

Méthodologie statistique

M2016/03

**Exploitation de l'enquête
expérimentale
Vols, violence et sécurité**

Tiaray Razafindranovona

Document de travail



Institut National de la Statistique et des Études Économiques

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES

Série des documents de travail « Méthodologie Statistique »

de la Direction de la Méthodologie et de la Coordination Statistique et Internationale

M 2016/03

Exploitation de l'enquête expérimentale Vols, violence et sécurité

Tiaray Razafindranovona*

Ce document regroupe différents travaux autour de l'expérimentation Vols, violences et sécurité et de sa comparaison avec l'enquête Cadre de vie et sécurité réalisés au Département des méthodes statistiques et à la Division conditions de vie des ménages de l'Insee. Je remercie Bruno Dietsch, Marine Guillerm, Carine Burricand, Michel Duée et Lucile Jamet pour leurs contributions à l'exploitation de cette expérimentation.

Je remercie également Gaël de Peretti et Olivier Sautory pour leur relecture attentive et leurs commentaires constructifs. Je reste seul responsable des erreurs ou omissions qui pourraient rester.

* DMCSI – DMS (Département des Méthodes Statistiques) – Division Recueil et traitement de l'information
18, bd Adolphe Pinard – 75675 PARIS CEDEX 14

Direction de la méthodologie et de la coordination statistique et internationale -Département des Méthodes Statistiques - Timbre L101
18, bd Adolphe Pinard - 75675 PARIS CEDEX - France -
Tél. : 33 (1) 41 17 66 33 - Fax : 33 (1) 41 17 61 97 - CEDEX - E-mail : -DG75-L001@insee.fr - Site Web Insee : <http://www.insee.fr>

*Ces documents de travail ne reflètent pas la position de l'Insee et n'engagent que leurs auteurs.
Working papers do not reflect the position of INSEE but only their author's views.*

Exploitation de l'enquête expérimentale Vols, violence et sécurité

Tiaray Razafindranovona *

Résumé

Afin d'accumuler de la connaissance et de bien cerner les difficultés méthodologiques que pose la collecte par Internet ou multimode auprès des ménages, l'Insee a programmé sur plusieurs années un plan d'expérimentations. L'enquête web/papier « Vols, violence et sécurité » (VVS) est l'une de ces expérimentations, réalisée en 2013, en parallèle de l'enquête de victimation en face à face « Cadre de vie et sécurité » (CVS). La convergence espérée des taux de victimation de ces deux enquêtes n'est pas obtenue, et ce, même au prix de calages qui posent question quant à leur applicabilité en production. Les écarts relèveraient non pas des effets de mode « purs », mais plutôt de la sélection non contrôlée.

Outre cette question centrale, différents points sont abordés dans ce document comme le déroulement de la collecte, les déterminants et la correction de la non-réponse ou encore la qualité des données recueillies et le *satisficing*.

Mots-clés : enquête multimode, collecte par internet, effets de mode, effets de sélection, enquête de victimation, taux de victimation, méthodologie d'enquête.

Abstract

Insee decided to launch a series of experimental households surveys to increase knowledge in the field of internet data collection and to identify methodological issues. One of these experimental surveys (web and paper), entitled "Vols, violence et sécurité (Thefts, violence and safety)" (VVS) took place in 2013, in parallel of the French victimisation survey (face-to-face) "Cadre de vie et sécurité (Living environment and safety, close to EU Safety Survey project)" (CVS). Convergence of victimisation rates is not achieved for these two surveys even with calibration techniques that would be questionable in a real production process. Differences between rates are probably more due to uncontrolled selection effects than "pure" mode effects.

Besides this main question, this document tackles subjects such as progress of data collection, nonresponse factors and correction or quality of collected data and satisficing.

Keywords : mixed-mode survey, web data collection, mode effects, selection effects, victimisation survey, victimisation rates, survey methodology.

* INSEE, Département des méthodes statistiques, tiaray.razafindranovona@insee.fr

Table des matières

1	Enquêtes de victimation, plan d'expérimentations multimode de l'Insee et objectifs	5
1.1	Les enquêtes de victimation	5
1.2	L'enquête Cadre de vie et sécurité	5
1.3	Le plan d'expérimentations d'enquêtes multimode à l'Insee	6
1.3.1	Protocole standard des expérimentations	6
1.3.2	Enquêtes expérimentales de l'Insee	7
1.4	L'enquête expérimentale Vols, violence et sécurité	7
1.5	Auto-sélection et protocoles spécifiques de l'enquête VVS	7
2	Bilan de la collecte et premières exploitations brutes de VVS	8
2.1	Déroulement de la collecte et taux de réponse	8
2.1.1	Collecte des réponses web	8
2.1.2	Le web : un mode plus "flexible" pour les actifs occupés ?	9
2.1.3	Ressenti des répondants web sur le questionnement	10
2.2	Non-réponse partielle, qualité de réponse	11
2.2.1	Non-réponse partielle	11
2.2.2	Qualité des réponses	12
3	Enseignements après apurements et imputations	16
3.1	Apurement et imputations	16
3.2	Taux de réponse	17
3.2.1	Caractéristiques des répondants	17
3.2.2	Déterminants de la non-réponse et de la non-réception des courriers	18
3.3	Premières estimations sur variables clés	21
3.3.1	Taux de victimation	21
3.3.2	Retour sur les questionnaires écartés	21
3.4	Comparaison entre les deux types de protocoles	22
3.4.1	Nombre de réponses	22
3.4.2	Victimations selon le protocole	23
4	Calages et comparaison des taux de victimation	24
4.1	Correction de la non-réponse totale	24
4.2	Principes des calages sur marges successifs	25
4.2.1	Le calage pour contrôler la sélection	25
4.2.2	L'enquête CVS comme bon étalon de la victimation	25
4.2.3	Les calages sur marges successifs	25

4.3	Quelle convergence des taux de victimation ?	27
4.4	Le problème de la sélection non contrôlée	28
4.4.1	La sélection non contrôlée comme principale explication des écarts entre CVS et VVS	28
4.4.2	Une illustration de la sélection non contrôlée	29
5	Des effets de mode ?	31
5.1	De la difficulté à distinguer effets de sélection et effets de mode	31
5.2	Quelques constats sur la possible mauvaise compréhension de certains concepts .	32
5.2.1	Des poly-victimes plus fréquents dans VVS	32
5.2.2	Analyse des double comptes	32
5.3	Elements sur l'utilisation de l'analyse de données multivariée (AFM duale) . . .	33
A	Comparaison des taux de victimation web / papier	40
A.0.1	Différences brutes	40
A.0.2	Principes du contrôle de la sélection par <i>kernel matching</i> sur score de propension	41
A.0.3	Résultats	44
B	Enquête sur les non-répondants à l'expérimentation VVS	47

Introduction

Depuis quelques dizaines d'années, l'analyse quantitative de la délinquance ne s'appuie pas uniquement sur les remontées des administrations mais aussi sur des enquêtes en population générale, appelées enquêtes de victimation. Ces enquêtes interrogent un échantillon d'individus sélectionnés aléatoirement sur les atteintes à la personne ou aux biens qu'ils auraient subis. Le nombre de personnes victimes d'atteintes est ainsi aujourd'hui estimé à partir des données de l'enquête annuelle "Cadre de vie et sécurité" réalisée par l'Insee en face à face donc via l'intermédiation d'enquêteurs.

Quels résultats sur les victimations seraient obtenus si les enquêtés étaient interrogés sans l'interface avec un enquêteur ? En l'absence d'enquêteur, la motivation à répondre est moindre : *in fine* non seulement la proportion de répondants peut être plus faible mais il peut également y avoir des différences dans les caractéristiques, observables ou non, des répondants selon le type de protocole de contact mis en oeuvre. Un autre résultat possible est celui d'une différence dans les réponses collectées qui relève strictement du mode de collecte : dans certains cas, un individu ne répond pas de la même manière face à un enquêteur et seul face à son écran d'ordinateur ou son questionnaire au format papier.

Pour essayer de mieux comprendre ces différents phénomènes, une enquête de victimation expérimentale et auto-administrée, appelée "Vols, violences et sécurité", a été réalisée par l'Insee en 2013. Ce document regroupe différentes contributions autour de cette expérimentation et de sa comparaison avec l'enquête en face à face réalisées au sein du Département des méthodes statistiques de l'Insee¹ en collaboration avec la Division conditions de vie des ménages².

1. Tiaray Razafindranovona, Bruno Dietsch et Marine Guillerm

2. Carine Burricand, Michel Duée et Lucile Jamet

Chapitre 1

Enquêtes de victimation, plan d'expérimentations multimode de l'Insee et objectifs

Cette partie fournit quelques éléments de contexte sur les enquêtes de victimation : depuis quand existent-elles et qu'apportent-elles par rapport aux remontées administratives sur les faits délictueux ? Le contexte général des expérimentations d'enquêtes multimode (offrant la possibilité de répondre par internet) à l'Insee sera également présenté avant de préciser les objectifs et le protocole de l'enquête expérimentale Vols, violences et sécurité.

1.1 Les enquêtes de victimation

Les enquêtes de victimation sont apparues aux États-Unis dans les années 1960 et offrent une approche alternative de la mesure de la délinquance à celle fondée sur les seuls comptages administratifs (Zauberman et Robert, 2011). En effet, les données administratives présentent certaines lacunes bien connues comme leur sensibilité aux pratiques d'enregistrement ainsi qu'aux modifications d'ordre législatif. Ainsi, les évolutions mesurées à partir d'enquêtes de victimation sont *a priori* moins sensibles aux modifications du cadre législatif ou administratif que celles résultant de l'exploitation des remontées provenant des services de police ou de gendarmerie. Ces comptages administratifs peuvent en effet être considérées par certains plus comme un reflet de l'activité des services les produisant (Desrosières, 2004) que de la réalité de la délinquance. Mais plus qu'une opposition frontale entre ces deux instruments de mesure, c'est la complémentarité de ces différentes sources qui permet d'offrir un cadre d'analyse global, un panorama des phénomènes délictueux.

1.2 L'enquête Cadre de vie et sécurité

En France, les enquêtes de victimation sont utilisées depuis environ 40 ans et depuis 1996, l'Insee est un acteur majeur dans la mesure des victimations à partir de ce type d'enquêtes. Cette mesure était auparavant réalisée via le module de victimation des enquêtes EPCV (Enquête permanente sur les conditions de vie des ménages) remplacé depuis 2007 par l'enquête de victimation, à part entière, appelée Cadre de vie et sécurité (CVS).

L'enquête de victimation CVS est une enquête obligatoire en population générale, qui vise à connaître les faits de délinquance dont les ménages et leurs membres ont pu être victimes dans les deux années précédant l'enquête. Elle permet donc de mesurer la prévalence et l'incidence de certaines atteintes aux personnes et aux biens. Elle porte également sur l'opinion des personnes concernant leur cadre de vie et la sécurité.

La passation du questionnaire est réalisée pour sa plus grande partie en face à face, par un enquêteur équipé d'un ordinateur (Capi pour *Computer Assisted Personal Interviewing*). Pour les questions les plus sensibles (violences sexuelles et violences au sein du ménage), il est demandé au répondant de saisir lui-même ses réponses sur l'ordinateur au moyen d'un protocole sous casque (Casi pour *Computer Assisted Self Interviewing*). Après verrouillage du questionnaire par l'enquêté, les réponses à ces questions sensibles ne sont plus accessibles à l'enquêteur ce qui garantit une complète confidentialité sur cette partie du questionnaire.

1.3 Le plan d'expérimentations d'enquêtes multimode à l'Insee

L'utilisation d'Internet comme mode de recueil des données privilégié ou complémentaire est une solution envisagée à plus ou moins long terme par les Instituts nationaux de statistique pour répondre à la demande toujours plus exigeante en termes de qualité et de diversité des enquêtes auprès des ménages, dans un contexte général de restriction budgétaire (de Peretti et Razafindranovona, 2014). En particulier, le recours à la collecte multimode pourrait apparaître comme la solution théorique à mettre en oeuvre que l'on se place dans le cadre de l'erreur d'enquête totale (Groves et Lyberg, 2010) ou dans celui, plus général, de la qualité totale (Lyberg, 2012).

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*, etc. Aussi, avant de développer ou généraliser l'utilisation du multimode, l'Insee s'est lancé dans une vaste opération d'expérimentations afin d'étudier ces différentes questions méthodologiques. En particulier, la nécessité d'expérimenter spécifiquement pour chaque type d'enquête auprès des ménages s'est imposée car les résultats de la littérature ne sont pas toujours facilement généralisables (Razafindranovona, 2015).

1.3.1 Protocole standard des expérimentations

Le protocole « standard » utilisé pour les enquêtes expérimentales auto-administrées à l'Insee est le suivant :

- envoi d'une lettre-avis avec les données de connexion ;
- 1ère relance (3 semaines après) avec envoi d'une lettre contenant les données de connexion et un questionnaire papier (+ enveloppe T) ;
- 2ème relance (3 semaines après la 1ère relance) avec les données de connexion.

La possibilité de répondre sur papier est offerte afin de ne pas exclure d'emblée de l'enquête une partie de la population, d'autant plus quand la réponse est obligatoire, comme dans l'enquête principale en face à face. En revanche, cette possibilité n'est offerte que dans un deuxième temps afin de favoriser la réponse par internet sans que cela n'ait d'incidence négative sur le taux de réponse global.

1.3.2 Enquêtes expérimentales de l’Insee

Une première expérimentation de type multimode incluant une collecte par internet a été réalisée en 2010 et portait sur le thème du logement (Amiel et Denoyelle, 2012). Depuis 2013, plusieurs expérimentations de collecte multimode sur la base du protocole décrit plus haut ont été réalisées ou sont programmées : Qualité de vie au travail (2013), Vols, violences et sécurité (2013), Logement (2014), Patrimoine (2015), etc.

1.4 L’enquête expérimentale Vols, violence et sécurité

L’enquête Vols, violences et sécurité (VVS) s’est déroulée au premier trimestre 2013. Cette enquête méthodologique par internet et papier selon le protocole standard décrit précédemment, sur le thème de la victimation, a pour objectif principal de comparer les résultats obtenus avec ceux de l’enquête Cadre de vie et sécurité, réalisée chaque année, depuis 2007, par l’Insee, en face à face.

L’enquête expérimentale VVS reprend une partie des thématiques de CVS en s’appuyant pour l’essentiel sur les mêmes questions. Que ce soit dans le questionnaire web ou sa version papier, il n’y a pas de questionnement sur les violences au sein du ménage. On considère, en effet, que ces protocoles n’offrent pas des garanties suffisantes de confidentialité et de sécurité au répondant.

1.5 Auto-sélection et protocoles spécifiques de l’enquête VVS

L’enquête VVS a concerné un échantillon de 40 000 individus en France métropolitaine. Les personnes enquêtées ont été réparties en 3 sous-échantillons ; l’un cible les individus les plus jeunes, âgés de 14 à 19 ans (sous-échantillon 3), les deux autres sous-échantillons concernent les 20 ans et plus avec 2 façons différentes d’essayer de contrôler que c’est bien le destinataire du courrier qui répond afin de limiter les phénomènes d’auto-sélection au sein du ménage. La première méthode consiste à demander à la personne ciblée de répondre avec un premier identifiant-mot de passe et à donner la possibilité à une seconde personne du ménage de répondre en fournissant un autre couple identifiant-mot de passe (sous-échantillon 2). La deuxième méthode consiste à intégrer dans le questionnaire un tableau des habitants du logement demandant l’année de naissance, le sexe et le recensement de quatre types de victimations (sous-échantillon 1). La volonté est pour ces protocoles de maximiser les chances que cela soit bien l’individu ciblé qui réponde, puisqu’une autre victime potentielle du ménage pourrait, par ailleurs, déclarer un fait qu’elle a subi.

Enfin, une vérification *a posteriori* de certaines données (âge et sexe) recueillies dans le questionnaire par comparaison avec les informations disponibles dans la base de sondage permettent également de limiter l’auto-sélection : les observations pour lesquelles la concordance n’a pas été jugée suffisante ont été écartées.

Chapitre 2

Bilan de la collecte et premières exploitations brutes de VVS

Un premier bilan peut être réalisé à partir des données brutes, avant redressements ou corrections. L'intérêt d'un tel bilan est de fournir des éléments quant au déroulement de la collecte ainsi que quelques premiers résultats ciblés à partir des méta-données ou de variables spécifiques à la passation de questionnaire (ressenti sur la difficulté ou la longueur du questionnaire). Ce bilan est également l'occasion de faire un point sur la qualité brute des réponses fournies et de donner de premiers éléments sur le nombre de questionnaires remplis.

2.1 Déroulement de la collecte et taux de réponse

Environ 14 500 questionnaires ont été remplis avec une répartition plutôt équilibrée entre les deux modes de collecte que sont le papier et le web : 7 500 questionnaires ont été retournés par voie postale et 7 000 questionnaires remplis sur internet.

2.1.1 Collecte des réponses web

La collecte s'est déroulée lors du premier trimestre 2013, de mi-janvier à début avril. La distribution des réponses validées sur internet est fournie dans le graphique 2.1. Les pics correspondent à la réception des courriers : lettre avis puis première et deuxième relances éventuelles. Environ 55 % des réponses arrivent avant la réception de la première relance, et les 45 % de réponses qui arrivent ultérieurement se répartissent de manière équilibrée sur les deux dernières périodes délimitées par les relances.

L'une des questions, relative à la collecte par internet, que les instituts nationaux de statistiques commencent à se poser est celle de l'accès aux questionnaires via les plateformes mobiles. Lors de cette expérimentation, les personnes ont accédé au questionnaire très majoritairement par ordinateur, d'une part, car la majorité des personnes n'avaient que leur ordinateur comme moyen d'aller sur internet et d'autre part, car la non adaptation des questionnaires internet à un visionnage par tablette ou smartphone n'incite sans doute pas à répondre par ce biais.

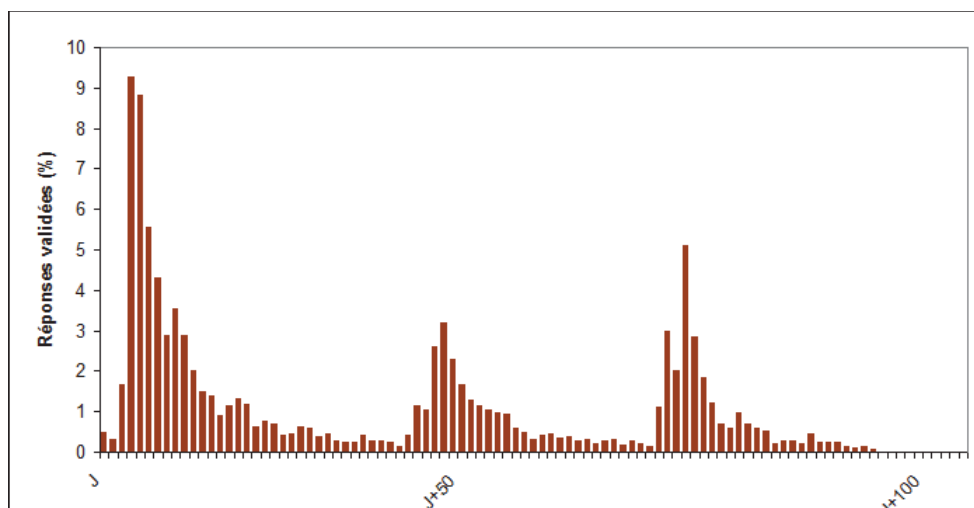


FIGURE 2.1 – Réponses web selon la date de validation

2.1.2 Le web : un mode plus "flexible" pour les actifs occupés ?

Un avantage de la collecte par internet auprès des ménages est la flexibilité offerte : le ménage peut choisir le moment qui lui convient le mieux pour répondre au questionnaire. Ainsi, pour certains instituts, offrir la possibilité de répondre par internet permettrait de faire participer certains ménages d'actifs occupés qui ne répondent pas aux sollicitations des enquêteurs, faute de temps.

En effet, les graphiques 2.2 et 2.3¹ indiquent que les répartitions des réponses selon l'heure et le jour de la semaine diffèrent quelque peu selon que l'on soit actif occupé ou non. Les actifs occupés utilisent plus souvent les plages horaires de soirée que les autres (étudiants, chômeurs, retraités ou inactifs) pour répondre à l'enquête ainsi que le week-end. Un tel constat ne préjuge en rien de ce qu'aurait été le comportement de ces ménages si l'enquête s'était déroulée en face-à-face : ces ménages auraient probablement dans leur très large majorité participé à l'enquête et il est ainsi difficile de déterminer l'apport effectif d'une telle flexibilité. Néanmoins, cela illustre, à titre indicatif, que la temporalité des réponses sur internet est liée à la temporalité des activités des individus enquêtés.

1. Précisons que le graphique 2.3 ne doit pas être interprété en termes de jour de la semaine le plus favorable à la réponse : rappelons que le graphique 2.1 indique que les réponses sont fortement concentrées autour des dates de réception des courriers et que donc, les fréquences de validation selon le jour de la semaine dépendent fortement du jour de la semaine où le courrier est reçu. En revanche, comparer des sous-populations sur cette caractéristique semble licite.

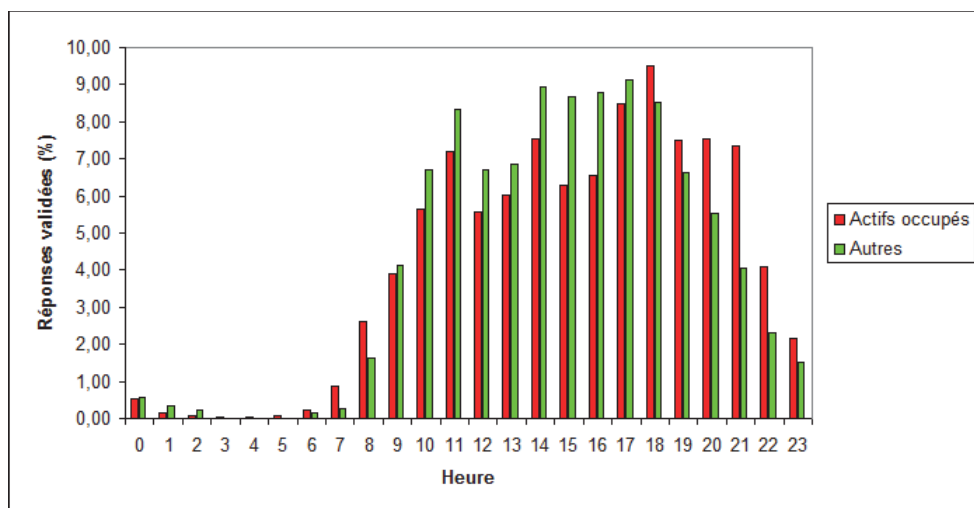


FIGURE 2.2 – Réponses web selon l'heure de validation

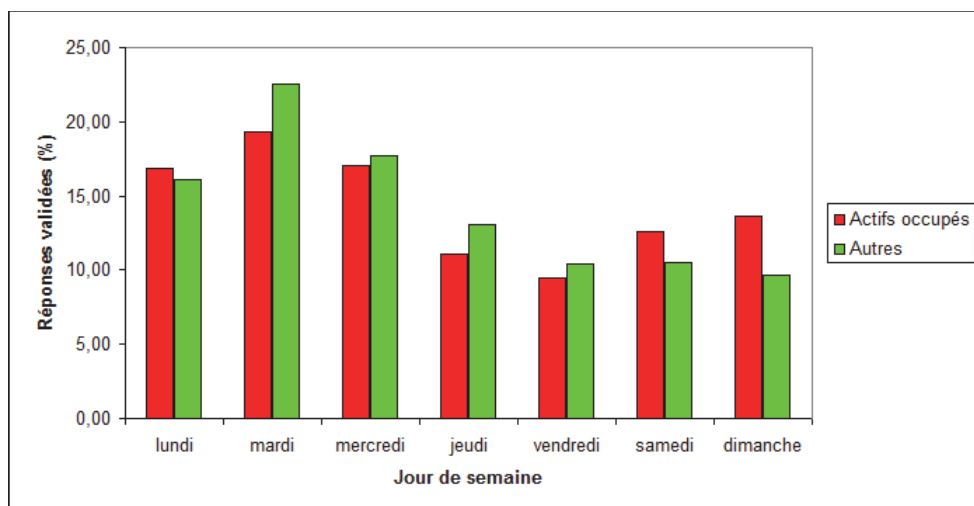


FIGURE 2.3 – Réponses web selon le jour de validation

2.1.3 Ressenti des répondants web sur le questionnement

Quelques questions posées en fin d'interrogation (uniquement dans la version web) portent sur le ressenti des répondants en ce qui concerne la longueur, la difficulté et la confidentialité du questionnaire. 87 % des répondants jugent le questionnaire comme étant de longueur raisonnable alors qu'ils sont 10 % à le trouver trop long et 3% qui, au contraire, le trouvent trop court. Pour environ les deux tiers des répondants, les questions sont faciles à comprendre, pour l'autre tiers, les questions sont jugées "normales"². La confidentialité semble assurée pour une très large majorité des répondants : 99 % affirment avoir pu répondre en toute confidentialité.

2. La modalité "difficiles à comprendre" est choisie par moins d'1 % des répondants.

Comme cela leur était proposé, une bonne partie des quelques rares personnes qui ont répondu négativement à cette question ont pu préciser leur situation dans la zone de commentaire libre : certaines sont méfiantes vis-à-vis des interrogations sur internet, d'autres précisent que le questionnaire a été rempli avec l'aide d'une tierce personne.

La zone de commentaire libre est par ailleurs utilisée par un nombre non négligeable de répondants : environ 20 % des répondants (soit 1 200 zones de commentaire) écrivent au moins un caractère dans cette zone. Pour certains d'entre eux, le commentaire précise qu'il n'y a rien à signaler : l'acronyme RAS ou sa version développée est utilisée dans une centaine de commentaires.

Sans faire une analyse textuelle approfondie de ces commentaires, nous pouvons relever les mots les plus fréquemment cités. Les mots autour du questionnement, de ses composantes et de l'enquête reviennent assez souvent : questions (326 occurrences), questionnaire (172), réponses (96), répondre (84), enquête (50), courrier (16) ou encore sondage (10). Les institutions que sont la police (89), la justice (48) et la gendarmerie (37) apparaissent assez souvent dans ces commentaires. Divers mots autour de la thématique de l'enquête sont utilisés : violence (122), insécurité (48) et sécurité (33), vol (93), agression (35) ou encore délinquance (22). Notons pour finir que l'adjectif indiscret est cité dans une cinquantaine de commentaires et qu'il y a également environ 50 occurrences sur les revenus et 30 sur le salaire.

2.2 Non-réponse partielle, qualité de réponse

La qualité de réponse est généralement considérée comme meilleure sur internet que sur papier en raison des divers contrôles en ligne qui peuvent être implémentés. En particulier, la non-réponse partielle est beaucoup plus faible car il est souvent plus aisé pour le répondant de sauter une question sur papier que sur un questionnaire en ligne.

2.2.1 Non-réponse partielle

Alors que sur papier, hormis le fait de proposer un questionnaire le plus clair et lisible possible, il n'existe guère de moyens de motiver l'enquêté à répondre à une question, sur internet plusieurs options sont possibles. Des contrôles bloquants peuvent être implémentés afin de forcer la réponse pour les questions les plus importantes : ils doivent néanmoins être utilisés avec parcimonie car le répondant peut choisir d'abandonner définitivement si ces sollicitations l'exaspèrent. Pour limiter la non-réponse partielle, une possibilité est de ne proposer la modalité de non réponse que dans un deuxième temps, lorsque l'enquêté tente de sauter la question. C'est ce choix qui a été fait pour le questionnaire en ligne de VVS.

Ainsi, la non-réponse partielle est quasiment inexistante sur les questionnaires en ligne alors qu'une proportion non négligeable de personnes sautent les questions sur papier. Ainsi, sur internet, la non-réponse est inférieure à 0,1 % pour les questions sur les victimations (voir tableau 2.1). Elle varie de 2,7 % à 4,3 % pour ces mêmes questions dans la version papier ; remarquons par ailleurs, que la non-réponse augmente au fur et à mesure que l'on progresse dans le questionnaire pour ces questions de victimation (les questions de victimation - vols, violences, menaces - sont présentées dans le tableau suivant leur ordre d'apparition dans le

questionnaire) indiquant peut-être un effet de lassitude. Sur la version web du questionnaire, l'une des seules questions pour laquelle la non-réponse n'est pas négligeable est celle du revenu ce qui n'est guère suprenant étant donné son caractère sensible. La non-réponse à cette question est néanmoins beaucoup moins élevée sur le web (2,5 %) que sur papier (10,6 %).

	NR VVS web (%)	NR VVS papier (%)
Victimations		
Vols dans le logement	0,07	2,69
Vols de véhicule	0,04	3,26
Vols avec violence	0,01	3,66
Vols sans violence	0,03	4,16
Violences physiques	0,01	4,25
Menaces	0,06	4,52
Insécurité		
Insécurité au domicile	0,03	3,88
Insécurité dans le quartier	0,01	3,94
Questions sociodémographiques		
Année de naissance	0,00	5,57
Sexe	0,00	4,56
Couple	0,00	5,58
Diplôme	0,00	7,13
Situation vis-à-vis du travail	0,14	6,12
Revenus (tranche)	2,45	10,63

TABLE 2.1 – Non-réponse partielle dans les questionnaires web et papier de VVS

2.2.2 Qualité des réponses

Dans l'ensemble, les questions posées dans l'enquête expérimentale VVS ne présentent *a priori* pas de difficultés insurmontables ; en particulier, celles qui sont posées à l'ensemble des enquêtés ne devraient pas être trop touchées par les phénomènes de type *satisficing*³ (Razafindranovona, 2015).

Comparaison sur un enchaînement de questions semblables dans leurs modalités

Deux questions proches dans leur formulation et semblables dans leurs modalités de réponse sont posées dans l'enquête : l'une porte sur l'efficacité de la police "À propos de l'action en général de la police ou de la gendarmerie nationale dans la société française actuelle, vous diriez personnellement qu'elle est", l'autre porte sur l'efficacité de la justice "À propos de l'action de la Justice et des tribunaux dans le traitement de la délinquance, vous diriez personnellement

3. Le *satisficing* renvoie à la faible implication du répondant. Ce phénomène peut s'appréhender à travers les modèles théoriques de réponse aux enquêtes. D'après la théorie du *satisficing* de Krosnick (1991), les répondants n'ont pas toujours la motivation et l'attention nécessaires au bon déroulement du processus cognitif qu'engage le répondant face à une question posée lors d'une enquête. Ainsi, les réponses fournies ne sont pas toujours optimales.

qu'elle est" et les modalités de réponse pour ces deux questions sont les mêmes, à savoir, dans l'ordre "Très satisfaisante", "Satisfaisante", "Peu satisfaisante", "Pas du tout satisfaisante" et "Pas d'opinion" (la non réponse est également possible).

Une forme de *satisficing* pourrait être de choisir la même modalité pour ces deux questions et que ce choix tienne plus de la solution de facilité que du report exact de son opinion.

Un premier constat porte sur la proportion de réponses "Pas d'opinion". On aurait pu penser que cette proportion aurait été importante sur internet car l'option de non réponse est rarement prise⁴ dans ce cas, mais cette proportion de "Pas d'opinion" est du même ordre voire inférieure à celle observée dans les réponses papier : 10,2 % contre 12,5 % pour la question sur la police, 11,8 % contre 12,3 % pour la question sur la justice. Sur papier, outre ces réponses "Pas d'opinion" il y a également 4 % de non réponses à chacune des deux questions quand la non réponse est quasi nulle sur internet. Il faut néanmoins prendre garde aux effets de sélection avant toute conclusion hâtive : il est possible que les modes de collecte sélectionnent des populations différentes vis-à-vis de ces questions et que ces différences dans la sélection expliquent en partie les comportements de réponse selon le mode⁵.

L'autre point est de savoir ici si les proportions de réponses identiques aux deux questions sont du même ordre selon le mode de collecte : les proportions sont en fait relativement proches avec 42 % de réponses identiques sur internet et 38 % sur papier⁶.

Ordre aléatoire de l'affichage des modalités

Pour deux des questions posées dans la version web du questionnaire, l'ordre d'affichage des modalités de réponse est aléatoire. Cet affichage aléatoire est implémenté pour des raisons de comparabilité avec les réponses en face à face : dans CVS, l'enquêteur lit les modalités des réponses possibles à ces deux questions également dans un ordre déterminé aléatoirement. Afficher aléatoirement l'ordre des modalités a pour objectif de minimiser les effets d'ordre (par exemple ceux de type *primacy*⁷ ou les effets de type *recency*⁸). Comme dans les données de l'enquête VVS sont disponibles à la fois les réponses et l'ordre proposé, il peut être intéressant d'observer si de tels effets d'ordre semblent exister ou non pour le mode de collecte internet⁹. La première question qui utilise l'ordre aléatoire des modalités est "De la présence de la police ou de la gendarmerie dans votre quartier (ou votre village), vous diriez ?" avec comme modalités de réponses "Qu'elle est suffisante", "Qu'elle est excessive", "Qu'elle est insuffisante", "Qu'elle est inexistante alors qu'elle serait nécessaire" et "Qu'elle est inexistante mais qu'elle n'est pas nécessaire". Si on numérote ces différentes modalités par respectivement 1, 2, 3, 4 et 5, l'affiche des modalités s'effectue selon 6 ordres possibles : 1-2-3-4-5, 1-3-4-5-2, 3-4-5-1-2, 3-4-5-2-1, 2-1-3-4-5 et 2-3-4-5-1.

4. Il est possible que certains internautes ne soient pas conscients que cette possibilité existe puisqu'elle n'est proposée que dans un deuxième temps.

5. Le contrôle de la sélection pourrait se faire en mobilisant des techniques comme le matching (Razafindranovona et al., 2013).

6. Notons que dans CVS, la proportion de réponses identiques à ces deux questions est de 39 %. En revanche, le recours à la modalité "Pas d'opinion" est bien moindre : 6 % des enquêtés choisissent cette modalité pour répondre à la question sur la police, 8 % à la question sur la justice.

7. Il s'agit de la tendance à choisir plutôt les premières modalités proposées que les dernières.

8. Il s'agit de la tendance à choisir plutôt les dernières modalités proposées que les premières.

9. Pour des raisons pratiques liées au support, l'ordre des modalités est toujours le même dans la version papier du questionnaire.

	Tous	1-2-3-4-5	1-3-4-5-2	3-4-5-1-2	3-4-5-2-1	2-1-3-4-5	2-3-4-5-1
Modalités (%)							
1.suffisante	42,9	43,9	42,0	44,1	40,5	44,2	42,5
2.excessive	2,0	2,6	1,0	2,0	2,2	1,4	2,7
3.insuffisante	26,6	29,4	27,4	24,3	26,8	26,6	25,2
4.inexistante alors qu'elle serait nécessaire	9,5	8,8	8,3	10,1	10,3	9,3	10,2
5.inexistante mais qu'elle n'est pas nécessaire	18,7	15,1	21,0	19,2	19,7	18,2	19,0

TABLE 2.2 – Réponses sur la présence de la police selon l'ordre des modalités (VVS)

Un type d'effet d'ordre que l'on peut observer est que les individus semblent moins souvent choisir la modalité lorsqu'elle est placée en dernière position que lorsqu'elle s'affiche dans une autre configuration. On peut ainsi constater cela pour toutes les modalités proposées en dernière position, à savoir les modalités 1, 2 et 5. Ceci est relativement cohérent avec la théorie : pour les questionnaires en ligne, qui s'appuient sur une dimension visuelle (et non orale), la tendance est de plutôt choisir les modalités en début de liste (Krosnick et Alwin, 1987).

Pour comparaison, le même tableau pour CVS 2013 est présenté (tableau 2.3) : une différence est que la non réponse ne figure pas dans le tableau pour plus de lisibilité. Si elle est quasi nulle dans VVS, elle représente environ 5 % des choix de modalités dans les différentes configurations.

Lorsque le questionnement s'appuie essentiellement sur la dimension orale, la tendance pour l'enquête peut être de se raccrocher aux dernières modalités (Krosnick et Alwin, 1987). Ceci peut arriver en particulier lorsque les listes de modalités sont longues et que le rythme d'interrogation est intense. Au vu des résultats à cette question de CVS, il ne semble pas que les effets de ce type soient particulièrement flagrants : les modalités placées en dernière position ne sont pas plus choisies que dans les autres configurations et celles placées en première position ne sont pas moins choisies.

	Tous	1-2-3-4-5	1-3-4-5-2	3-4-5-1-2	3-4-5-2-1	2-1-3-4-5	2-3-4-5-1
Modalités (%)							
1.suffisante	46,7	47,2	45,3	48,2	46,0	46,5	46,8
2.excessive	1,4	1,5	1,1	1,1	2,0	1,4	1,5
3.insuffisante	19,7	20,8	21,1	18,9	18,1	20,2	19,1
4.inexistante alors qu'elle serait nécessaire	10,2	10,1	10,2	9,5	10,6	10,2	10,8
5.inexistante mais qu'elle n'est pas nécessaire	17,0	15,7	17,6	17,2	18,1	16,8	16,5

TABLE 2.3 – Réponses sur la présence de la police selon l'ordre des modalités (CVS)

Une autre question applique un ordre aléatoire pour les modalités de réponse : "De l'action de la police ou de la gendarmerie en matière de lutte contre la délinquance dans votre quartier (ou votre village), vous diriez ?" avec comme modalités "Qu'elle est très efficace", "Qu'elle est efficace", "Qu'elle est peu efficace", "Qu'elle n'est pas efficace du tout". Les modalités s'affichent aléatoirement selon deux configurations : l'une dans le sens dans lequel les modalités sont listées,

l'autre configuration est dans le sens inverse.

	1-2-3-4	4-3-2-1
Modalités (%)		
1. Qu'elle est très efficace	2,7	2,1
2. Qu'elle est efficace	48,1	49,7
3. Qu'elle est peu efficace	36,7	36,4
4. Qu'elle n'est pas efficace du tout	9,1	7,5

TABLE 2.4 – Réponses sur l'efficacité de la police selon l'ordre des modalités (VVS)

Plus encore que sur l'autre question, il ne semble pas y avoir d'effets d'ordre particulièrement remarquables. Pour information, le tableau 2.5 donne les mêmes informations pour CVS 2013 (avec 28 % de non réponse sur cette question contre 4 % dans VVS). Les effets d'ordre ne semblent pas être particulièrement significatifs.

	1-2-3-4	4-3-2-1
Modalités (%)		
1. Qu'elle est très efficace	3,3	2,4
2. Qu'elle est efficace	44,8	45,6
3. Qu'elle est peu efficace	17,7	18,7
4. Qu'elle n'est pas efficace du tout	5,3	5,6

TABLE 2.5 – Réponses sur l'efficacité de la police selon l'ordre des modalités (CVS)

Chapitre 3

Enseignements après apurements et imputations

Après de premières exploitations brutes des données d'enquête, un travail d'apurement et d'imputations doit être réalisé avant de procéder à la comparaison de l'enquête expérimentale VVS avec l'enquête en face à face CVS. Une fois ce travail effectué, un bilan sur le comportement de réponse des enquêtés peut être réalisé en tenant compte du fait que tous n'ont pas forcément reçu le courrier avisant de l'enquête. À ce stade de l'analyse, de premiers résultats sur les variables clés des enquêtes que sont les victimations sont donnés. Quelques éléments visant à comparer les 2 types de protocoles mis en oeuvre pour limiter les phénomènes d'auto-sélection au sein du ménage sont également fournis.

3.1 Apurement et imputations

L'enquête a permis de récupérer environ 14 500 questionnaires. Parmi ces questionnaires, un bon nombre a dû être exclu. Ainsi ont été écartés les répondants autres que les personnes ciblées, les questionnaires trop lacunaires et enfin ceux dont les données ne correspondaient pas à celles de la base de sondage, soit dans le détail :

- 150 questionnaires, essentiellement papier, présentant trop de valeurs manquantes (toutes les questions de victimations, les questions sur les sentiments d'insécurité et des caractéristiques individuelles telles que le sexe, l'âge, le diplôme)
- 600 questionnaires obtenus avec les identifiants correspondant aux questionnaires volontaires, hors échantillon (second identifiant proposé dans la lettre-avis)
- 400 questionnaires correspondant au sous-échantillon de jeunes et pour lesquels l'âge du répondant est supérieur à 20 ans. Il est possible que ce soit dans certains cas des erreurs de saisie mais il se peut aussi qu'un parent ait répondu à la place de l'enfant. Par sécurité, ces questionnaires ont été écartés.
- 250 questionnaires des sous-échantillons 1 et 2 pour lesquels il y a une différence concernant le sexe et la date de naissance entre les données de l'échantillon et les données du questionnaire : il semble donc que ce ne soit pas la bonne personne qui ait répondu. Par ailleurs, on constate que globalement, ils déclarent plus souvent des victimations individuelles et qu'ils ressentent plus souvent de l'insécurité que les autres répondants.

- 150 questionnaires pour lesquels soit le sexe différait entre les données d'enquête et la base de sondage, soit la différence d'âge entre les deux sources était supérieure à 2 ans : après examen de leur profil qui correspondait à des taux de victimation sensiblement supérieurs à la moyenne, le choix a également été fait de les écarter.

Le travail d'imputations a quant à lui essentiellement concerné la variable de diplôme qui joue un rôle dans les différents calages ultérieurs. Une imputation déterministe a été réalisée sur la base de tris croisés entre les variables d'âge, de sexe et de groupe social.

3.2 Taux de réponse

Au final, environ 13 000 questionnaires sont considérés comme exploitables, soit un taux de réponse de 32 % avec une répartition entre Internet et papier plutôt équilibrée. Ce taux de réponse est très inférieur à ceux des enquêtes traditionnelles, en face à face, de l'Insee, mais se situe plutôt dans la norme des taux de réponse observés dans les enquêtes auto-administrées en population générale (Betts et Lound, 2010). Par ailleurs, une enquête auprès des non-répondants a été réalisée : elle interroge ces personnes sur les raisons pour lesquelles elles n'ont pas répondu à l'enquête. Les principales raisons évoquées sont le manque de temps, le fait de ne pas avoir accès à Internet et la non réception de la lettre-avis (voir annexe B).

3.2.1 Caractéristiques des répondants

L'exploitation des informations de la base de sondage indique des tendances déjà observées dans les enquêtes (expérimentales ou non) auto-administrées de l'Insee : le taux de réponse global est plutôt croissant avec le revenu et les répondants internautes sont plus jeunes et plus à l'aise financièrement que la population cible de l'échantillon et que la population de répondants papier (Gombault et Duée, 2012; Razafindranovona et al., 2013).

	Echantillon	VVS	VVS internet	VVS papier
Age (médian)	48	51	44	58
Revenu annuel (en k€) du foyer P25	15,8	20,3	23,9	17,8
Revenu annuel (en k€) du foyer (médian)	27,7	32,6	36,5	28,9
Revenu annuel (en k€) du foyer P75	42,3	47,8	53,1	42,7

TABLE 3.1 – Comparaison des répondants VVS avec la population échantillonnée (âge et revenu)

	VVS (%)	VVS internet (%)	VVS papier (%)
Diplôme de niveau baccalauréat ou inférieur	66,6	54,7	77,9
Diplôme de l'enseignement supérieur	33,4	45,3	22,1

TABLE 3.2 – Niveau de diplôme selon le mode de réponse

Une autre caractéristique des répondants internautes souvent soulignée dans la littérature sur les enquêtes en ligne est le niveau d'éducation plus élevé que celui de la population : 45

% des répondants internautes de VVS sont diplômés du supérieur contre 24 % en population générale d'après les données de l'enquête Emploi.

3.2.2 Déterminants de la non-réponse et de la non-réception des courriers

Les tableaux précédents donnent quelques premiers éléments sur les caractéristiques des personnes qui répondent à l'enquête VVS. Mais pour bien comprendre ce qui peut déterminer le fait de répondre ou non à une telle enquête, il faut bien avoir en tête que deux mécanismes sont en jeu : d'un côté, la réception ou non du courrier envoyé par l'institut pour inviter la personne sélectionnée à répondre, de l'autre la réponse ou participation en tant que telle.

En intégrant dans l'analyse les retours pour NPAI¹, on se rend compte que si pour certaines populations la proportion de questionnaires transmis est faible, cela peut être en partie parce que le courrier initial n'a pas été reçu. D'où l'importance d'avoir une base de sondage la plus fiable et la plus fraîche possible lorsque l'on souhaite enquêter les ménages de cette manière.

Par exemple, si l'on observe les réponses selon l'âge (tableau 3.3), on remarque que si la proportion de questionnaires retournés est plus faible chez les jeunes, cela tient probablement davantage des problèmes de non réception des courriers car cette population est mobile que du mécanisme de non réponse en tant que tel. Ainsi, la proportion de courriers revenus en NPAI est par exemple plus de 2 fois plus forte chez les 20-29 ans que chez les 40-49 ans (24 % contre 11 %), en revanche si l'on se restreint à ceux qui ont reçu les courriers, le taux de réponse est relativement similaire pour ces 2 tranches d'âge (38 % contre 40 %).

	NPAI (%)	Répondants (NPAI inclus) (%)	Répondants (hors NPAI) (%)
20-29 ans	23,7	29,0	38,0
30-39 ans	18,1	31,0	37,9
40-49 ans	11,0	35,9	40,3
50-59 ans	7,0	39,7	42,6
60-69 ans	5,6	42,8	45,3
70 ans et +	6,9	31,9	34,2

TABLE 3.3 – Réponses et NPAI selon l'âge

Et si l'on s'intéresse à l'estimation des déterminants de la non réponse en tant que telle, une solution pour limiter les biais liés à la sélection de première étape (réception ou non des courriers) est de mettre en oeuvre un modèle du type à la Heckman. Le tableau 3.4 donne les coefficients estimés pour certaines variables de la base de sondage dans l'équation de non-réponse selon deux méthodes : la première colonne fournit les coefficients estimés d'un modèle probit simple, la deuxième colonne les coefficients d'un modèle du type à la Heckman (heckprobit)².

Pour le modèle heckprobit, les équations sont les suivantes (Van de Ven et Van Praag, 1981) :

1. N'habite pas à l'adresse indiquée.

2. Pour le modèle heckprobit, la variable de sexe intervient dans l'équation de non réponse mais pas dans l'équation de sélection.

– équation latente

$$y^* = x_j\beta + u_1$$

– équation de la variable observée avec $y = 1$ si l'individu ne répond pas à l'enquête

$$y = 1_{y^* > 0}$$

– équation de sélection (on n'observe "vraiment" l'outcome que pour certaines des observations) avec $\tilde{y} = 1$ si l'individu a reçu le courrier (non NPAI)

$$\tilde{y} = z\gamma + u_2$$

Avec comme hypothèses $u_1 \sim N(0, 1)$, $u_2 \sim N(0, 1)$ et $\text{corr}(u_1, u_2) = \rho$.

Le tableau 3.4 montre que les coefficients estimés des déterminants de la non réponse ne sont pas tellement modifiés lorsque l'on passe d'une modélisation probit à un modèle à la Heckman. Néanmoins, on retrouve dans une certaine mesure un résultat observé dans l'approche descriptive (tableau 3.3), à savoir que l'effet d'être "jeune" joue un peu moins sur la non réponse lorsque l'on tient compte d'une première phase de sélection liée à la réception de la lettre-avis.

	Probit "simple"	Heckprobit
Constante	-0,15 (0,000)	-0,15 (0,000)
Sexe		
Homme		
Femme	-0,06 (0,000)	-0,06 (0,000)
Age		
20-29 ans	0,18 (0,000)	0,16 (0,000)
30-39 ans	0,20 (0,000)	0,19 (0,000)
40-49 ans		
50-59 ans	0,13 (0,000)	0,13 (0,000)
60-69 ans	0,09 (0,000)	0,09 (0,000)
70 ans et +	0,16 (0,000)	0,16 (0,000)
Statut matrimonial		
Célibataire	-0,07 (0,002)	-0,07 (0,001)
Marié		
Divorcé	-0,05 (0,060)	-0,05 (0,050)
Veuf	0,22 (0,000)	0,21 (0,000)
Pacsé	-0,13 (0,006)	-0,16 (0,005)
Non disponible	0,19 (0,011)	0,10 (0,395)
Enfants de moins de 13 ans dans le ménage		
Oui	0,08 (0,000)	0,09 (0,000)
Non		
Statut d'occupation		
Propriétaire ou autres		
Locataire	0,30 (0,000)	0,28 (0,000)
Zone urbaine sensible		
Commune avec Zus	0,05 (0,002)	0,04 (0,005)
Autres		
Décile de revenu		
Décile 1	0,63 (0,000)	0,64 (0,000)
Décile 2	0,54 (0,000)	0,55 (0,000)
Décile 3	0,47 (0,000)	0,47 (0,000)
Décile 4	0,39 (0,000)	0,40 (0,000)
Décile 5	0,30 (0,000)	0,30 (0,000)
Décile 6	0,20 (0,000)	0,19 (0,000)
Décile 7	0,18 (0,000)	0,18 (0,000)
Décile 8	0,11 (0,000)	0,10 (0,000)
Déciles 9 et 10		

TABLE 3.4 – Déterminants de la non réponse

3.3 Premières estimations sur variables clés

3.3.1 Taux de victimation

Rappelons que les principales variables d'intérêt de ces enquêtes sont les victimations subies pas les répondants. Nous nous intéressons ici tout particulièrement aux faits délictueux où la victime est l'individu : les vols avec ou sans violence (hors cambriolages, vols de véhicules), les violences physiques et les menaces. Pour des raisons de comparabilité entre les différentes versions des questionnaires l'analyse se concentrera essentiellement sur ces victimations, même si d'autres sont abordées dans CVS ou dans la version internet de VVS.

Les taux de victimation de VVS (bruts, avec pondérations de tirage) sont présentés dans le tableau 3.5.

	VVS internet-papier
Victimations (%)	
Vols avec violence	2,3
Vols sans violence	6,2
Violences physiques	2,4
Menaces	7,0

TABLE 3.5 – Taux de victimations bruts dans l'enquête VVS

Comme nous le verrons plus tard lors de la comparaison avec les taux de l'enquête CVS, les taux de VVS³ apparaissent comme plutôt largement plus élevés que ceux de l'enquête en face à face, en particulier pour ce qui concerne les vols.

3.3.2 Retour sur les questionnaires écartés

Questionnaires volontaires

Environ 600 questionnaires ont été écartés car ils ont été remplis par des répondants "hors échantillon" à savoir que ce ne sont pas les personnes sélectionnées dans l'échantillon mais des personnes appartenant *a priori* au même ménage qu'une personne sélectionnée : plus encore que pour les autres répondants on peut suspecter une très forte corrélation entre la participation de ces personnes et les variables d'intérêt.

Les taux de victimation de cette population de répondants volontaires s'avèrent en effet très largement supérieurs à ceux des individus sélectionnés avec pour les 4 victimations étudiées, des taux pratiquement deux fois supérieurs.

3. Les réponses web et papier ne sont pas différenciées : elles forment un tout puisque l'objectif essentiel est la comparaison de réponses auto-administrées avec les résultats de CVS. Pour quelques éléments de comparaison entre les taux de victimation "papier" et "web", se référer à l'annexe A

	VVS internet-papier (retenus)	Questionnaires "hors échantillon"
Victimations (%)		
Vols avec violence	2,3	4,3
Vols sans violence	6,2	11,8
Violences physiques	2,4	3,9
Menaces	7,0	13,2

TABLE 3.6 – Taux de victimations des questionnaires volontaires

Questionnaires écartés faute de concordance avec la base de sondage

Lorsque les réponses à l'enquête diffèrent des variables de la base de sondage (âge et/ou sexe), si l'erreur de saisie ne peut pas être complètement écartée, on peut également suspecter que le protocole ne soit pas respecté : une autre personne que celle sélectionnée répond à l'enquête. Le tableau 3.7 montre que les taux de victimation des vols avec ou sans violence sont légèrement supérieurs pour ces quelques 800 questionnaires écartés.

	VVS internet-papier (retenus)	Questionnaires écartés
Victimations (%)		
Vols avec violence	2,3	3,6
Vols sans violence	6,2	7,4
Violences physiques	2,4	2,5
Menaces	7,0	6,9

TABLE 3.7 – Taux de victimations des questionnaires écartés

3.4 Comparaison entre les deux types de protocoles

Pour rappel, deux protocoles différents ont été testés dans cette expérimentation pour limiter les phénomènes d'auto-sélection. L'un (protocole 1) propose dans le questionnaire un tableau où le répondant peut renseigner certaines victimations subies par les autres personnes de son ménage, l'autre (protocole 2) donne la possibilité à une autre personne du ménage de remplir un questionnaire.

3.4.1 Nombre de réponses

Le nombre de questionnaires retournés est similaire pour les deux protocoles avec dans les deux cas un peu plus de 6 700 questionnaires. La différence est que ce chiffre comprend pour le protocole 2 les questionnaires des répondants volontaires qui seront écartés lors de l'exploitation. Il faut donc retrancher ces quelques 600 questionnaires volontaires pour obtenir la proportion de ménages ayant spontanément répondu : cette proportion est de 36,5 % pour le protocole 1 et de 33,3 % pour le protocole 2. Le taux de réponse spontané est donc légèrement en faveur du protocole 1. Et cet avantage se confirme si l'on tient compte de l'apurement : un peu plus de questionnaires sont finalement écartés pour le protocole 2 que pour le protocole 1

(350 questionnaires vs 300 questionnaires). Le taux de réponse est donc au final de 34,9 % pour le protocole 1 et de 31,4 % pour le protocole 2.

Avec un taux de réponse spontané moins élevé et une proportion de questionnaires écartés un peu plus importante, il semblerait que le protocole 2 se soit révélé légèrement plus complexe pour le répondant que le protocole 1 avec entre autres des possibilités de confusion liés à l'utilisation des différents identifiants.

3.4.2 Victimations selon le protocole

La comparaison des taux de victimation selon le protocole ne révèle pas de grosses différences ⁴ (tableau 3.8). Dans tout ce qui va suivre, la différenciation entre les différents protocoles ne sera plus faite, les réponses seront globalement agrégées.

	Protocole 1	Protocole 2
Victimations (%)		
Vols avec violence	2,2	2,0
Vols sans violence	6,1	5,8
Violences physiques	2,0	2,3
Menaces	6,9	6,6

TABLE 3.8 – Victimations et protocoles

4. Pour les 4 victimations étudiées, un test d'égalité des proportions a été effectué : on ne rejette pas l'hypothèse nulle d'égalité des proportions au seuil de 10 %.

Chapitre 4

Calages et comparaison des taux de victimation

Après avoir observé les taux bruts de victimation dans l'enquête expérimentale multimode VVS, l'objectif est maintenant de comparer, après corrections (non-réponse totale, calages) ces taux avec ceux de l'enquête en face à face CVS. Dans quelle mesure parvient-on à faire converger les taux de victimation de ces deux enquêtes ? Pour cela, nous utilisons des méthodes dans l'esprit des *back-door models* (Vannieuwenhuyze et al., 2014) pour essayer de rendre comparables les répondants selon les différents modes de collecte, en contrôlant par des caractéristiques observables. Nous essaierons également de donner quelques éléments pour expliquer les différences qui persistent dans les taux de victimation, y compris après contrôles : l'hypothèse privilégiée est que ces différences relèvent ici essentiellement d'effets de sélection non contrôlée, effets qui rendent par ailleurs la conclusion difficile quant à l'existence ou non d'effets de mode intrinsèquement liés à la mesure (Razafindranovona, 2015).

4.1 Correction de la non-réponse totale

La correction de la non réponse totale s'inspire de ce qui se fait pour l'enquête Technologies de l'information et de la communication : 170 groupes homogènes de réponse (Little, 1986) sont définis par 5 tranches d'âge et 7 tranches de revenu avec, lorsque les effectifs le permettent, un croisement avec la présence de zone urbaine sensible (ZUS) dans la commune ou dans l'unité urbaine (la non réponse est plus importante dans ces zones).

Une correction alternative intégrant comme variable supplémentaire le niveau de délinquance enregistrée par les services de police et gendarmerie (nombre de faits rapporté à la population, par déciles) a également été réalisée mais ne modifie pas substantiellement les résultats : c'est donc la correction décrite plus haut qui a été retenue au final.

Cette correction de la non-réponse totale augmente les taux de victimation : c'est assez logique puisque la correction aboutit à augmenter le poids des personnes vivant dans une commune avec ZUS et le poids des jeunes.

	VVS avant CNR	VVS après CNR
Victimations (%)		
Vols avec violence	2,3	2,6
Vols sans violence	6,2	6,4
Violences physiques	2,4	2,7
Menaces	7,0	7,3

TABLE 4.1 – Taux de victimation avant et après correction de la non réponse totale

4.2 Principes des calages sur marges successifs

4.2.1 Le calage pour contrôler la sélection

Les premières constatations de comportements de réponse assez différenciés selon certaines caractéristiques sociodémographiques indiquent une nécessité de corriger a minima la déformation de l'échantillon. Mais ces corrections n'assurent pas la "représentativité" de notre échantillon final dès lors que les variables utilisées pour corriger ne sont pas suffisamment liées aux variables d'intérêt (ici les variables de victimation), ou que des éléments inobservables et corrélés aux variables d'intérêt participent grandement au processus de sélection des individus. Les calages successifs ont ici pour objet de tester la convergence possible entre les deux enquêtes en termes de taux de victimation globaux : les différences qui resteront à la fin du processus de calage relèveraient alors soit de l'effet de mode (lié à la mesure) soit de la sélection non contrôlée. L'objet n'est pas tant ici d'utiliser le calage en tant que méthode *stricto sensu* d'évaluation d'un effet causal sur un paramètre d'intérêt comme dans l'*entropy balancing* (Hainmueller, 2011) mais plutôt d'observer de manière empirique comment et dans quelle mesure on parvient ou non à réconcilier les taux de victimation de ces deux enquêtes, et éventuellement à quel prix.

4.2.2 L'enquête CVS comme bon étalon de la victimation

Lorsqu'une comparaison est effectuée entre les résultats de deux enquêtes différentes, se pose la question de la référence. Est-ce que l'une des deux enquêtes est *a priori* mieux qualifiée que l'autre pour constituer l'étalon ? Dans ce cas, il semble tout à fait raisonnable de plutôt prendre CVS comme étalon. Plusieurs raisons peuvent ainsi être évoquées : les taux de victimation sont globalement assez stables au cours du temps, les résultats sont plutôt en adéquation avec les données administratives recueillies par ailleurs pour des victimations dont le taux de plainte est élevé (vols de voiture par exemple) et le taux de réponse plutôt élevé dispense d'émettre des hypothèses trop fortes lors des phases de correction et de repondération.

4.2.3 Les calages sur marges successifs

Plusieurs calages sur marges (Sautory, 1993) successifs ont été effectués par la suite afin de voir comment les taux de victimation des deux enquêtes peuvent ou non être réconciliés.

Pour rappel du principe général, étant donné une certaine variable d'intérêt Y et connaissant

des totaux X_j sur l'ensemble de la population, on cherche à obtenir un estimateur de la forme

$$\hat{Y}_w = \sum_{k \in s} w_k y_k$$

où les poids w_k sont "proches" des pondérations initiales d_k (ici issues de l'étape de correction de la non réponse) et vérifient des équations de calage de type

$$\forall j = 1 \dots J \quad \sum_{k \in s} w_k x_{jk} = X_j \quad (4.1)$$

On va alors chercher un système de poids w_k solution de

$$\min_{w_k} \sum_{k \in s} d_k G\left(\frac{w_k}{d_k}\right)$$

où G est une fonction de distance, sous les contraintes des équations de calage (4.1).

Un premier calage fait intervenir des variables sociodémographiques classiques (via les marges utilisées pour CVS 2013) dont la sélection a été guidée par une étude préliminaire de corrélations (sur la base de régressions logistiques) entre caractéristiques sociodémographiques et victimations.

Un deuxième calage intègre le taux de plainte comme marge supplémentaire avec pour hypothèse sous-jacente que CVS constitue toujours la référence y compris pour les taux de plainte. Les personnes sollicitées pour VVS seraient d'autant plus enclines à répondre à l'enquête qu'elles auraient été victimes, et qu'elles n'auraient pas porté plainte, cet effet étant beaucoup moins marqué dans une enquête en face à face que dans une enquête auto-administrée du fait du rôle joué par l'enquêteur ; et on vérifie en effet que les taux de plainte apparaissent sous-estimés dans VVS. Une piste voisine consisterait à effectuer le calage, non plus sur le taux de plainte, mais directement sur les nombres de plaintes. La première idée serait d'utiliser pour ce faire une source extérieure, administrative, ce qui s'avère impossible compte tenu des unités de compte différentes entre les sources¹. C'est pourquoi on se basera plutôt sur le nombre de plaintes issu de CVS pour les différents types de victimations. Des analyses préliminaires ayant montré que le calage sur le taux de plainte n'avait que peu d'impact sur les taux de victimation, c'est finalement le calage sur nombre de plaintes qui a été retenu dans le processus de calages successifs. Un troisième calage fait intervenir comme marge supplémentaire le sentiment d'insécurité. Il s'avère en effet que les réponses à cette question divergent sensiblement entre les deux enquêtes, avec de l'ordre de 30 % de personnes éprouvant un sentiment d'insécurité dans leur quartier dans VVS, contre 12 % dans CVS 2013. Ces personnes, qu'elles aient été ou non victimes, auraient une plus forte propension à répondre aux enquêtes de victimation et cet effet de sélection jouerait à plein dans VVS et pas dans CVS. Comme, d'autre part, il y a une forte corrélation entre le sentiment d'insécurité dans le quartier et la victimation, un calage intégrant ce sentiment d'insécurité permettrait peut-être de faire converger les taux de victimation entre les deux enquêtes.

Enfin, l'échantillon a été redressé en intégrant comme cale supplémentaire le taux d'accès à

1. Plus précisément, il existe des différences à la fois de concept et de champ qui limitent les possibilités d'utilisation directe d'une source administrative.

Internet en utilisant une donnée issue de l'enquête Technologies de l'information et de la communication auprès des ménages.

Si l'on reprend les notations utilisées plus haut, nous effectuons 4 calages sur marges, donc nous obtenons 4 jeux de poids :

(w_{k1}) tels que

$$\min_{w_{k1}} \sum_{k \in s} d_k G\left(\frac{w_{k1}}{d_k}\right) \quad sc \quad \forall j = 1 \dots J \quad \sum_{k \in s} w_{k1} x_{jk} = X_j$$

(w_{k2}) tels que

$$\min_{w_{k2}} \sum_{k \in s} d_k G\left(\frac{w_{k2}}{d_k}\right) \quad sc \quad \forall j = 1 \dots J \quad \sum_{k \in s} w_{k2} u_{jk} = U_j$$

(w_{k3}) tels que

$$\min_{w_{k3}} \sum_{k \in s} d_k G\left(\frac{w_{k3}}{d_k}\right) \quad sc \quad \forall j = 1 \dots J \quad \sum_{k \in s} w_{k3} v_{jk} = V_j$$

(w_{k4}) tels que

$$\min_{w_{k4}} \sum_{k \in s} d_k G\left(\frac{w_{k4}}{d_k}\right) \quad sc \quad \forall j = 1 \dots J \quad \sum_{k \in s} w_{k4} z_{jk} = Z_j$$

avec

- X =[marges sociodémographiques],
- U =[X , nombre de plaintes],
- V =[U , sentiment d'insécurité],
- Z =[V , accès à internet].

La méthode exponentielle est utilisée dans les programmes d'optimisation donnant (w_{k1}) et (w_{k2}) alors que pour essayer de borner *a minima* les rapports de poids, c'est la méthode logit qui est utilisée dans les programmes d'optimisation donnant (w_{k3}) et (w_{k4}) .

4.3 Quelle convergence des taux de victimation ?

Après ces calages successifs, on ne peut conclure à la réconciliation des taux de victimation de VVS et CVS 2013. C'est pour les violences physiques que la convergence est la plus satisfaisante puisque l'estimation VVS est dans l'intervalle de confiance à 95 % estimé pour la référence CVS 2013. La convergence est moindre pour les menaces où l'on frôle néanmoins la borne supérieure de l'intervalle de confiance. Mais surtout, les différences restent très importantes pour les vols avec ou sans violence.

Non seulement ces calages ne permettent pas de faire totalement converger les taux de victimation, mais en plus, la distorsion des poids induite obère sans doute leur applicabilité. En effet, le calage sur les marges sociodémographiques, le taux de plainte et le sentiment d'insécurité ne converge, par la méthode logit, que pour des bornes de rapports des poids fixées

à 0,12 et 3,73 . La distorsion est encore plus considérable lorsque l'on rajoute dans le calage l'accès à Internet avec une convergence obtenue pour des bornes à 0,005 et 6 . De plus, le calage sur le sentiment d'insécurité qui est explicitement l'une des variables d'intérêt des enquêtes de victimation peut poser question. Ces différents constats semblent ainsi plutôt condamner de manière quasi définitive l'utilisation directe du seul protocole auto-administré type VVS pour estimer les prévalences de victimation.

	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

4.4 Le problème de la sélection non contrôlée

4.4.1 La sélection non contrôlée comme principale explication des écarts entre CVS et VVS

Les différences constatées entre les taux de victimation de VVS et de CVS 2013 peuvent relever essentiellement de deux phénomènes. Ces différences peuvent être pour partie des effets de mode "purs", liés à la mesure, c'est-à-dire au fait que sont utilisées différentes manières de collecter l'information sur le paramètre d'intérêt. Plus que des phénomènes liés à la désirabilité sociale ou au satisficing (Roberts, 2007), souvent avancés pour expliquer les différences de réponses entre un questionnaire avec enquêteur et un questionnaire auto-administré, dans ce cas, il s'agit ici plutôt de problèmes de compréhension du questionnaire et de confusion entre certains concepts : vols avec violence et violences physiques d'un côté, vols avec violence et vols sans violence de l'autre. Des premières analyses rendent compte d'une plus forte tendance à la multi-victimation de ces faits sans que l'on puisse conclure de manière définitive à la mauvaise compréhension du questionnaire. Par ailleurs, si ces problèmes de mesure existent, ils ne suffisent pas à expliquer entièrement les écarts constatés entre VVS et CVS 2013. L'explication la plus vraisemblable relève plutôt de la sélection non contrôlée : il existe des facteurs inobservables qui participent au processus de sélection des individus et ces facteurs sont fortement liés aux variables d'intérêt (Razafindranovona, 2015). En effet, l'intérêt et la motivation pour le thème de l'enquête sont des déterminants essentiels de la participation des enquêtés (Biemer et Lyberg, 2003). Or, ces caractéristiques ne peuvent pas parfaitement être captées par les variables collectées ; en particulier les méthodes classiques de repondération ne parviennent pas toujours à corriger les biais liés à ces phénomènes, d'autant plus importants que ces caractéristiques inobservables sont vraisemblablement fortement corrélées aux variables d'intérêt (problème de

non réponse non ignorable). En particulier, les différents calages ont mis en évidence le rôle du sentiment d'insécurité dans l'effet de sélection.

4.4.2 Une illustration de la sélection non contrôlée

Le protocole auto-administré ne parvient pas à faire participer les moins motivés par le sujet. Pour donner un exemple illustratif des biais liés aux problèmes de motivation, l'idée est d'utiliser des parodonnées et d'observer sur les données de l'enquête VVS, restreintes aux réponses des internautes, les victimations selon la date de validation du questionnaire, une approche assez semblable aux travaux de Voogt (Voogt, 2004) sur l'intérêt pour la question politique. Plus précisément, deux catégories temporelles ont été constituées selon que la réponse de l'individu arrive avant ou après la première relance postale, les réponses se répartissant environ pour moitié dans chacune des catégories. L'hypothèse sous-jacente est que la proportion de personnes "motivées" est moindre pour les répondants après relance que pour ceux qui répondent au début de la collecte ... une proportion qui serait sans doute encore moindre pour la population non-répondante, que le protocole en face à face parvient, lui, dans une certaine mesure à convaincre de participer ! Et si l'on considère également que la motivation est positivement liée aux victimations, alors la tendance attendue est d'observer des taux de victimation supérieurs au début de la collecte.

	Avant relance	Après relance
Victimations (%)		
Vols avec violence	2,9	2,4
Vols sans violence	8,2	6,4
Violences physiques	2,8	2,7
Menaces	10,3	7,2

TABLE 4.3 – Taux de victimations avant et après relance

L'hypothèse émise semble confirmée par ces résultats : les taux de victimation sont supérieurs lorsque la réponse arrive en début de la collecte avec des différences plus ou moins marquées selon les faits considérés. Pour contrôler des caractéristiques sociodémographiques et avoir une idée de l'ampleur de l'effet motivation en comparaison des effets de ces caractéristiques, des régressions logistiques ont été effectuées et confirment la tendance observée. L'effet de la date de validation du questionnaire est significatif pour les 4 types de victimations et massif pour au moins trois d'entre elles : la relation entre rapidité à répondre et victimation semble plutôt robuste.

Ces constats semblent bien illustrer de quelle manière la motivation peut participer au processus de sélection, en particulier quand les taux de réponse sont faibles. En revanche, une utilisation directe de ces résultats pour corriger les estimations de taux de victimation semble périlleuse : une idée pourrait être d'implémenter des estimateurs similaires à ceux proposés par Hansen et Hurwitz (1946) où l'on ferait l'hypothèse que les non répondants ont des comportements sur les variables d'intérêt proches de ceux qui répondent après la relance. Vu que les taux de victimation sont plus faibles après relance, cela tendrait à rapprocher les estimations

de VVS de celles de CVS. Mais au vu des taux de victimation constatés après relance, encore largement supérieurs à ceux de CVS, une telle hypothèse semble vraiment trop forte et cela tient en grande partie à la nature de la nouvelle sollicitation de l'enquêté (courrier par voie postale).

Chapitre 5

Des effets de mode ?

Si l'hypothèse avancée d'effets de sélection non contrôlée, sur caractéristiques inobservables, semble plausible peut-on néanmoins dire quelque chose en termes d'effets de mode liés à la mesure ? C'est toute la question qui peut se poser une fois que l'on considère que les différences dans les taux de victimation relèvent essentiellement de la sélection non contrôlée. Car il est important d'avoir une idée sur l'ampleur des effets de mode si l'on envisage d'agrèger des réponses provenant de différents modes de collecte. Est-ce que certains constats peuvent quand même être potentiellement robustes à ces problèmes de sélection ?

5.1 De la difficulté à distinguer effets de sélection et effets de mode

Les résultats du chapitre précédent montrent que les différences dans les taux de victimation de CVS et VVS restent importantes y compris après contrôle par des caractéristiques observables. Lorsque ces différences sont constatées, une question est de savoir si elles relèvent plutôt de la sélection non contrôlée ou de l'effet de mode (Razafindranovona, 2015). La réponse à cette question n'est pas toujours évidente mais dans ce cas, la piste privilégiée est plutôt celle de la sélection non contrôlée. En effet les victimations qui ont été considérées ne sont pas des victimations pour lesquelles il y aurait un fort soupçon d'erreur de mesure dans l'un ou l'autre des modes de collecte même si cette hypothèse ne peut être complètement écartée.

Les conclusions quant aux effets de mode et aux problèmes de sélection non contrôlée dépendent en partie des *a priori* théoriques, néanmoins il semble plausible dans certains cas de fournir certaines pistes d'interprétation. Lorsque les variables d'intérêt ne semblent *a priori* pas trop susceptibles d'être sujettes aux erreurs de mesure liées aux modes de collecte, les différences constatées devraient relever essentiellement de la sélection non contrôlée. À l'inverse, lorsque des différences constatées sont massives et que le processus de sélection semble n'intervenir qu'à la marge pour les variables considérées, on peut sans doute avancer l'existence d'effets de mode.

5.2 Quelques constats sur la possible mauvaise compréhension de certains concepts

Certains constats effectués dans la comparaison des réponses fournies dans CVS et VVS montrent qu'il existe bien des différences intrinsèquement liées aux modes de collecte.

5.2.1 Des poly-victimes plus fréquents dans VVS

Parmi les victimes, 71 % n'ont été victimes que d'un seul type d'atteinte selon l'enquête CVS 2013, contre 66 % des victimes de VVS. Deux hypothèses sont envisageables :

- VVS surestime les taux de victimation et de multi-victimation car les victimes et en particulier les poly-victimes sont davantage incitées à répondre à une enquête les concernant
- Le questionnaire auto-administré génère davantage de doublons. La succession des modules dans le questionnaire de CVS fonctionne par exclusion. Dans un premier temps, les vols personnels avec violence doivent être comptabilisés en excluant tout ce qui a pu se produire au cours d'un vol dans le logement ou d'un vol de voiture. Ensuite les vols sans violences sont comptabilisés en excluant tout ce qui a pu se produire au cours d'un vol avec violence. Puis, les violences physiques doivent être recensées en excluant tout ce qui a pu se produire au cours d'un vol. Enfin, les menaces sont recensées lorsqu'elles se sont produites sans vol ni violences physiques. Afin de ne pas surestimer la victimation, ces concepts et exclusions de champ sont bien définis et expliqués lors de la formation des enquêteurs. Le questionnaire de VVS reprend ces exclusions, mais elles peuvent ne pas être bien comprises quand les enquêtés sont seuls pour répondre. Par ailleurs, dans l'enquête CVS, le recensement des faits subis se fait à la suite et la description des faits renseignés intervient ensuite. Dans l'enquête VVS, le fait est recensé puis est décrit immédiatement avant de passer au recensement d'un autre type de fait.

5.2.2 Analyse des double comptes

Pour répondre à l'hypothèse de doublons générés par le questionnaire auto-administré, une analyse des poly-victimations a été réalisée en étudiant les caractéristiques des faits.

Cumul de violences physiques et de vols avec violences

Dans VVS, 37 répondants (soit 46 150 en effectifs pondérés) ont déclaré avoir subi à la fois un ou des vol(s) avec violence et des violences physiques, contre 16 dans CVS 2013 (70 215 en effectifs pondérés). Pour 7 % de ces répondants, les faits déclarés ont lieu la même année, le même mois et au même endroit, quand ce n'est jamais le cas dans CVS.

Cumul de vols avec violences et de vols sans violences

Dans VVS, 114 répondants (soit 146 860 en effectifs pondérés) ont été victimes de vol(s) avec violences et de vol(s) sans violence, contre seulement 9 dans CVS 2013 (36 340 en effectifs pondérés).

Pour près de la moitié de ces victimes, les différents vols ont eu lieu la même année et le même

mois dans VVS. Pour 44 %, ils ont lieu la même année, le même mois et au même endroit. Pour environ 2 répondants sur 10, les vols se sont produits la même année, le même mois, au même endroit et les mêmes objets ont été volés. Dans ce dernier cas, lorsque le montant des objets est renseigné, le montant est le même dans les deux faits renseignés. Dans l'enquête CVS 2013, de même que pour le cumul agressions physiques / vol(s) avec violences, on ne rencontre jamais ce cas de cumul.

Il y a donc clairement une forte suspicion de doubles comptes dans les cas de cumuls où les mêmes objets ont été volés. Comme on le signalait plus haut, la description des faits subis se fait dans VVS après chaque recensement de fait. Ainsi, concernant les vols avec ou sans violences, il est demandé aux enquêtés de décrire le dernier fait subi : « Concernant le dernier vol (ou la dernière tentative de vol) que vous avez subi(e) ». Mais, en ce qui concerne les vols sans violences, la description du dernier fait ne rappelle pas que l'on souhaite spécifiquement parler du dernier vol sans violence et les enquêtés pourraient être amenés à décrire le vol avec violence subi s'il s'agit du vol le plus récent plutôt que du vol sans violence. Nous n'avons cependant pas de moyen de savoir s'il s'agit d'un réel double compte ou s'il s'agit d'une mauvaise compréhension de la description du dernier vol.

Ces exemples montrent qu'il existe des effets liés au mode de passation du questionnaire : plus qu'un effet de mode purement lié à l'instrument, ces effets sont liés aux adaptations qui ont été effectuées pour transposer un questionnaire en face à face en un questionnaire auto-administré de longueur raisonnable. Ces constats permettent d'envisager des pistes d'amélioration pour une nouvelle version du questionnaire auto-administré si une opération de ce type est reconduite dans le futur.

5.3 Elements sur l'utilisation de l'analyse de données multivariée (AFM duale)

Cette partie décrit brièvement un cas d'utilisation de techniques d'analyse de données multivariée pour détecter des effets de mode. Pour des développements plus conséquents, on pourra se référer à l'article présenté aux Journées de méthodologie statistique (Guillerm et Razafindravona, 2015).

Il est classique dans une analyse du multimode de s'intéresser aux différentes variables de manière séparée. On montre alors que l'occurrence d'une certaine modalité est plus fréquente dans un mode plutôt qu'un autre. Ici, une autre optique est proposée en s'intéressant aux liaisons entre variables et en analysant comment elles évoluent quand on passe d'un mode de collecte à un autre. Pour des variables catégorielles, il s'agit donc de mettre en évidence des ensembles de modalités renseignées simultanément par un grand nombre d'individus ou au contraire rarement renseignées conjointement, et à analyser les différences de liaisons constatées selon le mode de collecte.

L'analyse factorielle multiple duale (AFMD) proposée par Lê et Pagès est particulièrement adaptée à cette analyse (Lê et Pagès, 2010). Elle répond à l'analyse factorielle de multi-tableaux

se présentant sous la forme d'un empilement de tableaux de données. Cette configuration se retrouve quand on dispose de plusieurs tableaux de données individus x variables (ici un tableau pour chaque mode de collecte) dont les variables sont communes. L'AFMD construit un unique jeu de composantes principales et assure que la contribution des différents groupes d'individus dans sa construction est équilibrée. Comme dans une analyse factorielle classique, ces composantes principales permettent de représenter graphiquement les liaisons entre variables (nuage des variables). L'apport de l'AFMD est de disposer d'un repère unique adapté à la représentation des corrélations entre variables aux niveaux global (pour l'ensemble des individus) d'une part et partiels (au sein de chaque groupe d'individus) d'autre part. Ces nuages peuvent être superposés. Chaque variable est ainsi représentée par quatre points (un point global et trois points partiels associés chacun à un mode de collecte, papier, internet, face à face). La trajectoire de ces points homologues permet d'analyser les différences de corrélations entre variables d'un groupe d'individus à un autre (étude de l'inter-groupe).

Une AFMD menée sur les données du dispositif CVS/VVS permet de mettre en évidence des différences de comportements de réponse selon le mode de collecte. Les premiers résultats (voir graphique page suivante) montrent un premier axe factoriel opposant les individus globalement satisfaits des institutions (police, gendarmerie, justice) et se sentant en sécurité à leur domicile, aux individus globalement insatisfaits et exprimant un sentiment d'insécurité. Les opinions exprimées apparaissent plus tranchées quand l'individu répond de manière auto-administrée (par internet ou par papier) par rapport à une interrogation en face à face. Sur le deuxième axe, la non-réponse aux différentes questions apparaît liée : ne pas répondre à une question va de pair avec la non-réponse aux autres questions. Cette liaison est plus forte dans une interrogation sur questionnaire papier par rapport à une interrogation en face à face. Un contrôle de la sélection sur caractéristiques observables a par ailleurs été réalisé et ne semble pas remettre en question ces résultats.

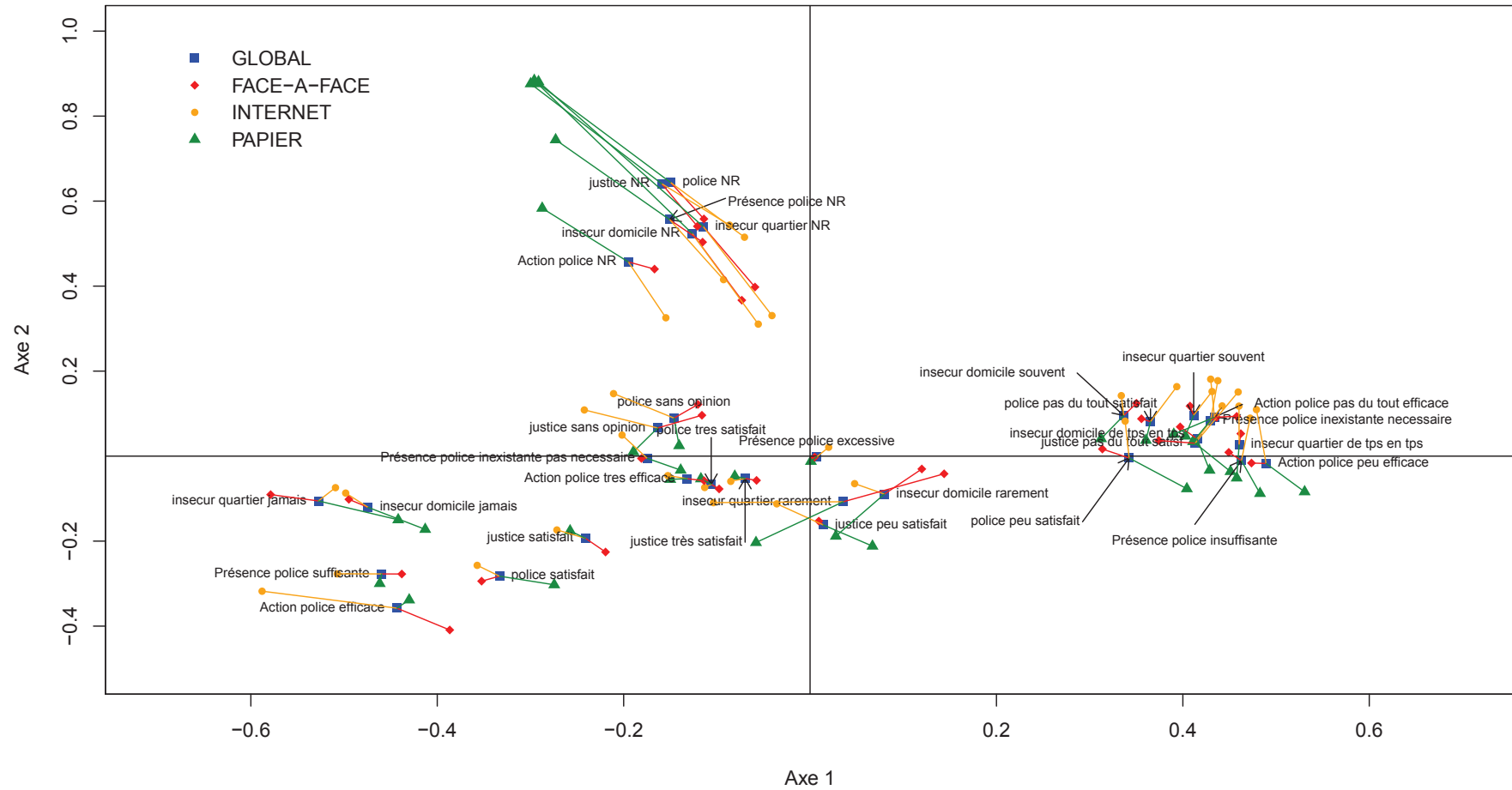


FIGURE 5.1 – AFMD : superposition des nuages global et partiels des modalités (Guillerm et Razafindranovona, 2015)

Conclusion

Quelques enseignements pratiques peuvent être tirés de l'exploitation de l'expérimentation Vols, violences et sécurité. En particulier, nous pouvons insister sur deux d'entre eux : l'un porte sur l'échantillonnage, l'autre sur le protocole de réponse.

Concernant l'échantillonnage, lorsque l'enquête vise nominativement des personnes sélectionnées, il est important de disposer de la base de sondage la plus fraîche possible : lorsque ce n'est pas le cas, la proportion de courriers qui ne parviennent pas à atteindre leur cible peut être massive en raison de la mobilité des personnes. Cela peut conduire à biaiser les estimations lorsque le comportement de mobilité est lié aux variables d'intérêt et que cette mobilité n'est qu'imparfaitement appréhendée par les variables de correction disponibles. Rappelons par ailleurs que les phénomènes de non réponse en tant que telle et de non contact ne reposent pas sur les mêmes déterminants : si les jeunes sont sous-représentés parmi les répondants, ce n'est pas tant parce qu'ils sont moins enclins à répondre au questionnaire mais surtout parce qu'il sont difficiles à contacter car mobiles. Dans certains cas, pour limiter ces problèmes, il peut même être judicieux de décaler le calendrier de l'enquête si cela permet de rafraîchir la base de sondage.

Pour limiter les problèmes d'auto-sélection au sein des ménages, deux protocoles ont été testés : s'ils sont proches du point de vue des résultats sur les taux de victimation, celui qui repose sur la fourniture de plusieurs identifiants présente un taux de réponse un peu plus faible et a également généré plus de déchets en apurement. Cela renvoie à l'idée que dans un cadre d'auto-administration, le protocole doit être le plus fluide possible, le plus naturel, le plus évident du point de vue de l'enquêté. Et ce, en particulier, lors des étapes cruciales du processus qui peuvent amener des enquêtés à décrocher. Une fois que l'enquêté est engagé dans l'auto-questionnaire, des efforts cognitifs ponctuels peuvent lui être demandés s'ils sont jugés nécessaires. Mais en amont, il faut probablement limiter au maximum les efforts à fournir par l'enquêté : le cheminement vers le questionnaire doit être le plus aisé possible.

L'enseignement majeur de l'expérimentation porte sur le problème de la sélection contrôlée. Il est important de rappeler qu'un protocole en face à face comme celui de CVS semble parvenir, au vu des taux de réponse de près de 70 %, à convaincre dans une large mesure l'essentiel de la population, y compris celle peu motivée par la thématique de l'enquête : on peut considérer que les biais liés à ces phénomènes de non-réponse non ignorable sont en grande partie limités. En revanche, un protocole auto-administré tel que celui de VVS, au taux de réponse bien moins élevé, ne parvient pas bien à persuader la population non motivée à participer. Tant que les taux de réponse des enquêtes auto-administrées seront aussi faibles, il est difficile d'envisager de fournir des estimations non biaisées des paramètres d'intérêt, ici les taux de victimation, qui reposent uniquement sur ces enquêtes. En effet, même après des calages qui peuvent sembler un peu forcés et un nombre de variables d'intérêt cibles assez limité, la convergence entre l'enquête

en face à face et l'expérimentation auto-administrée n'est pas satisfaisante.

Ce constat porte sur VVS mais ne se généralise pas forcément à toutes les enquêtes auto-administrées et n'est pas figé *ad vitam aeternam*. Néanmoins, dès lors que les taux de réponse sont faibles et que la thématique abordée peut engendrer une forte sélection des répondants selon la variable d'intérêt, il n'est sans doute pas raisonnable d'estimer directement des paramètres d'intérêt à partir d'enquêtes totalement et uniquement auto-administrées.

Par ailleurs, les problèmes de sélection contrôlée font que l'exploitation n'apporte pas de conclusion définitive sur l'effet de mode lié à la mesure pour les principales variables d'intérêt : on ne peut pas affirmer que les victimations soient sur-déclarées ou au contraire sous-déclarées en auto-administré. En revanche, certains constats, même s'ils sont relativement marginaux, comme la proportion plus importante de poly-victimes en auto-administré, ou encore une certaine tendance à choisir les premières modalités en auto-administré, laissent à penser que le passage d'un questionnaire long, détaillé, administré par un enquêteur à une version plus ramassée sans intermédiation d'enquêteur ne se fait pas sans heurts. Malheureusement, quelles que soient les précautions prises en termes de questionnaire, d'instructions, d'aide à la réponse ou de contrôles, les phénomènes de *satisficing* peuvent exister : si l'effort cognitif demandé est trop important, si le questionnaire est trop long et peu motivant, le répondant risque de faire des choix de réponse non optimaux. Tout doit être fait pour éviter ces écueils si l'on souhaite intégrer de nouveaux modes de collecte, mais cela est loin d'être simple.

Bibliographie

- Amiel, M.-H. et T. Denoyelle. 2012, «Enquêtes en ligne : comparaison de modes de questionnement sur le thème du logement», dans *Journées de Méthodologie Statistique*. http://jms.insee.fr/files/documents/2012/945_4-JMS2012_S26-3_AMIEL-ACTE.PDF.
- Betts, P. et C. Lound. 2010, «The application of alternative modes of data collection on UK government social surveys», *ONS, London*.
- Biemer, P. P. et L. E. Lyberg. 2003, *Introduction to survey quality*, vol. 335, John Wiley & Sons.
- Desrosières, A. 2004, «Enquêtes versus registres administratifs : réflexions sur la dualité des sources statistiques», *Courrier des statistiques*, vol. 111, p. 3–16.
- Givord, P. 2010, «Méthodes économétriques pour l'évaluation de politiques publiques», *INSEE, Document de travail de la DESE*. http://www.insee.fr/fr/publications-et-services/docs_doc_travail/G2010-08.pdf.
- Gombault, V. et M. Duée. 2012, «Un exemple d'enquête multimode à l'Insee : l'enquête TIC auprès des ménages», dans *7ème colloque francophone sur les sondages*, SFdS.
- Groves, R. M. et L. Lyberg. 2010, «Total survey error : Past, present, and future», *Public Opinion Quarterly*, vol. 74, n° 5, p. 849–879.
- Guillerm, M. et T. Razafindranovona. 2015, «L'apport de l'analyse factorielle multiple duale pour l'analyse de la collecte multimode», dans *Journées de Méthodologie Statistique*. http://jms.insee.fr/files/documents/2015/SO3_1_ACTE_V2_GUILLERM_JMS2015.PDF.
- Hainmueller, J. 2011, «Entropy balancing for causal effects : A multivariate reweighting method to produce balanced samples in observational studies», *Political Analysis*.
- Hansen, M. H. et W. N. Hurwitz. 1946, «The problem of non-response in sample surveys», *Journal of the American Statistical Association*, vol. 41, n° 236, p. 517–529.
- Heckman, J. J., H. Ichimura et P. Todd. 1998, «Matching as an econometric evaluation estimator», *The Review of Economic Studies*, vol. 65, n° 2, p. 261–294.
- Heckman, J. J., H. Ichimura et P. E. Todd. 1997, «Matching as an econometric evaluation estimator : Evidence from evaluating a job training programme», *The review of economic studies*, vol. 64, n° 4, p. 605–654.
- Krosnick, J. A. et D. F. Alwin. 1987, «An evaluation of a cognitive theory of response-order effects in survey measurement», *Public Opinion Quarterly*, vol. 51, n° 2, p. 201–219.
- Lê, S. et J. Pagès. 2010, «Dmfa : Dual multiple factor analysis», *Communications in Statistics - Theory and Methods*, vol. 39, n° 3, p. 483–492.
- Little, R. J. 1986, «Survey nonresponse adjustments», *International statistical review*, vol. 54, n° 1, p. 3.

- Lugtig, P., G. J. Lensvelt-Mulders, R. Frerichs et F. Greven. 2011, «Estimating nonresponse bias and mode effects in a mixed mode survey», *International Journal of Market Research*, vol. 53, n° 5, p. 669–686.
- Lyberg, L. 2012, «La qualité des enquêtes», *Techniques d'enquête*, vol. 38, n° 2, p. 115–142.
- de Peretti, G. et T. Razafindranovona. 2014, «Les enquêtes multimode : attention aux effets de mode», *Statistique et société*, vol. 2.
- Razafindranovona, T. 2015, «La collecte multimode et le paradigme de l'erreur d'enquête totale», *INSEE, Document de travail de la DMCSI*. http://www.insee.fr/fr/publications-et-services/docs_doc_travail/la%20collecte%20multimode-b.pdf.
- Razafindranovona, T., G. de Peretti et A. Barrau. 2013, «The philosophy of French experiments on internet and mixed-mode data collection», dans *UNECE Seminar on statistical data collection*.
- Roberts, C. 2007, «Mixing modes of data collection in surveys : A methodological review», cahier de recherche, ESRC NCRM.
- Rosenbaum, P. R. et D. B. Rubin. 1983, «The central role of the propensity score in observational studies for causal effects», *Biometrika*, vol. 70, n° 1, p. 41–55.
- Rubin, D. B. 1977, «Assignment to treatment group on the basis of a covariate», *Journal of Educational and Behavioral statistics*, vol. 2, n° 1, p. 1–26.
- Sautory, O. 1993, «La macro CALMAR», *Redressement d'un échantillon par calage sur marges, Série des documents de travail de la Direction des Statistiques Démographiques et Sociales*, vol. 55.
- Vannieuwenhuyze, J., G. Loosveldt et G. Molenberghs. 2014, «Evaluating mode effects in mixed-mode survey data using covariate adjustment models», *Journal of Official Statistics*, vol. 30, n° 1, p. 1–21.
- Van de Ven, W. P. et B. Van Praag. 1981, «The demand for deductibles in private health insurance : A probit model with sample selection», *Journal of econometrics*, vol. 17, n° 2, p. 229–252.
- Voogt, R. J. J. 2004, «I'm not interested : nonresponse bias, response bias and stimulus effects in election research», .
- Zauberman, R. et P. Robert. 2011, «Les enquêtes de victimation en europe», *Economie et statistique*, vol. 448, n° 1, p. 89–105.

Annexe A

Comparaison des taux de victimation web / papier

L'objectif principal de l'expérimentation est la comparaison des taux de victimation de CVS et de VVS. Néanmoins, au sein du protocole de VVS, deux modes de collecte sont utilisés et la question de la comparaison de ces modes peut se poser. Comme le web et le papier sont tous les deux des modes auto-administrés, la question est sans doute moins cruciale que lorsque les modes de collecte sont vraiment de nature différente.

Nous allons essayer de voir, par exemple, dans quelle mesure la proportion d'individus victimes de vols sans violence est différente chez les répondants web ou papier et si la différence persiste après un contrôle de la sélection.

A.0.1 Différences brutes

L'analyse est restreinte ici aux individus des sous-échantillons 1 et 2. Sur cette population, la proportion de victimes de vols sans violence est plus importante chez les répondants web. Néanmoins cette différence est moindre (pour des raisons liées à la sélection par la motivation) lorsque l'on compare uniquement les répondants après relance, car rappelons que la réponse par papier n'est possible qu'après avoir reçu le questionnaire contenu dans la lettre de relance.

	Web (tous)	Web (après relance)	Papier (après relance)
Vols sans violence (pondéré)	7,0 %	6,0 %	5,1 %
Vols sans violence (non pondéré)	7,1 %	6,2 %	5,0 %

TABLE A.1 – Taux de vols sans violence

Dans ce qui suit, nous n'allons pas tenir compte des pondérations de tirage. Le tableau A.1 montre en effet que l'impact des pondérations est somme toute assez limité sur les écarts de taux de victimation lorsque sont comparées réponses par papier et par internet. Par prudence, nous allons néanmoins nous assurer d'intégrer dans les variables contrôlant la sélection, le principal facteur de surreprésentation du tirage de l'échantillon : les revenus du ménage¹.

1. Notons par ailleurs que la littérature sur la technique d'évaluation que nous allons mettre en oeuvre, le

Il existe donc une différence d'environ 1 % dans les proportions de victimes de vols sans violence chez les répondants web et papier après relance et cette différence monte à 3 % lorsque l'on ne fait pas de restriction temporelle : dans quelle mesure les différences persistent après un contrôle *a minima* de la sélection (qui seraient donc potentiellement des effets de mode) corrigeant donc en partie les différences de structure entre ces sous-populations ? Le contrôle de la sélection sera réalisé en utilisant des techniques d'appariement (*matching*) (Lutig et al., 2011; Razafindranovona et al., 2013).

A.0.2 Principes du contrôle de la sélection par *kernel matching* sur score de propension

Le contexte est ici proche de celui de l'évaluation de l'effet causal d'un traitement où le traitement serait le fait de répondre par l'un des deux modes de collecte considérés. Pour un panorama complet des méthodes pour évaluer un effet causal, on pourra se référer à Givord (2010).

Rappelons tout d'abord brièvement le cadre canonique de Rubin (1974) de l'évaluation économétrique.

Le cadre canonique de Rubin

Dans le cadre de Rubin, le traitement est représenté par la variable D : si la personne est traitée, alors $D = 1$, sinon $D = 0$. Soit une variable d'intérêt Y . Chaque personne peut être dans deux états possibles, 0 ou 1, avec comme variables d'intérêt "potentielles" respectivement Y_0 et Y_1 . Le gain provenant du traitement est $\Delta = Y_1 - Y_0$.

Le problème de l'évaluation est que l'on ne connaît Δ pour personne étant donné que l'on observe $Y = DY_1 + (1 - D)Y_0$, c'est-à-dire, soit Y_0 , soit Y_1 .

Ainsi, pour une personne ayant reçu le traitement, on observe Y_1 : mais quelle serait la valeur de Y si la personne n'avait pas reçu le traitement ? L'essence des méthodes d'évaluation consiste à estimer ces situations contrefactuelles.

Une mesure de l'impact d'un programme que l'on souhaite classiquement obtenir est l'effet moyen du traitement dans la population des individus traités (*average treatment effect on the treated*) à savoir :

$$E(Y_1 - Y_0 | D = 1)$$

Un estimateur naïf consisterait à comparer la variable d'intérêt des traités avec celle des non traités mais il existe un biais de sélection dès lors que Y_0 et D ne sont pas indépendantes.

Effet du traitement conditionnel aux caractéristiques observables

L'effet moyen du traitement dans la population des individus traités, conditionnellement aux caractéristiques observables X est :

$$E(Y_1 - Y_0 | D = 1, X)$$

kernel matching sur score de propension, ne fournit pas de recommandations définitives sur l'utilisation des poids de sondage aux différentes étapes du processus.

La moyenne $E(Y_1 - Y_0|D = 1, X)$ est estimée à partir des données des participants aux programmes. En revanche, des hypothèses doivent être formulées pour estimer la moyenne contre-factuelle $E(Y_0|D = 1, X)$. En l'absence d'expérience contrôlée, $E(Y_0|D = 0, X)$ est utilisée pour approcher $E(Y_0|D = 1, X)$. Le biais de sélection provenant de cette approximation est :

$$B(X) = E(Y_0|D = 1, X) - E(Y_0|D = 0, X)$$

Pour construire le contrefactuel, l'hypothèse souvent invoquée est $Y_0 \perp D|X$. En fait, pour identifier l'effet du traitement, l'hypothèse est que le conditionnement par X élimine le biais ($B(X) = 0$) c'est à dire que :

$$E(Y_0|D = 1, X) = E(Y_0|D = 0, X)$$

Principes généraux du matching

Pour estimer l'effet du traitement conditionnellement aux caractéristiques observables X , une méthode d'évaluation très populaire est celle du matching. Le matching est basé sur l'idée intuitive de comparer la variable d'intérêt des traités avec la variable d'intérêt d'individus comparables (par leurs caractéristiques observables) qui eux ne sont pas traités. La différence entre ces variables d'intérêt est alors attribuée au traitement. Cette méthode fait donc l'hypothèse forte que le biais de sélection peut être purgé en conditionnant sur un ensemble de caractéristiques observables. Un avantage cette méthode est qu'elle ne suppose pas de spécifier la forme fonctionnelle de la relation entre variable d'intérêt et caractéristiques observables.

Un point important concernant l'utilisation de ces méthodes est que pour pouvoir comparer un individu traité avec un non traité aux caractéristiques similaires X , il faut s'assurer de l'existence de cet individu non traité mais similaire par ses caractéristiques X . Ce point se formalise à travers une hypothèse dite de support commun qui doit être vérifiée :

$$0 < P(D = 1|X) < 1$$

Nous notons dorénavant I_0 et I_1 l'ensemble des individus respectivement non-traités et traités et N_0 , N_1 leur nombre.

La méthode initialement proposée par Rubin (1977) consiste à apparier chaque individu traité i à un individu non traité comparable $\tilde{i}(i)$, au sens des caractéristiques observables X . Par exemple, on peut prendre pour contrefactuel l'individu non traité dont l'indice vérifie :

$$\tilde{i}(i) = \operatorname{argmin}_{j \in I_0} \|X_i - X_j\|$$

où $\| \cdot \|$ est une distance ; on peut prendre par exemple, la distance de Mahalanobis.

L'estimateur de l'effet sur les traités est alors :

$$\frac{1}{N_1} \sum_{i \in I_1} [Y_i - Y_{\tilde{i}(i)}]$$

Dans cette version simple du matching, le contrefactuel de l'individu traité se fait à partir d'un seul individu comparable. Il est également possible d'utiliser plusieurs individus comparables et de pondérer alors selon la distance à l'individu traité selon les caractéristiques observables.

Ainsi, une forme plus générale de l'estimateur de matching est proposée par Heckman, Ichimura et Todd (1997 et 1998) :

$$\sum_{i \in I_1} \omega_{N_0, N_1}(i) [Y_{1i} - \sum_{j \in I_0} W_{N_0, N_1}(i, j) Y_{0j}]$$

$W_{N_0, N_1}(i, j)$ est une pondération définie telle que pour tout $i \in I_1$, $\sum_{j \in I_0} W_{N_0, N_1}(i, j) = 1$.

Le choix d'une fonction de pondération reflète le choix d'une distance particulière dans la méthode de matching et les poids sont basés sur les distances dans l'espace des X .

Les $\omega_{N_0, N_1}(i)$ servent à sélectionner différents domaines de l'espace des X ou à tenir compte de l'hétéroscédasticité dans l'échantillon des traités.

Ainsi les différentes méthodes de matching correspondent à des choix de différentes fonctions de pondération $\omega_{N_0, N_1}(i)$ et $W_{N_0, N_1}(i, j)$.

Par ailleurs soit le voisinage $C(X_i)$ pour une personne de l'échantillon des traités. Les voisins de i sont les personnes $j \in I_0$ pour lesquelles $X_j \in C(X_i)$. Les personnes appariées à i sont celles de l'ensemble A_i tel que $A_i = \{j \in I_0 | X_j \in C(X_i)\}$.

Par exemple l'estimateur de matching de Rubin, dit du plus proche voisin présenté précédemment, consiste à prendre $\omega_{N_0, N_1}(i) = 1/N_1$, $C(X_i) = \min_{j \in I_0} \|X_i - X_j\|$ et où le schéma de pondération consiste à assigner tout le poids au seul appariement, soit $W_{N_0, N_1}(i, j) = 1$ si $j \in A_i$ et $W_{N_0, N_1}(i, j) = 0$ sinon.

Score de propension et matching

Pour que l'hypothèse d'indépendance conditionnelle soit vérifiée, on peut souhaiter effectuer l'appariement sur un grand nombre de variables. Mais les proches voisins peuvent alors être difficiles à trouver. En fait, les estimateurs sont d'autant plus biaisés que le nombre de variables de conditionnement est élevé. Un moyen de résoudre ce problème de dimensionnalité est d'utiliser le score de propension.

Rosenbaum et Rubin (1983) ont montré que si la propriété d'indépendance conditionnelle $Y_0 \perp D | X$ est vérifiée alors, en notant $P(X) = P(D = 1 | X)$, si on a également $0 < P(X) < 1$, on a l'indépendance $Y_0 \perp D | P(X)$. L'estimation de l'effet sur les traités peut alors s'effectuer en appariant sur le score de propension. En pratique, si après avoir conditionné par l'estimation du score de propension, la distribution des X dépend encore du traitement D , cela suggère une mauvaise spécification.

Précisons que si le problème de la dimensionnalité semble résolu, en pratique, le score de propension devra être estimé (car le score exact n'est que rarement connu) souvent par des spécifications *logit* ou *probit*. Alors qu'un des intérêts du matching est son caractère *a priori* non paramétrique, le matching sur score de propension va souvent reposer en partie sur des

hypothèses de spécification.

Kernel matching

Les méthodes de matching sur les plus proches voisins n'utilisent qu'une partie de l'information disponible : on peut souhaiter utiliser pour chaque individu traité l'ensemble de l'échantillon des non traités. Les méthodes de *kernel matching* reposent ainsi sur le fait que chaque individu non traité participe au contrefactuel de l'individu i avec une importance qui varie selon la distance avec l'individu considéré.

En reprenant les notations précédentes, l'estimateur de *kernel matching* s'écrit avec :

$$\omega_{N_0, N_1}(i) = \frac{1}{N_1} \quad , \quad A_i = I_0 \quad \text{et} \quad W_{N_0, N_1}(i, j) = \frac{G_{ij}}{\sum_{k \in I_0} G_{ik}} \quad \text{avec} \quad G_{ik} = G\left(\frac{X_k - X_i}{a_{N_0}}\right)$$

G étant une fonction noyau (positive, symétrique par rapport à 0 et telle que $\int_{-\infty}^{+\infty} G(u)du = 1$) et a_{N_0} un paramètre de fenêtre.

Des exemples de noyaux souvent utilisés sont le noyau gaussien et le noyau d'Epanechnikov.

Par définition de la fonction noyau, plus l'individu du groupe de contrôle est "proche" de l'individu traité plus sa contribution est importante dans la construction du contrefactuel. Le choix de la fenêtre relève d'un arbitrage classique entre biais et variance : plus la fenêtre est grande, plus la variance diminue mais en contrepartie le biais est également plus élevé. Par ailleurs, avec ce type d'estimateur où tous les individus du groupe de non-traités sont utilisés pour construire le contrefactuel de l'individu traité, l'hypothèse de support commun est très importante et le praticien devra y prêter une certaine attention.

Dans la version où le matching s'effectue sur le score de propension, cet estimateur peut s'écrire avec $P_i = P(D = 1|X_i)$:

$$\frac{1}{N_1} \sum_{i \in I_1} \left[Y_{1i} - \frac{\sum_{j \in I_0} G\left(\frac{P_j - P_i}{a_{N_0}}\right) Y_{0j}}{\sum_{k \in I_0} G\left(\frac{P_k - P_i}{a_{N_0}}\right)} \right]$$

A.0.3 Résultats

Une question intéressante que l'on peut se poser est de savoir si, dans notre contexte, il est préférable de comparer la population de l'ensemble des répondants papier à celle de l'ensemble des répondants web ou vaut mieux-t-il, pour cette dernière, se restreindre à celle qui répond après la relance postale. Pour illustrer les phénomènes qui peuvent être liés à la sélection non contrôlée, les deux configurations sont analysées. La méthode qui consiste à comparer les deux sous-populations répondant après relance (à partir de ce moment, la possibilité de répondre par l'un ou l'autre mode est offerte) semble être la plus pertinente dès lors que l'on souhaite mesurer

des effets de mode et "contrôler" au mieux la sélection sur inobservables². Essayer de mesurer l'effet causal dans l'autre configuration va donner quelques éléments en termes de sélection non contrôlée.

Les variables de contrôle X intégrées sont le sexe, l'âge, le niveau de diplôme, les revenus et le type de ménage. Une première étape consiste à modéliser le fait de répondre par l'un ou l'autre des modes selon ces variables de contrôle, par un modèle logistique. Cette modélisation sera utilisée pour calculer le score de propension. Les *odds-ratios* sont présentés dans le tableau A.2 (la variable à expliquer est le fait de répondre par papier) dans la configuration où l'on compare les sous-populations après relance postale.

	Odds-ratios
Femme vs Homme	1,3
18-24 ans vs 70 ans et plus	0,1
25-39 ans vs 70 ans et plus	0,1
40-49 ans vs 70 ans et plus	0,2
50-59 ans vs 70 ans et plus	0,3
60-69 ans vs 70 ans et plus	0,4
Aucun diplôme vs 2ème cycle universitaire et supérieur	2,6
CEP vs 2ème cycle universitaire et supérieur	2,8
CAP, BEP vs 2ème cycle universitaire et supérieur	2,6
Brevet vs 2ème cycle universitaire et supérieur	1,7
Bac vs 2ème cycle universitaire et supérieur	2,0
Bac+2 vs 2ème cycle universitaire et supérieur	1,5
Personne seule vs Couple avec enfants	1,6
Famille monoparentale vs Couple avec enfants	1,6
Couple sans enfants vs Couple avec enfants	1,1
Revenus du foyer inférieurs à 10k€ vs Revenus du foyer supérieurs à 60k€	1,6
Revenus du foyer entre 10k€ et 20k€ vs Revenus du foyer supérieurs à 60k€	1,5
Revenus du foyer entre 20k€ et 30k€ vs Revenus du foyer supérieurs à 60k€	1,4
Revenus du foyer entre 30k€ et 40k€ vs Revenus du foyer supérieurs à 60k€	1,3
Revenus du foyer entre 40k€ et 50k€ vs Revenus du foyer supérieurs à 60k€	1,3
Revenus du foyer entre 50k€ et 60k€ vs Revenus du foyer supérieurs à 60k€	1,1

TABLE A.2 – *Odds-ratios* de la régression logistique où la variable à expliquer est la réponse par papier

Le graphique A.1 donne la distribution des scores de propension selon que l'individu appartient au groupe de contrôle (répondants web) ou au groupe des traités (répondants papier) et permet d'illustrer l'existence d'un support de la distribution des scores commun.

2. Dans le verso de la lettre-avis, il est dit qu'une version papier du questionnaire sera envoyée dans quelques semaines. Il est donc possible que certaines personnes "motivées" pour répondre attendent la réception du questionnaire papier parce que ce mode de réponse a leur préférence. Le contrôle aurait été encore plus pertinent dans le cas où cette précision n'aurait pas été indiquée.

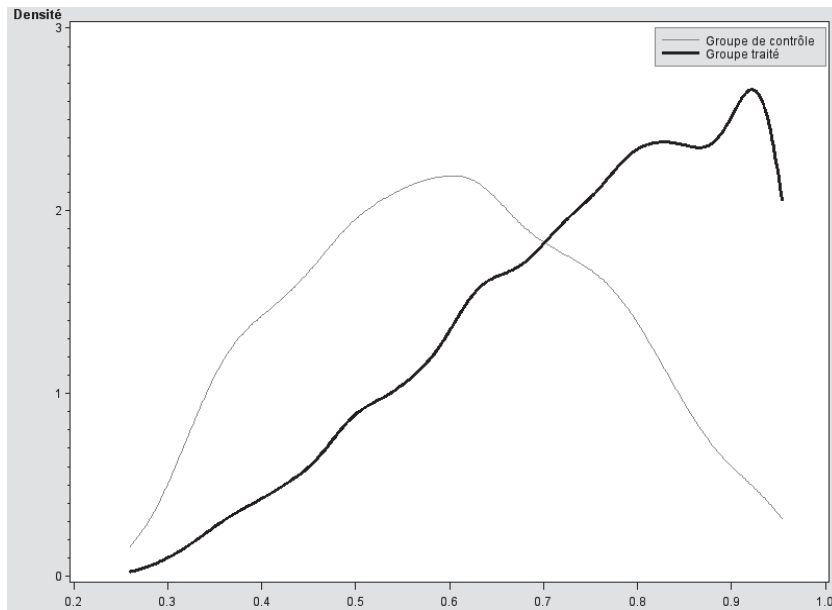


FIGURE A.1 – Distribution des scores de propension conditionnels

Enfin, le tableau A.3 montre que l'effet causal du mode de réponse est nul une fois que la sélection est contrôlée : il semble donc que pour cette question portant sur les vols, il n'y a pas de différence liée à la mesure.

	Vols sans violence (%)
Estimation directe $Y_{web} - Y_{papier}$	1,2 %
Estimation par <i>kernel matching</i>	0,0 %

TABLE A.3 – Différences avant/après contrôle par matching (répondants après relance)

En revanche, la tableau A.4 montre qu'il existe un effet d'environ 1 % lorsque l'on ne restreint pas aux réponses arrivées après la relance postale. Cela tient vraisemblablement au fait que les populations sont moins "comparables" car elles diffèrent sur un aspect de la sélection qui n'est pas pris en compte dans le contrôle.

	Vols sans violence (%)
Estimation directe $Y_{web} - Y_{papier}$	2,1 %
Estimation par <i>kernel matching</i>	1,1 %

TABLE A.4 – Différences avant/après contrôle par matching (tous répondants)

Annexe B

Enquête sur les non-répondants à l'expérimentation VVS

Une opération téléphonique de rappel de non-répondants à l'expérimentation VVS a été réalisée. Pour cette opération, l'échantillon était composé de 1 000 non-répondants. Environ 600 personnes ont répondu à cette enquête dont l'objectif est de recueillir les raisons pour lesquelles ces personnes n'ont pas participé à l'expérimentation VVS. L'intitulé exact de la question posée est " Pour quelles raisons n'avez-vous pas répondu à l'enquête?". L'enquêteur laisse tout d'abord la personne s'exprimer avant de lui énumérer les modalités possibles. Les résultats sont présentés dans le tableau suivant (plusieurs réponses sont possibles).

	Raisons (%)
Vous n'avez pas accès à Internet	17,1 %
Vous n'avez pas reçu la lettre-avis	16,8 %
Vous n'aviez pas le temps	16,0 %
Vous ne répondez pas aux enquêtes	10,9 %
Vous ne vous sentez pas concerné par le thème de cette enquête	12,0 %
Vous ne souhaitez pas aborder ce thème sur Internet	3,1 %
Vous avez essayé de répondre à l'enquête mais vous n'avez pas réussi	1,5 %
Vous avez oublié de répondre	8,6 %
L'enquête est trop longue	0,7 %
Il ne vous était pas possible de répondre (santé)	6,4 %
Autres	21,3 %

TABLE B.1 – Résultats de l'enquête auprès des non-répondants

Série des Documents de Travail « Méthodologie Statistique »

- 9601** : Une méthode synthétique, robuste et efficace pour réaliser des estimations locales de population.
G. DECAUDIN, J.-C. LABAT
- 9602** : Estimation de la précision d'un solde dans les enquêtes de conjoncture auprès des entreprises.
N. CARON, P. RAVALET, O. SAUTORY
- 9603** : La procédure **FREQ** de **SAS** - Tests d'indépendance et mesures d'association dans un tableau de contingence.
J. CONFAIS, Y. GRELET, M. LE GUEN
- 9604** : Les principales techniques de correction de la non-réponse et les modèles associés.
N. CARON
- 9605** : L'estimation du taux d'évolution des dépenses d'équipement dans l'enquête de conjoncture : analyse et voies d'amélioration.
P. RAVALET
- 9606** : L'économétrie et l'étude des comportements. Présentation et mise en œuvre de modèles de régression qualitatifs. Les modèles univariés à résidus logistiques ou normaux (**LOGIT**, **PROBIT**).
S. LOLLIVIER, M. MARPSAT, D. VERGER
- 9607** : Enquêtes régionales sur les déplacements des ménages : l'expérience de Rhône-Alpes.
N. CARON, D. LE BLANC
- 9701** : Une bonne petite enquête vaut-elle mieux qu'un mauvais recensement ?
J.-C. DEVILLE
- 9702** : Modèles univariés et modèles de durée sur données individuelles.
S. LOLLIVIER
- 9703** : Comparaison de deux estimateurs par le ratio stratifiés et application aux enquêtes auprès des entreprises.
N. CARON, J.-C. DEVILLE
- 9704** : La faisabilité d'une enquête auprès des ménages.
1. au mois d'août.
2. à un rythme hebdomadaire
C. LAGARENNE, C. THIESSET
- 9705** : Méthodologie de l'enquête sur les déplacements dans l'agglomération toulousaine.
P. GIRARD
- 9801** : Les logiciels de désaisonnalisation **TRAMO & SEATS** : philosophie, principes et mise en œuvre sous **SAS**.
K. ATTAL-TOUBERT, D. LADIRAY
- 9802** : Estimation de variance pour des statistiques complexes : technique des résidus et de linéarisation.
J.-C. DEVILLE
- 9803** : Pour essayer d'en finir avec l'individu Kish.
J.-C. DEVILLE
- 9804** : Une nouvelle (encore une !) méthode de tirage à probabilités inégales.
J.-C. DEVILLE
- 9805** : Variance et estimation de variance en cas d'erreurs de mesure non corrélées ou de l'intrusion d'un individu Kish.
J.-C. DEVILLE
- 9806** : Estimation de précision de données issues d'enquêtes : document méthodologique sur le logiciel **POULPE**.
N. CARON, J.-C. DEVILLE, O. SAUTORY
- 9807** : Estimation de données régionales à l'aide de techniques d'analyse multidimensionnelle.
K. ATTAL-TOUBERT, O. SAUTORY
- 9808** : Matrices de mobilité et calcul de la précision associée.
N. CARON, C. CHAMBAZ
- 9809** : Échantillonnage et stratification : une étude empirique des gains de précision.
J. LE GUENNEC
- 9810** : Le Kish : les problèmes de réalisation du tirage et de son extrapolation.
C. BERTHIER, N. CARON, B. NEROS
- 9901** : Perte de précision liée au tirage d'un ou plusieurs individus Kish.
N. CARON
- 9902** : Estimation de variance en présence de données imputées : un exemple à partir de l'enquête Panel Européen.
N. CARON
- 0001** : L'économétrie et l'étude des comportements. Présentation et mise en œuvre de modèles de régression qualitatifs. Les modèles univariés à résidus logistiques ou normaux (**LOGIT**, **PROBIT**) (version actualisée).
S. LOLLIVIER, M. MARPSAT, D. VERGER
- 0002** : Modèles structurels et variables explicatives endogènes.
J.-M. ROBIN
- 0003** : L'enquête 1997-1998 sur le devenir des personnes sorties du RMI - Une présentation de son déroulement.
D. ENEAU, D. GUILLEMOT
- 0004** : Plus d'amis, plus proches ? Essai de comparaison de deux enquêtes peu comparables.
O. GODECHOT
- 0005** : Estimation dans les enquêtes répétées : application à l'Enquête Emploi en Continu.
N. CARON, P. RAVALET
- 0006** : Non-parametric approach to the cost-of-living index.
F. MAGNIEN, J. POUGNARD
- 0101** : Diverses macros **SAS** : Analyse exploratoire des données, Analyse des séries temporelles.
D. LADIRAY
- 0102** : Économétrie linéaire des panels : une introduction.
T. MAGNAC
- 0201** : Application des méthodes de calages à l'enquête EAE-Commerce.
N. CARON
- C 0201** : Comportement face au risque et à l'avenir et accumulation patrimoniale - Bilan d'une expérimentation.
L. ARRONDEL, A. MASSON, D. VERGER
- C 0202** : Enquête Méthodologique Information et Vie Quotidienne - Tome 1 : bilan du test 1, novembre 2002.
J.-A. VALLET, G. BONNET, J.-C. EMIN, J. LEVASSEUR, T. ROCHER, P. VRIGNAUD, X. D'HAULTFOEUILLE, F. MURAT, D. VERGER, P. ZAMORA
- 0203** : General principles for data editing in business surveys and how to optimise it.
P. RIVIERE
- 0301** : Les modèles logit polytomiques non ordonnés : théories et applications.
C. AFSA ESSAFI
- 0401** : Enquête sur le patrimoine des ménages - Synthèse des entretiens monographiques.
V. COHEN, C. DEMMER
- 0402** : La macro **SAS** **CUBE** d'échantillonnage équilibré
S. ROUSSEAU, F. TARDIEU
- 0501** : Correction de la non-réponse et calage de l'enquêtes Santé 2002
N. CARON, S. ROUSSEAU

0502 : Correction de la non-réponse par répondération et par imputation
N. CARON

0503 : Introduction à la pratique des indices statistiques - notes de cours
J-P BERTHIER

0601 : La difficile mesure des pratiques dans le domaine du sport et de la culture - bilan d'une opération méthodologique
C. LANDRE, D. VERGER

0801 : Rapport du groupe de réflexion sur la qualité

des enquêtes auprès des ménages

D. VERGER

M2013/01 : La régression quantile en pratique

P. GIVORD, X. D'HAULTFOEUILLE

M2014/01 : La microsimulation dynamique : principes généraux et exemples en langage R

D. BLANCHET

M2015/01 : la collecte multimode et le paradigme de l'erreur d'enquête totale

T. RAZAFINDROVONA

M2015/02 : Les méthodes de Pseudo-Panel

M. GUILLERM

M2015/03 : Les méthodes d'estimation de la précision pour les enquêtes ménages de l'Insee tirées dans Octopusse

E. GROS – K.MOUSSALAM

M2016/01 : Le modèle Logit Théorie et application.

C. AFSA

M2016/02 : Les méthodes d'estimation de la précision

de l'Enquête Emploi en Continu

E. GROS – K.MOUSSALAM

M2016/03 : Exploitation de l'enquête expérimentale Vols, violence et sécurité.

T. RAZAFINDROVONA