

Méthode non-paramétrique de détection et de traitement des individus atypiques : cas de l'enquête EVE

(Enquête auprès des visiteurs résidant habituellement hors de France)

Claire JACOD¹ et Julien VALENTINO²

La ligne « Voyages » de la Balance des paiements recense, en débit, les dépenses touristiques des résidents français à l'étranger et, en crédit, les dépenses touristiques des non-résidents lors de leur séjour en France. Avec l'arrivée de l'euro en 2001, la méthodologie de la ligne « Voyages » de la Balance des Paiements a nécessité une complète révision. Après une phase transitoire, la nouvelle méthodologie a été définitivement adoptée en juin 2007 lors de la publication du rapport annuel de la Balance des Paiements.

Concernant le poste des recettes de la ligne « Voyages », une enquête complexe a été mise en place : l'Enquête auprès des voyageurs étrangers (EVE). Enquête trimestrielle, elle est réalisée en partenariat avec la direction du Tourisme. Les objectifs de cette enquête sont donc multiples : estimer les dépenses touristiques des non-résidents en séjour en France, mais aussi estimer par nationalité le nombre de touristes, le volume de nuitées... L'enjeu est de taille, la France étant le premier pays au monde en termes d'arrivées de touristes.

L'enquête EVE se distingue par son originalité et sa complexité : absence de base de sondage des non-résidents, réalisation de l'enquête pour chaque mode de transport utilisé par les non-résidents pour quitter la France, questionnaires auto-administrés et disponibles dans une dizaine de langues, multiplicité des sources de données de trafic...

L'important biais haussier observé dans cette enquête a nécessité la mise en place de différents traitements statistiques.

L'utilisation d'une méthode de détection des dépenses atypiques permet d'isoler les questionnaires fortement contributifs aux dépenses totales, qui correspondent à des réponses hors du champ de l'enquête (achat de bien immobilier, d'œuvres d'art...) ou à des erreurs (de saisie, de lecture, de conversion de monnaie...). Cette méthode intervient très en amont de la chaîne de production, quand toutes les informations ne sont pas encore disponibles. Elle n'utilise que peu de variables et se veut simple dans son application, consistant à calculer différents seuils de dépenses. Les questionnaires dont les dépenses se situent au-delà de ces seuils sont alors imputés. La méthode est basée sur une estimation de la densité des dépenses et fait intervenir un estimateur à noyau. Afin de garantir une certaine stabilité et résistance de

¹ Direction de la Balance des Paiements, Banque de France - Claire.Jacod@banque-france.fr

² (prédécesseur de Claire JACOD) - Ministère de l'Écologie, de l'Énergie, du Développement durable et de l'Aménagement du Territoire - Commissariat général au Développement durable - Julien.Valentino@developpement-durable.gouv.fr

ces seuils, la méthode est itérative et permet de déterminer des seuils « médians » après prise en compte d'un facteur aléatoire multiplicatif.

Cependant, la méthode des seuils ne détecte que les dépenses globalement atypiques. Un deuxième traitement est donc réalisé, qui s'intéresse aux questionnaires atypiques au sein de chaque strate (pays d'origine, mode de transport...), et qui consiste en un lissage polynomial des poids intra-strates. Ce traitement prend en compte une « distance » aux dépenses médianes au sein de chaque strate pour repondérer les questionnaires, sans que les poids de redressement s'en trouvent trop modifiés, et conduit à une réduction de la variance des estimateurs.

La présentation visera à exposer rapidement les grandes lignes et les spécificités de l'enquête EVE, puis à détailler les deux principaux traitements explicités précédemment : la détection des dépenses atypiques, et la méthode d'estimation retenue afin de limiter l'impact des atypismes restant sur le poste recettes de la ligne « Voyages ».