

**N° de catalogue 94-11**

**UTILISATION DES DONNÉES FISCALES DANS  
L'ENQUÊTE SUR LA DYNAMIQUE DU TRAVAIL ET DU  
REVENU : RAPPORT SOMMAIRE**

Septembre 1994

Ruth Dibbs, Division des enquêtes-ménages

Susan Poulin, Division des enquêtes-ménages

Maryanne Webber, Division des enquêtes-ménages

La série de documents de recherche de l'EDTR est conçue en vue de communiquer les résultats des études ainsi que les décisions importantes ayant trait à l'Enquête sur la dynamique du travail et du revenu. Ils sont offerts en français et en anglais. Pour obtenir une description sommaire des documents disponibles ou un exemplaire de ces documents, communiquez avec Philip Giles, EDTR, par la poste à Édifice Jean-Talon, 11<sup>ème</sup> étage, section D8, Statistique Canada, Ottawa (Ontario), Canada, K1A 0T6; par INTERNET: GILES@STATCAN.CA; par téléphone au (613) 951-2891; ou par télécopieur au (613) 951-3253.



## **SOMMAIRE**

L'interview sur le revenu de l'EDTR a pris pour modèle l'Enquête sur les finances des consommateurs (EFC). Les répondants doivent donner le revenu qu'ils ont reçu au cours de l'année civile précédente, pour chacune des 20 à 25 sources identifiées. La plupart de ces informations sont aussi déclarées par ces individus lorsqu'ils complètent leur déclaration de revenus. Conséquemment, l'EDTR et l'EFC prennent toutes deux avantage du fait que les répondants ont une connaissance accrue de leurs revenus à ce moment de l'année et recueillent donc les données au printemps.

Une étude a été réalisée pour évaluer la faisabilité de consulter les dossiers fiscaux plutôt que de recueillir les données lors d'une enquête. Ce rapport évalue ces questions. Il recommande que l'EDTR procède avec cette approche pour la collecte des données de 1995, même si d'autres études seront nécessaires pour vérifier si l'approche devrait continuer.



## TABLE DES MATIÈRES

	Page
1. Introduction	1
2. Scénario opérationnel	2
3. Conséquences pour le répondant	5
4. Conséquences sur la qualité des données	6
4.1 Principales utilisations des données sur le revenu de l'EDTR	7
4.2 Non-réponse et erreurs de réponse dans les données d'enquête	8
4.3 Couverture	11
4.4 Erreurs de réponse dans les dossiers fiscaux	11
4.5 Érosion de l'échantillon	12
4.6 Intégration des données d'enquête et des données fiscales	12
4.7 Évaluation globale de l'effet de la méthode mixte sur la qualité des données	16
5. Coûts	16
6. Stratégie de mise en oeuvre	17
6.1 Considérations temporelles	18
6.2 Élaboration et essai des textes à l'intention des répondants	19
6.3 Simulation des résultats de la méthode mixte à l'aide des données du premier cycle	19
6.4 Processus d'approbation	20
7. Conclusions et recommandations	20



## 1. INTRODUCTION

Pour l'Enquête sur la dynamique du travail et du revenu (EDTR), on communique avec les répondants deux fois par année. Le but du second contact, qui a lieu en mai, est de recueillir des données sur le revenu touché pendant l'année précédente. Avril et mai sont les meilleurs mois pour la collecte de ces données, car il y a plus de chances à ce moment-là que les répondants aient à portée de la main leur déclaration de revenus. Les données sur le revenu sont de meilleure qualité quand les répondants consultent leur déclaration.

Au lieu de recueillir les données sur le revenu au moyen d'une interview, nous pourrions le faire par appariement de microdonnées, c'est-à-dire en couplant les données de l'interview sur l'emploi aux données sur le revenu que renferment les dossiers fiscaux. Cette méthode pourrait à la fois accroître la qualité des données et réduire le fardeau de réponse.

Si nous adoptons cette méthode, il faudra obtenir des répondants l'autorisation de consulter leur dossier fiscal. Nous avons de très bonnes raisons de penser que les répondants qui nous donneront leur autorisation ne seront pas suffisamment nombreux pour que nous puissions tout simplement ne pas tenir compte de ceux qui nous la refuseront (ou recourir à l'imputation dans leur cas). Par conséquent, dans cette étude de faisabilité, nous évaluons en fait une méthode mixte qui consisterait à recueillir les données sur le revenu par la méthode du couplage des données pour les répondants qui autoriseraient la consultation de leur dossier fiscal et par la méthode de l'interview pour les autres. Les données obtenues par interview et celles tirées des dossiers fiscaux seraient intégrées dans un même ensemble de données.

Des préoccupations de deux ordres nous amènent à évaluer l'utilisation des dossiers fiscaux pour la collecte des données sur le revenu : la non-réponse dans les enquêtes sur le revenu et le manque de précision des données sur le revenu recueillies par interview. Ces problèmes sont potentiellement plus graves dans une enquête longitudinale que dans une enquête transversale. En raison du peu de temps dont nous disposons pour les travaux d'élaboration, nous n'avons pu évaluer l'utilisation des dossiers fiscaux avant la tenue du premier cycle de l'EDTR. Dans les pages qui suivent, nous examinons donc la faisabilité de l'utilisation des dossiers fiscaux en vue de la mise en oeuvre éventuelle de cette approche dans le second cycle de l'enquête.

## **2. SCÉNARIO OPÉRATIONNEL**

Quand et comment devrions-nous demander aux répondants s'ils nous autorisent à consulter leurs dossiers fiscaux? Nous avons évalué diverses options et conclu que le scénario que nous décrivons dans les paragraphes qui suivent était le meilleur. (Dans ce scénario, nous nous réservons la possibilité de poser les questions sur le patrimoine en mai 1995, mais, comme nous le verrons, nous serons peut-être obligés de faire un compromis sur ce point.)

- En janvier 1995, on effectuerait les interviews sur le travail du second cycle de l'enquête. À ce moment-là, l'intervieweur ne parlerait pas du projet d'utilisation des dossiers fiscaux au répondant; il lui dirait seulement qu'il rappellerait en mai.
- Avant l'interview de mai, on enverrait aux répondants un questionnaire (visant à faciliter l'interview téléphonique) ainsi qu'un document dans lequel nous leur expliquerions le projet d'utilisation des dossiers fiscaux. Dans le

questionnaire, il y aurait une case que les répondants cocheraient pour indiquer qu'ils consentent à ce que l'on consulte leur dossier fiscal. Pour ce qui est des répondants qui préféreraient fournir les renseignements demandés dans une interview téléphonique, on leur demanderait de remplir le questionnaire avant l'appel de l'intervieweur et de le garder près du téléphone jusqu'à cet appel.

- Quand l'intervieweur appellerait, il demanderait au répondant s'il autorise la consultation de son dossier fiscal pour la collecte des données sur le revenu. Si le répondant refusait, l'intervieweur ferait l'interview sur le revenu. En principe, les questions sur le patrimoine pourraient être posées et aux personnes ayant autorisé la consultation de leur dossier fiscal et à celles l'ayant refusée.
- Dans les cycles suivants, on n'entrerait pas en communication en mai avec les ménages dans lesquels tous les membres auraient autorisé la consultation de leur dossier fiscal. Si des données sur le patrimoine devaient être recueillies (à l'heure actuelle, nous recommandons de recueillir ces données aux trois ans, dans le deuxième et le cinquième cycle de chaque panel), elles le seraient à l'occasion de l'interview de janvier.
- En ce qui concerne les nouveaux membres du ménage, on leur demanderait au moment de la première interview sur le travail s'ils autorisent la consultation de leur dossier fiscal. Dans les panels suivants, on demanderait aux répondants longitudinaux s'ils autorisent la consultation de leur dossier fiscal au moment de la première interview sur le travail. Par exemple, dans le cas des répondants du second panel, cette autorisation serait demandée en janvier 1997.

Si nous posons les questions sur le patrimoine en mai 1995, nous aurons probablement moins de chances d'obtenir l'autorisation des répondants. En effet, les questions sur le patrimoine et les questions sur le revenu portant toutes sur des sommes d'argent, nous risquons d'embrouiller et d'irriter les répondants ayant donné leur consentement, et ces derniers pourraient changer d'avis. Une façon de contourner le problème serait de poser les questions sur le patrimoine dans une interview ultérieure, par exemple celle de janvier 1998 (autrement dit, il s'agirait de supprimer la première observation sur le patrimoine pour les membres du premier panel). Les diverses options possibles en ce qui a trait au patrimoine sont en train d'être étudiées dans une évaluation distincte. On y examinera de plus près l'inclusion de questions sur le patrimoine dans l'EDTR.

Pouvons-nous trouver une façon de bien expliquer l'utilisation des dossiers fiscaux aux répondants de manière à ce que ces derniers comprennent ce qu'engage cette option? Nous avons rédigé certains textes à l'intention des répondants et des intervieweurs et nous pensons qu'il est possible de trouver des formulations claires, qui ne risquent pas d'être irritantes, pour expliquer aux répondants les options qui s'offrent à eux. Voici les mesures que nous proposons :

- Envoi d'une brochure aux répondants avant l'interview de mai 1995 (la brochure serait envoyée en même temps que le questionnaire sur le revenu).
- Inclusion dans le questionnaire sur le revenu d'un énoncé autorisant la consultation du dossier fiscal et, en regard de cet énoncé, d'une case à cocher.

- Inclusion dans l'application d'IAO (interview assisté par ordinateur) de la question que l'intervieweur poserait au répondant pour lui demander l'autorisation de consulter son dossier fiscal.
- Distribution des documents didactiques et de questions et réponses aux intervieweurs et au personnel des services consultatifs.

Les essais faits jusqu'à maintenant ont montré qu'entre 50 et 60 % des répondants consentiraient à l'utilisation de leur dossier fiscal. Toutefois, tous les essais ont été faits en posant des questions hypothétiques aux répondants et sans explication préalable de cette option. Pour obtenir le taux de consentement le plus élevé possible, nous devons apporter beaucoup de soin à l'élaboration des textes destinés aux répondants, puis en faire l'essai. Nous prévoyons devoir faire appel à des consultants de l'extérieur pour la conception de textes efficaces.

### **3. CONSÉQUENCES POUR LE RÉPONDANT**

Un des principaux avantages de l'utilisation des dossiers fiscaux est la réduction du fardeau de réponse. Cet allègement se ferait de trois façons :

- Pour le premier panel, le nombre d'interviews serait réduit de 13 à 8. Pour les panels suivants, le nombre d'interviews ne serait que de 7 (l'interview préliminaire et 6 interviews sur le travail).
- Nous n'aurions plus à demander au répondant de *se préparer* pour l'interview, c'est-à-dire de consulter ses documents et de remplir à l'avance le questionnaire.

- En supprimant l'interview sur le revenu, nous supprimerions des questions que beaucoup de répondants trouvent délicates et difficiles. La preuve que ces questions sont délicates, c'est que les taux de réponse aux enquêtes sur le revenu sont inférieurs aux taux pour d'autres enquêtes. Pour ce qui est de la difficulté des questions, les répondants qui ont participé à l'essai du recensement national et qui ont rempli le questionnaire de compte rendu ont eux-mêmes déclaré qu'ils trouvaient difficile de répondre à des questions sur le revenu.

Pour les répondants, le fait d'avoir un choix est un élément positif. La déclaration de données sur le revenu suscite des craintes de divers ordres. Si nous pouvons offrir un choix aux répondants, ils seront probablement mieux disposés à collaborer. Ce serait également bon pour les intervieweurs d'avoir un choix à proposer.

#### **4. CONSÉQUENCES SUR LA QUALITÉ DES DONNÉES**

Outre la réduction du fardeau de réponse, nous nous attendons que l'utilisation des dossiers fiscaux améliore nettement la qualité des données. Comme les données de l'EDTR servent essentiellement à des analyses longitudinales, il est important de réduire au maximum les erreurs de réponse qui auraient pour effet de gonfler artificiellement le nombre de familles ayant une variation de revenu. C'est parce que l'on craignait la surdéclaration de changements d'états que l'on a appliqué la technique de la rétro-information dans l'interview sur le travail. Toutefois, le recours à cette technique (il s'agit pour l'intervieweur de rappeler au répondant les réponses fournies par ce dernier un an plus tôt) ne serait pas approprié dans une interview sur le revenu.

L'utilisation des dossiers fiscaux comporte aussi certains inconvénients. Nous verrons plus loin le pour et le contre de cette approche. Mais au préalable, pour situer dans son contexte la question de la qualité des données, nous examinerons dans la section 4.1 les principales utilisations auxquelles les données se prêteront.

#### **4.1 PRINCIPALES UTILISATIONS DES DONNÉES SUR LE REVENU DE L'EDTR**

Nous ne pouvons pas prévoir exactement comment les données seront utilisées. Toutefois, d'après les discussions que nous avons eues avec des chercheurs et compte tenu des utilisations qui ont été faites des données d'autres enquêtes longitudinales sur le revenu, nous prévoyons que les données de l'EDTR serviront à l'étude des questions suivantes.

Il se fera probablement des études sur la *mobilité économique des familles*, soit des études semblables à celles faites par le passé à partir des résultats de deux enquêtes américaines, la *Panel Study of Income Dynamics (PSID)* et la *Survey of Income and Program Participation (SIPP)*. Dans ces études, on examine à un niveau micro-comparatif les changements dans le revenu familial, on classe les familles selon la stabilité de leur revenu et on cherche à déterminer les corrélats des augmentations ou des diminutions de revenu. Parmi les corrélats étudiés, mentionnons l'activité et les événements relatifs à l'activité (p. ex. perte d'emploi et changement dans la quantité globale de travail fait par l'unité familiale) de même que les caractéristiques et les événements démographiques (p. ex. dissolution de la famille, familles reconstituées, mobilité géographique).

La *dynamique du faible revenu* est un sujet de recherche connexe. Certains chercheurs voudront quantifier les mouvements d'entrée et de sortie par rapport

aux seuils de faibles revenus. Par exemple, à quelle rythme se renouvelle la population à faible revenu? Y a-t-il une sous-strate de la population qui souffre de "pauvreté chronique"? Du point de vue de la politique sociale, ce sujet de recherche peut présenter beaucoup d'intérêt, et il se peut donc que l'on veuille définir les caractéristiques des familles qui vivent sous le seuil de faible revenu pendant de longues périodes. En outre, on voudra probablement déterminer dans quelle mesure les paiements de transfert (en particulier les prestations d'assurance-chômage et d'aide sociale) permettent de prévenir les épisodes de faible revenu ou aident à y mettre fin.

Enfin, nous prévoyons qu'il se fera des études sur l'interaction entre le revenu de la famille et l'activité des membres de la famille. Par exemple, le revenu de la famille influence-t-il le mouvement des membres entre travailleurs autonomes et travailleurs rémunérés? Quel est l'effet du revenu sur la transition entre l'activité et la retraite? Certains chercheurs étudieront peut-être le comportement des bénéficiaires de l'assurance-chômage ou de l'aide sociale face à leur activité sur le marché du travail.

À la lumière de ce qui précède, il est clair que les données qui susciteront le plus d'intérêt sont celles portant sur le revenu de la famille et sur la variation du revenu dans le temps.

#### **4.2 NON-RÉPONSE ET ERREURS DE RÉPONSE DANS LES DONNÉES D'ENQUÊTE**

La première interview sur le revenu de l'EDTR a eu lieu en 1994. Le taux de réponse obtenu était de 83 %. De plus, 6 % des enquêtés, bien que n'ayant pas refusé catégoriquement de fournir des données, n'ont répondu à aucune des questions sur le revenu. Lorsque l'on tient compte de ces enquêtés, le taux de

réponse n'est plus que de 77 %. L'utilisation des dossiers fiscaux pourrait nous permettre d'accroître le taux de réponse et, de cette façon, d'améliorer la qualité des données.

Outre le problème de la non-réponse, l'examen des données sur le revenu recueillies au moyen d'une interview indique qu'il y a une sous-déclaration de certaines sources de revenu et sous-déclaration des montants (par rapport à ce qui se passe lorsque les données proviennent des dossiers fiscaux). Par exemple, les chiffres recueillis au moyen de l'EFC correspondent à 80 % du montant réel des prestations d'assurance-chômage, comparativement à 94 % pour les dossiers fiscaux. On observe également une tendance à la sous-déclaration et un biais dans le cas du revenu de placements.

Une autre lacune des données d'enquête est que les distributions du revenu constituent parfois de "pointes", lesquelles sont attribuables au fait que les répondants donnent des estimations approximatives. Des études faites à partir d'enquêtes canadiennes et américaines révèlent que ces pointes peuvent être importantes. Ce genre d'erreur peut, entre autres, poser des problèmes en ce qui a trait à l'établissement de catégories de revenu.

Mais ce qui importe davantage, c'est que, pour étudier la dynamique du faible revenu et la mobilité économique des familles à un micro-niveau, il faut disposer de données fiables sur les variations du revenu de la famille. Les erreurs de déclaration aléatoires dans n'importe laquelle des interviews auprès d'un ménage peuvent entraîner une surestimation de la variation du revenu. C'est ce que montrent les données de l'EDTR recueillies auprès d'un échantillon d'environ 1 400 ménages interviewés deux années d'affilée. Pour les unités de cet échantillon (interviewées dans le cadre de l'EFC, puis réinterviewées un an plus tard lors du test 3 de l'EDTR), nous avons effectué un appariement de microdonnées : nous

avons comparé, pour les années 1991 et 1992, les données recueillies par interview aux données contenues dans les dossiers fiscaux. Voici les résultats en ce qui concerne les personnes ayant participé aux deux enquêtes et pour lesquelles on a réussi le couplage données d'enquête/données fiscales pour les deux années considérées :

- D'après les données d'enquête, 14 % des répondants ont touché un revenu au cours d'une année, mais pas de l'autre. D'après les dossiers fiscaux de ces mêmes répondants, le pourcentage de ceux ayant connu un tel changement de situation n'était que de 3 %.
- De fortes baisses du revenu total sont plus fréquemment observées dans les données d'enquête que les données fiscales. Les données des deux enquêtes indiquaient que plus de 15 % des répondants avaient subi une baisse de revenu supérieure à 25 % entre 1991 et 1992, alors que, d'après les données fiscales, la proportion ayant subi une baisse de cet ordre était de 10,5 % seulement. En outre, d'après les données d'enquête, 21 % des ménages avaient subi une baisse de revenu d'au moins 2 500 \$ en chiffres absolus, alors que, d'après les données fiscales, cette proportion n'était que de 15 %.
- Selon les données d'enquête, 16 % des répondants avaient touché un salaire ou un traitement en 1991 ou en 1992, mais pas les deux années, comparativement à 9 % selon les données fiscales.

Ces quelques résultats donnent du poids aux craintes exprimées au sujet de l'effet des erreurs de déclaration dans les analyses longitudinales<sup>1</sup>.

---

<sup>1</sup> Ces résultats doivent être interprétés avec prudence parce que l'échantillon était petit et parce que l'EDTR (mais non l'EFC) a été

### **4.3 COUVERTURE**

La couverture des dossiers fiscaux est-elle adéquate? Dans le passé, le sous-dénombrement de certaines parties de la population (en particulier les faibles revenus, les personnes âgées et les conjoints qui ne travaillent pas) aurait constitué un obstacle majeur au couplage que nous envisageons actuellement. Et, bien que la couverture ne soit toujours pas complète, elle est bien meilleure qu'elle ne l'était, peut-être en partie en raison de l'introduction des crédits d'impôt. En 1994, les dossiers fiscaux couvraient 94 % environ de la population de 20 ans ou plus, comparativement à 85 % en 1984.

Lorsque l'on proposera le couplage données d'enquête/données fiscales aux répondants, on leur demandera s'ils ont produit une déclaration d'impôt l'année précédente. Dans la négative, l'intervieweur fera l'interview sur le revenu. Cette façon de faire permettrait de compenser dans une large mesure le sous-dénombrement des dossiers fiscaux.

### **4.4 ERREURS DE RÉPONSE DANS LES DOSSIERS FISCAUX**

Ici, le problème porte essentiellement sur l'évasion fiscale. Certains chercheurs nous ont recommandé de ne pas nous servir de données fiscales ni même de demander aux répondants de consulter leurs dossiers fiscaux parce que cela les inciterait à ne pas déclarer les revenus tirés de l'économie souterraine. Nous estimons pour notre part que, quelle que soit la méthode employée, nous ne

---

menée selon la méthode d'interview assistée par ordinateur. En outre, il y avait des différences entre les questions sur le revenu de l'EFC et celles de l'essai de l'EDTR. Les résultats que nous avons examinés portent sur des sources de revenu comparables du point de vue des concepts. Comme l'objet de l'analyse était d'examiner la réponse, nous avons utilisé des données non vérifiées.

pouvons pas recueillir de données fiables sur le travail au noir. De façon générale, les gens ne nous fourniront pas de données sur un travail (ou des revenus) qu'ils ont l'intention de dissimuler dans leur déclaration de revenus. Cette opinion est confirmée par les résultats de groupes de concentration organisés pour l'EDTR de 1992.

#### **4.5 ÉROSION DE L'ÉCHANTILLON**

L'érosion de l'échantillon constitue l'un des principaux problèmes dans les enquêtes par panel. Dans l'EDTR, le fait que le taux de non-réponse soit plus élevé pour l'interview sur le revenu que pour l'interview sur le travail pourrait aggraver le problème. Autrement dit, les personnes qui auront refusé de participer à l'interview sur le revenu vont peut-être être davantage portées, dans le cycle suivant, à refuser de participer à l'interview sur le travail. En réduisant le nombre d'appels et en supprimant le volet délicat des questions sur le revenu, on pourrait réduire l'érosion et ainsi accroître la qualité des données.

#### **4.6 INTÉGRATION DES DONNÉES D'ENQUÊTE ET DES DONNÉES FISCALES**

La méthode mixte de collecte des données comporte indéniablement des inconvénients. Les voici, en résumé :

- Certains répondants pourraient accepter que l'on consulte leur dossier fiscal, puis changer d'avis. Il s'ensuit qu'à un niveau micro-comparatif, nous aurions des données provenant de deux sources pour une même personne.

- Les catégories de revenu établies pour les interviews ne correspondent pas tout à fait à celles figurant dans les déclarations de revenus, et ce manque de cohérence dans les concepts peut compromettre la qualité des données. Les dossiers fiscaux ne contiennent pas de données sur certains types de revenus non imposables qui peuvent être tout particulièrement importants dans le cas des familles à faible revenu.
- Les chercheurs se sont habitués à composer avec les lacunes de divers ensembles de données, mais l'intégration de données recueillies selon deux méthodes différentes est pour eux une nouveauté et une inconnue.

Toutefois, il faut souligner que la qualité des données recueillies par interview n'est pas uniforme non plus. On demande aux répondants de consulter leur déclaration de revenus pour se préparer à l'interview. Certains d'entre eux le font soigneusement, et les renseignements qu'ils fournissent en interview sont aussi précis que ceux figurant dans leur dossier fiscal. En revanche, les intervieweurs ont également affaire à des personnes qui ne fournissent que des estimations approximatives, et cela *par réponse interposée*. Parmi les répondants qui se sont prêtés à l'interview sur le revenu du test 3 de l'EDTR, 37 % ont rempli le questionnaire à l'avance, 17 % ont consulté leur déclaration de revenus pendant l'interview et 46 % ont fait l'interview au pied levé. Comme on pouvait s'y attendre, il y avait des différences entre ces trois groupes de répondants pour ce qui est du revenu moyen et du nombre de sources de revenu déclarées. Ainsi, la qualité des données serait loin d'être uniforme même si tous les renseignements étaient recueillis par interview.

Un autre point important est le suivant : le succès de la méthode mixte dépend de la mesure dans laquelle nous réussissons à bien intégrer les données de l'une et

l'autre source. Comment nous y prendrions-nous pour harmoniser les concepts? Dans les paragraphes qui suivent, nous donnons un résumé des travaux faits sur cette question jusqu'à maintenant.

En gros, l'approche que nous proposons pour la présentation des microdonnées est la suivante :

- Les données (qui feront l'objet de certaines vérifications) seront présentées selon la méthode utilisée pour les recueillir. Ainsi, pour chaque répondant, on présentera soit des données fiscales, soit des données d'enquête.
- En outre, pour chaque répondant, on présentera des données intégrées sur le revenu. (c'est-à-dire, un nouvel ensemble de données qui ne sera ni des données d'enquête ni des données fiscales mais plutôt quelque chose que l'on définira)
- Le nombre de catégories de revenu sera un peu moins élevé dans le cas des données intégrées que dans le cas des données produites à partir de l'une et l'autre source de données.
- Dans toute la mesure du possible, les données intégrées seront produites en fonction des concepts de l'enquête. Autrement dit, nous allons tâcher d'adapter les concepts qui sous-tendent les données fiscales aux concepts de l'enquête, plutôt que de faire l'inverse<sup>2</sup>.

---

<sup>2</sup> Nous continuerons toutefois d'utiliser les données fiscales à des fins d'imputation (sauf, évidemment, pour les sources de revenu qui ne figurent pas dans les dossiers fiscaux).

- Les données intégrées feront l'objet de corrections pour en améliorer la comparabilité.

À la lumière des résultats d'un premier examen, les données intégrées pourraient rentrer dans les catégories que voici :

- salaires et traitements
- revenu d'un travail autonome agricole
- revenu d'un travail autonome non agricole
- revenu de placements
- paiements de transferts
  - prestation fiscale pour enfants
  - PSV/SRG/AC
  - RPC/RRQ
  - prestations d'assurance-chômage
  - prestations d'aide sociale et supplément de revenu provinciaux
  - indemnités d'accident du travail
  - crédit pour taxe sur les produits et services
  - crédits d'impôt provinciaux
  - autres revenus de sources gouvernementales
- revenu de pension
- pension alimentaire, allocation de séparation et allocations de soutien d'un enfant
- autre revenu
- revenu total
- impôt payé sur le revenu

Si nous adoptons le projet proposé, nous allons devoir définir les critères d'intégration des deux ensembles de données et rédiger des documents dans

lesquels nous décrivons ces critères pour les chercheurs. Nous avons l'intention de simuler le processus d'intégration à partir des données du premier cycle. Nous sommes actuellement en train de coupler ces données aux données fiscales pour évaluer la qualité des résultats.

#### **4.7 ÉVALUATION GLOBALE DE L'EFFET DE LA MÉTHODE MIXTE SUR LA QUALITÉ DES DONNÉES**

Par rapport aux autres méthodes, la méthode mixte est plus complexe et comporte une quantité inconnue, en ce sens qu'elle réunit les point forts et les points faibles des méthodes de collecte de données via interview et via le couplage avec les données fiscales. Malgré ces désavantages, la méthode mixte permettrait probablement d'améliorer la qualité des données, grâce à la réduction de l'érosion de l'échantillon, à la possibilité d'une collaboration accrue des répondants, à l'amélioration du moral des intervieweurs et à la plus grande précision des données tirées des dossiers fiscaux.

#### **5. COÛTS**

Parce qu'il faut recourir à une méthode mixte pour la collecte des données, les économies possibles sont limitées. Les frais généraux associés à la tenue d'interviews sur le revenu resteraient virtuellement les mêmes, mais il y aurait réduction de la durée des interviews. Cette économie serait annulée par l'augmentation des coûts du traitement des données (dont le traitement sur ordinateur central) pour le couplage des données de l'EDTR et des données fiscales.

Si l'on suppose que 50 % des ménages consentiraient au couplage, et c'est là une hypothèse assez prudente, l'économie réalisée au poste de la collecte des données serait d'environ 20 000 \$ en 1995. (Le coût total de la collecte des données pour l'interview sur le revenu de 1994, c'est-à-dire le premier cycle d'interviews auprès des répondants du premier panel, était de 375 000 \$ environ). L'économie prévue pour 1995 serait petite parce qu'il faudrait communiquer avec chaque ménage.

En 1996, on ne communiquerait qu'avec les ménages ayant refusé la consultation de leur dossier fiscal et l'on estime que, cette année-là, l'économie serait de 130 000 \$ environ. En 1997, en supposant que l'on introduise dans l'échantillon un panel de la même taille, l'économie serait plus importante, de 235 000 \$ environ, et elle se maintiendrait à ce niveau dans les années suivantes.

Il est difficile d'estimer les dépenses supplémentaires qu'engagerait la méthode mixte. Avec le temps, le couplage de microdonnées au fichier d'impôt se ferait plus efficacement. Si les hypothèses au sujet des économies possibles sont justes, cela ne suffirait pas, en soi, à justifier l'adoption de cette méthode. Toutefois, ces économies ont été calculées à partir d'hypothèses prudentes et, si nous réussissons à amener une forte majorité des ménages à consentir au couplage, les économies seraient relativement plus élevées.

## **6. STRATÉGIE DE MISE EN OEUVRE**

Dans les paragraphes qui suivent, nous examinons les questions sur lesquelles nous devons nous pencher avant d'aller plus avant dans ce projet.

## 6.1 CONSIDÉRATIONS TEMPORELLES

On ne peut pas connaître les résultats d'une expérience avant de l'avoir faite! C'est la même chose pour la méthode du couplage : pour connaître le nombre réel de ménages qui vont consentir au couplage données d'enquête/données fiscales, pour définir les caractéristiques des ménages consentants et des ménages non consentants et pour déterminer l'effet du couplage sur l'érosion de l'échantillon, il faut obligatoirement appliquer la méthode. Toutefois, à partir du moment où nous proposons au répondant de recueillir les données sur le revenu par couplage des enregistrements et que ce dernier nous donne son consentement, nous ne pouvons pas faire marche arrière. Autrement dit, si nous proposons aux enquêtés la méthode du couplage en 1995 et que 30 % seulement de ces derniers acceptent, nous sommes obligés de conserver cette même méthode pour ces 30 % des répondants jusqu'à ce que le premier panel soit remplacé.

En appliquant la méthode en 1995, nous pourrions évaluer les réactions des répondants et régler diverses difficultés de traitement avant l'introduction du second panel dans l'échantillon. Nous disposerions alors des données empiriques voulues pour faire un choix parmi les options suivantes :

- Conserver la méthode mixte -- si cette dernière accroît effectivement le taux de réponse et réduit l'érosion de l'échantillon.
- Retenir la méthode de l'interview seulement -- si les gains du point de vue de la qualité des données ne sont pas suffisants pour justifier la complexité supplémentaire que présente l'adoption de la méthode mixte.
- Envisager des mesures plus radicales, par exemple prélèvement d'une partie de l'échantillon (ou de l'échantillon complet) du second panel dans les

dossiers fiscaux si ni la méthode de l'interview ni la méthode mixte ne permettent d'obtenir des résultats de la qualité voulue.

## **6.2 ÉLABORATION ET ESSAI DES TEXTES À L'INTENTION DES RÉPONDANTS**

Si nous adoptons la méthode du couplage, il est indispensable d'avoir une bonne documentation à offrir aux répondants. Les documents d'information qui seront envoyés aux répondants doivent être suffisamment attrayants pour que ces derniers aient envie de les lire et suffisamment clairs pour que nous puissions être sûrs que nous obtenons un «consentement éclairé». Si l'on adopte la méthode du couplage, il faudra donner, dès cet automne, la priorité absolue à l'élaboration de ces documents.

## **6.3 SIMULATION DES RÉSULTATS DE LA MÉTHODE MIXTE À L'AIDE DES DONNÉES DU PREMIER CYCLE**

On procède actuellement au couplage des données du premier cycle de l'EDTR aux données fiscales de 1993 pour évaluer la qualité des résultats. Le fichier des données couplées pourrait être utilisé pour simuler la méthode mixte :

- en substituant des données fiscales à certaines données d'enquête;
- en appliquant les algorithmes nécessaires pour estimer les éléments manquants;

- en dérivant des variables à partir de l'ensemble intégré des données d'enquête et des données fiscales.

Si l'on faisait cet essai, cela nous permettrait d'acquérir une expérience pratique du traitement des données obtenues selon la méthode mixte et d'évaluer les effets de la méthode sur les données. Toutefois, comme la décision d'adopter ou non la méthode doit être prise prochainement (dans l'hypothèse où la mise en oeuvre se ferait en 1995), nous manquerions de temps pour mener à terme ce travail.

#### **6.4 PROCESSUS D'APPROBATION**

Si la méthode mixte est adoptée, nous allons devoir obtenir l'autorisation du Comité du couplage des enregistrements. Nous supposons qu'il n'y a pas d'autres empêchements à la mise en oeuvre du projet.

### **7. CONCLUSIONS ET RECOMMANDATIONS**

Voici les principaux avantages que l'on prévoit tirer de l'utilisation des dossiers fiscaux des répondants qui nous autoriseront à les consulter :

- la réduction du fardeau de réponse;
- l'amélioration des rapports avec les répondants et du moral des intervieweurs du fait que l'on aurait un choix de méthode de collecte des données sur le revenu à proposer;
- la réduction possible du taux de non-réponse à la partie de l'EDTR ayant trait au revenu et de l'érosion de l'échantillon en général;

- l'accroissement de la précision des données sur le revenu dans le cas des répondants autorisant la consultation de leur dossier fiscal;
- une certaine réduction des coûts;
- la possibilité de bien évaluer les effets de la méthode mixte avant la mise en oeuvre dans le second panel.

Les principaux désavantages sont les suivants :

- une planification d'enquête et des fonctions de traitement des données plus complexes, ce qui viendra accroître le risque d'erreurs;
- un ensemble de données plus complexe, ce qui pourrait poser des difficultés au chapitre de l'interprétation des données et miner la confiance des utilisateurs, la méthode mixte n'étant pas une méthode de collecte établie;
- l'impossibilité dans les prochains cycles de revenir complètement à la méthode actuelle : une fois l'option du couplage acceptée par le répondant, on ne peut pas faire marche arrière.
- la probabilité de ne pas pouvoir recueillir les données sur le patrimoine à deux occasions dans le premier panel.

En résumé, la méthode mixte et l'utilisation des dossiers fiscaux comportent tant d'inconnues qu'il est difficile d'en prévoir les résultats de façon précise. Pour connaître ces résultats, il faut appliquer la méthode. Dans une enquête longitudinale comme l'EDTR, où l'érosion de l'échantillon et le manque de cohérence des réponses dans le temps sont des causes d'inquiétude, l'adoption de la méthode mixte pourrait être très avantageuse du point de vue de la qualité des données.

Toutefois, il ressort clairement que cette approche comporte des risques et que, si nous décidons de l'adopter, il faudra faire un travail énorme pour en maximiser les chances de succès. La production des documents d'information pour les répondants et des textes dont les intervieweurs se serviront pendant les interviews va nécessiter beaucoup de planification. Il est par ailleurs tout aussi important de bien intégrer les données des deux sources, et ce travail risque de prendre beaucoup de temps. Il faudra discuter de cette approche avec les utilisateurs et leur remettre les résultats obtenus à partir de l'ensemble de données créés lors de la simulation de la méthode mixte. Il serait également important d'avoir de la bonne documentation sur les méthodes d'intégration.

Si la mise en oeuvre de la méthode mixte comporte des risques et un surcroît de travail, il ne faut pas oublier qu'un changement fondamental au niveau de la méthode de collecte devient de plus en plus difficile à faire au fur et à mesure que l'enquête prend de l'âge et qu'augmentent les sommes et les efforts consacrés à une méthode particulière. Il serait donc important de régler cette question le plus tôt possible, plus précisément avant l'introduction du second panel.