

Enquête nationale sur les Transports et Déplacements : application de méthodes sur « petits domaines » à des estimations régionales

Josiane LE GUENNEC¹

L'enquête nationale sur les Transports et Déplacements (ENTD) est réalisée tous les dix ans sur l'ensemble du territoire. La dernière édition a été réalisée en 2006-2007. Enquête à plusieurs facettes, elle distingue les déplacements motivés par les activités habituelles quotidiennes, les déplacements de loisir du week-end, les voyages de longue distance. Elle a pour but de mesurer la fréquence des déplacements de la population selon leurs motifs et les modes de transport utilisés, les distances parcourues et le temps qui leur est consacré. Dans cette étude, on s'est intéressé aux déplacements de courte distance au cours d'une journée de semaine.

Organisée pour produire une statistique nationale, l'enquête n'est pas directement exploitable au niveau de chacune des 22 régions par des estimateurs directs de type Horvitz-Thomson, en raison de tailles trop faibles des échantillons régionaux. C'est pourquoi, pour un nombre limité d'indicateurs, on a tenté une modélisation destinée à réduire la variance des estimateurs régionaux.

Pour chacun des paramètres sélectionnés, quatre méthodes d'estimation ont été testées :

- un estimateur direct par calage séparé dans chaque région sur les totaux régionaux X_R de variables auxiliaires :

$$\hat{Y}_{R,cal} \approx \hat{Y}_{R,reg} = \hat{Y}_{R,HT} + \hat{\beta}_R' (\mathbf{X}_R - \hat{\mathbf{X}}_{R,HT}) \quad (1)$$

- un estimateur direct modifié par régression dans l'échantillon national complet :

$$\tilde{Y}_{R,regnat} = \hat{Y}_{R,HT} + \hat{\mathbf{b}}' (\mathbf{X}_R - \hat{\mathbf{X}}_{R,HT}) \quad (2)$$

Contrairement à l'estimateur (1), les coefficients b sont ici estimés au niveau national.

- un estimateur synthétique de type régression :

$$\hat{Y}_{R,regsyn} = \hat{\mathbf{b}}' \mathbf{X}_R \quad (3)$$

Les coefficients b sont estimés au niveau national, puis appliqués aux structures régionales.

- un estimateur par modélisation explicite, au moyen d'un modèle de niveau individus, avec un effet aléatoire prenant en compte l'effet spécifique de la région sur les comportements :

¹ Josiane Le Guennec (INSEE-Rennes)
Tél 02 99 29 25 04
e-mail : josiane.le-guennec@insee.fr

$$Y_{R,i} = \beta' X_{R,i} + v_R + e_{R,i} \quad (4)$$

Un programme d'estimation de variance des estimateurs directs par calage a été développé, de façon à situer les estimateurs modélisés dans les intervalles de confiance des estimateurs directs sans biais. Cinq régions ayant financé des renforcements d'échantillon ont permis de tester la fiabilité des résultats obtenus. Dans ces zones, l'estimateur direct avait un coefficient de variation assez faible pour constituer une référence et hiérarchiser les estimateurs de type « petits domaines ».

La comparaison des résultats montre le caractère biaisé de l'estimateur synthétique, et conclut à une plus grande efficacité de l'estimation par modélisation explicite avec effet aléatoire.

Bibliographie :

P. Ardilly, « Panorama des principales méthodes d'estimation sur petits domaines », INSEE-UMS, 2006.

Techniques d'enquête, volume 29, n°1, juin 2003 : numéro spécial de la revue de Statistique Canada consacré à l'estimation sur petits domaines.

M.A. Hidioglou et Z. Patak, « Estimation par domaine par la régression linéaire », Techniques d'enquête, volume 30 n°1, juin 2004, pp.73-85.

C. Lavergne, C. Trottier, « sur l'estimation dans les modèles linéaires généralisés à effets aléatoires », Revue de statistique appliquée, 2000, XLVIII, pp.49-67.

D. Ravaille, « Mise en œuvre d'estimations sur « petits domaines » pour l'estimation de taux de chômage et de taux d'activité par région, à partir de l'enquête Emploi », INSEE 2005, rapport de stage ENSAE.

« La mobilité des Français. Panorama issu de l'enquête nationale transports et déplacements 2008 », revue du CGDD, décembre 2010.