

*CENSUS FAMILIES**FAMILLES DE RECENSEMENT***SAMPLE DESIGN AND IMPLICATIONS**

The sample employed for the **Survey of Consumer Finances** is the Labour Force Survey sample. This multi-stage stratified clustered probability sample is described in detail in **Methodology of the Canadian Labour Force Survey, 1984-1990**, Catalogue No. 71-526 and the feature articles in **The Labour Force (October 1994, December 1994 and January 1995)** Catalogue No. 71-001. For a more general description, see **Guide to Labour Force Survey Data**, Catalogue No. 71-528.

Briefly, this sample is designed to represent approximately 98% of the population. Excluded population groups are:

1. Residents of the Yukon and Northwest Territories;
2. Residents of Indian reserves;
3. Residents of military barracks; and
4. Inmates of institutions such as prisons, penitentiaries, jails, reformatories, mental hospitals, tuberculosis hospitals, sanatoria, orphanages and homes for the aged.

Because of varying sampling and response rates, each record on the file has a weight attached to it. This weighting factor reflects the sample design and incorporates the inverse of the sampling ratio (which varies significantly by geographic area) and differential response rates for households, among other things. For further information on the weighting procedure, see Catalogue No. 13-207, as well as Catalogue Nos. 71-526, 71-528 and the issues of Catalogue No. 71-001 mentioned above.

The sample design and weighting procedure have important implications for data use. Each variable on the file must be multiplied by the weight in order to produce valid statistical estimates of population parameters such as means, medians, modes, estimated numbers and aggregates. However, the appropriateness of using these sample survey weights in regression analysis is another issue. A useful summary of this issue is presented in DuMouchel and Duncan (1983) "Using Sample Survey Weights in Multiple Regression Analysis of Stratified Samples". *Journal of The American Statistical Association*, Volume 78, December 1983, pp. 535-543.

PLAN DE SONDAGE ET IMPLICATIONS

L'échantillon utilisé pour l'**Enquête sur les finances des consommateurs** est celui de l'Enquête sur la population active. Cet échantillon aléatoire à grappes stratifiées à plusieurs degrés est décrit en détail dans la publication **Méthodologie de l'Enquête sur la population active du Canada, 1984-1990**, n° 71-526 au catalogue et dans les articles vedettes du mois parus dans **La population active (octobre 1994, décembre 1994 et janvier 1995)** n° 71-001 au catalogue. Pour une description plus générale, voir le **Guide d'utilisation des données de l'Enquête sur la population active**, n° 71-528 au catalogue.

En somme, cet échantillon est conçu de manière à représenter environ 98% de la population. Sont exclus les groupes suivants:

1. Les résidents du Yukon et des Territoires du Nord-Ouest;
2. Les résidents des réserves indiennes;
3. Les résidents de baraquements militaires; et
4. Les pensionnaires d'institutions comme les prisons, les pénitenciers, les maisons de correction, les hôpitaux pour malades mentaux, les hôpitaux pour tuberculeux, les sanatoriums, les orphelinats et les foyers pour personnes âgées.

En raison de la variation des fractions de sondage et de taux de réponse, nous avons attribué un poids à chaque dossier du fichier. Ce coefficient de pondération tient compte du plan de sondage et comporte l'inverse de la fraction de sondage (qui varie beaucoup d'une région géographique à l'autre) et, entre autres, différents taux de réponse pour les ménages. Pour obtenir de plus amples renseignements sur la méthode de pondération, voir la publication n° 13-207 au catalogue et la présente documentation ainsi que les publications n°s 71-526, 71-528 au catalogue et les parutions de n° 71-001 au catalogue cité plus haut.

Le plan de sondage et la méthode de pondération comportent d'importantes implications à l'égard de l'utilisation des données. Il faut multiplier chaque variable du fichier par le poids pour obtenir des estimations statistiques valides des paramètres démographiques, comme des moyennes, des médianes, des modes et des estimations d'agrégats. Quoi qu'il en soit, l'intérêt que présente l'utilisation de ces poids de sondage dans l'analyse de régression est un autre sujet. Un sommaire utile sur la question est présenté dans l'article intitulé "Using Sample Survey Weights in Multiple Regression Analysis of Stratified Samples", de DuMouchel et Duncan (1983). *Journal of the American Statistical Association*, Volume 78, December 1983, pp. 535-543.

*CENSUS FAMILIES**FAMILLES DE RECENSEMENT***SAMPLE DESIGN AND IMPLICATIONS
- concluded**

Since a multi-stage stratified clustered sample design was employed, standard error estimates based on the usual simple random formula generally underestimate the true standard errors. For an illustration of the differences, see the uncatalogued paper, "**An Investigation of Variance Estimation in Family Expenditure Surveys**" and the attached bibliography. See also the standard errors of average incomes (which estimate the sampling error) published in Catalogue No. 13-208. Standard errors of percentage distributions for family units by size are published in Catalogue No. 13-208.

The information necessary to estimate the standard errors is not given on this file. Most regression programs are not capable of considering design effects, even if the relevant information were available. However, generalized programs such as SAS, BMDP and SPSS produce coefficients using weighted least squares. The standard errors calculated by SPSS are much too low because the sum of the weights is used for the degrees of freedom. This can be corrected by dividing the weight on each record by the sum of the weights and multiplying by the sample size. This transformation will produce standard errors equivalent to those produced by SAS and BMDP. Specific requests to produce standard errors of estimates can be handled on a cost-recoverable basis. If this procedure is not followed, we suggest that users of regression packages be generally conservative in interpreting measures of statistical significance from these packages - perhaps using a factor of 2 or 3 to inflate the standard errors produced, since the standard errors in these packages are generally of the simple random type.

For further information on sampling and non-sampling errors associated with the survey, see Catalogue No. 13-207.

PLAN DE SONDAGE ET IMPLICATIONS - fin

Comme nous avons utilisé un échantillon à grappes stratifié à plusieurs degrés, les estimations de l'erreur type calculées selon la formule habituelle d'échantillonnage aléatoire simple sous-représentent généralement les erreurs types vraies. Pour une illustration des différences voir le document hors catalogue intitulé "**Estimation de la variance dans les enquêtes sur les dépenses des familles**" et la bibliographie qui s'y rattache. Voir également les erreurs types des revenus moyens (qui estiment l'erreur d'échantillonnage) publiées dans la publication n° 13-208 au catalogue. Les erreurs types des répartitions en pourcentage pour les unités familiales selon la taille figurent dans la publication n° 13-208 au catalogue.

Les renseignements dont nous avons besoin pour estimer les erreurs types ne figurent pas dans ce fichier. La plupart des programmes de régression ne peuvent pas prendre en considération les effets du plan, même si les renseignements pertinents sont connus. Toutefois, les programmes généraux comme le SAS, le BMDP et le PSSS produisent des coefficients au moyen des moindres carrés pondérés. Les erreurs types calculées d'après le PSSS sont beaucoup trop faibles parce qu'on y utilise la somme des poids pour établir les degrés de liberté. On peut remédier à la situation en divisant le poids de chaque dossier par la somme des poids et en multipliant par la taille de l'échantillon. On obtiendra ainsi des erreurs types équivalentes à celles que produisent le SAS et le BMDP. Nous pourrions répondre à ceux qui nous demanderont de produire des erreurs types particulières et ce, moyennant le recouvrement des frais. S'ils ne se prévalent pas de cette possibilité, nous suggérons aux utilisateurs qui se servent de programmes généraux de faire preuve de conservatisme lorsqu'il s'agit d'interpréter les mesures de la signification statistique à partir de ces programmes - et d'appliquer peut-être un facteur de 2 ou 3 pour gonfler les erreurs types produites - puisque les erreurs types dans ces programmes sont généralement de caractère aléatoire simple.

Pour de plus amples renseignements sur les erreurs d'échantillonnage et les erreurs non dues à l'échantillonnage dans l'enquête, voir la publication n° 13-207 au

CENSUS FAMILIES

FAMILLES DE RECENSEMENT

catalogue.

CENSUS FAMILIES

FAMILLES DE RECENSEMENT

COMPARISON OF SCF ESTIMATES TO NATIONAL ACCOUNTS ADJUSTED PERSONAL INCOME

COMPARAISON DES ESTIMATIONS DE L'EFC AUX ESTIMATIONS AJUSTÉES DES COMPTES NATIONAUX

Census Families - 1994

Familles de recensement - 1994

This detailed reconciliation is given for the convenience of researchers and must not be released or published without prior approval from the Director, Household Surveys Division, Statistics Canada. For a general description of differences in concepts and coverage, see Catalogue No. 13-207.

Ce rapprochement détaillé de données est mis à la disposition des chercheurs et ne peut être diffusé ou publié sans l'autorisation du directeur de la Division des enquêtes-ménages, Statistique Canada. Une description générale des différences sur le plan des concepts et du champ d'observation figure dans la publication n° 13-207 au catalogue.

SCF/NATIONAL ACCOUNTS COMPARISON

COMPARAISON ENTRE

EFC/COMPTES

NATIONAUX

Item	1994 Income	SCF NA Adjusted EFC CN ajustés	SCF ^a EFC ^a	NA Adjusted CN Ajusté
Poste	Revenu de 1994			
		%	Estimates in millions of dollars Estimations en millions de dollars	
1.	Wages and salaries • Salaires et traitements [§]	107.4	375,328	349,364
2.	Farm Income • Revenu net d'un emploi autonome agricole	132.0	3,596	2,724
3.	Non-Farm Income from Self-Employment including Net Income from Roomers and Boarders • Revenu net d'un emploi autonome non agricole et revenu net provenant de chambreurs, pensionnaires			
		98.4	26,350	26,785
4.	Earnings (Items 1+2+3) • Gains (Postes 1+2+3)	107.0	405,274	378,873
5.	Interest, Dividends and Miscellaneous Investment Income Intérêt, dividendes et autre revenu de placement	53.1	18,736	35,288
6.	Government transfers (Sum of 7 to 12) • Transferts gouvernementaux (le total de 7 à 12)	87.9	74,619	84,933
7.	Child Tax Benefit* • Crédit d'impôt pour enfants*	94.2	5,517	5,860
8.	GST Credit* • Crédit pour la TPS*	95.3	2,694	2,826
9.	Old Age Security and Guaranteed Income Supplement • Sécurité de la vieillesse et supplément de revenu garanti	102.2	19,263	18,840
10.	Canada/Quebec Pension Plan Benefits • Prestations du régime de pensions du Canada/Régime des rentes du Québec	91.8	17,539	19,098
11.	Unemployment Insurance • Prestations d'assurance-chômage	87.1	13,014	14,944
12.	Other Government Transfers ^b • Autres revenus provenant du gouvernement ^b	71.0	16,592	23,365
13.	Total Income (Items 1 + 2 + 3 + 5 + 6)			
	Revenu total (Postes 1 + 2 + 3 + 5 + 6)	99.9	498,629	499,094

a. SCF amounts include receipts for families whose members were all full-time members of the armed forces. These family units are not included in published tables.
Les chiffres de l'EFC comprennent la somme des montants déclarés pour les familles où tous les membres étaient membres à temps plein des forces armées. Ces familles sont exclus des tableaux publiés.

b. SCF = "Social Assistance and Provincial Income Supplements" plus "Other Income from Government Sources" plus "Provincial Tax Credits".
NA = Item 6 - (Items 7 + 8 + 9 + 10 + 11).
EFC = "L'aide sociale et les suppléments de revenu provinciaux" plus les autres revenus de sources gouvernementales plus les crédits d'impôt provinciaux.
CN = Poste 6 - (Postes 7 + 8 + 9 + 10 + 11).

* The Child Tax Benefit and the Goods and Services Tax Credit figures were obtained from Revenue Canada rather than National Accounts.
Les chiffres du crédit d'impôt pour enfants et le taxe de produits et services sont obtenus de Revenu Canada plutôt que des Comptes Nationaux.

PROCEDURES FOR MANUSCRIPT REVIEW

PROCÉDURES POUR FAIRE RÉVISER UN MANUSCRIT

CENSUS FAMILIES

Because of the complexities inherent in data from Household surveys, users of these data are advised to consult with the Household Surveys Division prior to publication of the data. A manuscript review service is provided by that Division to assist users. Users of microdata tapes are also advised prior to publishing their research to consult the relevant pages of the tape documentation provided regarding data reliability and the use of small samples.

The subscriber should file with Statistics Canada a copy of published research or other results based on the use of these microdata. The subscriber shall use the following form of accreditation in reference to the microdata.

"This analysis is based on the Statistics Canada microdata tape (Census Families - 1994 Income) which contains data collected in the 1995 Survey of Consumer Finances. All computations on these microdata were done by (Name of user organization) and the responsibility for the use and interpretation of these data is entirely that of the author(s)".

Manuscripts can be forwarded to:

Director
Household Surveys Division
Statistics Canada
Jean Talon Building, 5th floor
Tunney's Pasture
Ottawa, Ontario
K1A 0T6

The user should distinguish clearly in the manuscript between data derived directly from the tape on the one hand and all added assumptions, adjustments and modifications introduced by him/her on the other. For example, true wage rates cannot be calculated from the data on the tape and the researcher must specify his assumptions (for which the survey does not provide direct evidence) that would make the averages calculated from the tape appropriate proxies for such a variable.

FAMILLES DE RECENSEMENT

En raison des complexités inhérentes aux données provenant des enquêtes-ménages, il est recommandé de consulter la Division des enquêtes-ménages avant la publication de ces données. Un service de révision de manuscrits est offert par la Division pour assister les utilisateurs. Ceux qui utilisent les bandes de microdonnées pour fins de publication, devraient consulter les pages de la documentation qui traitent de la fiabilité des données et de l'usage de petits échantillons.

L'abonné devrait remettre à Statistique Canada un exemplaire de toute publication portant sur les recherches ou d'autres résultats basés sur l'utilisation des bandes de microdonnées. Dans le texte du manuscrit, il faudra indiquer la source des données de la façon suivante:

"L'analyse est basée sur la bande de microdonnées de Statistique Canada (Familles de recensement - Revenu de 1994) qui renferme des données recueillies dans le cadre de l'enquête sur les finances des consommateurs de 1995. Tous les calculs ont été effectués par (nom de l'organisme), et la responsabilité pour l'usage et l'interprétation de ces données est entièrement celle de ou des auteur(s)".

Les manuscrits peuvent être envoyés au:

Directeur
Division des enquêtes-ménages
Statistique Canada
Immeuble Jean-Talon , 5e étage
Parc Tunney
Ottawa, Ontario
K1A 0T6

Dans son manuscrit, l'utilisateur devra faire clairement la distinction entre les données tirées directement de la bande et les hypothèses, rajustements et modifications qu'il a apportés. Par exemple, les données tirées de la bande ne permettent pas de calculer les taux réels de salaire; le chercheur doit préciser les hypothèses (pour lesquelles l'enquête ne renferme pas de preuves directes) qui feront des moyennes calculées à partir des données de la bande des valeurs de substitution appropriées pour cette variable.

*CENSUS FAMILIES**FAMILLES DE RECENSEMENT***ESTIMATES OBTAINED FROM
MICRODATA FILE**

The following tables contain counts and aggregates obtained from this tape and are given for your use in checking computer output from the tape. On files prior to the 1984 income file, census families with exceptionally high incomes, large income losses from self-employment or unusual characteristics were excluded from the file. As of 1984, records with these conditions have had selected characteristics zeroed out rather than removing them from the file. Counts on the following pages include these records.

We suggest that users not release or analyse data based on a sample of less than 100 records. In fact, due to our procedure of income imputation we feel that a slightly larger minimal sample would be more appropriate. Users employing small samples may wish to consult the **Income and Housing Surveys Section, Household Surveys Division, Statistics Canada, Telephone (613) 951-4633**.

**ESTIMATIONS PROVENANT DU FICHIER
DE MICRODONNÉES**

Les comptes et les agrégats de revenu provenant de cette bande sont présentés selon diverses caractéristiques dans les tableaux qui suivent et vous sont fournis pour vous permettre de vérifier les sorties d'ordinateur produites à partir de cette bande. Les fichiers de microdonnées antérieurs au fichier sur le revenu de 1984 n'englobaient pas les familles de recensement dont le revenu était exceptionnellement élevé, celles qui ont déclaré des pertes considérables découlant d'un travail autonome, celles présentant des caractéristiques inhabituelles. À partir de 1984 on a annulé certaines caractéristiques des enregistrements portant sur ces familles au lieu de supprimer ces enregistrements du fichier. Les montants de contrôle aux pages suivantes comprennent ces enregistrements.

Nous recommandons aux utilisateurs de ne pas publier ni analyser des données fondées sur un échantillon de moins de 100 enregistrements. De fait, en raison de la méthode d'imputation utilisée pour le revenu, nous croyons que le nombre minimum d'enregistrements devrait être plus élevé. Les utilisateurs qui se servent de petits échantillons désireront peut-être consulter la **Section des enquêtes sur le revenu et le logement, Division des enquêtes-ménages, Statistiques Canada, Téléphone (613) 951-4634**.