

ESTIMATION SUR PETITS DOMAINES À L'INSEE : PREMIER BILAN ET PISTES DE PROGRÈS

Pascal Ardilly ¹

¹ *Insee, Direction de la méthodologie, 65 Bd Garibaldi, 69003 Lyon, pascal.ardilly@insee.fr*

L'Insee n'a qu'une expérience limitée en matière d'estimation sur petits domaines. Les "commandes" passées à l'Ums-ménages restent récentes mais on peut néanmoins déceler un frémissement traduisant un intérêt grandissant pour le sujet. La distance qui sépare la France d'autres pays sensiblement plus utilisateurs de ces techniques reste importante mais elle pourrait se réduire assez vite si une dynamique d'ampleur suffisante s'engageait au sein de la toute nouvelle Direction de la méthodologie. Quelques enquêtes ont donné lieu à des estimations localisées - avec plus ou loin de succès et la France a participé par ailleurs à un groupe de travail européen (dit Essnet) sur le sujet, qui vient de s'achever. Ces opérations sont l'occasion de faire le point sur les difficultés rencontrées et, en corollaire, sur les pistes de progrès que l'on pourrait envisager.

On peut distinguer des obstacles de nature organisationnelle ou informatique, devoir faire face à des difficultés liées à la disponibilité des données ou tout simplement rencontrer des problèmes de technique statistique. L'organisation renvoie à l'investissement de l'Insee (ou plus généralement de la statistique publique) pour traiter ce type de question, étant considéré qu'il s'agit de techniques complexes voire très complexes et qui ne se prêtent pas à un traitement mécanique de masse. Les obstacles informatiques sont pour leur part beaucoup moins liés à la stratégie interne mais ce sont plutôt les limites des logiciels disponibles qu'il faut prendre en compte.

La disponibilité de données "explicatives" locales est centrale : il existe en effet des contraintes de précision, d'homogénéité, de cohérence de ces données avec l'information du questionnaire. En ce qui concerne les enquêtes sur les personnes physiques et les ménages, le recensement est évidemment la source à laquelle on pense en priorité mais il peut présenter lui-même dans certains cas des incertitudes appréciables à un niveau fin (communal par exemple).

Quant aux difficultés techniques - parfois il s'agit d'ailleurs plutôt de "curiosités", ou d'interrogations voire même de perplexité face à des points encore obscurs - elles sont assez diverses. La conséquence de l'utilisation des techniques d'estimation sur petits domaines est l'introduction d'un biais d'estimation systématique - en tout cas selon l'approche classique des sondages. En effet, si on exclut la mise en oeuvre des techniques de calage sur marges locales (qu'il ne faut pas oublier!), les estimateurs sont typiquement modèle-dépendants. La question centrale de la qualité des estimations finales se pose à trois niveaux (au moins) : comment valider un "modèle", comment apprécier le biais du résultat final et enfin comment savoir si la méthode finalement retenue est la meilleure, ou du moins surpasse l'approche directe classique - ces trois composantes étant d'ailleurs largement liées. Au-delà de ces dimensions essentielles, on peut éprouver quelques réserves face à la complexité du traitement des modèles non-linéaires (modèle de Poisson, modèle binomial-logistique,...) qui servent de base, le plus souvent, aux estimations de proportions et aux dénombrements. Des questions plus philosophiques peuvent surgir quant au sens de l'introduction d'effets aléatoires locaux dans les modèles stochastiques, et on retrouve aussi les très récurrentes interrogations sur l'impact de la prise en compte des poids dans les estimateurs. La présentation abordera l'ensemble de ces aspects, de manière synthétique.

Bibliographie

Rao, J.N.K. (2003). *Small Area Estimation*, New-York, Wiley.