



# La Dynamique

DU TRAVAIL ET DU REVENU

Volume 3 Numéro 3

Septembre 1994

SSSSSSSSSSSSSSSS

## DONNÉES DU PREMIER CYCLE

§ Même s'il reste encore beaucoup à faire avant que les premières données ne «sortent», une étape importante est maintenant franchie : les données pour une première année complète ont été recueillies et sont en cours de traitement. Il est encore trop tôt pour parler en détail des résultats, mais nous avons déjà une idée de ce qui s'est passé sur le terrain et nous pouvons vous en faire part. Il s'agissait de notre première collecte à grande échelle sur le terrain, de sorte que nous devrions être en mesure d'améliorer encore les résultats, même s'il sont déjà très bons.

### Réponse

Il faudra attendre la fin du traitement pour connaître précisément les taux de réponse. Par exemple, nous devons peut-être transformer certains «répondants» en «non-répondants» parce que l'interview n'aura pas permis d'obtenir suffisamment de renseignements concrets. Il n'est jamais simple de calculer le taux de réponse d'une enquête par panel. Mais les autres enquêtes par panel pourraient fort bien nous paraître simples à côté de l'EDTR parce qu'il s'agit d'une enquête à deux phases, les renseignements sur le travail étant recueillis en janvier et ceux sur le revenu, en mai.

### Note de la rédaction

L'été a beau être la saison des loisirs, cela ne semble pas empêcher les gens de s'intéresser à l'EDTR. Chaque semaine, plusieurs demandes d'abonnement à *La Dynamique* nous parviennent. De nombreuses personnes achètent également notre série de Documents de recherche, soit sur papier (50 \$ par année), soit sur disquette (15 \$ par année). Nous sommes heureux d'accueillir nos nouveaux abonnés et espérons que l'EDTR répondra à leurs attentes.


Quant aux anciens, nous savons qu'ils attendent patiemment nos données. Qu'ils se réjouissent, car notre premier produit de données de l'EDTR sera bientôt disponible. Nous vous invitons à lire le présent numéro de *La Dynamique* pour en savoir plus long. Nous abordons avec enthousiasme cette étape importante.

Philip Giles

L'échantillon de l'interview sur le travail comprend environ 42 000 personnes, adultes et enfants, dont 40 000 répondants longitudinaux. Ces derniers sont les personnes vivant dans les logements sélectionnés pour le premier panel en janvier 1993. Les 2 000 autres répondants, soit 5 % du total, sont des *cohabitants*. C'est ainsi que l'on désigne les gens qui cohabitent présentement avec des répondants longitudinaux, mais qui ne cohabitaient pas avec eux lors de la sélection de l'échantillon. Le taux de réponse est de l'ordre de 86 %. Dans 6 % des cas, la non-réponse ne fait aucun doute (par exemple, le répondant était impossible à dépister ou il a refusé de répondre). Les 8 % qui restent représentent les questionnaires qui n'ont pas été remplis en entier, ou qui n'ont pu être extraits en raison de problèmes techniques du système d'interview assistée par ordinateur. Dans certains cas, il sera peut-être possible de récupérer les renseignements sur les disquettes de sauvegarde des intervieweurs.

Pour la seule interview de mai sur le revenu, le taux de réponse a été de 83 % environ. La proportion de cas non reçus pour cause de problèmes techniques est tombée à 3 %; ici encore, certains seront peut-être récupérables. Les refus et les cas «non dépistés» représentent environ 9 % du total. Les non-réponses pour d'autres raisons totalisent 3 %. Moins de 1 % des répondants n'étaient pas admissibles à l'interview sur le revenu parce qu'ils avaient, par exemple, déménagé dans un établissement ou à l'étranger, ou parce qu'ils étaient décédés.

L'interview sur le travail et celle sur le revenu portent sur des sujets connexes, bien que différents, et partagent la même période de référence. Nous avons choisi de faire deux interviews simplement pour optimiser la qualité des données. Plus précisément, nous croyons qu'il vaut mieux recueillir les renseignements sur le travail peu de temps après la fin de l'année de référence pour réduire au minimum les erreurs de mémoire, et les renseignements sur le revenu au moment des déclarations de revenus pour permettre aux répondants de consulter leurs dossiers. Ensemble, les interviews sur le travail et le revenu constituent un cycle.



Un répondant qui répond à l'interview sur le travail, mais non à celle sur le revenu, ou vice versa, nous donne une réponse partielle à l'ensemble du cycle. Sur les 42 000 répondants longitudinaux et cohabitants, 33 000 environ étaient admissibles aux interviews sur le travail et le revenu (parce qu'ils étaient âgés de 16 ans ou plus le 1<sup>er</sup> janvier 1994) d'après nos meilleures estimations. Environ 72 % de ces derniers ont répondu à une interview sur le travail et à une autre sur le revenu, 14 % ont répondu aux questions sur le travail mais non à celles sur le revenu, 6 % ont fait l'inverse, et 8 % n'ont répondu à aucune des deux interviews. Par conséquent, notre *taux de réponse au cycle d'enquête* est d'environ 92 %. Rappelons qu'il s'agit là des tout premiers résultats, qui pourront changer au fur et à mesure des derniers rappels.

### **À venir**

En mai, quelque 500 nouveaux membres (nouveaux cohabitants) ont été recensés dans les ménages de l'échantillon. Parce que ces personnes ne faisaient pas partie des ménages au début de l'année, elles n'étaient pas admissibles cette année aux interviews sur le travail et le revenu. Toutefois, elles ont été inscrites comme nouveaux membres et devront fournir l'information sur le travail et le revenu lors du deuxième cycle, si elles cohabitent toujours avec un répondant longitudinal.

Mot de la fin : Il n'y a que deux courtes périodes de collecte de données par année, mais bien du travail sur le terrain se fait dans l'intervalle. Par exemple, les répondants recevront un bulletin avant le deuxième cycle, qui les renseignera sur les données recueillies au premier cycle. Une lettre de remerciement a également été envoyée aux répondants du premier cycle. Une carte de changement d'adresse était jointe, et quelque 250 répondants nous l'ont retournée pour nous indiquer qu'ils avaient déménagé ou étaient sur le point de le faire. Un tel taux de réponse est inattendu et encourageant, et cela va sûrement nous faciliter la tâche pour le deuxième cycle de collecte des données.

SSSSSSSSSSSSSSSS

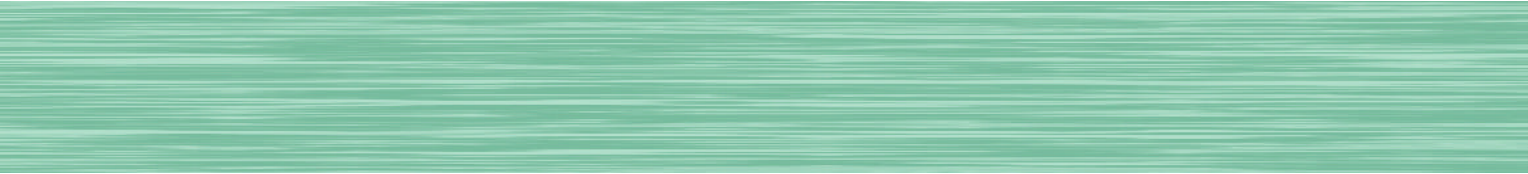
**PREMIÈRE PUBLICATION SUR  
LES DONNÉES DE L'EDTR :  
RÉSULTATS DE L'INTERVIEW  
PRÉLIMINAIRE**

§ Cet automne, l'équipe de l'EDTR lancera la première publication contenant les données de l'enquête, en prévision de la diffusion, au printemps 1995, du fichier de microdonnées complet du premier cycle et des résultats d'analyses détaillées. Cette publication sera par conséquent d'envergure modeste; on y trouvera les principales catégories de renseignements recueillis dans le cadre de l'interview préliminaire en janvier 1993. Lors de cette interview, nous avons enregistré les caractéristiques de base des répondants de notre premier panel. Les utilisateurs auront donc la possibilité, grâce à la publication, de «faire connaissance» avec les répondants de l'EDTR, en particulier de découvrir leur profil sur le plan du capital humain et des caractéristiques démographiques, avant d'utiliser le large éventail de données sur l'activité et le revenu qui seront mises à leur disposition plus tard.

De plus, nous espérons que les quelques courts articles présentés dans la publication inciteront les utilisateurs à lire davantage sur le sujet ou à réaliser leurs propres études ultérieurement. Dans certains cas, les variables de classification de base de l'interview préliminaire diffèrent de celles d'autres enquêtes, par exemple l'expérience de travail, le niveau d'instruction des parents et le nombre d'enfants mis au monde et/ou élevés. En ce qui concerne les variables plus courantes, on fera référence dans cette publication préparatoire à des résultats repères provenant d'autres enquêtes, afin d'évaluer la qualité des données de l'EDTR et les restrictions imposées par la taille de l'échantillon. Voici les sujets que nous avons l'intention d'aborder :

1. Écart salarial selon le sexe

Dans cette analyse, on croisera les renseignements relatifs à l'activité au moment de l'interview (janvier 1993) avec les variables de classification de base pour examiner l'écart de salaire entre les hommes et les femmes, tout en faisant référence à des travaux de recherche existant sur le sujet. On tirera profit du fait que l'EDTR comporte un large éventail de variables qui permettent de tenir compte des causes «légitimes» des écarts salariaux entre les hommes et les femmes, notamment l'âge, la scolarité, la profession, les heures de travail et la durée d'emploi. On



présentera également dans cette étude des résultats relatifs à une autre variable de contrôle, dont on ne dispose pas habituellement... (à vous de lire!).

## 2. Expérience de travail

L'expérience de travail n'est pas déterminée simplement par l'âge et les années de scolarité, mais c'est habituellement de cette façon que les chercheurs l'estiment (âge - années de scolarité - 5). Autrement dit, ils supposent qu'une personne commence à travailler de façon continue immédiatement après avoir quitté l'école. En quoi les résultats obtenus à partir des réponses aux questions sur l'expérience de travail dans l'EDTR diffèrent-ils de cette mesure approximative, et quelles sont les implications évidentes de ces résultats sur le plan de la recherche? Existe-t-il des sous-populations (mises à part celles des femmes et des enfants, qui sont évidentes) pour lesquelles la formule est quelque peu tendancieuse? En jetant une certaine lumière sur cet aspect du capital humain, il est à espérer que la variable «expérience de travail» non seulement améliorera la qualité des études fondées sur les données de l'EDTR, mais offrira également d'autres possibilités de recherche.

## 3. Mobilité intergénérationnelle du point de vue de l'instruction

Les parents espèrent toujours que leurs enfants réussiront mieux qu'eux ne l'ont fait, et la tendance à la hausse du niveau d'instruction porte à croire qu'à ce chapitre c'est généralement ce qui se produit. Les données de l'interview préliminaire de l'EDTR permettront de mesurer plus précisément cette mobilité intergénérationnelle (dans les deux sens). De plus, nous pourrons examiner les caractéristiques des adultes qui ont atteint un niveau d'instruction plus élevé que celui de leurs parents et les caractéristiques des adultes qui ont fait des études moins longues que leurs parents (après avoir tenu compte de l'effet de la hausse générale du niveau d'instruction).

#### 4. Données relatives à l'équité en matière d'emploi

Dans cet article, on décrira les renseignements que l'on peut tirer de l'EDTR relativement à deux des quatre groupes désignés aux fins de l'équité en matière d'emploi : les autochtones et les minorités visibles. Les chercheurs qui s'intéressent à la question de l'équité en matière d'emploi voudront connaître la qualité des données ainsi que les possibilités d'analyse qu'offre l'EDTR par rapport à d'autres enquêtes. On évaluera la qualité des données en les comparant à des données de référence. On examinera, en outre, comment les données relatives aux groupes minoritaires obtenues dans le cadre d'autres enquêtes longitudinales ont été analysées.

Afin de renseigner les personnes qui s'intéressent à l'EDTR sur tous les sujets abordés dans l'interview préliminaire, la publication contiendra deux notes brèves qui résumeront les résultats relatifs aux antécédents matrimoniaux et au nombre d'enfants mis au monde et/ou élevés. Enfin, elle comportera un aperçu du contenu, des utilisations et des avantages prévus des données.


Nos lecteurs se souviendront peut-être que, dans le numéro de février du bulletin *La Dynamique*, nous avons laissé entendre qu'un fichier de microdonnées établi à partir des résultats de l'interview préliminaire serait diffusé cet été. La publication analytique qui paraîtra cet automne vient modifier ces plans, ce qui nous permettra de respecter la date de diffusion prévue du fichier de microdonnées complet et des principaux résultats au printemps 1995.

SSSSSSSSSSSSSSSSSS

#### **UTILISATION DES DONNÉES FISCALES ) ÉTUDE DE FAISABILITÉ**

§ Plusieurs membres de l'équipe de l'EDTR ont examiné une façon assez prometteuse de réduire les coûts liés à l'enquête et d'alléger le fardeau du répondant tout en améliorant la qualité des données. La publication n° 94-11, «**L'utilisation des données d'impôt dans l'Enquête sur la dynamique du travail et du revenu : Rapport sommaire**» de la série des documents de recherche de l'EDTR, présente une analyse complète de la faisabilité et des avantages, pour l'EDTR, de se servir des fichiers





de données de l'impôt sur le revenu de Revenu Canada plutôt que de recueillir des données sur le revenu chaque année.

Un fait demeure nettement certain : l'utilisation des données fiscales ne remplacera pas complètement la méthode traditionnelle de l'interview pour certains éléments du revenu. En effet, tout le monde ne produit pas une déclaration de revenus et certains répondants choisiront une autre façon de fournir les renseignements demandés dans l'enquête. Au début donc, seule une partie des données sera touchée, et la marche à suivre sera la même pour de nombreux répondants, exception faite de la nouvelle question sur le consentement, qui sera posée à tous les répondants.

Combien de répondants autoriseraient Statistique Canada à consulter les données qu'ils ont déjà fournies à Revenu Canada? Nous avons tenté de le savoir, du moins de façon exploratoire, en posant une question hypothétique simple aux membres du premier panel de répondants de l'EDTR lors de l'interview sur le revenu, en mai 1994. À peine plus de la moitié (53 %) ont dit qu'ils autoriseraient Statistique Canada à consulter leurs dossiers fiscaux afin de recueillir les données sur le revenu. Ce résultat rejoint sensiblement celui qu'avait donné un essai semblable, en août 1993, lors de l'interview de l'Enquête sur la population active. On s'attend à ce que le taux de consentement augmente lorsque la question ne sera plus hypothétique, parce qu'on renseignera mieux alors les répondants sur les avantages de répondre à l'enquête de cette façon.

La question relative au consentement est importante parce que les répondants doivent savoir exactement quels renseignements sont demandés, même lorsque c'est un autre membre du ménage qui répond pour eux. Nous avons l'intention d'envoyer une trousse d'information avant l'interview sur le revenu. Les répondants auraient en outre la possibilité de retirer leur consentement en tout temps.

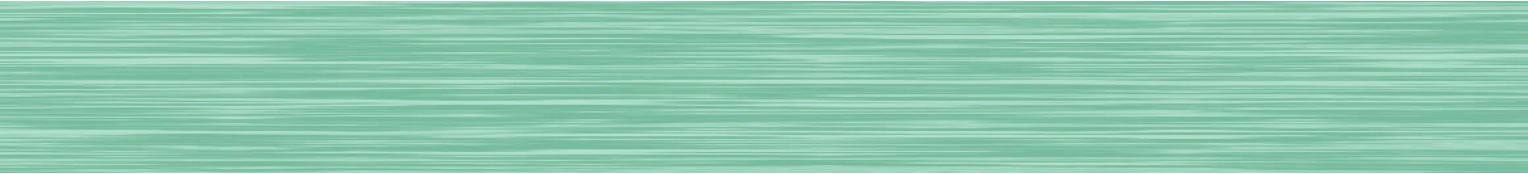
Quels avantages retire-t-on du fait d'utiliser les dossiers fiscaux de Revenu Canada chaque fois que possible? D'abord, on considère, pour un certain nombre de raisons, que les données qui proviennent de cette source sont de meilleure qualité que les

renseignements fournis par les répondants. La non-réponse aux questions sur le revenu, du moins à certaines d'entre elles, pourrait être réduite, et l'érosion de l'échantillon pendant la durée de l'enquête (six ans) serait moins forte. De plus, il est reconnu que les répondants ont tendance à sous-déclarer leurs revenus lors d'interviews auprès des ménages, comparativement aux montants qu'ils déclarent pour à l'impôt.

Des membres de l'équipe de l'EDTR ont mené des études afin d'estimer la différence de qualité des données entre celles qui proviennent des dossiers fiscaux et celles qui sont recueillies directement lors d'une enquête. Pour les seuls besoins de cette évaluation de la qualité des données, ils ont obtenu des données fiscales et des données d'enquête (essai sur le terrain de l'EDTR de 1993 et Enquête sur les finances des consommateurs (EFC) de 1992) se rapportant à un petit échantillon de répondants et visant deux années consécutives. Pour évaluer la qualité des données longitudinales, les auteurs ont examiné les changements dans le revenu calculés à partir des données de chacune des deux sources. Les données fiscales ont servi de données repères. Les auteurs ont noté que dans les données d'enquête, en général, les changements en ce qui concerne la perception de revenus étaient surestimés et que cette surestimation était probablement attribuable au fait que les répondants ont plus souvent tendance à oublier de déclarer des sources de revenu. Dans l'ensemble, nous croyons que la qualité des données serait meilleure si l'on utilisait les données fiscales chaque fois que possible.

Certains autres avantages de l'utilisation des données fiscales pour obtenir des renseignements sur le revenu dans l'EDTR sont assez évidents et peuvent être plus ou moins quantifiés. La possibilité d'alléger le fardeau du répondant est très attrayante dans une enquête longitudinale où les mêmes personnes sont interviewées de nombreuses fois en six ans. Dans le cas des répondants qui acceptent que l'on consulte leur dossier fiscal, le nombre d'interviews en six ans pourrait passer de 13 à 7. La réduction possible des coûts a, bien sûr, été examinée. Par exemple, on estime que le fait de réduire de moitié le nombre de ménages à interviewer entraînerait une réduction des coûts de l'ordre de 40 % en ce qui concerne l'interview sur le revenu.





Sur le plan de la qualité des données, il faut considérer le fait que l'on réunit des données sur le revenu qui proviennent de sources différentes. Premièrement, il existe des différences de concepts ou de contenu. Par exemple, dans les enquêtes-ménages comme l'EDTR, on utilise habituellement le concept de «revenu en espèces» qui, par définition, exclut de nombreux avantages sociaux qui sont une composante des salaires et des traitements, ce qui n'est pas le cas dans les données fiscales. Dans l'EDTR, on demande aux répondants de déclarer des éléments de revenu non imposables qui ne figurent pas sur la déclaration de revenus, mais dont certains peuvent aussi être estimés. Deuxièmement, si les différences de contenu sont minimales ou qu'elles peuvent être éliminées, le fait de combiner des données recueillies selon des méthodes différentes pose-t-il d'autres problèmes? Il est possible que les données d'enquête ne soient pas aussi homogènes qu'elles puissent paraître, puisque parmi ces données certaines auront été obtenues en faisant appel à la mémoire des répondants, d'autres auront été tirées de déclarations de revenus et d'autres encore auront été fournies par un enquêté substitut. Par conséquent, les résultats ne seront pas affaiblis par l'ajout de données de meilleure qualité tirées intégralement de déclarations de revenus. Cependant, la question mérite d'être examinée plus à fond et nous aimerions savoir ce qu'en pensent les utilisateurs.

§§§§§§§§§§§§§§§§

**PROTECTION DE LA  
CONFIDENTIALITÉ DES  
MICRODONNÉES DE L'EDTR**

§ En prévision de la diffusion des données, certaines recherches de fond ont été faites sur les principes de la confidentialité des données de l'EDTR diffusées sous forme de fichiers de microdonnées. L'identité de tous les répondants doit rester confidentielle, et nous devons avoir la certitude qu'elle ne pourra être dévoilée, directement ou indirectement. La confidentialité des données des enquêtes longitudinales est plus difficile à garantir parce que les exigences de confidentialité doivent être respectées non seulement au moment de chaque collecte, mais aussi pendant un certain nombre d'années.

Les fichiers de microdonnées de l'EDTR constituent certes un produit important de l'enquête. Toutefois, nous sommes conscients que certains utilisateurs auront besoin de données encore plus détaillées. Ces utilisateurs devront choisir entre deux solutions :

- Un service de produits personnalisés sera offert. Avant d'être envoyés aux clients, ces produits seront filtrés pour éliminer les éléments d'information confidentiels.
- Des travaux préparatoires sont en cours pour donner aux utilisateurs un moyen quelconque d'accéder à distance aux fichiers de microdonnées non filtrés de l'EDTR. Cette approche diffère de la précédente par le fait que c'est l'utilisateur qui écrit et présente le programme d'extraction des données. Ces dernières seront quand même filtrées avant d'être diffusées au client.

La formulation d'un plan de protection de la confidentialité des fichiers de microdonnées a pour but d'éviter que l'identité d'une personne ne soit divulguée par une variable, ou une combinaison de variables, de l'EDTR. Deux approches de base sont généralement proposées pour protéger la confidentialité : la modification des données (qui consiste à modifier les vraies valeurs pour protéger l'identité d'une personne tout en compromettant le moins possible la validité des analyses) et la réduction des données (qui consiste à réduire le contenu, par exemple en groupant des réponses ou en supprimant carrément une variable). Comme la modification des données tend à biaiser avec le temps la mesure des changements dans les caractéristiques personnelles, la réduction des données sera utilisée pour éliminer la plupart des risques de divulgation dans le cas de l'EDTR.

La réduction des données peut prendre bien des formes. Pour qu'un maximum de données soient disponibles, différentes approches seront combinées, les principales étant :

- l'exclusion de certaines variables du fichier, lorsque le risque de divulgation est élevé ou lorsque les variables en cause ont un intérêt analytique limité. Par exemple, le nom du répondant et l'emplacement précis de sa résidence.
- le recodage selon une structure anonyme. Les variables liées au plan d'échantillonnage, par exemple la strate, sont nécessaires pour un calcul précis de la variance. Pour les besoins de l'analyse, il suffit de savoir quels répondants font partie d'une même strate. Toutefois, la structure de codage laisse filtrer certains renseignements géographiques. Le recodage de la strate selon une structure «sans signification» ne fait rien perdre à l'analyste, mais réduit le risque de divulgation.
- l'imposition d'une population minimale. Une taille minimale de l'échantillon est établie pour diverses sous-populations visées par des sources externes de données ou facilement reconnaissables par quelqu'un de l'extérieur.
- le groupement. Les valeurs numériques sont groupées de sorte que la variable qui en résulte est catégorique (ou qualitative).
- le fusionnement des catégories. Approche semblable à la précédente, sauf qu'il s'agit de réduire le nombre de catégories en groupant celles qui comportent des variables catégoriques.
- le recodage des valeurs extrêmes. Les valeurs extrêmes sont habituellement rares dans la population, ce qui augmente le risque de divulgation. Toutes les valeurs extrêmes reçoivent un nouveau code qui indique simplement que la vraie valeur est extrême.
- l'arrondissement. Les valeurs, en particulier les valeurs monétaires, peuvent être arrondies à un niveau supérieur. Un tel arrondissement a peu d'incidence sur l'analyse, mais réduit considérablement le risque de divulgation.

L'équipe de l'EDTR cherche à définir des procédures en matière de confidentialité pour l'ensemble des données de l'enquête, qui s'étalera sur six ans, même si le premier fichier de microdonnées ne contiendra que les données de l'interview préliminaire et la première année de données sur le travail et le revenu. Les prochains fichiers de données cumulatives nécessiteront quelques manipulations avant la diffusion, mais il ne sera pas nécessaire de refaire toutes les évaluations chaque année. Pour plus de détails, voir le document de recherche n° 94-14 ou communiquer avec Pierre Lavallée au (613) 951-2892.

SSSSSSSSSSSSSSSS

**DOCUMENTS DE RECHERCHE**

§ Voici la liste des documents de recherche publiés récemment et qui peuvent être commandés séparément ou reçus d'office. Pour plus de détails, communiquer avec Anne Palmer, par téléphone au (613) 951-2903, par télécopieur au (613) 951-3253, ou par la poste à l'adresse suivante : 11-D8, immeuble Jean-Talon, Parc Tunney, Ottawa, K1A 0T6. Les utilisateurs du réseau Internet peuvent nous joindre en tapant GILES@STATCAN.CA.

*94-08 Le questionnaire et les procédures de collecte des données sur le revenu de l'EDTR - Mai 1994  
Élaine Fournier, Susan Poulin*

En mai 1994, une interview assistée par ordinateur a permis de recueillir les données sur le revenu de l'EDTR auprès du premier panel de répondants. Contrairement à d'autres interviews de l'EDTR, les répondants reçoivent par la poste un questionnaire papier avant l'appel téléphonique de l'intervieweur, ce qui leur donne le temps de consulter leurs dossiers financiers. Ce document décrit les procédures de collecte des données et le libellé des questions sur le revenu, et justifie dans une certaine mesure les décisions prises. Il contient un exemplaire du questionnaire posté aux répondants et résume les procédures suivies par l'intervieweur pour recueillir l'information.

94-09 *«Questionnaire» de 1994 : modules sur le contact et sur les caractéristiques démographiques*

*Ruth Dibbs, Bob Loverock, Alison Hale*

94-10 *«Questionnaire» de 1994 : interview préliminaire*

*Alison Hale, Debbie Lutz, Mike Brule*

Ces deux documents présentent le «questionnaire» papier, avec le libellé des questions, les catégories de réponse et l'enchaînement des questions pour la collecte des données de l'EDTR par interview assistée par ordinateur. Les questions des modules sur le contact et sur les caractéristiques démographiques sont comprises dans les deux interviews annuelles, soit celle sur le travail et celle sur le revenu. L'information sur l'appartenance au ménage est recueillie par le module sur le contact. Le module sur les caractéristiques démographiques comprend des questions sur la date de naissance, le sexe, l'état matrimonial et les liens entre les membres du ménage. L'interview préliminaire est administrée aux nouveaux membres du ménage et aux personnes qui ont eu 15 ans durant l'année de référence.

94-14 *Confidentialité des microdonnées de l'EDTR : approche générale pour la mise en oeuvre*

*Pierre Lavallée, Chantal Grondin*

Ce document donne un aperçu général des principes régissant la confidentialité et montre comment ces principes sont mis en application. L'orientation générale proposée pour l'EDTR est également décrite. Les faits nouveaux à cet égard seront exposés en détail dans d'autres documents de recherche.