

Comparaison des indices de prix des vêtements et des chaussures à partir de données de caisse et de données moissonnées sur le Web*

Antonio G. Chessa et Robert Griffioen

Question clé

Le moissonnage de prix sur le Web est de plus en plus considéré comme un mode de collecte alternatif possible. Les instituts de statistique peuvent en effet avoir des difficultés pour acquérir des données de caisse, acquisition qui est aussi un processus assez long. En outre, le moissonnage du Web permet de collecter les prix d'assortiments entiers de façon automatique, un plus par rapport à la collecte traditionnelle. En revanche, les données moissonnées ne donnent pas d'information sur les dépenses de consommation, alors que cette information est essentielle pour calculer les poids des produits. La question est donc de savoir si les données moissonnées sur le Web peuvent être utilisées pour construire les indices de prix à la consommation.

Méthodologie

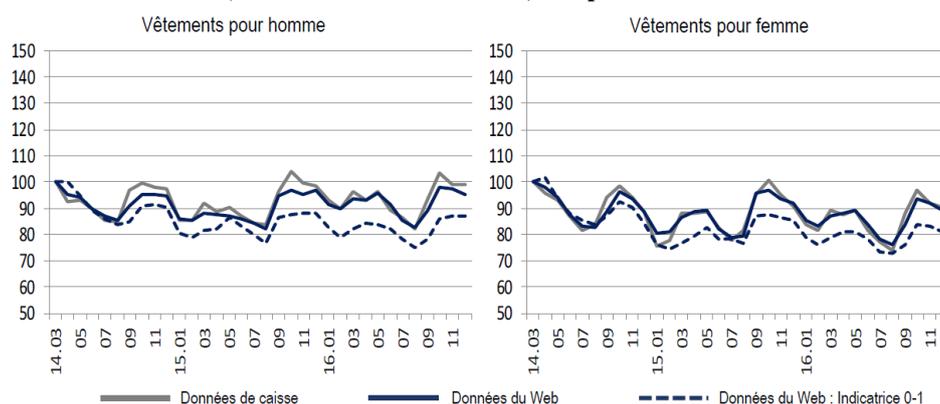
Statistics Netherlands reçoit les données de caisse et moissonne les prix en ligne et les métadonnées d'un même vendeur néerlandais. Les prix et les quantités des deux sources de données sont comparés pour 16 catégories de vêtements et de chaussures pour hommes et femmes, puis nous utilisons une méthode multilatérale (Geary-Khamis) pour calculer les indices des prix avec les deux ensembles de données.

Principaux résultats

Pour les catégories de produits étudiées :

- les données de caisse et les prix du web sont le plus souvent égaux, mais ces derniers sont en moyenne plus élevés ;
- les corrélations du nombre de produits vendus et du nombre de prix moissonnés sont remarquablement élevées au fil du temps ;
- bien que les prix moissonnés soient en moyenne plus élevés, les indices de prix calculés avec les deux sources sont peu différents pour la plupart des catégories de produits ;
- les différences au niveau COICOP sont seulement de 0.3 point de pourcentage pour les indices d'année en année ;
- remplacer le nombre de prix moissonnés par une indicatrice binaire (moissonné/non moissonné) débouche sur des différences importantes et un biais à la baisse de l'indice des prix moissonnés (figure). Cela suggère qu'il faut éviter de dé-dupliquer les prix des articles moissonnés plusieurs fois.

Comparaison des prix à partir des données de caisse et des données du Web en utilisant une indicatrice (moissonné/non moissonné) des prix du Web



Message

Ces résultats semblent prometteurs, mais il faut garder à l'esprit qu'ils ne portent que sur un seul détaillant. Deux questions émergent des comparaisons qui ont été faites. L'une porte sur les fortes corrélations des quantités vendues et du nombre de prix moissonnés au fil du temps. L'autre sur la stratégie du vendeur pour organiser son site Web : est-ce qu'il promeut les articles qui se vendent le mieux ? La détermination de profils caractéristiques sur la base des données de caisse pourrait aider à évaluer la pertinence des données du Web pour le calcul de l'indice des prix. Cela devrait encourager les instituts de statistique à investir dans des analyses statistiques des données de caisse et des données moissonnées sur le Web.