Techniques d'enquête, juin 1987
Vol. 13, n° 1, pp. 35-51
Statistique Canada

Utilisation des dossiers administratifs à des fins statistiques

G.J. BRACKSTONE[1]

RÉSUMÉ

Les demandes pour les statistiques sur tous les aspects de nos vies, de notre société et de notre économie continuent à augmenter. En même temps, les organismes statistiques, ainsi que beaucoup de leurs répondants, s'inquiètent de plus en plus du fardeau de réponse imposé par les enquêtes. L'utilisation des dossiers administratifs à des fins statistiques s'est présenté comme un des moyens alternatifs pour satisfaire ces demandes statistiques. Cet article expose l'expérience récente à Statistique Canada dans ce domaine, et discute des entraves à une plus grande exploitation de sources administratives. On examine aussi des méthodes possibles pour rendre les systèmes administratifs plus utiles à des fins statistiques, ainsi que certaines questions importantes liées à la protection des renseignements et au jumelage d'enregistrements.

MOTS CLÉS: Estimation indirecte; base de sondage; évaluation des enquêtes; accès; confidentialité.

1. INTRODUCTION

La demande de statistiques se rapportant à de nombreux aspects de notre vie, de notre société, de notre économie et de notre environnement ne cesse d'augmenter. Le fait que nous sommes en mesure, de nos jours, de traiter et de manipuler de vastes ensembles de données, puisque nous entrons dans ce qu'il est convenu d'appeler l'ère de l'information, explique peut-être en partie ce phénomène; celui-ci pourrait également découler de la complexité croissante des systèmes sociaux et économiques et de notre désir de mieux comprendre ces derniers. Quelle que soit la raison qui motive la demande, les organismes statistiques publics doivent y répondre, et ce dans un climat de grandes compressions budgétaires. Ces organismes sont en même temps conscients du fardeau de réponse accru que les répondants auraient à subir s'il fallait mener davantage d'enquêtes pour obtenir les renseignements voulus.

Ces facteurs nous ont amenés à chercher d'autres façons de procéder, la principale étant un plus grand recours aux systèmes administratifs existants. L'idée n'est pas nouvelle. Depuis longtemps déjà les renseignements relatifs à l'état civil, aux importations, aux exportations, aux soins de santé et à l'éducation nous parviennent sous forme de produits dérivés des processus administratifs. Nous verrons plus loin que l'utilisation des données administratives s'est étendue ces dernières années au domaine des statistiques sur les entreprises, les familles et les particuliers.

Les quatre prochaines sections de l'article décrivent les différents types de dossiers administratifs et l'utilisation qui en est faite, notamment à Statistique Canada. Il ressortira de cet exposé que l'appareil statistique canadien est fortement tributaire des dossiers administratifs. La section 6 discute de la question de l'accès aux dossiers administratifs et les moyens de rendre ces dossiers mieux adaptés aux usages statistiques. Enfin, un survol des préoccupations que l'exploitation des dossiers administratifs suscite à l'égard de la protection des renseignements personnels est fourni.

---

[1] G.J. Brackstone, Statisticien en chef adjoint, Statistique Canada, 26-J Immeuble R.H. Coats, Tunney's Pasture, Ottawa, Ontario K1A 0T6.

2. TYPES DE DOSSIERS ADMINISTRATIFS

Des dossiers administratifs, il y en a de toutes sortes. Il est important de faire la distinction entre ceux qui sont tenus au niveau national (habituellement par le gouvernement fédéral) et ceux qui le sont à un niveau inférieur (par les provinces ou les municipalités, par exemple). Pour que ces derniers soient utiles au niveau national, il faut que divers secteurs de compétence s'entendent notamment sur les définitions, les normes, la présentation des dossiers et la marche à suivre. Il n'est pas toujours facile de conclure de telles ententes, surtout dans des domaines qui sont constitutionnellement du ressort des provinces.

Les dossiers administratifs diffèrent par leur objet, lequel détermine en grande partie l'univers visé et leur qualité, et par conséquent leur utilité statistique. On peut distinguer six grandes catégories de dossiers administratifs selon leur raison d'être:

(1) *Les dossiers tenus pour contrôler la circulation des biens et des personnes aux frontières.*

Ces dossiers incluent notamment les dossiers sur les importations, les exportations, l'immigration et l'émigration. L'univers visé et le contenu de ces dossiers administratifs dépendent des lois et des règlements à faire respecter et du succès avec lequel ces lois et règlements sont appliqués. Dans l'ensemble, ils sont bien appliqués. Les dossiers administratifs sur l'immigration ne contiennent pas d'information concernants les immigrants illégaux, par définition, mais sont complets à d'autres égards. Toutefois, comme l'émigration n'est pas contrôlée, il n'existe pas de dossiers administratifs concernant l'émigration. Les dossiers administratifs sur les importations canadiennes sont en général plus précis que les dossiers sur les exportations parce que le besoin d'établir les droits de douane frappant les importations requiert l'information plus détaillée.

(2) *Les dossiers tenus en vertu de l'obligation légale d'enregistrer certains événements.*

À titre d'exemple de ces événements, citons les naissances, les décès, les mariages, les divorces, les constitutions d'entreprises en société ou les fusions d'entreprises et l'octroi de licences. De façon générale, l'univers visé et la qualité des dossiers recueillis à cette fin sont très élevés au Canada, car tout citoyen doit fournir une preuve que ce genre d'enregistrement a été fait pour obtenir des droits ou des avantages.

(3) *Les dossiers nécessaires à la prestation d'avantages et à l'administration d'engagements d'assistance.*

Citons par exemple les impôts, l'assurance-chômage, les pensions, l'assurance-maladie et les allocations familiales. L'univers visé et le contenu de ces dossiers dépendent beaucoup des programmes auxquels ils se rapportent. L'univers visé peut être très bien couvert, mais, pour des raisons politiques ou administratives, sa définition risque de ne pas être la plus utile pour les travaux d'analyse.

(4) *Les dossiers nécessaires à l'administration des établissements publics.*

Il s'agit notamment des dossiers concernant les écoles, les universités, les établissements de soins de santé, les tribunaux et les prisons. De façon générale, ces dossiers portent principalement sur le nombre de cas dont l'établissement a la responsabilité et non sur chaque personne qui y fait un séjour. Par contre, ils donnent habituellement des statistiques agrégées très complètes sur la population qui fréquente ces établissements. Au Canada, beaucoup de dossiers administratifs de cette catégorie sont du ressort des provinces.

(5) *Les dossiers qui découlent de la réglementation de certaines activités économiques par le gouvernement.*

Citons les dossiers relatifs au transport, aux affaires bancaires, à la radiodiffusion et aux télécommunications ainsi qu'à la gestion des approvisionnements ou des prix de certaines marchandises, particulièrement dans le secteur agricole.

(6) *Les dossiers liés à la prestation de services publics.*

Ces services comprennent l'électricité, le téléphone et l'eau. Le taux de couverture des abonnés et la qualité des renseignements relatifs à la prestation des services et à la facturation sont bons en général. Plusieurs de ces services sont administrés au niveau provincial ou municipal.

La façon dont les dossiers administratifs sont constitués varie également beaucoup d'un organisme à un autre. La plupart des systèmes administratifs dont l'univers visé est grand sont à présent automatisés, mais le fait que le matériel utilisé et la présentation des données diffèrent non seulement d'un secteur de compétence à un autre mais aussi entre l'organisme administratif et l'organisme statistique pose un problème qu'il reste à régler. L'automatisation de plus en plus répandue entraîne également le problème des nombreuses modifications apportées aux dossiers initiaux par l'organisme administratif intéressé avant qu'ils soient transmis à l'organisme statistique. Alors qu'un meilleur contrôle de la qualité des formulaires reçus aurait peut-être de bonnes chances d'améliorer la qualité finale du fichier administratif, l'organisme statistique est obligé, lui, de procéder à des travaux supplémentaires pour connaître et évaluer l'incidence de toute opération de traitement préliminaire que l'organisme administratif a pu exécuter. Dans certains services administratifs, les dossiers individuels demeurent à l'endroit où ils ont été constitués et seules les données agrégées sont conservées dans un système central. Cette pratique restreint la capacité de l'organisme statistique d'évaluer la qualité des données et limite les possibilités en matière d'analyse.

Enfin, les dossiers diffèrent du point de vue de leur accessibilité. Il y a des lois et des règlements qui régissent l'accès à certains dossiers administratifs ainsi que leur utilisation à des fins secondaires, c'est-à-dire autres que celles pour lesquelles ils ont été constitués, notamment à des fins statistiques. Nous reparlerons de cette question à la section 6.

3. UTILISATION DES DOSSIERS ADMINISTRATIFS

Les données administratives peuvent se prêter à quatre catégories d'utilisation statistique. La plupart des applications qui sont faites de ces dossiers à des fins statistiques se rattachent à une de ces catégories, en représentent une combinaison ou en constituent une variante.

(1) *Totalisations directes*

On inclut ici le comptage des unités dans les fichiers, leur classement recoupé selon certaines caractéristiques et le regroupement des variables quantitatives associées à chaque unité. Les statistiques sur l'état civil et le commerce extérieur sont des exemples importants. Mentionnons également la publication des chiffres mensuels sur le nombre de demandeurs et de bénéficiaires de l'assurance-chômage ventilés selon la province, l'âge, le sexe, la durée des prestations et leur nature ainsi que la production des sommaires annuels sur la répartition du revenu dans chaque comté d'après les données du fichier de l'impôt sur le revenu des particuliers.

(2) *Estimations indirectes*

Cette catégorie comprend les travaux où l'on se sert de données tirées des dossiers administratifs pour effectuer une estimation. Un exemple serait le couplage de deux déclarations d'impôt consécutives d'un particulier dans le but de produire des estimations partielles de la migration, lesquelles peuvent être pondérées à l'aide des données de recensement utilisées comme données de référence. Ces estimations de la migration sont ensuite utilisées dans le programme d'estimations démographiques de Statistique Canada, au même titre que les données administratives sur les naissances, les décès et l'immigration. Un deuxième exemple serait l'utilisation des données fiscales sur

les petites entreprises au lieu de la collecte des données requises au moyen d'une enquête. Les données fiscales, corrigées s'il y a lieu, sont combinées aux données d'enquête sur les grandes entreprises pour produire des données agrégées sur la branche d'activité.

On classe également dans cette catégorie d'utilisation le couplage de différents fichiers administratifs ou statistiques dans le but de produire des estimations. Par exemple, le couplage du registre des décès avec les fichiers des personnes exposées à certains risques dans le but d'estimer les taux de mortalité différentielle, ou encore le couplage de dossiers fiscaux avec des dossiers de l'assurance-chômage et de la formation de la main-d'oeuvre dans le but d'analyser la participation et l'adaptation au marché du travail.

(3) *Bases de sondage*

Dans cette catégorie, nous incluons l'utilisation de dossiers administratifs pour créer, compléter ou mettre à jour les bases de sondage servant aux recensements ou aux enquêtes. Un bon exemple serait l'utilisation des renseignements que les employeurs fournissent à Revenu Canada au sujet des retenues sur la paye. Le questionnaire que les titulaires d'un nouveau compte de retenues sur la paye doivent remplir permet de savoir exactement si de nouvelles entreprises ont été crées ou si des modifications ont été apportées à la structure d'entreprises existantes. Même s'il n'y a pas au Canada de registre de logements, un deuxième exemple serait l'utilisation des permis de bâtir ou des listes de nouveaux abonnés au téléphone ou au service d'électricité pour déterminer qu'un logement est nouveau.

(4) *Évaluation des enquêtes*

Cette catégorie comprend l'utilisation de dossiers administratifs à des fins de vérification, de validation ou d'évaluation de données produites au moyen d'une enquête, et ce au niveau de chaque unité ou à un niveau agrégé. Dans le passé, on s'est servi des dossiers de l'immigration et de l'impôt pour évaluer les questions du recensement sur l'immigration et les revenus, et des dossiers des allocations familiales pour vérifier le taux de couverture du recensement en ce qui a trait aux enfants.

Il y a un facteur important qui permet de déterminer l'utilisation qui sera faite d'une source de renseignements administratifs, et c'est la qualité présumée des données administratives comparativement à celle des données d'enquête correspondantes. Certaines études d'évaluation utilisent les dossiers administratifs pour évaluer les réponses aux questions des enquêtes, alors que d'autres se servent de données d'enquête comme données de référence pour évaluer les estimations produites à partir de données administratives. La qualité des dossiers administratifs doit être évaluée cas à cas. Du point de vue de l'utilisation statistique qu'on va en faire, elle dépend en général d'au moins trois facteurs:

(i) les définitions utilisées pour les besoins du système administratif;
(ii) le taux de couverture voulu du système administratif;
(iii) la qualité des données recueillies et de leur traitement par les services administratifs.

Si un de ces trois facteurs présente des faiblesses, l'utilité statistique des dossiers administratifs risque d'être compromise. La rapidité avec laquelle ces dossiers peuvent être obtenus est un autre aspect important à considérer. Les contraintes qui peuvent éventuellement limiter cette utilisation et qu'il faut par conséquent considérer avant toute décision ont été traitées ailleurs (voir par exemple Brackstone, 1984). Le tableau 1 compare les points forts et les points faibles des dossiers administratifs et ceux des recensements et des enquêtes.

Pour illustrer l'utilisation qui est faite au Canada des dossiers administratifs, nous allons décrire deux domaines d'application à Statistique Canada. Le premier a trait à la production de statistiques sur les entreprises et le second, à la production de statistiques sur les particuliers et les familles.

Tableau 1

Comparaison des recensements, des enquêtes et des dossiers administratifs comme sources de données statistiques

Éléments	Recensements	Enquêtes	Dossiers administratifs
1. Couverture	Visent à une couveture à 100% de la population	Certaines enquêtes ne comprennent pas certains secteurs de la population (par exemple, les réserves indiennes, les régions éloignées)	Les populations visées sont définies par les exigences administratives
2. Contenu	La grande diversité des éléments de données permet beaucoup de classements recoupés	Couvrent habituellement un petit nombre de sujets mais de façon plus approfondie qu'un recensement	Limités aux variables requises pour des fins administratives
3. Concepts/définitions	Peuvent être établis en fonction des besoins de l'analyse sociale et économique	Peuvent être établis en fonction des besoins de l'analyse sociale et économique	Établis en fonction des exigences administratives
4. Estimations relatives aux petites régions	Disponibles par suite des efforts déployés pour arriver à une couverture complète	Ne peuvent être produites dans la plupart des cas	Peuvent être produites, pourvu que chaque dossier ait un code géographique correspondant à une petite région
5. Contrôle qualitatif	Peut être conçu pour réduire au minimum les erreurs	Étant de plus petite taille que les recensements, ils permettent un contrôle encore plus serré que dans les recensements	Les dossiers administratifs sont soumis au contrôle de l'organisme administratif et le contrôle qualitatif n'attire pas l'attention, sauf peut-être pour les variables clés
6. Coût	Très élevé	Coût relativement faible par enquête, même si le coût cumulatif d'une enquête habituelle portant sur une période intercensitaire de cinq ans peut être élevé	Relativement peu élevé si les coûts initiaux de collecte sont imputés au programme administratif
7. Fréquence	Tous les 5 ou 10 ans (selon les sujets)	Peuvent être annuelles, trimestrielles ou mensuelles selon les sujets	Peuvent être annuels ou mensuels selon le programme administratif
8. Actualité des données	Les données peuvent être obtenues six mois à 2½ ans après le jour du recensement	Les enquêtes habituelles répétées produisent des résultats en quelques semaines, alors que les enquêtes spéciales peuvent nécessiter plusieurs mois	Selon le processus administratif, alors qu'il se peut qu'une forme utilisable d'un fichier annuel ne soit pas prête avant qu'une bonne partie de l'année suivante soit écoulée
9. Stabilité	Les changements sont sous le contrôle des statisticiens qui répondent aux besoins des utilisateurs	Dans des enquêtes répétées, les changements sont peu fréquents pour permettre les comparaisons temporelles	Des changements peuvent être apportés en raison de modifications législatives ou de modifications de règlements ou en raison de modifications des pratiques administratives
10. Fardeau du répondant	Lourd mais peu fréquent; on le réduit par l'utilisation d'échantillonnage	Léger en moyenne, bien que lourd pour ceux qui ont été choisis	Aucun fardeau supplémentaire

4. DONNÉES ADMINISTRATIVES ET ENQUÊTES-ENTREPRISES

Un projet de remaniement complet de l'infrastructure et de la stratégie sur lesquelles s'appuie le programme d'enquêtes-entreprises est actuellement en cours à Statistique Canada. Il s'agit en particulier de procéder à la refonte du registre d'entreprises (qui sert de base de sondage aux enquêtes-entreprises), de repenser le rôle et l'utilisation des données fiscales et d'élaborer une stratégie cohérente pour la conception d'enquêtes-entreprises tenues une ou plusieurs fois par an. Ce remaniement a été entrepris car il fallait:

(a) éliminer les faiblesses évidentes qui existent actuellement sur le plan de la qualité des données;
(b) harmoniser les données provenant de différentes enquêtes;
(c) réduire au minimum le fardeau de réponse en utilisant autant que possible les données fiscales;
(d) réduire la quantité de ressources nécessaires à la mise à jour des bases de sondage des enquêtes.

Colledge (1987) donne une description détaillée de ce projet.

Les données fiscales et celles qui sont extraites des comptes de retenues sur la paye jouent un rôle primordial dans la réalisation des enquêtes-entreprises. En vertu de la Loi sur la statistique, Statistique Canada a accès aux déclarations d'impôt annuelles des sociétés (T2) et des particuliers (T1). Les retenues d'impôts que les employeurs font sur la paye de leurs salariés sont également à sa disposition. Ces données remplissent deux fonctions distinctes à Statistique Canada:

(i) elles servent à la mise à jour de la base de sondage (liste d'entreprises);
(ii) elles remplacent les données concernant l'impôt sur le revenu qui pourraient être recueillies au moyen d'une enquête.

4.1 Mise à jour de la base de sondage

La mise à jour de la liste d'entreprises est une tâche complexe pour deux raisons principales: premièrement, la structure d'un grand nombre d'entreprises de même que les liens qui existent entre elles sont loin d'être simples et,deuxièmement, il est difficile de se tenir au courant du nombre élevé de "naissances" et de "morts" qui surviennent dans le secteur des petites entreprises. La définition du terme "entreprise" est elle-même une tâche délicate. Il faut établir une distinction entre la structure juridique (entreprise constituée en société), la structure opérationnelle (comment l'entreprise est organisée pour mener ses activités) et la structure statistique (les unités sur lesquelles on a besoin de données pour effectuer les analyses). À l'intérieur de chaque structure, il est possible d'établir une hiérarchie. Ainsi, la structure statistique se compose, de haut en bas, des entreprises, des sociétés, des établissements et des emplacements. Pour tenir à jour la base de sondage, il ne suffit pas d'y ajouter les nouvelles unités (naissances) et de supprimer celles qui ont mis fin à leurs activités (morts); il faut également tenir compte des changements qui se produisent dans les liens entre les unités des entreprises à structure complexe, et cela veut aussi dire le lien entre la hiérarchie statistique et la hiérarchie opérationnelle.

La stratégie proposée consisterait à faire une mise à jour continue de la structure des entreprises qui dépassent une certaine taille (le seuil variant selon la branche d'activité) et du lien entre cette structure et les unités déclarantes du point de vue de l'impôt. L'activité économique des entreprises touchées représenterait environ 70% de l'activité du secteur auquel elles appartiennent. L'"établissement du profil" est un travail qui permet de déterminer la structure interne des entreprises complexes. Il consiste à interviewer des dirigeants d'entreprise pour prendre connaissance de la structure opérationnelle et établir les unités statistiques adéquates. Le système

de retenue sur la paye (RP) de Revenu Canada est une source importante de renseignements sur les changements qui surviennent dans les entreprises (naissances et réorganisations). La mise en activité d'un nouveau compte de RP par un employeur est le signal qu'il y a du nouveau. On procède alors à un suivi qui permet de savoir si une mise à jour s'impose. Les autres signaux proviennent des déclarations d'impôt annuelles, des réponses données aux enquêtes normales et de l'établissement routine des profils.

On n'essaye pas de déterminer les unités qui composent les petites entreprises ni le lien qui existe entre ces unités, car il y a trop de mouvement dans ce secteur. On se sert carrément des données administratives. Deux listes d'entreprises peuvent servir de base de sondage pour les enquêtes: la première est établie à partir des déclarations d'impôt annuelles les plus récentes et la seconde, à partir des comptes actifs de RP. Toutes les entreprises dont la taille dépasse le seuil fixé sont enlevées de ces listes. Ces dernières se superposent et, pour chaque enquête, on se sert de celle qui convient le mieux. Pour les enquêtes infra-annuelles, on privilégie la liste établie à partir des comptes de RP, qui est plus à jour car il est possible d'ouvrir ou de fermer un compte à n'importe quel moment de l'année. Par contre, on n'y trouve pas les entreprises qui n'ont pas d'employés.

4.2 Remplacement des données d'enquête

Afin de réduire le fardeau de réponse et les coûts, on remplace autant que possible les données d'enquête par les données fiscales. Les notions de base et les définitions sur lesquelles ces dernières sont fondées ne coïncident pas toujours avec celles que les enquêtes doivent utiliser pour des raisons d'uniformité avec le système de comptabilité nationale ou pour les besoins d'analyses. Il faut donc prendre soin d'extraire des déclarations d'impôt seulement les éléments d'information qui correspondent le plus aux définitions de l'enquête. En outre, les données fiscales ne couvrent pas l'éventail de variables que de nombreuses enquêtes-entreprises cherchent à étudier. Notamment, on n'y trouve pas de statistiques sur la production.

Un autre problème que pose l'utilisation des données fiscales concerne l'établissement du lien qui existe entre l'unité pour laquelle on a une déclaration d'impôt et l'unité ou les unités visées par l'enquête. On rencontre ce problème particulièrement dans le cas des grosses entreprises à structure complexe dont il a été question plus tôt.

La stratégie élaborée relativement aux enquêtes annuelles consiste à utiliser les données fiscales principalement pour les petites entreprises où il y a, en général, une déclaration d'impôt par unité d'exploitation. Le fardeau de réponse des petites entreprises est ainsi nettement réduit, et la qualité des données finales n'en souffre pas trop puisqu'on obtient la plus grande partie des renseignements dont on a besoin sur l'activité économique au moyen des déclarations que les grandes entreprises font dans le cadre des enquêtes.

Il ressort clairement de ce bref aperçu des nouvelles stratégie et infrastructure établies pour les enquêtes-entreprises que la poursuite de ce programme est fondamentalement tributaire des données fiscales. Statistique Canada et Revenu Canada doivent donc travailler en collaboration très étroite afin que l'incidence de toute réforme, administrative ou autre, du système fiscal puisse être évaluée avant qu'elle n'entre en vigueur et afin que les mesures nécessaires soient prises à temps.

5. DONNÉES SOCIO-ÉCONOMIQUES TIRÉES DE DOSSIERS ADMINISTRATIFS

À la fin des années 70, on a commencé à faire un effort concerté pour utiliser les données sur les particuliers, les familles et les ménages qui se trouvaient dans les dossiers administratifs. Cette démarche était motivée à l'origine par la croissance du coût des enquêtes et la nécessité

de trouver des solutions de rechange moins onéreuses. Il est vite devenu évident que la meilleure façon de réaliser, au Canada, le potentiel statistique des dossiers administratifs sur les particuliers serait de s'en servir non pas pour remplacer le recensement quinquennal, mais pour le compléter en allant chercher des données pour les petites régions dans ces dossiers pendant la période intercensitaire. Comparativement au recensement, les systèmes de dossiers administratifs actuels ne permettent pas d'obtenir le même niveau de couverture, une aussi grande précision géographique ni un éventail comparable de renseignements sur les caractéristiques des particuliers, des familles et des ménages. On continue néanmoins de chercher à atteindre le même taux de couverture que le recensement en combinant plusieurs systèmes de dossiers administratifs et de voir s'il est possible de remplacer les réponses fournies à certaines questions du recensement par des données tirées de sources administratives.

Nous allons examiner l'utilisation des dossiers administratifs pouvant compléter les données de recensement pendant la période intercensitaire. Les travaux expérimentaux réalisés jusqu'à présent ont porté surtout sur les systèmes administrés à l'échelle nationale (p. ex., l'impôt sur le revenu, l'assurance-chômage (A.-C.), les allocations familiales et la sécurité de la vieillesse) plutôt que sur les systèmes administrés au niveau provincial ou infraprovincial (p. ex., l'assurance-maladie, les permis de conduire et les évaluations municipales). Ces derniers posent, en plus des problèmes propres à l'utilisation de tous les dossiers administratifs, celui du manque d'uniformité entre les secteurs de compétence.

Le fichier des déclarations d'impôt annuelles des particuliers (T1) est la principale source de statistiques sur les Canadiens. Au départ, on s'en est servi pour des totalisations directes qui ont permis de produire des statistiques sur le revenu et l'activité selon l'âge et le sexe, aux niveaux provincial et infraprovincial. Le lieu de résidence des déclarants est déterminé au moyen du code postal indiqué dans le dossier. Un fichier de conversion permettant d'établir la correspondance entre les codes postaux et les diverses divisions géographiques de recensement (province, comté, municipalité, circonscription électorale, etc.) a été créé. Il est également possible de produire des totalisations spéciales pour les régions définies par les utilisateurs en fonction des codes postaux.

Les données ainsi obtenues sont, bien sûr, fondées sur des notions, des définitions et des règlements découlant de la Loi de l'impôt sur le revenu. Les définitions ne sont donc pas nécessairement les mêmes que celles qui permettent d'effectuer les analyses requises (p. ex., il n'y a pas de données sur certains régimes d'aide sociale qui ne sont pas assujettis à l'impôt). Le revenu peut être ventilé selon la source, par exemple le revenu d'emploi peut être séparé. Les variables pouvant faire l'objet d'un classement recoupé sont limitées; (p. ex., l'âge, le sexe et l'état matrimonial). Bien qu'il y ait une question sur la profession dans le formulaire de déclaration d'impôt, la qualité des réponses et de leur codage n'est pas assez bonne pour qu'on puisse s'en servir à des fins statistiques. La couverture se limite aux personnes qui sont tenues de produire une déclaration d'impôt : les personnes à faible revenu et les personnes à charge sont donc sous-représentées. Il se peut que les modifications apportées à la législation fiscale aient, avec le temps, une incidence marquée sur le taux de couverture, lequel a nettement augmenté, entre 1977 et 1978, lorsqu'il a été décidé que les personnes à faible revenu devraient faire une déclaration si elles voulaient obtenir le nouveau crédit d'impôt pour enfants.

Malgré toutes ces réserves, la totalisation directe des données tirées des fichiers fiscaux permet de produire, pendant la période intercensitaire, des données pour les petites régions sur le revenu qui sont fort utiles. Une publication récente de Statistique Canada fait état de statistiques ainsi établies pour les régions de tri d'acheminement, c.-à-d., les régions regroupant les adresses dont les trois premiers caractères du code postal sont les mêmes (Statistique Canada 1987). Étant donné qu'une principale préoccupation, lorsqu'on publie des données pour les petites régions, est de s'assurer qu'il est impossible d'en déduire des renseignements concernant des particuliers, on ne donne pas de statistiques pour les régions où moins de 100 personnes ont fait des déclarations d'impôt.

Les fichiers des déclarations d'impôt des particuliers sont également utilisés pour produire les estimations de la migration annuelle. Pour cela, les dossiers d'une même personne se rapportant à deux années consécutives sont appariés et le code attribué à la division de recensement (ou au comté) dans chaque dossier est comparé. S'il n'est pas le même, on en déduit que le déclarant a déménagé. Les données démographiques et les renseignements relatifs aux exemptions permettent d'estimer le nombre total de personnes qui ont déménagé avec le déclarant. Dans une dernière étape, comme le fichier fiscal ne couvre pas l'ensemble de la population, on procède à un redressement pour produire l'estimation annuelle du nombre total de migrants. Depuis 1981, les chiffres produits en se servant du fichier fiscal sont utilisés dans le cadre du programme des estimations démographiques de Statistique Canada. Norris et Standish (1983) décrivent en détail la méthode utilisée pour estimer la migration au moyen des dossiers fiscaux.

Nous avons vu plus tôt qu'il était possible d'établir des statistiques sur le revenu des particuliers grâce aux données fiscales; le revenu des familles présente encore plus d'intérêt tant du point de vue des possibilités d'analyse que du point de vue des programmes sociaux qui y sont liés. Pour déduire le revenu des familles au moyen des dossiers fiscaux individuels, il faut pouvoir déterminer quels titulaires de dossiers appartiennent à une même famille et ensuite apparier les données qui les concernent. Des travaux de cette nature sont en cours et les résultats sont prometteurs. Pour une description des méthodes employées et des résultats obtenus, voir Auger (1987).

Le système d'assurance-chômage (A.-C.) constitue la deuxième source importante de données administratives sur les particuliers. Statistique Canada a accès au fichier des demandeurs et à celui des bénéficiaires. Ces deux fichiers contiennent les enregistrements de personnes qui, pour diverses raisons, ont peut-être droit aux prestations. Il ne s'agit pas toujours de chômeurs au sens de la définition internationale de chômage retenue par l'enquête sur la population active (EPA), d'où sont tirées les données publiées sur le sujet. S'il y avait moyen de trouver dans le système de l'A.-C. une catégorie se rapprochant de celle de l'EPA, il serait possible de produire à partir de ses fichiers des données sur le nombre de chômeurs dans les petites régions. Cependant, comme dans le meilleur des cas les catégories de l'A.-C. ne sont pas identiques à celles de l'EPA, il faut donc plutôt chercher à harmoniser ou concilier ces deux sources de données. Notamment, les chiffres mensuels de l'A.-C. sur les petites régions pourraient être pris comme indicateurs de la variation du chômage à l'échelon local et ils pourraient ensuite être corrigés en fonction d'estimations fiables de l'EPA correspondant à un niveau géographique supérieur (p. ex. une province). Plusieurs méthodes d'estimation de ce genre ont été examinées (comme la méthode de l'estimation par régression et celle de l'estimation par le quotient avec maintien de la structure), mais il n'a pas encore été possible de déterminer laquelle est la meilleure. Trottier et Choudhry (1985) donnent une description des travaux effectués dans ce domaine et Feeney (1987) parle d'une approche semblable utilisée dans les travaux australiens. L'utilisation de modèles à séries chronologiques exploitant la structure corrélée de l'erreur dans le temps est une méthode qui semble très prometteuse (Choudhry et Hidiroglou 1987).

Les exemples que nous venons de donner montrent qu'on utilise principalement les données tirées des dossiers administratifs sur les particuliers pour effectuer des totalisations directes ou des estimations, contrairement aux données tirées des dossiers administratifs concernant les entreprises, dont on se sert surtout pour mettre à jour la base de sondage ou carrément remplacer les données d'enquête.

Bien que les domaines d'activité auxquels ces deux exemples ont trait soient en pleine expansion et revêtent beaucoup d'importance à Statistique Canada, ils ne touchent qu'une petite fraction du nombre de fichiers administratifs dont on se sert au Bureau. À titre d'exemple, les dossiers administratifs sont depuis longtemps universellement utilisés en statistique sociale

(éducation, santé, justice) pour la création de bases de sondage et pour l'obtention de données. Les responsables des recherches en cours sur les enquêtes téléphoniques et les registres d'adresses utilisent les dossiers administratifs pour dresser les listes de logements ou de ménages. Une enquête interne récente a dénombré plus d'une cinquantaine de systèmes administratifs actuellement exploités à des fins statistiques. Tous les types de dossiers et toutes les catégories d'utilisation mentionnés aux sections 2 et 3 sont représentés, y compris des registres de maladies et des dossiers renfermant des données sur les véhicules automobiles immatriculés, les mouvements d'aéronefs, la commercialisation du lait, la taxe de vente sur le carburant, les permis de construire municipaux et les droits de douane et d'accise.

6. ACCÈS AUX SYSTÈMES ADMINISTRATIFS ET DROIT DE REGARD SUR LEUR CONTENU

Le tour d'horizon que nous venons de faire de l'utilisation des dossiers administratifs à des fins statistiques montre bien que ces dossiers constituent une source d'information indispensable pour de nombreux programmes de Statistique Canada. Cela nous amène aux mesures que le Bureau peut prendre pour protéger ces sources de données et peut-être rendre ces dernières plus utiles du point de vue statistique. Cette section porte sur les deux principaux problèmes, c'est-à-dire l'accès aux dossiers administratifs et la manière dont on peut influer leur contenu, leur conception et la marche à suivre à leur égard.

6.1 Accès

Le fondement juridique de l'accès aux dossiers administratifs est l'article 12 de la Loi sur la statistique (1971) qui se lit comme suit:

> Une personne ayant la garde ou la charge de documents ou archives conservés dans un département ou dans un bureau municipal, une corporation, entreprise ou organisation et dont on pourrait tirer des renseignements que l'on cherche à obtenir pour les objets de la présente loi ou qui aideraient à compléter ou à corriger ces renseignements, doit en permettre l'accès, à ces fins, à une personne autorisée par le statisticien en chef à obtenir ces renseignements ou cette aide pour le complètement ou la correction de ces renseignements.

Cette disposition, qui semble donner un droit assez étendu à l'accès, comporte cependant des restrictions. Dans certains cas, les lois régissant le processus administratif limitent l'accès aux données administratives ou leur utilisation secondaire. Il en résulte une incompatibilité entre les lois qui, au mieux, a pour effet de retarder les négociations au sujet de l'accès. Dans d'autres cas, l'accès à des fins statistiques est permis de façon explicite.

L'adoption de lois facilitant l'accès aux dossiers administratifs est une condition nécessaire mais non suffisante de l'utilisation productive des dossiers administratifs. Il serait probablement beaucoup plus efficace, lorsqu'on cherche à obtenir l'accès aux dossiers administratifs, d'essayer d'y arriver par la voie de la collaboration (sur le plan de l'élaboration et de l'utilisation de ces dossiers dans un but statistique) que d'avoir recours à des dispositions ou à des sanctions légales. En effet, une fois l'accès obtenu, il est possible de parvenir à l'étape suivante (qui consiste à influer sur la conception ou les méthodes d'utilisation) seulement s'il existe un esprit de collaboration entre l'organisme administratif et l'organisme statistique.

L'accès aux dossiers administratifs dont jouit le Bureau est strictement un phénomène à sens unique. Les microdonnées individuelles vont de l'organisme administratif à l'organisme statistique, et seulement des données agrégées, protégées par la confidentialité, peuvent retourner en sens inverse. La seule exception technique à cette règle se produit lorsque l'organisme administratif dépend de l'organisme statistique pour organiser, préparer, mettre en forme, traiter ou restructurer ses dossiers et que les microdonnées originales sont renvoyées sous une autre forme à l'organisme fournisseur.

6.2 Droit de regard sur les changements

Nous avons déjà mentionné l'incidence que des modifications apportées aux pratiques ou règlements administratifs auraient sur les statistiques produites. Il suffirait qu'on change la clientèle d'un programme, qu'on introduise une mesure incitant les personnes visées par un programme à y participer ou au contraire à ne plus y participer ou bien qu'on modifie la procédure de manière à altérer la qualité ou l'exhaustivité des dossiers, et les séries chronologiques seraient interrompues. L'organisme statistique doit donc être sur ses gardes et réagir chaque fois qu'un changement imposé de l'extérieur risque de se produire.

Il y a, en revanche, des changements que l'organisme statistique aimerait bien voir se réaliser. Le statisticien qui souhaite utiliser des dossiers administratifs éprouve souvent un sentiment de frustration en pensant que ceux-ci seraient beaucoup plus utiles si seulement de petites modifications y étaient apportées. Par exemple, l'ajout d'une nouvelle question, l'utilisation d'une notion différente, l'inclusion de nouveaux sous-groupes ou l'exécution d'une nouvelle vérification de la qualité pourraient améliorer sensiblement la valeur statistique des dossiers. Par contre, pourquoi l'organisme administratif envisagerait-il d'apporter des changements qui ne sont pas requis par le processus administratif, quand ces changements contribuaient probablement dans une certaine mesure à augmenter les coûts et la complexité de ce processus?

L'organisme statistique a donc un défi à relever : celui de persuader les administrateurs que les avantages découlant de ces changements seraient plus grands que les coûts supplémentaires générés. La difficulté vient de ce que ce n'est pas nécessairement le ministère responsable du système administratif qui profite des avantages en question, mais plutôt les ministères décisionnaires ou d'autres utilisateurs qui se servent des données à des fins statistiques.

Il est habituellement plus facile de tenir compte des exigences statistiques au moment de concevoir un système que de modifier un système déjà en cours d'exploitation. Par conséquent, un mécanisme qui permettrait de tenir compte des besoins en données statistiques au stade de la conception ou de la restructuration en profondeur d'un système administratif serait préférable à un mécanisme qui tenterait seulement d'adapter les systèmes en place. En ce moment, au Canada, on parle beaucoup de la réforme de la taxe que le gouvernement songe à instituer. Cette réforme pourrait changer significativement la collecte des données sur les entreprises au Canada. La participation de statisticiens à la conception d'un système comme celui-là ajouterait beaucoup aux avantages statistiques qu'on pourrait en retirer. Bien entendu, l'établissement d'un nouveau système administratif est un événement relativement rare, de sorte qu'il faut également adapter les systèmes en place si l'on veut obtenir des avantages statistiques à court terme. Par ailleurs, le fait qu'il soit relativement exceptionnel qu'on conçoive ou qu'on restructure des systèmes administratifs importants renforce l'argument selon lequel il ne faut rater aucune occasion de prendre part à ce genre de travaux.

6.3 Mécanismes

Il existe divers moyens ou mécanismes permettant à l'organisme statistique d'avoir un certain droit de regard sur les sytèmes administratifs et d'y avoir accès; certains de ces mécanismes sont essentiellement bilatéraux, mettant en jeu l'organisme statistique et un service administratif quelconque, tandis que d'autres concernent l'ensemble du secteur public. En voici quelques exemples:

(i) créer des comités bilatéraux constitués de hauts fonctionnaires qui examineraient les questions intéressant les deux organismes, notamment les problèmes relatifs à la fourniture de données administratives;

(ii) fournir à l'organisme administratif les données statistiques afin de montrer à la fois l'utilité des données et, le cas échéant, les lacunes attribuables aux pratiques administratives;

(iii) offrir des conseils ou des services techniques au service statistique de l'organisme administratif;

(iv) adopter une politique officielle en matière de collecte de renseignements exigeant que tout projet de collecte de données (dans un but statistique ou administratif) soit soumis à l'examen d'un organisme central;

(v) demander que chaque nouvelle proposition de programme soit assortie d'un plan prévoyant la façon d'obtenir les renseignements statistiques nécessaires pour contrôler et évaluer le programme en question;

(vi) promouvoir l'utilisation des définitions statistiques standard (p. ex., ce que l'on entend par les termes "famille", "établissement commercial" ou "chômeur") dans les systèmes administratifs;

(vii) faire en sorte que les vérificateurs des activités gouvernementales recommandent l'utilisation des dossiers administratifs comme moyen économique par excellence de collecte de renseignements;

(viii) suivre une orientation politique favorisant un recours accru à certains systèmes administratifs ou la recherche de solutions de rechange aux enquêtes;

(ix) supprimer les obstacles législatifs qui limitent l'accès aux dossiers administratifs ou à leur utilisation à des fins statistiques.

L'expérience vécue par le Bureau dans ses relations avec les autres ministères fédéraux a été particulièrement fructueuse lorsque des ententes bilatérales étroites ont été conclues. La création de comités supérieurs bilatéraux au début des années 80 a favorisé et parfois assuré l'élaboration de telles ententes. Les mesures prises à l'échelle de l'appareil gouvernemental, par exemple la gestion de l'information et la planification statistique, ont moins bien réussi à favoriser l'utilisation des dossiers administratifs. Les vérifications des opérations gouvernementales et les directives du cabinet ont bien donné une certaine impulsion à des activités visant à accroître l'utilisation des données administratives, mais l'augmentation de cette utilisation dépend elle-même, encore une fois, de l'existence de liens de travail étroits avec certains ministères. Bien qu'il convienne de décrire l'organisme statistique comme un organisme progressif qui tente de briser les barrières irrationnelles faisant obstacle à l'utilisation des données administratives, il faut quand même également reconnaître qu'il puisse exister au sein de l'organisme même une certaine résistance au changement. Les membres du personnel dont les carrières ont été vouées à la conception et à la réalisation d'enquêtes peuvent avoir besoin d'arguments convaincants pour admettre que les restrictions budgétaires et les besoins en données nous obligent à présent à combiner les enquêtes et d'autres méthodes.

Comme les remarques qui ont précédé s'appliquent uniquement aux systèmes administrés par des ministères fédéraux, il convient d'ajouter quelques mots au sujet des dossiers provinciaux. Certaines des mesures mentionnées conviendraient aussi aux dossiers administrés par les provinces; toutefois les rapports avec les administrations infranationales posent un problème fondamental, celui de la conformité à des normes communes. La différence entre les besoins et les priorités des provinces, accentuée par des possibilités technologiques croissantes, entraînera une diversité de plus en plus grande des systèmes administratifs s'il n'existe pas de force centralisatrice. Dans le passé, le Bureau a eu recours à divers mécanismes pour essayer de favoriser une certaine uniformisation, mais avec plus ou moins de succès. Comme dans le cas des services fédéraux qui ont la garde de dossiers administratifs, l'avantage réciproque doit être l'argument principal. Il y a des comités fédéraux-provinciaux dans plusieurs domaines. Le Conseil de la statistique de l'état civil, composé d'officiers provinciaux de l'état civil et de représentants du Bureau, est un exemple d'une longue et fructueuse collaboration. Des comités comme celui-là ont, dans le passé, élaboré des conventions pour la déclaration de certains éléments de données et ils en ont suivi l'application. Par exemple, le système de déclaration des finances municipales a été conçu à la suite de réunions fédérales-provinciales sur les statistiques financières des municipalités.

7. CONFIDENTIALITÉ, PROTECTION DES RENSEIGNEMENTS PERSONNELS ET RELATIONS AVEC LE PUBLIC

Même si l'organisme statistique a le pouvoir légal d'exploiter les dossiers administratifs et même si les organismes administratifs fournissent ces dossiers dans un esprit de grande collaboration, il faut faire très attention à la perception que le public a de l'utilisation des dossiers administratifs à des fins autres que celles pour lesquelles ces dossiers ont été conçus. Comme l'efficacité, sinon la survie, d'un organisme statistique dépend en très grande partie du maintien de la collaboration et de la confiance des répondants, l'organisme en question ne doit pas s'engager à la légère dans une activité qui pourrait saper cette collaboration ou cette confiance.

Dans beaucoup de pays, le public est, depuis quelques années, de plus en plus sensible à tout ce qui concerne la protection des renseignements personnels et, par conséquent, à l'accès à l'information et au contrôle de cet accès. Au Canada, l'adoption de la Loi sur la protection des renseignements personnels en 1982 témoigne de cette préoccupation grandissante. Cette loi dit, entre autres choses et avec certaines exceptions, qu'un répertoire de toutes les banques de renseignements personnels relevant d'institutions gouvernementales fédérales doit être publié périodiquement, que toutes les personnes ont le droit d'avoir accès aux renseignements les concernant contenus dans ces banques de renseignements et que les renseignements personnels ne peuvent servir qu'à des usages qui sont compatibles avec les fins pour lesquelles ils ont été obtenus. Une des exceptions prévues à cette dernière disposition est que des renseignements personnels peuvent être divulgués

> "... à toute personne ou à tout organisme, pour des travaux de recherche ou de statistique, pourvu que ... les fins auxquelles les renseignements sont communiqués ne peuvent être normalement atteintes que si les renseignements sont donnés sous une forme qui permette d'identifier l'individu qu'ils concernent, (et) la personne ou l'organisme (s'engagent) par écrit (auprès du responsable de l'institution) à s'abstenir de toute communication ultérieure des renseignements tant que leur forme risque vraisemblablement de permettre l'identification de l'individu qu'ils concernent." (Loi sur la protection des renseignements personnels 1982 alinéa 8(2) (j)).

Cette disposition régit l'utilisation des dossiers administratifs à des fins statistiques du point de vue de la Loi sur la protection des renseignements personnels. L'article en question est toutefois formulé sous réserve de toute autre loi du Parlement, de sorte que si une disposition d'une loi régissant le processus administratif interdisait une telle utilisation, celle-ci aurait préséance.

Bien qu'il soit reconnu dans la Loi sur la protection des renseignements personnels et dans d'autres lois que les travaux statistiques constituent, dans certaines conditions, une utilisation secondaire légitime des dossiers administratifs, cela ne dissipe pas pour autant les préoccupations du public au sujet d'une utilisation possible des banques de données qui irait à l'encontre des intérêts des particuliers. Il est douteux que les citoyens moyens fassent la différence entre l'utilisation statistique, où l'identité du titulaire de chaque dossier ne présente plus aucun intérêt, et l'utilisation administrative, où il est essentiel de savoir à qui chaque dossier se rapporte. Il serait plus facile d'expliquer cette différence et de s'en servir si l'on pouvait dire sans équivoque que, pour le genre de travaux qui sont faits en statistique, on n'a jamais besoin d'identificateurs. Malheureusement, ce n'est pas le cas. Plusieurs techniques statistiques tout aussi légitimes exigent qu'on se serve d'identificateurs au cours des manipulations intermédiaires des données. Ces techniques supposent toutes une forme quelconque d'appariement des données provenant de différents fichiers ou de différentes sources, et les identificateurs permettent d'apparier les bons dossiers ensemble. Une fois cette opération effectuée, les dossiers peuvent être rendus anonymes si aucun couplage subséquent n'est prévu. Citons, à titre d'exemples de cas où une forme d'identification est nécessaire, le recensement de la population (pour veiller à ce que la couverture soit exhaustive et s'en

assurer par la suite), les études longitudinales pour lesquelles on utilise des dossiers administratifs, les enquêtes épidémiologiques et enfin les études d'évaluation visant à vérifier les réponses dans des questionnaires d'enquête en les comparant à des données de sources administratives. Le fait de devoir expliquer pourquoi on a besoin d'identificateurs alors que l'identité des personnes auxquelles les données se rapportent ne présente aucun intérêt est un défi de taille pour l'organisme statistique.

Une autre source de préoccupation a trait à l'engagement même de respecter la confidentialité. Même si Statistique Canada a toujours été irréprochable en matière de protection de la confidentialité, il y a sans doute des répondants qui sont sceptiques quant à la protection accordée aux renseignements qui les concernent. Il est possible que l'utilisation de recenseurs connus des répondants, en particulier dans les petites localités, ajoute à leur scepticisme. Certains répondants semblent occuper comme principe qu'il se fait beaucoup d'échanges de renseignements entre les ministères fédéraux et, dans certains cas, ne font de distinction entre les différents ministères.

Une autre inquiétude encore aurait trait non pas à la confiance dont seraient dignes les services ayant actuellement la garde de banques de renseignements, mais découlerait de la crainte de n'avoir aucune protection contre une violation éventuelle de la confidentialité, soit par des moyens illégaux soit par des personnes élues qui auraient un point de vue différent sur la protection des renseignements personnels. Pour protéger le public contre cette éventualité, il faudrait qu'on élimine des bases de données statistiques toute information permettant d'identifier les personnes dont il y est question.

Cette préoccupation du public au sujet de la protection et du traitement des renseignements personnels oblige l'organisme statistique à penser à des mesures susceptibles de prévenir ou de réduire les réactions défavorables du public à l'utilisation légitime des dossiers administratifs pour des travaux statistiques. Comme il s'agit là essentiellement d'une question de perception du public, il importe que l'organisme statistique ne cache rien de ses pratiques, et que le public soit bien au courant des mesures en application, notamment:

a) L'information publique destinée aux répondants et aux utilisateurs doit toujours mettre l'accent sur l'importance accordée à la confidentialité de toutes les données individuelles (microdonnées) acquises par l'organisme statistique.
b) Le fait que les microdonnées ne circulent que dans un sens devrait être souligné. L'organisme statistique reçoit des microdonnées, mais il n'en ressort que des données agrégées ou résumées, dont la confidentialité est garantie par leur nature. Cela est vrai aussi bien pour les données de recensement ou les données d'enquête que pour les données tirées de dossiers administratifs.
c) Il faudrait faire valoir les avantages qui découlent de l'utilisation des dossiers administratifs : réduction du fardeau de réponse et allègement du fardeau fiscal pour les contribuables. Ces affirmations devraient s'appuyer sur des chiffres et des faits.
d) Des règles explicites en matière de jumelage d'enregistrements énonçant les conditions dans lesquelles l'organisme statistique peut exercer cette activité permettrait, d'une part, de montrer que le jumelage d'enregistrements ne se fait ni à la légère ni sans contrôle et, d'autre part, de ne pas donner suite aux demandes de jumelage d'enregistrements qui ne rempliraient pas les conditions requises.
e) En vertu de la Loi sur la protection des renseignements personnels, les individus doivent être informés des fins auxquelles les renseignements personnels les concernant sont recueillis. Il faudrait inciter les organismes administratifs à indiquer de façon précise l'utilisation statistique dont ces renseignements pourraient faire l'objet. Même si l'utilisation de données à des fins statistiques constitue une utilisation secondaire acceptable des dossiers administratifs, la mention explicite de ces fins sur le formulaire de collecte permettrait d'éviter les mauvaises surprises.

f) Il faudrait que les moyens pris pour assurer la protection matérielle dont l'utilisation des dossiers administratifs délicats fait l'objet soient évidents, et cette protection devrait peut-être même être plus grande que celle qui est habituellement exercée au Bureau. À Statistique Canada, les divisions qui sont les principales détentrices des données fiscales sont logées dans des locaux à accès limité, situés dans des bâtiments à l'entrée desquelles on effectue des vérifications de sécurité.

g) Le fait que les fichiers statistiques ne sont pas soumis à l'examen des services de sécurité ou de renseignement permet en grande partie d'entretenir le sentiment de confiance que le public éprouve à l'égard de la confidentialité absolue des données fournies à l'organisme statistique. Lorsque le Service canadien du renseignement de sécurité a été créé en 1983, Statistique Canada a obtenu une telle exemption pour ses données (la seule accordée à un organisme fédéral).

Les points que nous venons de mentionner correspondent à des mesures précises qui peuvent être prises pour éviter que le public ne réagisse mal à l'utilisation des dossiers administratifs ou pour contrer une telle réaction une fois qu'elle a eu lieu, mais ce dont l'organisme statistique a surtout besoin, c'est d'un appui politique fort. Le capital politique à gagner d'une réduction visible des coûts et du fardeau de réponse, conjugué à des garanties politiques fermes de protection des données des particuliers, constitue une base solide sur laquelle les politiciens peuvent s'appuyer pour dissiper les inquiétudes du public au sujet de l'utilisation des dossiers administratifs à des fins statistiques. En même temps, ils doivent démentir immédiatement et de façon non équivoque toute suggestion selon laquelle les dossiers statistiques sont utilisés à des fins administratives.

8. CONCLUSION

Les dossiers administratifs sont et demeureront une source de plus en plus utile de données statistiques. Les points forts et les points faibles relatifs des données tirées des systèmes administratifs, du point de vue des coûts, de l'univers visé, et de leur qualité, utilité et actualité par rapport aux données du recensement ou aux données d'enquête dictent la façon la plus efficace d'utiliser ces différentes sources de données. Entre autres utilisations courantes, il y a les totalisations directes, les estimations indirectes, le remplacement des réponses d'enquête, la constitution et la mise à jour de bases de sondage et l'évaluation de données. Ces utilisations sont aujourd'hui assez répandues dans la plupart des programmes statistiques et elles le seront probablement encore davantage à l'avenir.

Au Canada, les dossiers administratifs font à présent partie intégrante de l'appareil statistique. C'est en partie grâce à leur utilisation que Statistique Canada a pu maintenir ses programmes malgré la compression de ses budgets. Du fait même, le fardeau de réponse a été réduit et on a commencé à produire de nouvelles séries de données ou à produire les anciennes plus souvent. Comme nous ne disposons pas de registres administratifs, il nous a fallu porter une attention très particulière aux questions de couverture et d'utilisation combinée de données administratives et de données d'enquête pour nous assurer que les estimations de totaux pour un univers soient valides. Le recours à des techniques de jumelage d'enregistrements, bien que nécessitant un contrôle rigoureux, s'est avéré d'une très grande utilité, notamment en ce qui a trait aux données sur les entreprises, aux études longitudinales du marché du travail et aux enquêtes épidémiologiques.

En raison de leur utilisation croissante des dossiers administratifs, les organismes statistiques sont de plus en plus tributaires des autres organismes pour la fourniture continue des données dont ils ont besoin pour leurs programmes. Quels que soient les textes de loi et les politiques qui régissent les activités de l'organisme statistique, l'établissement d'ententes de

collaboration étroite avec les organismes fournisseurs est extrêmement important. La capacité de l'organisme statistique d'influer sur la conception ou la refonte de systèmes administratifs repose sur une compréhension mutuelle des besoins des deux organismes intéressés. Si l'on pouvait adopter, à l'échelle de l'appareil gouvernemental, une politique ou un principe donnant à Statistique Canada voix au chapitre de la conception des systèmes administratifs ou, plus généralement, de la façon de satisfaire les besoins statistiques des nouveaux programmes, cela faciliterait la tâche de l'organisme statistique mais ne remplacerait pas la création d'un climat de collaboration étroite avec les organismes administratifs.

Il est possible d'envisager divers mécanismes qui permettraient à l'organisme statistique d'avoir accès aux dossiers administratifs et d'influer sur leur conception à l'intérieur de l'appareil gouvernemental. L'applicabilité et l'efficacité de chaque mécanisme dépendront du contexte législatif et politique ainsi que du mandat et du statut de l'organisme statistique au sein de l'appareil gouvernemental. D'après l'expérience de Statistique Canada, l'établissement de relations de travail bilatérales étroites avec les ministères administratifs, fondées sur le principe des avantages réciproques, est l'approche la plus efficace. Il importe que l'utilisation des dossiers administratifs reçoive un appui politique, et cet appui s'est manifesté dans les décisions récentes du gouvernement relativement aux compressions budgétaires.

REMERCIEMENTS

Pour cette communication, nous nous sommes inspirés des idées et des travaux d'un bon nombre de personnes à Statistique Canada. Nous y avons notamment incorporé des passages extraits du document rédigé en collaboration avec Michael Colledge, Ivan Fellegi et John Leyes.

BIBLIOGRAPHIE

AUGER, E. (1987). Family data from the Canadian personal income tax file. Article présenté à la réunion annuelle de 1987 de l'American Statistical Association.

BRACKSTONE, G.J. (1984). Répercussions des changements technologiques sur le recensement. *Estadistica*, 36, 43-60.

CANADA, Loi sur la statistique (1971). *Statuts du Canada 1970-71-72*, c.15.

CANADA, Loi sur la protection des renseignements personnels (1982). *Statuts du Canada 1980-81-82*, c.11.

CHOUDHRY, G.H., et HIDIROGLOU, M.A. (1987). Small area estimation: Some experiences at Statistics Canada. *Proceedings of the 46th Session of the International Statistical Institute*, (en voie de rédaction).

COLLEDGE, M.J. (1987). The Business Survey Redesign Project: Implementation of the new strategy at Statistics Canada. *Proceedings of the Third Annual Research Conference, US Bureau of the Census*, (en voie de rédaction).

FEENEY, G.A. (1987). The estimation of the number of unemployed at the small area level. Dans *Small Area Statistics, An International Symposium*, (eds. R. Platek, J.N.K. Rao, C. Särndal et M.P. Singh), New York: John Wiley.

NORRIS, D.A., et STANDISH, L.D. (1983). A technical report on the development of migration data from tax records. Document de travail, Statistique Canada.

ROWEBOTTOM, L.E. (1978). The utilization of administrative records for statistical purposes. *Survey Methodology*, 4, 1-15.

STATISTIQUE CANADA (1987). *Tableaux sommaires pour RTA urbaines et codes postaux ruraux.* No. de catalogue 17-602, Statistique Canada.

TROTTIER, I., et CHOUDHRY, G.H. (1985). Model based unemployment estimates for small areas. Dans *Small Area Statistics, An International Symposium ('85 Contributed Papers)*, (éds. R. Platek et M.P. Singh), Laboratory for Research in Statistics and Probability, Carleton University/ Université d'Ottawa.