

Conférence spéciale

L'EMPLOI DE CALMAR II : CALAGE GÉNÉRALISÉ ET APPLICATION A LA NON-REPONSE.

Jean-Claude DEVILLE (),
Josiane LE GUENNEC, Olivier SAUTORY (**)*

(*) CREST / ENSAI
(**) CEPE

La correction de la non-réponse totale se fait généralement par une ré pondération des unités élémentaires. Celle-ci résulte logiquement de l'estimation, à partir des données, d'un modèle paramétrique de réponse spécifié en utilisant les variables qui ont la plus grosse influence sur la probabilité de réponse postulée. Les corrections de poids apparaissent alors comme les inverses des probabilités de réponses estimées.

Cette façon de voir les choses amène à se poser des questions sur les méthodes d'estimation à mettre en œuvre pour parvenir à ces fins. Il apparaît que l'utilisation d'un principe de calage arrange singulièrement bien les choses. Cependant on a intérêt à caler plutôt sur des variables corrélées aux variables d'intérêt, qui n'ont pas de raison d'être les mêmes que celles qui influent sur la propension à répondre.

On est donc amené à développer une forme de calage un peu plus générale que celle de la théorie habituelle et qui présente un intérêt autonome. Sous sa forme linéaire, cette généralisation du calage se ramène à une régression avec variables instrumentales.

Le calage généralisé produit une réduction de la variance en même temps qu'une diminution du biais de réponse. Il autorise une correction de la non-réponse même quand les variables qui la causent ne sont pas entièrement observées et sont, éventuellement, des variables d'intérêt. La précision des estimateurs calés s'estime avec la technique habituelle des résidus.

La méthode du calage généralisé est disponible dans la macro de calage CALMAR II, dont elle constitue la principale innovation méthodologique par rapport à la version précédente. Les autres nouveautés de la macro sont :

- la possibilité de réaliser des "calages simultanés" dans le cas où la collecte d'informations s'opère à différents niveaux d'observation (par exemple ménages et individus), permettant d'assurer une cohérence entre les statistiques obtenues à partir des différents fichiers de l'enquête ;
- l'ajout d'une nouvelle fonction de distance ;
- une meilleure prise en compte des colinéarités entre variables de calage ;
- quelques améliorations dans l'ergonomie du programme.