

# Les enquêtes multimode : attention aux effets de mode

Gaël de Peretti<sup>1</sup>  
et Tiaray Razafindranovona<sup>2</sup>  
Insee, Département des méthodes statistiques



Les organismes producteurs d'information statistique, et en particulier la Statistique publique (« Official statistics » dans le monde anglo-saxon) doivent faire face à une demande d'enquêtes toujours plus exigeante en termes de qualité, qualité au sens large à savoir précision, pertinence, comparabilité, cohérence, clarté, fraîcheur, dans un contexte général de restriction budgétaire. L'utilisation d'internet comme mode privilégié ou complémentaire de recueil des données paraît une bonne solution. En effet, il est a priori peu coûteux, que ce soit dans ses dimensions financière ou de durée. Ainsi, il y a une volonté commune à ces organismes pour recourir de façon plus systématique à la collecte par internet, et plus généralement multimode, pour faire face aux problèmes de couverture, d'échantillonnage, de non-réponse ou de mesure, tout en respectant des contraintes budgétaires. Il s'agirait de limiter conjointement l'erreur d'enquête totale et les coûts d'enquête. Cependant, si la collecte par internet est un mode peu coûteux, elle pose des problèmes méthodologiques non négligeables : couverture, auto-sélection ou biais de sélection, non-réponse et les difficultés de sa correction, « satisficing »<sup>3</sup>, etc. Aussi, avant de développer ou généraliser l'utilisation du multimode, les effets de mode de collecte doivent être étudiés pour voir dans quelle mesure la collecte multimode peut être effectivement intégrée dans le processus de production des enquêtes.

Les informations statistiques sont produites à partir de deux grands types de sources : les données administratives et les enquêtes. Ces dernières se sont développées en France à partir des années 1950 suite à la diffusion en France des techniques d'échantillonnage. Depuis deux décennies, le recours aux données administratives se multiplie afin de limiter la charge de réponse qui pèse sur les entreprises et les ménages, mais aussi de profiter d'informations plus ou moins exhaustives que l'on peut ventiler de façon détaillée. Toutefois, la statistique d'enquête reste une source importante de la statistique publique compte tenu de sa capacité à s'adapter à des questionnements complexes, des populations spécifiques, voire les combinaisons des deux.

1. [gael.de-peretti@insee.fr](mailto:gael.de-peretti@insee.fr)

2. [tiaray.razafindranovona@insee.fr](mailto:tiaray.razafindranovona@insee.fr)

3. Voir plus loin

Cependant, si les enquêtes ont été développées à l'origine pour éviter des recensements coûteux, ces opérations, compte tenu de la complexité croissante des protocoles et de la taille importante des échantillons nécessaire pour produire des informations sur des sous-populations plus ou moins fines, sont elles-aussi coûteuses. Or, les instituts nationaux de statistique (INS) et les organismes publics ou parapublics producteurs de données statistiques doivent faire face à la fois à des restrictions budgétaires mais aussi à une baisse tendancielle du taux de réponse, ce qui n'est pas sans incidence sur la qualité des enquêtes.

De ce fait, ces organismes réfléchissent à la mise en place de nouveaux processus de production pour les enquêtes en recourant à une collecte multimode, et plus particulièrement à la collecte par internet afin de réduire les coûts de production sans diminuer la qualité des informations produites, c'est-à-dire en limitant autant que possible l'erreur d'enquête totale. Il faut aussi avoir en tête que la qualité d'une enquête ne se réduit pas à la seule précision des estimations qu'elle permet de produire mais doit aussi prendre en compte d'autres dimensions comme la pertinence, la comparabilité (dans le temps ou l'espace), la cohérence (avec d'autres sources), l'accessibilité et la clarté pour les utilisateurs potentiels, la fraîcheur, etc.

## L'erreur d'enquête totale, un paradigme ancien de nouveau à la (multi)mode

Si le terme erreur d'enquête totale est relativement récent (Andersen, Kasper, Frankel et al., Total Survey Error, 1979), le concept est né en même temps que les enquêtes. En effet, dès le début, des travaux portent sur les différentes dimensions possibles de l'erreur d'enquête et de multiples typologies ont été développées pour distinguer les sources d'erreur.

Ainsi, les principales sources d'erreur d'une enquête seraient :

- l'erreur de spécification : conception du questionnaire pas adapté aux objectifs de l'enquête ;
- l'erreur de couverture : base de sondage incomplète ou incluant à tort des unités que l'on ne souhaite pas enquêter ;
- l'erreur d'échantillonnage : erreur due au fait que l'on interroge qu'une fraction de la population visée ; c'est l'erreur la plus documentée et qui possède l'avantage de pouvoir en général être calculée ;
- l'erreur de non-réponse : elle peut être totale si l'unité ne répond à aucune question de l'enquête ou partielle si seule une partie du questionnaire est concernée. Là encore, il est possible dans la plupart des cas de corriger et de quantifier les effets de cette non-réponse ;
- l'erreur de mesure ou d'observation : erreur dont les sources sont multiples. Elle peut provenir de l'enquêté. Dans ce cas, les deux sources les plus étudiées sont relatives à des notions de psychologie : la désirabilité sociale, à savoir la tendance à fournir des réponses conformes aux normes sociales ou donnant une bonne image de l'enquêté, de façon consciente ou pas ; le « satisficing » c'est-à-dire le fait de choisir une réponse « dont on se satisfait » mais qui n'est pas la réponse optimale. Cet effet de satisficing dépend à la fois de la difficulté de la question mais aussi des compétences de l'enquêté et de sa motivation. Par exemple, lors d'une question avec plusieurs modalités, les personnes sujettes au satisficing peuvent avoir tendance à choisir la première des modalités ("primacy") ou la dernière ("recency"). De façon générale, elles sont enclines à répondre de façon machinale. Mais les erreurs d'observation peuvent aussi provenir de l'enquêteur ou de l'enquête ;
- l'erreur de traitement des données : là encore les sources sont multiples (saisie, codification, pondération, tabulation, imputation, etc.).

Parmi ces erreurs, comme nous l'avons écrit précédemment, certaines sont susceptibles d'être quantifiées et sont prises en compte via des techniques statistiques plus ou moins sophistiquées : erreur d'échantillonnage, erreur de non-réponse. Pour d'autres, les résultats sont empiriques et les solutions quand elles existent ne sont pas généralisables mais adaptées à la source analysée. Parfois, il est même possible d'identifier ou de constater des erreurs sans

être en mesure de pouvoir les corriger. Tout au plus, le statisticien pourra préciser les limites de l'information qu'il a produite.

Ainsi, aussi séduisant soit le paradigme de l'erreur d'enquête totale, il n'est pas encore totalement opérationnel pour les producteurs de données. Il est toutefois utile pour identifier les sources potentielles d'erreur, pouvoir apporter des informations sur le poids relatif de ces différentes composantes et ainsi orienter sur quels champs il faut faire des progrès méthodologiques pour une opération donnée et enfin réfléchir a priori au moyen de les éviter.

S'il n'est pas nouveau, ce paradigme est à l'origine de nombreux travaux méthodologiques et est au cœur de nombreux colloques, comme par exemple le séminaire annuel ITSEW (International Total Survey Error Workshop, premier en 2005 puis annuel depuis 2008) du NISS (National Institute for Statistical Science). Par ailleurs, dès lors que l'on souhaite modifier un protocole de collecte, cette approche conceptuelle paraît tout à fait adaptée pour réfléchir aux effets de ces modifications d'un point de vue global. Aussi, le renouveau de ce paradigme n'est-il pas sans lien avec le développement des enquêtes multimode.

## **« Il y a plus de quarante ans que je fais des enquêtes multimode sans que j'en susse rien »**

La notion de mode intervient à plusieurs étapes dans une enquête : contact, collecte, suivi, etc. Par mode, nous entendons façon de faire ou de procéder pour réaliser le contact, la collecte, le suivi, etc., d'une enquête. Le mode de contact recouvre les moyens employés pour solliciter la participation d'un ménage ou d'une personne. Le mode de collecte correspond à la façon dont on va recueillir l'information. Le mode de suivi concerne plutôt les opérations longitudinales, c'est-à-dire avec plusieurs interrogations, et les relations post-enquêtes avec les répondants comme l'envoi de premiers résultats, le suivi d'adresse, etc.

En toute rigueur, une enquête est dite multimode dès lors que le mode de contact, le mode de suivi, le mode de collecte ou les modes de collecte, etc., sont différents, ce qui est quasiment toujours le cas dans les enquêtes de la Statistique publique. En effet, une fois l'échantillon constitué, les personnes sont souvent contactées via un courrier ou lettre-avis leur expliquant les objectifs de l'enquête et l'importance de leur participation, puis ils sont approchés directement ou par téléphone par les enquêteurs afin de fixer un rendez-vous, rendez-vous qui se passera généralement en face à face avec une collecte assistée par un ordinateur. Par la suite les enquêtés pourront être recontactés soit parce qu'ils font partie d'un panel ou d'une enquête en plusieurs vagues, soit parce que l'on souhaite contrôler les informations qu'ils ont fournies, etc. En fait, le terme enquête multimode fait plutôt référence à la notion de collecte multimode. Généralement, trois cas de collecte multimode sont distingués : le multimode intégré ; le multimode séquentiel ; le multimode concurrent. Dans le premier cas, il s'agit d'interroger la même personne selon différents canaux. Le cas le plus classique consiste à recueillir une partie de l'information en face à face via un enquêteur et une autre à l'aide d'un questionnaire auto-administré. Il peut s'agir d'un carnet recensant les dépenses du ménage sur une période donnée (enquête Budget des familles par exemple) ou les activités d'une personne (enquête Emploi du temps), mais aussi d'un questionnaire spécifique par exemple sur des questions sensibles afin de limiter les biais de désirabilité sociale déjà évoqués précédemment. Dans le deuxième cas, il s'agit au fur et à mesure de la collecte de proposer des modes différents de collecte afin de limiter la non-réponse. Par exemple, l'organisme va contacter une entreprise par internet, puis en cas de non-réponse par téléphone, puis toujours en cas de non réponse envoyer un enquêteur sur place. Enfin, le multimode concurrent consiste à offrir dès le début de la collecte le mode de son choix à l'enquêté, souvent avec le même souci de vouloir limiter la non-réponse.

Les modes de collecte les plus classiques sont : le face à face ; le téléphone ; le questionnaire auto-administré qui peut être un questionnaire papier, un questionnaire en ligne sur internet (plus rarement un questionnaire audio intégré au sein d'une enquête en face à face qui permet d'assurer la confidentialité des réponses malgré la présence d'un enquêteur). Toutefois, on peut aussi parler de collecte passive, quand par exemple, des données d'enquête sont appariées avec des sources externes. C'est le cas de l'enquête de l'Insee sur les Revenus fiscaux et sociaux qui permet, entre autres, de calculer le taux de pauvreté : on réalise un appariement entre les données de l'enquête Emploi, des fichiers fiscaux et des fichiers sociaux afin d'améliorer la qualité des informations recueillies sur les revenus et les aides sociales perçues.

Comme nous l'avons énoncé en introduction, les producteurs de statistique doivent faire face à une baisse tendancielle des taux de réponse, en particulier pour les enquêtes Ménages (par la suite, l'essentiel des résultats exposés portera sur les enquêtes Ménages). Cette baisse des taux de réponse peut s'expliquer à la fois par un accès plus difficile au logement des enquêtés (digicode, interphone, etc.), un moindre sens civique ou plus simplement une multiplication des enquêtes qui peuvent lasser les enquêtés. Ainsi, il devient de plus en plus difficile pour un simple citoyen de distinguer entre une enquête menée par un institut privé souvent à but lucratif et une enquête de la Statistique publique censée produire des informations pour éclairer le débat public. Pour lutter contre cette baisse, le recours à la collecte multimode paraît une solution séduisante.

## Un exemple d'enquête multimode

Dans le cadre d'un règlement européen, l'Institut national de la statistique et des études économiques (Insee) réalise chaque année depuis 2007 une enquête téléphonique sur les Technologies de l'information et de la communication (TIC). Compte tenu de l'absence de base de sondage téléphonique exhaustive, l'Insee a lancé plusieurs opérations méthodologiques entre 2009 et 2011 pour limiter l'erreur de couverture avant de mettre en place un nouveau protocole. La collecte est réalisée en partie par téléphone (personnes retrouvées dans l'annuaire téléphonique, échantillon S1), en partie par internet/papier avec deux sous-populations : une sous-population dont les personnes sont présentes dans l'annuaire (échantillon S21) ; une sous-population non-présente dans l'annuaire (échantillon S22). Il est donc possible après standardisation des échantillons S1 et S21 d'identifier a priori un effet de mode, les personnes ayant été affectées aléatoirement à un mode de collecte. Cet effet de mode est corrigé en partant du principe que l'enquête téléphonique est le bon étalon. Cette correction est appliquée aux répondants internet/papier non présents dans l'annuaire. Cette méthodologie a permis d'obtenir des résultats cohérents avec d'autres enquêtes de la Statistique publique réalisées en face à face (ne souffrant donc pas de problème de couverture) sur le taux d'équipement (enquête Statistiques sur les ressources et conditions de vie) et l'utilisation d'internet (enquête Emploi du temps). Cette intégration de la collecte par internet/papier dans le processus de production est possible du fait de la grande simplicité du questionnaire.

## La collecte par internet a des avantages indéniables mais il ne faut pas céder sans réflexion à cet effet de mode

Compte tenu de la diffusion des ordinateurs et d'internet au sein de la population, le recours à la collecte par internet est une alternative intéressante à plus d'un titre pour les instituts producteurs d'information statistique. Tout d'abord, c'est le mode de collecte le moins coûteux, le coût marginal d'une enquête étant quasi nul. Ensuite, il est rapide et la gestion de la collecte peut se faire en continu. De même, il peut être interactif et apporter un soutien à l'enquêté via des liens hypertexte, des info-bulles, des Foires aux questions (FAQ), des contrôles intégrés qui peuvent éviter des erreurs de saisie. Enfin, c'est un mode peu intrusif qui permet à l'enquêté de répondre au moment où il le souhaite et peut aussi limiter les biais de désirabilité sociale puisque a priori, il est toujours plus facile de se confier à un ordinateur qu'à un enquêteur.

Mais, il faut aussi avoir à l'esprit qu'une partie de la population n'a pas accès à internet et qu'elle possède des caractéristiques socioéconomiques particulières. Par ailleurs, les choix techniques retenus par les instituts doivent être compatibles avec la plupart des navigateurs web et s'adapter dans une certaine mesure aux nouveaux outils d'accès à internet comme les tablettes, les smartphones, etc. Mais le défaut majeur de la collecte par internet est son faible taux de réponse. Ainsi, dans les expérimentations en cours à l'Insee depuis le début des années 2010, le taux de réponse des enquêtes internet est de 15 à 20 % (on atteint 35 à 40 % en intégrant la collecte papier associée). Cela n'est évidemment pas sans conséquence sur la qualité de l'information produite.

Aussi de nombreux travaux méthodologiques sont réalisés depuis les années 2000 sur le multimode pour voir dans quelle mesure, il est possible de mettre en place de tels protocoles sans détériorer l'erreur totale d'enquête. Car, ce qu'il faut savoir, c'est que lorsque l'on souhaite agir sur une composante de l'erreur d'enquête totale, cela peut avoir des conséquences sur les autres composantes. Ainsi, on peut offrir à des enquêtés difficilement joignables ou récalcitrants à une enquête en face à face la possibilité de répondre par internet afin de diminuer l'erreur de non-réponse. Il faudra s'assurer que les erreurs de mesure liées par exemple à un fort « satisficing » de ces personnes ne viennent pas limiter le gain en erreur totale d'enquête.

### À la recherche des effets de mode

De nombreuses études de méthodologie d'enquête s'intéressent à la mesure des effets de mode. Il s'agit d'analyser les effets du mode de collecte sur l'estimation du phénomène étudié. Le principe générique est de comparer les réponses fournies par deux modes différents. Cet effet de mode global est généralement séparé en deux effets : effet de sélection ; effet de mesure souvent appelé effet de mode dans la littérature.

Pour essayer de contrôler le premier, les techniques utilisées sont multiples mais le principe général est identique : il s'agit de rendre les deux échantillons comparables en s'assurant de la similarité sur des caractéristiques sociodémographiques observables dans les deux modes de collecte. L'hypothèse forte sous-jacente est que la prise en compte de ces variables de contrôle suffit à supprimer les effets de sélection. Une fois cet effet de sélection éliminé, les écarts constatés entre les deux modes de collecte sont qualifiés d'effet de mode ou d'effet de mesure (voir exemple en encadré). Évidemment, il est probable que nous ne disposons pas forcément de toutes les variables nécessaires à la bonne prise en compte des effets de sélection. Ceci est d'autant plus préjudiciable si les liens entre une variable omise et la sélection sont forts et que de plus le pouvoir explicatif de cette variable est important pour notre variable d'intérêt.

## Un exemple d'effet de mode

Fin 2010, l'Insee a réalisé une enquête expérimentale sur le logement et la mobilité résidentielle sur internet auprès de 10 000 personnes tirées dans les fichiers fiscaux, ce qui permettait, entre autres, d'avoir des informations sur le revenu fiscal et le logement (surface, nature, c'est-à-dire appartement ou maison, statut d'occupation). Ces personnes ont reçu une lettre-avis début octobre. Les non-répondants ont reçu une première lettre de relance deux semaines après contenant un coupon réponse leur permettant d'obtenir une version papier du questionnaire (sans enveloppe préaffranchie), puis une deuxième un mois plus tard. Le taux de réponse est de 24 % (20 % si l'on ne prend pas en compte les répondants papier). La structure des répondants à cette expérimentation, en termes de caractéristiques sociodémographiques et de conditions matérielles du logement, était très différente de celle de l'enquête Logement 2006. Afin de corriger de ces effets de structure (ou de sélection), les données ont été standardisées (on parle de calage sur marges) pour que la structure, en termes d'âge, de sexe, de diplôme, de statut d'occupation, de revenu fiscal, et enfin de tranche d'unité urbaine, surface, nombre de pièces, nombre d'habitants et statut d'occupation du logement, soit identique à celle observée en population générale en 2010. Après cette standardisation, on constate des réponses presque toujours plus négatives sur internet que dans l'enquête Logement sur toutes les questions d'opinion sur le logement, comme les problèmes de bruit, de température (excès de froid ou de chaud), de relation avec le voisinage, de qualité de l'air, etc. Deux explications sont avancées. Certains considèrent que malgré les efforts réalisés pour prendre en compte les effets de sélection, les personnes souhaitant se plaindre de leur logement seraient plus enclines à répondre (diminution de la qualité de l'enquête). D'autres estiment que les biais de désirabilité sociale liés à la présence d'un enquêteur diminueraient. En effet, il peut paraître délicat de reconnaître devant l'enquêteur une privation de confort, ce qui n'est pas le cas par internet.

Mais l'identification des effets de mode n'est pas une fin en soi. Dès lors que l'on souhaite réaliser une enquête multimode, il faut se poser la question de l'agrégation des résultats produits par les différents modes de collecte. Pour certains auteurs, l'objectif principal est de repérer les effets de mode pour pouvoir à terme les réduire. Pour d'autres, il existe un étalon parmi les modes et il s'agit de s'approcher des résultats produits par cet étalon. Par ailleurs, si les estimateurs de paramètres classiques comme les totaux, les moyennes, les quantiles sont identiques, il faut aussi contrôler que les corrélations entre la variable d'intérêt et les covariables sont semblables quel que soit le mode.

Parmi les résultats établis, il est clair qu'il existe une hiérarchie entre les différents modes que sont le face à face, le téléphone, les questionnaires auto-administrés papier puis internet en termes de taux de réponse et de coût : ils sont décroissants du face à face à internet). Par ailleurs, d'autres résultats semblent faire consensus sur les erreurs de mesure. Ainsi, les biais de désirabilité sociale seraient moins forts pour les questionnaires auto-administrés que pour le téléphone puis le face à face. En revanche, le biais de satisficing serait plus important sur internet puis sur papier, au téléphone et enfin en face à face. Ainsi, il y a beaucoup plus de non-réponse partielle dans les questionnaires auto-administrés que dans les enquêtes téléphoniques ou en face à face. De même, lorsqu'il y a des batteries de question (souvent d'opinion), avec des modalités de réponse identiques (par exemple tout à fait d'accord, plutôt d'accord, plutôt pas d'accord ; pas du tout d'accord), les répondants internet ont tendance à

choisir plus systématiquement la même réponse. Enfin, les enquêtés seraient plus pessimistes en l'absence d'enquêteur. Toutefois, de nombreux auteurs posent la question de la difficile généralisation d'un résultat obtenu lors d'une expérimentation à l'ensemble des enquêtes. Dans une certaine mesure, cette prudence incite à construire des expérimentations spécifiques à une enquête donnée pour étudier dans quelle mesure l'introduction d'une collecte multimode, et plus particulièrement de la collecte par internet peut permettre de réduire les coûts sans détériorer l'erreur d'enquête totale, voire en la diminuant. Pour l'instant, ce qui ressort des premières expérimentations d'enquête internet/papier lancées par l'Insee depuis le début des années 2010, c'est que compte tenu des faibles taux de réponse, il est difficile d'affirmer que les effets de mode mesurés sont « nets » des effets de sélection. Aussi à court terme, tant que les taux de réponse ne se seront pas améliorés, il paraît plus sage de voir internet comme un mode complémentaire plutôt que le mode principal des enquêtes Ménages.

## Références

- [1] Andersen R., Kasper J., Frankel M. R. and associates "Total survey error" San Francisco, Jossey-Bass Publisher, 1979.
- [2] Couper M. P. "The Future of Mode Data Collection", *Public Opinion Quarterly*, 75 (5), 889-908, 2011.
- [3] Fripiat, D., Marquis, N. « Les enquêtes par internet en sciences sociales : un état des lieux » *Population*, 65(2), 309-338, 2010.