

## Détails techniques sur les estimations de l'Enquête mensuelle sur les services de restauration et débits de boissons (EMSRDB)

Pour l'EMSRDB, on estime les ventes totales à l'aide de deux types de méthodologies d'estimation. La première est l'estimateur Horvitz-Thompson et consiste à multiplier chaque réponse en données par son poids d'échantillonnage. Pour les strates à tirage complet, le poids est de un puisque toutes les entreprises de la strate sont sélectionnées dans l'échantillon. Pour les strates à tirage partiel, un échantillon aléatoire est sélectionné et le poids correspond à l'inverse de la probabilité de sélection.

Le second genre de méthodologie d'estimation est l'estimation par la méthode du ratio. Elle est utilisée dans la population représentée par les strates à tirage partiel des industries 7221 et 7222 du système de classification des industries de l'Amérique du Nord (SCIAN), au Québec, en Ontario, au Manitoba, en Saskatchewan, en Alberta et en Colombie-Britannique. Elle est également utilisée pour les strates à tirage nul de toutes les catégories du SCIAN et de toutes les provinces.

Dans l'EMSRDB, l'estimation par la méthode du ratio améliore la qualité de l'estimation en tirant parti de la grande corrélation entre les données d'enquête et l'information auxiliaire. Les données d'enquête sont les ventes déclarées par les répondants et l'information auxiliaire est le revenu déclaré au programme de la taxe sur les produits et services (TPS) administré par l'Agence du revenu du Canada (ARC). Pour calculer l'estimateur par le ratio :

- Les données sur le revenu provenant de la TPS et les données des ventes de l'enquête proviennent d'un échantillon d'unités; en outre, les données sur le revenu provenant de la TPS sont disponibles pour toutes les unités non échantillonnées dans la population.
- On calcule l'estimation des ventes à partir des données d'enquête en utilisant le poids d'échantillonnage (ventes\_est). Ceci est l'estimé Horvitz-Thompson de ventes.
- On calcule l'estimation du revenu provenant de la TPS à partir des données sur la TPS, uniquement en fonction des unités échantillonnées, à l'aide du poids d'échantillonnage (TPS\_est). Ceci est l'estimé Horvitz-Thompson du revenu provenant de la TPS.
- On additionne les données sur le revenu provenant de la TPS pour toutes les unités, échantillonnées ou non, afin d'obtenir le total connu du revenu provenant de la TPS (TPS\_total).
- On calcule le ratio (TPS\_total / TPS\_est), que l'on appelle le poids g.
- L'estimation des ventes par la méthode du ratio correspond aux ventes\_est multipliées par le poids g.

C'est-à-dire que l'on obtient l'estimateur par le ratio(ER) comme suit :

$$ER = \text{ventes\_est} * (\text{TPS\_total} / \text{TPS\_est})$$

À partir de l'échantillon, l'estimation de TPS\_total (connu) est TPS\_est. Si TPS\_est est plus grand que TPS\_total, nous nous attendons (puisque'il y a une étroite relation entre les données sur les ventes et les données sur la TPS) à ce que ventes\_est soit plus grand que les ventes totales (réelles et inconnues). Cela donne un poids  $g$  ( $TPS\_total / TPS\_est$ ) de moins de 1 et un ER dont la valeur est inférieure à celle de ventes\_est. Si, par contre, TPS\_est est inférieur à TPS\_total, le poids  $g$  est supérieur à 1 et la valeur de l'ER est plus grande que celle de ventes\_est.

La méthode de l'estimation par le ratio sert actuellement à produire des estimations pour l'EMSRDB à compter de celles de janvier 2009 et remplace la méthode par modèle ratio qui était utilisée auparavant.

Le nouvel estimateur par le ratio offre l'avantage de déceler tôt les entreprises qui ne sont plus en exploitation (« décès ») puisque les fermetures d'entreprise pour les unités échantillonnées sont décelées immédiatement et servent à calculer l'estimation par le ratio. Ces unités d'enquête donneraient une valeur de zéro, ce qui réduirait l'estimation globale et le poids de ces unités représenterait d'autres fermetures d'entreprises qui n'auraient pas été décelées dans les données auxiliaires sur la TPS.

Dans l'ancienne méthode par modèle ratio, on utilisait des données préliminaires ou initiales de la TPS pour lesquelles l'information sur la fermeture d'entreprises n'était pas aussi actuelle. Cela s'explique par le fait qu'au départ, on supposait que les entreprises non déclarantes étaient « vivantes » dans le fichier mensuel de la TPS pour tenir compte de la remise tardive des déclarants mensuels et de la remise des déclarants trimestriels et annuels. Après une période prédéterminée et quand plus aucune autre remise n'était reçue ou attendue, l'entreprise finissait par être considérée comme étant « fermée » à la date de la dernière remise.

Même si les premières versions des données sur la TPS ne représentent pas immédiatement les décès réels, à mesure que des versions mises à jour et ultérieures des données de la TPS sont fournies par l'ARC, ces données représentent plus exactement les réponses de l'enquête tenant finalement compte de la fermeture d'entreprises. Par conséquent, puisque les estimations révisées et déjà publiées avant 2009 – qui reposent sur l'ancienne méthode par modèle ratio – se servent de la dernière version à jour des données sur la TPS qui tiennent compte des fermetures d'entreprises, les anciennes estimations sont compatibles avec les estimations qui découlent de la nouvelle méthode d'estimation par le ratio et, ainsi, aucune révision n'est nécessaire.

Tirage nul :

Il n'y a pas d'échantillon pour les strates à tirage nul. On estime plutôt les ventes à l'aide de la méthode de l'estimation par le ratio pour toutes les provinces et SCIAN en fonction des données provenant des strates à tirage partiel.

Mesures de l'exactitude :

L'erreur-type et le coefficient de variation (CV) des estimations sont calculés à partir du plan d'échantillonnage et de la méthode d'estimation à l'aide des données d'enquête recueillies.