

La collecte multimode à l'Insee : d'un effet de mode à un effet sur nos travaux

**Un point d'étape
dans le plan d'expérimentations Insee**

François Beck, Gaël de Peretti

Plan

- Enquêtes multimode : définitions et enjeux
- Quelques éléments de comparaison entre les modes
- Contexte de la collecte par internet des enquêtes ménages
 - La couverture internet
- Résultats des enquêtes internet ménages en production
- Les enquêtes expérimentales menées à l'Insee
 - VVS
 - Logement web
 - Qualité de vie au travail
- Point sur les conclusions des GT lancés autour du développement de la collecte internet dans les enquêtes ménages
- Conclusion

Enquêtes multimode : définitions et enjeux

- Utilisation de canaux de communication variés au cours d'une même enquête
 - Multimode intégré
 - Multimode concurrentiel
 - Multimode séquentiel
 - Adaptative design
- Tirer le meilleur parti de chaque mode
 - Faire face à la baisse des taux de participation aux enquêtes, atteindre les individus qui ne veulent plus répondre aux enquêtes
 - Recours à internet pour répondre aux attentes de certains enquêtés
 - Optimiser le travail des enquêteurs
 - Des modes plus adaptés aux questions sensibles

Comparaison entre les modes

- Taux de réponse :
 - Face à face > téléphone > papier > internet
- Coûts :
 - Face à face > téléphone > papier > internet
- Des effets contrastés :
 - En l'absence d'enquêteur, moins de désirabilité sociale mais plus de *satisficing*
 - Une gestion des filtres très difficile sur questionnaire papier
 - Baisse de la non réponse partielle sur internet (vs papier ; Razafindranovona, 2015)
 - Sur les sujets sensibles : des modes genrés et générationnels ?
 - Pas vraiment de consensus sur la manière de mesurer l'ampleur des effets de mode
 - Résultats très dépendants des thèmes abordés, de la population cible

Avantages et inconvénients de la collecte par internet

- Avantages :
 - Rapidité
 - Coûts
 - Interactivité
 - Possibilité d'introduire des filtres complexes
 - Moment du remplissage à discrétion du répondant
 - Recours à des contenus variés (vidéos, son, image...)
 - Mode bien adapté aux 15-30 ans qui sont difficiles à « accrocher »
- Mais...
 - Problèmes de non réponse, d'auto-sélection
 - Moins d'accompagnement pour motiver l'enquêté à répondre, pour expliciter les instructions (problèmes de littératie)
 - Problèmes de couverture
 - Problèmes techniques (compatibilité des navigateurs, des logiciels de lecture des vidéos et du son)
 - Nécessité de plus en plus criante d'un *responsive design*

Quelques éléments sur la couverture (TIC 2007-2015)

- Le taux d'équipement des individus en internet (84%) est en nette hausse par rapport à 2007 (54%)
- La hausse entre 2013 et 2016 est particulièrement forte pour les plus de 60 ans
- Le taux d'utilisation quotidienne ou quasi-quotidienne d'internet est passé de 36 % en 2007 à 65 % en 2016
- Là encore la hausse la plus forte entre 2013 et 2016 est observée pour les plus de 60 ans
 - 60-74 ans : (13 % en 2007 ; 40 % en 2013 ; 45 % en 2016)
 - 75 ans et + : (3 % en 2007 ; 11 % en 2013 ; 14 % en 2016)

Contexte de la collecte par internet des enquêtes Ménages

- Enquêtes internet intégrées dans le processus de production : EAR, TIC, EVA, etc.
- Enquêtes expérimentales :
 - Enquête internet/papier en parallèle de l'enquête face à face (ou téléphone)
 - Protocole standard : échantillonnage dans les fichiers fiscaux, courrier avis, 2 relances dont une avec le questionnaire papier, ce dernier est un extrait de l'enquête référente
 - Exploitation standardisée avec un objectif de comparaison des deux enquêtes et d'identification des effets de mesure

Résultats des enquêtes internet ménages en production

- Enquête annuelle de recensement : internet proposé depuis 2015 ; hausse annuelle du TR ; plus de la moitié des individus en 2017 (54 % des BI et 46 % des FL) ; baisse de la NR partielle
- Technologies de l'information et la communication : limiter les défauts de couverture de l'enquête téléphonique et les coûts de collecte ; TR 70 % par tél vs 60 par internet/papier ; effet de mode concentré sur population spécifique (60 ans et plus, peu qualifiés)
- Entrée dans la vie adulte : dispositif mixte séquentiel /concurrentiel tél /internet / papier ; population de jeunes entrés en 6^e en 2007 ; trois sous-populations élèves du secondaire, du supérieur, dans la vie active ; taux de collecte 36 % tél, 23 % internet, 8 % papier

Les enquêtes expérimentales

- Vols, violence et sécurité : comparaison avec CVS (enquête victimation) ; contrôle de l'auto-sélection avec plusieurs dispositifs (2 réponses possibles, enquête ménages au départ, mineurs)
- Logement web : capacité à produire de l'information de qualité sur un sujet complexe, le loyer (par ex. difficulté à séparer « loyer pur » des charges)
- Qualité de vie au travail : comparaison avec Conditions de travail ; protocole internet pur ; test de questions sur les risques psychosociaux ; test lettre-avis

Vols violence et sécurité (1)

- Réponse : 40 000 courriers envoyés dont 4 500 PND, 14 500 répondants (6 800 internet), 1 500 questionnaires inexploitable (pas le kish, questionnaire très incomplet) soit 32 % de taux de collecte
- Répondants internet : 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%)
- Surestimation des victimations (4 calages sur marges sociodémographiques CMS, taux de plainte TP, sentiment d'insécurité SI, accès internet)

Vols violence et sécurité (2)

	Vols avec violence (%)	Vols sans violence (%)	Violences physiques (%)	Menaces (%)
Après CNR	2,6	6,4	2,7	7,3
Calage sur marges sociodémographiques	2,9	6,4	2,9	7,7
+nombre de plaintes	2,4	5,4	2,8	7,4
+nombre de plaintes + sentiment d'insécurité	2,1	4,6	2,2	5,6
+nombre de plaintes + sentiment d'insécurité + accès à Internet	1,8	4,0	2,0	5,2
CVS 2013	1,0	2,8	2,2	4,5
Intervalle de confiance - CVS 2013	[0,8 : 1,2]	[2,5 : 3,1]	[1,9 : 2,5]	[4,1 : 5,0]

TABLE 4.2 – Taux de victimation de VVS après calages successifs

Vols violence et sécurité (3)

- Effet de sélection ? Étude de la « motivation »

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.

Logement web

- Réponse : 40 000 courriers envoyés dont 7 500 PND, 12 000 répondants (5 000 internet)
- NRP plus importante et problème filtre sur papier
- Effet arrondi plus fort pour les personnes n'ayant pas eu recours à un document
- Écart de loyer :
 - CNR et apurement : 31€ sur le loyer moyen (+6 % en relatif)
 - Correction sélection observable : 25€ soit 5 %
- Effet de sélection ou effet de mesure ?

Qualité de vie au travail (1)

- Taux de collecte: internet plus qualifiés, revenus plus élevés ; lettre-avis : taux de réponse identique mais personnes sélectionnées différentes

SSECH	Internet	Papier	NPAI	Non-réponse	Autres	Total
1	2 924	1 527	1 164	3 796	589	10 000
2	2 756	1 638	1 125	3 863	618	10 000
3	3 738	1	1 247	4 055	959	10 000
4	3 751	1	1 178	4 046	1 024	10 000
Total	13 169	3 167	4 714	15 760	3 190	40 000

Sous-échantillon 1 – lettre-avis A, internet et papier modes de réponse possibles

Sous-échantillon 2 - lettre-avis B, internet et papier modes de réponse possibles

Sous-échantillon 3 - lettre-avis A, internet seul mode de réponse possible

Sous-échantillon 4 - lettre-avis B, internet seul mode de réponse possible

Autres : non équipés, refus, enquête non-répondants

(la lettre-avis B insiste sur les risques psychosociaux)

taux de réponse : papier+internet 52 % (dont 32 pts internet) ; internet seul 44 %

Taux de réponse plutôt meilleur que d'habitude mais population de personnes en emploi

Qualité de vie au travail (2)

- Questions test ont permis d'améliorer le questionnaire de l'enquête RPS 2015
- Calages successifs (sociodémographique, Groupe Social, bien-être OMS, santé ressentie) :
 - Factuel mieux « géré » par sociodémographique, subjectif par bien-être
 - Effet de mesure : ordonnancement des questions différent, formulation différente, *satisficing* pour les batteries de questions ; 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

Premiers enseignements du plan d'expérimentations

- Nécessité d'améliorer les taux de réponse : importance de la phase de contact et donc de la capacité à mobiliser plusieurs vecteurs (courrier, tél, internet) ; adaptative design pour limiter les coûts
- Mieux concevoir les questionnaires : limiter la complexité, la durée ; uniformiser les questionnaires web d'un point de vue ergonomique ; revoir le protocole (Panel web) ; créer un portail enquêtes ménages adapté pour inciter, motiver, intéresser les enquêtés ; prévoir des variables de contrôle pour mieux corriger les effets de sélection
- Développer des méthodes standards d'agrégation des données : lien avec protocole (séquentiel ou concurrentiel) ; existence d'un mode étalon ; taille des échantillons de répondants selon les modes

Point sur les conclusions des GT

- GT « Amont » :
 - Approche pragmatique : Développer la collecte par internet sans abandonner les autres modes, à cause des nombreuses limites du CAWI
 - Les enquêtes doivent devenir plus modulaires
 - Le niveau de qualité cible est à définir enquête par enquête
- GT « Aval » :
 - Constat d'un « effet de mode », se décomposant en un biais de sélection et un effet de mesure
 - Le biais de sélection sur variables observables peut être corrigé par des méthodes classiques de correction de la non-réponse et/ou de calage
 - Malgré ces corrections, subsiste encore un effet de mode « résiduel » intégrant un effet de mesure **et** un biais de sélection sur variables non observables
 - Nécessité de mettre en œuvre des techniques de détection de ce type de biais et des effets de mesure « purs »
 - Nécessité d'augmenter les taux de réponse par internet et de continuer les expérimentations
 - Travailler sur les méthodes d'agrégation

Point sur les conclusions des GT

- GT « Maîtrise d'œuvre organisationnelle » :
 - Nécessité de mettre en place un guide de mise en œuvre du multimode destiné aux concepteurs et CPOS
 - Nécessité de développer le rôle du CPOS référent internet, en tant que référent pour les autres CPOS des enquêtes ayant un volet internet et en l'associant aux COPIL des enquêtes concernées
- GT « Outils et infrastructure informatiques » :
 - Démarrage en mars 2017
 - Mutualisation des développements communs à toutes les enquêtes : Réflexion autour de la filière Pogues/ENO/ RMÉS
 - Un portail d'enquêtes ménage unique et attractif

Conclusion

- Expérimentations à exploiter : enquêtes sur l'estimation du prix de la résidence principale, la conjoncture auprès des ménages, l'emploi...
- Un important développement d'outils informatiques de la conception à la gestion de la collecte
- Réflexions à venir sur l'utilisation des parodonnées
- Vaste chantier qui va alimenter les travaux des méthodologues d'enquête
- Le face à face a encore de belles heures devant lui (William Thorn, OCDE, mai 2017)

Bibliographie

- 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.
- Razafindranovona T. (2016), Exploitation de l'enquête expérimentale Vols, violence et sécurité, *Document de travail Méthodologie statistique*, n° M2016/03, juin, Insee.
- Razafindranovona T. (2016), Exploitation de l'enquête expérimentale Logement internet/papier, *Document de travail Méthodologie statistique*, n° M2016/08, décembre, Insee.
- Razafindranovona T. (2017), Exploitation de l'enquête expérimentale Qualité de vie au travail, *Document de travail Méthodologie statistique*, n° M2017/01, février, Insee.