



N° 31-533-XIF au catalogue

# Enquête mensuelle sur les industries manufacturières : utilisation des données administratives



## Comment obtenir d'autres renseignements

Toute demande de renseignements au sujet du présent produit ou au sujet de statistiques ou de services connexes doit être adressée à : Section de marketing et diffusion, Division de la fabrication, de la construction et de l'énergie, Statistique Canada, Ottawa, Ontario, K1A 0T6, (téléphone: (613) 951-9497, sans frais: 1-866-873-8789, télécopieur. : (613) 951-9499, courriel : [manufact@statcan.ca](mailto:manufact@statcan.ca)).

Pour obtenir des renseignements sur l'ensemble des données de Statistique Canada qui sont disponibles, veuillez composer l'un des numéros sans frais suivants. Vous pouvez également communiquer avec nous par courriel ou visiter notre site Web.

<b>Service national de renseignements</b>	<b>1 800 263-1136</b>
<b>Service national d'appareils de télécommunications pour les malentendants</b>	<b>1 800 363-7629</b>
<b>Renseignements concernant le Programme des bibliothèques de dépôt</b>	<b>1 800 700-1033</b>
<b>Télécopieur pour le Programme des bibliothèques de dépôt</b>	<b>1 800 889-9734</b>
<b>Renseignements par courriel</b>	<b><a href="mailto:infostats@statcan.ca">infostats@statcan.ca</a></b>
<b>Site Web</b>	<b><a href="http://www.statcan.ca">www.statcan.ca</a></b>

## Renseignements sur les commandes et les abonnements

Le produit n° 31-533-XIF au catalogue est gratuit sur Internet. Les utilisateurs sont priés de se rendre à [http://www.statcan.ca/cgi-bin/downpub/freepub\\_f.cgi](http://www.statcan.ca/cgi-bin/downpub/freepub_f.cgi)

## Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois, et ce, dans la langue officielle de leur choix. À cet égard, notre organisme s'est doté de normes de service à la clientèle qui doivent être observées par les employés lorsqu'ils offrent des services à la clientèle. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1 800 263-1136.



Statistique Canada

Division de la fabrication, de la construction et de l'énergie

Section de l'enquête mensuelle sur les industries manufacturières

## **Enquête mensuelle sur les industries manufacturières : utilisation des données administratives**

Publication autorisée par le ministre responsable de Statistique Canada

© Ministre de l'Industrie, 2004

Tous droits réservés. Il est interdit de reproduire ou de transmettre le contenu de la présente publication, sous quelque forme ou par quelque moyen que ce soit, enregistrement sur support magnétique, reproduction électronique, mécanique, photographique, ou autre, ou de l'emmagasiner dans un système de recouvrement, sans l'autorisation écrite préalable des Services de concession des droits de licence, Division du marketing, Statistique Canada, Ottawa, Ontario, Canada K1A 0T6.

Octobre 2004

N° 31-533-XIF au catalogue

Périodicité: hors-série

ISBN: 0-662-78023-X

Ottawa

This publication is available in English (Catalogue no. 31-533-XIE)

### **Note de reconnaissance**

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population, les entreprises, les administrations canadiennes et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques précises et actuelles.

## **L'Enquête mensuelle sur les industries manufacturières : utilisation des données administratives**

L'Enquête mensuelle sur les industries manufacturières (EMIM) figure parmi plusieurs enquêtes-entreprises menées par Statistique Canada (SC) qui produisent des estimations évaluant l'activité économique au Canada. Ces renseignements sont essentiels au processus décisionnel des gouvernements et du secteur privé. Le fardeau du répondant est un problème important auquel SC doit faire face pendant une enquête. Pour atténuer ce fardeau, SC a déployé de grands efforts au cours des dernières années afin de rationaliser les données recueillies et de substituer les données administratives aux données d'enquêtes, dans la mesure du possible.

L'EMIM est une des enquêtes dans le cadre desquelles on a tenté d'utiliser des données administratives existantes (données fiscales) comme substitut, données qui se sont avérées être des remplaçantes convenables des données d'enquêtes. L'objectif du projet de la taxe sur les produits et services (TPS) a été d'exploiter le plein potentiel des données administratives existantes afin de modéliser les données d'enquête afin de réduire le fardeau du répondant et de réduire les coûts associés à la collecte.

À ce titre, à partir du mois de référence d'août 2004, l'EMIM utilisera les données d'enquête et les données administratives modélisées à partir des déclarations de TPS afin d'estimer les niveaux de livraison de produits manufacturiers. Voici une brève description de la méthode utilisée dans ce processus.

### **À propos de la TPS et de l'Enquête mensuelle sur les industries manufacturières**

Mise en place en 1991, la TPS est une taxe fédérale prélevée sur la consommation de produits et de services au Canada. Dans toutes les provinces, sauf le Québec, elle est perçue par l'Agence du revenu du Canada (ARC). Toutes les provinces, sauf Terre-Neuve-et-Labrador, la Nouvelle-Écosse et le Nouveau-Brunswick, perçoivent cette taxe en ajoutant 7 % de la valeur des ventes à la facture. À Terre-Neuve-et-Labrador, en Nouvelle-Écosse et au Nouveau-Brunswick, la taxe est imposée sous forme d'une taxe de vente harmonisée (TVH) de 15 % qui comprend la TPS et la taxe provinciale de chaque province.

Toutes les entreprises, à l'exception de celles dont le revenu est inférieur à 30 000 \$, doivent effectuer des versements de la TPS. Les entreprises qui affichent des revenus annuels supérieurs à 6 millions de dollars annuellement doivent effectuer des versements mensuels. Celles dont les revenus se situent entre 500 000 \$ et 6 millions de dollars doivent le faire à chaque trimestre. Celles qui génèrent des revenus se situant entre 30 000 \$ et 500 000 \$ font des paiements annuels. Les entreprises qui effectuent des versements aux mois et aux trimestres doivent le faire au plus tard 30 jours après la fin de la période, tandis que celles qui font des envois annuels disposent d'un délai de trois mois après la fin de la période.

L'ARC envoie le fichier de la TPS à la Division des données fiscales (DDF) de Statistique Canada. La division réalise un traitement supplémentaire, seulement pour des raisons statistiques. Le traitement fait en sorte que la base de données à laquelle accéderont les différents programmes d'enquêtes-entreprises soit propre et complète. Le traitement effectué par la Division des données fiscales comprend la correction des données inexactes, la détection d'observations aberrantes et le remplacement de données manquantes au moyen d'une calendrialisation et d'une extrapolation.

## Enquête mensuelle sur les industries manufacturières <sup>1</sup>

L'EMIM est une enquête par sondage qui permet d'obtenir des renseignements sur les livraisons, les stocks et les commandes de tous les établissements manufacturiers canadiens. Chaque mois, des données sont recueillies sur :

- les livraisons (produits de propre fabrication (PPF))
- les stocks (matières premières (MP), produits en cours (PEC), produits finis (PF))
- les commandes (commandes en carnet (CC))

Afin de diminuer le fardeau de réponse et de réduire les coûts de collecte, les plus petites unités de la population observées (qui représentaient la marge inférieure des livraisons totales de chaque province, soit 2 %) étaient exclues. On a ainsi pu réduire la base de sondage, qui est passée d'environ 104 000 établissements à 35 000 établissements.

L'échantillon de l'EMIM est réparti en fonction de l'industrie, de la province et de la taille (basée sur la valeur totale des livraisons). Un échantillon d'environ 10 000 unités est tiré sur les 35 000 unités qui restent dans la base de sondage. L'échantillon reste le même de mois en mois, sauf dans le cas des nouvelles unités (naissances), qui sont échantillonnées au moyen de la même probabilité qui est utilisée pour les unités de la population originale.

### Remplacement possible des données de l'enquête par celles provenant de la TPS

L'échantillon de l'EMIM comporte deux types d'unités : les unités simples et les unités complexes. En ce qui concerne les unités complexes (unités formées de plus d'un établissement), il arrive souvent qu'une seule valeur de la TPS corresponde à tous les établissements, ne permettant donc pas une répartition par province ou par classe du SCIAN. L'échantillon de l'EMIM comprenait 4 100 unités complexes en vie qui ne sont pas éligibles pour le remplacement avec les données TPS. Les unités qui restent dans l'échantillon sont simples (représentant qu'un seul établissement) et sont éligibles d'être remplacés (approximativement 5 008 unités simples en vie).

La définition d'une unité simple s'applique à la fois aux petits et aux grands établissements. Par conséquent, ce ne sont pas toutes les unités désignées comme simples qui peuvent être remplacées. Pour faire en sorte que l'utilisation des données de la TPS n'ait pas d'incidence considérable sur les estimations en général, seules les unités qui sont simples, en vie et qui ne sont pas comprises dans les unités qui représentent la portion supérieure, soit 80 %, du SCIAN par province, en ce qui concerne les livraisons, peuvent être remplacées. Ainsi, les établissements dominants ou de grande taille ne seront pas exclus de l'enquête.

### Lien entre l'Enquête mensuelle sur les industries manufacturières et la base de données de la TPS

Il fallait établir un lien entre le fichier de la TPS et l'échantillon de l'EMIM afin de pouvoir remplacer les données sur la livraison de l'EMIM par les données fiscales modélisées. On a déterminé que ce lien était le Numéro d'entreprise (NE) et on a effectué un certain nombre de tests pour découvrir combien d'unités de l'échantillon étaient liées de façon individuelle à un dossier de la TPS. Il en a été ainsi pour la plupart des unités de l'échantillon. Il y a cependant eu des situations où un établissement n'avait accumulé aucune recette de la TPS alors que ses données de livraisons étaient supérieures à zéro, et vice versa. Ces cas ont été étudiés en détail afin de trouver la cause de la différence.

---

1. Pour plus d'information sur la méthodologie de l'EMIM consultez **Définitions, source de données et méthodes : numéro d'enquête [2101](#)**.

Les unités de substitution d'août 2004 ont été choisies de sorte qu'elles correspondent toutes à un dossier de la TPS.

### **Corrélations entre les livraisons de l'EMIM et les recettes de la TPS**

Afin d'utiliser les données modélisées des recettes de la TPS comme substitut convenable aux données sur la livraison de l'EMIM, il fallait trouver une bonne corrélation entre les livraisons manufacturières et les recettes de la TPS. Comme l'illustre le tableau 1, les corrélations sont de bonne qualité et s'améliorent même lorsqu'on élimine les valeurs nulles et les observations aberrantes.

**Tableau 1. Corrélations entre les livraisons (juillet 2004) et les recettes (juillet 2004)**

Types des unités	Corrélation %
Total	0,826
Unités rapportant 0 \$ enlevés	0,834
Unités aléatoires et rapportant 0 \$ enlevés	0,984

En raison de contraintes de temps associées aux dates de publication de l'EMIM et à l'obtention des données fiscales auprès de l'ARC, les données fiscales de la TPS ne sont pas disponibles à temps pour être utilisées par l'EMIM pour le mois courant. Les données de la TPS pour le mois d'avant le mois courant ne sont pas reçues à temps non plus pour permettre une revue et implémentation adéquate. Les données de la TPS qui sont deux mois avant le mois courant (par exemple données pour le mois du mois de la TPS pour le mois de référence de juillet pour l'EMIM) sont reçues avec assez de délais pour être incorporées dans le processus de traitement des données de l'EMIM. Il a été décidé de vérifier si un modèle peut être développé en utilisant les données de la TPS de deux mois d'avance par rapport au mois de référence de l'EMIM. Comme on peut le voir dans le tableau 2, la corrélation entre les livraisons du mois actuel et les recettes d'il y a deux mois (M-2) était de bonne qualité.

**Tableau 2. Corrélations entre les livraisons (juillet 2004) et les recettes (mai 2004)**

Types des unités	Corrélation %
Total	0,779
Unités rapportant 0 \$ enlevés	0,8
Unités aléatoires et rapportant 0 \$ enlevés	0,953

### **Simulation du processus de remplacement par les données fiscales**

Même si la corrélation entre les livraisons de l'EMIM et les recettes de la TPS était très bonne, la principale préoccupation concernant l'utilisation de données de la TPS était la préservation des tendances et des niveaux actuels de production de l'EMIM. Afin de s'assurer de ceci, une simulation de la méthode a été effectuée pour la période allant d'octobre 2002 au plus récent mois dont les données étaient disponibles au moment de la simulation.

### **Détermination des unités devant être remplacées**

Il fallait premièrement déterminer quels établissements pouvaient être remplacés. Ces unités comme simples, en vie, et ne figurant pas parmi les unités qui forment les premiers 80 % du domaine d'intérêts. Dans le cadre de la simulation, le domaine était formé des établissements des classes à trois chiffres du SCIAN, réparties selon la province. La seule exception était la

classe 336 du SCIAN (Fabrication de matériel de transport) car ses classes réparties par provinces étaient codées à quatre chiffres du SCIAN. Les unités simples de taille importante ou qui importent à l'industrie sont également exclues de la liste en raison de l'incidence sur leur industrie ou sur leur province.

On a donc considéré 4 453 établissements comme propices à la substitution au cours du premier mois de la simulation (octobre 2002). Vu que l'impact du remplacement de ces unités n'était pas connu, et que certaines unités similaires étaient nécessaires pour l'échantillon afin de supporter le modèle TPS, il fut décidé que 50% de ces unités étaient éligibles pour remplacer la TPS. Conséquemment, un échantillon simple aléatoire de la moitié de ces établissements a été choisi dont 2,241 unités choisies pour remplacement.

## **Modélisation**

### **Modélisation des livraisons**

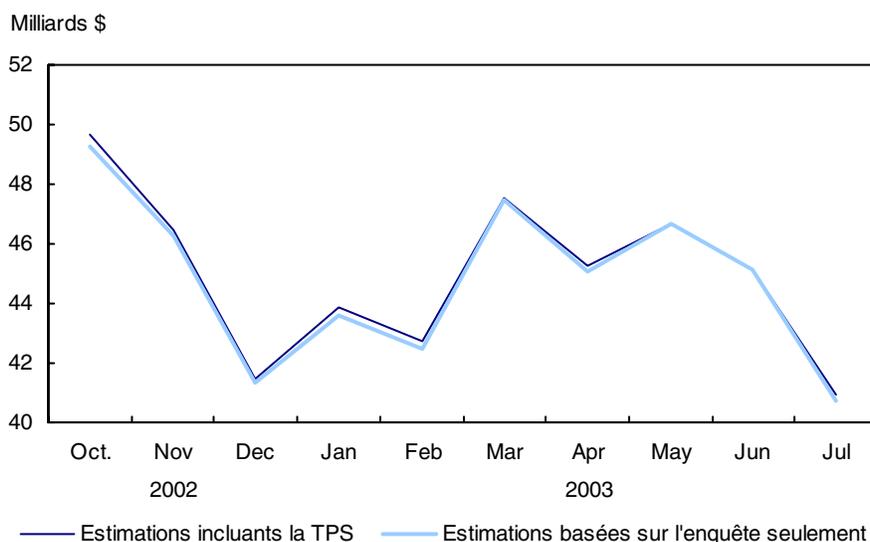
Pour chaque mois de la simulation, on a créé un modèle basé sur les données des livraisons du mois actuel ainsi que sur les données des recettes de la TPS de deux mois auparavant (M-2). Le modèle était fondé sur la relation entre les livraisons et les recettes de la TPS des unités qui pouvaient être remplacées mais qui sont demeurées dans la portion enquêtée et qui étaient toujours en vie. Ce modèle a ensuite été appliqué aux données des recettes de la TPS des unités devant être remplacées.

De façon à s'assurer de la création d'un modèle convenable, les observations aberrantes ont été retirées du groupe modèle. Il s'agissait d'observations dont les recettes de la TPS ne correspondaient pas aux données rapportées des livraisons. Il y a eu des situations où un établissement n'avait accumulé aucune recette de la TPS, tandis que les estimations de la valeur des livraisons de l'enquête étaient assez importantes, et vice versa.

Après avoir été produit, le modèle a été appliqué aux unités devant être remplacées. On a également éliminé les observations aberrantes (unités qui sont entièrement hors de portée lorsque comparées avec toutes les unités du modèle) de cette population. Les dossiers d'observations aberrantes générées par le modèle ont été remplacés par le processus normal d'imputation.

L'un des principaux avantages de cette approche était qu'une fois les valeurs modélisées au niveau micro-économique, il était possible d'utiliser le processus d'estimation actuel. Comme le montre le graphique 1, les résultats de la simulation révèlent peu de différences entre les estimations contenant des données administratives et celles qui ont été calculées uniquement à partir de l'enquête. L'intervalle de confiance de 95 % laisse supposer que la nouvelle ligne demeure bien à l'intérieur des limites de l'estimé original. Cette hypothèse peut être appliquée à des estimations plus détaillées, mais il faut noter qu'avec davantage de détails de l'industrie, la modélisation n'est pas aussi fiable et les observations aberrantes peuvent avoir une plus grande influence.

## Résultats de la simulation des livraisons au niveau canadien



### Modélisation des stocks et des commandes en carnet

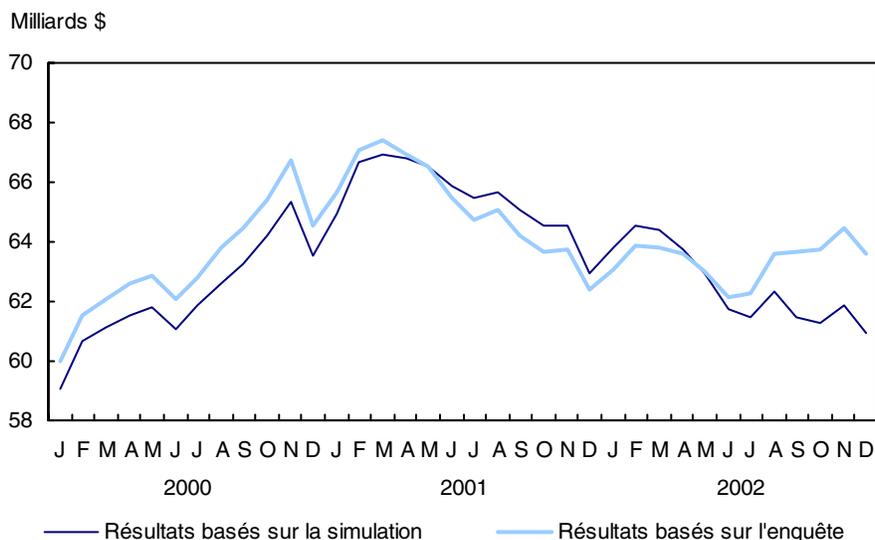
À l'heure actuelle, les données manquantes sur les stocks et les commandes en carnet en raison d'une absence ou d'un retard des réponses sont imputées par le système de traitement de l'EMIM. La tendance des établissements du même groupe du SCIAN ou provincial est utilisée aux fins de l'imputation pour les stocks et les commandes en carnet de ces unités.

Différentes enquêtes préliminaires ont démontré peu ou aucune corrélation entre les stocks et les variables disponibles dans le fichier de la TPS. Le système de traitement actuel servira à imputer les stocks et les commandes en carnet de tous les établissements dont les données doivent être remplacées par des données de la TPS.

Bien que la qualité des données imputées soit convenable pendant une période de quelque mois, au fil du temps, on s'attend à une diminution de la qualité des données des stocks et des commandes en carnet imputés. Ainsi, il a fallu prendre certaines précautions en ce qui a trait à l'imputation à long terme. Comme les strates de l'échantillon de l'EMIM sont mises à jour tous les trois ans, il a été décidé que trois ans serait la période maximale pendant laquelle une unité de TPS serait conservée avant d'être complètement éliminée de l'échantillon ou réinsérée dans la population observée. À des fins d'essai, les séries des stocks et des commandes en carnets de la période 2000-2002 ont été imputées et comparées aux séries publiées. Les résultats des tests sont illustrés dans le graphique 2. Des résultats similaires ont été obtenus pour les commandes en carnet. La qualité des données imputées était semblable à celle de la série publiée, ce qui a permis de conclure que l'imputation est une méthode convenable de remplacement des données d'enquête concernant les stocks et les commandes en carnet pour les unités de la TPS.

En ce qui concerne les unités qui deviennent inactives (dont la valeur des exportations est zéro), on cessera l'imputation des stocks et des commandes en carnet. Actuellement, lorsqu'un établissement dont les données sont recueillies au moyen de l'enquête devient inactif, les analystes peuvent décider de réduire les stocks de l'unité sur une période de quelques mois. Lorsque des unités dont les données sont obtenues au moyen de la TPS deviennent inactives, il est également possible d'utiliser cette stratégie, si les stocks sont importants, bien qu'on ne s'attende pas à ce que cela se produise régulièrement car la plupart de ces unités sont de très faible taille. Il est donc possible que certaines unités de la TPS conservent des stocks pendant quelques mois après être devenues inactives. Autrement, on cessera l'imputation des stocks et des commandes en carnet de la plupart des établissements.

## Résultats de la simulation pour les inventaires au niveau canadien



### Conclusion

Il a été démontré que la substitution de données d'enquêtes par des données administratives, particulièrement par des données de la TPS, conservait la qualité élevée des estimations de l'EMIM tout en réduisant le fardeau du répondant, surtout en ce qui a trait aux petites entreprises. Cette nouvelle initiative permet également une plus grande flexibilité par rapport à la sélection d'unités à être interviewées car celles-ci peuvent être remplacées par des données administratives sans qu'il y ait d'incidence sur les estimations en général.

Dès le mois d'août 2004, quelques 2,200 unités (représentant 25% de l'échantillon de l'EMM) ne recevront pas de questionnaire. Plutôt, leurs valeurs seront modelées en utilisant les données de la TPS (pour les livraisons) et les données imputées (inventaires et commandes en carnet). Les résultats obtenus basés sur le modèle courant seront suivis de près pour ainsi déterminer s'il est possible d'utiliser d'avantage les données administratives.