

---

## COMBINAISON ET CORRECTION DE LA NON-RÉPONSE D'UN ÉCHANTILLON MULTIMODE ET MONOMODE

*Khaled LARBI (\*)*

*(\*) Insee, Direction de la méthodologie et de la coordination statistique et internationale*

*khaled.larbi@insee.fr*

**Mots-clés :** Correction de la non-réponse, multimode, partage des poids.

**Domaines :** Théorie des sondages après collecte - Traitement de la non-réponse totale.

---

### Résumé

L'utilisation de différents modes de collecte au sein d'une enquête s'est développée ces dernières années. Cette stratégie permet d'augmenter les taux de réponse en permettant d'obtenir des informations sur des individus qui n'auraient pas répondu via un premier mode. Nous considérons dans la suite une approche par multimode séquentiel : un premier mode - généralement moins coûteux comme Internet - est proposé. Si l'individu ne répond pas alors il lui est proposé de répondre via un mode plus intermédié comme le téléphone (ou le face-à-face, voire une combinaison séquentielle des deux).

L'existence d'un mode de collecte à faible coût – ici Internet – offre la possibilité de tirer, en sus de l'échantillon principal multimode, un échantillon de taille plus importante qui sera collecté en mobilisant uniquement ce mode de collecte peu onéreux. Cet échantillon de taille plus importante pourrait ainsi permettre d'effectuer des analyses à une échelle plus locale.

Cependant, le mécanisme de réponse sur cet échantillon monomode risque d'être différent de celui affectant l'échantillon multimode : en particulier, du fait de taux de réponse (nettement) plus faibles dans le protocole monomode Internet que dans l'approche multimode, une sélection des répondants sur variables inobservables dans l'échantillon monomode pourrait avoir un impact plus fort que dans le cas multimode. Ce type de mécanisme de non-réponse ne permet pas d'assurer une correction de la non-réponse conduisant à des estimations consistantes des totaux.

Dans ces travaux, nous considérons une enquête mobilisant deux échantillons :

- un échantillon multimode  $S_{\text{multimode}}$  dans lequel le mode Internet puis le mode téléphone sont proposés de manière séquentielle. Nous supposons par ailleurs que la non-réponse est ignorable dans cet échantillon,
- un échantillon monomode  $S_{\text{monomode}}$  dans lequel seul un mode de collecte est proposé. Cet échantillon peut être tiré de manière à surreprésenter certains domaines (par exemple, certaines régions).

L'échantillon multimode permet donc de construire des estimateurs approximativement sans biais du total. Néanmoins, l'échantillon monomode permet d'accroître la taille de l'échantillon de l'enquête. De plus, la surreprésentation de certains domaines de cet échantillon permet *a priori* une estimation plus précise sur ces domaines. Toutefois, l'échantillon monomode risque d'être sujet à une sélection des répondants sur variables inobservées. Deux interrogations se dessinent :

- Existe-t-il une méthode de combinaison de  $S_{\text{multimode}}$  et  $S_{\text{monomode}}$  et de correction de la non-réponse permettant de construire des estimateurs plus efficaces que ceux basés sur  $S_{\text{multimode}}$  y compris en cas de problème de sélection ?
- Si oui, existe-t-il des gains pour l'estimation restreinte au domaine surreprésenté dans  $S_{\text{monomode}}$  ?

Cet article présente plusieurs stratégies de combinaison d'échantillons et de correction de la non-réponse et compare leur efficacité à l'aide de simulations. Dans ces travaux, les tailles des échantillons ainsi que les plans de sondage permettant d'obtenir  $S_{\text{multimode}}$  et  $S_{\text{monomode}}$  sont des simplifications des échantillons tirés dans l'enquête Logement 2023.

Nous comparons les différentes méthodes dans différents cas : cas de non-réponse ignorable pour les deux échantillons, présence de sélection sur inobservable uniquement lors de la réponse par Internet, présence de sélection sur inobservable Internet et totale.