

UNE MÉTHODE DE MESURE DE L'EFFET ENQUÊTEUR

Catherine Berthier, Jean-Claude Deville, Bernard Néros

Le contexte des enquêtes auprès des ménages

La plupart des enquêtes auprès des ménages à l'Insee se déroulent en face à face par visite d'enquêteur. L'échantillon propre à une enquête est issu d'une base de logements, par un plan de sondage complexe, à plusieurs degrés, qui aboutit à confier à chaque enquêteur des logements pas trop éloignés les uns des autres, au nombre de 25 environ pour un échantillon de taille moyenne (soit une dizaine de milliers de logements).

Pour une enquête donnée, on cherche s'il existe une part de variation des réponses due au fait que tous les ménages ne sont pas interviewés par le même enquêteur. Autrement dit, l'effet enquêteur existe si un ménage est susceptible de donner des réponses différentes à des enquêteurs différents. A défaut d'interroger deux fois chaque ménage, on mesure cette variation sur deux échantillons de ménages équivalents, qui ont répondu à des enquêteurs différents pour un même questionnaire. La première vague du panel européen a fourni la possibilité de mettre en place cette expérience.

Il s'agit d'une mesure globale car elle ne conduit pas à isoler le travail d'un enquêteur particulier : on s'intéresse à la différence entre les résultats tirés de deux enquêtes en tout point semblables, excepté le fait qu'elles ont été menées par deux groupes d'enquêteurs.

Cette préoccupation autour de l'effet enquêteur redevient actuelle au moment où la collecte assistée par ordinateur (système Capi) se généralise. Cette transformation du mode de collecte peut conduire à faire intervenir un nombre plus limité d'enquêteurs, parce que leur formation doit être accrue. Or cette concentration joue sur l'effet enquêteur, et on court le risque qu'elle ne l'amplifie.

La première vague du panel européen ne s'est pas déroulée sous le système Capi, mais de manière traditionnelle, par remplissage de questionnaires sous forme papier. Un certain nombre d'erreurs peuvent subsister, commises par l'enquêteur, (filtres mal respectés, modalités impossibles....) qui seraient évitées dans une collecte sous Capi. L'effet enquêteur n'est pas entaché par ce type d'erreurs, puisqu'on le mesure à partir des données apurées.

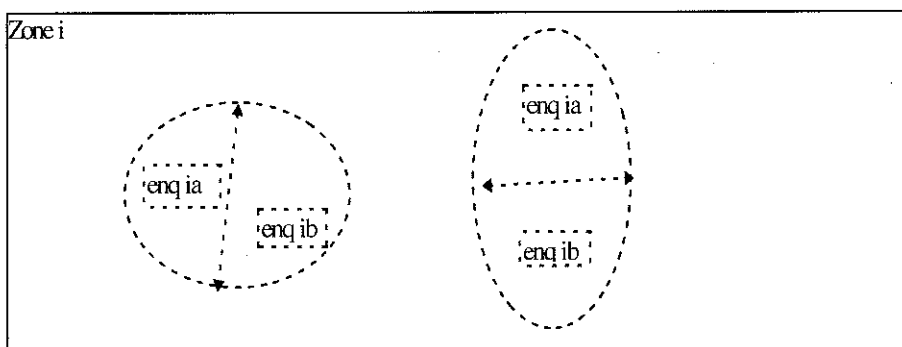
Le dispositif expérimental

Un dispositif spécifique a été mis en place sur la première vague de l'enquête panel européen, permettant d'isoler l'effet enquêteur en neutralisant l'effet du terrain d'enquête. Un certain nombre de zones ont été sélectionnées avant le tirage de l'échantillon propre à l'enquête, chaque zone regroupant deux terrains couverts par un lot de deux enquêteurs. Dans chaque zone, les deux terrains ont été jugés ressemblants et présentant une homogénéité interne assez forte. Les zones sont pour la plupart urbaines parce qu'on y trouve plus facilement des terrains équivalents.

Les deux enquêteurs correspondant à une zone se sont partagé par moitié les échantillons des deux terrains. L'opération a porté sur 58 zones et 116 enquêteurs, pour 1400 ménages.

Graphique 1 :

Le partage d'une zone sur une paire d'enquêteurs désignés par enq ia et enq ib



Estimation de l'effet enquêteur

Grâce à ce dispositif qui respecte l'organisation de la collecte, on considère disposer de I zones sur lesquelles $2I$ échantillons de ménages indépendants ont été répartis aléatoirement sur $2I$ enquêteurs. On peut donc former deux groupes d'enquêteurs, notés groupe a et groupe b, qui ont mené deux enquêtes en parallèle dans l'ensemble des I zones.

Graphique 2 :
Deux enquêtes en parallèle

	groupe a		groupe b
zone i	<i>enquêteur 'ia'</i> ménage n°1 ménage n°19 ($n_{ia}=19$)		<i>enquêteur 'ib'</i> ménage n°20 . .. ménage n°35 ($n_{ib}=16$)

Dans chaque zone i , on dispose des deux enquêtes a et b. Pour une variable d'intérêt X , on estime la moyenne dans chaque zone \bar{x}_{ia} et \bar{x}_{ib} , à l'aide de deux échantillons, de taille n_{ia} et n_{ib} . Au mieux, on dispose de la mesure de X sur tous les ménages des deux échantillons, qu'on a choisis de même taille. Mais le plus souvent il y a une non-réponse totale et une non-réponse partielle, et n_{ia} et n_{ib} sont les tailles en général différentes des deux sous-échantillons de répondants propres à X . Lorsqu'on estime l'effet enquêteur à partir des seuls répondants, on se limite au cas où la non-réponse obéit à un mécanisme homogène à l'intérieur de la zone.

En l'absence d'effet, on noterait m_i et σ_i^2 la moyenne et la variance de X dans la zone i . S'il y a un effet enquêteur, cette moyenne et cette variance intègrent une erreur de mesure individuelle propre à chaque enquêteur. On modélise cette erreur de la manière suivante :

$$E\left(\bar{x}_{ia}/ia \text{ est tiré}\right) = m_i + \mu_{ia} \quad E\left(\bar{x}_{ib}/ib \text{ est tiré}\right) = m_i + \mu_{ib}$$

$$V\left(\bar{x}_{ia}/ia \text{ est tiré}\right) \approx \frac{\sigma_{ia}^2}{n_{ia}} \quad V\left(\bar{x}_{ib}/ib \text{ est tiré}\right) \approx \frac{\sigma_{ib}^2}{n_{ib}}$$

(on assimile le sondage à un sondage aléatoire simple à taux de sondage négligeable)

Les μ_{ia} et μ_{ib} sont 2I variables aléatoires indépendantes de moyenne nulle, parce qu'on ne peut pas identifier de biais, et de variance commune σ_{enq}^2 .

On cherche à estimer σ_{enq}^2 :

En constatant que $E\left(\frac{\bar{x}_{ia} - \bar{x}_{ib}}{n_{ia,ib}}\right) = \mu_{ia} - \mu_{ib}$, et en utilisant la formule approchée de variance ci-dessus, il vient :

$$(\mu_{ia} - \mu_{ib})^2 = E\left(\frac{(\bar{x}_{ia} - \bar{x}_{ib})^2}{n_{ia,ib}}\right) - \frac{\sigma_{ia}^2}{n_{ia}} - \frac{\sigma_{ib}^2}{n_{ib}}$$

Or $(\mu_{ia} - \mu_{ib})$ est de moyenne nulle et de variance $2\sigma_{enq}^2$.

On en déduit que, en notant s_{ia}^2 et s_{ib}^2 les variances corrigées, dans chaque zone i, $(\bar{x}_{ia} - \bar{x}_{ib})^2 - \frac{s_{ia}^2}{n_{ia}} - \frac{s_{ib}^2}{n_{ib}}$ estime sans biais $2\sigma_{enq}^2$.

Un estimateur sans biais de σ_{enq}^2 est donc :

$$\hat{\sigma}_{enq}^2 = \frac{I}{2I} \sum_{i=1}^I \left[(\bar{x}_{ia} - \bar{x}_{ib})^2 - \frac{s_{ia}^2}{n_{ia}} - \frac{s_{ib}^2}{n_{ib}} \right]$$

Cet estimateur de variance peut conduire à des valeurs négatives.

Comment tester l'existence de l'effet enquêteur

L'absence d'effet enquêteur revient à : $\sigma_{enq}^2 = 0$

On cherche à tester la nullité de σ_{enq}^2 , pour accepter ou rejeter l'existence d'un effet enquêteur, contre l'hypothèse alternative $\sigma_{enq}^2 > 0$. Mais on ne dispose pas de la loi de l'estimateur $\hat{\sigma}_{enq}^2$ sous l'hypothèse d'absence d'effet enquêteur.

Pour passer outre, on *simule* la loi de $\hat{\sigma}_{enq}^2$, sous l'hypothèse privilégiée, celle d'absence d'effet enquêteur. Sous cette hypothèse, les deux groupes a et b ont les

mêmes propriétés statistiques que d'autres groupes de ménages, pourvu que ces groupes présentent la même répartition par zones.

Pour obtenir ces autres groupes de ménages, l'idée est de réutiliser les mêmes données d'enquêtes, collectées sur les mêmes ménages, en *mélangeant* ces ménages pour reformer deux autres groupes. Ces mélanges s'effectuent en respectant les frontières des zones : à partir des $n_{ia} + n_{ib}$ ménages de la zone i, on peut former par mélange aléatoire deux nouveaux groupes i1 et i2, qui ne respectent plus la séparation par enquêteur, tout en obéissant à la contrainte de répartition, n_{i1} ménages d'un côté, et n_{i2} de l'autre.

On choisit par exemple de classer tous les ménages enquêtés de la zone i suivant un aléa, on affecte les données collectées auprès des n_{i1} premiers ménages au groupe i1, et les suivantes au groupe i2. On empile ensuite ces mélanges sur les I zones et on calcule à partir des deux nouveaux groupes 1 et 2 la statistique

$$\frac{I}{2I} \sum_{i=1}^I \left[(\bar{x}_{i1} - \bar{x}_{i2})^2 - \frac{s_{i1}^2}{n_{i1}} - \frac{s_{i2}^2}{n_{i2}} \right]$$

On obtient *une autre* valeur de l'estimateur de variance.

Pour obtenir toute une distribution, on crée autant de fois qu'on veut deux groupes. Mille « mélanges » de ce type simulent une distribution de cet estimateur de variance.

On dispose de la *valeur observée* $\hat{\sigma}_{enq}^2$, celle calculée à partir des deux groupes d'enquêteurs a et b. On teste la nullité de σ_{enq}^2 en plaçant cette valeur observée dans la distribution simulée. Le fractile auquel correspond la valeur observée fournit une p-valeur à partir de laquelle on accepte ou on rejette l'hypothèse d'absence d'effet enquêteur.

Plutôt que de retenir pour une variable X donnée la seule conclusion du test pour un niveau fixé, on préfère conserver le fractile auquel correspond la valeur observée : si celui-ci est proche de la médiane, on conclut à l'absence d'effet enquêteur. Si ce fractile est décalé vers la droite de la distribution, correspondant à une valeur élevée, (il s'agit d'un test unilatéral), on soupçonnera un effet enquêteur. Si on constate un décalage pas très fort, mais se produisant sur plusieurs variables, on y verra une confirmation de l'effet enquêteur, alors qu'en fixant un niveau de test, on aurait rejeté l'hypothèse d'effet enquêteur sur chacune des variables examinées séparément.

En outre, ce procédé de test se justifie parce que *s'il y a effet enquêteur*, le fait de mélanger les ménages pour qu'ils ne soient plus regroupés par groupes d'enquêteurs, *atténue* l'effet enquêteur, mais ne l'annule pas complètement. De ce

fait en comparant la vraie valeur à la distribution simulée, on sous-estime le décalage dû à l'effet enquêteur.

Comment figurer une différence entre les statistiques provenant des deux groupes d'enquêteurs

Pour voir une différence éventuelle entre les statistiques tirées de chacune des enquêtes menées par les deux groupes d'enquêteurs, on étend cette méthode à une fonction T des observations. T peut être un total, une moyenne, un quantile, un ratio, une variance..... A partir de groupes i1 et i2 obtenus selon le même principe de mélanges, on calcule les statistiques T_{i1} et T_{i2} . Du point de vue de cette statistique, une distance $D_{1,2}$ entre les deux groupes 1 et 2 formés en empilant les groupes i1 et i2 peut être évaluée comme :

$$D_{1,2} = \sum_{i=1}^I |T_{i1} - T_{i2}|$$

(On pourrait d'ailleurs généraliser en introduisant $\sum_{i=1}^I f(|T_{i1} - T_{i2}|)$, f possédant certaines propriétés)

La distribution de $D_{1,2}$ est simulée en formant mille fois deux groupes 1 et 2, par mélanges.

On examine comment se situe sur la distribution *la valeur observée*, celle qui mesure la distance entre les deux groupes d'enquêteurs $D_{a,b} = \sum_{i=1}^I |T_{ia} - T_{ib}|$

Le fait de former deux groupes par hasard, et non plus par enquêteur, entraîne une compensation des effets individuels des deux enquêteurs à l'intérieur de chaque groupe. On attribue donc à l'effet enquêteur le fait que la distance entre les deux groupes d'enquêteurs paraisse élevée dans la distribution des distances entre groupes formés par hasard.

Application de la méthode

Cette méthode a tout d'abord été appliquée sur la phase qui précède l'interview, qui comprend le repérage et le contact des deux enquêtes.

L'information tirée de cette phase a été résumée d'une part par la distinction entre répondants et non-répondants à l'enquête (pour les logements déclarés dans le champ de l'enquête), d'autre part par le statut de l'enquête (cette fois pour l'ensemble des logements de l'échantillon).

Extrait du questionnaire de l'enquête panel européen

Le statut de l'enquête

L'enquêteur doit placer le ménage dans l'une des catégories suivantes :

- le logement est hors champ
- le logement est dans le champ, mais on n'obtient pas de réponse du ménage.
- le logement est dans le champ, le ménage accepte de répondre à l'enquête.

Le montant du revenu :

1- Pour résumer :

En considérant l'ensemble des revenus de tous les individus du ménage actuellement, quel est le montant mensuel des revenus nets (de contributions sociales et CSG) dont votre ménage dispose ?

2- Si vous ne pouvez donner un montant précis, pouvez-vous au moins en donner une estimation ? (on propose à l'enquêté une liste de 9 tranches de revenu)

L'opinion du ménage sur son niveau de vie :

Si on considère à présent les ressources mensuelles de votre ménage, diriez-vous qu'elles vous permettent de vivre :

1. très difficilement
2. difficilement
3. assez difficilement
4. assez aisément
5. Aisément
6. très aisément

En s'intéressant à la phase précédant l'interview, on cherche en particulier à savoir si l'hypothèse d'homogénéité du comportement de réponse est acceptable ou non.

Sur les seuls ménages répondants, la méthode est appliquée sur la durée de l'interview, et sur deux questions suivantes portant sur le montant du revenu mensuel, et sur l'appréciation du niveau de vie. La déclaration du revenu a donné lieu à deux traitements, l'un portant sur le montant de revenu libellé en chiffres, le deuxième sur le mode de déclaration choisi, en chiffres ou en tranches.

Sur ces questions, l'effet enquêteur sur la non-réponse partielle n'a pas été testée, parce que les non-réponses sont rares.

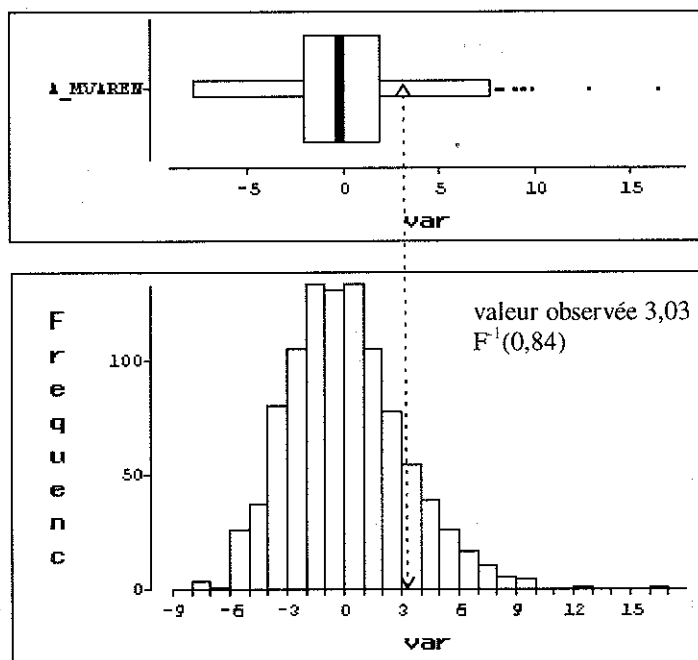
Le test d'un lien entre déclaration de revenu et l'opinion du ménage sur son propre niveau de vie a été construit de la manière suivante : tout d'abord, un revenu par unité de consommation a été calculé en appliquant une échelle d'équivalence (le

poids du premier adulte est 1, le poids de chaque personne âgée de plus de quatorze ans est 0,5, celui de chaque enfant est 0,3). Puis des tranches de revenu ont été déterminées pour présenter les mêmes effectifs par tranche que l'opinion sur le niveau de vie. Ensuite une variable de cohérence a été calculée pour chaque ménage comme la différence entre les niveaux d'échelle du ménage pour le revenu et le niveau de vie.

Pour une variable donnée, deux distributions ont été simulées : celle de la variance enquêteur $\hat{\sigma}_{enq}^2$, et celle de la différence des moyennes des groupes 1 et 2.

Pour la variance enquêteur comme pour la différence entre groupes, si la valeur observée est décalée sur la droite de la distribution simulée, on lit dans ce décalage l'impact de l'effet enquêteur sur la statistique choisie. Par contre, le fait que cette valeur observée se situe vers le centre de la distribution permet d'exclure un effet enquêteur (la distribution simulée est à peu près normale, le centre correspond à la fois à la moyenne, au mode, et à la médiane, donc au fractile 0,5).

Graphique 3 :
Distribution de la variance enquêteur du revenu mensuel du ménage



Les résultats de l'ensemble des simulations réalisées sont résumés dans le tableau ci-dessous par la donnée du fractile auquel correspond la valeur observée, c'est-à-dire celle propre aux deux groupes d'enquêteurs.

Tableau : Fractiles des valeurs observées dans les distributions simulées

Champ retenu	Variable	Fractile* de la variance enquêteur observée	Fractile* de la différence observée
Tous les logements tirés	Statut de l'enquête	0,62	0,63
Tous les logements hormis les hors champ	Non réponse totale	0,64	0,54
Les ménages répondants à l'enquête	Durée de l'interview	1	1

* Lecture : Pour le statut de l'enquête, la variance enquêteur observée se trouve placée dans la distribution simulée de telle sorte que 62% des valeurs de la distribution lui sont inférieures.

Champ retenu	Variable	Fractile* de la variance enquêteur observée	Fractile* de la différence observée
Les ménages donnant leur revenu en chiffres	Revenu mensuel du ménage	0,84	0,91
Les ménages donnant leur revenu en chiffres ou en tranche	Mode de déclaration du revenu	0,99	0,99
Les ménages répondants à la question sur le niveau de vie	Opinion du ménage sur son niveau de vie	0,69	0,48
Les ménages répondants aux questions niveau de vie et revenu	Cohérence de classement entre niveau de vie et revenu	0,75	0,80

En résumé

L'effet enquêteur ne joue ni sur le statut de l'enquête ni sur la non-réponse totale quand on en confond les différentes causes possibles. On légitime ainsi l'hypothèse d'homogénéité des comportements de réponse, qui autorise à modéliser l'effet enquêteur à partir des seuls répondants.

Pour le statut de l'enquête en trois modalités, la moyenne est utilisée comme un résumé de la distribution. La nature qualitative de la variable rendrait préférable un test portant sur une distance entre distributions, plutôt que sur une différence de moyennes. Ce travail n'a pas encore été réalisé.

Avec la même limite portant sur l'utilisation de la moyenne, l'effet enquêteur a été testé sur le statut de l'enquête plus détaillé, en différenciant les différentes causes de non-réponse (ménage impossible à joindre, absent de longue durée, déclaré inapte à répondre, ou refus de répondre). Ce statut détaillé, comme la durée d'interview déclarée, font apparaître un fort effet enquêteur. Ces deux variables ont en commun d'être renseignées directement par l'enquêteur. Elles ne sont que très peu sensibles à une interaction entre l'enquêteur et l'enquêté. Elles sont entachées d'un fort effet individuel de l'enquêteur.

L'effet paraît nul sur l'appréciation du niveau de vie. Il n'est pas très fort sur le montant du revenu. Sur la cohérence de classement entre échelle de revenu et

échelle de satisfaction du niveau de vie, l'effet est un peu plus élevé que sur le seul niveau de vie. Puisque l'effet touche un décalage entre les deux échelles calculé en moyenne par enquêteur, il laisse soupçonner un décalage à tendance un peu systématique par enquêteur.

Quant au mode de déclaration du revenu, il peut être considéré comme déterminé par l'enquêteur.

BIBLIOGRAPHIE

Cochran W. G., *Sampling Techniques*, Wiley.

Deville J.-C.(1994), *Quelques éléments pour l'analyse de l'effet enquêteur dans le dispositif « europanel »*, Note Insee N°780/F401.

Särndal C.-E., Swensson B., Wretman J., *Model Assisted Survey Sampling*, Springer-Verlag.