
ESTIMATION ET DÉCOMPOSITION DE L'EFFET DE MODE DANS LES ENQUÊTES MULTIMODES (INTERNET/TÉLÉPHONE)

Gaëlle DABET, Zora MAZARI, Ines OUJIA (*)

(*) Céreq, Équipe ingénierie et gestion d'enquêtes

melanie.vignale@cereq.fr

Mots-clés : Multimode, agrégation, matching, estimation, économétrie

Domaines concernés : multimode, effet de mode

Résumé (entre 350 et 900 mots environ)

Depuis la fin des années 90, le Céreq a mis en place un dispositif d'enquêtes original qui permet d'étudier l'accès à l'emploi des jeunes sortis du système éducatif une même année. L'enquête Génération, historiquement administrée par téléphone, a été rénovée et intègre désormais un nouveau mode de collecte : internet. La première édition de l'enquête produite en multimode est la Génération 2017, interrogée en 2020.

Des expérimentations multimodes ont été conduites pour tester plusieurs protocoles de collecte. Elles ont mis en avant les enjeux liés à l'introduction d'un mode de collecte en auto-administré : la formulation des questions, les consignes, la disposition des éléments sur une page doivent permettre de collecter des données comparables (téléphone/internet). Ces expérimentations ont également permis des premières estimations des effets de mode sur les variables d'intérêts de l'enquête. Le multimode induit à la fois des effets de sélection – le choix du mode de collecte diffère selon le profil des répondants – et des effets de mesure – un même enquêté peut répondre différemment à la même question selon le mode. Par la suite, le questionnaire Génération a été retravaillé, tant sur le fond que sur la forme, de manière à limiter ces effets.

Cet article s'inscrit ainsi dans la poursuite des travaux présentés lors des deux dernières JMS^{[1][2]}. Il propose une analyse des effets de mode au travers de deux enquêtes multimodes distinctes dont les protocoles de collecte sont sensiblement différents : la première interrogation la « Génération 2017 », réalisée en 2020, et l'enquête Génération « Covid et après ? », ré-interrogation de la « Génération 2010 ». Cette dernière dispose d'un échantillon embarqué avec affectation aléatoire du mode de collecte (deux sous échantillons

monomodes) permettant de contrôler de l'effet de sélection. Pour chaque enquête, différentes méthodes seront implémentées et testées.

Concernant l'enquête Génération 2017, une méthode de matching sur score de propension sera mise en œuvre. Les différentes étapes de son application seront présentées : estimation du score de propension et mesure de sa qualité, choix de la stratégie d'utilisation du score dans l'estimation des effets. Plusieurs méthodes d'appariements seront aussi évaluées (en fonction des contraintes d'équilibrage) afin d'obtenir des résultats satisfaisants.

Du côté de l'enquête Génération « Covid et après ? », dans la mesure où les échantillons sont indépendants, l'effet de sélection est *a priori* maîtrisé. Il va s'agir de repérer les éventuels biais de mesure en utilisant par exemple une méthode de calage.

Les variables présentant un effet de mesure pourront être comparées aux variables détectées lors des expérimentations. S'agit-il des mêmes variables ? Les ajustements du questionnement ont-ils permis de limiter les effets de mode ? Nous nous interrogerons également sur les préconisations à fournir pour l'utilisation des données d'enquête. Quelles analyses et quelles interprétations seront possibles lors des exploitations ?

Bibliographie

- [1] Barret C., Dzikowski C., « La collecte par Internet est-elle l'avenir des enquêtes Génération du Céreq ? », *12^{èmes} Journées de Méthodologie Statistique*, 2015.
- [2] Cissé M., Barret C., « Agrégation de données multimode : Impact sur la modélisation des variables présentant un effet de mesure », *13^{èmes} Journées de Méthodologie Statistique*, 2018
- [3] Legleye, S., « Effets de sélection, imputations et effets de mode : les dernières tendances en matière de multimode », *Séminaire de Méthodologie Statistique*, 2017.
- [4] Quantin S., « Estimation avec le score de propension sous R », *Méthodologie statistique*, document de travail, INSEE, 2018