



BD/MSPS 

Guide d'utilisation

Le présent guide traite en profondeur des fonctions offertes par la BD/MSPS en mode “boîte noire”. Il contient aussi une description générale des variables et des paramètres. On trouvera des renseignements élémentaires sur ces rubriques dans le document *Introduction et survol*.



Statistics
Canada

Statistique
Canada

Canada

Table des matières

Introduction.....	2
Paramètres.....	2
Fichiers de paramètres	2
Fichiers des paramètres de commande	3
Fichiers des paramètres d'ajustement de la base de données	3
Fichiers des paramètres fiscaux/de transfert	3
Types de paramètres	3
Chaînes.....	4
Scalars	4
Vectoriels	5
Tables de recherche.....	5
Tableaux.....	6
Édition des paramètres.....	7
Variables	7
Variable de classe et variable d'analyse	7
Niveau de la famille	8
Variables de base de données et variables modélisées	9
Variables de variante et variables de base	9
Règles d'affectation des noms des variables	9
Expressions	10
Constantes	10
Variables	11
Opérateurs	11
Spécificateurs de niveau de famille	11
L'opérateur de différence variante-base	12
Opérateurs mathématiques.....	12
Opérateurs logiques et opérateurs de comparaison.....	12
Fonctions.....	12
Instructions.....	13
Instructions d'affectation	13
Instructions de définition	13
Instructions de commande de flux	15
Paramètres de description	15
Fichiers de la base de données	16
Fichiers de ménage/d'individu.....	16
Fichiers des dépenses.....	16
Fichiers des poids.....	16
Contrôle de la pondération.....	16
Fonction Valeurs de référence	16
Création de variables utilisateur d'une exécution antérieure.....	17
Modification des variables de base de données	18
Ajustement de la base de données	19
Contrôle de l'imputation	19
Variante et base.....	20
Contrôle des variantes.....	20

Contrôle de la base.....	21
Fonction Sous-échantillonnage.....	22
Fonction Sélection	23
Fonction Taux marginal d'imposition	25
Fonction Variable définie par l'utilisateur.....	27
Fonction Variable utilisateur.....	27
Fonction Sortie texte	29
Fonction Sortie SAS	31
Fonction Totalisation intégrée	36
Fonction Tableaux croisés	39
Spécification des totalisations définies par l'utilisateur.....	39
Éléments d'une demande de totalisation.....	39
Le niveau de totalisation d'une demande de totalisation	40
Éléments propres à la totalisation	40
Les expressions impliquent des fonctions de variables agrégées	40
Imposition de qualificatifs aux éléments objet d'une totalisation	41
Syntaxe du qualificatif	41
Niveaux de classification	42
Création d'une catégorie "Toutes" avec le suffixe "+"	42
Ordre des niveaux et aspect de la table	43
Spécification d'une table sur plusieurs lignes.....	43
Paramètres de commande connexes.....	43
Questions de niveaux d'analyse.....	44
Les totalisations sont fonction des paramètres de sélection.....	45
Fonction Analyse de la distribution	45
Mesures de l'inégalité et polarisation de la répartition du revenu	48
Introduction aux mesures de l'inégalité	48
Mesures des agrégats	50
Mesure de la polarisation de la répartition du revenu.....	54
L'activation du calcul des mesures de l'inégalité et polarisation	56
Fonction Points de retournement	57
Introduction.....	57
Fonctionnement de la fonction Points de retournement	59
Exemple	62
Fonctions diverses.....	65
Fonction Nombres aléatoires	65
Analyse du seuil de faible revenu	65
Ajustement à l'aide sociale et à l'assurance-emploi	66
Aide sociale.....	66
Assurance-emploi	67

Introduction

Le présent document décrit en détail la façon d'utiliser le MSPS en "boîte noire". On suppose que le lecteur a lu le document *Introduction et survol*, et qu'il connaît à fond les concepts qui y sont exposés. L'utilisation du MSPS en mode "boîte de verre" est exposée dans le document *Guide de programmation*.

La première section du présent guide contient ce qui suit :

- Une description de l'utilisation des divers types de paramètres employés pendant l'exploitation du MSPS.
- Une description des variables qui peuvent contenir de l'information de la BDSPS et des résultats de l'application du modèle fiscal/de transfert pour les individus des familles.
- Une brève description des types de fichiers de bases de données fournis par Statistique Canada comme entrée dans le MSPS.
- Une description du processus d'ajustement de la base des données utilisé par le modèle pour permettre à la BDSPS de représenter différentes années civiles.
- Une explication des dispositions spécifiques qui permettent la comparaison de deux scénarios par le MSPS.

Les sections suivantes exposent, à tour de rôle, chacune des grandes fonctions du MSPS. L'ordre de présentation de ces fonctions dans le guide suit l'ordre dans lequel les paramètres qui les commandent figurent dans les fichiers des paramètres de commande du MSPS.

Les utilisateurs qui sont particulièrement intéressés par la fonction Tableaux croisés du MSPS trouveront, à la section Fonction Tableaux croisés, une description complète, mais très concise de cette fonction. Ils trouveront une description plus complète, avec des exemples pratiques, dans la publication *Guide d'utilisation des tableaux croisés*.

Paramètres

C'est par les paramètres que l'utilisateur commande tous les aspects du MSPS. Un paramètre est une valeur qui est spécifiée avant l'exécution réelle du modèle et qui détermine la façon dont le modèle s'exécute. Il est spécifié dans l'un des trois types de fichiers de paramètres d'entrée lus par le MSPS. Le paramètre est représenté dans ces fichiers sous forme d'un symbole en majuscules suivi d'une valeur. Les trois types de fichiers de paramètres sont étudiés à la section Fichiers de paramètres du guide.

Les paramètres peuvent contenir divers genres d'information, depuis une simple valeur numérique jusqu'à une chaîne compliquée s'étendant sur plusieurs lignes, qui servira à préciser une demande de tableaux croisés. Les diverses formes de paramètres sont étudiées à la section Types de paramètres. Les paramètres sont modifiés directement dans MSPS Visuel pour changer les valeurs dans le fichier de paramètres approprié. Dans la version MSPS Classique, il est possible de modifier les valeurs des paramètres de façon interactive en utilisant la fonction d'édition des paramètres décrite à la section Fonction Édition des paramètres. La source finale de chaque paramètre du MSPS est le document *Guide des paramètres*. Ce guide est donné en direct sous deux formes : en ordre alphabétique ou en ordre des domaines. Pour accéder à la description alphabétique des paramètres, cliquer deux fois sur "Listes alphabétiques", puis sur "Paramètres". Pour accéder à la description des paramètres classés en ordre des domaines, cliquez deux fois sur "Manuel de référence", puis sur "Guide des paramètres". Un utilitaire commode (compparm) qui signale les différences entre les fichiers des paramètres est décrit dans le document *Guide d'utilisation des outils*.

FICHIERS DE PARAMÈTRES

Comme nous l'avons mentionné, le MSPS comporte plus de 800 paramètres répartis entre trois groupes correspondant à autant de types de fichiers qui commandent trois opérations principales. Les **paramètres de commande** spécifient les fichiers d'entrée et de sortie d'une exécution particulière du modèle et servent à en activer les fonctions logicielles. Les **paramètres d'ajustement de la base de données** s'appliquent à la majoration des éléments monétaires de la base de données. Enfin, les **paramètres fiscaux/de transfert** régissent le fonctionnement spécifique des programmes fiscaux/de transfert. Pour les taxes sur les produits, il existe un jeu de paramètres implicites ou par défaut.

Les paramètres d'un fichier de paramètres sont ordonnés et mis en forme d'une façon bien précise. Le fichier de paramètres qui ne respecte pas les conventions de mise en forme sera remis en forme au besoin, si un nouveau fichier des paramètres est produit par le MSPS.

Fichiers des paramètres de commande

Chaque exécution du MSPS utilise un seul fichier des paramètres de commande. Ce fichier, qui a l'extension `.cpr`, donne les noms des fichiers d'entrée et de sortie qui doivent être utilisés à l'exécution et fournit des valeurs qui commandent toutes les fonctions du MSPS décrites dans les sections suivantes du guide.

Certaines entrées du fichier des paramètres de commande ne sont pas utilisées pour communiquer de l'information au MSPS. Ces paramètres de commande "d'information" servent plutôt à fournir à l'utilisateur certains renseignements sur l'exécution. Un exemple est le paramètre `SAMPLE` qui indique la fraction de la BDSPS qui a été lue pendant l'exécution.

Fichiers des paramètres d'ajustement de la base de données

Chaque exécution du MSPS utilise un seul fichier des paramètres d'ajustement de la base de données dont l'extension est `.apr`. La première fonction de ce fichier consiste à spécifier les paramètres qui "font vieillir" les données jusqu'à une année ultérieure quelconque en fournissant à la BDSPS des facteurs de croissance propres à des variables. Le fichier `.apr` contient aussi des paramètres qui précisent des mécanismes structuraux de l'année en question, comme les taux de chômage propres aux régions.

Fichiers des paramètres fiscaux/de transfert

Les fichiers de paramètres fiscaux/de transfert précisent tout le fonctionnement des algorithmes fiscaux/de transfert. Une exécution particulière du MSPS peut exiger des fichiers de paramètres fiscaux/de transfert 0, 1 ou 2. Le nombre de fichiers de paramètres fiscaux/de transfert dépend de la valeur donnée aux paramètres de commande `BASMETH` et `VARMETH`.

TYPES DE PARAMÈTRES

Les paramètres fournissent divers genres d'information au MSPS. Il y a par conséquent divers genres de paramètres, et chacun convient à un type particulier d'information. Les paramètres de chaîne fournissent de l'information textuelle au MSPS. L'information textuelle peut être aussi simple qu'une étiquette décrivant un fichier des paramètres, ou être aussi compliquée qu'une demande de tableaux croisés formulée par l'utilisateur. Les paramètres scalaires fournissent des nombres uniques, comme le montant de l'allocation familiale par enfant. Le paramètre vectoriel fournit un ensemble de nombres unidimensionnel. Un exemple est l'ensemble de seuils de

revenus utilisé pour déterminer les colonnes du tableau 2 intégré. Les paramètres de tables de recherche fournissent de l'information sur des fonctions linéaires ponctuelles utilisées dans le système fiscal/de transfert. Un exemple est la table d'impôt fédéral. Le paramètre de tableau fournit un ensemble de nombres bidimensionnel. On en a un exemple dans les taux de taxe de vente actuels des provinces, par catégorie de dépenses et par province. Les sections suivantes étudient chacun de ces types de paramètres.

Chaînes

Les paramètres de chaîne sont utilisés pour fournir de l'information textuelle (et non numérique) au MSPS. Chaque paramètre de chaîne a une longueur maximale permise. Si la longueur maximale est dépassée, un message d'erreur pertinent est généré et la valeur du paramètre est tronquée.

L'exemple suivant est extrait du fichier de paramètres fourni \spsd\ba02.mpr.

```
MPRDESC          Current values for 2002
                  # Source: Given as LABEL=...
                  # Update: Value=LABEL
```

Dans cet exemple, le paramètre MPRDESC attribue un titre général au jeu de paramètres fiscaux/de transfert.

La source et la ligne de mise à jour automatiquement générés par le MSPS pour rendre le fichier de paramètre plus compréhensible. L'utilisateur ne devrait pas tenter d'insérer ses propres commentaires sur la même ligne que la valeur de paramètre, parce que le MSPS élimine ces commentaires et les remplace par des étiquettes de paramètres lorsque le fichier de paramètres est écrit. Veuillez consulter la section Commentaires des paramètres pour obtenir plus d'information sur les commentaires définis par l'utilisateur dans les fichiers de paramètres.

La paramètre de chaîne peut être réparti sur plusieurs lignes s'il est trop long, ou si l'utilisateur désire une présentation plus claire. L'exemple suivant illustre ce genre de présentation.

```
XTSPEC                                # Spécifications des tableaux croisés
IN: {ex5,
     ex7:S=3,
     ex8:S=3}
    * c10+;
IN: c11+ * {ex5, ex5/ex7:L="Average Benefits", ex7:S=3, ex8:S=3};
IN: c12+ * {ex5, ex7:S=3, ex8:S=3, scfrecs };
```

Dans cet exemple, le paramètre XTSPEC, qui est utilisé pour générer des tableaux croisés définis par l'utilisateur, a été établi en une chaîne de sept lignes. Les quatre premières lignes appartiennent à la même demande de tableau. L'utilisateur a aligné les variables pour que le fichier soit facile à lire et que les lignes soient faciles à corriger par la suite. Les trois dernières lignes affichent d'autres façons acceptables d'écrire des demandes de tableaux. La dernière demande peut terminer ou non par un point-virgule; cependant, il est suggéré de toujours utiliser le point-virgule dans la dernière demande de tableaux.

Scalaire

Les paramètres scalaires sont utilisés pour fournir des valeurs numériques uniques au MSPS. Elles sont données sur la même ligne que la valeur du paramètre, comme l'indique l'exemple suivant.

```
STDFA          418.56          # Allocations familiales fédérales normales par
enfant
```

Dans cet exemple, extrait du fichier de paramètres fiscaux/de transfert \spsd\ba92.mpr, le montant de l'allocation familiale par enfant a été fixé au montant indiqué. Les valeurs scalaires, comme tous les paramètres numériques, sont remises en forme lorsque le fichier des paramètres est écrit.

Vectoriels

Les paramètres vectoriels sont utilisés pour fournir un ensemble ordonné de valeurs numériques. Les valeurs réelles du paramètre sont indiquées une par ligne. Les utilisateurs peuvent ajouter de nouvelles dimensions au paramètre à la dernière ligne (à côté de l'astérisque). L'exemple suivant illustre ces points.

```
DISTP          13              # Points d'arrêt pour le tracé des
histogrammes
    1
    5
   10
   20
   30
   40
   50
   60
   70
   80
   90
   95
   99
   *
```

On attribue au paramètre vectoriel DISTP un ensemble de 13 valeurs se situant entre 1 et 99. Ce paramètre précise les points d'arrêt en centile qui définissent l'axe horizontal du tracé d'histogramme produit par la fonction d'analyse de la distribution (voir la section Fonction Analyse de la distribution).

Tables de recherche

Les paramètres de table de recherche sont utilisés pour fournir de l'information définissant les fonctions continues par morceaux. Une fonction de ce genre est indiquée dans le MSPS par un ensemble de 3-uplets ordonnés. Chaque 3-uplet correspond à un "point d'arrêt" ou à un changement de sens de la fonction. Le nombre de 3-uplets est indiqué sur la même ligne que le nom du paramètre, tandis que les 3-uplets suivent, un sur chaque ligne qui suit.

- Le premier nombre de chaque 3-uplet donne la valeur X de la fonction.
- Le deuxième nombre donne la valeur Y correspondante.
- Le troisième donne la pente du segment de la courbe qui commence en (X, Y).

Il y a une certaine redondance dans cette information. Les valeurs Y (sauf la première) peuvent être dérivées de la pente. Par ailleurs, les pentes peuvent être dérivées des valeurs X et Y. Ces

deux formes de redondances correspondent à deux types de paramètres de recherche, appelés paramètres de table de recherche “Point et pente” et “X-Y”. Les deux types sont utilisés dans le MSPS parce que les tables de recherche particulières sont plus naturellement représentées dans une forme ou dans l’autre.

Prenons l’exemple suivant d’une table de recherche “Point et pente” :

FTX	4		# Table d’impôt fédéral
0	0	0.160	
31677	(5068)	0.220	
63354	(12037)	0.260	
103000	(22345)	0.290	

Source: Federal Income Tax 2002 - Schedule 1

Comme le commentaire l’indique, il s’agit de la table d’impôt fédéral de 2002. La table comporte quatre taux d’imposition distincts et on spécifie la fonction en fournissant les valeurs de X et le taux d’imposition correspondant (c.-à-d. la pente) pour chaque taux. La valeur Y de départ (0 dans l’exemple) doit aussi être indiquée. Les autres valeurs de Y sont calculées à partir des valeurs et des pentes et elles sont indiquées entre parenthèses pour souligner le fait qu’elles sont calculées. Ces valeurs Y dérivée sont calculées par le MSPS et elles sont produites lorsque le fichier de paramètres est écrit. L’utilisateur doit fournir une valeur pour les trois colonnes d’un paramètre de table de recherche Point et pente, mais les valeurs de Y n’ont pas besoin d’être précises (sauf pour la première), puisqu’elles sont calculées correctement par le MSPS à la lecture du paramètre.

Voici un exemple de table de recherche “X-Y”.

SPAT	3		# Taux de participation à l’allocation au conjoint selon la tranche de revenu
0	0.550	(0.0001)	
3316	0.720	(0.0003)	
4270	1.000	(0.0003)	

Le paramètre SPAT donne la proportion des personnes admissibles qui participent au programme, comme fonction du niveau d’allocation. Ainsi, 72 % des personnes admissibles au programme AC participeront au programme si leur allocation est de 3 316 \$. La troisième colonne donne la pente de la fonction continue par morceaux correspondante. Les valeurs de pente sont trop petites pour être représentées avec précision dans cet exemple. Quoi qu’il en soit, les valeurs de pente sont calculées avec précision par le MSPS et sont données dans le fichier de paramètres à des fins d’information seulement.

Tableaux

Les tableaux sont des ensembles bidimensionnels de nombres. Un exemple est le paramètre PTF qui est reproduite ci-dessous.

PTF	9			# Seuil de pauvreté pour le tableau 4
16186	13883	13787	12829	11186
20233	17354	17234	16036	13982
25163	21583	21433	19943	17390
30460	26126	25945	24142	21050
34049	29205	29002	26986	23531
37638	32284	32059	29830	26012

41227	35363	35116	32674	28493
41227	35363	35116	32674	28493
41227	35363	35116	32674	28493

Les indices utilisés pour déterminer une valeur dans la table commence toujours à zéro. Le paramètre PTF est un tableau bidimensionnel donnant un seuil de revenu fourni par l'utilisateur pour les familles : les lignes représentent le nombre de personnes moins un - ceci fixe l'indice à zéro pour une famille d'une personne - et la colonne, la région urbaine (hdurb).

ÉDITION DES PARAMÈTRES

Le MSPS Classique contient une fonction d'édition des paramètres qui offre une méthode pour l'utilisateur à utiliser afin de modifier les valeurs des paramètres dans un fichier de paramètres d'entrée sans utiliser un éditeur de texte. Dans MSPS Visuel, les paramètres peuvent être ouverts et mis en forme interactivement.

Si le paramètre qui est édité est un vecteur, une table de recherche ou un tableau, le dialogue est légèrement différent. Si la dimension peut-être modifiée, il y a une invite qui permet cette modification. Ensuite, si le paramètre est un tableau, le système demande la colonne à modifier. Enfin, une invite est produite pour tous les éléments du paramètre, à tour de rôle. Toutes les touches d'édition indiquées ci-dessus peuvent être utilisées.

Variables

Cette section contient de l'information de référence sur les variables utilisées dans la BD/MSPS. Une variable contient de l'information sur un ménage, une personne ou une famille en particulier dans la BD/MSPS. Ceci est distinct d'un paramètre qui contient habituellement des données utilisées pour spécifier le système fiscal/de transfert utilisé dans une simulation. L'information définitive sur chacune des variables de la BD/MSPS se trouve dans le document *Guide des variables* dont l'accès est aussi possible dans l'ordre alphabétique en cliquant deux fois sur "Listes alphabétiques", puis sur "Variables".

Il est possible d'utiliser la BD/MSPS sans référer à des variables. Ainsi, les paramètres d'une simulation fiscale/de transfert pourraient être entrés dans le MSPS et il est possible d'utiliser les rapports intégrés habituels pour l'analyse des résultats. Cependant, de nombreuses fonctions plus puissantes de la BD/MSPS exigent que l'on utilise des variables. Ces fonctions, qui sont décrites plus en détails ailleurs dans le présent manuel, comprennent la sélection des enregistrements, les tableaux croisés, les sorties SAS, les rapports de cas et l'analyse de la distribution. Il y a aussi une fonction qui permet à l'utilisateur de créer de nouvelles variables selon les valeurs des variables existantes.

Les variables ont plusieurs caractéristiques qu'il faut comprendre si l'on veut comprendre le contenu des sections suivantes. Ces caractéristiques sont décrites dans les sections suivantes.

VARIABLE DE CLASSE ET VARIABLE D'ANALYSE

Les variables de la BD/MSPS peuvent être réparties en deux grandes catégories. Les variables dont la valeur indique l'appartenance à des catégories distinctes porte le nom de variables « de classe » dans le présent document. On dit aussi que ce sont des variables de classification. Un

exemple de variable de classe est idsex, qui enregistre le sexe d'une personne. Les variables dont la valeur contient des valeurs numériques qui peuvent servir dans des opérations mathématiques portent le nom de variables « d'analyse ». Un exemple de variable d'analyse est idiemp, qui enregistre le revenu d'emploi d'un individu.

La distinction qu'il y a entre les variables de classe et les variables d'analyse est importante. Certaines des fonctions du MSPS exigent que l'on spécifie, dans certains contextes, une variable de classe ou une variable d'analyse. Les variables de classe peuvent être converties en variables d'analyse, ou les variables d'analyse en variables de classe, les fonctions de variables étant spécifiées par l'utilisateur.

Une variable, l'âge, doit souvent être utilisée soit comme variable de classe, soit comme variable d'analyse. Par conséquent, deux variables ont été définies pour l'âge : idage est la version de variable « de classe », et idnage, est pour la version de variable « d'analyse ». idage peut être utilisé pour générer un rapport sur la distribution par âge d'un certain groupe de personnes, tandis que idnage pourrait être utilisé pour générer un rapport sur l'âge moyen des personnes de divers groupes.

NIVEAU DE LA FAMILLE

La BDSPPS a une structure hiérarchique en ce sens que les individus conservent leur contexte de famille. Cela signifie que la BDSPPS peut être considérée comme un fichier d'individus, un fichier de familles (définies de façons diverses) ou un fichier de ménages. Chaque variable de la BD/MSPS est défini à un niveau naturel dans cette hiérarchie de famille. Ainsi, la variable hdprov (province de résidence) est définie au niveau du ménage, tandis que idiemp (revenu d'emploi) est défini au niveau de l'individu. Les diverses fonctions d'établissement de rapports du MSPS permettent à l'utilisateur de spécifier le niveau de la famille de l'analyse désirée pour une fonction de sortie en particulier. Cependant, l'utilisateur n'est pas limité à l'utilisation des variables dont le niveau naturel est le même que celui qui est spécifié. Un certain nombre de règles servent à interpréter la signification de ces demandes "entre familles".

Si le niveau naturel de la variable est "plus élevé" que le niveau précisé par l'utilisateur, la valeur de la variable est la valeur qui se trouve à ce niveau plus élevé. Ainsi, une référence hdprov quand on spécifie le niveau d'analyse "individu" renvoie simplement à la province associée au ménage dont l'individu fait partie.

Si le niveau naturel de la variable est "plus bas" que le niveau spécifié par l'utilisateur, il se produit deux cas, selon que la variable est une variable d'analyse ou une variable de classe. S'il s'agit d'une variable d'analyse, alors la valeur de la variable est la somme des valeurs du niveau "le plus bas" contenues dans l'unité de la famille au niveau spécifié. Par exemple, une référence à idiemp lorsque l'on spécifie le niveau d'analyse comme "ménage" fait référence à la somme des revenus d'emploi de toutes les personnes d'un ménage. S'il s'agit d'une variable de classe, alors la valeur de la variable est la valeur associée à la première unité contenue (appelée l'individu ou la famille de référence) au niveau naturel de la variable. Ainsi, une référence à idage lorsque l'on spécifie le niveau d'analyse "ménage" fait référence à l'âge de la première personne qu'il y a dans le ménage.

L'utilisation des variables de classe de cette façon exige de toute évidence une connaissance de l'ordre dans lequel les individus et les familles sont disposés. Les individus sont disposés à l'intérieur des familles en commençant par le conjoint le plus âgé, se poursuit avec l'autre conjoint (s'il est présent), le tout étant suivi des enfants dans l'ordre croissant d'âge. Les familles sont arrangées à l'intérieur des ménages avec la famille "principale" en premier. Les pensionnaires, par exemple, constituent des familles "secondaires" contenant seulement un individu seul.

Le tout se complique encore lorsque la sélection a été activée avec les paramètres de commande SELFLAG, SELUNIT et SELSPEC. La sélection se fait, en bout de ligne, au niveau de l'individu, et, par conséquent, les remarques suivantes s'appliquent aux variables dont le niveau naturel est "individu". Si une telle variable fait l'objet d'une référence quand on spécifie un niveau d'analyse plus élevé, seuls les individus sélectionnés font l'objet du traitement. Ainsi, si la fonction de sélection a été configurée de façon à sélectionner seulement les personnes dont l'âge se situe entre 18 ans et 55 ans, alors, une référence à `idiemp` lorsque l'on spécifie le niveau d'analyse "ménage" réfère à la somme des revenus d'emploi de toutes les personnes du ménage dont l'âge se situe entre 18 et 55 ans. De même, une référence à `idage` lorsque l'on spécifie le niveau d'analyse "ménage" réfère à l'âge de la première personne du ménage dont l'âge se situe entre 18 et 55 ans.

VARIABLES DE BASE DE DONNÉES ET VARIABLES MODÉLISÉES

Les variables de la BD/MSPS peuvent être réparties en deux groupes selon qu'elles sont lues (ou dérivées directement) de la base de données ou qu'elles résultent du processus de modélisation. Les deux premières lettres du nom de la variable indiquent si la variable est une variable modélisée ou non. Comme l'indique la section Règles d'affectation des noms de variable, ci-dessous, les préfixes `im`, `ub` et `ct` indiquent des variables modélisées. Tous les autres préfixes indiquent des variables de base de données.

VARIABLES DE VARIANTE ET VARIABLES DE BASE

Le MSPS permet d'accéder à deux ensembles distincts de variables modélisées, les valeurs de base et les valeurs de variantes. Si l'utilisateur exécute une simulation unique, les valeurs de variante de toutes les variables sont définies. Si l'utilisateur exécute deux simulations simultanément, les valeurs de toutes les variables de base sont définies. Si l'utilisateur lit les valeurs de base d'un fichier de résultats produit antérieurement, les valeurs de base des variables sauvegardées seront définies et toutes les autres variables de base recevront la valeur zéro.

Pour référer à une variable de base, il suffit de précéder le nom de la variable correspondante du caractère de soulignement. Ainsi, `_immicons` réfère à la variante Revenu consommable, tandis que `immicons` réfère à la valeur de base du Revenu consommable.

RÈGLES D'AFFECTATION DES NOMS DES VARIABLES

Contrairement à ce qu'il en est pour les paramètres, qui sont toujours en majuscules, les variables de la BD/MSPS sont toujours en minuscules. À quelques exceptions près, les variables sont soumises aux règles d'affectation de noms qui exigent que les deux premières lettres du nom (le préfixe) indiquent le niveau de famille de base de la variable, qu'elle soit lue de la base de données ou produite par le modèle. Le reste des lettres du nom (la souche) décrit la variable elle-

même. Les variables de demande de prestations d'assurance-chômage ont un chiffre inséré entre le préfixe et la souche, qui indique la demande à laquelle la variable fait référence. Voici un tableau des préfixes valides et de leur signification.

hh	Données sur la structure des ménages
hd	Caractéristiques des ménages
ef	Caractéristiques des familles économiques
cf	Caractéristiques des familles de recensement
nf	Caractéristiques des familles nucléaires
id	Données sur l'individu, de la base de données
im	Données sur l'individu, du modèle
uc	Données sur les demandes de prestation d'assurance-chômage, de la base de données
ub	Données sur les demandes de prestation d'assurance-chômage, du modèle
fx	Données sur le régime de dépenses, de la base de données
ct	Données sur les taxes à la consommation, du modèle

EXPRESSIONS

Les capacités des expressions du MSPS, qui sont utilisées dans un certain nombre de paramètres de commande, ont été considérablement étendues. Plutôt que de décrire les modifications qui ont été apportées, les paragraphes suivants décrivent en entier la syntaxe des expressions du MSPS.

Les expressions sont utilisées dans un certain nombre de fonctions du MSPS. Plus précisément, on trouve les expressions dans les paramètres SELSPEC, MARSPEC, TPSPEC, UVAR et XTSPEC. Une expression du MSPS est une suite de constantes, de variables, d'opérateurs et de fonctions dont la syntaxe est valide et qui évaluent en vue de donner une quantité à virgule flottante. La syntaxe est semblable à celle qui est utilisée en langage de programmation C. Comme en C, les commentaires (délimités par /* et */) peuvent être insérés quand bon vous semble et l'ordre des opérations peut être modifié par l'utilisation des caractères de parenthèse, (et). En outre, comme en C, une valeur zéro représente la valeur logique "faux", tandis que toute valeur qui n'est pas zéro représente la valeur logique "vrai". Les opérateurs logiques retournent la valeur 1 pour représenter une valeur "vrai".

Constantes

Les constantes sont des nombres décimaux qui peuvent contenir une partie fractionnelle facultative. La notation exponentielle n'est pas reconnue. Un certain nombre de constantes couramment utilisées dans le MSPS peuvent être entrées sous forme de symboles qui sont les synonymes en majuscules donnés dans le tableau suivant :

Mnemonic	Valeur	Signification
NFLD	0	Codes des provinces utilisés dans la variable hdprow
PEI	1	
NS	2	
NB	3	
QUE	4	
ONT	5	
MAN	6	
SASK	7	

ALTA	8	
BC	9	
MALE	0	Codes utilisés dans les variables idsex.
FEMALE	1	
HEAD	0	Codes utilisés dans la variable idcfrh.
SPOUSE	1	
CHILD	2	

Variables

Les variables consistent en un caractère de soulignement (_) facultatif suivi d'une lettre minuscule, puis, si on le désire, d'autres lettres minuscules et chiffres. Les variables doivent être des variables de base de données de la BDSPS, des variables modélisées de MSPS ou des variables utilisateur créées par la Fonction Variable utilisateur. Les variables sont décrites plus en détail dans le document *Guide des variables* ou peuvent aussi être expliquées en direct si l'on clique sur Index puis sur Variables.

Opérateurs

Les opérateurs prennent un ou deux arguments et utilisent la hiérarchie d'exécution suivante :

Opérateurs	Description
IN: SP: NF:	spécificateurs de niveau de famille
CF: EF: HH:	
- ! @	moins unaire, négation logique, différence variante – base
**	Exponentiation
* /	multiplication, division
+ -	addition, soustraction
< <= > >=	comparaison de nombres
LT LE GT GE	
== !=	égalité, inégalité
EQ NE	
&&	conjonction logique
AND	
	disjonction logique
OR	

Spécificateurs de niveau de famille

Les expressions sont évaluées au niveau de l'individu. Les opérateurs du spécificateur de niveau de famille permettent la référence à des variables touchant les membres d'autres familles. Les spécificateurs de niveau de famille précèdent immédiatement une variable ou une expression, faisant la référence à la valeur de la variable ou de l'expression sur l'unité de famille contenant l'individu et en faisant le cumul.

Spécificateur	Signification
IN:	Individu actuel
SP:	Conjoint de l'individu actuel
NF:	Famille nucléaire dont fait partie l'individu actuel
CF:	Famille de recensement dont fait partie l'individu actuel
EF:	Famille économique dont fait partie l'individu actuel
HH:	Ménage dont fait partie l'individu actuel

Voici un certain nombre d'exemples :

SP:idiemp	valeur du revenu d'emploi du conjoint
SP:(idage>70)	est 1 si le conjoint a plus de 70 ans et 0 autrement
CF:(idisa>0)	nombre de personnes dans la famille de recensement recevant l'aide sociale
CF:immicons/HH:immicons	proportion du revenu consommable du ménage associée à la famille de recensement dont fait partie l'individu actuel
CF:(idage>=3 && idage<=5)	nombre de personnes de la famille de recensement dont l'âge se situe entre 3 et 5

Les spécificateurs de niveau de famille touchant des variables utilisateur exigent une attention particulière. L'analyste doit veiller à ce que la variable utilisateur ait été calculée pour toutes les personnes du ménage avant que l'opérateur du spécificateur du niveau de famille soit appliqué soit à elle, soit à une expression qui la contient.

L'opérateur de différence variante-base

L'opérateur de différence variante-base calcule la différence entre la valeur de la variante et la valeur de base d'une variable modélisée. Par exemple, l'expression @immicons est synonyme de immicons - _immicons et donne la variation survenue entre le revenu consommable de base et la variante.

Opérateurs mathématiques

Les opérateurs mathématiques conventionnels +, -, *, / suivent les règles de hiérarchie conventionnelles et n'exigent aucune explication. Le symbole ** signifie l'exponentiation et a un haut niveau dans la hiérarchie.

Opérateurs logiques et opérateurs de comparaison

Les opérateurs logiques fonctionnent de façon habituelle. Le MSPS permet les synonymes pour certains des opérateurs, comme l'indique la table ci-dessus. Comme il est mentionné, toute valeur autre que zéro a une valeur logique "vrai". Tous les opérateurs logiques et tous les opérateurs de comparaison utilisent 1 pour indiquer une valeur "vrai".

Fonctions

Les fonctions sont spécifiées avec la notation attendue et elles figurent dans le tableau suivant.

Fonction	Description
abs(a)	valeur absolue
nneg(a)	valeur non négative : équivalent à max(0,a)
min(a,b)	minimum de deux valeurs
max(a,b)	maximum de deux valeurs
ln(a)	logarithme naturel
exp(a)	Exponentiation
split(a,b1,b2,...)	Détermine le fractionnement dans lequel une valeur se retrouve
trunc(a)	Partie entière d'une valeur

La fonction split retourne 0 si $a \leq b1$, 1 si $b1 < a \leq b2$, etc.

INSTRUCTIONS

La fonction Variable utilisateur utilise les instructions, qui sont similaires aux instructions du langage C. La syntaxe de ces instructions est décrite dans la présente section. Les instructions visent principalement à affecter des valeurs d'expression aux variables utilisateur.

Instructions d'affectation

L'affectation se fait par le symbole =. Le MSPS est délibérément plus restrictif que le C dans la syntaxe de cet opérateur, ce qui empêche certains genres d'erreurs qu'il est facile de faire dans ce langage. La syntaxe est la suivante :

```
<variable utilisateur> = <expression> ;
```

La variable qui reçoit l'affectation ne peut être une variable de base de données ou une variable modélisée. Si la variable utilisateur a déjà une valeur (créée soit par une instruction antérieure, soit par la fonction Valeur de référence), la valeur est remplacée. Une forme spéciale permet l'affectation au conjoint de l'individu actuel :

```
SP:<variable utilisateur> = <expression> ;
```

Instructions de définition

Une variable utilisateur a un certain nombre de caractéristiques en plus de la valeur qu'elle contient pour chaque individu. Elle peut avoir une étiquette et être soit une variable de classe, soit une variable d'analyse. S'il s'agit d'une variable du type analyse, elle peut avoir une partie fractionnaire qui peut avoir un niveau de précision spécifié pour l'impression. S'il s'agit d'une variable de classe, elle a un nombre fixe de niveaux permis et chaque niveau peut avoir une étiquette associée utilisée pour l'impression. Par défaut, une variable utilisateur est une variable d'analyse sans partie fractionnaire (sauf si elle est générée par la fonction split), mais ces caractéristiques peuvent être modifiées par les instructions suivantes.

Instruction label

Comme son nom l'indique en anglais, l'instruction label permet d'associer une étiquette textuelle à une variable utilisateur. La syntaxe est la suivante :

```
label(a) = "étiquette textuelle" ;
```

Dans cet exemple, la variable utilisateur a reçoit l'étiquette indiquée. Cette étiquette sera utilisée par les diverses fonctions de sortie du MSPS.

Instructions levels

L'instruction `levels` donne le nombre de niveaux pour une variable utilisateur et fournit une étiquette pour chaque niveau. Elle détermine implicitement que la variable utilisateur sera du type classe.

```
levels(a) = "niveau #0", "niveau #1", "niveau #3" ;
```

L'exemple détermine que `a` est une variable de classe à trois niveaux et ces trois niveaux ont les étiquettes indiquées. Comme c'est le cas dans toutes les variables de classe du MSPS, le premier niveau a la valeur 0, le deuxième a la valeur 1, etc.

Instruction type

L'instruction `type` force la variable utilisateur à être du type classe ou du type analyse. Cela n'est habituellement pas nécessaire parce que les variables utilisateur se voient en général attribuer implicitement un type approprié.

```
type(a) = analysis ;  
type(b) = class ;
```

Les exemples ci-dessus déterminent que la variable utilisateur `a` est une variable d'analyse et que `b` est une variable de classe.

Instruction precision

L'instruction `precision` indique que la variable utilisateur a une partie fractionnaire et indique que le nombre de chiffres à inscrire à droite de la virgule de décimale à l'impression des valeurs contenues dans la variable. La plupart des variables du MSPS, puisqu'elles sont des nombres de dollars, n'ont aucune partie fractionnaire. Ceci fait en sorte qu'elles peuvent être stockées efficacement dans les fichiers de résultats du MSPS. La précision doit être fixée à une valeur autre que zéro si l'on désire conserver les parties fractionnaires dans les fichiers de résultats. Cela fera en sorte que l'on stockera la variable comme une quantité simple précision à virgule flottante, avec un total dont la précision est d'environ 6,5 chiffres.

```
precision(a) = 3 ;
```

La précision de la variable d'analyse a été fixée à 3. Si l'on imprime `a`, il y aura 3 chiffres à droite de la virgule de décimale.

Si l'instruction de précision est utilisée (et est plus grande que 0), la variable est stockée dans des fichiers de résultats sous forme de nombre à double précision.

Affectation de la fonction de fractionnement split

Si une variable se voit affecter le résultat de la fonction `split`, il se produit un certain nombre de déclarations implicites au sujet de la variable. Plus précisément, la variable sera du type classe avec le nombre approprié de niveaux et les étiquettes de niveau par défaut seront générées pour cette variable, si cela est possible.

```
b = split(idiemp, 1000, 2000, 3000) ;
```

L'exemple ci-dessus donne à `b` des attributs qui équivalent aux instructions suivantes :


```
label(b) = "Groupe de revenu d'emploi" ;  
levels(b) = "Min. - 1000", "1001 - 2000", "2001 - 3000", "3001 - max." ;
```

Si le premier argument du fractionnement avait été une expression, aucune étiquette n'aurait été générée pour b. Si l'un ou l'autre des arguments suivants du fractionnement avait été des expressions, les étiquettes de niveau n'auraient été générées.

Instructions de commande de flux

Il y a trois instructions que l'on peut utiliser pour influencer sur le déroulement des commandes qui déterminent l'exécution des instructions.

Groupe d'instructions

Il est possible de faire en sorte qu'un ensemble d'instructions agisse comme un groupe en le faisant précéder et suivre d'accolades : { et }. Ainsi, des groupes d'instructions peuvent être exécutés sous condition avec les instructions if et else.

```
{ a=1; b=2; }
```

Cet exemple regroupe les deux affectations en une seule instruction.

Instruction if

L'instruction if permet l'exécution conditionnelle d'une instruction (ou d'un groupe d'instruction)

```
if (idcfrh==0 && idspoflg==0 && cfnkids>0 ) benefit = 1000;
```

Cet exemple attribue au chef des familles monoparentales une allocation de 1 000 \$. Noter que les parenthèses précèdent et suivent la condition.

Instruction if else

L'instruction if else fait en sorte que deux instructions soient exécutées selon qu'une condition soit vraie ou fausse.

```
if (imigist == 0) benefit = 500 ;  
else benefit = 100 ;
```

Cet exemple attribue une allocation de 500 \$ aux personnes qui n'ont aucun bonus SRG et 100 \$ à toutes les autres personnes. L'instruction if else peut être utilisée à répétition pour spécifier une action selon une condition d'un certain nombre de conditions, comme dans l'exemple suivant :

```
if (idcfrh==0) {  
  if (cfnkids==0) benefit = 0;  
  else if (cfnkids==1) benefit = 100;  
  else if (cfnkids==2) benefit = 300;  
  else if (cfnkids==3) benefit = 600;  
  else benefit = 600 + 400 * (cfnkids-3);  
}
```

Cet exemple attribue une allocation au chef d'une famille de recensement selon le nombre d'enfants que compte la famille. L'allocation par enfant augmente avec la parité de l'enfant.

PARAMÈTRES DE DESCRIPTION

Le paramètre de commande CPRDESC permet à l'utilisateur de donner une courte description générale de l'objet d'une exécution particulière du MSPS.

Le paramètre de commande ALGDESC a une fonction d'information. Sa valeur est créée par le MSPS et elle est stockée dans le fichier des paramètres de commande à des fins de documentation. Elle contient de l'information sur les algorithmes standard et de rechange qui peuvent être utilisés dans la version du MSPS exécutée. Si l'utilisateur suit la procédure donnée dans le document *Guide de programmation* lorsqu'il utilise le système en mode "boîte de verre", une description appropriée des algorithmes fiscaux/de transfert seront donnés dans ALGDESC.

Fichiers de la base de données

L'information contenue dans le BDSPS est répartie dans trois genres de fichiers distincts. Veuillez vous référer au *Guide Introduction et survol* pour une liste détaillée des fichiers de la BD/MSPS. Les fichiers de la base de données sont nommés par Statistique Canada et en devraient pas être changés. La présente section décrit le contenu des divers types de fichiers de la BDSPS et les paramètres de commande correspondants.

FICHIERS DE MÉNAGE/D'INDIVIDU

Le paramètre de commande INSPD précise le nom du fichier d'entrée BDSPS principal. Les fichiers de ce type ont une extension .spd et contiennent de l'information sur le revenu, la démographie et la structure des familles et des ménages.

FICHIERS DES DÉPENSES

Le paramètre de commande INPFVXV précise le nom du fichier de la BDSPS qui contient les données sur les dépenses des ménages. Étant donné que ce ne sont pas tous les utilisateurs qui sont intéressés à ces données (dont la fonction principale consiste à permettre la simulation des taxes à la consommation), la lecture de ce fichier n'est pas nécessaire. La désactivation du paramètre de commande FXVFLAG désactive la lecture du fichier des données sur les dépenses.

FICHIERS DES POIDS

Le paramètre de commande INPWGT précise le nom du fichier de la BDSPS qui contient les poids des familles. Un certain nombre de fichiers des poids ont été fournis avec la BD/MSPS. Ces fichiers ont été construits de façon à reproduire la population estimée ou projetée du Canada par âge, par sexe et par province, pour un certain nombre d'années.

Contrôle de la pondération

Le paramètre de commande WGTFLAG détermine si l'échantillon de la base de données utilisé sera pondéré ou non. Il est habituellement laissé activé, ce qui fait que les poids sont appliqués. Si désactivée, le fichier de poids spécifié dans le paramètre INPWGT (le cas échéant) ne sera pas lu et tous les poids seront fixés à la valeur 1.0. Cette fonction peut être utilisée pour produire les tabulations sans pondération de la BDSPS et elle est aussi utile si les "ménages hypothétiques" générés avec l'utilitaire bldspd (voir le document *Guide d'utilisation des outils*) sont utilisés plutôt que ceux qui sont fournis avec la BDSPS.

FONCTION VALEURS DE RÉFÉRENCE

Les valeurs des variables modélisées produites dans l'exécution du MSPS peuvent être celles auxquelles on accédera dans une exécution subséquente en utilisant les fichiers de résultats pour créer les valeurs modélisées de base, de la façon décrite à la section Variable de variante et Variable de base (utilisant OUTMRSFLAG, OUTMRSVARS, OUTVARMRS et OUTMRSFRAC pour créer le fichier des résultats, et BASMETH et INPBASMRS pour lire le fichier des résultats dans une exécution subséquente). La fonction Valeur de référence permet une forme différente d'accès aux variables créées dans une exécution antérieure. La grande différence réside dans le fait que les variables produites par la fonction Valeurs de référence doivent être soit des variables utilisateur, soit des variables de base de données, plutôt que des variables de base modélisées.

Trois paramètres de commande sont utilisés dans la fonction Valeurs de référence. La fonction est activée en déclenchant REFFLAG. INPREF doit recevoir le nom d'un fichier de résultat (.mrs) produit antérieurement. Le paramètre REFVARS donne la liste des noms des variables que l'utilisateur désire lire du fichier des résultats et peut, si on le désire, être utilisé pour renommer ces variables aussi.

Les diverses opérations que la fonction Valeurs de référence peut exécuter sont décrites en détail dans les sections suivantes.

Création de variables utilisateur d'une exécution antérieure

L'utilisation la plus courante de la fonction Valeurs de référence est probablement l'accès aux variables utilisateur produites dans une exécution antérieure. Pour ce faire, l'utilisateur doit d'abord créer un fichier des résultats qui contient au moins une variable utilisateur en fixant la valeur des paramètres de commande de la même façon que dans l'exemple suivant.

```
UVARFLAG      1
UVAR
    etr = (CF:immtot - CF:immicons) / CF:immtot;
    label(etr) = "Taux de taxe actuels";
    precision(etr) = 3;
OUTMRSFLAG    1
OUTMRSVARS    immtot immicons etr
OUTVARMRS     refvals.mrs
```

Dans cet exemple, une variable utilisateur nommée etr est créée. Cette variable contient le taux d'imposition réel sur le revenu de la famille de recensement résultant de l'exécution d'un MSPS en particulier. Cette variable, ainsi que le revenu consommable et le total modélisé, est enregistré dans un fichier des résultats nommé refvals.mrs. Comme l'instruction de précision a été utilisée, la variable etr sera enregistrée dans le fichier de résultats sous forme d'un nombre à trois décimales.

Pour accéder à la variable etr dans une exécution subséquente, il faudrait utiliser les paramètres de commande suivants :

```
REFFLAG      1
INPREF       refvals.mrs
REFVARS      etr
```

Cela fait que la variable `etr` est récupérée du fichier `refvals.mrs`. La variable `etr` peut alors être utilisée par toute fonction de sortie ou d'établissement de rapport du MSPS. La variable `etr` pourrait être renommée `refetr` si l'on spécifie plutôt ce qui suit :

```
REFFLAG      1
INPREF       refvals.mrs
REFVARS      refetr=etr
```

Cela permettrait que le taux d'imposition réel de la deuxième exécution soit créé et nommé `etr` (avec le paramètre `UVAR`) sans qu'il y ait conflit. Dans l'exemple suivant, les valeurs des variables modélisées `immtot` et `immicons` du fichier `refvals.mrs` sont aussi converties en variables utilisateur :

```
REFFLAG      1
INPREF       refvals.mrs
REFVARS      refetr=etr reftot=immtot refcons=immicons
```

Les variables modélisées doivent être renommées en d'autres variables utilisateur si elles sont mentionnées dans `REFVARS`, comme le montre l'exemple précédent. Il faudrait aussi noter que les variables utilisateur créées avec la fonction Valeurs de référence peuvent être écrasées par la fonction Variables utilisateur.

Les valeurs des variables de tout nombre d'exécutions du MSPS peuvent être combinées en un seul fichier de résultats par la lecture à l'entrée des variables avec la fonction Valeurs de référence, puis dans l'écriture de ces variables, avec les nouvelles produites par l'exécution en cours, dans le fichier des résultats `OUTVARMRS`. Lorsque l'on utilise cette technique, il faut noter que le fichier `OUTVARMRS` ne peut pas avoir le même nom que `INPBASMRS` ou `INPREF`.

Modification des variables de base de données

La fonction Valeurs de référence peut être utilisée pour remplacer les valeurs des variables de la BDSPS. Cette utilisation est illustrée par l'exemple hypothétique suivant, qui a été simplifié et dans lequel un utilisateur tente de modéliser un certain genre de comportement en réaction à une diminution des revenus tirés de l'AC. Dans la première exécution, on suppose que les variables de base existent et les valeurs de variantes proviennent d'un scénario dans lequel les prestations d'AC ont été touchées par des changements de paramètres. Supposons que l'utilisateur désire faire une exécution subséquente dans laquelle les personnes réagissent au manque à gagner des prestations d'AC en accroissant leur revenu d'emploi. Pour ce faire, l'utilisateur produirait un fichier de résultats de la première exécution en utilisant les paramètres suivants :

```
UVARFLAG      1
UVAR
    newemp = idiemp + nneg(_imiuib - imiuib);
    label(newemp) = "Revenu d'emploi (après réaction)";
OUTMRSFLAG     1
OUTVARMRS      run1.mrs
OUTMRSVARS     newemp
```

La variable `newemp` contient le revenu d'emploi après le comportement de réaction (on a supposé qu'une augmentation des prestations d'AC ne diminuera pas le revenu d'emploi). Dans une exécution subséquente, la valeur de base de données du revenu d'emploi peut être remplacée

par le revenu d'emploi survenant après le comportement de réaction en utilisant les paramètres suivants :

```
REFFLAG      1
INPREF       run1.mrs
REFVARS      idiemp=newemp
```

Il faut noter que l'ajustement de la base de données se produit avant que la fonction Valeurs de référence remplace toute variable de base de données. En d'autres mots, les variables de base de données remplacées ne connaissent pas de croissance après leur lecture du fichier INPREF.

Les idées tirées de cet exemple peuvent être combinées avec la fonction Sortie fichier texte et les utilitaires de création de base de données pour créer des fichiers de base de données BDSPS modifiés. Ainsi, on pourrait créer une nouvelle variable utilisateur basée sur les variables de la BD/MSPS en utilisant la fonction Variables utilisateur et la sortie vers un fichier de résultats. Cette variable pourrait être lue au cours d'une exécution subséquente et être affectée à la variable de base de données supplémentaire idext0 en utilisant REFVARS. Les fichiers /spsd/bldspd.cpi et /spsd/bldspd.api pourraient être utilisés pour créer une sortie ASCII de la BDSPS (y compris la nouvelle valeur de idext0) qui servirait alors à l'entrée de l'utilitaire bldspd pour créer un nouveau fichier .spd qui incorpore idext0 en permanence. Veuillez vous référer au *Guide d'utilisation des outils* pour plus de renseignements.

Ajustement de la base de données

L'ajustement de la base de données est une phase distincte des calculs du MSPS. Il s'agit d'un processus par lequel les variables de la BDSPS peuvent être modifiées de façon à refléter les changements qui se sont produits entre l'année de base des données de la BDSPS et l'année qui intéresse l'auteur de l'analyse. L'utilisateur peut désirer modifier les variables de la BDSPS pour d'autres fins aussi. Il est possible de contrôler le processus d'ajustement de la base de données en donnant les valeurs appropriées aux paramètres contenus dans le fichier des paramètres d'ajustement de la base de données (donné dans le paramètre de commande INPAPR).

Il y a deux types distincts de paramètres dans le fichier des paramètres d'ajustement de la base de données. Les paramètres structuraux constituent le premier type et ils visent à représenter les changements qualitatifs apportés à des variables en particulier. Les facteurs de croissance sont du second type et ils servent à accroître les variables de revenus pour tenir compte de l'inflation et (ou) de la croissance économique, ou encore pour compenser la sous-déclaration d'éléments de dépenses. Veuillez consulter le document *Guide des paramètres* pour obtenir plus d'information sur les paramètres d'ajustement de la base de données.

CONTRÔLE DE L'IMPUTATION

Afin de corriger des problèmes connus de sous-déclaration des enquêtes, on a ajouté une série de variables. Les variables s'appliquent aux prestations d'assurance-emploi, aux prestations d'aide sociale, et aux prestations du RPC/RRQ. Pour chacun de ces éléments de revenus, il y a deux variables possibles.

1. Aucune conversion Utiliser les valeurs déclarées à l'origine dans l'EFM

2. Conversion de lignes : Utiliser les valeurs générées par la procédure de conversion des lignes

Veuillez consulter le document *Guide de création de bases de données* pour obtenir plus d'information sur la procédures de conversion réelles. L'utilisateur a le contrôle sur la sélection des variables par l'intermédiaire de trois paramètres d'ajustement de base de données (.apr). Les paramètres sont donnés ci-dessous avec leur valeur par défaut.

IMPUIBOPT	2	# Méthode d'imputation, Assurance-chômage # Source: Use rank conversion
IMPSAOPT	2	# Méthode d'imputation, Aide sociale # Source: Use rank conversion

Variante et base

Dans la terminologie utilisée dans le MSPS, "Variante" représente l'ensemble des variables, algorithmes et paramètres fiscaux/de transfert qui correspondent à l'un des deux calculs de modèles simultanés possibles. L'autre ensemble de calculs possible est décrit par le terme "Base". Les résultats qui intéressent principalement l'utilisateur seront habituellement les valeurs de variante. Les valeurs de base sont habituellement utilisées pour les seules fins de comparaison. La présente section décrit les paramètres de commande utilisés pour spécifier la façon dont les valeurs de base et les valeurs de variante sont produites.

CONTRÔLE DES VARIANTES

Le paramètre de commande VARMETH fournit de l'information sur l'algorithme utilisé pour produire des valeurs de variantes pour les variables. Ces valeurs et leur signification sont données dans le tableau ci-dessous :

VARMETH	Interprétation
0	Aucune valeur de variable de variante
1	Non utilisé
2	Variables de variante produites avec l'algorithme standard et le fichier des paramètres fiscaux/de transfert INPVARMPR.
3	Variables de variante produites avec un autre algorithme et le fichier des paramètres fiscaux/de transfert INPVARMPR.

Les paramètres utilisés pour générer des valeurs de variante sont donnés dans un fichier dont le nom est fourni dans le paramètre de commande INPVARMPR. L'emplacement exact du fichier peut être spécifié au complet, comme c:\spsd\ba02.mpr, ou on peut supposer que le fichier se trouve dans le répertoire courant, p. ex. var1.mpr. Pendant l'interaction du MSPS avec l'utilisateur, il est possible de changer les valeurs des paramètres de variante. Si ces changements sont apportés, le MSPS copiera un fichier de paramètres contenant les changements apportés au fichier dont le nom est spécifié dans le paramètre de commande OUTVARMPR. Ceci permet d'utiliser un fichier (INPVARMPR) comme point de départ pour les valeurs des paramètres de variante, avec un autre fichier (OUTVARMPR) contenant les valeurs réelles qui ont été utilisées pour produire les valeurs des variables de variante.

Le paramètre de commande VARALG vise à informer. Il donne le nom global de l'algorithme (soit standard, soit de rechange, selon ce qui est spécifié dans VARMETH).

Il est possible de préserver les résultats des variables de variante pour utilisation dans des exécutions futures du MSPS. Ceci peut entraîner des économies de temps substantielles si les mêmes résultats sont utilisés plusieurs fois dans des exécutions subséquentes du MSPS. Trois paramètres de commande doivent être spécifiés pour générer un fichier de résultats. OUTMRSFLAG doit déclencher pour activer la fonction en vue de l'enregistrement des variables de variante. OUTVARMRS doit recevoir le nom du fichier dans lequel les résultats seront conservés (le fichier devrait avoir une extension .mrs). Enfin, le paramètre de commande OUTMRSVARS doit contenir une liste des variables pour lesquelles il faut une sortie. Les noms des variables qu'il y a dans OUTMRSVARS sont séparés par des espaces. Les variables utilisateur peuvent être inscrites sous forme d'une liste d'OUTMRSVARS et elles seront produites dans le fichier des résultats. Ces variables peuvent alors être récupérées pendant une exécution subséquente avec la fonction Valeurs de référence.

Comme c'est le cas avec tous les paramètres de chaîne, les valeurs longues peuvent être continuées sur des lignes subséquentes (voir la section Chaînes). La section Fonction Sélection décrit la façon dont le fichier de résultats peut être utilisé pour fournir des valeurs des variables de base dans une exécution subséquente de MSPS.

Par défaut, chacune des variables qu'il y a dans le fichier de résultats verra sa valeur arrondie au chiffre entier le plus proche, avant l'écriture. Si l'instruction de précision a été utilisée à la création de la variable utilisateur, le niveau de précision sera conservé dans le fichier de résultats. Si l'on active OUTMRSFRAC, toutes les variables seront enregistrées à double précision. Bien que plus précis, ce format accroît de beaucoup la taille du fichier.

L'utilitaire bldmrs.exe peut construire un fichier de résultats du MSPS à partir d'un fichier ASCII. Cet utilitaire et un nouvel utilitaire connexe, spsdinfo, sont décrits dans le document *Guide d'utilisation des outils*.

CONTRÔLE DE LA BASE

Le paramètre de commande BASMETH détermine l'algorithme qui est utilisé pour générer les valeurs des variables de base. Il faut noter que, à l'encontre de VARMETH, la valeur 1 permet à l'utilisateur d'utiliser un fichier de résultats produit auparavant pour générer les valeurs de variable de base. Ce processus est décrit plus en détail ci-dessous. Le tableau suivant donne les valeurs valides de BASMETH et leur signification.

BASMETH	Interprétation
0	Aucune valeur de variable de base
1	Les variables de base proviennent du fichier des résultats existant INPBASMRS.
2	Variables de base produites avec l'algorithme ordinaire et les fichiers de paramètres fiscaux/de transfert INPBASMPR.
3	Variables de base produites avec l'algorithme de rechange et le fichier des paramètres fiscaux/de transfert INPBASMPR.

Le paramètre BASALG, comme le paramètre VARALG, vise à informer. Il contient une chaîne décrivant l'algorithme global utilisé pour produire les valeurs des variables de base.

Si la valeur de BASMETH a été fixée à 1, il faut spécifier une valeur pour INPBASMRS. INPBASMRS donne le nom du fichier contenant les valeurs qui doivent être utilisées pour les variables de base. Le paramètre de commande d'information INPMRSVARS contiendra une liste de variables obtenues de INPBASMRS. Les valeurs de base de toutes les autres variables seront zéro.

Si la valeur de BASMETH a été fixée à 2 ou 3, il faut spécifier une valeur pour INPBASMPR. INPBASMPR donne le nom du fichier contenant les paramètres fiscaux/de transfert qui doivent être utilisés pour générer les valeurs des variables de base. Le MSPS ne fournit pas de fonction qui permet de faire la modification interactive de ces paramètres. Au besoin, on peut utiliser un éditeur de texte pour cette fin.

Fonction Sous-échantillonnage

Le MSPS s'exécute rapidement sur les ordinateurs modernes, donc, le sous-échantillonnage n'est pas nécessaire habituellement. Par contre, il est toujours possible d'exécuter le MSPS avec un échantillon partiel.

Dans le MSPS Visuel, le sous-échantillonnage est contrôlé en donnant à un paramètre de commande la valeur désirée pour obtenir un sous-échantillon donné. Le MSPS lit et traite alors la BDSPS jusqu'à ce que le sous-échantillon désiré soit obtenu.

Dans le MSPS Classique, il est possible de commander le sous-échantillonnage de deux façons. Avec la première méthode, l'utilisateur affecte le paramètre de commande SAMPLEREQ au sous-échantillon désiré. Le MSPS lit alors et traite la BDSPS jusqu'à ce que le sous-échantillon désiré soit atteint. Avec la seconde méthode, l'utilisateur appuie sur CTRL-C pour interrompre l'exécution après le traitement d'une certaine fraction de la BDSPS. Dans un cas comme dans l'autre, le sous-échantillon vraiment obtenu est enregistré dans le paramètre de commande d'information SAMPLE, et les tables de sortie sont établies par la réciproque de cette valeur.

La lecture de la BDSPS fonctionne toujours dans l'ordre, mais un sous-échantillon relativement non faussé est quand même obtenu étant donné que les ménages contenus dans la BDSPS sont vraiment classés au hasard. L'ordre n'est pas entièrement aléatoire, cependant. Pour améliorer

la précision des résultats obtenus lorsque l'on utilise le sous-échantillonnage, la BDSPS a été organisée comme un échantillon stratifié dynamique. Les strates utilisées sont la province et le revenu des ménages et les échantillons de stratification des sous-ensembles sont produits pour des valeurs 5 %, 10 % et 25 % de SAMPLE. Des échantillons de ces grosseurs ont été organisés de façon que l'on ait des distributions représentatives de provinces et de revenus de ménage.

La fonction de sous-échantillonnage fonctionne indépendamment de la fonction de sélection (voir la section Fonction Sélection). Cela signifie que, pour une valeur donnée de SAMPLE, les mêmes ménages font l'objet d'une lecture et d'un traitement, quelle que soit la sélection. Si l'utilisateur désire échantillonner explicitement un groupe qui l'intéresse en particulier, il est possible d'activer une série de nombres aléatoires et le nombre aléatoire correspondant peut être utilisé dans l'expression SELSPEC (voir la section Fonction Nombres aléatoires pour obtenir plus d'information sur les nombres aléatoires du MSPS).

Fonction Sélection

La fonction Sélection du MSPS donne un moyen pour restreindre l'ensemble d'individus et de familles qui sont traités par les fonctions Sortie et Établissement de rapports du MSPS. La sélection comble deux besoins distincts. En premier lieu, elle permet à l'analyse de porter sur un sous-groupe d'intérêt en particulier. En second lieu, elle permet au MSPS de s'exécuter plus rapidement en restreignant le nombre de ménages traités à ceux qui intéressent l'utilisateur.

Le paramètre de commande SELFLAG doit être déclenché pour activer la fonction Sélection. Lorsque la fonction est désactivée mais que les autres paramètres qui commandent la fonction (SELSPEC et SELUNIT) contiennent des valeurs, ils conservent leurs valeurs. Par conséquent, en changeant la valeur de SELFLAG, l'utilisateur peut momentanément désactiver la spécification d'une sélection compliquée et la réactiver facilement par la suite.

Le paramètre SELSPEC est une chaîne contenant une expression (voir la section Expressions pour obtenir une description des expressions du MSPS) qui est évaluée pour chaque individu du ménage. Si le résultat de l'évaluation n'est pas 0, l'individu est considéré comme sélectionné. Si un individu est sélectionné, toutes les personnes de la même unité de famille (comme l'indique SELUNIT) est aussi sélectionnée, quelle que soit la valeur de SELSPEC pour ces autres individus. En d'autres mots, si au moins une personne de l'unité de famille (indiqué par SELUNIT) est sélectionnée, toute l'unité de famille est sélectionnée. Il faut noter que, si SELUNIT est fixé à 0, alors chaque individu est sélectionné selon la valeur de SELSPEC pour cet individu. Dans ce cas, il n'y a aucune propagation aux autres membres de la famille.

Les valeurs de SELUNIT et leurs significations sont données dans le tableau ci-dessous. SELUNIT détermine le type d'unités qui doit être sélectionné, mais n'a aucune influence sur le niveau de l'analyse utilisée pour calculer SELSPEC pour chaque individu. Le niveau de l'analyse des variables dans l'expression SELSPEC est toujours le niveau de l'individu. Il faut cependant noter que les variables individu dans l'expression SELSPEC peuvent être contournées à un niveau supérieur d'analyse si l'on utilise les préfixes appropriés de niveau de famille. Par exemple, la variable idiemp dans SELSPEC réfère au revenu d'emploi de l'individu, tandis que

CF:idiemp réfère au revenu d'emploi total de la famille de recensement dont l'individu fait partie.

SELUNIT	Niveau de propagation de la sélection
0	Individu (aucune propagation)
1	Famille nucléaire
2	Famille de recensement
3	Famille économique
4	Ménage

Tableau 1 : Interprétation des valeurs SELUNIT

La combinaison de la propagation de la sélection avec SELUNIT, des expressions générales utilisant SELSPEC et des préfixes de niveau d'analyse à l'intérieur de SELSPEC crée une puissante fonction de sélection générale. Les exemples suivants illustrent le fonctionnement de ces mécanismes.

Ex. n°	SELUNIT	SELSPEC	Explication
1	0	hdprov == 0	Individus à Terre-Neuve.
2	4	hdprov == 0	Ménages avec des individus à Terre-Neuve (plus efficace que l'exemple précédent).
3	3	idage >= 65	Familles économiques comptant au moins une personne âgée.
4	3	idage >= 65 && idiemp > 0	Familles économiques comptant des personnes âgées avec revenu d'emploi.
5	0	idiemp > 10000	Individus ayant un revenu d'emploi de plus de 10 000 \$.
6	2	idiemp > 10000	Familles de recensement comptant au moins une personne ayant un revenu d'emploi de plus de 10 000 \$.
7	2	CF:idiemp > 10000	Familles de recensement ayant un revenu d'emploi de plus de 10 000 \$.
8	0	imiuib > 0 && (EF:idiemp / efnpers) > 10000	Individus recevant des prestations d'A.C. dans des familles économiques dont le revenu d'emploi par personne dépasse 10 000 \$.
9	2	(immdisp - _immdisp) > 1000	Familles de recensement comptant au moins une personne dont le revenu disponible a augmenté d'au moins 1 000 \$ entre la base et la variante.

Tableau 2 : Exemples de fonctions Sélection

La fonction Sélection fonctionne essentiellement au niveau de l'individu. Le résultat final de la sélection est que les individus d'un ménage sont ou non marqués comme sélectionnés. Cependant, les fonctions Sortie et Établissement des rapports du MSPS fonctionnent aux niveaux d'analyse de familles précisées par l'utilisateur. Les règles suivantes indiquent si une unité de famille est ou non considérée comme une observation pour des fins d'établissement d'un rapport.

Si tout individu d'une unité de famille est marqué comme sélectionné, l'unité de famille est considérée comme une observation pour l'établissement des rapports. Par contre, si aucun individu d'une unité de famille est marqué comme sélectionné, l'unité de famille n'est pas portée au rapport.

L'utilisateur doit faire preuve de prudence lorsque SELUNIT est inférieur au niveau d'analyse spécifié dans l'une des fonctions de sortie ou d'établissement des rapports du MSPS. Dans ces cas, les unités de famille qui seront données dans le rapport devraient être des unités partielles. Par exemple, si SASUNIT est fixé à 2 et SASVARS à idiemp, et que les paramètres de la fonction Sélection sont ceux qui sont indiqués dans l'exemple 5 du tableau ci-dessus, alors chaque enregistrement dans le SAS résultant correspondra à une famille de recensement et contiendra le revenu d'emploi de tous les membres de la famille de recensement dont le revenu d'emploi a dépassé 10 000 \$.

Pour calculer correctement les taxes, impôts et transferts, le MSPS simule toujours les ménages au complet. De ce fait, la sélection n'a aucun effet sur les valeurs de toute variable, modélisée ou de base de données, au niveau de l'individu. Cependant, si un examen des variables de base de données dans l'expression SELSPEC indique qu'aucun individu du ménage ne pourrait être sélectionné (quelle que soit la valeur de toute variable modélisée dans SELSPEC), le MSPS passe alors immédiatement au ménage suivant. Ceci peut entraîner une diminution considérable du temps d'exécution du MSPS. Quoi qu'il en soit, il faut calculer tous les trains de nombres aléatoires activés afin de vérifier que leurs valeurs seront reproductibles d'exécution en exécution, quelle que soit la sélection.

L'utilisateur peut aussi désirer restreindre le nombre d'individus ou de familles sélectionnés. Le paramètre SELMAX détermine le nombre maximal d'individus ou de familles qui seront traités. Ce dispositif est particulièrement utile lorsque l'on utilise la fonction Sortie texte ou la fonction Points de retournement.

Fonction Taux marginal d'imposition

Le taux marginal d'imposition est défini comme la proportion d'un dollar supplémentaire de revenu qui est imposée. Il s'agit là d'un concept utile, car il permet de déterminer dans quelle mesure l'incitation à obtenir des revenus supplémentaires est réduite par le système fiscal/de transfert. Le MSPS possède une fonction qui permet de calculer le taux marginal d'imposition. Lorsque MARFLAG est déclenché, la fonction est activée et le système fiscal/de transfert s'applique deux fois à chaque ménage, une fois aux revenus originaux et une fois aux revenus augmentés, et la variation qui survient dans le revenu consommable est notée.

La définition de taux marginal d'imposition donné ci-dessus n'est pas complète. Pour calculer un taux marginal d'imposition, on doit en outre préciser la source de revenu qui est touché par l'augmentation, le montant de l'augmentation du revenu, ainsi que les individus qui doivent recevoir l'augmentation. La variation du revenu consommable au niveau de l'individu peut alors être agrégée pour produire les taux marginaux d'imposition à différents niveaux de famille d'analyse. Chacun de ces points est étudié ci-dessous.

Les taux marginaux d'imposition varient selon la source de revenus, car le système fiscal/de transfert traite les revenus différemment selon la source. Ainsi, un dollar supplémentaire de revenu en dividendes est imposé différemment d'un dollar supplémentaire de revenu d'emploi. Le paramètre de commande MARVAR donne le nom de la variable de base de données de revenu qui doit être augmentée pour le calcul des taux marginaux d'imposition. Pour poursuivre l'exemple ci-dessus, si MARVAR contient la valeur idiemp, le revenu d'emploi est augmenté, tandis que, si elle contient la valeur idivid, le revenu en dividendes est augmenté.

Si le montant de l'augmentation du revenu est petit, il est alors possible de faire la mesure précise de la pente locale de la fonction fiscale/de transfert (équivalente au taux marginal d'imposition). Cependant, cela n'est pas toujours souhaitable. La fonction fiscale/de transfert peut changer rapidement sur une petite plage de revenus. Par exemple, si le revenu auquel le supplément de revenu garanti est réduit à 0 (le revenu de seuil), il y a changement instantané du taux marginal d'imposition de 50 %. En outre, il n'est pas clair que les individus ont la possibilité (ou le désir) de changer leurs revenus d'un petit montant en réaction aux fluctuations locales du taux marginal d'imposition. Le paramètre MARAMT détermine le montant de l'augmentation de revenu utilisé pour le calcul des taux marginaux d'imposition.

Lorsque l'on calcule les taux marginaux d'imposition pour les familles, il est nécessaire de préciser non seulement la source et le montant du revenu qui doit être augmenté, mais il faut aussi indiquer quels sont les individus dont le revenu est augmenté. Ainsi, lorsqu'on évalue le taux marginal d'imposition du revenu d'emploi, il n'est pas souhaitable d'augmenter le revenu d'emploi d'individus non employables (p. ex. les enfants et les personnes âgées). Le paramètre MARSPEC indique les individus qui doivent recevoir une augmentation. Si la valeur de l'expression MARSPEC n'est pas zéro pour un individu, alors l'individu reçoit l'augmentation de MARAMT aux fins du calcul du taux marginal d'imposition. Dans l'exemple ci-dessus, en fixant MARSPEC à $\text{idage} > 17 \ \&\& \ \text{idage} < 65$, on restreindrait l'opération d'augmentation au sous-ensemble désiré d'individus.

Les taux marginaux d'imposition diffèrent selon le niveau d'analyse de la famille. Si, par exemple, on prend un couple marié, dont une personne a un revenu nul. Si un montant de 500 \$ était donné à l'individu qui n'a pas de revenu, le taux marginal d'imposition de cette personne serait de zéro. Le revenu consommable du conjoint de la personne diminuerait cependant du fait de la réduction de la déduction/du crédit d'impôt de personne mariée. Par conséquent, le taux marginal d'imposition du conjoint serait calculé à une valeur définie, puisque les impôts du conjoint ont augmenté, même s'il n'a reçu aucun revenu supplémentaire. Cependant, si l'on considère le cas du point de vue de la famille, le taux marginal d'imposition ferait une évaluation à une valeur raisonnable.

Pour permettre le calcul des taux marginaux d'imposition à différents niveaux d'analyse de la famille, le MSPS affecte des valeurs aux variables de niveau individu immaramt et immartax lorsque la fonction Taux marginal d'imposition est activée. immaramt enregistre le montant de l'augmentation de revenu reçu par l'individu. Sa valeur est égale à MARAMT si MARSPEC a évalué la valeur comme non zéro, autrement sa valeur est zéro. immartax enregistre la différence entre immaramt et la variation survenue dans le salaire consommable de l'individu du fait de l'augmentation des revenus. Elle représente le montant de l'impôt perçu de l'individu du fait de l'augmentation de revenus. Il faut noter que de désactiver le paramètre fiscal/de transfert

CTFLAG désactive le calcul des taxes à la consommation, ce qui fait que le calcul de taux marginal d'imposition se fait selon le revenu disponible plutôt que selon le revenu consommable.

En utilisant `immartax` et `immaramt`, l'utilisateur peut calculer les taux marginaux d'imposition de divers niveaux d'analyse de famille par la fonction Variable définie par l'utilisateur (voir la section Fonction Variable définie par l'utilisateur). Par exemple, l'expression `immartax/immaramt` calculera le taux marginal d'imposition.

Le paramètre `MARBASEFLAG` permet à l'utilisateur d'examiner plus en détail les variables modélisées changées ainsi que l'ampleur de ces changements du fait de l'augmentation du revenu. Lorsque le paramètre est activé, toutes les valeurs des variables modélisées associées au revenu qui a changé sont enregistrées sous les noms de variables habituellement réservés aux valeurs d'exécution de base (p. ex., `_immdisp`).

Fonction Variable définie par l'utilisateur

Le MSPS a deux modes d'utilisation distincts, la "boîte noire" et la "boîte de verre". Le mode "boîte de verre" permet à l'utilisateur une grande souplesse dans la construction de nouveaux algorithmes et la création de nouvelles variables, mais il peut être quelque peu compliqué à utiliser et exige une certaine connaissance de la programmation. La fonction Variable définie par l'utilisateur permet à l'utilisateur de créer de nouvelles variables d'établissement de rapports en mode "boîte noire". Elle permet à l'utilisateur de faire de nombreuses analyses qui, autrement, exigeraient qu'on apporte des changements dans la programmation du MSPS. Voici une description de la fonction Variable définie par l'utilisateur.

FONCTION VARIABLE UTILISATEUR

La fonction Variable utilisateur permet de créer de nouvelles variables définies par l'utilisateur dont les valeurs sont créées par des instructions du MSPS. Les variables utilisateur sont définies au niveau de l'individu, bien que des expressions puissent référer à des constructions de niveaux d'analyse plus élevés. Le paramètre `UVARFLAG` doit être déclenché pour que cette fonction soit activée. Le paramètre `UVAR` contient une liste des instructions du MSPS qui permettent de créer et d'affecter les variables utilisateur. Il est possible de créer jusqu'à 75 variables utilisateur. Il faut noter que les affectations dans `UVAR` peuvent remplacer les valeurs de variable utilisateur portant le même nom créé par la fonction Valeurs de référence.

Dans l'exemple suivant, l'utilisateur désire créer un nouveau fichier des poids dans lequel la population de Terre-Neuve a été augmentée de 3 % par rapport à sa valeur de base. Prenons l'ensemble suivant de paramètres de commande :

```
VARMETH      0
SEED          1
              42
UVARFLAG     1
UVAR
    if (hdprov==NFLD) {
        fltwgt = 1.03 * hdwgthh;
        intwgt = trunc(fltwgt);
        if (idrand0 < (fltwgt-intwgt)) intwgt = intwgt + 1;
```

```

    }
    else intwgt = hdwgthh;
ASCFLAG      1
ASCUNIT      4
ASCVARS      intwgt
ASCSTYLE     4
OUTASC       newwgts.prn

```

VARMETH a été fixé à zéro pour éviter le calcul inutile de variables modélisées. Un des trains de nombres aléatoires a été activé, avec 42 comme valeur de départ arbitraire. Puisque VARMETH a été fixé zéro, le fait que les nombres aléatoires du modèle ont été désactivés n'aura aucun impact sur les résultats. Un certain nombre de variables utilisateur sont créées dans UVAR et l'une d'elles (intwgt) est inscrite dans un fichier mixte pour entrée subséquente dans l'utilitaire bldwgt (décrit dans le document *Guide d'utilisation des outils*).

Les instructions contenues dans UVAR sont interprétées comme suit. En premier, intwgt est fixé à hdwgthh si la province n'est pas Terre-Neuve. Si la province est Terre-Neuve, hdwgthh est augmenté de 3 %, ce qui crée la variable fltwgt. Cette variable ne peut pas être produite telle quelle à la sortie, puisqu'elle contient une partie fractionnaire, tandis que les poids de la BDSPS doivent toujours être entiers. La partie fractionnaire de fltwgt est par conséquent éliminée pour produire intwgt. La ligne suivante ajoute 1 à intwgt par un mécanisme probabiliste selon la valeur de la partie fractionnaire : plus la partie fractionnaire est proche de 1, plus il est probable que intwgt soit augmenté.

UVAR comprend une liste d'instructions. Chaque instruction est évaluée, pour chaque membre du ménage, avant que l'instruction suivante soit évaluée. Cela signifie que les exemples, comme la suivante, fonctionneront comme prévu :

```

UVAR
    income = idiemp + idisenf + idisefm;
    incrat = income / CF:income;

```

Dans cet exemple, incrat est la part de l'individu du revenu de la famille de recensement. Cela fonctionne parce que la première instruction est évaluée pour tous les individus avant l'évaluation de la seconde instruction. Il y a des coûts de calcul associés à l'évaluation des instructions de cette manière, parce que le MSPS doit traiter tous les membres d'un ménage pour chaque instruction contenue dans UVAR. Il est possible d'éviter ces coûts en regroupant les instructions dans UVAR de la façon donnée dans l'exemple suivant :

```

UVAR
{
    income = idiemp + idisenf + idisefm ;
    tax    = idftax ;
    taxrate = tax / income ;
}
rate = taxrate / (CF:tax / CF:income);

```

Dans cet exemple, UVAR contient deux instructions, la première étant une instruction composée. income, tax et rate peuvent être calculés sans référence aux autres membres de la famille; on les a donc regroupés dans une instruction composée pour des fins d'efficacité. Le calcul du taux ne peut être regroupé dans les instructions précédentes parce qu'il faut que l'impôt et le revenu soient calculés auparavant pour tous les membres de la famille. rate est le rapport du taux d'imposition d'un individu par rapport au taux d'imposition global de la famille de l'individu.

Fonction Sortie texte

La fonction Sortie texte du MSPS crée un fichier texte contenant des microdonnées. Il y a deux classes d'utilisation auxquelles on peut donner les fichiers de ce genre. On peut soit l'imprimer ou le consulter directement, soit l'utiliser comme entrée dans d'autres progiciels, comme une base de données ou un chiffrier.

La fonction est activée par le déclenchement de ASCFLAG et en spécifiant le nom du fichier à produire dans OUTASC. Comme c'est le cas avec les autres paramètres contenant des noms de fichier de sortie, OUTASC sera généré automatiquement s'il n'est pas spécifié, selon le nom de fichier des paramètres de commande.

Les variables dont les valeurs doivent être affichées sont indiquées dans ASCVARS. Chaque cas qui est produit à la sortie correspond au niveau d'analyse de données dans ASCUNIT, comme l'indique le tableau suivant.

ASCUNIT	Niveau de la famille à la sortie du cas
0	Individu
1	Famille nucléaire
2	Famille de recensement
3	Famille économique
4	Ménage

Tableau 3 : Interprétation des valeurs données dans ASCUNIT

Le nombre de décimales pour la précision qui doit être donnée dans le fichier de sortie est déterminé par le paramètre ASCEXTPRC (Nombre de décimales pour plus de précision). Si on laisse ASCEXTPRC à zéro, la fonction Sortie texte donnera seulement des nombres entiers des variables. Lorsque ASCEXTPRC est fixé à un nombre positif, elle ajoute le nombre requis de décimales aux nombres (mais non aux variables entières). La précision supplémentaire est utile pour le calcul des taux marginaux d'imposition quand on utilise la fonction Points de retournement. L'utilisateur devrait faire preuve de prudence et regrouper les variables entières et les variables de nombres dans ASCVARS car cela peut avoir de fortes répercussions sur l'aspect du fichier de sortie.

Un certain nombre de styles de sortie sont acceptés. On peut donner un style particulier à la sortie en affectant une valeur appropriée au paramètre ASCSTYLE. Il y a cinq styles de sortie différents qui correspondent aux valeurs 1, 2, 3, 4 et 5 de ASCSTYLE. Le format de sortie de chaque style est plus facile à illustrer qu'à expliquer. L'aspect de la sortie, dans ASCUNIT est fixé à 0 (les cas des individus) et ASCVARS à hdprov idage idsex idiemp, est donné ci-dessous pour chaque style.

Une valeur 1 de ASCSTYLE produit un rapport conçu de façon à être facile à lire. Il y a un ménage par page et une variable par ligne. Le nom et l'étiquette de la variable sont imprimés et les valeurs de chaque unité sont indiquées dans des colonnes alignées. Dans cet exemple, les unités sont des individus, parce que ASCUNIT a été fixé à 0. Les variables qui existent au niveau du ménage sont données dans la première colonne, puisque l'on sait que leurs valeurs sont identiques pour toutes les unités du ménage. La fonction Sélection (voir la section Fonction

Sélection) est en général utilisée de concert avec ce style, puisque, autrement, on pourrait obtenir de très gros fichiers.

Dans un cas spécial appliquant une valeur 1 à ASCSTYLE, si un tiret (“-”) est trouvé dans ASCVARS, une ligne de séparation est générée dans le rapport. Le fichier `detsum.cpi`, dans le répertoire `/spsd`, illustre ce dispositif, et contient des valeurs pour les paramètres de fonction de sortie texte qui créent un ensemble utile de variables pour utilisation générale. Il est facile de l’activer par la fonction Lecture de paramètre.

hdseqhh	Numéro d’ordre du ménage	61		
idefseq	Numéro d’ordre de la famille économique	0	0	0
idcfseq	Numéro d’ordre de la famille de recensement	0	0	0
hdprov	Province	6		
idage	Âge	28	24	0
idsex	Sexe	0	1	1
idcfrh	Lien avec le chef de la famille de recensement	0	1	2
immmkt	Revenu d’emploi	27953	17350	0
immtran	Tous les revenus de transferts.....	194	371	0
<saut de page>				
hdseqhh	Numéro d’ordre du ménage	62		
idefseq	Numéro d’ordre de la famille économique	0	0	0
idcfseq	Numéro d’ordre de la famille de recensement ...	0	0	0
hdprov	Province	6		
idage	Âge	53	50	18
idsex	Sexe	0	1	0
idcfrh	Lien avec le chef de la famille de recensement.	0	1	2
immmkt	Revenu d’emploi.....	36457	3750	4274
immtran	Tous les revenus de transferts	0	1222	0

Une valeur 2 de ASCSTYLE produit un rapport conçu de façon à être consulté dans une fonction Importation de chiffrier. L’agencement est semblable à celui que l’on obtient avec une valeur 1 de ASCSTYLE, mais les étiquettes de variable et tous les espaces superflus ont été éliminés et le caractère de saut de page d’imprimante a été remplacé par une chaîne de caractères vide.

```

" "
"hdseqhh" 61
"idefseq" 0 0 0
"idcfseq" 0 0 0
"hdprov" 6
"idage" 28 24 0
"idsex" 0 1 1
"idcfrh" 0 1 2
"immmkt" 27953 17350 0
"immtran" 194 371 0
" "
"hdseqhh" 62
"idefseq" 0 0 0
"idcfseq" 0 0 0
"hdprov" 6
"idage" 53 50 18
"idsex" 0 1 0
"idcfrh" 0 1 2
"immmkt" 36457 3750 4274
"immtran" 0 1222 0

```


Une valeur 3 de ASCSTYLE produit un rapport conçu de façon à pouvoir être consulté dans un système de chiffrier ou de base de données. Chaque unité (dans l'exemple, chaque individu) est enregistré sur une seule ligne, avec un seul espace entre chaque valeur variable. La première ligne du fichier contient une liste des noms des variables, dans l'ordre où elles sont inscrites dans chaque ligne.

```
"hdseqhh" "idefseq" "idcfseq" "hdprov" "idage" "idsex" "idcfrh" "immmkt"
"immtran"
61 0 0 6 28 0 0 27953 194
61 0 0 6 24 1 1 17350 371
61 0 0 6 0 1 2 0 0
62 0 0 6 53 0 0 36457 0
62 0 0 6 50 1 1 3750 1222
62 0 0 6 18 0 2 4274 0
```

Une valeur 4 de ASCSTYLE produit un rapport identique à celui qu'on obtient avec la valeur 3, sauf que la première ligne, qui pourrait être très longue ou ne pas convenir pour certaines fins, est éliminée.

```
61 0 0 6 28 0 0 27953 194
61 0 0 6 24 1 1 17350 371
61 0 0 6 0 1 2 0 0
62 0 0 6 53 0 0 36457 0
62 0 0 6 50 1 1 3750 1222
62 0 0 6 18 0 2 4274 0
```

Une valeur 5 de ASCSTYLE produit un rapport conçu en vue d'une conversion dans un format compressé, qui peut être lu par le MSPS. La sortie générée quand on utilise cette valeur est utilisée par les utilitaires bldspd, bldfxv et bldwgt. On trouvera plus d'information dans le document *Guide d'utilisation des outils*.

Voici un exemple où ASCSTYLE=5 pour un ménage (toutes les valeurs de variables ne sont pas données).

```
000001 47 217 2 1 2 1 3 4 1
0 0 0 0 0 0 0 0 42 0 3 11 13 99 1 2 3 0 52 1 0 0 0 0 0 0 0 0 0 1 0 0 0 ...
```

Un rapport produit avec ASCSTYLE=5 donne un format fixe qui contient toutes les variables demandées, où le délimiteur est un caractère vide et qui contient tous les enregistrements par cas commençant avec un enregistrement de ménage qui est suivi par les enregistrements individuels.

Fonction Sortie SAS

SAS Institute produit une version pour micro-ordinateur de son système d'analyse statistique sur ordinateur central avec des fonctionnalités substantiellement identiques. SAS possède des fonctions étendues d'analyse, d'établissement de rapports et de manipulation de fichiers de microdonnées. Afin de tirer profit de ces fonctions, le MSPS incorpore une interface avec le système SAS pour PC. Cette interface produit des fichiers binaires à documentation intégrée du même format que SAS lui-même utilise pour stocker les données. Le moteur de bibliothèque utilisé est 6.03. Donc, les données peuvent être rendues accessibles pour SAS sans les problèmes qu'impliquent le traitement de la disposition des fichiers texte, les étiquettes et les définitions appropriées de variables, etc. Un autre avantage est que les exigences d'espace disque et de

temps d'exécution sont fortement réduites, puisque SAS devrait habituellement lire un grand fichier texte, puis le convertir en gros fichier binaire avant que toute opération ne puisse être effectuée sur les données.

De nombreuses opérations qui peuvent être exécutées avec SAS, comme la sélection des enregistrements, la création de nouvelles variables, les tableaux croisés et les analyses de distribution peuvent être exécutées avec les fonctions intégrées du MSPS. Ces fonctions sont beaucoup plus rapides que leurs équivalents SAS et n'exigent pas beaucoup d'espace disque. Par conséquent, en règle générale, les fonctions intégrées du MSPS devraient, dans la mesure du possible, être utilisées de préférence à une solution SAS. L'utilisateur ne devrait utiliser SAS que si la fonction qu'il doit exécuter ne peut l'être avec le MSPS. En voici des exemples : analyse de régression, diagramme de dispersion et analyse de microdonnées sur au moins trois modèles exécutés simultanément.

La fonction Sortie SAS est activé en déclenchant SASFLAG. Par défaut, la BD/MSPS établit OUTSAS à "spsmtemp.ssd". L'utilisateur peut remplacer "spsmtemp.ssd" par tout autre nom s'il le désire en utilisant la procédure PROC DATASETS, en SAS. Chaque enregistrement du fichier SAS produit correspond au niveau d'analyse de la famille donnée SASUNIT, comme l'indique le tableau ci-dessous. Les variables qui doivent être produites à la sortie sont données dans le paramètre de chaîne SASVARS. Le paramètre de chaîne SASTITLE peut être utilisé pour fournir un titre fixe pour le fichier SAS produit.

SASUNIT	Niveau de la famille de la sortie SAS
0	Individu
1	Famille nucléaire
2	Famille de recensement
3	Famille économique
4	Ménage

Tableau 4 : Interprétation des valeurs de SASUNIT

Le MSPS crée une variable spéciale, hdwgthhs, conçue de façon à être utilisée avec la fonction Sortie SAS lorsque la fonction Sous-échantillonnage (voir la section Fonction Sous-échantillonnage) est aussi active. La valeur de hdwgthhs est la même que le poids du ménage hdwgthh, sauf qu'elle a été accrue par la réciproque de SAMPLEREQ, la grosseur de l'échantillon demandé par l'utilisateur. Ceci permet aux tabulations SAS des sous-échantillons d'avoir des totaux pondérés globaux corrects. Il faut noter que hdwgthhs est représenté avec une précision restreinte dans le fichier SAS, de façon que les totalisations ne correspondront pas exactement aux tables produites par le MSPS. Les résultats seront cependant identiques si SAMPLEREQ est fixé à 1.0.

Il y a une restriction à l'utilisation du paramètre OUTSAS. Étant donné que les fichiers SAS contiennent une structure d'en-tête chiffrée, le MSPS doit utiliser l'en-tête d'un fichier SAS existant ayant un nom identique pour créer un nouveau fichier SAS nommé OUTSAS. La structure d'un tel fichier (p. ex. les variables ou le nombre d'enregistrements qu'il contient) n'a pas d'importance - tout ce qu'il faut, c'est un fichier qui est produit par SAS et qui a le bon

moteur de bibliothèque (6.03 ou 6.04). Il est facile de créer un fichier dans SAS en utilisant des instructions de la façon suivante :

```
LIBNAME CD `.`;  
DATA CD.MYRUN;  
RUN;
```

Il faut noter que, si vous utilisez une version ultérieure de SAS, vous pouvez aussi avoir besoin de spécifier un moteur de bibliothèque dans LIBNAME. Consultez la documentation de SAS pour obtenir plus d'information. Dans la version 6.12 de SAS, par exemple, la syntaxe appropriée était :

```
LIBNAME CD V603 `.`;
```

Par contre, le MSPS sait comment créer un fichier nommé spsmtemp.ssd, et il utilisera ce nom si un fichier de gabarit SAS existant nommé OUTSAS n'existe pas (dans ce cas, le MSPS changera le nom du fichier de OUTSAS en spsmtemp.ssd). L'utilisateur pourra alors changer ce nom en un nom qui convient mieux dans SAS en utilisant les instructions suivantes :

```
LIBNAME CD `.`;  
PROC DATASETS LIBRARY=CD;  
    CHANGE SPSMTEMP=MYRUN;  
RUN;
```

Si le SPSMTEMP est utilisé, le fichier est créé en format SAS 6.03. Pour alléger la tâche d'écriture des fichiers SAS à partir de ce fichier, l'utilisateur devrait créer une bibliothèque vide dans un autre répertoire et utiliser la méthode de copie (COPY) ressemblant à ce qui suit :

```
LIBNAME CDA `***`;  
PROC COPY IN=CD OUT=CDA;  
    SELECT MYRUN;  
RUN;
```

Si cela n'est pas fait, l'écriture d'un fichier traité dans un format SAS permanent peut exiger de 60 à 100 fois plus de temps qu'il n'en faut habituellement parce que SAS tentera d'écrire votre fichier SAS en format 6.03 plutôt que dans le format associé à la version SAS que vous avez.

Les fichiers SAS ne contiennent pas toute l'information nécessaire pour leur utilisation. Ce qui manque, c'est l'information sur les chaînes utilisées pour afficher les variables de niveau de classification (p. ex. l'association des codes numériques 0 à 9 aux chaînes des noms de province pour la variable hdprov). Dans SAS, cette information est donnée par l'association d'un format SAS nommé avec chaque variable de classe, et par la définition de ce format dans la bibliothèque de format SAS. Une bibliothèque de formats SAS est en retour produite quand on fournit les instructions à la procédure SAS PROC FORMAT. Lorsque le MSPS produit un fichier SAS qui contient les variables de classe, il produit aussi un fichier associé d'instructions SAS qui créera la bibliothèque de formats SAS dont on a besoin. Le fichier qui contient ces instructions est le même que OUTSAS, mais il a l'extension .sfm. Ces points sont illustrés dans l'exemple suivant :

Supposons qu'un fichier SAS nommé sasexam1.ssd était déjà créé dans le répertoire courant et que les paramètres de commande suivants étaient inclus dans le traitement MSPS :

```
SAMPLEREQ      0.049919527      # Taille de l'échantillon demandé
```

```

SASFLAG          1          # Drapeau de déclenchement de la fonction de
sortie SAS
OUTSAS           sasexam1.ssd      # Nom du fichier SAS (sortie) des résultats
SASUNIT          0          # Niveau de famille pour les sorties SAS
SASVARS          # Variables choisies pour sortie SAS
[chaîne]
                hdwgthhs hdurb idind idiemp
SASTITLE Exemple 1          # label de fichier SAS [chaîne]

```

Le MSPS écraserait le fichier sasexam1.ssd par un fichier contenant les variables demandées et, en outre, produirait un fichier sasexam1.sfm dont le contenu serait le suivant :

```

PROC FORMAT;

VALUE $URB
    '0' = '>500,000'
    '1' = '100,000-499,999'
    '2' = '30,000-99,999'
    '3' = '<30,000'
    '4' = 'Rural'
;
VALUE $IND
    '0' = 'Jamais actif'
    '1' = 'Agriculture'
    '2' = 'Autres secteurs primaires'
    '3' = 'Industries manuf., (Non durables)'
    '4' = 'Industries manuf., (Durables)'
    '5' = 'Construction'
    '6' = 'Transports & Communications'
    '7' = 'Commerce de gros'
    '8' = 'Commerce de détail'
    '9' = 'Intermédiaires financiers, etc.'
    ':' = 'Éducation et services connexes'
    ';' = 'Santé, bien-être, religion'
    '<' = 'Loisirs, logement, alimentation'
    '=' = 'Services aux entreprises et services divers'
    '>' = 'Administration publique'
    '?' = 'Travaillait il y a plus de 5 ans'
;
RUN;

```

Un traitement SAS typique de ces fichiers et la production d'une totalisation d'échantillons pourraient se traduire comme suit :

```

OPTIONS PS=66 LS=100;
LIBNAME cd '.';

%INCLUDE 'sasexam1.sfm';

PROC TABULATE DATA=cd.sasexam1;
    FREQ hdwgthhs;
    VAR idiemp;
    CLASS hdurb idind;
    TABLE idind , hdurb * idiemp * MEAN /RTS=32;
RUN;

```

La sortie journal du traitement SAS ressemble à ce qui suit (les commentaires d'en-tête du code de source ont été supprimés) :

NOTE: Copyright(c) 1985,86,87 SAS Institute Inc., Cary, NC 27512-8000, U.S.A.
NOTE: SAS (r) Proprietary Software Release 6.03
Licensed to STATISTICS CANADA, Site 11250001.

NOTE: AUTOEXEC processing completed.

```
10
11     OPTIONS PS=66 LS=100;
12     LIBNAME cd `.`;
13
14     %INCLUDE `sasexam1.sfm`;
NOTE: Format $URB has been output.
NOTE: Format $IND has been output.
NOTE: The PROCEDURE FORMAT used 4.00 seconds.
43
44     PROC TABULATE DATA=cd.sasexam1;
45         FREQ hdwgthhs;
46         VAR idiemp;
47         CLASS hdurb idind;
48         TABLE idind , hdurb * idiemp * MEAN /RTS=32;
49     RUN;
```

NOTE: The PROCEDURE TABULATE used 28.00 seconds.

NOTE: SAS Institute Inc., SAS Circle, PO Box 8000, Cary, NC 27512-8000

La sortie Liste du travail SAS (le résultat de PROC TABULATE) ressemble à ce qui suit :

SAS 2:38 Thursday, November 3, 1988

1

	Taille de la région urbaine				
	>500,000 Revenu d'emploi MOYEN	100,000- 499,999 Revenu d'emploi MOYEN	30,000- 99,999 Revenu d'emploi MOYEN	<30,000 Revenu d'emploi MOYEN	Rural Revenu d'emploi MOYEN
Industrie					
Jamais actif	6.76	0.00	0.00	0.13	0.00
Agriculture	2698.67	2817.82	10818.05	2061.59	5553.86
Autres secteurs primaires	31381.33	15286.83	23136.33	34467.59	12669.33
Industries manuf., (Non durables)	15810.92	19891.68	14169.33	11600.21	15971.44
Industries manuf., (Durables)	19260.93	19184.97	18454.21	12534.72	16083.21
Construction	18343.82	6477.05	12288.02	9963.10	13942/21
Transports et Communications	22447.44	19853.87	19917.40	18425.47	15658.44
Commerce de gros	18825.63	18380.04	17831.44	20956.16	11500.33
Commerce de détail	10105.57	8703.40	10764.16	6511.38	7975.06
Intermédiaires financiers, etc.	16439.42	18389.81	22450.45	9194.94	18690.63
Éducation et services connexes	17711.69	21707.24	14946.91	13942.98	18353.11
Santé, bien-être, religion	16600.30	20511.94	15668.01	14343.19	12162.36
Loisirs, logement, alimentation	7338.29	6572.39	3373.93	8054.35	3443.01
Services aux entreprises et services divers	11549.08	9165.25	8003.49	9348.16	4496.60
Public Administration	20048.38	14224.60	21265.92	14206.35	14602.06
Travaillait il y a plus de 5 ans	162.66	47.73	233.73	215.25	1026.99

Fonction Totalisation intégrée

Le MSPS peut générer un certain nombre de tables prédéfinies qui contiennent de l'information générale sur les résultats d'une simulation. Les tables contiennent l'information sur les caractéristiques démographiques, de revenus, de fiscalité et de transfert des unités de famille.

Pour les activer, il suffit d'activer un paramètre de drapeau. En outre, les paramètres auxiliaires commandent des dispositifs de certaines des tables. Parce qu'elles sont si faciles à activer et qu'elles sont conçues de façon à fournir une image globale du système fiscal/de transfert, elles sont souvent utilisées comme sortie initiale pour une analyse. Ces tables sont aussi utilisées dans la fonction Interface de chiffrier, qui est décrite dans le document *Guide d'utilisation des outils*, qui prolonge l'utilité des tables intégrées en permettant une gamme plus large de mesures de comparaisons entre les traitements.

Toutes les tables intégrées ont un ensemble identique de lignes, dont la plupart correspondent aux variables définies dans le document *Guide des variables*. Cinq formats de colonnes différents sont disponibles, et correspondent aux numéros de table 0, 1, 2, 3 et 4. Chacune de ces tables a une table auxiliaire qui peut être sélectionnée séparément et qui contient le nombre sous-jacent d'observations non-zéro pour chaque cellule de la table originale. Ces tables auxiliaires (numérotées 0A, 1A, 2A, 3A et 4A) peuvent être utilisés pour le calcul des moyennes ou des proportions de participation. Le tableau suivant décrit les caractéristiques de chacun des cinq types de tableau.

Table	Dimension de la colonne	Paramètres de commande
0	Total du Canada (une colonne)	T0FLAG, T0AFLAG, pour activer.
1	Provinces, total	T1FLAG, T1AFLAG, pour activer.
2	Groupes de revenus, total	T2FLAG, T2AFLAG, pour activer INCVAR, pour spécifier la variable retenue INCGP, pour spécifier les groupes de revenus.
3	Type de famille (6), total	T3FLAG, T3AFLAG, pour activer.
4	Groupes de rapports de faible revenu	PVRAT, pour spécifier les groupes de rapport de pauvreté INCVAR, pour spécifier la variable de revenu. PTF (paramètre de structure) pour spécifier le seuil de pauvreté par grosseur de famille et urbanisation.

Il y a trois autres paramètres qui influent sur toutes les tables intégrées activées.

- OUTTBL contient le nom du fichier qui contiendra les tables générées. Si OUTTBL n'est pas spécifié, un nom de fichier basé sur le nom du fichier des paramètres de commande, mais ayant l'extension .tbl, sera utilisé.
- TABUNIT définit le niveau d'analyse de famille qui doit être utilisé pour toutes les tables intégrées activées, comme l'indique le tableau ci-dessous. Si l'on spécifie une valeur zéro pour TABUNIT (niveau de l'individu) pour les tables 3 ou 4, les individus sont classés selon le type de la famille nucléaire dont ils font partie.

TABUNIT	Niveau d'analyse des tables intégrées
0	Individu
1	Famille nucléaire
2	Famille de recensement
3	Famille économique
4	Ménages

- TABDELTA définit le seuil de revenu consommable de l'unité utilisé pour définir les lignes de gagnants ou de perdants des tables.

Tableau 5 : Interprétation des valeurs de TABUNIT

Le graphique suivant utilise la table intégrée 0 (activée en déclenchant T0FLAG) pour documenter la signification de chacun des lignes des tables intégrées. On trouvera dans le document *Guide des variables* une description de la signification de chacune des variables.

Table 0 : Résultats pour les familles de recensement

Variable (x1,000,000)	TOTAL	
Unités familiales (x1000)	11921.4	units
Personnes (x1000)	27433.5	persons
Enregistrements EDM	45371.6	scfrecs
Enregistrements BDSPS	79764.0	spsdrecs
Revenu modélisé (base)	359305.8	_immicons
Revenu modélisé (variante)	359305.8	immicons
Changement	0.0	(immicons-_immicons)
Nombre de gagnants (x1000)	0.0	
Nombre de perdants (x1000)	0.0	
Aucun changement (x1000)	11921.4	
Gains du gagnant	0.0	
Perte du perdant	0.0	
Revenu total	520266.3	immtot
Revenu gagné	437703.4	immmkt
Revenu d'emploi	341723.5	imiemp
Revenu d'un emploi indépendant	27159.4	imiself
Revenu de placements	41430.8	imminv
Autre revenu	27395.1	immoth
Revenu de transferts	82554.7	immtran
Impôt total	160955.0	immtax
Transferts nets	-78401.8	imnettr
Revenu disponible	408230.1	immdisp
Revenu consommable	359305.8	immicons
Taxes et impôt fédéraux	100153.4	imftax
Impôt fédéral (revenu)	59603.5	imtxf
Cotisation à l'ass.-ch.	8173.4	imuic
Recouvrement (d'ass.-ch.)	131.7	imuiibr
Cotisations au RPC/RRQ	6068.5	imcqppc
Autres recouvrements	644.4	imothrep
Taxes à la consommation	25532.0	imtxfc
Transferts fédéraux	72064.5	imftran
Allocations familiales	2855.3	imffa
Crédit d'impôt (enfants)	2287.3	imctcben
SV	14354.2	imioas
SRG	4318.9	imigis
SPA	565.3	imispa
Prestations d'ass.-ch.	18789.6	imiuib
RAPC (fédérale)	5014.2	imfsa
Revenu - RPC/RRQ	16493.3	imicqp
Autres transferts	4697.3	imfothtr
Crédit d'impôt (ventes)	2688.0	imfstc
Abattement rem. (Québec)	0.7	imqtar

Solde fédéral net	28089.5	imfedbal
Taxes/impôts provinciaux	60802.9	imptax
Impôt provincial	37415.1	imtxp
Taxe de vente prov.	23388.2	imtxpc
Transferts provinciaux	10490.9	imptran
Allocations familiales	519.8	impfp
Prestations âgées	611.7	imigist
RAPC (Provinciale)	7853.3	impsa
CR&D. d'impôt & autres tr.	1506.2	impalltc
Solde provincial net	50312.3	imprvbal
+-----+-----+		
-		

Fonction Tableaux croisés

Le concept du MSPS reconnaît que les tables intégrées ne combleront pas toujours les besoins des analystes. Par conséquent, il fournit une puissante fonction qui permet à un utilisateur de demander les tables qu'il désire en particulier. Un guide auxiliaire distinct, *Guide d'utilisation des tableaux croisés*, donne une documentation détaillée sur la fonction Tableaux croisés sous une forme complète qui convient aussi à ceux qui font l'apprentissage de la fonction Tableaux croisés. Cette section, par contre, donne l'information essentielle de façon plus concise.

SPÉCIFICATION DES TOTALISATIONS DÉFINIES PAR L'UTILISATEUR

Les totalisations définies par l'utilisateur sont activées en déclenchant XTFLAG. L'analyste spécifie les totalisations qu'il désire par l'élément XTSPEC du fichier des paramètres de commande. XTSPEC est l'un des paramètres de chaîne du MSPS. Le contenu du paramètre consiste en demandes de totalisation séparées par des points-virgule. L'exemple suivant contient deux demandes de totalisation et générerait deux tables.

```
XTSPEC
EF: hdprov * {units};
CF: cftype+*{immtot/units};
```

La première table afficherait le nombre de familles économiques par province. La seconde ferait la totalisation du revenu total moyen par famille de recensement pour plusieurs catégories différentes de familles de recensement et, aussi, pour toutes les familles de recensement prises ensemble.

La dernière ligne de la demande accepte maintenant le point-virgule. Il est fortement recommandé d'utiliser le point-virgule.

ÉLÉMENTS D'UNE DEMANDE DE TOTALISATION

Une demande de totalisations comprend de multiples éléments. Elle commence habituellement avec la spécification de l'unité d'analyse pertinente, c'est-à-dire individu (IN:), famille nucléaire (NF:), famille de recensement (CF:), famille économique (EF:) ou ménage (HH:). Si la demande de totalisation ne comprend pas de spécification d'unité, l'unité par défaut individu "IN:" s'applique. Le reste de la demande de totalisation comprend au moins un "niveau" et il y a un astérisque qui sépare chaque niveau d'un autre. Un des niveaux, le niveau de totalisation,

indique l'élément ou les éléments qui doivent être totalisés. Tous les niveaux qui restent sont des niveaux de classification. Ils précisent le genre de catégorisation que l'on désire en particulier dans la table. Par conséquent, dans la première des deux demandes de totalisation indiquées ci-dessous, le niveau d'analyse est la famille économique (EF:). En outre, la quantité qui est totalisée est le nombre de ces familles (unités) et l'analyste désire voir ces comptes de familles ventilés par province (hdprov).

LE NIVEAU DE TOTALISATION D'UNE DEMANDE DE TOTALISATION

Dans toute demande de totalisation donnée, le MSPS détermine le niveau de totalisation unique en plaçant des accolades, { et }, au début et à la fin. Un niveau de totalisation peut préciser des éléments multiples qui seraient totalisés dans la même table. S'il y a plusieurs éléments dans le niveau de totalisation, ils doivent être les séparés par des virgules. La spécification suivante dans XTSPEC ferait la totalisation du revenu total

```
CF:{immtot,  
    units,  
    immtot/units}  
*hdtenur+;
```

qui revient aux familles de recensement, le nombre de ces familles ainsi que le revenu moyen par famille de recensement, tous ventilés par catégorie de logement habité (loyer, propriété avec hypothèque, etc.). La demande peut aussi être inscrite sur deux lignes, comme ci-dessous, mais nous suggérons fortement d'utiliser l'autre forme pour facilement suivre les variables et leur format de sortie.

```
CF:{immtot,units,immtot/units}*hdtenur+;
```

ÉLÉMENTS PROPRES À LA TOTALISATION

Naturellement, il y a certaines restrictions sur ce que la fonction Tableaux croisés peut totaliser, mais la manière dont le MSPS a été conçu le rend passablement souple. Pour les débutants, l'analyste peut faire la totalisation de toute variable d'analyse disponible dans la BDSPP ou calculée dans le modèle (MSPS). Des listes complètes de ces variables se trouvent dans le document *Guide des variables*. En outre, comme l'illustrent les exemples précédents, l'analyste peut définir "à la volée" les expressions de totalisation qu'il désire, les construire à partir de tout type de variable antérieure et des opérateurs mathématiques appropriés. La section Expressions du présent guide étudie en profondeur l'utilisation des expressions par le MSPS.

LES EXPRESSIONS IMPLIQUENT DES FONCTIONS DE VARIABLES AGRÉGÉES

Lorsque la fonction Tableaux croisés fait la totalisation correspond à une expression, elle commence par créer une valeur de "somme" agrégée pour chaque variable d'expression. Ce n'est qu'à ce moment qu'elle exécute, en utilisant les valeurs de somme, les opérations (additions, multiplications, divisions, etc.) de l'expression. C'est en général ce que l'analyste a précisément l'intention de faire. C'est-à-dire qu'un analyste qui fait la totalisation {immtot/persons} désire diviser une somme de revenus par le compte pertinent de personnes du total. Le résultat est exactement ce que donnerait la fonction Tableaux croisés.

De même, l'analyste qui fait la tabulation {immtot/persons} ne veut généralement pas faire la somme, entre les unités, des revenus par habitant de ces unités, le résultat qui serait obtenu si la fonction Tableaux croisés avait évalué l'expression pour chaque enregistrement plutôt que

d'utiliser les valeurs agrégées. Si l'analyste a vraiment besoin de calculs à l'intérieur des enregistrements avant la totalisation, il utilisera une méthode comportant deux étapes. En premier lieu, une variable utilisateur, définie au niveau d'analyse approprié, exécutera les calculs à l'intérieur des enregistrements. En second lieu, la variable utilisateur se retrouvera dans la demande de totalisation.

IMPOSITION DE QUALIFICATIFS AUX ÉLÉMENTS OBJET D'UNE TOTALISATION

En l'absence d'instructions explicites de l'analyste, la fonction Tableaux croisés fait des choix éclairés quant aux étiquettes, aux ajustements, au nombre de décimales, etc. Cependant, elle fournit un mécanisme qui permet à l'analyste d'exercer, lorsqu'il le désire, un contrôle explicite sur ces dispositifs. Plus particulièrement, l'analyste peut imposer un ou plusieurs qualificatifs sur toute variable ou expression qui fait l'objet de la totalisation. Les quatre qualificatifs disponibles sont S, P, L et M et ils sont étudiés à tour de rôle ci-dessous.

La qualificatif S détermine l'ajustement de la variable soumis à la totalisation. Par conséquent, un qualificatif S=6 donnerait une entrée de table en millions. La fonction Tableaux croisés indiquerait cette sélection en incluant la chaîne "(M)" dans l'étiquette de la variable. L'analyste devrait restreindre la valeur de S à la plage [-6, 9].

Le qualificatif P termine le nombre de décimales affichées pour une entrée de la table. Ainsi, un qualificatif P=2 donnerait les entrées les tables qui comprendraient la virgule suivie de deux décimales. Un qualificatif P=0 donnerait des entrées de table sans décimale. L'analyste devrait restreindre la valeur de P à la plage [0, 8].

Le qualificatif L permet à l'analyste de ne pas tenir compte de l'attribution d'étiquettes par défaut aux entrées de table. L'attribution d'étiquette par défaut est déjà passablement perfectionnée -- la fonction Tableaux croisés utilise une étiquette de variable quand elle est disponible; quand il n'y a pas d'étiquette, la fonction Tableaux croisés établit l'étiquette à partir du nom de la variable visée par la totalisation ou du texte de l'expression de totalisation. Le texte d'étiquette fourni par la qualification L remplace l'étiquette par défaut de la fonction Tableaux croisés. Donc, par exemple, un analyste qui ferait la totalisation {scfrecs } pourrait spécifier une étiquette "Enregistrements - EFM" pour faciliter la consultation. Tout comme l'utilisateur s'y attendrait, la fonction Tableaux croisée ajoute à l'étiquette tous les "suffixes" appropriés, par exemple "(%)" pour les entrées ajustées à -2.

Le qualificatif M permet à l'analyste de demander une normalisation le long de l'une des marges de la table. L'application la plus courante de ce qualificatif est le calcul de la distribution en pourcentage; en effet, l'utilisation de M établit à -2 le facteur d'ajustement par défaut. La demande de totalisation suivante en est un exemple.

```
CF: {units:M=hdprov P=1} * cftype+ * hdprov+;
```

La demande fournirait une table donnant la distribution en pourcentage, entre les provinces, des nombres de familles de recensement ventilés par type de famille de recensement. Le qualificatif M=hdprov indique que la normalisation se ferait sur les dimensions de province. Le qualificatif P=1 fait en sorte que les entrées de la table donne des résultats au dixième de 1 %.

SYNTAXE DU QUALIFICATIF

L'analyste indique la présence des qualificatifs par un caractère deux-points après l'expression ou la variable visée par la totalisation. Les qualificatifs suivent les deux-points et précèdent la variable de totalisation suivante ou l'accolade à la fin du niveau de totalisation. Un qualificatif prend la forme de Q=C, où Q représente le caractère du qualificatif (S, P, L ou M) et C représente le contenu du qualificatif (un nombre, une étiquette ou le nom d'une variable, selon le type de qualificatif utilisé). S'il y a plusieurs qualificatifs, ils sont séparés l'un de l'autre par des espaces. À l'intérieur d'un qualificatif, cependant, il ne devrait y avoir aucun espace d'un côté ou de l'autre du caractère "égale" (=). Il faut noter que chacune des expressions ou des variables de totalisation du niveau de totalisation d'une demande de table peut avoir son propre ensemble de qualificatifs. La demande de totalisation

```
EF: {units: S=6 P=1,
      units :L="Distr. des FE" M=hdprov }
      *hdprov+ * hdenur+;
```

illustre quelques-unes de ces possibilités. La table qui résultera devrait afficher le nombre de familles économiques par province et type de logement, ainsi que leur distribution entre les provinces.

NIVEAUX DE CLASSIFICATION

Une demande de totalisation comprend habituellement au moins un niveau de classification. Un niveau de classification consiste dans le nom d'une variable de classe, c'est-à-dire une variable "entière" qui classe l'unité pertinente en catégories complètes et s'excluant mutuellement. Ces niveaux sont séparés du niveau de totalisation ainsi que les uns des autres par des astérisques. En pratique, un niveau de classification précise une dimension dans laquelle la fonction Tableaux croisés doit ventiler les éléments au niveau de totalisation. Par conséquent, dans la demande de totalisation donnée en exemple :

```
IN: {scfrecs: L="Enregistrements" P=0 S=0} * idsex * idmarst+;
```

l'analyste compte les nombres d'enregistrements EFM pour des individus selon le sexe et l'état matrimonial.

La variable de classe provient en général de deux sources : les variables de classe de base de données ou les variables de classe modélisées. Les utilisateurs peuvent aussi définir leurs propres variables de classe avec la fonction UVAR.

CRÉATION D'UNE CATÉGORIE "TOUTES" AVEC LE SUFFIXE "+"

Utilisée seule comme niveau de classification, une variable de classe génère dans la table un ensemble de catégories mutuellement exclusives. Cependant, l'analyste désirera souvent avoir des entrées "Somme toutes catégories" supplémentaires dans la table. Le suffixe + donne cette capacité. Lorsque l'analyste ajoute le suffixe + au nom de la variable d'un niveau de classification, la fonction Tableaux croisés génère une nouvelle catégorie "Toutes" égale à la somme pour toutes les catégories de la variable. L'analyste peut utiliser le suffixe + indépendamment entre les niveaux de classification d'une demande de totalisation.

Par conséquent, notre exemple précédent

```
IN: {scfrecs: L="Enregistrements" P=0 S=0} * idsex * idmarst+;
```

créera une catégorie “Toutes” pour la dimension État matrimonial, mais non pour la dimension Sexe.

ORDRE DES NIVEAUX ET ASPECT DE LA TABLE

La commande que donne l’analyste relativement aux niveaux de totalisation et de classification dans sa demande de totalisation détermine l’aspect de la table produite. En général, plus le niveau est loin à droite, plus le cycle des catégories est répété. Les éléments de totalisation multiples d’un niveau de totalisation sont effectivement traités comme des catégories. Par conséquent, le niveau le plus à droite d’une demande détermine les catégories de colonnes de la table. L’avant-dernier niveau détermine les catégories de lignes, le niveau qui précède l’avant-dernier niveau détermine les catégories de segments, etc. Cette règle donne à l’analyste un bon contrôle sur l’aspect de la table produite.

Pour poursuivre avec l’exemple ci-dessus,

```
IN: {scfrecs: L="Enregistrements" P=0 S=0} * idsex * idmarst+
```

nous voyons que la table produite donnera les catégories d’état matrimonial dans les colonnes et le sexe comme variables de commande pour les segments.

Il y a une exception à la règle générale lorsqu’il y a seulement un élément dans le niveau de totalisation. Pour ce cas en particulier, la fonction Tableaux croisés traite la demande de totalisation comme si le niveau de totalisation était le premier de la demande. Ceci empêche la production de tables à segments multiples qui n’ont qu’une colonne ou une ligne dans les segments.

SPÉCIFICATION D’UNE TABLE SUR PLUSIEURS LIGNES

Le paramètre XTSPEC est l’un des paramètres de chaîne d’un fichier des paramètres de commande du MSPS. Le MSPS s’attend à ce qu’une demande se termine par un point-virgule. La demande peut s’étendre sur autant de lignes qu’il en faut pour rendre la demande facilement lisible. Il importe donc de mettre un retour de chariot à la fin de chaque ligne de la demande. L’exemple suivant donne une demande de sortie à variables multiples.

```
IN: imfiler+ *  
  {imitot:S=0,  
   imdedfn:S=0,  
   imdedft:S=0,  
   imitax:S=0,  
   imbft:S=0,  
   imfsur:S=0,  
   imtxf:S=0,  
   imqtaa:S=0,  
   imtxp:S=0,  
   units:S=0} * hdprov+;
```

PARAMÈTRES DE COMMANDE CONNEXES

De tous les paramètres du MSPS qui déterminent la production des tables, XTSPEC est la locomotive. Cependant, trois autres paramètres obligatoires ont un effet direct sur la production des tables.

Le paramètre XTFLAG est un drapeau qui indique au MSPS s'il faut produire ou non les tables spécifiées par XTSPEC. Les tables sont produites seulement si XTFLAG est déclenché.

Le paramètre XTLINES indique au MSPS combien de lignes entrent sur une page. Il aide la fonction Tableaux croisés à éviter tant la répartition d'un segment donné sur des pages multiples, que le gaspillage de papier, en commençant chaque nouvelle table ou chaque nouveau segment sur une nouvelle page. XTLINES doit avoir une valeur qui se situe entre 0 et la 32767. Une valeur typique est de 66 lignes, ce qui permet de loger la sortie sur une page standard de 11 pouces à raison de six lignes au pouce.

Le paramètre XTCOLS détermine le compromis entre le fait d'avoir plusieurs lignes dans les en-têtes de colonne et le fait d'avoir des tables plus larges. XTCOLS doit accepter une valeur se situant entre 80 et 32767. L'interprétation générale de la valeur du paramètre est celle d'un nombre maximal souhaité de colonnes d'impression dans le fichier de sortie. Cependant, l'algorithme particulier par lequel XTCOLS influe sur la sortie imprimée est passablement complexe et ne peut être résumé avec précision. Une valeur typique de XTCOLS est 132; cette valeur correspond au nombre maximal de positions d'impression disponibles sur de nombreuses imprimantes.

QUESTIONS DE NIVEAUX D'ANALYSE

La BD/MSPS et la fonction Tableaux croisés fonctionnent dans un contexte hiérarchique. Bien qu'il y ait cinq niveaux d'unités d'analyse acceptés (IN, NF, CF, EF et HH), les données sous-jacentes n'existent qu'aux niveaux des ménages et des individus. Les variables des niveaux intermédiaires doivent donc être définies par propagation des variables des ménages vers le bas ou par cumul des variables définies au niveau de l'individu. Les variables de décimales sont plutôt simples à cet égard étant donné qu'elles peuvent être agrégées vers le haut depuis le niveau des individus. Par exemple, le revenu disponible d'une famille économique est simplement la somme des revenus disponibles de tous ses membres. Cependant, l'utilisation des variables de dépenses définies au niveau du ménage dans une table des familles nucléaires exige que l'on exerce une certaine prudence quand on en fait l'interprétation.

Les variables de classe, primordiales pour les tableaux croisés, posent encore plus de défis. Bien que la propagation vers le bas, par exemple la province de résidence pour une famille de recensement, ne présente aucun problème, les opérations de cumul peuvent être passablement compliquées. Par exemple, idsex est, mécaniquement, acceptable comme variable de classe pour une table définie au niveau de famille de recensement. Cependant, qu'est-ce qu'il représente? L'analyste doit bien comprendre que la famille de recensement sera classée selon le sexe d'un individu membre représentatif. Par conséquent, il devrait connaître les règles qui déterminent l'identification de ces individus par le MSPS.

En somme, même si le MSPS prévoit des valeurs par défaut raisonnables pour les choses typiques qu'un analyste peut faire, le cumul est une question qui peut se révéler complexe. La

section Niveaux de famille du Guide d'utilisation présente une étude exhaustive des questions de cumul dans la BD/MSPS.

LES TOTALISATIONS SONT FONCTION DES PARAMÈTRES DE SÉLECTION

Étant donné que les entrées des tables dépendent des agrégats des variables ou des cas (pondérés ou non) pertinents dans la BDSPS, ils dépendent nécessairement des critères utilisés pour sélectionner les cas qui doivent faire l'objet d'un traitement. En particulier, les entrées d'une table, et l'interprétation de ces entrées, peuvent dépendre très fortement des critères de sélection utilisés pour une exécution du MSPS. Il y a trois paramètres de contrôle pertinents. (1) Le paramètre SELFLAG détermine s'il y a une quelconque sélection. (2) Le paramètre SELUNIT détermine le type d'unité qui sera sélectionné. (3) Le paramètre SELSPEC définit la condition que les cas doivent respecter pour être inclus dans l'analyse; il offre la possibilité de spécifier des critères qui dépendent simultanément de plusieurs niveaux d'unité d'analyse.

Des problèmes d'interprétation risquent en particulier de se produire lorsque le niveau de SELUNIT réfère à un sous-ensemble du niveau associé à la demande de totalisation. Les difficultés surviennent parce que l'on peut n'avoir inclus que des sous-ensembles d'une famille dans les tables produites. La section Fonction Sélection du Guide d'utilisation donne une étude plus complète des critères de sélection et de leurs répercussions sur les sorties du MSPS.

Fonction Analyse de la distribution

La Fonction Analyse de la distribution du MSPS permet à l'analyste d'évaluer visuellement les propriétés statistiques d'une variable de la BD/MSPS. Un histogramme, utilisant des points de découpage en centiles fournis par l'analyste et incorporant la troncature des extrémités si désiré, est généré à partir d'un échantillon d'observations de la variable. Parce que les observations sont sujettes à la sélection (voir la section Fonction Sélection) et que la variable peut être définie par la fonction Variable définie par l'utilisateur (voir la section Fonction Variables définie par l'utilisateur), la fonction d'analyse de la distribution peut être un outil puissant d'exploration. On peut trouver sous l'onglet **Tableaux** de MSPS Visuel la sortie de la fonction Analyse de la distribution. De plus, toute sortie est transférée au fichier .TBL pour l'exécution du programme.

Une fonction secondaire indépendante consiste à enregistrer les valeurs extrêmes de la variable demandée pour toutes les observations et à identifier les ménages où surviennent ces valeurs extrêmes. En utilisant les numéros de ménages qui sont produits, la fonction Sélection, et la fonction Sortie des cas (voir la section Fonction Sortie texte), l'analyste peut explorer les propriétés de ces ménages et comprendre ce qui, dans leur cas, peut entraîner une valeur extrême pour la variable en question.

La fonction Analyse de la distribution est activée quand le paramètre DISTFLAG est déclenché. La variable à être analysée est donnée dans le paramètre de chaîne DISTVAR et les observations sont faites dans le niveau d'analyse DISTUNIT, comme l'indique le tableau suivant.

DISTUNIT	Niveau de famille de l'analyse de distribution
0	Individu
1	Famille nucléaire
2	Famille de recensement
3	Famille économique
4	Ménage

Tableau 6 : Interprétation des valeurs de DISTUNIT

Pour générer l'histogramme, jusqu'à DISTSAMP observations seront faites et les observations avec une valeur zéro sont comprises si le drapeau DISTZERO est déclenché. Ces observations et les poids de ménage associés sont conservés en mémoire et triés après que le MSPS a terminé le traitement de tous les ménages. L'augmentation de DISTSAMP accroît donc les besoins en mémoire.

Le paramètre vectoriel DISTP contient les points de découpage en centiles à utiliser pour la génération de l'histogramme. Le premier et le dernier élément de DISTP sont interprétés comme des valeurs à utiliser pour tronquer la distribution au-dessous et au-dessus, respectivement. Si ces valeurs sont fixées à zéro et à 100, l'échantillon entier sera alors utilisé pour générer l'histogramme. Cependant, il est souvent utile de faire quelques troncations de valeurs extrêmes puisque l'inclusion de ces valeurs extrêmes réduit les détails qui peuvent être illustrés dans le gros des observations.

Il reste à décrire deux paramètres. DISTPWID est le nombre de positions d'impression utilisées pour produire l'histogramme. Une valeur de 79 convient pour l'affichage à l'écran de votre PC, mais une valeur plus grande peut être donnée si la sortie doit se faire sur imprimante pouvant sortir un document plus large. DISTPHGT donne le nombre de lignes utilisées pour produire l'histogramme. Une valeur de 20 convient pour l'affichage à l'écran.

L'exemple suivant illustre les points ci-dessus. La fonction Analyse de la distribution est utilisée, au niveau de la famille de recensement, pour l'analyse des résultats. Voici les paramètres de commande pertinents :

```

DISTFLAG          1
DISTUNIT          2
DISTVAR           ex0
DISTSAMP          4000
DISTZERO          1
DISTP             13
                  1
                  5
                  10
                  20
                  30
                  40
                  50
                  60
                  70
                  80
                  90
                  95
                  99

```


DISTPWID 70
DISTPHGT 17

La sortie produite par la fonction d'analyse de la distribution est la suivante :

Rapport de répartition : Taux d'imposition moyen pour les familles de recensement

Nombre total d'observations = 3278
Observations nulles = 798
Les statistiques suivantes sont basées sur toutes les 3278 observations.

Statistiques descriptives :

Somme des poids = 10053381
Somme pondérée = 1088932.76
Somme des carrés pondérée = 202166.59
Moyenne pondérée = 0.11

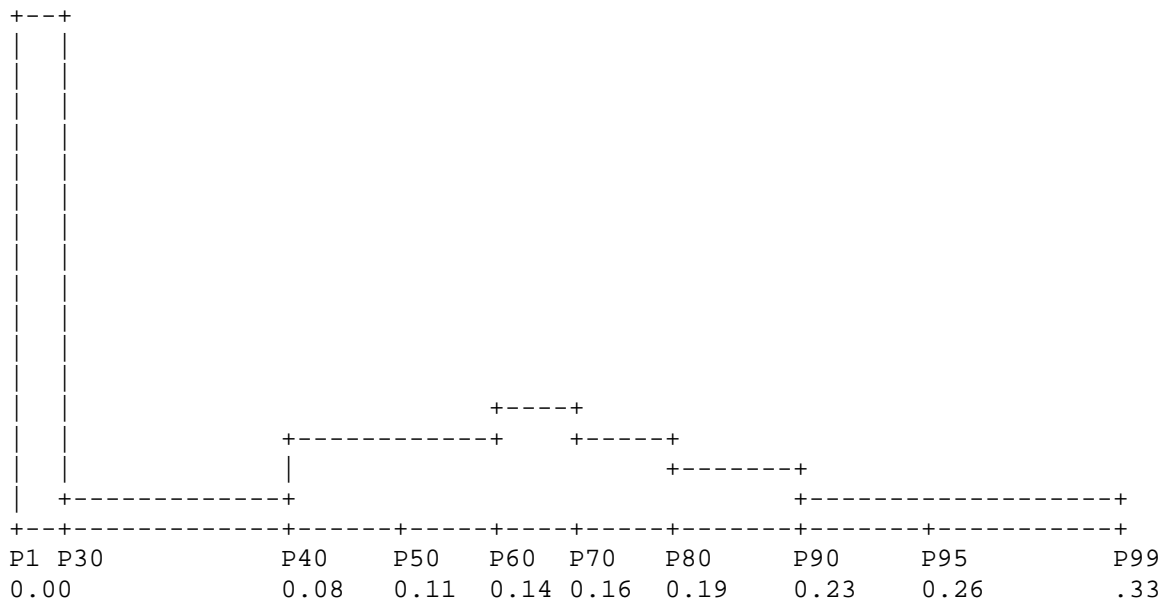
Valeurs extrêmes (avec les numéros de ménage associés):

Minimum	hdseqhh	Maximum	hdseqhh
-0.00	2111	0.57	2354
0.00	1	0.57	1853
0.00	2	0.57	2113
0.00	3	0.57	2114
0.00	4	0.57	1034

Quantiles choisis:

Q1 =0.00	P1=0.00	P90=0.23	P20=0.00	P60=0.14
Med=0.11	P2=0.00	P95=0.26	P30=0.01	P70=0.16
Q3 =0.18	P5=0.00	P98=0.30	P40=0.08	P80=0.19
	P10=0.00	P99=0.33		

Tracé d'histogramme:



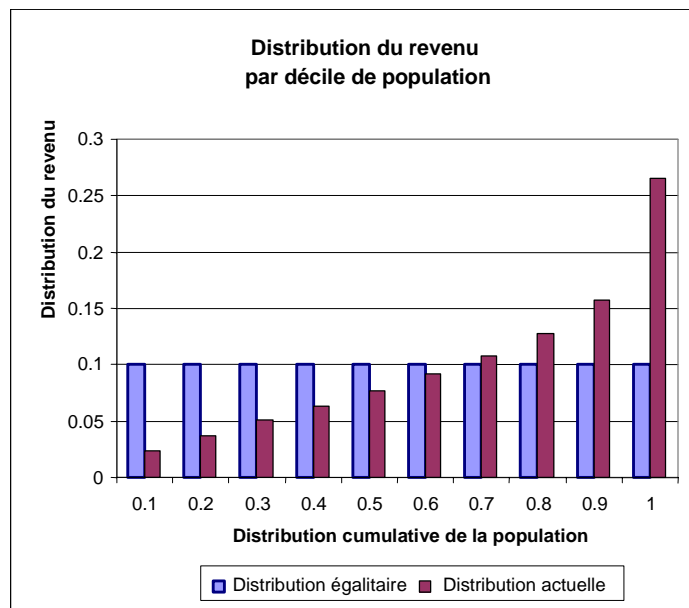
La première partie du rapport est basée sur tous les enregistrements (selon la sélection, si SELFLAG a été fixé à 1). Elle comprend une section de statistiques descriptives sur la variable demandée, elle est suivie d'une section qui donne les cinq observations les plus petites et les cinq observations les plus grandes, avec les numéros (variable hdseqhh) des ménages dans lesquels les valeurs extrêmes ont été observées.

La seconde partie du rapport est basée sur les premières observations DISTSAMP. Elle contient une estimation des quantiles choisis de la distribution et l'histogramme. Quand on interprète l'histogramme, il faut noter que la surface correspond au nombre d'observations (pondéré). Les centiles demandés dans DISTP sont imprimés immédiatement sous l'histogramme, tandis que la ligne suivante donne les valeurs de la variable correspondant aux centiles. S'il n'y a pas suffisamment d'espace, certaines des étiquettes sont supprimées.

MESURES DE L'INÉGALITÉ ET POLARISATION DE LA RÉPARTITION DU REVENU

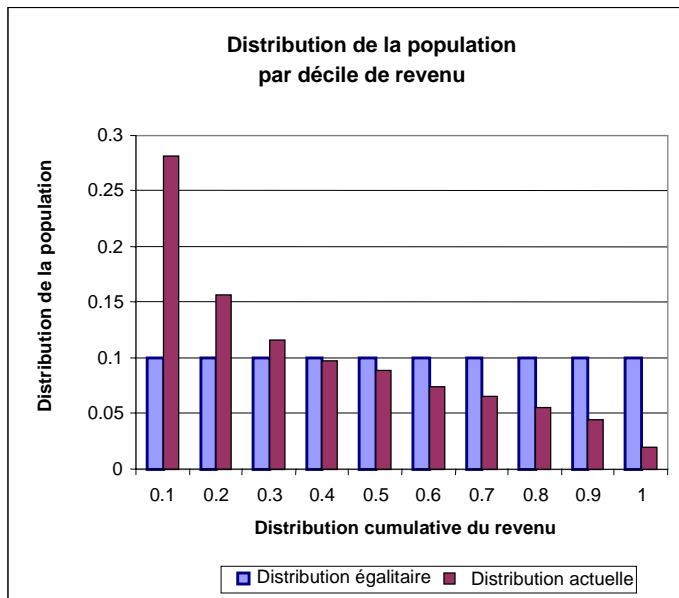
Introduction aux mesures de l'inégalité

Dans de nombreuses études, les mesures d'inégalité sont plutôt décrites comme part du revenu total par décile de la population, comme l'illustre la figure suivante. Nous observons que 10 % des gens les plus pauvres détiennent seulement 2 % du revenu total, tandis que 10 % des gens les plus riches détiennent 26 % du revenu total. Avec une répartition égale, chaque tranche de 10 % de la population totale détient 10 % du revenu total. Puisqu'il est facile à obtenir, certaines études utilisent le ratio des déciles (9^{ème} décile/1^{er} décile) comme mesure scalaire des inégalités.



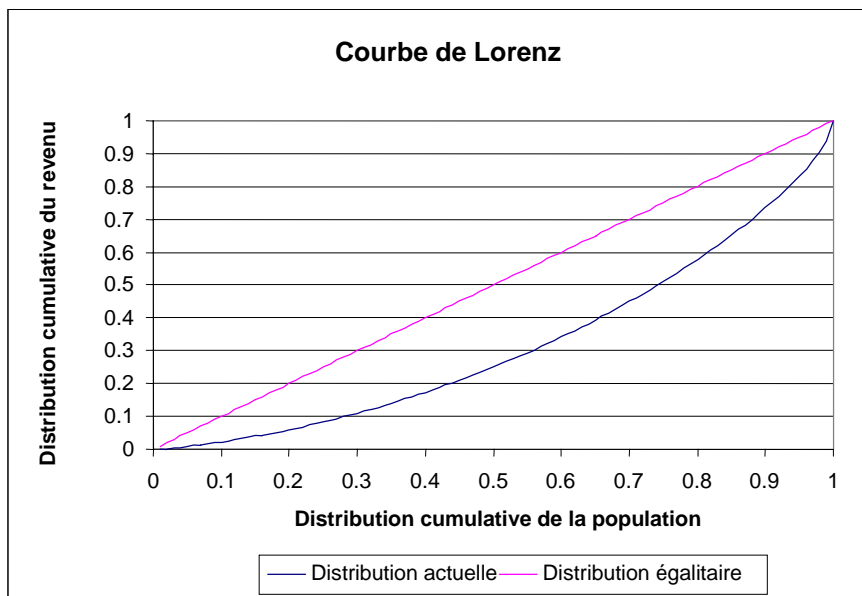
Une autre méthode consiste à considérer la partie de la population détenant une partie donnée du revenu comme l'illustre la figure suivante. Nous observons alors que la première tranche de 10 % du revenu total est partagée par 27,5 % de la population totale, tandis que la dernière tranche de 10 % est partagée par 2 % de la population totale. Cette méthode est utilisée par de nombreuses mesures de polarisation de la répartition du revenu. Ainsi, Thurow, dans la nouvelle vague de documents portant sur la disparition de la classe moyenne, utilise la proportion des salariés entre 75 % et 125 %, ou 75 % et 150 % du revenu médian; Blackburn et Bloom, pour leur part,

utilisent 60 % à 200 %, ou 60 % à 225 % du revenu médian.



La courbe de Lorenz permet de représenter la relation entre les parts du revenu détenues par une partie de la population. Pour produire la courbe de Lorenz, nous trions toutes les personnes par revenu, puis nous commençons à cumuler leur part du revenu total et leur part de la population totale.

Dans la figure suivante, la diagonale représente l'équirépartition où 0,1 de la population totale détient la même part du revenu total. La répartition réelle du revenu est la courbe de Lorenz où la première tranche de 0,3 (30% plus pauvre) de la population totale détient seulement 0,1(10%) du revenu total.



Les mesures d'inégalité et polarisation calculées dans MSPS sont basées sur les estimations de Lorenz.

Mesures des agrégats

Habituellement liée à la courbe de Lorenz, il y a la classe des mesures d'inégalité qui sont linéaires dans les revenus. Cette classe comprend les coefficients de GINI, de S-GINI et de GINI généralisé. Ces mesures peuvent aussi être décomposées entre des facteurs contributifs.

Toutes les mesures qui sont linéaires dans les revenus peuvent être interprétées comme une estimation pondérée de l'aire comprise entre la courbe d'équirépartition et la courbe de Lorenz. Dans le cas du coefficient de GINI [C. Gini, *Variabilità e Mutabilità*, Studi Economico-giuridici, Université de Cagliari, 1912], le poids est égal à 2. Cela soulève une interprétation géométrique du coefficient de GINI : si A est l'aire comprise entre la courbe d'équirépartition et la courbe de Lorenz, et B est l'aire qu'il y a en dessous de la courbe de Lorenz, alors le coefficient de GINI est le rapport $A/(A+B)$. Le coefficient de GINI varie entre zéro et un. Plus le coefficient de GINI est près de zéro, plus la répartition du revenu est égale. Un coefficient de GINI égal à un signifie que seulement un individu détient tout le revenu de l'économie.

Une mesure de la polarisation du revenu dérivée du coefficient de GINI est plus synthétique, cette mesure étant l'indice de polarisation de Foster-Wolfson (FW). L'indice FW mesure l'aire de la surface comprise entre la tangente de la courbe au point milieu de la population (ligne orange) et la courbe de Lorenz (dans la figure suivante, il s'agit de l'aire en grisé) par rapport à l'aire de la surface comprise entre la courbe d'équirépartition et la tangente. L'indice FW est une transformation du coefficient de GINI comme l'indique la relation suivante :

$$FW = 2 (2T - GINI) / mtan,$$

où

mtan est la pente de la tangente au 50^e centile de la population ou

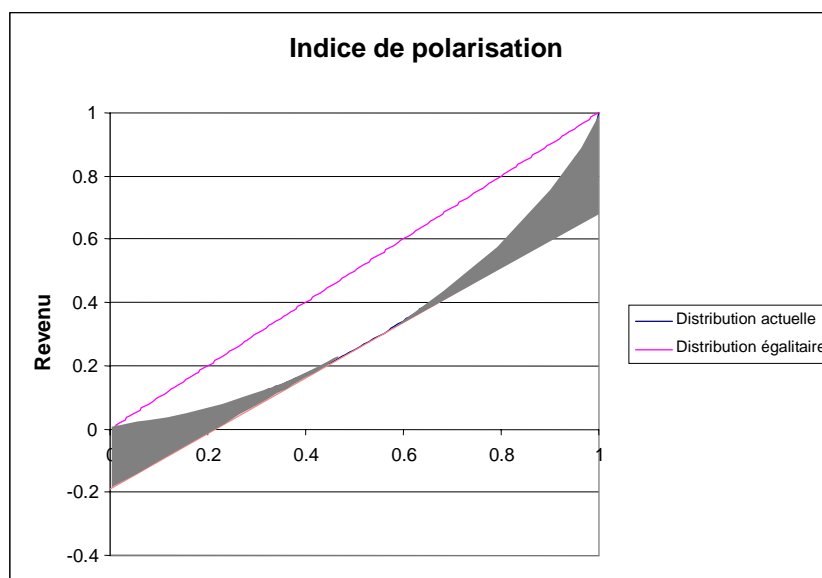
mtan = revenu médian/revenu moyen,

et

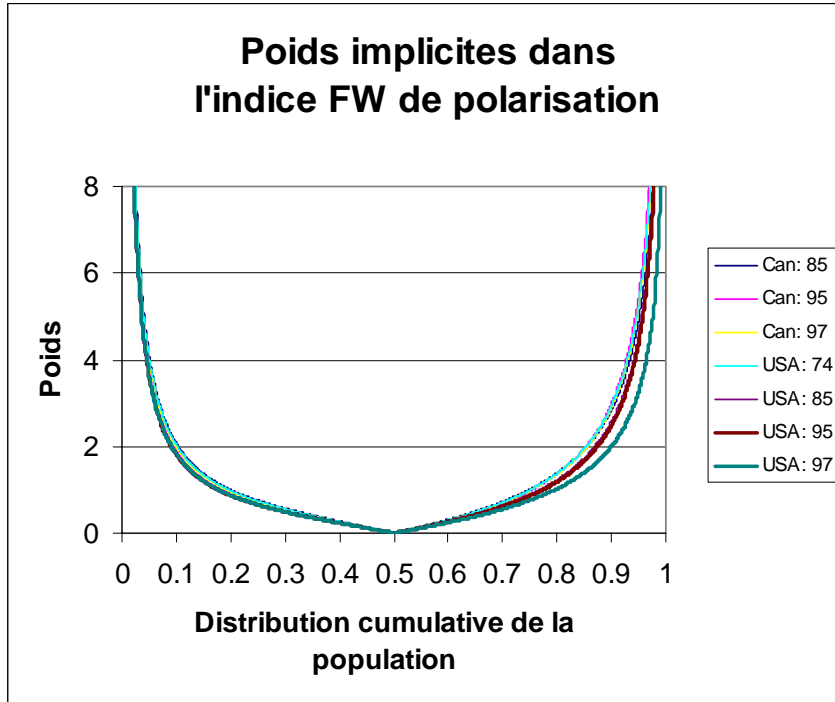
T est l'aire qu'il y a entre la courbe d'équirépartition et la tangente.

$T = 0,5 - \text{part du revenu au } 50^{\text{e}} \text{ centile de la population, } (L(0,5)),$

$T = 0,5 - L(0,5).$



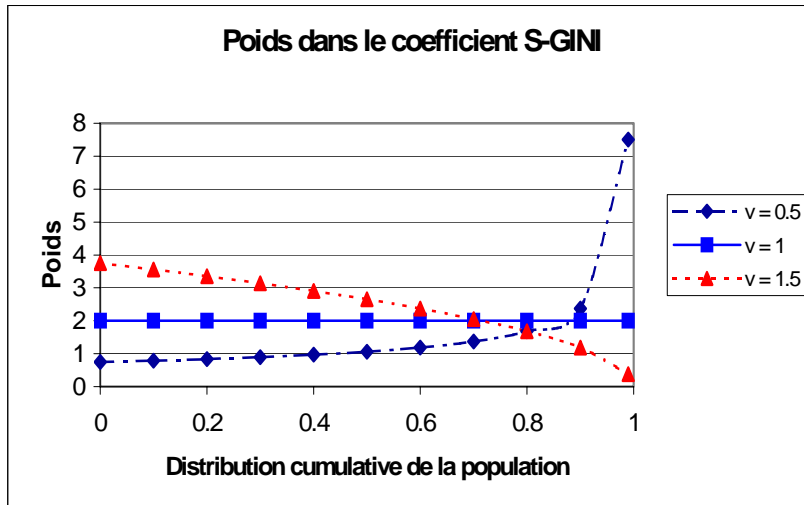
Cette mesure fait la pondération implicite de la distance entre la courbe d'équirépartition et la courbe de Lorenz avec un poids égal à zéro au point médian et augmente progressivement le poids à mesure qu'on se déplace vers les deux extrémités de la répartition du revenu, comme un bol ayant un fond relativement grand et presque plat. Cette mesure place implicitement sur le côté inférieur de la répartition du revenu plus de poids que sur le côté des revenus élevés parce que la distance entre la courbe d'équirépartition (ligne 11) est inférieure dans le premier centile à ce qu'elle est dans le dernier centile. La figure suivante illustre le genre de poids implicitement généré par l'indice de polarisation dans sept répartitions du revenu étudiées dans Wolfson, Michael et Brian Murphy, *New Views on Inequality Trends in Canada and the United States*, Monthly Review, avril 1998. La figure illustre que des poids implicites sont au minimum au point médian de la répartition de la population et augmentent lentement jusqu'à devenir des poids très grands aux deux extrémités de la distribution du revenu.



Le coefficient de S-GINI [Donaldson, D. et J. A. Weymark (1980), “*A Single Parameter Generalization of the GINI Indices of Inequality*”, *Journal of Economic Theory*, 22, pp 67-86, and Yitzhaki, S. (1983), “*On an extension of the GINI Index*”, *International Economic Review*, 24, pp.617-628] utilise un jeu de poids différent qui reflète les préférences de la société. Si v est la préférence de la société, p est la répartition cumulative de la population et $k(p)$ le poids associé à p , nous avons alors :

$$k(p) = v(v+1)(1-p)^{(v-1)}.$$

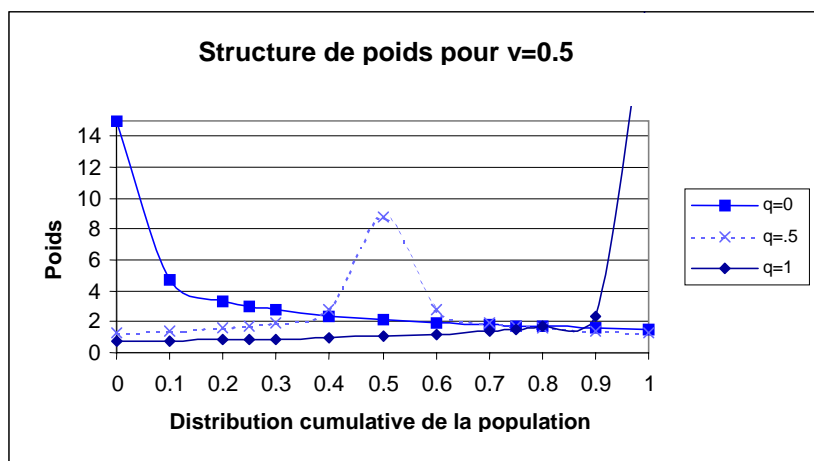
Si $v=1$, alors le poids associé à p est toujours 2 et le coefficient de S-GINI est le coefficient de GINI. Si $v<1$, il y a plus de poids imposé sur la partie supérieure de la répartition et, si $v>1$, il y a plus de poids imposé sur la partie inférieure de la répartition, comme l’illustre la figure suivante pour $v=0,5$, 1, et 1,5.



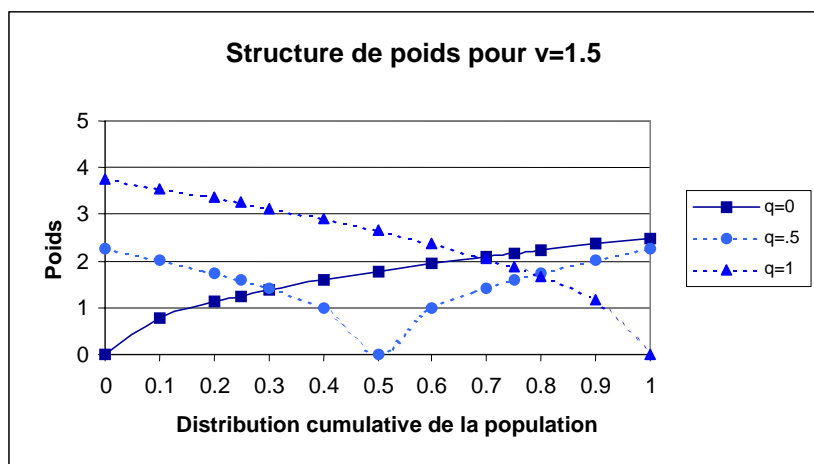
Plus récemment, Duclos, Jean-Yves et Martin, Tabi (1996), “*Linear Inequality Measures and the Redistribution of Income*”, cahier de recherche 96-08, CRÉFA, Université Laval, proposaient une classe de mesures de l’inégalité linéaire à deux paramètres, le coefficient de GINI généralisé. Si v est encore la préférence de la société et q produit des poids symétriques autour de sa valeur, pour $0 < q < 1$, alors les poids $k(p)$ sont :

$$k(p) = \frac{v(v+1)|q-p|^{(v-1)}}{q^{(v+1)} + (v+q)(1-q)^v},$$

L’aspect intéressant de cette mesure est sa capacité de reproduire l’indice de S-GINI lorsque q reçoit la valeur 1. Tous les poids sont symétriques autour de q . Si v est inférieur à 1, il y a plus de poids imposé au centre de la répartition des poids, q , comme l’illustre la figure suivante pour $v=0,5$. Il semble que $q=0, 0,5$, et 1 soient les mesures les plus intéressantes. Lorsque $q=0$, les poids les plus importants sont appliqués pour les revenus les plus bas et ensuite les poids diminuent; l’autre extrême est quand $q=1$, qui est le coefficient S-GINI. Une valeur intéressante est $q=0,5$, où les poids sont symétriques autour du point médian de la population. L’utilisation de ces valeurs de paramètres pour le coefficient de GINI généralisé (G-GINI) est plus intéressante que celle des statistiques descriptives parce qu’elle tient compte de la répartition du revenu au complet et non seulement d’un segment.



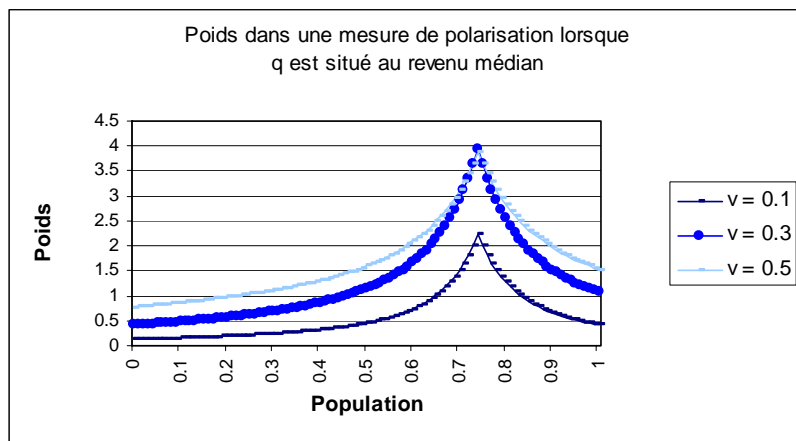
Lorsque $v > 1$, il y a moins de variation dans la structure des poids et les poids sont quand même symétriques autour de q . Dans ce cas, il y a plus de poids imposés aux deux extrémités de la répartition de la population et ils décroissent jusqu'à la valeur zéro lorsque $q = p$. Lorsque $q = 1$, l'indice de S-GINI, les poids sont plus importants à l'extrémité la plus basse de la répartition de la population. Lorsque $q=0$, il y a plus de poids imposés à l'extrémité supérieure de la répartition du revenu, mais les poids s'accroissent à un taux plus lent que lorsque $v < 1$. Lorsque $q = 0,5$, les poids décroissent jusqu'à zéro quand ils atteignent $p=0,5$. Il semble que, lorsque $v > 1$, les valeurs les plus intéressantes de q soient 1,5, une valeur fréquemment utilisée dans S-GINI et 0,5, à cause de la forme en V ou en U autour du point médian de la répartition de la population, ces formes sont intéressantes pour la mesure de la polarisation du revenu.



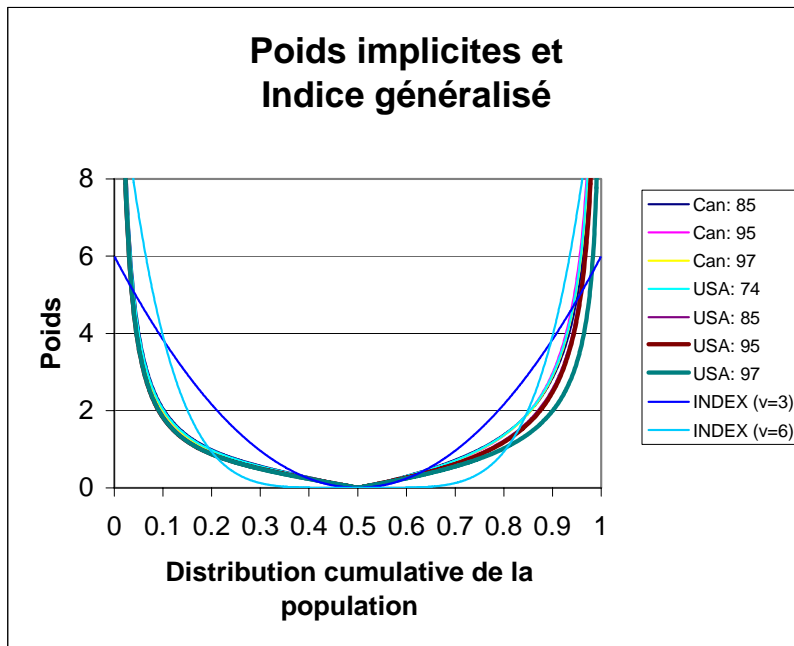
Mesure de la polarisation de la répartition du revenu

L'utilisation du coefficient de G-GINI comme mesure de rechange exige que q reçoive la valeur du revenu médian, $q=0,74$. Alors, l'attention se porte non pas sur la partie de la population comprise dans un intervalle arbitraire autour de la valeur du revenu médian, mais elle se porte plutôt sur la part du revenu détenue par la population entourant la valeur médiane de revenu. Les

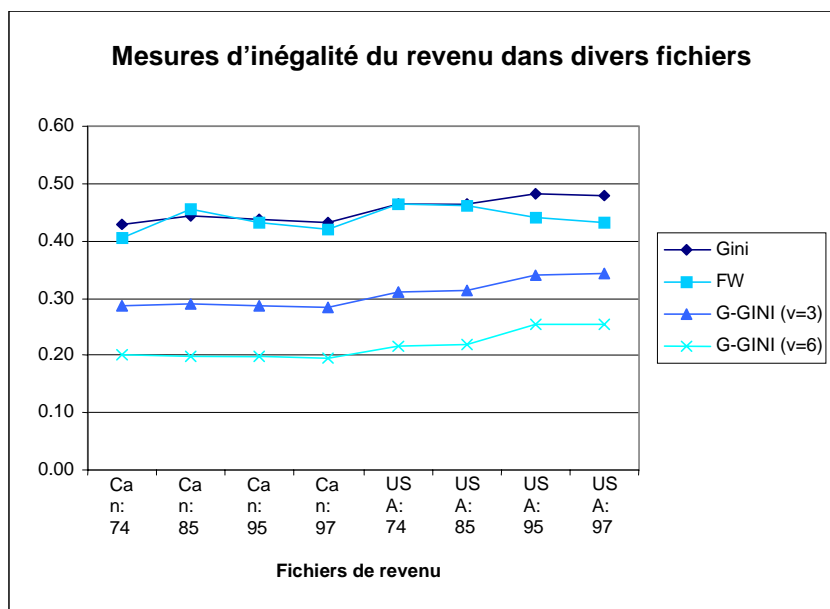
poids servant à l'évaluation avec le coefficient de G-GINI décroissent des deux côtés de q si $v < 0$. L'utilité du coefficient de G-GINI est qu'il tient compte de la répartition de l'ensemble du revenu et non seulement d'une petite partie de ce revenu. Parce que la part du revenu détenue par les déciles les plus élevés augmente, le choix de v est critique pour l'évaluation de la polarisation du revenu médian. La figure suivante représente les poids pour q à la valeur médiane du revenu et pour des valeurs de $v=0,05, 0,1$ et $0,3$.



La figure suivante illustre les poids implicites associés à l'indice de polarisation FW de répartition de sept salaires qu'étudient Wolfson, Michael et Brian Murphy, *New Views on Inequality Trends in Canada and the United States*, Monthly Review, avril 1998. Le coefficient de G-GINI ayant la valeur $v=3$ et le centre étant à la population médiane ($q=0,5$) peuvent reproduire les poids du centre entre 45 % et 55 % de la répartition de la population, mais le coefficient de G-GINI place plus de poids sur les populations entre 5 % et 45 %, et 55 % et 95 %, et moins de poids sur les tranches de 5 % des deux extrémités de la courbe de répartition du revenu. Si v reçoit la valeur 6, la plupart des poids se situant entre 35 % et 65% de la population sont près de zéro et les poids à l'extérieur de cette plage sont plus proches des poids implicites observés.



La figure suivante illustre la relation qu'il y a entre l'indice de polarisation FW et le coefficient de G-GINI lorsque v a la valeur 3 et la valeur 6. Il y a seulement un facteur d'ajustement entre les valeurs de G-GINI lorsque $v=3$ et $v=6$. Comme le coefficient de GINI décroît aux États-Unis entre 1995 et 1997, et que le FW décroît depuis 1985, les valeurs du coefficient de G-GINI augmentent toutes deux en 1995 et 1997. Cela indique que l'inégalité s'accroît dans ces années pour la population se situant à l'extérieur de l'intervalle médian de la population implicitement définie par v .



L'activation du calcul des mesures de l'inégalité et polarisation

Quand DISTFLAG est activé, le calcul des mesures d'inégalité est activé en déclenchant le paramètre INEQFLAG. Pendant le calcul, les valeurs nulles et négatives sont éliminées.

Normalement, les valeurs de v et q fixées dans INEQMEASURE produisent le coefficient de GINI, les valeurs du coefficient de S-GINI avec $v=0,5$ et $1,5$. L'utilisateur peut fixer jusqu'à dix valeurs d'inégalité de revenu. Lorsque le coefficient de GINI est demandé, l'indice de polarisation de FW est automatique produit.

Comme fonction particulière, l'utilisateur peut demander de donner à q la valeur de la partie de la population tirant un revenu médian en donnant à q la valeur 99,0. Habituellement q peut prendre toute valeur se situant entre 0 et 1, ce qui comprend ces valeurs extrêmes. Lorsque q a la valeur 1, il génère le coefficient de S-GINI.

Fonction Points de retournement

INTRODUCTION

La fonction Points de retournement du MSPS fournit un moyen permettant de faire l'analyse des points auxquels le taux marginal d'imposition du ménage d'un individu change quand son revenu augmente. Le taux marginal d'imposition est la proportion de l'impôt payé sur le dollar supplémentaire de revenu. Les points de retournement représentent les changements qu'il y a dans la pente de la fonction fiscale/de transfert entière (le taux marginal d'imposition). La fonction calcule les divers niveaux de revenu où le ménage d'un individu, selon ses caractéristiques, verrait des changements dans son taux marginal d'imposition, puis elle produit de l'information sur la valeur d'un certain nombre de variables à chacun de ces niveaux de revenu.

Les niveaux de revenu où les points de retournement surviendront en général dans une fonction fiscale/de transfert d'un individu comprend le niveau de revenu où les contributions du RPC sont déduites en premier, le niveau de revenu où les cotisations d'AC sont faites en premier, le niveau de revenu où l'impôt sur le revenu du gouvernement fédéral devient payable, etc. Ici, la fonction fiscale/de transfert de 1988 réelle d'une personne seule résidant en Nouvelle-Écosse, qui a travaillé pendant les 52 semaines de 1988 et dont la seule source de revenu provient d'un emploi salarié, présenté dans le diagramme suivant. **Il faut noter que cette simulation a été créée avec une version précédente du MSPS et que le diagramme a été produit avec un autre logiciel à partir des résultats obtenus du MSPS.**

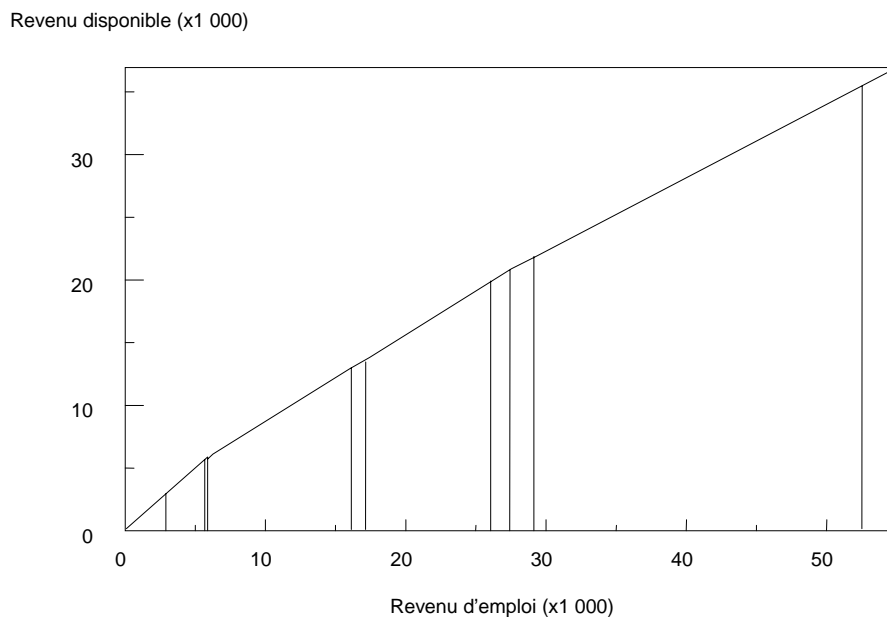


Figure 1. Fonctions fiscales pour le chef de ménage, année d'imposition 1988

La fonction Points de retournement a déterminé le niveau de revenus à chaque point dans la fonction fiscale/de transfert où la courbe de la fonction fiscale changeait, c'est-à-dire que le taux marginal d'imposition changeait. Chaque point de retournement est marqué par une ligne verticale dans le graphique ci-dessus. À partir du produit de la fonction Points de retournement pour cet individu, il est possible de calculer les taux marginaux d'imposition et de les présenter dans le graphique suivant. La raison de chaque point de retournement est enregistrée dans les notes qui suivent le graphique :

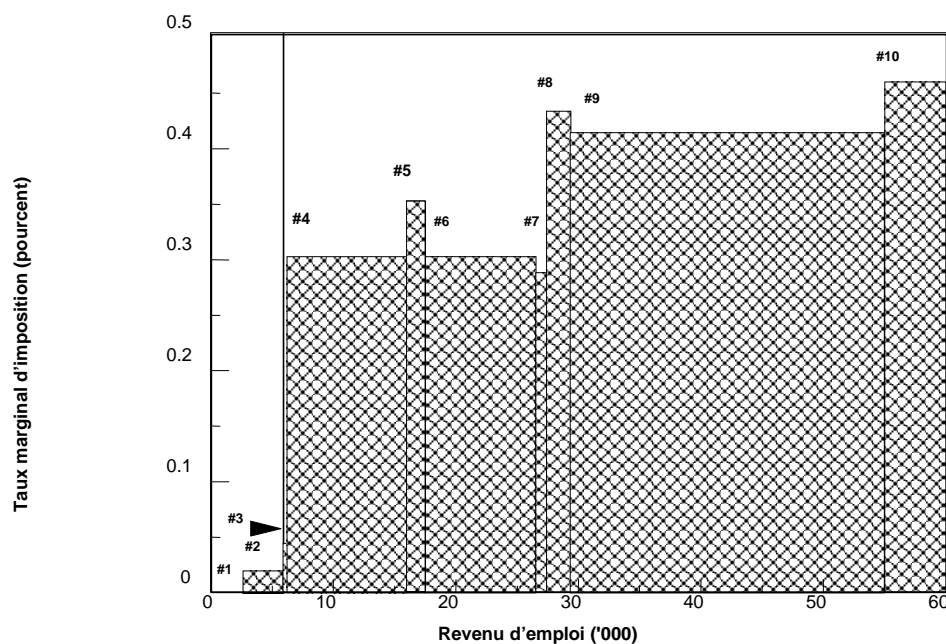


Figure 2. Points de retournement pour le chef de ménage, année d'imposition 1988

- [1] L'individu reçoit le revenu d'emploi
- [2] Le niveau de revenu où les contribution du RPC sont payées
- [3] Le niveau de revenu où les contributions d'AC sont faites.
- [4] L'individu paie l'impôt de base sur le revenu du gouvernement fédéral et la surtaxe du gouvernement fédéral à ce niveau de revenu.
- [5] Le crédit de taxe de vente du fédéral commence à diminuer à ce niveau de revenu.
- [6] Le niveau de revenu où le crédit de taxe de vente du fédéral devient zéro.
- [7] Le niveau de revenu où les contributions au RPC atteignent leur valeur maximale de crédit d'impôt.
- [8] L'individu atteint le deuxième taux d'imposition.
- [9] Le niveau de revenus où les contributions d'AC atteignent leur valeur maximale de crédit d'impôt.
- [10] L'individu atteint le troisième taux d'imposition du gouvernement fédéral.

Il faut noter que le graphique a été produit avec le logiciel utilisant la sortie de la BD/MSPS. La section suivante explique la façon d'accéder à la fonction Points de retournement de la BD/MSPS. Après cette étude, un exemple pratique est exposé et décrit le fichier des paramètres de commande qui a produit les données représentées dans les graphiques ci-dessus. On trouve aussi un exemple de la sortie texte de la fonction Points de retournement.

FONCTIONNEMENT DE LA FONCTION POINTS DE RETOURNEMENT

La fonction Points de retournement du MSPS traite un ménage à la fois. La fonction identifie les divers niveaux de revenus, au cent près, où le ménage verrait un changement dans le taux marginal d'imposition du fait des changements dans l'applicabilité des taxes, des impôts ou des transferts, à partir de l'information contenue dans les déclarations de revenu originales de chacun des membres du ménages et tel qu'il est enregistré dans la base de données. [La fonction Points de retournement n'est pas un modèle d'optimisation; par conséquent, elle agit de façon à optimiser le niveau de revenus disponible du ménage pour chaque revenu donné seulement. Cependant, le MSPS optimise les valeurs de plusieurs paramètres fiscaux/de transfert, par exemple, les frais de garde pour enfants.] Pour activer la fonction Points de retournement, le paramètre de commande TPFLAG doit être déclenché.

Un ménage est sélectionné pour la modification des points de retournement avec la fonction Sélection (voir la section Fonctions de sous-échantillonnage). Le paramètre SELSPEC devrait contenir une expression qui définit le type de ménage que l'analyste désire examiner, par exemple la province du ménage ou le nombre d'enfants, etc. Le paramètre SELMAX, qui définit le nombre maximal de ménages sélectionnés, devrait être fixé à 1 pour sélectionner un seul

ménage pour modification avec la fonction Points de retournement. Il est possible de sélectionner plus qu'un ménage; cependant, plus le nombre de ménages sélectionnés est élevé, plus le traitement prend de temps. Le premier ménage examiné par le MSPS contient au moins un membre qui respecte les critères de sélection sera choisi de la base de données BDSPPS. Si aucun ménage de la base de données ne respecte les critères de sélection, il n'y aura pas de ménage choisi pour la modification des points de retournement. Étant donné que la fonction Points de retournements calcule les points de retournement au niveau du ménage seulement, le paramètre SELUNIT doit être fixé à 4, le niveau du ménage. Si un niveau d'analyse plus bas est choisi, l'analyste peut observer une sortie qui contient des points de retournement "inexplicables", ce qui entraîne des changements dans les paramètres fiscaux/de transfert des autres membres du ménage qui n'ont pas été sélectionnés pour la sortie.

L'analyste doit aussi spécifier quels revenus des membres du ménage seront augmentés par la fonction Points de retournement. Les membres du ménage non choisis pour l'augmentation du revenu conserveront leur revenu original à chaque point de retournement. Cette sélection se fait avec le paramètre TPSPEC, une chaîne contenant une expression (voir à la section Variables la description des expressions du MSPS) qui identifie la personne du ménage choisie dont le revenu sera modifié par la fonction Points de retournement. Ainsi, l'analyste pourrait donner à TPSPEC l'expression `idcfrh==0`, ce qui ferait que seul le revenu du chef de famille de recensement sera modifié par la fonction Points de retournement. Lorsque ce paramètre est fixé à 1, les revenus de tous les membres du ménage seront augmentés. L'analyste doit se rappeler que le nombre de points de retournement produits augmentera substantiellement à mesure que le nombre de membres dont le revenu sera modifié augmentera.

L'analyste doit aussi identifier quelles sources seront augmentées par la fonction Points de retournement. Le paramètre TPVARS précise les sources de revenus sélectionnées. La fonction peut modifier la variable d'analyse de la base de données (c.-à-d. contenant le préfixe `id`). S'il y a plus qu'une variable spécifiée, il faut alors entrer chacune sous forme de liste de variables séparées par des caractères vides. En général, l'analyste désirera modifier seulement une source de revenus, comme `idiemp` (revenu d'emploi), mais sa décision dépendra de la plage de revenus sélectionnée et des résultats désirés.

La plage de revenu, dans laquelle les points de retournement seront calculés, doit aussi être précisée. L'analyste aura fixé une limite inférieure et une limite supérieure de la plage de revenu, exprimées en dollars ou exprimées par rapport au revenu du ménage, selon la valeur de `TPMETH`. Par exemple, l'analyste peut choisir une plage de revenu se situant entre 0 \$ et 80 000 \$ du revenu du ménage ou une plage de revenu se situant entre la valeur actuelle du revenu du ménage et dix fois la valeur actuelle du revenu du ménage. Pour préciser une modification de la valeur des revenus en dollars, le paramètre `TPMETH` doit être fixé à 1. La modification par facteur d'ajustement est activée quand on fixe à 2 la valeur du paramètre `TPMETH`.

Lorsque l'on sélectionne la modification de la valeur en dollars, la valeur des paramètres de commande `TPLL` (limite inférieure des points de retournement) et `TPUL` (limite supérieure des points de retournement) est fixé à une valeur particulière, en dollars. Par exemple, si `TPLL` est fixé à zéro et `TPUL` est fixé à 80 000, les points de retournement qui se situent dans la plage de 0 à 80 000 \$ de revenus seront calculés. Si l'on sélectionne une limite supérieure plus grande que

80 000 \$, par exemple, il n'y aura pas plus de points de retournement, s'il n'y a aucun autre changement dans le taux marginal d'imposition à une valeur de revenu plus élevée. Lorsque la modification du facteur d'ajustement est sélectionnée, la valeur de TPUL devient un facteur d'ajustement; par exemple, en fixant à 10 la valeur de TPUL, on produit des points de retournement jusqu'à 10 fois la valeur de revenu originale du ménage. Dans ce cas, si la valeur de TPLL est fixée à 0, l'ajustement commencera à partir de 0 \$ de revenus; si TPLL est fixé à 1, alors l'ajustement commencera au niveau de revenu original et, si TPLL est fixé à 2, l'ajustement commencera à deux fois le niveau original du revenu. Les modifications par facteur d'ajustement fait qu'il est possible de modifier le revenu de plusieurs membres du ménage et d'observer les changements dans leurs impôts, leurs taxes et leurs transferts, quand le revenu augmente d'un certain pourcentage. Il faut se rappeler d'ajouter les paramètres TPLL et TPUL lorsque l'on passe de la modification de la valeur en dollars à la modification par facteur d'ajustement.

S'il y a plusieurs membres d'un ménage identifiés pour la modification des points de retournement et que TPMETH=2, la relation entre les revenus des membres du ménage sélectionnés est préservée tout au long de l'analyse. Si, par exemple, le revenu du premier membre sélectionné est deux fois plus élevé que le revenu du seul autre membre sélectionné, alors le premier membre soit les deux tiers de chaque cent d'augmentation de revenus. De même, s'il y a plusieurs sources de revenus sélectionnées en vue de l'augmentation et que TPMETH=2, alors la relation entre les deux sources de revenus est préservée. Si la première source de revenus est deux fois plus élevée que la seconde source de revenus, alors la première source reçoit les deux tiers de chaque augmentation de revenus. Si TPMETH est fixé à 1, alors les deux membres du ménage sélectionnés et (ou) les deux sources de revenus sélectionnées augmentent de la même quantité exacte à chaque point de retournement. Par conséquent, TPMETH=2 convient mieux comme limite d'estimation des points de retournement de plusieurs membres d'un ménage ou de plusieurs sources de revenus.

La formule et le contenu de la sortie de la fonction Points de retournements doivent être spécifiés. Les résultats peuvent être envoyés à la fonction de Sortie texte ou l'analyste peut exporter les résultats en utilisant la fonction Sortie SAS pour poursuivre l'analyse. **Une sortie texte concise peut être obtenue par l'utilisateur s'il inclut /detsum.cpi dans le fichier des paramètres de commande.** Pour accroître la précision de la sortie texte, le paramètre de commande ASCEXTPRC devrait être fixé à 2 ou plus. Le calcul précis des taux marginaux d'imposition peut exiger cette précision supplémentaire.

Il faut noter que, avant l'exécution de l'estimation des points de retournement, le paramètre de commande ROUNDFLAG devrait être désactivé. Si la fonction d'arrondissement du nombre n'est pas désactivée, les résultats donneront de nombreux points de retournements superflus. Le modèle désactivera automatiquement ROUNDFLAG si l'analyste ne le fait pas et il s'affichera un commentaire dans le journal. L'analyste devrait aussi désactiver le paramètre fiscal/de transfert CTFLAG. Si la fonction de calcul des taxes à la consommation demeure activée, alors les taxes à la consommation seront calculées selon un régime de dépenses constant, quel que soit le niveau de revenu. Par conséquent, le seul effet des taxes à la consommation que le modèle peut produire est une augmentation du taux marginal d'imposition d'un montant constant, à tous les niveaux de revenus.

Le taux marginal d'imposition réel affiché dans les graphiques précédents a été calculé avec la formule suivante :

$$MTR_{i(i+1)} = \frac{immicons_{i+1} - immicons_i}{idiemp_{i+1} - idiemp_i}$$

Cette formule représente le taux marginal d'imposition dans l'intervalle se situant entre les points de retournement i et $i+1$. Le dénominateur de cette fonction contient la variable spécifiée par TPVARS.

EXEMPLE

Dans cet exemple de simulation, créée avec une version précédente de la fonction Points de retournement du MSPS, a été activée afin de modifier le revenu d'emploi de 1988 d'une personne seule qui avait les caractéristiques suivantes :

Caractéristiques

- la personne est le seul membre de ce ménage
- la personne a un certain revenu
- la seule source de revenu de la personne est un emploi
- la personne est un employé à temps plein
- la personne a travaillé au moins 50 semaines au cours de 1988
- la personne réside en Nouvelle-Écosse

Expression MSPS

hhnin==1
immtot>0
immtot==idiemp
idlyfp==1
idlyww>=50
hdprov==2

La section pertinente du fichier des paramètres de commande (.cpr) qui sélectionne un individu respectant les spécifications ci-dessus est donnée ci-dessous.

```
###
## 2.1.7 Fonction de sélection d'enregistrements
###

SELFLAG      1    # Drapeau de déclenchement de fonction de sélection
SELUNIT      4    # Niveau de famille pour la fonction de sélection
SELSPEC      # Spécifications pour la sélection
hdprov==2 && hhnin==1 && idiemp>0 && idlyfp==1 && idlyww>=50
&& immtot==idiemp
SELMAX      1    # Fonction de sélection du nombre maximum de ménages

###
## 2.1.9 Sélection du point de retournement
###

TPFLAG      1    # Drapeau pour sélectionner le point de retournement
TPSPEC      1    # Expression identifiant les individus à changer
TPVARS      idiemp # Variables à modifier
TPMETH      1    # Méthode pour modifier les variables
TPLL        0.00  # Limite inférieure pour les variables modifiées
TPUL        60000.00 # Limite supérieure pour les variables modifiées
```

SELFLAG active la fonction Sélection. SELUNIT est fixé au niveau du ménage. SELSPEC identifie les ménages pertinents de la base de données. SELMAX = 1 indique que les points de

retournement seront calculés seulement pour le premier ménage qui correspond à SELSPEC. TPFLAG active la fonction Points de retournement. TPSPEC = 1 indique que les revenus de tous les membres du ménage seront augmentés; cependant, SELSPEC a restreint le nombre de membres du ménage à 1. La modification de la valeur en dollars a été choisie par TPMETH = 1; par conséquent, les points de retournement seront calculés pour une plage de revenus de 0 à 60 000 \$ de revenu d'emploi. Le paramètre de commande ROUNDFLAG est aussi fixé à zéro.

La fonction Sortie texte a été fixée de façon à produire la sortie appropriée. La configuration est donnée par le fichier \spsd\detsum.cpi. Enfin, le modèle de taxes à la consommation est désactivé par CTFLAG = 0 dans le fichier .mpr approprié.

Les utilisateurs noteront que, pour le calcul du taux marginal d'imposition dans l'exemple, le revenu total est égal au revenu d'emploi du fait de nos critères de sélection. En outre, le revenu consommable est égal au revenu disponible étant donné que le modèle de taxe à la consommation a été désactivé.

Voici une liste de la sortie de la fonction Sortie texte produite à l'exécution. La liste modifiée donne toutes les variables non-zéro pertinentes pour la situation fiscale de cet individu. La liste est un sous-ensemble de la liste produite par la configuration de la fonction Sortie que l'on trouve dans le fichier \spsd\detsum.cpi.

1	hdseqhh	Numéro d'ordre du ménage	217	217	217	217	217
2	hdprov	Province	2	2	2	2	2
3	idage	Âge	42	42	42	42	42
4	idsex	Sexe.....	0	0	0	0	0
5	idmarst	État matrimonial.....	3	3	3	3	3
6	idlyww	Domaines de travail	52	52	52	52	52
7	idiemp	Revenu d'emploi	0	2600	5876	5876	6267
8	imitot	Revenu total	0	2600	5876	5876	6267
9	imcqqpc	Cotisations au RPC/RRQ.....	0	0	66	66	73
10	imuic	Cotisations à l'assurance-chômage..	0	0	0	138	147
11	idothded	Autres déductions du revenu total	13	13	13	13	13
12	imdedft	Déductions du revenu total.....	13	13	13	13	13
13	iminet	Revenu net.....	0	2587	5863	5863	6254
14	idcharit	Dons de charité.(340).....	34	34	34	34	34
15	imitax	Revenu imposable	0	2587	5863	5863	6254
16	imfedtax	Impôt fédéral avant les crédits					
	d'impôt.....		0	440	997	997	1063
17	imbtc	Crédit d'impôt personnel de base.	1020	1020	1020	1020	1020
18	imchartc	Crédit d'impôt pour dons de charité	0	6	6	6	6
19	imcppctc	Crédit d'impôt pour cotisation au					
	RPC.....		0	0	11	11	12
20	imuictc	Crédit d'impôt pour cotisations					
	à l'assurance chômage.....		0	0	0	23	25
21	imtaxcr	Total des crédits d'impôt.....	1020	1026	1037	1060	1063
22	imatxcrt	Total des crédits d'impôt appliqués	0	440	997	997	1063
23	imfstc	Crédit pour taxe fédérale sur					
	les ventes.....		70	70	70	70	70
24	immemp	Tous les revenus d'emploi.....	0	2600	5876	5876	6267
25	immmkt	Revenu d'emploi.....	0	2600	5876	5876	6267
26	imftran	Revenu de transferts-fédéral.....	70	70	70	70	70
27	immtran	Tous les revenus de transfert...	70	70	70	70	70
28	immtot	Revenu total.....	70	2670	5946	5946	6337
29	immtax	Tous les impôts.....	0	0	66	204	221
30	imptax	Impôts provinciaux.....	0	0	0	0	0
31	imftax	Impôts fédéraux.....	0	0	66	204	221
32	immdisp	Revenu disponible	70	2670	5880	5742	6117
33	imtxfc	Taxes fédérales à la consommation.	0	0	0	0	0
34	immicons	Revenu consommable.....	70	2670	5880	5742	6117

Étant donné que TPLL était fixé à zéro, le niveau de revenu d'emploi initial est de zéro (colonne 1, ligne 7). Le premier point de retournement, revenu d'emploi de 2 600 \$ (colonne 2, ligne 7), se produit parce qu'une augmentation de 1 % du revenu au-dessus de 2 600 \$ entraîne la déduction des contributions au RPC (colonne 2, ligne 9). Le troisième niveau de revenus d'emploi, 5 876 \$, indique le seuil de revenu au-delà duquel les cotisations à l'AC sont retenues. Un changement du revenu d'emploi de moins d'un cent de plus que ce niveau de revenu, représenté aussi comme 5 876 \$ (colonne 4, ligne 7) entraîne la déduction de 138 \$ (colonne 4, ligne 10) des cotisations à l'AC. Le cinquième niveau de revenu d'emploi, 6 267 \$, indique le point où l'impôt fédéral sur le revenu de l'individu est égal au crédit d'impôt non remboursable

total (colonne 5, lignes 21 et 22). Après ce point, il paie de l'impôt fédéral sur le revenu. Les autres points de retournement de la fonction fiscale/de transfert de l'individu ne sont pas représentés ici, mais ils ont été étudiés plus haut.

Fonctions diverses

FONCTION NOMBRES ALÉATOIRES

Il est habituel que le programme fiscal/de transfert, même s'il peut être orienté vers une population en particulier, n'atteigne pas cette population. Ce phénomène de "participation" peut être modélisé avec des numéros pseudo-aléatoires si une probabilité de participation est donnée. Le MSPS fournit jusqu'à 20 séries indépendantes de nombres aléatoires qui sont modélisées par le paramètre vectoriel SEED. Le nombre d'éléments de SEED est le nombre de nombres aléatoires indépendants générés. Chaque élément de SEED fournit un nombre entier utilisé pour lancer le générateur de nombres aléatoires. Ces éléments sont habituellement fixés à des valeurs différentes pour générer des nombres aléatoires indépendants.

Les nombres aléatoires sont générés par un algorithme standard à entiers et congruence, sont stockés dans les variables de niveau de l'individu idrand0 à idrand19 et prennent des valeurs de 0.0 à 1.0, avec une probabilité uniforme. La fonction Nombres aléatoires a été mise en œuvre de façon à fonctionner de façon reproductible, indépendamment de toute sélection d'enregistrements. En d'autres mots, pour une valeur donnée de SEED, les nombres aléatoires idrand0 à idrand19 auront des valeurs identiques pour un individu en particulier dans la base de données, indépendamment de toute sélection qui peut avoir été demandée.

Il vaut mieux toujours augmenter le paramètre SEED (p. ex. ne jamais modifier SEED pour avoir moins de rangées que la valeur par défaut de cette version) car l'exécutable du MSPS utilise quelques-unes des séries de nombres aléatoires dans le code du modèle. Lorsqu'il y a réduction du nombre de nombres aléatoires générés, les résultats du modèle changeront.

ANALYSE DU SEUIL DE FAIBLE REVENU

La BD/MSPS facilite l'analyse du seuil de faible revenu en fournissant les moyens pour examiner le faible revenu avant et après impôt. Ceci peut être fait soit avec une table intégrée ou en utilisant les variables d'utilisateur UVAR, ou les deux.

Note importante : Dans cette version du MSPS, la variable pour urbanisation, hduurb, a été comprimée. Ce qui veut dire que des seuils de faible revenu inappropriés pourraient s'appliquer aux familles dans les régions comprimées. Veuillez être vigilant en utilisant les SFR.

Les seuils de faible revenu incluent avec la BD/MSPS sont les SFRs avant impôt et après impôt – PTF et PTFAT respectivement. Ils sont considérés comme paramètres d'ajustement de la base de données. Pendant l'exécution du modèle, le seuil de faible revenu approprié pour la taille et la classe d'urbanisation d'une famille donnée est placé dans les variables de niveau de famille efpovthr (avant impôt) et efpvthat (après impôt).

Pareillement, le revenu d'une famille qui sera comparé à ces seuils est cumulé et mémorisé dans les variables individuelles `impovinc` (la somme du revenu pour être comparé au SFR avant impôt) et `imatpinc` (la somme du revenu pour être comparé au SFR après impôt). Pour éviter le compte en double, ces variables contiennent le revenu de la famille économique, mais seulement pour la première personne de la famille. Il faut noter que `imatpinc`, la valeur de l'après impôt, est basé sur le revenu de consommation. Si un utilisateur veut comparer le revenu disponible au SFR après impôt, il faut qu'il rende invalide les calculs des taxes à la consommation avec le paramètre fiscal/de tranfert `CTFLAG`.

Ces quatre variables sont disponibles pour être utilisé dans la fonction variable utilisateur. Une expression de la forme `btppoor=EF:impovinc < efpovthr`; identifierait les individus des familles qui se trouvent sous le seuil de faible revenu avant impôt. L'expression `atppoor=EF:imatpinc < efvthr`; identifierait les individus des familles qui se trouvent sous le seuil de faible revenu après impôt. Alors, `XTSPEC` de `IN:{btppoor,atppoor}` donnerait un total d'individus sous le SFR à la fois avant et après impôt. Alternativement, l'expression `EF:btppoor=EF:impovinc < efpovthr`; nous ferait remarquer le chef de la famille économique à faible revenu et puis `XTSPEC` de `EF:{btppoor,btppoor/units}`; donnerait un total et une incidence de familles économiques sous le SFR.

La BD/MSPS peut aussi produire un tableau intégré qui montre la distribution de plusieurs variables par un groupement du ratio du revenu d'un ménage pour un seuil de faible revenu spécifié. Le tableau est demandé avec les paramètres de contrôle `T4FLAG` et `T4AFLAG`. De plus, on doit spécifier si les résultats du tableau 4 sont basés sur le SFR avant impôt ou le SFR après impôt. Cela est produit avec le paramètre de contrôle `LICOOPT`.

Il faut aussi noter que, parce que la BDSPS doit garder la confidentialité et tente de corriger certains problèmes de sous-déclaration qu'il y a dans les sources de données, ainsi que l'inclusion possible de taxes de commodités en considération du faible revenu de l'après impôt, les comptes et les chiffres de faible revenu peuvent différer de ceux qui sont publiés ailleurs.

AJUSTEMENT À L'AIDE SOCIALE ET À L'ASSURANCE-EMPLOI

Aide sociale

L'algorithme standard donne le nombre attendu de prestataires de l'AS lorsque `SAPFLAG` est désactivé. Lorsque `SAPFLAG` est activé, le nombre de prestataires de l'AS ou le total des dépenses de l'AS peut être ajusté si l'on suit les étapes suivantes :

- Activer `SAPFLAG`. La valeur standard de `SATARGET[xx]` (où `xx` est un numéro de province) est fixé par l'équipe de la BD/MSPS. Pour modifier le nombre total de cas ou le total des dépenses d'AS, il faut donner à `SATARGET[xx]` une valeur se situant entre 0,01 et 0,99. Cela permettra une variation d'environ $\pm 80\%$ dans le nombre de cas, dans la province `xx`.
- Pour atteindre une valeur cible, l'utilisateur devrait utiliser la fonction de recherche de but, `spsmter`, décrite dans le [Guide d'utilisation des outils](#). Le paramètre d'ajustement est `SATARGET[xx]`. Dans l'exemple suivant, nous supposons une diminution de 10 % du nombre de prestataires d'AS en Ontario en 2002.

Si l'utilisateur est intéressé à diverses modifications de l'AS par province, il faudrait appliquer les étapes ci-dessus à des provinces sélectionnées, une à la fois.

Exemple

Supposons une diminution de 10 % des prestataires de l'AS en Ontario en 2002.

La première étape consiste à calculer combien de ménages sont prestataires de l'AS en Ontario. Vous pouvez faire tourner la simulation saont sous le répertoire spsm\example\sa_adjustment pour créer le fichier de sortie. Dans cette simulation, le nombre total de ménages recevant l'AS en Ontario en 2002 est de 400 819. Une réduction de 10 % signifie une valeur cible de 360 737 ménages.

On exécute la fonction de recherche de but en utilisant ce nombre de bénéficiaires. Vu la nature de SATARGET, la méthode binaire devrait être employée. La contrainte applicable est de 0,01 à 0,5, puisque nous réduisons le nombre de ménages. Voir le Guide d'utilisation des outils pour se renseigner sur la façon d'utiliser la fonction de recherche de but.

Pour exécuter ce modèle dans le MSPS Classique, vous pouvez copier le répertoire \\spsm\example\sa_adjustment dans un répertoire de travail et l'exécuter avec le programme de traitement par lot (.bat).

Assurance-emploi

Il y a deux façons d'ajuster les prestations d'assurance-emploi. Vous pouvez soit changer UIEITKP[xx] ou UITARGET[xx] (où xx représente une province). Dans les exemples suivants, on explore deux scénarios : une réduction aléatoire de prestataires d'assurance-chômage et une réduction basée sur les caractéristiques des prestataires. Il est important de noter que:

- les valeurs cibles pourraient être des valeurs en dollars pour un nombre de prestataires;
- une hypothèse implicite est que le taux de chômage demeure le même et que les changements aux paramètres du programme sont déjà mis en œuvre.

Réduction aléatoire

Supposons que le nombre de prestataires de l'assurance-emploi est réduit de 5 % en Nouvelle-Écosse, et que la réduction n'est pas liée à quelque caractéristique personnelle que ce soit. Le paramètre UIEITKP devrait être utilisé pour réduire le nombre de bénéficiaires. Les étapes nécessaires à la production de ce résultat sont les suivantes :

La première étape est de calculer le nombre réel de prestataires de l'assurance-emploi. Vous pouvez tourner la simulation uirand trouver sous le répertoire spsm\example\ui_random pour créer le fichier de sortie. Dans cette simulation, le nombre total de prestataires de l'assurance-emploi pour 2002 en Nouvelle-Écosse est de 111 970 donc, une réduction de 5 % serait de 106 371.

On exécute la fonction de recherche de but en utilisant ce nombre de bénéficiaires. Vu la nature de UIEITKP, la méthode binaire devrait être employée. Voir le Guide d'utilisation des outils

pour se renseigner sur la façon d'utiliser la fonction de recherche de but.

Le tableau suivant illustre la valeur cible finale, le résultat, ainsi que la valeur qui en résulte pour UIEITKP :

Province	Cible	Résultat	UIEITKP
Nouvelle-Écosse	106 371	106 362	0,051758

Pour exécuter cet exemple dans MSPS Classique, vous devriez copier le répertoire [\\spsm\exemple\ui_random](#) dans un répertoire de travail et exécuter le fichier de traitement par lot (.bat).

Réduction systématique

Supposons que le nombre de prestataires de l'assurance-emploi est réduit de 5 % en Nouvelle-Écosse, et que cette réduction est liée à des caractéristiques individuelles. Dans un processus de réduction, les individus ayant le moins de probabilités de devenir prestataires d'assurance-emploi sont abandonnés en premier lieu. Le paramètre UITARGET devrait être utilisé pour réduire le nombre de prestataires.

La première étape consiste à calculer le nombre réel de prestataires de l'assurance-emploi. Vous pouvez tourner la simulation uibehav trouver sous le répertoire spsm\exemple\ui_behaviour pour créer le fichier de sortie. Dans cette simulation, le nombre total de prestataires de l'assurance-emploi pour 2002 en Nouvelle-Écosse est de 111 970 donc, une réduction de 5 % serait de 106 371.

On exécute la fonction de recherche de but pour ce nombre de bénéficiaires. Vu la nature de UITARGET[2], la méthode binaire devrait être employée. Voir le [Guide d'utilisation des outils](#) pour se renseigner sur la façon d'utiliser cette fonction.

Le tableau suivant illustre la valeur cible finale, le résultat, ainsi que la valeur qui en résulte pour UITARGET :

Province	Cible	Résultat	UITARGET
Nouvelle-Écosse	106 371	106 366	0,429688

Dans ce cas, la méthode aléatoire et la méthode de comportement produisent le même nombre de prestataires d'assurance-emploi dans chaque province, étant donné les erreurs d'itération.

La grande différence entre les deux scénarios réside dans la caractéristique des prestataires de l'assurance-emploi qui n'est pas la même dans les deux cas.

Pour exécuter cet exemple dans MSPS Classique, vous devriez copier le répertoire [\\spsm\\example\\ui_behaviour](https://spsm.example/ui_behaviour) dans un répertoire de travail et exécuter le fichier de traitement par lot (.bat).