussi un changement de portée. Les articles du présent numéro font état de l'intégration de l'EDTR et de l'Enquête sur les finances des consommateurs, une enquête annuelle de Statistique Canada servant à produire des estimations de revenu, y compris les taux de faible revenu. C'est pourquoi La Dynamique


comprendra désormais des renseignements au sujet de ce programme d'intégration. De plus, la série de documents de recherche, qui sera maintenant désignée sous le titre Série des documents de travail sur la dynamique du revenu et du travail connaîtra aussi une portée élargie.

Nathalie Noreau

suffisamment de chevauchement entre les objectifs et le contenu des deux enquêtes pour permettre des réductions de coûts, ou encore une utilisation plus efficace des ressources, grâce à la fusion. En deuxième lieu, l'intégration favorisera l'harmonisation des données, et sans elle, les estimations du revenu différeront pour des raisons difficiles à quantifier et à expliquer aux utilisateurs.

### Contenu

L'intégration des données démographiques, de même que celles sur le travail et le revenu de l'EDTR avec les données équivalentes de l'EFC semble assez aisée. Les données contextuelles disponibles pour l'analyse des données transversales sur le revenu seront élargies grâce à cette fusion, étant donné que les données démographiques et culturelles, ainsi que les données sur le marché du travail, seront tirées de l'interview actuelle sur le travail de l'EDTR, plutôt que de l'EPA. En outre, étant donné que la première enquête utilise une période de référence annuelle, elle semble mieux convenir à la production de données annuelles sur le revenu. Parmi les domaines prioritaires au sujet desquels les données seront enrichies figurent : la rémunération des hommes et des femmes, en vue d'obtenir davantage de renseignements sur les personnes qui ne travaillent pas à longueur d'année et à temps plein; la dynamique des familles, en vue d'établir des liens entre les changements familiaux et les données sur le revenu des familles; l'expérience de travail et la formation élargie, en vue d'analyser les salaires et les traitements; les périodes pendant lesquelles les répondants reçoivent des transferts gouvernementaux, par exemple des prestations d'assurance-chômage, des indemnités des accidents du travail et des prestations d'aide sociale (plutôt qu'un montant annuel uniquement).



Uniformité des séries chronologiques transversales  
Cette transition devrait entraîner une rupture des séries chronologiques transversales. Cela vient peut-être principalement du fait que, dans le cadre de l'EDTR, les répondants ont la possibilité de donner accès à leur déclaration de revenus, en vue de réduire l'attrition et les erreurs de réponse. Jusqu'à maintenant, pour le premier panel, plus de la moitié des répondants ont fourni les données sur le revenu au moyen de la "méthode de consultation de la déclaration de revenus" plutôt que par le biais d'une "interview". Cette option sera aussi intégrée aux données transversales, à compter de la collecte de cette année. Une fois les données recueillies, les deux méthodes de collecte sont fusionnées en un seul ensemble de données qui comprend une série de catégories de revenu raisonnablement équivalentes. À partir des études effectuées jusqu'à maintenant, il ressort clairement (ce qui n'est pas surprenant) que les modèles de déclaration des deux sources ne sont pas identiques. Ainsi, le recours à la "méthode de consultation de la déclaration de revenus" entraînera probablement une rupture de la série chronologique. Toutefois, les comparaisons des données de l'EFC et de l'EDTR au cours des deux prochaines années donneront un aperçu préalable des répercussions prévues de ce facteur important.

#### Actualité

Le recours aux données des déclarations de revenus a aussi un impact sur l'actualité des données. À l'heure actuelle, les premières estimations de revenu de l'EFC sont diffusées en novembre, et elles sont suivies par une série de diffusions échelonnées sur plusieurs mois. Du fait du recours aux données fiscales, la première diffusion de données transversales sur le revenu ne peut pas se faire aussi tôt, étant donné que le fichier Impôt de Revenu Canada n'est pas disponible avant l'automne. Par ailleurs, le lien plus étroit avec les


données des déclarations de revenus entraînera probablement une plus grande compression du calendrier de diffusion pour une année donnée. L'objectif consiste à fournir les données au plus tard 15 mois après la fin de la période de référence.

#### Plan d'échantillonnage et estimation

L'élément clé de l'intégration est le regroupement des échantillons, les mêmes répondants fournissant à la fois des données transversales et longitudinales. L'échantillon longitudinal sera augmenté chaque année au moyen d'un échantillon remis à niveau, afin de maintenir la représentativité transversale. Chacun des deux panels longitudinaux comprend environ 15 000 ménages, soit un échantillon d'environ 30 000 ménages sur une base annuelle. Une fois l'attrition prise en compte, un échantillon remis à niveau d'environ 10 000 ménages par année est nécessaire pour maintenir le niveau de fiabilité des données transversales sur le revenu. Les personnes faisant partie de l'échantillon "remis à niveau" répondront à l'interview préliminaire et à l'interview sur le travail de l'EDTR en janvier. Grâce aux procédures normalisées actuellement appliquées à l'EDTR, les membres de l'échantillon "remis à niveau" pourront choisir de permettre l'utilisation des données administratives sur le revenu les concernant ou de fournir des données sur le revenu dans le cadre d'une interview en mai. L'application des mêmes procédures de collecte que celles utilisées pour les membres de l'échantillon "régulier" permettra l'intégration des données des deux échantillons en une seule base de données.

#### Collecte des données et relations avec les répondants

L'intégration fournit l'occasion de revoir les procédures de collecte des données et l'approche utilisée en ce qui a trait aux relations avec les répondants. Compte tenu des quelques années d'expérience dont nous disposons relativement à l'interview



assistée par ordinateur, il s'agit d'une occasion de prendre du recul et d'examiner comment on peut améliorer l'application de la collecte, par exemple en réduisant les erreurs de déclaration et d'entrée des données. Le revenu est un sujet très délicat pour les répondants, mais nous pouvons peut-être en faire un sujet plus "facile à aborder" pour ces derniers et les intervieweurs. Étant donné que la fusion affectera de toute façon la continuité des séries chronologiques transversales, il convient de profiter de l'occasion qui est offerte d'innover quant aux techniques de collecte des données et aux approches utilisées relativement aux relations avec les répondants.

#### Traitement, vérification et imputation

Le système de traitement de l'EDTR poursuivra son évolution, mais l'intégration des programmes aura un effet sur lui, au fur et à mesure que les données transversales prendront de l'importance. Le système de l'EDTR intégrera les "pratiques les plus efficaces" de l'EFC. La présence de données longitudinales pour une majorité de répondants et l'utilisation des données administratives de l'impôt sur le revenu joueront un rôle fondamental quant au traitement des données.

#### Produits et services

Les produits de données transversales et longitudinales demeureront nécessaires, tout comme les produits démontrant comment ces deux genres de données se complètent. À court terme, la gamme de produits offerts actuellement dans le cadre des deux programmes sera maintenue. À long terme, on prévoit des consultations avec les utilisateurs des données quant aux changements qui pourraient être apportés à la gamme de produits, même si certaines orientations s'appliquant à l'ensemble de Statistique Canada seront maintenues. Parmi ces orientations

figurent : l'abandon graduel des produits en papier; l'harmonisation des données, la documentation des utilisateurs et l'amélioration de la présentation des produits; l'accroissement du contenu analytique des programmes de production; les projets de collaboration en matière d'analyse avec d'autres secteurs de Statistique Canada et de l'extérieur de Statistique Canada. Une des grandes préoccupations dans le domaine de la diffusion est l'impact de l'intégration d'échantillons sur la confidentialité des fichiers de microdonnées. À l'heure actuelle, les fichiers de microdonnées à grande diffusion produits à partir de l'EFC permettent d'établir des liens entre les membres des ménages. Une des mesures de protection de la confidentialité des fichiers de microdonnées à grande diffusion de l'EDTR est de supprimer ces liens.

## LE «PONT» VERS L'INTÉGRATION

Initialement, on prévoyait utiliser le système de production de l'EDTR pour les données transversales sur le revenu au moment du changement du mode de collecte (pour l'année de référence 1997). Toutefois, le système de l'EDTR n'est pas encore complètement établi et est encore en mode de "rattrapage". Il ne sera pas en mesure, à ce moment-là, de produire des données transversales sur le revenu répondant aux attentes des utilisateurs en ce qui a trait à l'actualité.


Par conséquent, pour les années de revenu 1997 et 1998, un système de traitement provisoire (appelé le «pont») sera élaboré pour produire des données transversales sur le revenu au moment opportun. Les éléments du «pont» comprendront des données vérifiées et non vérifiées de l'échantillon longitudinal, ainsi que des données non vérifiées de l'échantillon remis à niveau. Dans la mesure du possible, le «pont» utilisera une combinaison de codes existants des systèmes de production actuels de l'EFC et de



l'EDTR. On ne choisira qu'un nombre limité de variables sur le revenu, de variables démographiques et de variables sur le travail pour le «pont». Les variables choisies sont celles nécessaires pour produire les statistiques habituelles "haut de gamme" sur le revenu, y compris les revenus moyens, les taux de faible revenu, les ratios de rémunération des femmes et des hommes et les effets des paiements de transfert et des impôts sur le revenu des familles. Les données sur le revenu de 1997 traitées au moyen du système du «pont» devraient être diffusées en mars 1999. L'écart de 15 mois entre la fin de l'année de référence et la diffusion des données sera maintenu lors de la conversion à un système de production intégré.



## DONNÉES SUR LE LOGEMENT ET L'ÉQUIPEMENT MÉNAGER



Le programme actuel de statistiques sur le revenu et le logement comprend l'EFC, ainsi que l'Enquête annuelle sur l'équipement ménager (EEM), l'Enquête annuelle sur les réparations et les rénovations, qui est maintenant effectuée en régime de récupération complète des coûts, ainsi que d'enquêtes spéciales occasionnelles. Même s'il y a un décalage d'un mois, l'échantillon chevauche celui de l'EFC, ce qui permet l'utilisation de données totales et détaillées sur le revenu pour l'analyse des conditions de logement et d'autres caractéristiques. En fait, les données de l'EEM et de l'EFC et les données sur le loyer de l'EPA sont combinées en une base de données annuelle qui s'est révélée très populaire.

Les données sur le logement et l'équipement ménager continueront d'être recueillies, dans le cadre toutefois de l'Enquête sur les dépenses des familles (EDF). L'inconvénient de ce changement est la perte de données détaillées sur le revenu pouvant être utilisées parallèlement à celles de l'EEM. Toutefois, on était d'avis qu'une partie du contenu de l'EEM ne pouvait pas être utilisée dans un



contexte longitudinal et ajoutait au fardeau de réponse, sans que les avantages soient suffisants pour justifier les risques accrus. L'EEM comprend aussi des questions qui sont davantage orientées sur les “biens de consommation”, les électroménagers, par exemple. L'utilité d'une perspective longitudinale n'est pas évidente pour ces éléments, et ceux-ci s'intègrent mieux à l'Enquête sur les dépenses des familles.

## ÉCHELLES D'ÉQUIVALENCE


Aux fins de l'identification des familles à faible revenu, on convient généralement que la taille et la composition de chaque famille devraient être prises en compte. Par exemple, une personne seule qui a un niveau de revenu donné devrait être moins susceptible de figurer parmi les personnes à faible revenu qu'une famille de cinq ayant le même revenu monétaire. Des échelles d'équivalence sont utilisées comme un indice des besoins d'une famille donnée, par rapport à une famille ne comptant qu'une personne. Ainsi, les revenus des familles peuvent être comparés pour toutes les familles, à partir d'un revenu familial “corrigé” : le revenu familial réel divisé par l'échelle d'équivalence.

Il n'existe pas d'échelle d'équivalence acceptée de façon générale, et il faut toujours faire preuve de jugement à cet égard. Le fichier de la première vague de l'EDTR comprenait une échelle d'équivalence utilisée par l'OCDE (Organisation de coopération et de développements économiques). Ce n'est pas l'échelle d'équivalence utilisée par Statistique Canada pour la production de ses MFR (mesures de faible revenu). Cette deuxième échelle d'équivalence sera comprise dans les prochains fichiers de microdonnées à grande diffusion de l'EDTR.



CONSEILS À L'INTENTION  
DES UTILISATEURS DU  
FICHIER DE MICRODONNÉES  
À GRANDE DIFFUSION



Parmi les nombreux commentaires et questions que nous avons reçus depuis la diffusion du fichier de microdonnées de la première vague de l'EDTR et du produit sur CD-ROM, certains éléments reviennent fréquemment. Le présent article vise à souligner certains points importants.

#### Codes réservés

Même si elles ne sont pas toujours utilisées, toutes les variables de l'EDTR comportent quatre "codes réservés", qui ont une signification spéciale. Les utilisateurs doivent tenir compte de ces codes réservés, particulièrement dans le cas des variables numériques. Sauf pour quelques exceptions, les codes réservés correspondent aux quatre valeurs les plus élevées pour le champ, selon la longueur de celui-ci. Pour un champ à un caractère, ces valeurs sont 6, 7, 8, 9. Pour un champ à deux caractères, ces valeurs sont 96, 97, 98, 99, et ainsi de suite pour les champs plus longs. La signification de ces codes est la suivante :

6, 96, 996, ... = Ne fait pas partie de l'échantillon de l'EDTR  
(S'applique aux répondants de l'échantillon qui ne sont pas là pour toutes les années d'enquête.)

7, 97, 997, ... = Ne sait pas (En général, ces codes signifient que nous ne connaissons pas la valeur, mais certains sont le résultat d'une réponse Ne sait pas du répondant.)

8, 98, 998, ... = Refus (Personne qui refuse de fournir la valeur.)

9, 99, 999, ... = Sans objet (Personne non concernée par l'élément de données.)

Il ne faut pas effectuer d'opérations arithmétiques, par exemple le calcul des moyennes, sans supprimer les codes réservés.

#### Contraintes d'enregistrement d'IVISION

Par suite de ce qui précède, certains utilisateurs ont exprimé leur frustration relativement à l'aspect peu fonctionnel des contraintes

d'enregistrement du logiciel IVISION Browser. Les versions à venir du logiciel devraient permettre une plus grande souplesse. En attendant, pour rechercher une table comportant une contrainte d'enregistrement complexe, il faut calculer une nouvelle variable (à partir de «définir un recodage», «définir des bandes» ou «définir un champ dérivé»). Cette nouvelle variable doit servir à définir la table, puis on doit mettre l'accent sur cette section pour obtenir la table relative à la population visée.

#### Repérage de la population âgée de 16 à 69 ans

Le fichier de microdonnées comprend des dossiers pour toutes les personnes âgées de 16 ans et plus à la fin de 1993, soit la période de référence pour la plupart des données. Les données sur le travail ne sont recueillies que pour les personnes âgées de 16 à 69 ans, en raison du très faible taux d'activité sur le marché du travail des personnes âgées de 70 ans et plus. Malheureusement, il est difficile de repérer la population âgée de 16 à 69 ans dans le fichier de microdonnées à grande diffusion. Cela sera corrigé pour les fichiers à venir. Dans le cas du fichier de la première vague, on peut repérer la population âgée de 16 à 69 ans en n'utilisant que les dossiers dont la valeur de la variable STUDEF26b diffère de 9. [Cette variable est un indicateur des personnes qui étaient aux études au cours de l'année de référence. Toutefois, elle n'a pas d'importance comme telle dans ce contexte.]

#### SOMMAIRES DES DOCUMENTS DE RECHERCHE

Voici un aperçu des documents de recherche publiés récemment, qui peuvent être commandés individuellement (5 \$) ou dans le cadre d'un abonnement annuel (25 \$ sur disquette ou 50 \$ sur papier pour 12 à 15 documents).

96-09 Le niveau de scolarité : Facteur déterminant de  
l'autonomie et l'autorité en milieu de travail  
George Butlin, Jillian Oderkirk

Les résultats de l'analyse des données de l'Enquête sur la dynamique du travail et du revenu de 1993 et de l'Enquête sociale générale de 1994 indiquent que, à quelques exceptions près, l'éducation est l'un des meilleurs éléments prédictifs de la capacité à obtenir un emploi offrant de l'autonomie et permettant l'exercice d'une autorité en milieu de travail. Ce constat reste valide même lorsque l'incidence d'autres facteurs influant aussi sur l'accès à ces types d'emploi, notamment le sexe, l'âge, la taille de l'entreprise, le nombre d'années d'expérience et le secteur d'activité, est prise en considération.

96-11 Données sur la famille tirées de l'enquête sur la dynamique  
du travail et du revenu : situation en 1996  
Philip Giles

La «situation de famille» d'un individu est souvent un élément déterminant dans la prise de diverses décisions. Par exemple, une personne peut choisir de ne pas occuper un emploi rémunéré parce que les gains de son conjoint sont suffisants pour subvenir à leurs besoins. Les variables relatives à la famille sont donc importantes pour de nombreuses analyses, tout particulièrement pour les analyses longitudinales. L'utilisation des données longitudinales sur la famille est compliquée parce que la famille peut changer avec le temps, au rythme des arrivées et des départs, des naissances et des décès des membres. Une discussion accompagnée de plusieurs exemples sont présentés, montrant comment les données sur la famille peuvent être analysées dans une optique longitudinale.