

COMPARAISON DE DIFFÉRENTS SCÉNARIOS DE CALAGE POUR L'ENQUÊTE SRCV

Robin Chatelot (*), Maxime Bergeat (*)

(*) Insee, Département des ressources et des conditions de vie des ménages

robin.chatelot@insee.fr

Mots-clés (6 maximum) : Calage, marges de calage, SRCV, bootstrap

Domaine concerné : Théorie des sondages, après collecte - Calage

Résumé

L'ajout d'une marge de calage sur laquelle caler ses poids peut conduire à un dilemme. D'un côté, ajouter une marge permet de s'assurer que les poids permettent de retrouver la valeur de cadrage. Mais de l'autre, l'ajout de cette marge supplémentaire nécessite parfois de déformer fortement les poids. Ce dilemme est particulièrement prégnant dans le cas où le nombre de marges utilisé est trop élevé et/ou la taille de l'échantillon trop faible : faire converger le calage sans revoir les critères de qualité du calage à la baisse peut alors s'avérer impossible. Des contraintes trop fortes sur un calage peuvent donc amener à des ruptures de série (par exemple dans des cas où il est nécessaire de faire varier les bornes de calage d'une année sur l'autre) alors que l'un des intérêts majeurs du calage est de les éviter.

L'enquête « Statistiques sur les ressources et les conditions de vie » (SRCV) est réalisée chaque année par l'Insee. Elle porte sur environ 30 000 individus de plus de 16 ans, interrogés pendant 4 années consécutives. Cette enquête européenne permet notamment de calculer plusieurs indicateurs de pauvreté ou d'exclusion sociale, tels que la pauvreté monétaire et la privation matérielle et sociale. Aujourd'hui, l'enquête utilise des marges de calage :

- sociodémographiques (estimées à partir de l'enquête emploi en continu, ces marges sont communes à plusieurs enquêtes ménages de l'Insee) ;
- fiscales (masses de revenus estimées à partir des fichiers fiscaux avec lesquels l'enquête SRCV est appariée).

Par ailleurs, depuis le millésime 2021 de l'enquête, le calage est également réalisé sur des distributions de revenu déclaré par unité de consommation, estimés à partir du dispositif Filosofi ou de l'enquête Revenus fiscaux et sociaux.

Les marges utilisées dans l'enquête SRCV sont nombreuses et peuvent conduire à des distorsions de poids importantes. Le but de cet article est d'étudier des scénarios alternatifs de calage de l'enquête SRCV et d'analyser différents critères permettant de choisir les marges à utiliser pour le calage de SRCV. Les tests seront réalisés à partir de l'enquête SRCV 2024 en se concentrant sur les poids

transversaux, sur le champ de la France métropolitaine¹. Les revenus qui sont en partie imputés avec des mécanismes d'imputation qui dépendent des poids (en particulier les prestations sociales) ne seront pas modifiés dans cette étude comparative à des fins de simplification.

Les critères retenus pour comparer les scénarios alternatifs de calage sont de plusieurs natures pour déterminer quel est le scénario le plus prometteur :

- étude de critères sur la structure des poids issus du calage : répartition des rapports de poids après/avant calage, accumulation des rapports de poids au bord des bornes utilisées dans le calage², déformation des poids...
- étude de critères plus conceptuels comme la source des marges utilisées pour le calage, leur champ, ainsi que leurs potentielles variabilités dans le temps. Ces critères sont difficiles à quantifier mais sont essentiels à la construction d'un jeu de marges pertinent ;
- estimation de la variabilité des variables d'intérêt selon les différents poids calés testés à l'aide d'une méthode par *bootstrap*. Cela permettra d'avoir une idée de la « robustesse » des différents scénarios de calage testés. Il sera aussi important de comparer les résultats obtenus avec ceux des jeux de poids utilisés actuellement dans l'enquête, pour estimer les risques de ruptures de série.

Si le temps le permet, ces critères pourront être analysés sur plusieurs millésimes pour analyser la stabilité dans le temps du respect de ces différents critères.

Bibliographie

[1] Torsten Harms et Pierre Duchesne, « On Calibration Estimation for Quantiles », Statistics Canada, Catalogue No, 12-001-XPB, Vol, 32, No, 1, pp 37-52, June 2006

¹ Les départements d'Outre-mer (hors Mayotte) sont intégrés au champ de l'enquête SRCV depuis 2022. Le calage de l'enquête est réalisé séparément pour les ménages résidant en Outre-mer, avec des marges de calage moins détaillées que celles retenues pour la France métropolitaine compte tenu de la taille de l'échantillon.

² La méthode logit est utilisée pour le calage de l'enquête, ce qui permet de contrôler les rapports de poids après/avant calage.