

LA COLLECTE MULTIMODE DANS LES ENQUÊTES MÉNAGES : UN POINT D'ÉTAPE DANS LE PLAN D'EXPÉRIMENTATION DE L'INSEE

Gaël de Peretti¹ & Heidi Koumarios² & Stéphane Legleye³ & Tiaray Razafindranovona⁴

Insee,

*Département des méthodes statistiques (Auteurs 1, 2 et 4)
Département de l'emploi et des revenus d'activité (Auteur 3)
18 boulevard Adolphe Pinard, timbre L120,
75675 Paris cedex 14*

*gael.de-peretti@insee.fr ; heidi.koumarios@insee.fr ;
stephane.legleye@insee.fr tiaray.razafindranovona@insee.fr*

Résumé. L'Insee (Institut national de la statistique et des études économiques) s'est lancé depuis le début des années 2010 dans un plan d'expérimentations d'enquête ménages par internet. L'introduction de ce nouveau mode de collecte, et plus généralement, le souhait de développer la collecte multimode, est un trait commun à de nombreux instituts nationaux de statistiques, et a relancé l'intérêt pour les travaux de méthodologie d'enquête souvent en lien avec la notion d'erreur totale d'enquête. En effet, si la collecte multimode et l'introduction du support internet est séduisant de prime abord, cela n'est pas sans soulever de nombreux problèmes lors des phases de réalisation d'une enquête : échantillonnage, conception, collecte, redressement, estimation, etc. L'objectif de ce plan était donc d'accumuler des compétences et des connaissances sur ce sujet pour voir dans quelle mesure, il serait possible d'introduire la collecte par internet dans les enquêtes ménages, et plus généralement développer la collecte multimode. Il s'agit là de faire un premier bilan de ces expérimentations et de voir dans quelle mesure, les résultats déjà obtenus vont aider à reconfigurer la façon de concevoir, organiser, réaliser les enquêtes, mais aussi de produire les informations statistiques.

Mots-clés. Statistique publique, erreur totale d'enquête, multimode, enquête ménages, méthodologie d'enquêtes

Abstract. Since the beginning of 2010, INSEE (National Institute for Statistics and Economic Studies) has embarked on a plan for household surveys using the Internet. The introduction of this new mode of collection, and more generally the desire to develop mixed-mode data collection, is a feature common to many national statistical institutes, and has revived interest in survey methodology often linked to total survey error. Indeed, mixed-mode survey and the introduction of Internet is attractive at first sight, but it raises many problems during the phases of a survey: sampling, design, collection, adjustment, estimation, etc. The objective of this plan was therefore to accumulate skills and knowledge on this subject to see to what extent it would be possible to introduce Internet in household surveys and more generally to develop mixed-mode survey. This is a first assessment of these experiments and it allows to see to what extent the results already obtained will help to reconfigure how to design, organize and carry out the surveys, but also to produce the results.

Keywords. Official statistics, total survey error, mixed-mode, household survey, survey methodology

1 Les enquêtes ménages par internet à l'Insee

Les premières expériences d'enquête ménages par internet ont commencé à la fin des années 2000 avec la mise en place d'une enquête auprès des non-répondants de l'enquête Emploi par internet ou papier en 2008. Depuis, on peut distinguer deux types d'enquête par internet auprès des ménages : les enquêtes intégrées dans le processus de production ; les expérimentations à visée

méthodologiques.

Concernant les premières, la plus importante est évidemment l'enquête annuelle de recensement. Après quatre années d'expérimentation de 2011 à 2014, essentiellement sur les aspects techniques (gestion de la collecte), l'Insee a intégré la collecte par internet en 2015. Ce mode de collecte a été choisi par 33 % des ménages (et 38 % des personnes, les ménages plus nombreux ayant une propension à répondre par internet plus importante) en 2015 et par 39 % des ménages (et 46 % des personnes) en 2016. L'effet le plus notable est la baisse de la non-réponse partielle (effet classique, voir par exemple, Razafindranovona 2015).

Ensuite, vient l'enquête Technologie de l'information et de la communication. L'introduction d'internet s'est faite progressivement après plusieurs opérations méthodologiques débutées en 2010 (Gombault, Duée, 2012 ; de Peretti, Razafindranovona, 2014). Il s'agit ici de développer le recours à la collecte par internet pour faire face aux problèmes de couverture tout en respectant des contraintes budgétaires. L'objectif est de maximiser la qualité d'enquête (Lyberg, 2012) en réduisant un des termes d'erreur d'enquête totale (Groves, Lyberg, 2010), en limitant voire diminuant les coûts d'enquête. Le principe est de faire en parallèle une enquête par téléphone, une enquête par internet/papier, et d'avoir un échantillon de contrôle pour corriger des effets de mode si nécessaire. Les constats établis à partir des deux dernières enquêtes (2015 et 2016) sont intéressants : le taux de réponse de l'enquête téléphonique est supérieur à celui d'internet/papier (70 % vs 60 %) ; une relance des non-répondants téléphone par internet/papier permet de remonter le taux de collecte de plus de 8 points ; les répondants internet/papier répondent majoritairement par papier ; les effets de mode (sélection et mesure), observées sur les deux variables centrales de l'enquête – taux d'équipement et taux d'utilisation d'internet dans les 3 derniers mois – sont concentrés sur des populations spécifiques comme les personnes de plus de 60 ans peu qualifiées.

De même, l'Insee réalise l'enquête Entrée dans la vie adulte (EVA) auprès d'une cohorte d'élèves entrés en 6^e en septembre 2007 en collaboration avec les Services statistiques ministérielles de l'Éducation nationale et de l'Enseignement supérieur et de la recherche (DEPP pour Direction de l'évaluation, de la prospective et de la performance, SIES pour sous-direction des Systèmes d'information et des études statistiques). En 2013 et 2014, l'Insee a interrogé les jeunes sortis précocement du système scolaire. En 2015, le dispositif complet se met en place avec une enquête de la DEPP sur les conditions de vie des lycéens et apprentis et une enquête du SIES sur le suivi des bacheliers. Il s'agit d'un panel avec une collecte multimode plutôt séquentielle :

- les enquêtés sont contactés par courrier (lettre-avis) et par courriel pour ceux pour lesquels on dispose d'une adresse mail
- une relance courrier est faite un mois plus tard avec l'ajout du questionnaire papier
- un mois plus tard, les non-répondants sont enquêtés par téléphone mais la réponse par internet ou papier est toujours possible (partie collecte multimode concurrentiel)

En 2016, le taux de collecte est de 67 % dont 36 % par téléphone, 23 % par internet et 8 % par papier. La part de répondant par internet a fortement progressé par rapport à 2015 (38 % téléphone, 9 % internet et 12 % papier) et pour deux raisons essentielles : plus d'adresse mail pour réaliser la collecte ; modification du protocole avec envoi du questionnaire papier seulement lors de la relance et non dès l'envoi de la lettre-avis comme c'était le cas les années précédentes. Pour la première fois, en 2016, on peut interroger des personnes qui avaient participé en 2015 à l'une des trois enquêtes organisées par l'Insee, la DEPP ou le SIES. Les taux de réponse divergent fortement selon ces trois sous-échantillons : 61 % (Insee), 67 % (DEPP) et 91 % (SIES).

Il existe d'autres enquêtes ménages par internet, mais de moindre importance, soit qu'elles consistent à offrir une alternative à un questionnaire papier auto-administré papier qui est un complément d'un dispositif d'enquête classique en face à face (QAA sur le bien-être pour l'enquête Statistiques sur les ressources et conditions de vie ou SRCV, enquête non-répondants de l'enquête Emploi), soit qu'il s'agit d'une enquête à périodicité moins régulière (enquête sur les salariés de la

Fonction publique d'État, en complément de l'enquête sur le coût de la main d'œuvre et la structure des salaires ou Ecmoss).

2 Le plan d'expérimentations de l'Insee

Au-delà de ces dispositifs, l'Insee a mis en place un plan d'expérimentations en 2012 pour répondre aux questions méthodologiques que suscitent les enquêtes multimode, en particulier la difficulté à produire des résultats généraux sur ces enquêtes (Couper, 2011).

Le recours à la collecte multimode, et en particulier à la collecte par internet poursuit deux objectifs par forcément réconciliables : réduire les coûts d'enquête ; réduire l'erreur totale d'enquête (Andersen, Kasper, Frankel *et al.*, 1979 ; Groves, Lyberg, 2010) et donc améliorer la qualité des enquêtes (Lyberg, 2012).

Le plan d'expérimentations repose sur les principes suivants :

- une expérimentation pour chaque enquête réalisée en parallèle de l'enquête principale pour prendre en compte les spécificités de chaque enquête
- une limitation maximale des interférences entre l'enquête principale et l'expérimentation en termes d'échantillonnage, de calendrier et de charge de travail
- une volonté de répondre à chaque expérimentation à une nouvelle question afin d'accumuler progressivement des connaissances et de l'expérience sur ces questions
- un protocole standard : échantillonnage dans les fichiers fiscaux et premier contact via un envoi de courrier de masse ; questionnaire extrait de l'enquête principale et réduit à 20 minutes en moyenne ; l'envoi d'une lettre-avis, puis de deux relances espacées chacune de 3 semaines (avec le questionnaire papier équivalent au questionnaire web envoyé à la première relance)
- un plan d'exploitation standard : bilan sur les taux de réponse ; exploitation brute i.e. caractéristiques des répondants, qualité des réponses sur les variables cibles, comparaison si possible des informations collectées et des informations présente dans la base de sondage ; apurement et imputation ; calage/redressement pour séparer autant que faire se peut les effets de sélection des effets de mesure ; estimation des résultats et conclusions sur l'expérimentation

Ce plan a déjà été décrit (de Peretti et Razafindranovona, 2013 ; de Peretti, Razafindranovona, 2014) et ici nous ne reviendrons que sur les résultats des 3 dernières expérimentations exploitées (voir section 3)

3 Les principaux résultats des enquêtes « Vols violence et sécurité », « Qualité de la vie au travail », « Logement-web »

Les résultats présentés sont ici sont un rapide résumé de trois documents de travail détaillés sur chacune de ces expérimentations (Razafindranovona, 2016, 2016, 2017). Il s'agit de pointer ce qui semble les points saillants de ces expérimentations.

3.1 « Vols violence et sécurité (VVS) »

L'objectif central de cette enquête était de voir dans quelle mesure il était possible de produire des informations sur la victimation à partir d'un dispositif léger (papier/internet), tout en s'intéressant à l'étude sur l'auto-sélection des répondants grâce à un dispositif particulier : un sous-échantillon avec la possibilité de répondre pour deux individus ; un sous-échantillon avec un questionnaire ménages en début d'enquête ; un sous-échantillon de jeunes dont on ne connaît pas l'identité dans la base de sondage. Par la suite, l'individu à interroger sera nommé Kish. La comparaison est faite avec l'enquête Cadre de vie et sécurité (CVS) réalisée depuis 2004 en face à face. Compte tenu de la stabilité dans le temps des estimations, d'un bon taux de réponse (75 %) et de l'adéquation de cette enquête avec des données administratives sur les plaintes enregistrées après victimation, l'enquête CVS est considéré comme l'étalon auquel doit se comparer VVS.

Sur les 40 000 lettre-avis envoyés, 4 500 sont en NPAI (le destinataire n'est pas connu par La Poste). Sur les 35 500 destinataires, on constate 21 000 non-répondants et parmi les 14 500 répondants (53 % papier, 47 % internet), 1 500 questionnaires sont inexploitable soit que le répondant n'est pas le Kish, soit que le questionnaire est incomplet. Le taux de collecte est donc de 32 %, le taux de réussite de 37 % (élimination des NPAI). Globalement, les répondants internet sont plus jeunes (âge médian 44 vs 48), plus riches (revenu annuel médian de 36,5 k€ vs 27,7 k€), plus diplômés (45 % ont un diplôme du supérieur vs 24%).

Le résultat le plus saillant est la tendance à la surestimation des victimations, surestimation difficile à corriger malgré des calages sur marges nombreux et s'appuyant sur les données d'intérêt de l'enquête (Tableau 1). Quatre calages successifs sont réalisés : calage sur marges sociodémographiques (CMS) ; CMS et taux de plainte (CMS +TP) ; CMS +TP et sentiment d'insécurité (CMS + TP + SI) ; CMS +TP + SI et accès à internet (CM total).

Tableau 1 : Prévalence des victimations dans les enquêtes VVS et CVS

En %	Vols avec violence	Vols sans violence	Violences physiques	Menaces
Après CNR	2,6	6,4	2,7	7,3
CMS	2,9	6,4	2,9	7,7
CMS + TP	2,8	5,9	2,7	7,2
CMS + TP + SI	1,9	4,4	1,8	4,9
CM Total	1,8	4,2	1,7	4,8
<i>Référence CVS</i>	<i>1,0</i>	<i>2,8</i>	<i>2,2</i>	<i>4,5</i>

Source : VVS, CVS 2013.

Champ : personnes âgées de 14 ans ou plus et vivant dans un logement ordinaire en France métropolitaine.

Le calage sur les variables d'intérêt mais non centrales de l'enquête doit plutôt être vu comme un exercice que comme une pratique à adopter. Il y a par exemple une sous-estimation des plaintes dans VVS, ce qui pourrait être dû au fait que les victimes profitent de cette enquête pour parler de leur victimation pour laquelle ils n'ont pas jugé nécessaire ou pas osé porter plainte. De même, les personnes ayant un sentiment d'insécurité sont sur-représentées : peut-être est-ce ce sentiment d'insécurité qui a poussé les personnes à répondre. Dans tous les cas, ces calages successifs amènent à une très forte distorsion des poids, ce qui pose la question de la capacité à corriger les effets de sélection et à considérer les effets résiduels comme des effets de mesure.

Afin d'approcher cet effet motivation à répondre, on compare les résultats obtenus avant et après relance. Les personnes ayant répondu avant les relances sont toujours plus victimes. Ces personnes sont de fait toutes des personnes ayant répondu par internet, ce qui implique d'en disposer chez soi. Toutefois, cet effet persiste même après contrôle des caractéristiques sociodémographiques pour tous les types de victimation (Tableau 2).

Tableau 2 : Contrôle de l'effet motivation des répondants

	Vols avec violence	Vols sans violence	Violences physiques	Menaces
Avant relance (en %)	2,9	8,2	2,8	10,3
Après relance (en %)	2,4	6,4	2,7	7,2
Odd-ratios (Après vs Avant relance)	0,80	0,78	0,92	0,65

Source : VVS 2013.

Champ : personnes âgées de 14 ans ou plus et vivant dans un logement ordinaire en France métropolitaine, ayant répondu par internet.

3.2 « Logement-web (LW) »

Il s'agit là de tester la capacité à collecter une information de qualité sur les loyers sans

l'intermédiation d'un enquêteur pour expliquer certains concepts ou insister pour le recours à des documents.

Sur les 40 000 individus échantillonnés, 7 500 n'ont pas reçu la lettre-avis (NPAI), 12 000 ont répondu dont 5 000 par internet et 7 000 par papier. Classiquement, la qualité est moins bonne sur papier que sur internet : plus de non-réponse partielle sur papier, mauvaise gestion des filtres. Les personnes ayant déclaré avoir recours à un document pour répondre font moins d'arrondi (5 % à la centaine vs 16 %). Afin de pouvoir comparer les résultats de l'enquête en face à face avec l'enquête par internet/papier, une succession d'opération est réalisée : apurement ; correction de la non-réponse ; calage sur les totaux de la base de sondage ; corrections liées aux charges, garage, aides. À l'issue de ces traitements l'écart constaté sur le loyer moyen est de + 31 € (518 vs 487), soit une différence relative de 6 %.

Afin d'éliminer certains effets de sélection, on applique plusieurs méthodes intégrant de nombreuses variables de contrôles : surface du logement, région, taille de l'unité urbaine, nombre d'habitants du logement, nombre de pièces, type de propriétaire, date d'achèvement, étage, ascenseur, date d'emménagement, diplôme, entité à laquelle est versé le loyer. Les méthodes testées sont : la régression linéaire, la repondération par score de propension, le calage sur marge. Quelle que soit la méthode, elles réduisent l'écart entre les deux enquêtes, mais il reste conséquent de l'ordre de 25 € soit 5 %. Toutefois, compte tenu du faible taux de réponse, il est difficile de savoir s'il s'agit d'un effet de mode lié aux seuls effets de mesure ou s'ils perdurent des effets de sélection.

3.3 « Qualité de la vie au travail (QVT) »

L'enquête QVT est le pendant internet/papier de l'enquête Conditions de travail. Au-delà de la comparaison des deux enquêtes, QVT avaient d'autres objectifs : test d'un protocole internet pur (i.e. sans possibilité de répondre via un questionnaire papier) ; effet de l'argumentaire de la lettre-avis (l'un orienté conditions de travail, l'autre risques psychosociaux) ; tests de question pour la future enquête sur les risques psychosociaux.

Concernant les taux de réponse, le protocole internet pur atteint 44 %, le protocole internet/papier 52 % (dont 32 % par internet). On retrouve des résultats classiques sur les différences entre les répondants papier et internet : les derniers sont souvent plus qualifiés et perçoivent des revenus plus élevés. Les effets de l'argumentaire sont faibles en termes de taux de réponse mais semblent avoir sélectionnées des personnes différentes. Enfin, la Dares (Direction de l'animation de la recherche, des études et des statistiques, SSM du ministère du travail) a tiré profit des questions tests sur les risques psychosociaux pour créer des questions fermées grâce à la post-codification des questions ouvertes testées et ajouter des modalités à des questions fermées après l'exploitation des libellés des questions « précisez ».

Afin de comparer les deux enquêtes, différents types de calage ont été étudiées : sexe/âge/diplôme/nationalité ; le même + le groupe social ; le même + le score de bien-être de l'OMS ; le même + la santé ressentie. Globalement, ces calages ont tendance à réduire les écarts, le bien-être réduisant les écarts sur les questions plus subjectives, les caractéristiques sociodémographiques réduisent les écarts sur les variables factuelles. Les effets de mode sont là encore de deux ordres :

- effet de mesure : ordonnancement des questions différents, formulation différente, satisficing pour les batteries de question ; désirabilité sociale plus forte en face à face
- sélection non contrôlée : des similarités avec VVS ; l'absence de la variable secteur d'activité dont on connaît les liens avec les conditions de travail

4 Quelle suite donner à ces expérimentations

L'Insee a encore plusieurs expérimentations à exploiter comme celles sur l'estimation du prix de la

résidence principale, la conjoncture auprès des ménages et les tests autour de l'enquête Emploi.

Toutefois, au regard des premières exploitations, si l'on souhaite développer la collecte par internet, il semble important : d'améliorer les taux de réponse ; de mieux concevoir les questionnaires pour limiter les différences de mesure entre mode ; d'attacher du soin à l'introduction des bonnes variables de contrôle pour faciliter l'élimination des effets de sélection.

Concernant le taux de réponse, la phase de contact est primordiale. De ce fait, des pistes sont explorées pour améliorer les courriers et courriels adressés et pour multiplier les canaux de contact afin d'augmenter les chances de convaincre les personnes échantillonnées de répondre.

Concernant les questionnaires, au-delà des réflexions sur la conception du questionnaire, des travaux sont en cours pour uniformiser les questionnaires web d'un point de vue ergonomique, en s'appuyant sur les principes de Morrison (2008), en ayant un outil commun et en développant à terme un portail d'enquêtes ménages par internet qui permettent aux enquêtés de mieux comprendre par eux-mêmes les enjeux de ces enquêtes.

Enfin, des travaux méthodologiques sont engagés pour développer des méthodes standards qui permettent d'agréger des informations recueillies sur plusieurs modes. Les choix dépendent à la fois du protocole (multimode concurrentiel ou séquentiel), de la capacité ou non à avoir un mode étalon, des tailles des échantillons de répondant selon les modes, etc.

Bibliographie

- [1] Andersen R., Kasper J., Frankel M. R., and associates (1979), *Total survey error*, San Francisco, Jossey-Bass Publisher.
- [2] Couper, M. (2011), "The Future of Modes of Data Collection", *Public Opinion Quarterly*, 75(5), p. 889-908.
- [3] Gombault, V. et Duée, M. (2012), « Un exemple d'enquête multimode à l'Insee : l'enquête TIC », article présenté au 7^e colloque francophone sur les Sondages de la Société française de statistique, Rennes.
- [4] Lyberg L. (2012), « La qualité des enquêtes », *Techniques d'enquête*, 38(2), p. 115-142.
- [5] Morrison R. L. (2008), *Rédaction et révision des lignes directrices pour la conception de questionnaires*, Symposium Statistique Canada 2008.
- [5] de Peretti, G. et Razafindranovona, T. (2013), « Les enquêtes multimode : multi-problème ou multi-solution ? », article présenté aux 45^e Journées de Statistique de la SFdS, Toulouse.
- [6] de Peretti G., Razafindranovona T. (2014), *Les enquêtes multimode au prisme de l'erreur totale d'enquête*, Recueil du Symposium 2014 de Statistique Canada, Au-delà des méthodes traditionnelles d'enquête : l'adaptation à un monde en évolution.
- [7] Razafindranovona T. (2015), La collecte multimode et le paradigme de l'erreur d'enquête totale, *Document de travail Méthodologie statistique*, n° M2015/01, mars, Insee.
- [8] Razafindranovona T. (2016), Exploitation de l'enquête expérimentale Vols, violence et sécurité, *Document de travail Méthodologie statistique*, n° M2016/03, juin, Insee.
- [9] Razafindranovona T. (2016), Exploitation de l'enquête expérimentale Logement internet/papier, *Document de travail Méthodologie statistique*, n° M2016/08, décembre, Insee.
- [10] Razafindranovona T. (2017), Exploitation de l'enquête expérimentale Qualité de vie au travail, *Document de travail Méthodologie statistique*, n° M2017/01, février, Insee.