

N°22
Octobre 2023

Méthode

Methodological assessment report on the overhaul of victimization survey in France : transition to multimode

Authors: Hélène Guedj, Axelle Quiviger, Tiaray Razafindranovona, Agathe Rosenzweig, Sandra Zilloniz

Editor-in-chief : Olivier Filatriau

This publication was presented and discussed at the Scientific Council of the " Experience and Feelings concerning Security " survey (meeting on November 15, 2021) and at the Quality Label Committee (meeting on January 20, 2022).

CONTENTS

INTRODUCTION	5
PART I – The multimode switch: between a sense of history and experimental science	7
1. General information on multimode surveys and mode effects	8
1.1 Multimode surveys 7	
1.1.1 Contact modes and collection modes.....	8
1.1.2 Collection modes	7
1.1.3 Main multimode protocols.....	9
1.1.4 What are the objectives?	9
1.2 Mode effects 9	
1.2.1 Definition	9
1.2.2 Selection effect and measurement effect	10
1.3 Mode effects strictly linked to measurement 10	
1.3.1 Satisficing and social desirability	10
1.3.2 Questionnaires and multimode	10
1.3.3 Assessment	11
1.3.4 Main findings.....	11
2. Insee's multimode victimisation survey experiments.....	12
2.1 The development of the use of multimode and experiments within the public statistics department 13	
2.1.1 Multimode experiments at Insee.....	13
2.1.2 Promising work with the pooling of methods and tools	14
2.2 The 2013 “Theft, Violence and Security” experiment 15	
2.2.1 Presentation.....	14
2.2.2 Main findings.....	15
2.2.3 Conclusions	16
2.3 The 2019 CVS “Panel Multimode” experiment 18	
2.3.1 Presentation.....	17
2.3.2 Main findings.....	19
2.3.3 Conclusions	20
2.4 The Panel in SRCV experiment 22	
2.4.1 Presentation.....	21
2.4.2 Main findings.....	22
2.4.3 Conclusions	22
3. The Genese survey: the full-scale multimode experiment	25
3.1 The framework for analysis: data collected in phase 1 of the Genese survey	
Erreur ! Signet non défini.	

3.2	Characterisation of respondent profiles by mode and by eligibility profile for telephone follow-up	25
3.2.1	Characterisation by mode.....	24
3.2.1.1	Auxiliary variables generated by the sampling frame	Erreur ! Signet non défini.
3.2.1.2	Socio-demographic variables generated by the questionnaire	Erreur ! Signet non défini.
3.2.2	Determining factors in the choice between responding via the Internet and responding by telephone	28
3.3	Evaluation of mode effects on all the victimisation and opinion variables	Erreur ! Signet non défini.
3.3.1	Using regression models to evaluate the mode effect linked to measurement	Erreur ! Signet non défini.
3.3.1.1	Modelling used.....	31
3.3.1.2	Estimated mode effects derived from regressions upon variables of interest	Erreur ! Signet non défini.
3.3.2	Using matching models to evaluate the mode effect linked to measurement..	Erreur ! Signet non défini.
3.3.2.1	General principles of modelling by matching	34
3.3.2.2	The question of common support	34
3.3.2.3	Mode effects assessed using matching models	Erreur ! Signet non défini.
3.4	Elements for discussions on the interpretation of effects	38
PART II – The historical perspective: connecting the past to the future in a present marked by the health crisis		39
4.	General information on the overhaul of household surveys and managing serial breaks	Erreur ! Signet non défini.
4.1	Frequent overhauls of household surveys but very limited literature on their impact and correction	40
4.2	Theory on methods for managing serial breaks	42
4.2.1	Parallel operation of the current process and the updated process	Erreur ! Signet non défini.
4.2.1.1	General principle.....	41
4.2.1.2	Synthetic correction methods	42
4.2.1.3	The model approach in situations where numbers are too low	Erreur ! Signet non défini.
4.2.2	Time series approach	43
4.2.2.1	Principle	42
4.2.2.2	Implementation	42
5.	Victimisation measurement in France: discontinuities and connections	Erreur ! Signet non défini.
5.1	Frameworks	44
5.2	Connection work during previous overhauls	46
5.3	CVS 2007-2019: a fairly stable system but not devoid of discontinuities	46
5.3.1	Questionnaires 2007 to 2019	47

5.3.2	The overhaul of the adjustment methodology in 2016 and the back-casting of the series	50
5.4	The 2020-2021 turning point: a switch marked by the health crisis	Erreur ! Signet non défini.
5.4.1	The two previous editions of CVS	53
5.4.2	Auxiliary sources to compare for the year 2020	56
6.	Examples of corrections of serial breaks in the Employment survey and in Swedish and Dutch victimisation surveys, within the SSP and in neighbouring countries	58
6.1	Case studies involving the connection of indicator series generated by surveys repeated over time	58
6.1.1	Overhauls of the Employment survey: contexts and backcasting techniques	Erreur ! Signet non défini.
6.1.2	The backcasting of series derived from victimisation surveys carried out in the Netherlands	61
6.1.3	The backcasting of series derived from victimisation surveys carried out in Sweden	64
PART III – CVS/Genese comparisons: initial elements for analysis		64
7.	Victimisation indicators (prevalence/consistent multivictimisation /referral)	64
8.	Indicators for perceptions and opinions concerning security	65
PART IV Summary and outlook		71
9.	Key insights	71
9.1	Mode effects that are generally limited as regards victimisation but more marked in questions regarding opinions and perceptions concerning security.	71
9.2	The prevalence of victimisation derived from Genese is indeed consistent with that of the historical CVS series	72
9.3	Documenting overhauls: indispensable at the present time and for the near future, probably very useful even in the very distant future	72
10.	Outlook and road map for the future victimisation survey system.	Erreur ! Signet non défini.
10.1	Consolidating the mode effect findings	72
10.2	The question of comparability over time	73
10.3	Correcting and tailoring mode effects	74
Bibliography		73
APPENDICES		80

INTRODUCTION

The “*Cadre de vie et sécurité*” (“Living Environment and Security”) survey (CVS, cf. Appendix I) is an annual victimisation survey which was carried out nationally by Insee (the French National Institute of Statistics and Economic Studies) from 2007 to 2021, a project it jointly managed with the *Observatoire national de la délinquance et des réponses pénales* (ONDRP, the French National Observatory of Delinquency and Penal Responses, until it was shut down in 2020) and the *Service statistique ministériel de la sécurité intérieure* (SSMSI, the French Ministerial Statistics Department for Internal Security created in 2014). Its objective is to discover the facts regarding incidents of delinquency of which households and individuals may have been the victims during the two years prior to the survey, and whether or not these incidents have given rise to statements taken by the police or *gendarmeries*. It also allows the gathering of the entire population’s (victims and non-victims) opinions concerning their living environment and security, and an analysis of feelings of insecurity and levels of satisfaction with the work of the justice system and the security forces. The information obtained through the victimisation survey is separate from and complementary to data recorded by the police and the national *gendarmeries* because victims do not always file complaints. Combined, they provide valuable tools for assessing and analysing both delinquency and feelings of insecurity.

Like other public statistics surveys, the CVS survey has evolved over the years, with the modification of questions, additions and the modification of thematic modules (transport, fraud, corruption, discrimination, etc.), specific territorial extensions in the French overseas territories, etc. In April 2018, Insee informed the SSMSI and the ONDRP that, due to severe budgetary constraints in survey programming and associated resources in the form of interviewers, it would no longer be able to carry out the CVS survey in its current form (annual basis, mode of collection) from 2022 onwards. The Ministry of the Interior made a commitment to introduce financial and human resources in order to ensure, via the SSMSI, the sustainability of a victimisation survey system, taking advantage of the opportunities offered through combining various modes of collection (Internet, telephone, face-to-face, etc.). The new system’s objective is to address the need to understand both the level and structure of victimisation and perceptions concerning security, and it aims to maximise representative findings at a sub-national level at a sustainable cost.

In line with commitments made to the *Conseil national de l’information statistique* (CNIS, the French National Council for Statistical Information), since the spring of 2019 the SSMSI, with the support of Insee, has been leading discussions on the evolution of the CVS survey system’s overall structure between now and 2022. In particular, it must respond to the social demand expressed at the meetings organised by the CNIS in 2015 and 2016. Within this context, experiments were conducted at Insee in 2019. Prior to the consultation’s launch, the SSMSI conducted two written consultations in order to analyse the CVS survey’s strengths and weaknesses in relation to other existing systems and draw up a road map for the overhaul project. The initial consultation was carried out between August and November 2019, with experts in victimisation and perceptions concerning security who were associated with the CVS survey consultation. Then, in October 2020 (behind schedule due to the health crisis), a “thematic” consultation was conducted with departmental statistics services (SSM) and various entities with an interest in these subjects (administrations, researchers, etc.). This work highlighted the need to design a system capable of overcoming CVS’s weaknesses and blind spots, in particular the problem of the accuracy of assessments. In parallel with these consultations, from July 2019 to December 2020, the SSMSI introduced and led a methodological working group linked to Insee to guarantee the designing of the Genese (*Genre et sécurité*, Gender and Security) survey (cf. Appendix II) which constituted a full-scale field experiment in 2021.

On 14 December 2020, when this work had been completed, the SSMSI officially organised the launch meeting for the consultation on the overhaul of the victimisation survey system, which was attended by Insee, departmental statistics services, government bodies, research centres and trade union representatives, professional organisations and local councils, and CNIS members. During this meeting, the SSMSI presented the conclusions of the consultations that had been carried out as well as the project’s comitology, which is structured around a steering committee responsible for ensuring that objectives are met, monitoring the progress of the design, preparation and implementation phases, and

a consultation committee to collaboratively design the statistical methodology, the field of population of interest, the field of those affected, the questionnaires and the documents relating to the fields concerned by the survey. For the quality and scope of the work carried out by the SSMSI, it was also felt that it was very important to provide the survey overhaul project with a multidisciplinary Scientific Council (CS). The CS's role was to scientifically assess work carried out within the framework of the consultation committee, in particular by making decisions concerning methodological arbitration (questionnaires and protocol). The CS's opinion on the exploitation of the data produced is also eagerly awaited by the SSMSI. Indeed the SSMSI's ambition is develop a system capable of providing data that is widely disseminated and exploited by researchers.

Building on existing surveys, and in association with Insee, the SSMSI designed a multimode two-phase protocol that makes it possible, at a sustainable cost and in an innovative and reliable way, to meet the emerging need for more accurate, localised data, whilst still providing a detailed description of victims and offences experienced. On the occasion of its initial meeting, the Scientific Council recommended that the project team conduct a rigorous methodological evaluation of the impact of the change of system, both from the point of view of controlling the biases relating to the multimode switch and also concerning comparability in the timeliness of the findings prior to the 2022 survey review by the Quality Label Committee at the end of 2021.

This four-part report was jointly produced by the SSMSI's Statistical Surveys section and Statistical Methods and Short-Term Analyses section in response to the Scientific Council's recommendations. The first part uses a scientific approach to examine the challenges associated with multimode development and the main insights gained from the CVS multimodal experiments and the Genese survey, in terms of mode effects in the measurement of victimisation and perceptions concerning security. The second part of the report is devoted to the question of the continuity of the series: firstly it offers an overview of overhauls of household surveys within the public statistics service, and then presents elements of theory on the connection of statistical series. It then explores work carried out to connect victimisation series generated by the various observation systems that have been used since the beginning of the 1980s. Lastly, it examines concrete cases featuring the management of series breaks in the Employment survey and in Dutch and Swedish victimisation surveys. The third part offers some new initial elements of comparison in the main indicators resulting from CVS and the Genese survey. Finally the last part uses a prospective approach to propose a synthesis of all the elements presented in order to identify the main insights for the new victimisation survey.

This report is to be attached to the survey presentation file which is to be submitted for consideration by the Quality Label Committee in January 2021.

PART I – The multimode switch: between a sense of history and experimental science

1. General information on multimode surveys and mode effects

The development of new technologies and the possibility of collecting information, at a reduced cost, from households and businesses via online questionnaires has contributed to the enthusiasm for so-called “multimode” surveys. Yet multimode is not a complete novelty: for quite some time, certain statistical surveys have involved several modes of collection in their production processes. But the widespread use of self-administered online questionnaires raises certain questions, when the usual and validated paradigm was that of collection via the intermediary of an interviewer. In particular, a central question is that of mode effects. This chapter therefore furnishes some general information on these various notions, based on articles and reference works that may also supplement our remarks with subsequent developments (Couper, 2011; de Leeuw 2018; Schouten *et al.*, 2021).

Multimode surveys

1.1.1 Contact modes and collection modes

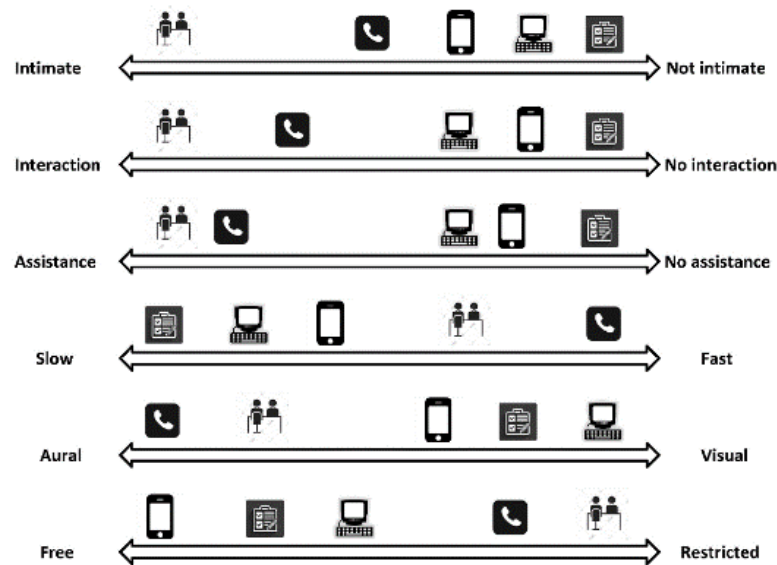
Typically, a “multimode survey” is defined as a survey that involves several modes in the collection of data from the units approached. More precisely, the multimode dimension usually refers to the collection modes, namely the various modes by which the units may respond to the survey. But multiplicity may also exist in the means of communication used to reach or send reminders to respondent units (de Leeuw, 2018): notification letters sent by post, emails pushing for responses via Internet, SMS, voicemail messages, etc. This contact dimension should therefore not be ignored in an overall analysis of a survey system.

1.1.2 Collection modes

Today, the main collection modes used to collect responses from surveys are (Groves *et al.*, 2011): computer-assisted face-to-face collection (CAPI, *computer-assisted personal interviewing*), computer-assisted telephone collection (CATI, *computer-assisted telephone interviewing*), collection via Internet (CAWI, *computer-assisted web interviewing*) and paper collection (PAPI, *paper and pencil interviewing*). The usual dichotomy is to distinguish modes using an intermediary in the form of an interviewer (CAPI, CATI) from self-administered modes (CAWI, PAPI).

The modes of collection may be characterised according to various dimensions and are represented graphically (figure 1.1) according to their positions on the underlying scales (Schouten *et al.*, 2021): degree of intimacy (intrusion of communications into personal life), human interaction, support options, pace of the questioning, presentation (oral or visual) and timing (when and where).

Figure 1.1: Collection modes and their characteristics



Source: Schouten et al., 2021

Smartphones and computers, as online response modes (CAWI), are represented separately.

Other collection modes include audio-CASI (*audio computer-assisted self-interviewing*: the respondent operates alone on a computer where questions are displayed on a screen and also audio-broadcast) and its IVR telephone counterpart (*interactive voice response*).

1.1.3 Main multimode protocols

The collection modes may be combined in various ways; the main ones are:

- concurrent multimode: respondents may choose between various modes that are offered simultaneously;
- sequential multimode: respondents are offered a new mode in the event of non-response after a certain period of time has elapsed;
- deferred concurrent multimode: when a new mode is offered, respondents may still respond using the mode previously offered;
- disjoint multimode: modes are offered to separate sub-samples;
- integrated multimode: several modes are used to complete a questionnaire or during various waves of interviews.

The stages of the collection process are also punctuated by reminders, which may be of various types depending on the populations targeted, the mode and the available coordinates. A point requiring attention is not to give the impression of harassing respondents, in particular when another collection mode is liable to be offered subsequently.

1.1.4 What are the objectives?

Within the context of the total survey error paradigm (Andersen et al., 1979), using a multimode system may be envisaged in various dimensions: to improve coverage, to encourage participation in the survey and thus reduce total non-response error, and to reduce measurement error. The development of these systems is also motivated by cost-efficiency. Modes may, for example, be combined to improve the quality of the survey without adding to costs, or to reduce costs without a loss of quality (Schouten et al., 2021).

Therefore, the search for an appropriate combination of the various collection modes needs to integrate these various dimensions into the analysis. Typically, the sequential offering of various collection modes usually begins with the least expensive mode and ends with the most expensive mode.

However, the costs associated with surveys and their potential overhauls are not always easy to quantify, in particular where essential developments and those that may be extended to other surveys are concerned (Olson *et al.*, 2020), since these may be considered as long-term investments.

Mode effects

1.1.5 Definition

A mode effect can be defined as the effect a collection mode has on an assessment of the variable of interest studied. Thus, when considering the variable of interest y , the collection mode M , y^M the variable y observed via the mode M , and R^M a response indicator variable via the mode M , the (absolute) mode effect associated with collection mode M would be:

$$E(y^M | R^M = 1) - E(y) = E(y^M - y | R^M = 1) + E(y | R^M = 1) - E(y)$$

This formula for the mode effect indicates that this generic term refers to two distinct components (Tourangeau *et al.*, 2013): on the one hand, the question of knowing who is responding and who is not responding depending on the mode, and on the other hand how respondents respond, depending on the mode.

Typically, one of the multimode protocol's interests is in involving as many respondents as possible by offering or assigning them several possible response modes. The collection modes thus select respondents in various ways but this is not a problem in itself: in most cases this even constitutes a desirable effect of the collection protocol. On the other hand, the fact that people respond differently to the questionnaire according to the collection modes is an effect that may be undesirable, in particular in the analysis and interpretation of all aggregated responses.

1.1.6 Selection effect and measurement effect

We will therefore break down the generic mode effect into two components: on the one hand, the selection effect, and on the other, the measurement effect.

The selection effect is characterised by a different composition of respondents from that of the target population. This effect differs according to the collection modes: equipment, cognitive capacity and intrinsic preferences tend to select respondents differently. Induced biases may be controlled in a fairly conventional way if the selection is largely due to individuals' observable characteristics. The situation is more complicated when participation is linked to variables of interest, resulting in unobservable selection bias (Castell & Sillard, 2021).

The mode effect's other component is the effect strictly linked to measurement: do individuals respond in the same way, according to the various collection modes? This effect is the one to which we may refer, incorrectly, when we talk about mode effects. Indeed, this effect may be undesirable: a multimode protocol allows different fractions of the population to be reached and this differentiated selection is desirable, but this then raises the question of the effect intrinsically linked to measurement.

Mode effects strictly linked to measurement

1.1.7 Satisficing and social desirability

In measurement error literature for survey data, two types of bias are mainly mentioned: those that come under *satisficing*, and biases linked to social desirability.

The *satisficing* bias refers to low respondent involvement: they do not invest a great deal of effort in responding to a question but may be satisfied with giving an approximate response. The main forms this may take include partial non-response, order effects (*primacy* or *recency*), non-differentiation on response scales, default acquiescence, rounding of numbers, etc. The tendency towards *satisficing* depends on several factors (Krosnick, 1991): the difficulty of the question, the respondent's abilities (cognitive capacity and familiarity with the subject of the survey) and lastly their intrinsic or extrinsic motivation for responding. The main recommendations for limiting *satisficing* biases involve asking short questions, limiting the number of modalities, beginning the questionnaire with easy and interesting questions, etc.

The social desirability bias corresponds to the tendency amongst respondents to provide disingenuous responses that give a good image of themselves or that comply with normative expectations (Edwards, 1953). This mechanism may be exercised implicitly or as a result of a respondent's conscious desire to manipulate his/her image in the eyes of other people. There are various types of social desirability determinants: the subject of the survey, the questioning, the conditions of anonymity and the social distance between respondent and interviewer.

1.1.8 Questionnaires and multimode

A significant amount of literature addresses the issue of questionnaire *design* through the lens of multimode surveys (Dillman *et al.*, 2014; Dillman & Edwards, 2016; Tourangeau *et al.*, 2013; Tourangeau, 2017; Tourangeau *et al.*, 2017). The general idea is to maximise the quality of mode-specific data whilst minimising data differences between modes (de Leeuw, 2018). Hence there are two main approaches: on the one hand, the so-called unimode approach which insists on the minimisation of the differences between modes by proposing a unified stimulus according to the modes; on the other hand, the so-called best practice approach, which seeks above all to minimise error within each mode even if the stimuli differ according to the modes.

1.1.9 Assessment

Considering 2 collection modes A and B, the (relative) mode effect corresponds to the difference in the assessments of the variables of interest observed in mode A and B respondents.

$$\begin{aligned} & E(y^A | R^A = 1) - E(y^B | R^B = 1) \\ &= E(y^A - y^B | R^A = 1) + E(y^B | R^A = 1) - E(y^B | R^B = 1) \end{aligned}$$

The measurement effect is the difference, for the same individual, in their responses according to the collection modes and the selection effect is the difference in the composition of respondents for each of the modes. Therefore the measurement effect may also be interpreted as the overall mode effect (of one mode compared to another), purged of compositional differences between modes.

The difficulty lies in distinguishing between what comes under the measurement effect and what comes under the selection effect (Vannieuwenhuyze, 2014). More precisely, there is a risk that the quantification of the effect intrinsically linked to measurement could be biased due to insufficient or inadequate selection control. The issue may be addressed as part of an evaluation of the causal effect of processing, drawing on all the approaches and techniques described in this literature (Givord, 2014; Imbens & Rubin, 2015).

In particular, selection control requires a sufficient number of variables (unaffected by the mode) for the conditional independence hypothesis to be credible: subject to the control variables, responding in one mode is independent of the variable of interest. Amongst the control variables that explain the selection differences between modes and are correlated to variables of interest, we may wish to provide socio-demographic characteristics derived from the sampling frame or the questionnaire (questions that are not really subject to mode effects), geolocation variables, context information, paradata, etc. Selection control may also be partly guaranteed by a suitable protocol, moving closer to an experimental framework in order to control differences in unobservable characteristics more efficiently, in particular those associated with the motivation to respond to the survey, potentially correlated to variables of interest.

The various available techniques for assessing the mode effect intrinsically linked to measurement include regression models, assessment using *matching*, and reweighting methods. Some of these methods are based on the preliminary assessment of a propensity score (to respond using one mode rather than another). An important underlying hypothesis is that of the probabilistic mechanism: in constructing the counterfactual, it is assumed that all individuals have a non-zero probability of responding to each of the collection modes, which may raise so-called common support issues.

According to theoretical assumptions and the experimental processing analysis framework, the assessment of a residual mode effect, purged of observable selection differences, may be interpreted differently. In cases where selection effects due to differences in observable and unobservable characteristics appear to be controlled, residual effects may then well be considered as mode effects that are intrinsically linked to measurement. But when this is not the case, the distinction then cannot be made: a (strong) hypothesis, which considers the measurement effect to be negligible, would attribute the residual effect to unobservable selection differences.

1.1.10 Main findings

The traditional dichotomy contrasts self-administered modes with modes using the intermediary of an interviewer: the main differences between, on the one hand, modes using interviewers (face-to-face, telephone) and self-administered modes (Internet, paper) have thus been observed empirically.

Self-administered modes are better suited to sensitive questions and generate fewer social desirability biases than responses collected face-to-face or by telephone (Tourangeau *et al.*, 2013). Moreover, as regards purely factual/objective questions, there are no real differences between self-administered modes and modes using interviewers, but these are more marked regarding questions of opinion/subjective questions (Couper, 2011; Schork *et al.*, 2021).

Regarding *satisficing*, conclusions are more tentative: certain studies indicate that this trend is less significant where modes using the intermediary of an interviewer are concerned, others do not report marked differences. But the tendency is that the more complex the question (length, concepts, reasoning), the more liable it is to generate mode effects. For questions where a large number of modalities are proposed, the order effects are of a different nature: respondents have a tendency to choose one of the first modalities (*primacy effect*) if they are presented visually, and conversely to choose one of the last modalities (*recency effect*) if they are read out.

Between the two main modes that use the intermediary of an interviewer (CAPI and CATI), there are no major differences between face-to-face and telephone. Some of these may be directly related to how questionnaires are administered: when this is face-to-face, for questions with a long list of modalities it is sometimes customary to provide the list visually (code card), so as to reduce the cognitive effort required and limit order effects. Publications also do not indicate very marked differences between the two main self-administered modes (CAWI and PAPI). Here too, differences may be directly related to the conditions under which questionnaires are administered: although there is nothing to prevent individuals responding to a questionnaire in paper form from skipping a question, the pressure to respond (obligation, further prompting in the event of non-response to a question, etc.) may be stronger via Internet.

2. Insee's multimode victimisation survey experiments

Insee's development of the use of multimode in household surveys has become a necessary element of Insee's strategic priorities for 2025¹, as it has for those of the main foreign national statistics institutes. The initial reason for this is the decline in the participation rate for traditional face-to-face and telephone surveys. The second reason is to improve the quality of the data collected: a self-administered mode may be preferable for the collection of personal and sensitive information. The last reason is cost.

All household surveys are affected by the transition to multimode. The EVA, TIC and FPE surveys already rely on mixed modes including a web-based component, and the annual census survey (EAR) has provided an online version since 2015. Since the beginning of the year 2020, as part of a more general overhaul, the landmark Employment survey has partially switched to multimodal collection (collection via Internet for follow-up interviews). Indeed, by 2023, other surveys will have conducted experiments aimed at implementing the multimode switch (Housing survey, *Emploi du temps* (Time Use) survey, etc.).

This section presents a brief overview of the development of multimode and of experiments carried out within the public statistics department. It provides further details concerning the experimental “*Vol, violences et sécurité*” (“Theft, Violence and Security”, VVS) survey carried out in 2013, as well as the CVS “panel” and “multimode” experiments conducted in 2019, which enabled knowledge to be acquired regarding multimode in general, and the measurement of victimisation and perceptions concerning security in particular.

The development of the use of multimode and experiments within the public statistics department

From 2010 and prior to its development of the use of multimode in household surveys with Internet as the preferred collection mode, Insee launched a vast experiment operation. At the same time, the use of this new collection mode for the annual census survey (EAR) and business surveys had increased considerably.

The aim was to reduce costs and adapt to new behaviours and expectations. But, as mentioned in 1.1 and 1.2, methodological difficulties may arise: coverage, self-selection and selection bias, non-response, correction problems, *satisficing*, etc. (Razafindranovona, 2015). The findings reported in the literature are not always generally applicable, since the guidelines adopted by Insee in their development of multimode in household surveys consisted in focusing on defining, for each survey, the best possible use of Internet.

2.1.1 Multimode experiments at Insee

An initial multimode-type experiment, which included collection via Internet, was carried out in 2010 and focused on the theme of housing (Amiel & Denoyelle, 2012). Then, from 2013, several multimode collection experiments were conducted: *Qualité de vie au travail* (Quality of Life at Work) (2013), Theft, Violence and Security (2013), *Patrimoine* (Heritage) (2015), *Logement* (Housing) web/paper (2014), the Muse project (Multimode on Employment) (2014-2018), and the Mélopée project (*Mise En Ligne et OPTimisation de l'Enquête Emploi*, Online Publication and Optimisation of the Employment Survey) (2017-2021) for the design and deployment of the multimode Employment survey from January 2021, etc.

Interpretations of these initial experiments revealed significant mode effects and rates of return that were clearly lower than those obtained via face-to-face collection (de Peretti & Razafindranovona., 2014; Razafindranovona, 2016a, Razafindranovona, 2016b, Razafindranovona, 2017). Selection bias in observable variables may be corrected using traditional non-response and/or calibration correction

¹<https://www.insee.fr/fr/statistiques/fichier/4130132/INSEE-2025.pdf>

methods. Despite such corrections, in most cases there is still a “residual” mode effect, incorporating a measurement effect and a selection bias in non-observable variables. It is therefore essential to implement techniques for detecting this type of bias and “pure” measurement effects.

Initial insights gained from the experiment plan also highlight:

- the need to improve response rates by multiplying the vectors for contacting respondents (significance of the contact phase) and using adaptive design to limit costs;
- the importance of improving the design of questionnaires by limiting their complexity and their duration, standardising the ergonomics of web questionnaires, creating a single and attractive household survey portal and providing control variables in order to improve the correction of selection effects;
- the need to develop standard data aggregation methods linked to the protocol in place and the pooling of IT tools and infrastructures between surveys.

2.1.2 Promising work with the pooling of methods and tools

In 2018, at the *10^e Colloque francophone sur les sondages* (10th Francophone Symposium on Surveys) during the *Journées de méthodologie statistiques* (Statistical Methodology Days, JMS), an entire session devoted to multimode was organised, reflecting the enormous challenge this subject represents for public statistics. These sessions provided an opportunity to discuss the progress, practices and findings of French and foreign teams working on multimode collection protocols.

In particular, it was shown that measurement effects appeared to be less marked in more objective and factual data, but that they may be very clear in more sensitive and subjective questions (Neumayr et al., 2018 on data relating to job satisfaction). One of CEREQ's (*Centre d'études et de recherches sur les qualifications*, French Centre for Studies and Research on Qualifications) presentations during the JMS also demonstrated, using experiments from the *Génération* survey, that the mode effect between telephone and Internet concerned only a minority of the tested variables (more precisely, opinion variables: Barret & Cissé, 2018). These findings were similar to those already described in international scientific literature. They provided reassurance regarding the relative robustness of web surveys, at the same time confirming the need to use appropriate methods to seek to correct variables subject to mode effects. When measurement effects are significant within a multimode survey, when compared to a traditional single-mode survey, several actions are possible: including as calibration margins the desired variable of interest total (using a reference mode); imputing the responses gathered into one of the samples. In real data, the Living Environment and Security (CVS) survey and its Internet version, the Theft, Violence and Security (VVS) survey, both conducted in 2013, constitute a simple method for implementing the second strategy by proposing a propensity and imputation score-matching method (Legleye et al., 2018). The various communications also highlighted the limitations of telephone surveys. Although, since the 1990s, the latter constituted an interesting alternative to face-to-face surveys, these are now contending with the low telephone reachability of households and decreasing levels of participation; this is encouraging telephone survey designers to explore the possibilities of multimode collection (Senaux, 2018).

As has been done within other national statistics institutes, Insee has developed a programme for pooling the developments required for the design of collection media. There are tools for designing and generating questionnaires, and an Internet collection platform, to address the requirements of surveys undertaken. To address the requirements of household surveys, the *Métallica* (*METadonnées Actives, Logiciels Libres et Infrastructures pour une Collecte Assistée*, Active Metadata, Free Software and Infrastructures for Assisted Collection) programme was launched, with the aim of optimising the survey production process and in particular questionnaire design. In particular it aims to complete the application components required for developing the multimode collection of household surveys. The TIC (IT and Communications) web/paper and FPE surveys have been integrated into the *Métallica* network. The TIC survey with its telephone component is to be integrated in 2022 and the Camme survey is to be tested at the end of 2022 then integrated into T1 in 2023. Other household surveys have been handled on a case-by-case basis, in considering the switch to multimode: the *Loyers et charges* (Rents

and Charges) survey then the Housing surveys. For the Employment, SRCV (Statistics on Income and Living Conditions) and *Histoire de vie* (Life History) / Heritage, Time Use, *Conditions de travail / risques psychosociaux* (Working Conditions) / Psycho-social Risks) surveys which are based on complex questionnaires: although some features have yet to be developed, the integration of most of the surveys mentioned is already envisaged.

Lastly, working groups have been established within Insee since 2017, to offer practical solutions for the design of collection media, adapted to online and self-administered modes of procedure, and for the measurement of mode effect adjustments. Ultimately, a newsletter provides those involved in the public statistics department with information about the latest developments in multimode collection.

The 2013 “Theft, Violence and Security” experiment

2.1.3 Presentation

Context & objectives

Within the context mentioned in the introduction of the development of multimode in household surveys, and alongside the face-to-face CVS survey, in 2013 Insee conducted the methodological Internet/paper “Theft, Violence and Security” (VVS) survey in order to measure the findings’ responsivity to the collection mode. It was presented on 4 July 2012 to the Quality Label Committee which issued a notice of review.

Time frame

The VVS collection took place between 14 January 2013 and 16 March 2013.

Methods

The VVS survey covers some of the CVS themes by largely focusing on the same questions: burglaries, vehicle theft, personal theft, physical violence, threats, opinions on living environment and security. This is therefore a short version of the CVS 2013 questionnaire transposed onto Internet. A 6-page paper version has also been designed: shorter than the Internet version, it excludes questions on sexual violence experienced outside the household. In both Internet and paper questionnaires, for reasons of confidentiality at the time, VVS did not include violence within the household (questions on violence posed via headphones for CVS).

In view of the anticipated low response rates (around 30 %) and the prevalence observed in the victimisation surveys, a sample of 40,000 individuals were drawn from the “FIP 2011” tax files available in 2012 (resulting from declarations in 2011 of income during 2010 and *taxe d’habitation* (Council Tax)).

In this type of self-administered protocol there is always a risk that a different individual to the person selected within a household might respond to the survey. In particular, given the theme, a different person to the person selected may be tempted to respond if, unlike the person selected, he/she has experienced theft or violence. Such situations constitute risks of bias that are liable to increase the levels of victimisation, and hard to identify. In order to study the self-selection of individuals within their households, the overall sample was divided into 3 according to the following protocol:

- in the 1st sub-sample composed of individuals aged 20+, a household questionnaire was offered at the start of the interviews (table of occupants including year of birth, sex, the identification of four incidents of victimisation: personal theft with violence, personal theft without violence, physical violence, threats). Since the introduction to these questions concerned all the members of the household, it was expected to maximise the chances of the targeted individual responding to the extended questionnaire, since the other potential victims in the household were given the opportunity to state within the table that they had been victims;

- in the 2nd sub-sample, two sets of logins and passwords were sent to the selected individuals (also aged 20+), thus offering another potentially more “concerned” member of the household the opportunity to respond to the questionnaire;
- the 3rd sub-sample was composed of young people aged 14 to 19 whose identity (surname/first name) was not noted in the sampling frame.

The contact protocol established with the respondents was intended to promote responses via Internet whilst allowing paper responses, so as to increase the overall rate of response. Accordingly, two reminders were sent roughly four weeks apart:

- the paper questionnaire was attached to the first reminder, along with a pre-paid return envelope, whilst the option to respond via Internet remained available;
- the second reminder simply consisted of a letter.

2.1.4 Main findings

The anticipated VVS findings were as follows:

- measurement of the response rate according to the mode (internet, paper) for the three sub-samples;
- respondent representativity and characteristics;
- the findings’ susceptibility to the collection mode;
- initial feedback from a victimisation survey via Internet (modular survey, questions concerning sexual violence outside the household)

Out of the 40,000 notification letters sent, 4,500 did not reach their intended recipients. Out of the 35,500 individuals who received their letters, 14,500 responded: 53 % on paper and 47 % via Internet. Approximately 1,500 questionnaires were excluded, either because the questionnaires were only partly completed, or because the respondent was not the individual selected. In total, around 13,000 usable questionnaires were obtained, i.e. a response rate of 32 % (63 % in CVS 2013). No significant difference between the two first sub-samples was revealed to argue in favour of employing one protocol rather than another to limit self-selection. The following findings aggregate all the respondents.

Compared to the initial sample’s characteristics, the VVS Internet respondents were, as in many other experiments carried out at the time, much younger (average age 44 vs 48), wealthier (average annual household income €37 K vs €28 K; [figure 2.1](#)). Those who responded on paper were, by contrast, older (58). Their average annual income was close to the level observed in the sample (€29 K) and their level of education was inferior to that of the CAWI respondents (22 % were university graduates vs 45 %).

Figure 2.1: Characteristics of VVS respondents according to response mode

	Sample	VVS Internet Respondents	VVS Paper Respondents
Average age (years)	48	44	58
Average annual household income (€K)	28	37	29
University graduates (%)	-	45	22

Source: VVS; Razafindranovona, 2016a

Coverage: individuals aged 14+ living in ordinary housing in metropolitan France

Interpretation: 45 % of VVS Internet respondents are graduates

Concerning the findings’ susceptibility to the collection mode, respondents to the self-administered VVS questionnaire (internet/paper) were more likely to report that they had been victimised than in CVS, despite a number of adjustments (or calibrations) taking into account the survey’s variables of interest. Four successive calibrations were implemented, all with a minimum of non-response correction (CNR): i/ calibration on socio-demographic margins (CMS); ii/ CMS and taking complaint rates into account

(CMS + TP); iii/ CMS + TP and taking sense of insecurity into account (CMS + TP + SI); iv/ CMS + TP + SI and taking Internet access into account (figure 2.2).

Figure 2.2: VVS victimisation rates following successive calibrations

	Theft with violence* (%)	Theft without violence* (%)	Physical Violence (%)	Threats (%)
After non-response correction (CNR)	2.6	6.4	2.7	7.3
Calibrations on socio-demographic margins	2.9	6.4	2.9	7.7
+ complaint rates	2.8	5.9	2.7	7.2
+ complaint rates + sense of insecurity	1.9	4.4	1.8	4.9
+ complaint rates + sense of insecurity + Internet access	1.8	4.2	1.7	4.8
CVS 2013	1.0	2.8	2.2	4.5

Source: VVS; Razafindranovona, 2016a

Coverage: individuals aged 14+ living in ordinary housing in metropolitan France.

Interpretation: following calibration on socio-demographic margins, the VVS level of victimisation involving theft without violence is 6.4 %.

Despite such processing, VVS prevalence assessments remained higher overall than those of CVS. To test the hypothesis of an uncontrolled selection effect associated with unobservable variables, a comparison was made between the victimisation rates of Internet respondents calculated before and after the first reminder. The rates before the reminder proved to be consistently higher than the rates after the reminder, which confirmed the idea that individuals motivated by a subject that concerns them are more inclined to respond to a survey. Logistic regressions confirmed the relationship between timeliness of response and victimisation by controlling socio-demographic characteristics. The questionnaire's validation date effect proved to be significant for the 4 types of victimisation tested (figure 2.3).

Figure 2.3: Monitoring of respondents' motivation effect on victimisation rates

	Victimisation Rates			
	Theft with violence * (%)	Theft without violence * (%)	Physical Violence (%)	Threats (%)
Before reminder (in %)	2.6	8.2	2.8	10.3
After reminder (in %)	2.4	6.4	2.7	7.2
Odds-ratios** "Before reminder" vs "After reminder"	0.80	0.78	0.92	0.65

** control socio-demographic variables: sex, age, marital status, qualifications, employment situation, type of urban unit.

Source: VVS; Razafindranovona, 2016a

Coverage: individuals aged 14+ living in ordinary housing in metropolitan France who responded via Internet.

Interpretation: of the individuals who responded via Internet to the VVS survey, 10.3 % of the individuals who responded before the reminder stated that they had been the victims of threats, as opposed to 7.2 % of the individuals who responded via Internet after the reminder.

2.1.5 Conclusions

An analysis of the findings of the VVS survey showed that, at the time, convergence between the two surveys was insufficient when it came to envisaging simply switching the CVS protocol for a solely self-administered protocol such as that of the VVS, without any specific adjustments; due, on the one hand, to the rather significant differences in the correlations observed between victimisations and socio-

demographic criteria, and on the other, to divergences regarding incidents of victimisation associated with theft, which attested to a possible selection effect through unobservable variables.

At the time, the experiment resulted in the conclusion that the integration of self-administered modes such as Internet required significant adjustments to the face-to-face questionnaire, in order to clarify the concepts and guarantee, as far as possible, that the respondents' relationship to the questionnaire remained the same. A discussion concerning modes of recruitment and the invitation to participate was also considered necessary. The risk of self-selection amongst respondents, determined by how closely the survey's topic concerns them, is indeed high when a survey's subject is sensitive. However this risk, which is known and affects all surveys, appeared to be particularly high in cases of mainly self-administered protocols with significantly lower collection rates than surveys where interviewers are involved.

The 2019 CVS "Multimode Panel" experiment

2.1.6 Presentation

Context

Since 2015 and following various sessions devoted to the CVS survey, in order to discuss, in particular, the possibility of an annual victimisation survey, the CNIS issued several recommendations including exploring options for the development of the collection system through dissociating short-term and structural information requirements and re-exploring the use of multimode. The project managers opened a debate concerning the collection system and its sampling plan, in association with the survey's consultation and operations group, and Insee's Collection and Information Processing division and Sampling division. To meet the need for short-term information that had been confirmed during the previous debate organised by the CNIS in September 2016, it was evident that the accuracy of certain indicators' measurements needed to be increased. A possible solution was the survey's – at least partial – panelisation. To test this solution, the project managers proposed introducing two experimental protocols alongside the "traditional" CVS survey: the "Panel in SRCV (Statistics on Income and Living Conditions)" experiment (cf. 2.4) and the "CVS Multimode Panel" experiment.

Objectives

The "CVS Multimode Panel" experiment consisted in re-interviewing, via Internet and telephone in 2019, those who had responded to the CVS survey in 2018, using a simplified questionnaire. Three objectives were pursued: i/ assessing the potential for greater accuracy associated with panelisation; ii/ measuring the collection mode effects; iii/ determining which multimode protocol would be the most productive for this type of survey. The project was examined in session by the Quality Label Committee on 4 April 2018.

Time frame

The "Multimode Panel" review experiment was conducted between 15 February and 30 April 2019 (i.e. 15 days later than CVS 2019).

Methods

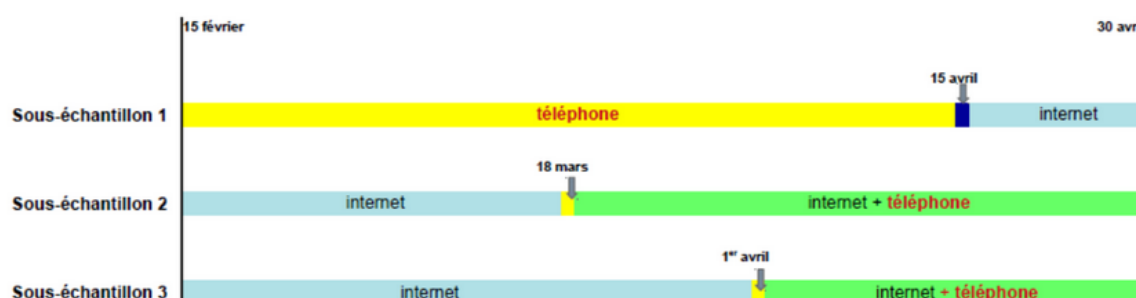
The questionnaire used in the experiment was a simplified version of CVS: essentially the reference period was restricted to the previous calendar year (N-1) instead of the two previous years (N-1 / N-2); almost all the open fields were deleted, with the exception of gathering clear descriptions of each offence; penultimate victimisation incidents were no longer gathered; filter blocks placed on victims were simplified; incidents of violence within households were no longer gathered.

The experiment concerned all the individuals selected for individual questionnaires in 2018, for whom a household questionnaire and an individual questionnaire had been completed (excluding QPV (urban policy priority district) over-samples), i.e. 13,123 “Panel” individuals aged 14+.

The sample re-interviewed was broken down into three sub-samples of 4,374 individuals subject to various protocols (figure 2.4):

- sub-sample 1 was interviewed by telephone only in the initial collection phase (2 months of collection). There was a complete switch to Internet mode for this sub-sample’s non-respondents 15 days before collection ended;
- sub-sample 2 was interviewed by Internet then telephone, with a “long” telephone phase. This sample’s individuals initially received a follow-up email, then a letter, then an email. A switch to telephone mode was made when the 4th reminder was sent (after one month), alerting respondents that an interviewer would be contacting him/her. As a result of this switch, the Internet mode was retained. The two modes coexisted for a month and a half.
- sub-sample 3 was interviewed by Internet then telephone, with a “short” telephone phase. The schedule for sub-sample 3’s reminders was similar to that of sub-sample 2 but the reminders were delayed by one week. The switch to telephone was later (15 days after sub-sample 2). The two modes coexisted for a month.

Figure 2.4: Protocol used in the 2019 CVS “Multimode and Panel” experiment



Source: Castell et al., 2021

Individuals were assigned to one or other of the sub-samples by random selection with unequal probabilities based on contact details supplied in 2018 (figure 2.5). Sub-sample 3 was composed of a greater proportion of individuals without contact details.

Figure 2.5: Assignment probabilities in the 2019 “Multimode Panel” experiment’s 3 sub-samples, based on contact details furnished in 2018

	Sub-sample 1	Sub-sample 2	Sub-sample 3
Telephone only	0.56	0.24	0.20
Email only	0.18	0.33	0.49
Telephone + email	0.18	0.44	0.38
No contact details	0.10	0.30	0.60

Source: 2019 CVS “Multimode Panel” experiment; Castell et al., 2021

Coverage: individuals aged 14+ living in ordinary housing in metropolitan France

Interpretation: the probability of being assigned to sub-sample 3 for panel individuals who had only supplied an email address was equivalent to 0.49.

2.1.7 Main findings

The findings presented within this section were communicated during the 11th Francophone Symposium on Surveys (Castell *et al.*, 2021). The information presented under this heading cites this study’s findings.

In total, 10,271 individuals responded to the survey: 5,490 by telephone and 4,781 via Internet, i.e. an overall collection rate of 78.3 %. Unsurprisingly, in sub-sample 1, the majority of responses were by telephone (91.1 %), whilst in sub-samples 2 and 3 responses were predominantly via Internet (62.8 % and 67.8 % respectively, [figure 2.6](#)).

Figure 2.6: Collection rate according to response mode

	Sub-sample 1	Sub-sample 2	Sub-sample 3	Together
Collection rate (in % of the sample)	77.7	80.7	76.4	78.3
By Telephone (in % of respondents)	91.1	37.2	32.2	53.4
Via Internet (in % of respondents)	8.9	62.8	67.8	46.5

Source: 2019 CVS “Multimode Panel” experiment; Castell et al., 2021

Coverage: individuals aged 14+ living in ordinary housing in metropolitan France

Interpretation: in sub-sample 1, 8.9 % of respondents responded via Internet.

The 3 sub-samples’ overall collection rates (without specifying the mode) are very close, depending on the type of contact details furnished in 2018, except amongst individuals who provided no contact details ([figure 2.7](#)).

Figure 2.7: Sub-sample collection rates according to available contact details (in %)

	Sub-sample 1	Sub-sample 2	Sub-sample 3	Together
Telephone + email	86.3	86.6	87.1	86.7
Telephone only	76.6	78.9	76.8	77.2
Email only	74.0	76.0	74.7	75.0
No contact details	47.4	57.8	54.0	54.5
With sub-sample 1's contact details frame	77.7	79.8	79.9	-
Overall	77.7	80.7	76.4	78.3

Source: 2019 CVS "Multimode Panel" experiment; Castell et al., 2021

Coverage: individuals aged 14+ living in ordinary housing in metropolitan France

Interpretation: the rate of collection reached 87.1 % within sub-sample 3, amongst respondents who could be contacted by telephone and email.

In total, almost 300 questionnaires were excluded because they had not been completed correctly (2.3 % of the questionnaires collected). Moreover, the self-administration of the CVS questionnaire, which was designed for face-to-face, required a number of post-collection reworkings in connection with double-counting and the misclassification of offences. Almost 180 incidents of double-counting in victimisation variables were detected and processed (81 % were found in Internet questionnaires) and almost 300 descriptions of offences were deemed dubious, and 200 corrected (79 % were found in Internet questionnaires). Such reworkings had an impact on measured victimisation rates, in particular vandalism (to homes and cars), threats and abuse. Internet responses were more concerned by such downstream modifications than telephone responses.

Concerning the representativity of respondent samples, a comparison was made between the 3 protocols, using two separate measurements: standardised bivariate differences and an overall multivariate indicator, R. The following control variables were introduced: type of housing, urban unit band, type of household, sex, age, qualifications, nationality, work status, socio-professional category, type of contact details available, and agreement to refer back to previous variables). The findings of such calculations indicate that Internet adds no real gains in representativity when it is used as a follow-up to the telephone (sub-sample 1). Conversely, using the telephone as follow-up for those who had not responded via Internet significantly improved the representativity of respondents (sub-samples 2 and 3), especially in sub-sample 3 in which telephone follow-ups were longer.

The raw analysis of victimisation rates shows a significant over-reporting via Internet of a number of victimisation variables compared to the telephone, even after post-collection clearance. To compare the mode of response effect on the victimisation rates, a propensity score analysis was carried out. Measurement biases were assessed using weighted bivariate logistic regressions comparing, on the one hand, Internet respondents with telephone respondents within sub-samples 2 and 3, and on the other, sub-samples 2 and 3's Internet respondents before the switch to telephone with sub-sample 1's telephone respondents. Control variables were successively added within the regressions: i/ socio-demographic variables (those taken into account within the propensity score); ii/ incidents of delinquency recorded at IRIS (French Institute of International and Strategic Relations); iii/ data on victimisation, opinions and perceptions concerning security gathered during CVS 2018. The addition of control variables was found to modify bivariate differences very marginally. Lastly, regardless of which respondent populations were compared, the conclusions were identical. The most significant biases were measured in burglaries and theft outside main residences, acts of vandalism (to homes and cars) and threats.

2.1.8 Conclusions

The “CVS multimode panel” experiment enabled several interesting conclusions to be established for the future:

- the findings highlighted the importance of designing questionnaires suitable for self-administration in advance, in order to minimise misclassification and double counting which require time-consuming and costly post-collection processing which is likely to be insufficient due a lack of available supporting information;
- the telephone appeared to be an essential alternative mode to supplement Internet in improving the representativity of respondents;
- the raw victimisation rates are systematically higher on Internet. The analyses introduced suggest that the measurement bias is generally more significant for incidents of “household” victimisation than for incidents of “individual” victimisation;

Preliminary findings concerning the use of panelisation seem to indicate that its interest has not been proven: very little selection in the variables gathered in initial interviews, and in the experiment the attrition of the rare profiles of victims of sensitive incidents of violence was higher than for the other profiles.

The experiment did not add anything to the conclusions of VVS concerning the mode effect in incidences of sexual violence. A number of interviewers did indeed decide not to pose such questions by telephone, which they had been authorised to do. Unfortunately they did not systematically, where relevant, correctly enter the variable for distinguishing cases where the question was not posed from cases where respondents had not experienced sexual violence, thereby compromising the interpretation of this partially-gathered information.

The Panel in SRCV experiment

2.1.9 Presentation

Context & objectives

Within the context mentioned in 2.3.1, the “Panel in SRCV” experiment’s objectives were as follows:

- to compare the levels of victimisation rates, sense of insecurity and the most worrying issues in French society measured in CVS and SRCV;
- to compare these indicators’ annual developments:
- to analyse the contribution of a panel review on the accuracy of the measurement of victimisation indicators.

The introduction of fifteen victimisation questions into the Statistical Panel on Resources and Living conditions (SRCV) from 2018 was accepted by the SRCV’s steering committee and the choice of questions approved by the survey’s operation and consultation group. The principle of this experiment had been validated by the Quality Label Committee during the 5 July 2017 session.

Time frame

The victimisation questions were posed in SRCV in 2018, in 2019 and were repeated in 2020. However in 2020, due to the health crisis, the victimisation questions were abandoned just when the administration of SRCV switched to the telephone and therefore were only posed to one part of the sample (around 7,000 address files).

Methods

The fifteen questions introduced into SRCV covered the main incidents of victimisation measured in CVS over the reference period N-1, excluding “sensitive” incidences of violence (sexual violence and incidents of violence within the household): 7 “household” level questions (burglary and attempted burglary, vehicle theft, and vandalism) and 8 “individual” level questions (personal theft and attempted

personal theft, with or without violence, physical violence, threats, abuse, sense of insecurity and security concerns). No adjustments to the SRCV protocol were introduced: it was administered face-to-face to all the individuals aged 16+ in the selected households. The proxy generally authorised in SRCV was not permitted for the questions taken from CVS.

2.1.10 Main findings

Initial feedback from the 2018 collection was communicated to the project managers in spring 2019. More advanced analyses were presented during one of the Victimisation survey overhaul project's "Survey Methodology" working group's sessions in January 2021. The findings presented below have not yet been published or communicated.

Initial analyses showed higher levels of victimisation in SRCV, variations in magnitude that are variable but always present at both individual and household levels. Trends in the victimisation indicators measured in the two surveys show significant variations at individual level, except for threats and abuse, and completely opposite trends for theft without violence and physical violence. The sense of insecurity was at a similar level but with differentiated trends (figure 2.8).

Figure 2.8: Levels and trends for incidents of victimisation and sense of insecurity – comparisons between CVS and SRCV

	2018		2019		Trend 2018-2019	
	CVS	SRCV	CVS	SRCV	CVS	SRCV
Burglary*	2.0	2.4	1.7	1.8	-16%	-27%
Theft without breaking and entering*	0.9	1.4	0.9	1.0	-4%	-29%
Car theft*	1.0	1.8	0.8	1.3	-17%	-28%
Theft of items in/on cars*	2.5	3.8	2.4	3.1	-4%	-18%
Theft of two-wheeled motor vehicles*	2.0	2.8	1.3	5.0	-37%	77%
Damage to homes*	2.2	2.9	2.0	2.4	-9%	-16%
Damage to cars*	5.2	7.0	4.5	5.9	-13%	-15%
Theft with physical violence or threats**	0.4	0.8	0.3	0.7	-24%	-14%
Threats without violence**	2.0	2.2	1.8	2.5	-13%	13%
Physical violence**	1.2	1.7	1.3	1.4	14%	-14%
Threats**	3.7	4.9	3.4	4.4	-9%	-10%
Abuse**	10.1	12.6	9.2	11.5	-9%	-9%
Insecurity in the home (frequent)**	2	1	2	1	-24%	-14%
Insecurity in the home (from time to time)**	6	6	6	5	-7%	-14%
Insecurity in the district or village (frequent)**	3	2	2	2	-23%	4%
Insecurity in the district or village (from time to time)**	10	12	10	10	0%	-15%

* In % of households ; ** in % of individuals in the SRCV age range

Source: SRCV 2018-2019 and CVS 2018-2019; Insee interpretations

Concerning the worrying nature of the problems facing society, the items proposed in the list are generally evaluated in the same way but nevertheless with a few differences in hierarchy and level of preoccupation, and their trends.

2.1.11 Conclusions

Several factors may help to explain the observed variations: the protocol, the questionnaires and the survey's methodology in general are different, in particular a proxy is authorised in SRCV whilst it is not permitted in CVS; this led to significant "collection gaps" amongst the younger members of society. Further interviews will be required in order to improve our understanding of the discrepancies. Nevertheless, the "Panel in SRCV" experiment's initial findings suggest that a system organised into two separate collections, with on the one hand an annual monitoring of indicators for victimisation and perceptions concerning security within an existing "omnibus" system (such as SRCV), and on the other an in-depth thematic survey at more frequent intervals, should not be proposed.

3. The Genese survey: the full-scale multimode experiment

The *Genre et Sécurité* (Gender and Security, Genese) survey carried out by the SSMSI in 2021 constituted a real full-scale multimode victimisation survey experiment (see Appendix II for a full presentation of the Genese survey). This system was designed in advance in order to sustain discussions on the future of victimisation survey systems and provide answers to the main methodological questions. In particular, Genese's protocol offers an appropriate framework for analysing mode effects, and more precisely quantifying effects intrinsically linked to measurement for questions concerning victimisation, and perceptions and opinions concerning security.

The framework for analysis: data collected in phase 1 of the Genese survey

The Genese survey can be broken down into two phases. The initial phase is a screening survey for collecting socio-demographic characteristics, identifying offences experienced and questioning perceptions and opinions concerning security. The second phase is a screened survey for re-interviewing phase 1 respondents, over-representing amongst them respondents who had stated that they had experienced gender-based victimisation. The analysis of mode effects presented in this part relates exclusively to questions collected during the phase 1 filter survey.

Phase 1 of the survey covers 169,060 individuals from ordinary households in metropolitan France aged 18 to 74 on 1st January 2021. Its protocol is deferred competitive multimode with Internet (CAWI) as the initial proposed response mode, then paper or telephone (CATI), according to the contact details available. More precisely, after 40 days of collection, 25,000 individuals randomly selected from amongst the 80,000 non-responding units for whom a telephone number was available in the sampling frame, or obtained after enrichment by the provider, were entered into the telephone follow-up process.

At the end of the phase 1 collection, in addition to the standard work of clearing and validating questionnaires, specific processing was carried out in advance of the analysis of mode effects. The idea was to eliminate questionnaires that were too incomplete (in variables of interest and socio-demographic characteristics) so that differentiated behaviour in terms of partial non-response did not interfere with an assessment of mode effects. Finally, 96,600 questionnaires were retained, of which 79,300 were collected via Internet, 5,300 by telephone and 12,000 in paper form.

Characterisation of respondent profiles by mode and by eligibility profile for telephone follow-up

The Genese survey's protocol, deferred competitive multimode, is differentiated according to the information available for each of the respondents and according to their response status during the collection. Therefore, not all respondents will be contacted by telephone whereas, on the other hand, all respondents are initially invited to respond via Internet. Hence, differences observed in respondents' profiles according to modes may be partly due to differentiation within the protocol. This initial overall characterisation according to response mode will then be supplemented by an analysis of the choice of collection mode determinants.

3.1.1 Characterisation by mode

The characterisation of respondents according to response mode will first be presented on the basis of auxiliary variables generated by the sampling frame, then according to the socio-demographic variables collected during the phase 1 questioning. The frequencies presented here are weighted according to individuals' initial draw weight in the Fideli sampling frame.

3.1.1.1 Auxiliary variables generated by the sampling frame

Characterising respondents according to the variables generated by the sampling frame makes it possible to get rid of any possible mode effects. Although such information may no longer be completely up to date at micro level due to a collection time lag, they are robust enough for general observations.

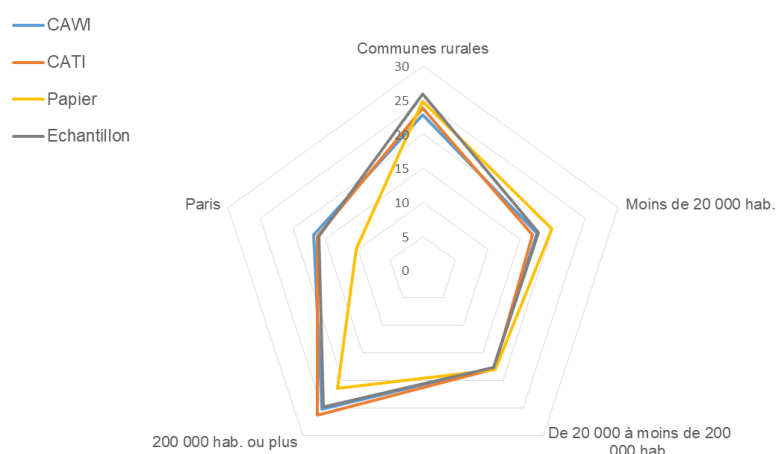
The findings were fairly consistent with the literature. Accordingly, Internet respondents were rather more affluent and had a relative tendency to live in the urban unit of Paris, compared to the other respondents. Telephone respondents differed by having a greater tendency, compared to the others, to belong to the lowest living standard deciles (figure 3.3); moreover, they tended to live in *quartiers de reconquête républicaine* (Republican Re-conquest Districts, QRR) and (Urban Priority Districts, QPV). Lastly, those who responded on paper were older (figure 3.1) and frequently lived in rural *communes* (municipalities) (figure 3.2).

Figure 3.1: Age of respondents according to response mode

	Minimum	1st quartile	Median	Average	3rd quartile	Maximum
CAWI	18	34	47	47	59	74
CATI	18	33	47	47	61	74
Paper	18	43	56	53	66	74
Sample	18	33	47	47	60	74

Source: Genese survey (N=96 600); SSMSI processing

Figure 3.2: Distribution by respondents' size of urban unit of residence according to response mode



Source: Genese survey (N=96 600); SSMSI processing

Note: the notion of urban units is based on the continuity of the buildings and the number of residents. Urban units are built in metropolitan France and in its overseas territories (DOM) according to the following definition: a municipality or group of municipalities with a continuous built-up area (no breaks of more than 200 metres between two constructions) and at least 2,000 residents.

Interpretation: 10.2 % of those who responded on paper live in the urban unit of Paris, as against 16.1 % in the sample.

Figure 3.3: Distribution of respondents by living standard decile, according to response mode

	CAWI	CATI	Paper	Sample
<1st decile	5.8	8.9	7.1	8.9
Between the 1st and 2nd deciles	6.3	9.6	8.7	8.7
Between the 2nd and 3rd deciles	7.0	10.0	9.9	8.7
Between the 3rd and 4th deciles	7.8	10.5	10.6	8.8
Between the 4th and 5th deciles	9.0	10.8	11.7	9.2
Between the 5th and 6th deciles	10.0	10.7	10.9	9.5
Between the 6th and 7th deciles	11.6	10.0	11.6	10.1
Between the 7th and 8th deciles	12.6	8.8	10.5	10.2
Between the 8th and 9th deciles	13.4	9.0	9.2	10.3
Between the 9th and 10th deciles	14.1	8.0	7.9	10.5
Unspecified	2.6	3.6	1.9	5.1

Source: Genese survey (N=96 600); SSMSI processing

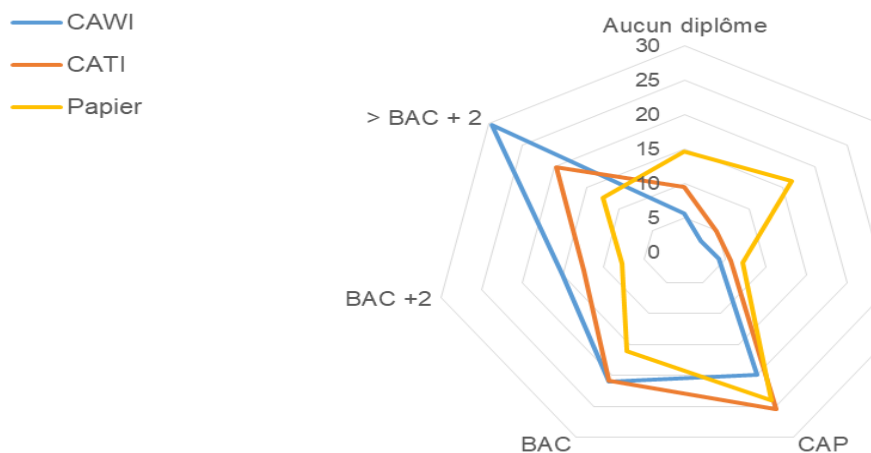
Note: a living standard is equal to a household's disposable income divided by the number of consumption units (uc). The living standard is therefore the same for all the individuals within the same household. Consumption units are generally calculated according to the so-called amended OCDE (French Organisation for Economic Cooperation and Development) equivalence scale which attributes 1 uc to the first adult in a household, 0.5 uc to the other individuals aged 14 + and 0.3 uc to children under 14." (Insee website) "The concept [...] of a living standard used here is to be used with precaution since it only takes monetary components into account and excludes benefits in kind." (Filosofi documentation)

Interpretation: 14.1 % of CAWI respondents (Internet) had living standards between the 9th and 10th deciles, as against 10.5 % in the sample.

3.1.1.2 Socio-demographic variables generated by the questionnaire

In general, respondents' profiles are heavily structured according to qualification levels (figure 3.4): Internet respondents are the most highly qualified (30 % have higher qualifications than *bac* +2 (2 years of higher education), as against 20 % for telephone respondents, and 13 % for those who responded on paper), whereas those who responded on paper are more likely to have no qualifications (15 % as against 6 % for Internet and 9 % for telephone). Differences also appear in the types of household: proportionately there are more single individuals amongst the telephone respondents (37 % as against 32 % for Internet and 23 % for paper). Finally, it should be noted that telephone respondents tend to be foreign nationals (7 % as against 5 % for the other modes).

Figure 3.4: Distribution of respondents by level of qualification according to response mode



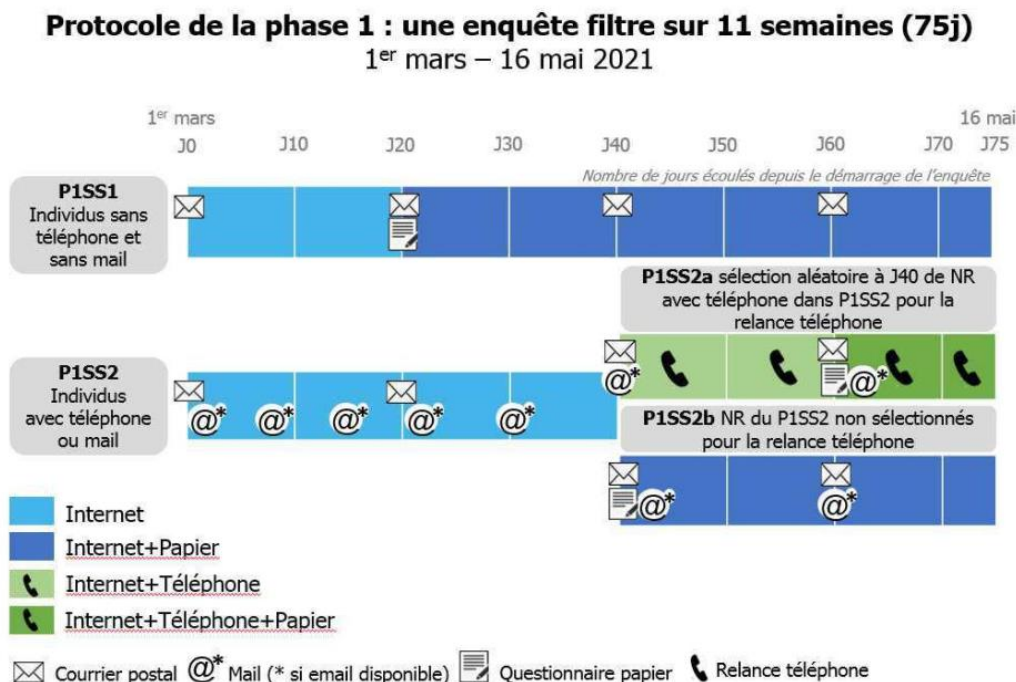
Source: Genese survey (N=96 600); SSMSI processing

Interpretation: 30 % of CAWI respondents (Internet) have a level of qualification higher than bac +2.

3.1.2 Determining factors in the choice between responding via the Internet and responding by telephone

Although the phase 1 survey protocol offers three collection modes overall, the way in which modes are proposed depends upon response status, availability of telephone contact information and telephone follow-up selection (figures 3.5 et 3.6). Thus, the great majority of respondents responded via Internet less than 40 days after the beginning of collection (68 %). The CAWI respondents who responded after Day 40, who were eligible and selected for telephone follow-up, constitute a significant proportion of the total sample (5 %). Since this sub-group was distinguished by its theoretical ability to choose between response modes, the determining factors for this choice (between Internet and telephone) were studied: the propensity to respond via Internet was greater amongst the most affluent, and amongst young women and students (figure 3.7). Also, it appears that individuals who did not state that they had moved house within the last 12 months had a much more significant relative risk ratio of responding via Internet.

Figure 3.5: Genese survey phase 1 protocol



Source: SSMSI

Figure 3.6: Respondent group numbers

	Number	(%)
CAWI before Day 40	65,738	68.1
CAWI after Day 40, eligible and selected	4,805	5.0
CAWI after Day 40, eligible and non-selected	7,259	7.5
CAWI after Day 40, non-eligible	1,381	1.4
Telephone	5,345	5.5
Paper before Day 40	719	0.7
Paper after Day 40, eligible and selected	1,460	1.5
Paper after Day 40, eligible and non-selected	7,446	7.7
Paper after Day 40, non-eligible	2,400	2.5
Total	96,553	100.0

Source: Genese survey; SSMSI processing

Interpretation: amongst the 96,553 questionnaires collected and integrated for analysis, 68.1 % were collected via Internet within the first 40 days of collection.

Figure 3.7: Determining factors for the Internet versus telephone response choice (respondents eligible for the telephone follow-up who were selected on Day 40)

	Relative risk ratio (odds-ratios)	P-value	Significance
Sex x Age (H, aged 18-29)	ref.	ref.	ref.
Sex x Age (H, aged 30-39)	0.87	0.51	
Sex x Age (H, aged 40-49)	0.93	0.74	
Sex x Age (H, aged 50-59)	1.06	0.79	
Sex x Age (H, aged 60-74)	1.03	0.90	
Sex x Age (F, aged 18-29)	1.53	0.03	*
Sex x Age (F, aged 30-39)	0.82	0.31	

Sex x Age (F, aged 40-49)	1.09	0.68	
Sex x Age (F, aged 50-59)	0.96	0.84	
Sex x Age (F, aged 60-74)	0.78	0.30	
TUU (rural municipalities)	ref.	ref.	ref.
TUU (2,000-4,999 res.)	0.87	0.41	
TUU (5,000- 9,999 res.	0.96	0.81	
TUU (10,000-19,999 res.)	1.33	0.09	
TUU (20,000- 49,999 res.)	1.00	0.98	
TUU (50, 000- 99,999 res.)	0.78	0.11	
TUU (100,000- 199,999 res.)	1.02	0.90	
TUU (200,000- 1,999 999 res.)	0.91	0.39	
TUU (Paris)	0.85	0.22	
Nationality (French by birth)	ref.	ref.	ref.
Nationality (French nationality obtained)	0.99	0.92	
Nationality (foreigner - stateless)	0.96	0.81	
Qualifications (few or no qualifications)	ref.	ref.	ref.
Qualification (CAP)	1.10	0.42	
Qualification (BAC)	1.32	0.03	*
Qualification (BAC+2)	1.45	0.01	**
Qualification (> BAC+2)	1.68	0.00	***
Status (Employed, self-employed)	ref.	ref.	ref.
Status (Employed, Public sector employee)	1.66	0.00	***
Status (Employed, Private sector employee)	1.48	0.00	**
Status (Employed, Employed by private individuals/Family)	1.90	0.02	*
Status (Unemployed)	1.36	0.10	
Status (Retired)	1.13	0.54	
Status (Disabled)	1.26	0.38	
Status (Student)	3.08	0.00	***
Status (Not working - other)	1.53	0.13	
Marital/conjugal status (Single)	ref.	ref.	ref.
Marital/conjugal status (Married)	1.33	0.01	*
Marital/conjugal status (Civil partnership)	1.29	0.08	
Marital/conjugal status (Widow)	0.60	0.03	*
Marital/conjugal status (Divorced)	0.93	0.61	
Type of housing (Flat)	0.92	0.34	
Type of household (Person living alone)	ref.	ref.	ref.
Type of household (Couple without children)	1.00	0.99	
Type of household (Couple with children)	1.13	0.38	
Type of household (Single-parent family)	0.95	0.75	
Type of household (Other)	1.00	1.00	
Living standard decile (d01)	ref.	ref.	ref.
Living standard decile (d02)	1.01	0.95	
Living standard decile (d03)	1.21	0.30	
Living standard decile (d04)	1.07	0.73	
Living standard decile (d05)	1.01	0.97	
Living standard decile (d06)	1.56	0.01	*

Living standard decile (d07)	1.38	0.07	
Living standard decile (d08)	1.60	0.01	**
Living standard decile (d09)	1.68	0.00	**
Living standard decile (d10)	1.36	0.09	
Living standard decile (NR)	0.87	0.56	
QPV resident (No)	ref.	ref.	ref.
QPV resident (Yes)	1.01	0.93	
QRR resident (No)	ref.	ref.	ref.
QRR resident (Yes)	0.94	0.73	
Has not moved house (No)	ref.	ref.	ref.
Has not moved house (Yes)	2.76	0.00	***
Victimisation reported block B1 (No)	ref.	ref.	ref.
Victimisation reported block B1 (Yes)	1.47	0.28	
Victimisation reported block B2 (No)	ref.	ref.	ref.
Victimisation reported block B2 (Yes)	1.62	0.05	
Victimisation reported block B4 (No)	ref.	ref.	ref.
Victimisation reported block B4 (Yes)	1.80	0.08	
Victimisation reported block B5 (No)	ref.	ref.	ref.
Victimisation reported block B5 (Yes)	1.22	0.77	
Victimisation reported block B6 (No)	ref.	ref.	ref.
Victimisation reported block B6 (Yes)	1.08	0.75	

Source: Genese survey (Internet respondents eligible and selected for telephone follow-up on Day 40, telephone respondents; N = 9 900); SSMSI processing

Interpretation: with other controlled characteristics, the relative risk ratio for responding via Internet, rather than by telephone, is 1.68 for an individual with a higher qualification than bac +2, compared to an individual with few or no qualifications.

*** significant effect at the 0.1 % threshold, ** significant at the 1 % threshold * significant at the 5 % threshold

Evaluation of mode effects on all the victimisation and opinion variables

An evaluation of mode effects will be based on a comparison between CAWI responses and CATI responses. Indeed, since the traditional dichotomy in the literature on mode effects compares self-administered modes and modes that use the intermediary of an interviewer, the main challenge here is to compare the Genese survey protocol's mainly self-administered mode with the protocol with interviewer's single mode.

The central issue is managing to quantify mode effects strictly linked to measurement by comparing Internet responses and telephone responses to various victimisation questions as well as questions regarding opinions and perceptions concerning security. For this, the analysis is part of a framework which, theoretically, is fairly similar to a random experiment where each participant is randomly assigned to the processed group or the control group, which constitutes the clearest quantification of the processing effect (Burtless, 1995).

Accordingly, Internet respondents who are eligible (40 days after the collection began) but not selected (by random selection) for the telephone follow-up are compared to telephone respondents (eligible and selected). The idea being to have access to the two most comparable sub-populations possible, including from the point of view of their unobservable characteristics: a mode effect purged of selection differences in observable characteristics may then be considered as a mode effect that is intrinsically linked to measurement.

3.1.3 Using regression models to evaluate the mode effect linked to measurement

To evaluate the mode effect intrinsically linked to measurement, each of the various variables for incidents of victimisation experienced in 2020 (block B of the questionnaire) and the various variables for opinions and perceptions concerning security (block C) are regressed upon a set of control variables generated by the sampling framework and the questionnaire, and upon a mode indicator variable.

3.1.3.1 Modelling used

The assessments presented are derived from *logit* modelling on all the variables of interest used. *Probit* models have also been introduced, which gave identical results with regard to the list of significant covariates, and very similar marginal effects. Since these two models are based on different hypotheses, the results obtained attest to the robustness of the link between the variable of interest and the significant covariates.

Control variables are ideally discriminating from the point of view of the mode of response, correlated to the variables of interest and not subject to mode effects. Those used for our regression models are as follows: sex x age group intersection (10 modalities), size of urban unit (9 modalities), nationality, level of qualification (5 modalities), employment status x social group intersection (9 modalities), marital status, type of housing, type of household, living standard decile, whether resident in QPV, whether resident in QRR and an indicator for not having moved house within the last 12 months.

3.1.3.2 Estimated mode effects derived from regressions upon variables of interest

For each variable of interest, the assessments obtained from the regression models, marginal effects and mode variable odds-ratios are provided. These parameters may then be interpreted in terms of the mode effects linked to measurement (purged of selection differences).

The main finding is that mode effects are generally of limited magnitude in the different victimisation variables (figure 3.8). Few are statistically significant, and moreover the effects do not systematically proceed in the same direction. Accordingly, the telephone effect is significantly negative for the question concerning acts of vandalism: on average and with controlled selection differences, people report acts of vandalism more often by telephone than via Internet. The mode effect, on the other hand, is significantly positive for physical violence and offensive comments or attitudes of a non-sexual nature. For such indicators, the effects may be related to a relative ambiguity in the wording of underlying issues, with a more pronounced *satisficing* tendency in CAWI respondents who are not necessarily prepared to make the effort to respond optimally or to self-censor in the event of hesitation.

Relative ambiguity in the wording of the questions and complexity in the concepts may also explain the opposite effect for burglary (positive effect) and theft without breaking and entering (negative effect), even if this observation needs to be qualified because the effects are not statistically significant.

Finally, it should be noted that, in accordance with theoretical expectations (social desirability), for the most sensitive questions (forced sexual intercourse, attempted forced sexual intercourse, fondling of the genitals and fondling of buttocks/breasts/hips/kissing), respondents seem to report fewer incidents by telephone than via Internet. The mode effects for these different indicators are, however, not statistically significant.

Figure 3.8: Mode variable effects (telephone vs Internet) for all the victimisation variables, assessed by *logit* regressions

	Average marginal effect	Odds-ratio	Evidence of the effect
Burglary	0.0031	1.42	
Attempted burglary	-0.0006	0.94	
Theft without breaking and entering	-0.0028	0.69	
Acts of vandalism	- 0.0056 (*)	0.70 (*)	-

Car theft	0.0014	1.42	
Theft of car accessories	-0.0014	0.93	
Theft of two-wheeled motor vehicles	0.0010	0.70	
Bike theft	0.0005	0.96	
Other offences targeting vehicles	0.0001	1.01	
Theft with physical violence or threats	-0.0005	0.88	
Attempted theft with physical violence or threats	-0.0017	0.64	
Theft without physical violence or threats	-0.0004	0.97	
Attempted theft without physical violence or threats	-0.0003	0.97	
Bank fraud	0.0059	1.10	
Physical violence	0.0078 (**)	1.65 (**)	+
Forced sexual intercourse	-0.0009	0.68	
Attempted forced sexual intercourse	-0.0015	0.61	
Fondling of genitals	-0.0013	0.63	
Fondling of buttocks/breasts/hips/kissing	-0.0022	0.89	
Offensive comments or attitudes of a sexual nature	0.0050	1.22	
Offences of a non-sexual nature	0.0083 (*)	1.32 (**)	+

Source: Genese survey (Internet respondents eligible but not selected for telephone follow-up on Day 40, telephone respondents; N= 12 300); SSMSI processing

Interpretation: with other controlled characteristics, the probability of reporting offences of a non-sexual nature on the telephone is increased by 0.83 % as compared to Internet; this effect is significantly positive and the relative risk ratio is 1.32 (for the telephone as compared to Internet).

*** significant effect at the 0.1 % threshold, ** significant at the 1 % threshold * significant at the 5 % threshold

On the other hand, mode effects linked to measurement are more marked and tend to be more significant for variables for opinions and perceptions concerning security (figure 3.9). Hence, the mode effect linked to telephone measurement is significantly negative for questions on sense of insecurity: with controlled selection differences, individuals report a sense of insecurity via Internet more often than by telephone. Doubtless also for issues linked to social desirability, mode effects are significantly positive and fairly marked concerning questions of opinions regarding the police, the *gendarmerie* and the justice system. Concerning more factual questions regarding the observation of delinquent phenomena, effects – even if potentially significant – are more measured.

Figure 3.9: Mode variable effects (telephone vs Internet) for opinion and perception variables, assessed by *logit* regressions

	Average marginal effect	Odds-ratio	Evidence of the effect
Delinquency is a worrying issue in France	-0.0200 (**)	0.84 (**)	-
Delinquency is a worrying issue within the district/village	-0.0767 (***)	0.66 (***)	-
Remote district/village	0.0237 (***)	1.19 (***)	+
Pleasant district /village	0.0308 (***)	1.50 (***)	+

Safe district /village	0.0857 (***)	1.75 (***)	+
Sense of insecurity at home	-0.0250 (***)	0.81 (***)	-
Sense of insecurity in the district /village	-0.0586 (***)	0.67 (***)	-
Sense of insecurity when using public transport	-0.1747 (***)	0.45 (***)	-
Damage to public facilities in the district /village	-0.0500 (***)	0.79 (***)	-
Damage to cars or two-wheeled motor vehicles in the district /village	-0.0379 (***)	0.81 (***)	-
Observation of phenomena associated with excessive alcohol consumption	-0.0088	0.96	
Observation of phenomena associated with the trafficking or consumption of drugs	-0.0133	0.93	
Observation of phenomena associated with other forms of trafficking (including procuring)	0.0068	0.83	
Observation of destructive acts	-0.0440 (***)	0.78 (***)	-
Observation of unwelcome or hostile behaviour	-0.0423 (***)	0.77 (**)	-
Observation of traffic offences	0.0549 (***)	1.27 (***)	+
Observation of theft	-0.0370 (***)	0.79 (***)	-
Observation of other forms of delinquency	-0.0413 (***)	0.72 (***)	-
Favourable opinion concerning the activities of the police/gendarmerie in protecting individuals and property	0.1149 (***)	1.61 (***)	+
Favourable opinion concerning the activities of the police/gendarmerie in combating serious crime	0.1089 (***)	1.56 (***)	+
Favourable opinion concerning the activities of the police/gendarmerie in combating terrorism	0.1023 (***)	1.56 (***)	+
	Average marginal effect	Odds-ratio	Evidence of the effect
Favourable opinion concerning the activities of the police/gendarmerie in maintaining law and order	0.1315 (***)	1.74 (***)	+
Generally favourable opinion concerning the operation of the justice system	0.0951 (***)	1.66 (***)	+
Favourable opinion concerning police/gendarmerie presence in the district/village	0.0857 (***)	1.45 (***)	+

Favourable opinion concerning police/gendarmerie activities in the district/village	0.1892 (***)	2.22 (***)	+
Good or fairly good relationship with the police/gendarmerie in the district /village	0.1139 (***)	1.63 (***)	+

Source: Genese survey (Internet respondents eligible but not selected for telephone follow-up on Day 40, telephone respondents; N= 12 300); SSMSI processing

Interpretation: with other controlled characteristics, the probability of expressing a favourable opinion concerning the operation of the justice system on the telephone is increased by 9.51 % as compared to Internet, and the relative risk ratio is 1.66 (for the telephone as compared to Internet).

*** significant effect at the 0.1 % threshold, ** significant at the 1 % threshold * significant at the 5 % threshold

3.1.4 Using matching models to evaluate the mode effect linked to measurement

On the basis of the same respondent sub-populations and the same selection control variables used in the regression models presented above, mode effects intrinsically linked to measurement were also assessed using *matching* models.

3.1.4.1 General principles of modelling by matching

To assess the average effect of the processing under consideration here (the response mode), propensity score *matching* was carried out, which may be defined as the probability of being processed (i.e. of responding by telephone), subject to all the socio-demographic control variables. Using such techniques is theoretically very similar to a framework for the construction of a counterfactual to assess the causal effect of processing.

According to Rosenbaum and Rubin's founding article, if the hypothesis that the processing assignment mechanism may be ignored is verified (*strongly ignorable treatment assignment*), then matching respondents by their propensity score enables the causal effect of processing to be found (Rosenbaum & Rubin, 1983).

In our case, the processing impact we are seeking to assess is CATI respondent group membership. We are seeking to measure the impact of such processing on reports of victimisation and opinions and perceptions concerning security. The hypothesis of the ignorability of mechanisms for assigning processing may mean that the two groups' propensity scores follow the same law; in the literature it is called *Conditional Independence Assumption*. The assignment mechanism would have been completely ignorable in the event that the individuals selected were obliged to respond by telephone. However, here the respondents' position is that of self-selection, since they may decide to respond via Internet, even though they have been selected for the telephone follow-up (and potentially contacted in order to respond by telephone).

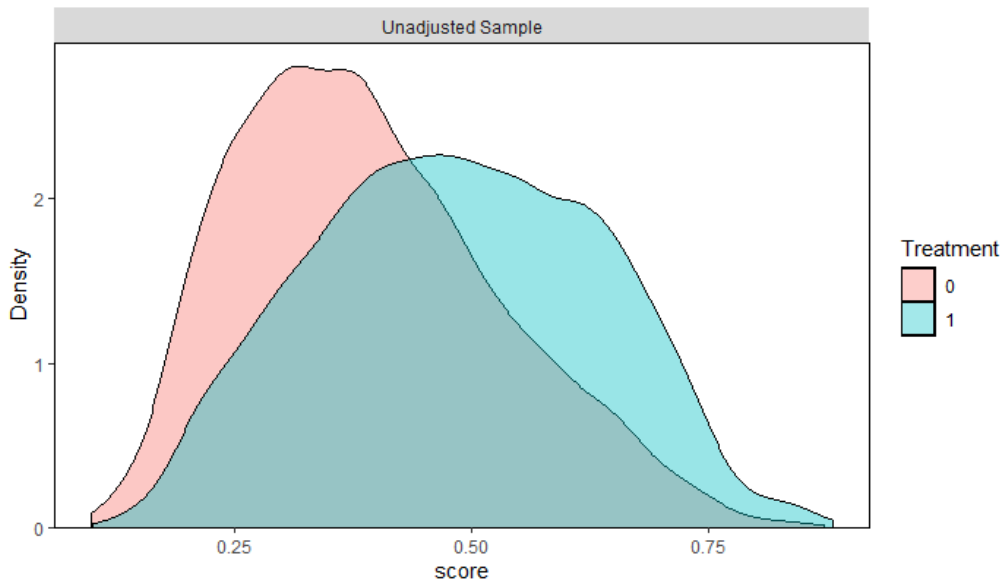
Using *matching* to assess the mode effect seems nevertheless to be largely justified here, even if there may be bias due to non-mandatory assignment. In addition, the construction of the counterfactual on the basis of the sub-populations studied (CAWI respondents that are eligible but not selected for the telephone follow-up vs CATI respondents) limits self-selection phenomena that would have been more problematic when comparing sub-populations responding at very different points in the collection process, with potentially great differences in their unobservable characteristics (motivation).

Reset matching was carried out (Abadie & Imbens, 2006) and the findings presented and interpreted, since our mode effects strictly linked to measurement relate to so-called "bias-corrected matched" estimators (Abadie & Imbens, 2011).

3.1.4.2 The issue of common support

An important issue when introducing models based on propensity scores is that of common support, the hypothesis here being that there are individuals with comparable observable characteristics for the various collection modes. The measurement effect is then quantified on this common support which must be as broad as possible for questions of external validity. As an illustration (figure 3.10), the propensity score distribution (derived from the same control variables as those used in the regression models) is provided for the CATI group (considered here as the processed group), and the eligible non-selected CAWI group (non-processed group), before matching.

Figure 3.10: Propensity score distribution for the processed (CATI) and non-processed (CAWI) groups



Source: Genese survey (Internet respondents eligible but not selected for telephone follow-up on Day 40, telephone respondents; N= 12 300); SSMSI processing

3.1.4.3 Mode effects assessed using matching models

The average effects of processing resulting from matching models, and interpreted as mode effects that are intrinsically linked to telephone measurements, are presented for all the variables of interest. Such effects are generally very similar to average marginal effects assessed using regression models, which seems to attest to the robustness of the findings. For a few rare variables of interest, the significance of the mode’s assessed effect may vary according to the modelling: for offences of a non-sexual nature, the assessed effect is significantly positive using the regression model but this is not the case using *matching* (figure 3.11); conversely, for the observation of phenomena associated with the excessive consumption of alcohol and with the trafficking and consumption of drugs, the assessed effects are significantly negative using the *matching* model but this is not the case using the regression model (figure 3.12).

Figure 3.11: Mode variable effects (telephone vs Internet) for all the victimisation variables (*matching* and *logit* summary of findings)

	Average effect of the <i>matching</i> process	Average <i>logit</i> marginal effect	Evidence of the effect
Burglary	0.0030	0.0031	
Attempted burglary	-0.0003	-0.0006	
Theft without breaking and entering	-0.0040	-0.0028	

Acts of vandalism	-0.0068 (*)	- 0.0056 (*)	-
Car theft	0.0008	0.0014	
Theft of car accessories	-0.0021	-0.0014	
Theft of two-wheeled motor vehicles	-0.0012	0.0010	
Bike theft	-0.0008	0.0005	
Other offences targeting vehicles	0.0003	0.0001	
Theft with physical violence or threats	-0.0010	-0.0005	
Attempted theft with physical violence or threats	-0.0020	-0.0017	
Theft without physical violence or threats	-0.0004	-0.0004	
Attempted theft without physical violence or threats	-0.0005	-0.0003	
Bank fraud	0.0006	0.0059	
Physical violence	0.0076 (*)	0.0078 (**)	+
Forced sexual intercourse	-0.0008	-0.0009	
Attempted forced sexual intercourse	-0.0013	-0.0015	
Fondling of genitals	-0.0009	-0.0013	
Fondling of buttocks/breasts/hips/kissing	-0.0029	-0.0022	
Offensive comments or attitudes of a sexual nature	0.0010	0.0050	
Offences of a non-sexual nature	0.0056	0.0083 (*)	

Source: Genesee survey (Internet respondents eligible but not selected for telephone follow-up on Day 40, telephone respondents; N= 12 300); SSMSI processing.

Interpretation: with other controlled characteristics, the probability of reporting physical violence on the telephone is increased by 0.76 % as compared to Internet.

Figure 3.12: Mode variable effects (telephone vs Internet) for all the victimisation variables (matching and logit summary of findings)

	Average effect of the matching process	Average logit marginal effect	Evidence of the effect
Delinquency is a worrying issue in France	-0.0199 (***)	-0.0200 (**)	-
Delinquency is a worrying issue within the district/village	-0.0825 (***)	-0.0767 (***)	-
Remote district/village	0.0294 (***)	0.0237 (***)	+
	Average effect of the matching process	Average logit marginal effect	Evidence of the effect
Pleasant district /village	0.0295 (***)	0.0308 (***)	+
Safe district /village	0.0787 (***)	0.0857 (***)	+
Sense of insecurity at home	-0.0264 (***)	-0.0250 (***)	-
Sense of insecurity in the district /village	-0.0679 (***)	-0.0586 (***)	-

Sense of insecurity when using public transport	- 0.1862 (***)	-0.1747 (***)	-
Damage to public facilities in the district /village	-0.0553 (***)	-0.0500 (***)	-
Damage to cars or two-wheeled motor vehicles in the district /village	-0.0403 (***)	-0.0379 (***)	-
Observation of phenomena associated with excessive alcohol consumption	-0.0248 (*)	-0.0088	-
Observation of phenomena associated with the trafficking or consumption of drugs	-0.0266 (**)	-0.0133	-
Observation of phenomena associated with other forms of trafficking (including procuring)	-0.0076	0.0068	
Observation of destructive acts	-0.0537 (***)	-0.0440 (***)	-
Observation of unwelcome or hostile behaviour	-0.0517 (***)	-0.0423 (***)	-
Observation of traffic offences	0.0564 (***)	0.0549 (***)	+
Observation of theft	-0.0413 (***)	-0.0370 (***)	-
Observation of other forms of delinquency	-0.0470 (***)	-0.0413 (***)	-
Favourable opinion concerning the activities of the police/gendarmerie in protecting individuals and property	0.1112 (***)	0.1149 (***)	+
Favourable opinion concerning the activities of the police/gendarmerie in combating serious crime	0.1111 (***)	0.1089 (***)	+
Favourable opinion concerning the activities of the police/gendarmerie in combating terrorism	0.1019 (***)	0.1023 (***)	+
Favourable opinion concerning the activities of the police/gendarmerie in maintaining law and order	0.1347 (***)	0.1315 (***)	+
Generally favourable opinion concerning the operation of the justice system	0.0931 (***)	0.0951 (***)	+
	Average effect of the matching process	Average <i>logit</i> marginal effect	Evidence of the effect
Favourable opinion concerning the police/gendarmerie presence in the district/village	0.0863 (***)	0.0857 (***)	+
Favourable opinion concerning police/gendarmerie activities in the district/village	0.1872 (***)	0.1892 (***)	+
Good or fairly good relationship with the police/gendarmerie in the district /village	0.1254 (***)	0.1139 (***)	+

Source: Genese survey (Internet respondents eligible but not selected for telephone follow-up on Day 40, telephone respondents; N= 12 300); SSMSI processing.

Interpretation: with other controlled characteristics (assessment by matching), the probability of expressing favourable opinions concerning the justice system on the telephone is increased by 9.31 % as compared to Internet.

Elements for discussions on the interpretation of effects

According to the literature, mode effects intrinsically linked to measurement are particularly marked in questions of opinion, in particular those relating to the activities of the police/*gendarmerie* and the justice system. The effects are also fairly clear in questions relating to a sense of insecurity. They are, on the other hand, less marked in more factual questions concerning the observation of delinquent behaviour. Such effects may be interpreted in terms of social desirability: social interaction may lead to greater compliance with normative expectations amongst the individuals being interviewed.

Generally, mode effects on the various victimisation variables directly generated by phase 1 questioning are of limited magnitude. The Genese survey's contact and targeted follow-up protocol, the communication tools made available to the respondents (website, hotline), and especially the short, simple, ergonomic phase 1 questionnaire, which was designed to be self-administered, have enabled a high level of participation to be achieved. This important advance design work probably helped to control the CAWI/CATI mode effects, which were a lot less marked than those observed in the VVS and CVS experiments carried out by Insee (cf. 2.2-2.4). These are very significant findings which nevertheless warrant consolidation for certain types of offences, the infrequency of which may compromise the performance of tests (typically for sensitive questions, theoretically subject to mode effects). The common characteristic of the few victimisation questions for which the mode effect is significant is the relative ambiguity of their wording. For victimisation questions where the telephone mode effect is positive (mainly physical violence), two main interpretations regarding response behaviour and social interactions may be offered. One relates to the tendency towards *satisficing* amongst Internet respondents, which leads them to make the minimum effort required to respond to the questionnaire, and hence to understate their reporting of victimisation in order to minimise their response time. Conversely, the intermediary of an interviewer's human voice diminishes the effort required from respondents to understand the questions. The other explanation concerning the mechanism involved may be the tendency towards self-censorship: when an individual is not certain that he/she has understood the question, he/she tends to self-censor and hence to understate the fact of having been a victim, whereas with an interviewer, he/she is able to ask for details, and therefore is more likely to class him/herself as a victim. Moreover, additional investigations could also be carried out into possible behaviour differentials according to modes in terms of the reporting of multiple incidents of victimisation, since the issue of the correct classification of offences, in the absence of multiple incorrect records, is a standard problem in the identification of victimisation (Castell *et al.*, 2021).

PART II – The historical perspective: connecting the past to the future in a present marked by the health crisis

4. General information on the overhaul of household surveys and managing serial breaks

A number of public statistics surveys, including Insee's flagship household surveys, have been subject to overhauls over the course of their history. Whether originally motivated by methodological innovation, the revision of their content, constraints imposed by regulations or international standards, the allocation of resources in the form of interviewers, or budgetary considerations, the challenge for the statistician is to design a high-quality system that brings new advantages to compensate for any losses resulting from the overhaul.

Modifications in the definition of concepts and nomenclatures may have consequences for the continuity of the indicator series derived from assessments based on survey data. However, this does not pose insurmountable problems, providing the underlying data is unmodified. Hence it may be sufficient to simply reassess the historical series by applying the new concepts: serious methodological investigations are not necessarily required, and it is rather a question of a communications exercise that needs to be anticipated as soon as possible.

On the other hand, within the context of a major overhaul of a survey process, data is no longer necessarily consistent with the past and other approaches must be introduced to ensure the continuity of the historical series of indicators. There may be several reasons for modifications that may have an impact on assessments derived from survey data: questionnaire upgrades, changes in collection protocols, modifications to a survey's scope, changes in the sampling frames used or in the sampling, or modifications in the calibration methods used.

Envisaged as a review of the literature, this chapter first makes the following observation: although survey overhauls are common within the public statistics department, processing associated with the issue of serial breaks is not extensively addressed in the literature, with the notable exception of the *Emploi en France* (Employment in France) survey (cf. 6.1). The recent context associated with the health crisis could be an opportunity for new methodological developments. Secondly, this chapter offers a brief theoretical introduction to serial connection methods.

Frequent overhauls of household surveys but very limited literature on their impact and correction

Giving rise to small changes in the formulation of questions or major upheavals in survey systems, overhauls do not always have the same consequences for the data produced. They are also not processed in the same way and depend on a system's required continuity or monitoring of indicators over time. Numerous public statistics surveys have undergone changes and overhauls in the course of their history but, in most cases, relatively few corrections have been made to data and to work in the literature on context, the impact of the changes and their measurement. Thus, we can cite recent overhauls concerning systems with serialised indicators:

- the *Statistiques sur les revenus et conditions de vie dans l'Union européenne* (Statistics on Income and Living Conditions in the European Union, SRCV/EU-SILC) system underwent several updates to its questionnaire over time. From 2004, the SRCV system was used for the annual collection of information via a survey and follow-up of the individuals interviewed during successive collections. The first of the SRCV survey's overhauls took place in 2008, during the course of which the entire system of information on living conditions was merged in the EU-SILC/ SRCV system. Some household resources were no longer collected in the survey

questionnaire but obtained by matching with fiscal and social sources. The system was subsequently overhauled a second time in 2020 in order to conform to the new European regulatory framework, shorten the questionnaire to focus on European demands, bring forward deadlines for the dissemination of findings and increase the main estimators' accuracy. Regarding the data produced, what makes this survey unique and a great asset is its capacity to conduct comparative analyses between European countries, the opportunities its rotating panel offers for the longitudinal monitoring of individuals, and its thematic approach to quality of life and living conditions (Burricand *et al.*, 2014).

- the *Technologies de l'information et de la communication* (Information and Communication Technologies, TIC) household survey underwent a number of modifications to its sampling frame. With the development of new technologies, the TIC survey has been in existence as a household survey since 2007. Between 2007 and 2010, its sampling frame consisted of the telephone numbers of subscribers listed in France Télécom's directory. From 2011 to 2018, the TIC survey sample for mainland France consisted of households taken from the *Taxe d'Habitation* (Council Tax). Since 2019, samples have been taken from Fidéli. Series are produced and distributed by Insee without documenting any corrections made to data (Insee, 2019).
- the *Conditions de travail* (Working Conditions) survey, which has documented the evolution of work for 40 years (Beque *et al.*, 2019), underwent a significant overhaul of its entire system in 2012. This survey, which is complementary to the Employment survey, was carried out once every 7 years from 1978 to 2005; since 2013, it has been a stand-alone Dares-Insee panel survey, conducted every three years, focusing alternately on "working conditions" (2013 and 2019) and "psycho-social risks" (CT-RPS, the first of which took place in 2015-2016). To ensure the survey editions' comparability over time, its designers checked a set of elements whilst they were preparing the 2012 edition and publishing its findings (unpublished work):
 - the face-to-face collection mode, despite its cost, was retained so that its findings are comparable (Algava, 2015) with surveys carried out by the same network of interviewers (Insee);
 - the sample size was maintained for a similar accuracy of findings;
 - the absence of seasonal variations in working conditions was verified;
 - the field's comparability has been analysed: the complementary survey concerns working individuals employed within the meaning of the Employment survey and this concept had to be rebuilt using questions from common household core questions;
 - the team also had to ensure that both the introduction of the "public service" and "private clinics" extensions, and the adjustment method used to integrate these extensions, did not distort the findings.

Within the context of multimode development and even though all the SSP's household surveys are affected by the transition (cf. 2.1), all its survey designers will have to investigate the question of continuity. In this regard, the population lockdowns in 2020 and 2021 associated with the coronavirus epidemic may have constituted a catalyst for change. Indeed, the lockdowns disrupted the collection of information for a certain number of household surveys, in particular those requiring a home interview face-to-face with an interviewer. No further interviews could be conducted during the months of March, April and May 2020, and all collections were switched to telephone interviews where possible, or suspended. During the second lockdown, the collection of household surveys continued, but exclusively by telephone or via Internet. Since administering certain long and complex surveys, such as *Histoire de vie* (Life History) and Heritage, solely by telephone proved to be a complicated process, a new collection mode via video-conferencing was introduced from mid-November.

Finally, to shed light on the impact of the health crisis, existing surveys such as the Camme survey, the Life History and Heritage survey and the SRCV survey – where questions or question modules were added – were upgraded.

Changes in collection modes may cause breaks in the data produced that need to be checked, analysed and possibly corrected. The period 2020-2021 also coincided with a number of multimode experiments at Insee and within the SSP. There was an increase in communications and the sharing of experiences in order to report on work carried out to manage crisis-induced breaks in the survey data and in administrative data: one example is the series of seminars on progress and innovation during the Covid crisis, organised by Insee's *Direction de la méthodologie et de la coordination statistique et internationale* (French Department of Methodology and Statistical and International Coordination, DMCSI) and also the “*Évolution de la pratique des enquêtes pendant l'épidémie de Covid-19*, Evolution of Survey Practice during the Covid-19 Epidemic” seminar held by the *Société française de statistique* (French Statistics Society, SFDS). It was too early for the Francophone Symposium on Surveys, held in October 2021, to be covered by these presentations. Working groups of survey designers are beginning to raise the issue in order to share good practices.

Theory on methods for managing serial breaks

To maintain uninterrupted and comparable series of indicators generated by repeated statistical surveys in the event of a major overhaul, two main approaches exist. These approaches are presented together in an article written by an author from Statistics Netherlands and two authors from the Office for National Statistics (van den Brakel *et al.*, 2008). The ideal framework for managing the transition is to simultaneously carry out the current process and the updated process. In the event that this approach cannot be implemented, the second best option would be to assess and correct breaks using statistical time series models.

4.1.1 Parallel operation of the current process and the updated process

4.1.1.1 General principle

The parallel operation of the current process and the updated process, with recovery in at least one period of the series, is the natural method used to assess the impact of a major overhaul on a survey process. This allows the main parameters of interest to be assessed according to the two processes and the significance of the differences between the assessments to be tested. In addition, this approach provides an ideal framework for the transition between processes, in particular from the point of view of the acceptability of the update.

To carry out such parallel surveys, the most direct frameworks are experimental protocols in which samples are divided into sub-samples in a totally random manner (*completely randomized design*) or, for improved efficiency, using block-based selection (*randomized block design*). Sample sizes are determined according to the strength of the tests to be carried out. One recovery period may suffice but logistical issues (number of deployable interviewers) may mean that the two processes need to be carried out over several periods.

4.1.1.2 Synthetic correction methods

These methods rely on simple modelling to adjust and reverse the series using the difference observed over the period when the two processes were carried out simultaneously.

T represents the period when the two processes were carried out in parallel; $\hat{y}_{R,T}$ and $\hat{y}_{N,T}$ represent the estimators for the observed parameter via the current process R and the updated process N respectively, for the period T.

An initial method is the additive approach:

$$\tilde{y}_{N,t} = \hat{y}_{R,t} + (\hat{y}_{N,T} - \hat{y}_{R,T}) = \hat{y}_{R,t} + \hat{\Delta}_T \quad \text{with } t = 1, \dots, T-1$$

The correction is then independent of the parameter value $\hat{y}_{R,t}$, which is not always credible. As the adjusted series may then take negative values, a multiplicative approach is logically preferable:

$$\tilde{y}_{N,t} = \hat{y}_{R,t} \frac{\hat{y}_{N,T}}{\hat{y}_{R,T}} \quad \text{with } t = 1, \dots, T-1$$

This approach is based on the hypothesis that the correction is proportional to $\hat{y}_{R,t}$. Finally, for the proportions, another correction is proposed:

$$\tilde{y}_{N,t} = \hat{y}_{R,t} + \hat{\Delta}_T \frac{\hat{y}_{R,t} (1 - \hat{y}_{R,t})}{\hat{y}_{R,T} (1 - \hat{y}_{R,T})} + \hat{\Delta}_T \quad \text{with } t = 1, \dots, T-1$$

Corrections may be applied to sub-populations by applying the estimated correction factor to the entire population. For example, for the multiplicative approach:

$$\tilde{y}_{N,t}^r = \hat{y}_{R,t}^r \frac{\hat{y}_{N,T}}{\hat{y}_{R,T}} \quad \text{with } t = 1, \dots, T-1$$

But such a correction is then based on a very strong hypothesis: the differences observed at the general level would also hold for the sub-populations of interest. Moreover, this may also pose consistency problems between adjusted series. Consistency may nevertheless be ensured by integrating a linear constraint and via optimisation techniques.

Therefore correction methods are based on hypotheses that are strong enough to extrapolate the observed difference over the period prior to recovery. These hypotheses are less and less credible the further the period of adjustment is from the period when the parallel processes are operated.

4.1.1.3 The model approach in situations where numbers are too low

The use of traditional survey inference techniques for synthetic correction analysis is relevant when, within the context of the parallel operation of the two processes, sample sizes are sufficiently large. But the limits of this approach become apparent when samples assigned to the updated process are not large enough. In such situations, a model approach (as opposed to *design* approach) may be used to assess the discontinuities associated with overhauling a survey process. This type of method is used, for instance, to assess discontinuities at sub-national level, along with small-range assessment techniques (Rao & Molina, 2016; van den Brakel *et al.*, 2016; van den Brakel *et al.*, 2021).

4.1.2 Time series approach

4.1.2.1 Principle

The time series approach is to be preferred if it has not been possible to operate the two processes in parallel, for instance for budgetary reasons. Moreover, even when both processes have been operated simultaneously, this approach allows additional points to be integrated into a series generated by an updated process, and breaks to be quantified more accurately. Certain recent works (van den Brakel *et al.*, 2017; van den Brakel *et al.*, 2020) therefore use both approaches together.

This approach requires sufficient perspective: assessments may be improved as data is added later. This may be envisaged within the framework of a well-defined revision policy.

4.1.2.2 Implementation

In the event that there is no recovery between the current process and the updated process, the idea is to explain when the update was operated in a time series model (*intervention analysis*). This approach assumes that the underlying model accurately describes the evolution of the indicator and that there is no structural change in the trend or the seasonal component when the overhaul is implemented. Indeed, should there be structural change during this period, it will then be impossible to distinguish structural changes in the indicator from the intrinsic effects of overhauling the process.

Implementation may be based on REGARIMA or TRAMO models, by integrating the moment when the process is updated and the duration of the transition. Another solution is the structural modelling of the time series, broken down into a trend, a seasonal component, a component predicted by explanatory variables and an irregular component. The explanatory vector then contains at least one variable, known as an intervention variable, which indicates the moment when the overhaul is implemented. The traditional method is to write such models in space-state form and obtain assessments using a Kalman filter (Durbin & Koopman, 2012). The value estimated for this intervention variable is then interpreted as the series' discontinuity due to the overhaul.

The measurement equation describes how the observed series depend on unobserved state variables α and errors ε . Within our context, state variables are the trend models' level and slope components and the intervention variables' regression coefficients.

$$\hat{y}_t = Z_t \alpha_t + \varepsilon_t$$

The transition equation describes how the state variables evolve over time.

$$\alpha_t = T \alpha_{t-1} + \eta_t$$

If the recovery period between the two processes is long enough, then a bivariate structural time series model can be considered. The vector components are then observed together solely during the recovery period.

5. Victimisation measurement in France: discontinuities and connections

Frameworks

Historically, the measurement of delinquency was based on administrative statistics including prison, judicial and, more recently, police statistics (Robert & Zauberman, 2011a). The limitations of data recorded by the police are well known (Zauberman *et al.*, 2009; Estival & Filatriau, 2019). It does not allow the comprehensive measurement of delinquent phenomena because it only identifies offences brought to the attention of the security forces or judicial authorities (complaint, statement, intervention, etc.). It therefore depends heavily on the propensity of victims to lodge complaints, on complaint recording practices and on enforcement. In addition, the typology upon which registered delinquency statistics are based must adapt to changes in legislation and reforms that define the boundaries of offences, petty crime, crime, etc. Moreover, it does not provide a detailed socio-demographic profile of victims or contribute to an understanding of the influence of delinquency on the representation and behaviour of individuals. Given these well-identified and documented limitations, another approach has been under development in the English-speaking world since the 1960s: victimisation surveys (Zauberman, 2015). They provide knowledge of direct victim delinquency, whether or not the individuals concerned have reported the incidents to the competent authorities. Their principle is simple: households and/or individuals are contacted and asked to indicate and describe incidents of delinquency (from a specified list) of which they had been victims during a reference period. They may also be asked about their opinions and perceptions concerning security and the activities of the public authorities in combating delinquency (suppression, prevention, judicial response, etc.).

In France, it is possible to roughly distinguish three periods of direct victimisation measurement via surveys (Robert & Zauberman, 2011b):

- **1980-1996 “the era of the initial surveys”**: in the early 1980s, the *Centre de recherche sociologique sur le droit et les institutions pénales* (Centre for Sociological Research on Law and Penal Institutions, Cездip) embarked on a sustained programme of local and national victimisation surveys. Following a qualitative survey conducted in 1982, in 1986 Cездip conducted an initial national survey on a wide range of incidents of victimisation (covering the years 1984-1985): burglary, theft, assault both sexual and between cohabitants, fraud, delinquent behaviour within businesses and delinquency relating to labour law. This first survey also included an important section on opinions and attitudes, especially those concerning delinquency and social change (Zauberman *et al.*, 1990; Zauberman & Robert, 1995).
- **1996-2006 “the EPCV era”**: from 1996, Insee included a victimisation module of around twenty questions on a limited number of incidents involving property and individuals within the fixed part of the permanent survey on household living conditions (EPCV), conducted annually in January. In the wake of the Caresche/Pandraud parliamentary report’s recommendations in 2002 and the creation of the *Observatoire national de la délinquance* (French National Delinquency Observatory, OND, a department of the INHESJ (French National Institute for Advanced Studies in Security and Justice) which was created in 2003 and then became the ONDRP until it was shut down in 2020), Insee upgraded the January 2005 EPCV victimisation module. The years 2005/2006 thus constituted a pivotal period between the EPCV era and the era that followed: the upgraded module entitled “Living Environment and Security (CVS)” made it possible to provide details of incidents of victimisation from the usual fixed part of the EPCVs using a sample of over 12,000 respondent households. This arrangement was continued in January 2006.
- **2007-2021 “the CVS era”**: in 2007, the introduction of the SRCV panel (Statistics on Income and Living Conditions – European EU-SILC system) led to the end of the EPCV series. The Insee-OND partnership continued from 2007 with the completion of an annual survey devoted to the study of victimisation, using the name of the prototype conducted in 2005 and 2006. Since 2007, fourteen editions of the CVS survey have been conducted in mainland France. There

have also been overseas versions (in Reunion in 2011, in the 3 French territories in America in 2015, and in Mayotte in 2020). Since its creation in 2014, the SSMSI has been associated with the joint management of the CVS survey.

In April 2018, Insee informed the SSMSI and the ONDRP of its severe budgetary constraints in terms of survey programming and associated resources in the form of interviewers, and announced that, from 2022, it would no longer be continuing with the CVS survey in its current form (annual issue, collection mode, etc.), hence the overhaul of the CVS survey constitutes the beginning of a new era for the statistical measurement of victimisation in France.

Given the importance of security issues in the national debate and in public policy, putting the findings generated by the new system into a historical perspective in order to provide trend analyses constitutes a major challenge which has been mentioned several times within the overhaul project's Consultation Committee and reaffirmed by the Scientific Council. It is indeed important to be able to compare one year's findings with those of previous years and also to be able to compare the evolution of delinquency with "extra-delinquent" data (youth unemployment, living standards, etc.). Although the CVS survey's overhaul is resolutely in line with this trend analysis objective, since it is aimed at improving the monitoring of victimisation and perceptions concerning security, it nevertheless represents a risk for the analysis of trends in phenomena measured before 2022. Therefore, in carrying out the overhaul, the historical depth of certain key indicators needs to be preserved as far as possible by drawing on previous work connecting victimisation series, by examining the CVS survey's discontinuities and the backcasting of series carried out in 2016, and finally by mobilising all auxiliary information that facilitates the connection of the two systems within a context marked by the health crisis.

Connection work during previous overhauls

The EPCV surveys were carried out face-to-face at national level between 1996 and 2006 on average samples of 6,000 to 7,000 respondent households and 11,000 to 12,000 respondent individuals. They included a module of around twenty questions on victimisation. In 1999, this module was updated at the request of the *Institut des hautes études de la sécurité intérieure* (French National Institute for Advanced Studies in Security, IHESI, which was created in 1989 and became INHESJ, French National Institute for Advanced Studies in Security and Justice). This annual measure has, for the first time in France, made it possible to offer an evolving measurement of delinquency other than that based on data recorded by police and *gendarmerie* forces.

The victimisation data generated by the initial 1996 EPCV survey was compared to data from the survey carried out in 1986 by Cездip, and developments between the two surveys were analysed in relation to recorded delinquency statistics (Robert *et al.*, 1999). The annual nature of the EPCVs provided the joint Cездip-OSC (*Observatoire sociologique du changement*, French Sociological Observatory of Change) team with an opportunity for new comparative work on "police" data (Lagrange *et al.*, 2004). The two last editions of the EPCV victimisation module were updated and offered to a household sample that was twice as large. All the victimisation data collected in the EPCVs, connected to the findings of the 1986 survey, was once again the subject of an annotated series in 2008 (Robert *et al.*, 2008). Finally, during the CVS switch, Cездip researchers focused on connecting the EPCVs to the new survey (Miceli *et al.*, 2009).

Comparisons were generally made possible by the relative homogeneity of collection within the various systems. In particular, the population range concerned, the reference period (the two previous years) and the inclusion of questions which allowed series to be compared according to the three following criteria: prevalence, consistent multi-victimisation (the number of times respondents reported that they had been the victims of a given type of victimisation) and referral to the authorities. The overlap between the various offences was fairly comfortable, but overall 4 large series of offences could be analysed as changes occurred: burglary, theft of and from vehicles, personal theft and assault. The data generated by the systems to be connected is then compared with the recorded delinquency series ("police statistics") and with other local or regional surveys (Île-de-France victimisation survey, and Cездip local

surveys conducted in Amiens in 1999 and in Aulnay-sous-Bois, Aubervilliers, Gonesse, Saint-Denis and in the Lyon conurbation in 2005, Robert *et al.*, 2006).

These various connection exercises have made it possible in practice to consolidate certain observations: a decline in the trend of damage to property, an opposite trend in assaults. Physical violence appears to be stagnating but lower-intensity violence, especially verbal violence (threats and insults) are more common and also more frequently repeated. A significant limitation highlighted by the authors concerns the size of the samples which often fail to accurately characterise incidents of victimisation and refine the analysis of changes in the rarer offences. Present in early surveys but not in EPCVs, sexual assault and domestic violence were not the subject of a series. For such offences, it should at least be possible to compare data from the new national survey with the CVS stacks as well as with the series of Paris surveys (IAURIF, 2020).

This methodology, repeated several times to evaluate the serialisation of French victimisation survey data, provides a grid for evaluating the connection of the new system to the CVS survey. Beyond the victimisation data, an analysis of the connection may be extended to data on opinions and perceptions concerning security, given the recovery of such questions in CVS and the new system. In addition, the new survey presents an interesting sample size that should enhance comparative work.

CVS 2007-2019: a fairly stable system but not devoid of discontinuities

The CVS survey (cf. Annexe I) was continued every year for 14 editions (with the exception of the year 2020 due to the health crisis). Although the system was fairly stable between 2007 and 2019, it has nevertheless undergone changes over time and in particular a revision of the series:

- changes in content: introduction or removal of modules or questions, in 2017 the modification of the question allowing the gathering of incidents of sexual violence in 2017 which engendered, within the #MeToo post-wave context, a possible series break;
- a significant overhaul of the adjustment methodology which led to a revision and backcasting of the series in 2016.

5.1.1 Questionnaires 2007 to 2019

The structure and content of the CVS questionnaires have a broad and stable common base throughout its editions, however certain modules may have appeared or disappeared between 2007 and 2019, and questions and response modalities may have been reformulated or added without an evaluation of the impact on the measurement of the main indicators.

The household questionnaire (QM) has undergone few major evolutions throughout its editions, apart from:

- the introduction of the distinction between acts attempted and acts committed in the identification of thefts of items in or on cars, from the 2009 edition;
- the introduction in 2010 of the module on bank fraud;
- the withdrawal, from the 2017 edition onwards, of the section entitled “consequences of household victimisation” which investigated the question of the impact of lack of security on household mobility requirements and school mapping avoidance strategies for households with a child/children aged from 3 to 18;
- the introduction of the distinction between acts attempted and acts committed in the identification of burglaries, added from the 2017 edition onwards.

In addition to modifications to the wording of questions or to response modalities, the individual IQ questionnaire underwent a number of significant developments throughout its editions, in particular:

- the introduction in 2012, in the “physical violence” and “threats” modules, of a question clarifying the discriminatory nature of the offence by distinguishing 3 (non-exclusive) motives based on

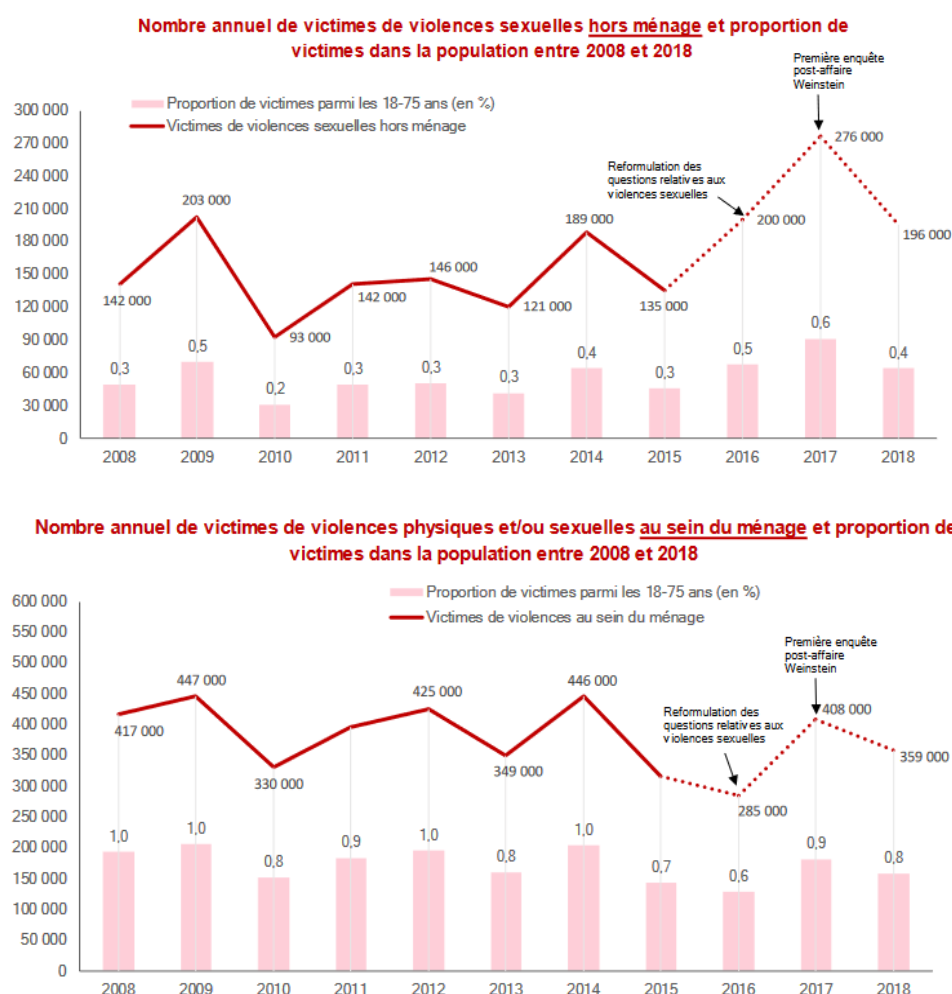
the question posed since 2007 in the “insults” module: 1. racist, antisemitic, xenophobic; 2.homophobic; 3. sexist;

- the removal, from 2014, of the “Public transport” section which included just under 20 questions and analysed sense of insecurity on public transport according to the type of transport, the frequency and reasons for its use (work, holidays, etc.). Since the survey’s beginnings, this section had been financed by the *Conseil Général de l’Environnement et du Développement Durable* (The French General Council for the Environment and Sustainable Development, CGEDD). The CGEDD had found the data difficult to use, and decided to end its financial contribution, which resulted in the section’s removal;
- the introduction in 2018 of three new modules: “scams”, “corruption” and “victimisation and discrimination” in order to meet new needs expressed, in particular, during meetings of the consultation committee. These 3 modules were continued in 2019. In 2020, the “scams” module was removed, the “corruption” module remained unchanged and the 3rd module was maintained in a short version that included only discrimination;
- the removal from 2018 of the “series” section identifying separate offences committed by the same author or group of authors. Since this section was rather difficult to use, and did not allow the measurement of phenomena involving harassment or “stalking”, it was removed so that 2018’s 3 new modules mentioned above could be introduced without lengthening the average time spent administering the questionnaire;
- for the same reasons, the “mobile telephones” section which provided the specific identification of telephone theft, was removed from 2018.

The QAA has undergone significant modifications:

- the introduction in 2009 of questions allowing victims in N-1 to be distinguished from victims in N-2;
- the introduction in 2010 of a series of questions on “household threats”: identification and description of threats made by a cohabiting individual at the time of the survey;
- the removal from 2014 of the “household threats” questions so as to introduce a new series of questions entitled “Psychological abuse by spouses or ex-spouses”;
- the 2017 edition had a great impact on the QAA:
 - o firstly, with the harmonisation of the questions identifying sexual violence outside the household and within the household. This reformulation was accompanied by a serial break that was difficult to quantify and interpret over time since subsequent measurements (2018 and 2019) had also been gathered within the post-Weinstein affair and #MeToo wave context. The significant fluctuations observed recently in the prevalence of sexual violence thus present multifactorial causes and must be discussed with caution ([figure 5.1](#)).
 - o then with the introduction of questions to distinguish between victims in N-1 and victims in N-2 for incidents of exhibitionism and imposed physical contact of a sexual nature, excluding sexual violence;
- likewise, the QAA was modified in 2018 for the time-constant integration of the 3 new IQ modules. Hence it was marked by:
 - o the removal of questions identifying sexual violence and physical violence per cohabitant prior to N-2
 - o the removal of questions on “psychological abuse by spouses or ex-spouses”.

Figure 5.1: Sexual violence outside the household and violence within the household in CVS 2009-2019



Coverage: individuals aged 18–75 living in ordinary housing in metropolitan France

Source: CVS Surveys 2009-2019; SSMSI processing, CVS 2019 survey report.

5.1.2 The overhaul of the adjustment methodology in 2016 and the back-casting of the series

Up until 2015, the survey's weighting method was based on simultaneous calibration between household and individual levels, to ensure consistency between estimators calculable from the household sample and those calculable from the sample of individuals. Each weighting was based on the units' sampling weights via a one-step margin calibration, without specific non-response modelling.

However this approach posed a certain number of problems;

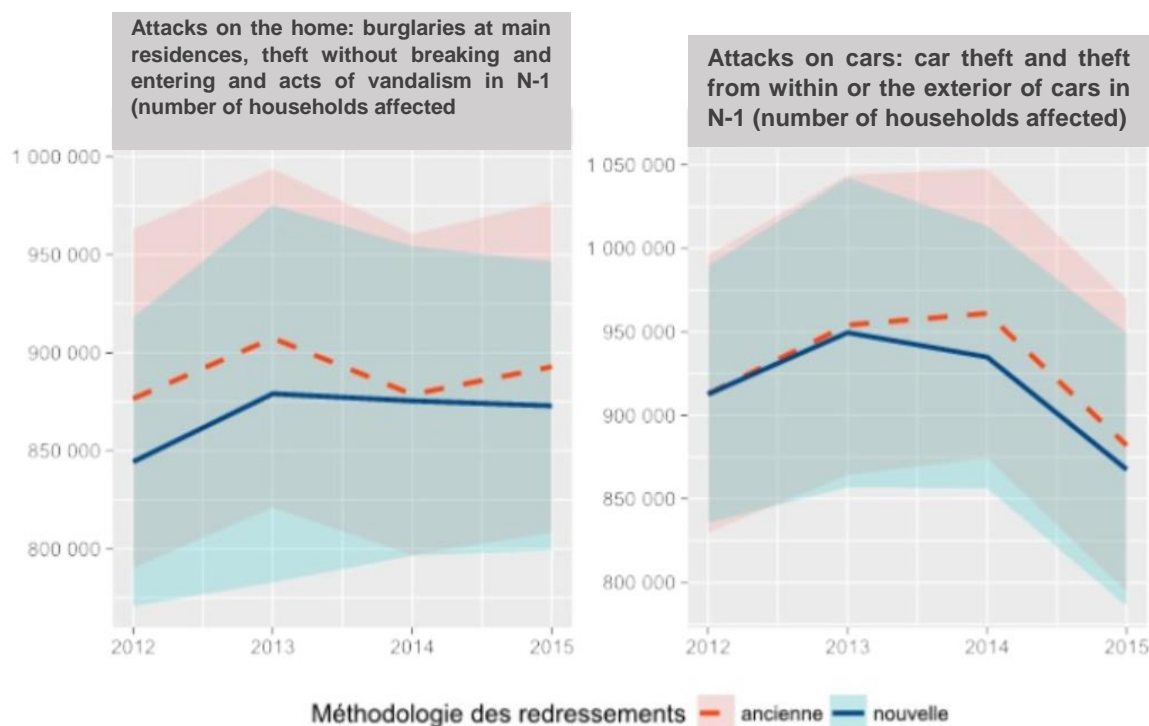
- on the one hand, each year, for around 500 households, solely responses to the household questionnaire were available. Since, for these households, the person selected for the individual questionnaire from amongst the eligible individuals in a household is a non-respondent, the information collected via these questionnaires cannot be used;
- the one-step non-response correction by margin calibration may, on the other hand, in certain very specific situations, result in a deterioration in estimator properties (Haziza & Lesage, 2006). Even if situations where one-step calibration leads to an increase in non-response bias are rare, a separation of the steps for correcting total non-response by reweighting and margin calibration is recommended by Insee's statistical methods department.

The method for adjusting and therefore calculating the weight was completely reviewed by Insee's Methodology Directorate in 2016. The decision was taken to produce three separate weights: a weight associated with respondent households (regardless of the response behaviour of the household's Kish individual), with each respondent-Kish individual, and with each Kish-individual responding to the questionnaire using headphones. As a result, all completed household questionnaires may be used. On the other hand, the discontinuation of simultaneous calibration has made it possible to reduce the number of constraints imposed by calibration and introduce new margins for improved accuracy.

The new adjustment method was replicated for the 2012–2016 surveys. It initially enabled 500 to 700 household questionnaires to be reintegrated per year. Then the Surveys Division compared the two adjustment methods. It appeared that the confidence intervals overlapped very broadly. Moreover, trends were in most cases retained (bearing in mind that such trends themselves generally evolve within the estimator's confidence interval). Nevertheless, evolutions in levels were observed: generally household victimisation was on the decline and victimisation measured in individual and self-administered questionnaires was on the increase (figures 5.2 à 5.4).

The years 2007 to 2011 underwent backcasting; the levels were adjusted to those of the period 2012–2015 but the evolutions retained. Weights were modified using margin calibration to ensure that the estimators calculated using individual data from surveys carried out between 2007 and 2011 were equal to those of the backcast series. This method is similar to the one used on the *Emploi en Continu* (Continuous Employment) Survey data between 2003 and 2012 in order to backcast the effect of the 2013 questionnaire change in the level of estimators generated by the survey (Insee, 2015).

Figure 5.2: Comparisons of assessments of the main offences derived from the household questionnaire before and after the overhaul of the CVS 2009-2019 adjustment methodology



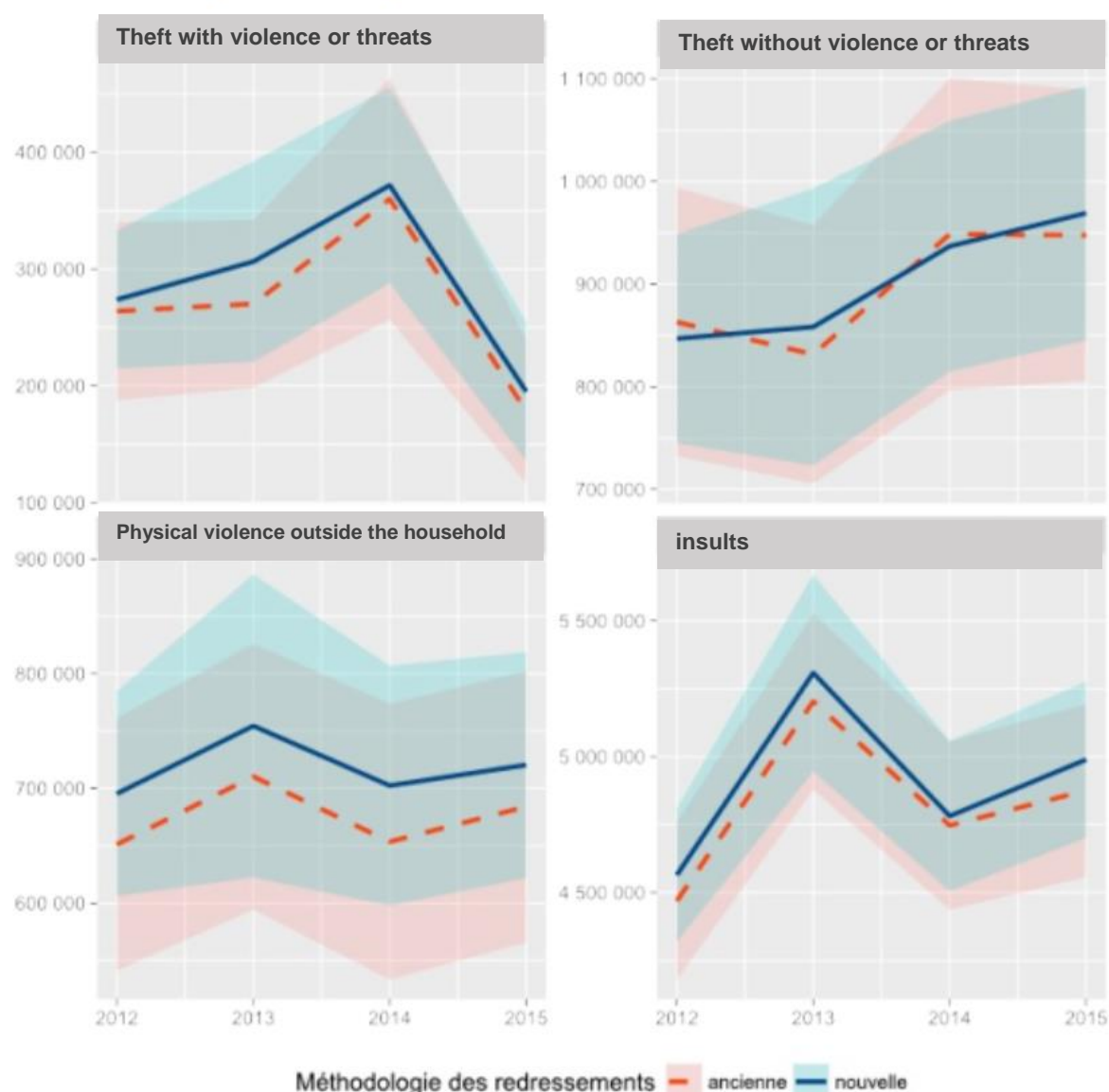
Source : Enquêtes Cadre de vie et sécurité, millésimes 2012-2015.

Champ : France métropolitaine, ménages ordinaires.

Lecture : Avec l'ancienne méthodologie des redressements, l'estimateur du nombre total de ménages déclarant une atteinte au logement l'année précédant l'enquête lors du millésime 2012 est d'environ 875 000 avec un intervalle de confiance à 95 % de [790 000 ; 960 000]. Avec la nouvelle méthodologie des redressements, cet estimateur est d'environ 840 000 avec un intervalle de confiance à 95 % de [770 000 ; 920 000].

Insee processing 2018

Figure 5.3: Comparisons of assessments of the main offences derived from the individual questionnaire before and after the overhaul of the CVS 2009-2019 adjustment methodology



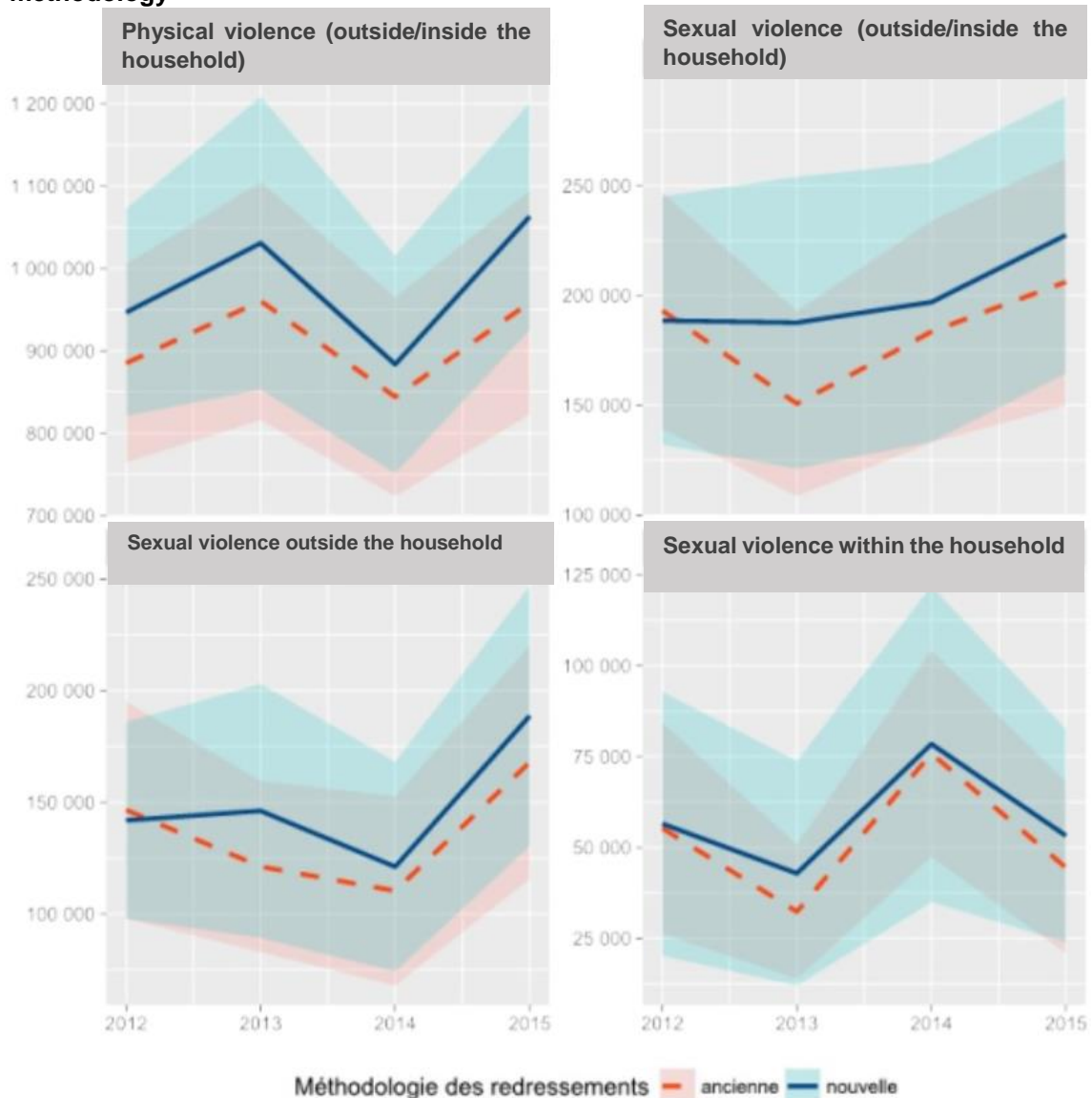
Source : Enquêtes Cadre de vie et sécurité, millésimes 2012-2015.

Champ : France métropolitaine, individus âgés de 14 ans ou plus au 1^{er} janvier de l'année d'enquête appartenant à un ménage ordinaire.

Lecture : Avec l'ancienne méthodologie des redressements, l'estimateur du nombre total d'individus déclarant avoir été victime d'un vol avec violence l'année précédant l'enquête lors du millésime 2012 est d'environ 265 000 avec un intervalle de confiance à 95 % de [185 000 ; 340 000]. Avec la nouvelle méthodologie des redressements, cet estimateur est d'environ 275 000 avec un intervalle de confiance à 95 % de [215 000 ; 335 000].

Insee processing 2018

Figure 5.4: Comparisons of assessments of the main offences derived from the self-administered questionnaire before and after the overhaul of the CVS 2009-2019 adjustment methodology



Source : Enquêtes Cadre de vie et sécurité, millésimes 2012-2015.

Champ : France métropolitaine, individus âgés de moins de 75 ans au 1^{er} janvier de l'année d'enquête, majeurs au moment de l'enquête et appartenant à un ménage ordinaire.

Lecture : Avec l'ancienne méthodologie des redressements, l'estimateur du nombre total d'individus déclarant avoir été victime de violences physiques l'année précédant l'enquête lors du millésime 2012 est d'environ 890 000 avec un intervalle de confiance à 95 % de [760 000 ; 1 000 000]. Avec la nouvelle méthodologie des redressements, cet estimateur est d'environ 950 000 avec un intervalle de confiance à 95 % de [820 000 ; 1 070 000].

Insee processing 2018

The 2020-2021 turning point: a switch marked by the health crisis

5.1.3 The two previous editions of CVS

The two previous editions of the CVS survey included modifications to the questionnaires in 2019 and two notable changes liable to induce serial breaks prior to the planned overhaul in 2022:

- a shift in the collection schedule: the Insee survey schedule for the first half of the year was revised in 2019 to allow data from the SRCV system to be sent according to the new European schedule from 2020. Hence the SRCV survey, that until 2019 was being administered from early May to the end of June, was scheduled for February – April, replacing the CVS survey. This shift in the collection dates may have an impact on recollections of events and may affect comparability, due both to the increased time lag between the collection date and the start of the reference period imposed by the questionnaire (“oversight” effect), and to the desire to report incidents that took place during the first months of the survey year despite the fact that they do not fall within the reference period (“telescoping” effect; Sudman & Bradburn, 1973; Gottfredson & Hindelang, 1977). A shift in the schedule may also have an impact on the activities of individuals and consequently on the prevalence of measured offences and the sense of insecurity. However, according to Insee, the impact of this shift should remain limited. Indeed, in the CVS survey, for each incident of victimisation, the respondent must specify the month, or failing that the season, during which the last incident took place. Depending on the incidents of victimisation, 77 % – 93 %² of respondents spontaneously provide the month in which the event took place. The most common incidents (insults, damage to cars) are the most difficult to date. Hence the memory effect appears to have the most impact on these less significant/violent incidents. Moreover, the collection of the CVS survey in Guyana took place in two waves, from early January 2015 to the end of June 2015. The data revealed very slightly lower rates of victimisation for individuals who had participated in the second wave, as regards certain incidents of victimisation. These small differences did not necessitate a correction of the victimisation rates.
- modifications in sampling (reduction of sample size and change of sampling frame): in 2020, the planned main CVS sample consisted of 16,300 address files and was to be accompanied by an “Urban Priority Districts” (QPV) over-sample of 3,000 address files, i.e. a total sample size of 19,500 address files as in 2019. The size of the 2021 survey sample was reduced by around 30 %: it is composed of 13,700 address files, of which 2,500 were in QPV. In addition, like all current survey samples, the 2021 sample was drawn from the new master-sample (*Nautille*, based on fiscal files, which has replaced *Octopusse*, which was based on the population census; Sillard & al., 2020), in accordance with Insee’s commitment to budget management. According to Insee, the risk of breaks induced by the change of sampling frame is low. However, this modification requires the non-response model to be adjusted for household-level weighting. Since the QPV over-samples are already drawn from fiscal sources (Fideli), a non-response model using this sampling frame has already been created. In addition, the change of frame could lead to greater accuracy. In fact, in *Nautille*, the primary unit draw was balanced on the variables on 1st January 2016, whereas in *Octopusse* balancing was carried out on variables derived from the 1999 population census. On the other hand, the rotating nature of the census required balancing on each of the five rotation groups (which led to the introduction of five balancing conditions in order to balance on one variable and therefore placed very great limitations on the number of characteristics to be balanced). The return to a comprehensive sampling frame thus makes it possible to significantly increase balancing possibilities, with a significant improvement in first-degree variance. These improvements in accuracy could (at least partly) offset losses induced by reducing the size of samples and help maintain the high quality of the findings disseminated in 2021.

Due to the Covid-19 health crisis, Insee was obliged to discontinue its face-to-face surveys from 16 March 2020 and only gradually resumed them from 15 July. The social distancing constraints and travel

² Insee operations.

restrictions associated with the health crisis prevented the CVS 2020 survey from being carried out face-to-face as usual. Given the length of the questionnaire and the subjects addressed, it was not possible to switch to telephone collection of the CVS survey within the specified schedule. Nor was it possible to carry this collection over to the second half of the year, to maintain other ongoing operations such as the *Trajectoires et Origines* (Trajectories and Origins) survey that had not been reissued for 10 years.

The 2021 collection was prepared and developed so that it could be adjusted to the health situation, with the option of “traditional” face-to-face administration and the alternative option of administration by telephone using a reduced form of the questionnaire, including the complete deletion of the self-administered part. Covid’s “3rd wave” and the government announcements of 31 March led to the collection’s being carried out exclusively by telephone between 19 April and 9 June. As of that date, interviewers who were happy to resume face-to-face collection were able to return to the field. Therefore the collection carried out between 9 June and 26 June 2021 was a mixture of telephone and face-to-face. In total, 66 % of the households contacted took part. Face-to-face interviews represented less than 2 % of the completed questionnaires.

At the time of writing this report, the CVS 2021 data is at the statistical processing phase, where the final survey weights are produced. Initial elements for analysis are offered in Part III (cf 7.1).

5.1.4 Auxiliary sources to compare for the year 2020

The absence of CVS 2020 and the carrying out of a CVS 2021 under conditions that were different to previous editions naturally complicated the examination of the connection of CVS editions to the new survey system that was to be in operation from 2022. The year 2020 was atypical for numerous categories of offences, as the data derived from procedures recorded by police and *gendarmerie* forces demonstrates (SSMSI, 2021).

In order to carry out this comparison work, it will be particularly important to compare all the available data in an attempt to distinguish the various effects induced by the change in methodology (collection modes, change of questionnaire, etc.) from contextual effects associated with the health crisis. To this end, the following data can be used:

- Genese 2021: the survey conducted by the SSMSI on a very large sample relates to lifetime victimisation, perceptions and opinions concerning security, distinguishing between incidents that occurred in 2021, 2020 (outside/during lockdowns) and previously (cf. Annexe II).
- Epicov wave 2: Inserm-Drees’s³ (French Directorate for Research, Studies, Evaluation and Statistics) major survey on the health status and living conditions of the population, in relation to the Covid-19 epidemic (Epicov) was conducted in several waves. In the autumn of 2020, the wave 2 questionnaire included questions on domestic violence and various types of victimisation (theft, burglary, fraud), to provide additional insights.
- SRCV 2020: the Statistics on Income and Living Conditions (SRCV) survey contained a module of questions on victimisation in 2019. However, this was administered to a reduced number of households since the survey’s collection was suspended due to health restrictions during the spring 2020 lockdown.
- Recorded data on delinquency for the year 2020 assessed by the SSMSI
- The Île-de-France victimisation survey: a comparison of the 2019 and 2022 CVS editions (to be confirmed) with the 2019 and 2021 editions and the Genese survey’s findings - representative at regional level – could undoubtedly provide a better understanding of the effect of methodological changes on the measurement of victimisation.
- The 2022 *Vécu et ressenti en matière de sécurité* (Experience and Feelings concerning Security, VRS) survey: the new SSMSI victimisation survey will allow incidents that occurred during the year 2020 (N-2) to be studied.

³ <https://drees.solidarites-sante.gouv.fr/sources-outils-et-enquetes/enquete-epicov-epidemiologie-et-conditions-de-vie-sous-le-covid-19>

6. Examples of corrections of serial breaks in the Employment survey and in Swedish and Dutch victimisation surveys, within the SSP and in neighbouring countries

This chapter contains practical illustrations, with a few methodological elements for managing continuity in indicator series generated by statistical surveys:

- work relating to various Employment survey overhauls (transition to continuous in 2003, the questionnaire's overhaul in 2013 and its integration into multimode from 2020). Since the Employment survey should enable labour market developments to be monitored over time, this work is valuable and particularly well documented;
- examples from neighbouring countries (the Netherlands and Sweden) of the connection of indicator series following major overhauls of their victimisation survey systems.

Case studies involving the connection of indicator series generated by surveys repeated over time

6.1.1 Overhauls of the Employment survey: contexts and backcasting techniques

Conducted since 1950, the Employment survey has undergone numerous changes of a conceptual nature throughout its history, notably in order to comply with the *Bureau international du travail* (International Labour Office, BIT) and Eurostat, or to improve the measurement of transformations in the labour market, but also in aspects relating to methodology (sampling, non-response processing, collection mode, etc.) and techniques, with the increased computerisation of data collection and processing (Goux, 2003). There is abundant literature on changes made and work done to preserve time series.

Without going into its history in detail, the last twenty years illustrate the main changes in the survey and their varied nature. From 2003, the survey became a quarterly rather than an annual one, and it was collected continuously every week during each quarter. This significant change meets Europe's desire to harmonise unemployment statistics. From July 1998 to June 2011 its implementation was preceded by a "lightweight system", a sort of prototype of the new continuous Employment survey in order, on the one hand, to test the organisation of the collection and work within the regional directorates, and on the other, to collect seasonal information on employment and unemployment that would enable the series of possible seasonal effects to be corrected. The sample surveyed was small and the questionnaire was reduced to a minimum. Alongside this system, tests were carried out until 2001. Before that, between 1992 and 1996, methodological operations in the form of quarterly survey experiments were carried out (Goux, 2003). In 2002, two simultaneous surveys on employment were introduced: an annual survey in March and a continuous quarterly survey. A comparison of the assessments provided by the annual survey during the month of March, and in the first quarter for the new continuous survey, enabled the gap to be measured and found minimal differences concerning the measurement of employment and unemployment (Givord, 2003).

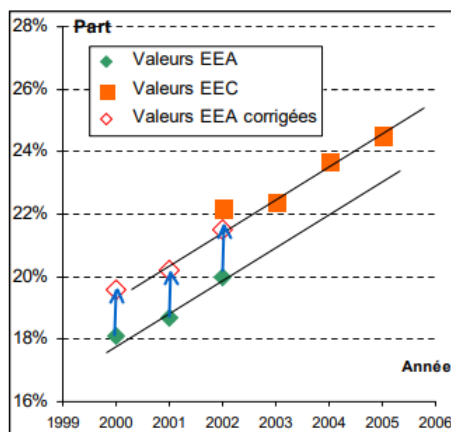
Then, in 2009 – another significant date –, in response to controversy about unemployment figures in 2006-2007, its sample was increased by 50 %. In 2013 its questionnaire was updated, including rewriting the questions to improve the switch to telephone, comply with European regulations and adjust to evolutions in the labour market. This overhaul was an opportunity to measure the extent to which modifications to the questionnaire may affect response behaviour, with significant implications for the survey's main indicators (in particular the unemployment rate expressed as a level) (Destandau *et al.*, 2015). In 2014, the survey also became continuous in the overseas territories (excluding Mayotte).

With each modification, the massive task of backcasting must be carried out to ensure the consistency of the series over a long period. For this reason, the survey evolves roughly every ten years, rather than with each edition, as part of planned overhaul work.

The backcasting following the 2003 overhaul was documented in the footnote “Correction of series breaks following the annual Employment survey’s transition in 2003 to the continuous Employment survey” (Insee, 2008).

The general principle (Figure 6.1) was to apply an additive coefficient intended to represent the mean deviation between the values in the continuous survey and in the annual survey to all the values obtained from the annual survey. For certain aggregates, a prior logarithmic transformation was applied.

Figure 6.1 General principle for the backcasting of the annual/continuous Employment survey



Source: Insee, 2008

Corrections are based on the hypothesis that each variable of interest is locally refined. The method consists in decreasing the observed values either side of the date of transition to the continuous Employment survey (for example from 2001 to 2005, with 2 annual observations in 2001 and 2002 and an average of 3 annual observations from 2003 to 2005) on a refined trend and an “annual survey” indicator.

$$VAR(t) = a + b * t + marche * (t < 2003) + résidu$$

So the corrected values prior to 2003 are:

$$VAR_{corrigée}(t) = VAR(t) - marche$$

Additional corrections are applied to ensure consistency between disaggregated series and their corresponding aggregates.

The backcasting carried out following the 2013 overhaul is documented in the footnote “The backcasting of the Insee series “Labour Market – Long Series” findings following the change of questionnaire in 2013” (Insee, 2015).

A two-stage method was adopted: the backcasting of the 2003 – 2012 quarterly series then the backcasting of the annual series prior to 2003. The backcasting of the quarterly series was divided into two main stages: firstly a backcasting of the main aggregates then a re-evaluation of new individual weightings based on the series produced during the initial stage.

Early in 2021, primarily for the purpose of responding to a new European regulation, a new survey was put into operation (Guillaumat-Tailliet *et al.*, 2021). Beyond the European requirements, the questionnaire’s overhaul also provided an opportunity to meet national expectations, take into account recent employment changes (teleworking, new forms of employment) and introduce the updated socio-professional nomenclature. In order to introduce a single measurement break, this overhaul provided an opportunity to modernise the collection protocol at a number of levels: by offering Internet as an additional response mode for the re-interviewing phases; by updating the weighting method (using the information on household earnings contained in the new sampling frame to correct non-response more

accurately); and by revising the “rules of attachment” to dwellings for the individuals⁴ determining a survey’s scope.

Since the objective was to have one sole questionnaire regardless of the collection mode, simple formulations suited to its self-administration had to be found. In particular, tools for successfully navigating the long lists of professions and qualifications were developed.

The new questionnaire was repeatedly tested: in June 2018, the preliminary questionnaire was tested face-to-face on 1,000 households; in December 2018, a 100% Internet test was conducted amongst volunteers; in 2019, a complete two-wave test (an initial interview and a follow-up interview) was conducted amongst 1,400 households, in accordance with the target protocol.

The protocol’s outline had been defined but in order to refine all the protocol’s parameters, several experiments were required: duration of Internet exclusivity, media (letter or email) to be used to communicate with households, follow-up dates, etc. In order to do this, from 2013 to 2018 the Muse (*MUltimode Sur l’enquête Emploi*, Multimode in the Employment Survey) project was set up at Insee; its purpose was to experiment with introducing Internet as an alternative collection mode to collection using an interviewer in the Employment survey (Garnero, 2019). Three types of consecutive experiments were carried out: from 2014 to 2015, qualitative experiments on small samples so as to develop a “fluid” questionnaire for the Internet version; in 2016, experiments with large samples of Internet users to test the operation of large-scale collection via Internet on Insee’s servers; and in 2017 and 2018, multimode protocol tests with several multimode Internet / telephone with an interviewer protocol variants. The insights gained from these experiments made it possible to define a survey compatible with the addition of Internet to the current collection modes.

The Muse project, piloted since March 2016 by the Employment Survey Evolution Programming team, then passed the baton to the Mélopée (*Mise En Ligne et Optimisation de l’Enquête Emploi*, Online Publishing and Survey Optimisation) project so that the new 2021 Employment Survey could be organised.

To quantify possible induced measurement breaks as accurately as possible, a large-scale pilot survey was introduced during five consecutive quarters in 2020, enabling a comparison to be made between the old and the new surveys (Insee, 2021; Guillaumat-Tailliet *et al.*, 2021). The principle of this pilot was to play out in advance what would happen when the new survey was put into production. It had therefore been built as a perfect replica of the new survey in terms of the questionnaire, the protocol and the downstream processing method.

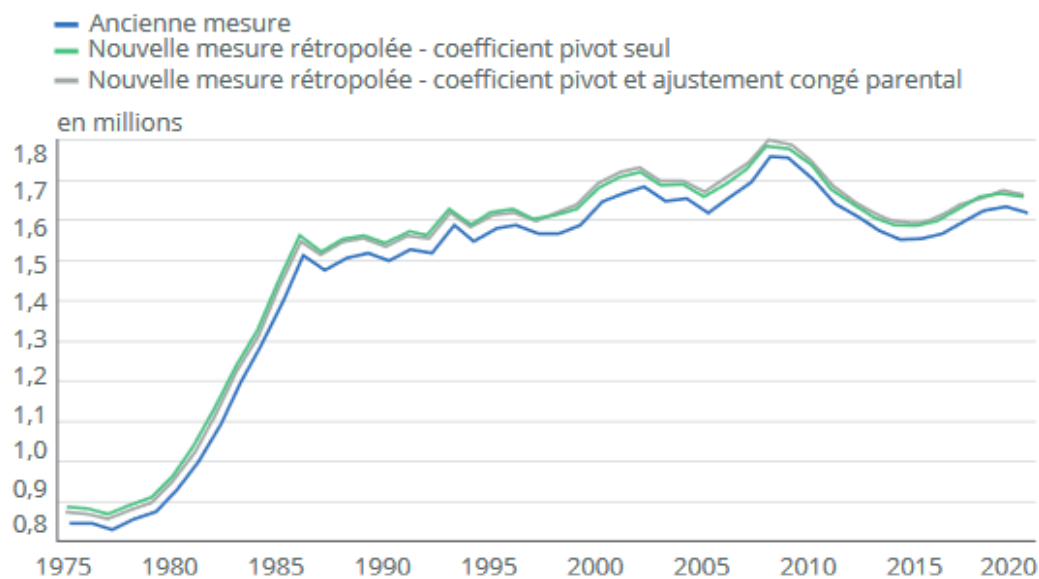
Ultimately, due to the various impacts the anticipated significant margin of error and the effects of the 2020 health crisis had on the new and the old surveys, it was deemed preferable to assess the measurement break overall by comparing the two versions of the survey without seeking to quantify the effects of each change (changes in employment concept and weighting, “proxy” effect, other effects (protocol, collection mode, etc.) or cross effects). Nevertheless, certain specific effects concerning for example conceptual changes and changes in weighting methods could be isolated (Insee, 2021). An analysis of the mode effects associated with responding via Internet, combining the more complex “measurement” effect and “selection” effects, may be the subject of future work.

In order to publish findings with methods and concepts that are constant over time, a long period of time was spent backcasting the series. Backcasting consists in adjusting past observations so that they correspond to a new measurement. For the Employment survey’s overhaul, this means adjusting the past as if the questionnaire had always been posed as in the first quarter of 2010, that the concepts used are identical, as well as the weighting method and the protocol (figure 6.2). This is a highly theoretical situation (Internet did not exist in 1975), but it is the only way to ensure the comparability of the findings over a long period of time. In practice, an average “pivot” coefficient is assessed over the

⁴ For the most part, this concerns adult students living both in student accommodation and in their parents’ homes. In the old survey, they were frequently attached to their parents’ homes. From 2021, they are interviewed for both their dwellings, whilst ensuring that they are not overrepresented.

five recovery quarters between the old and the new surveys. It measures the extent of the break over the recovery period by assuming that breaks are proportional to each year.

Figure 6.2 Employment of women aged 35–39, former measurement and new backcast



Lecture : en 2020, on comptabilisait 1 620 000 femmes de 35 à 39 ans en emploi avec l'ancienne version de l'enquête Emploi ; cette estimation a été rehaussée à 1 660 000 avec la nouvelle version de l'enquête.

Champ : France hors Mayotte, femmes vivant en logement ordinaire de 35 à 39 ans.

Source : Insee, enquête Emploi, séries longues sur le marché du travail.

measurement

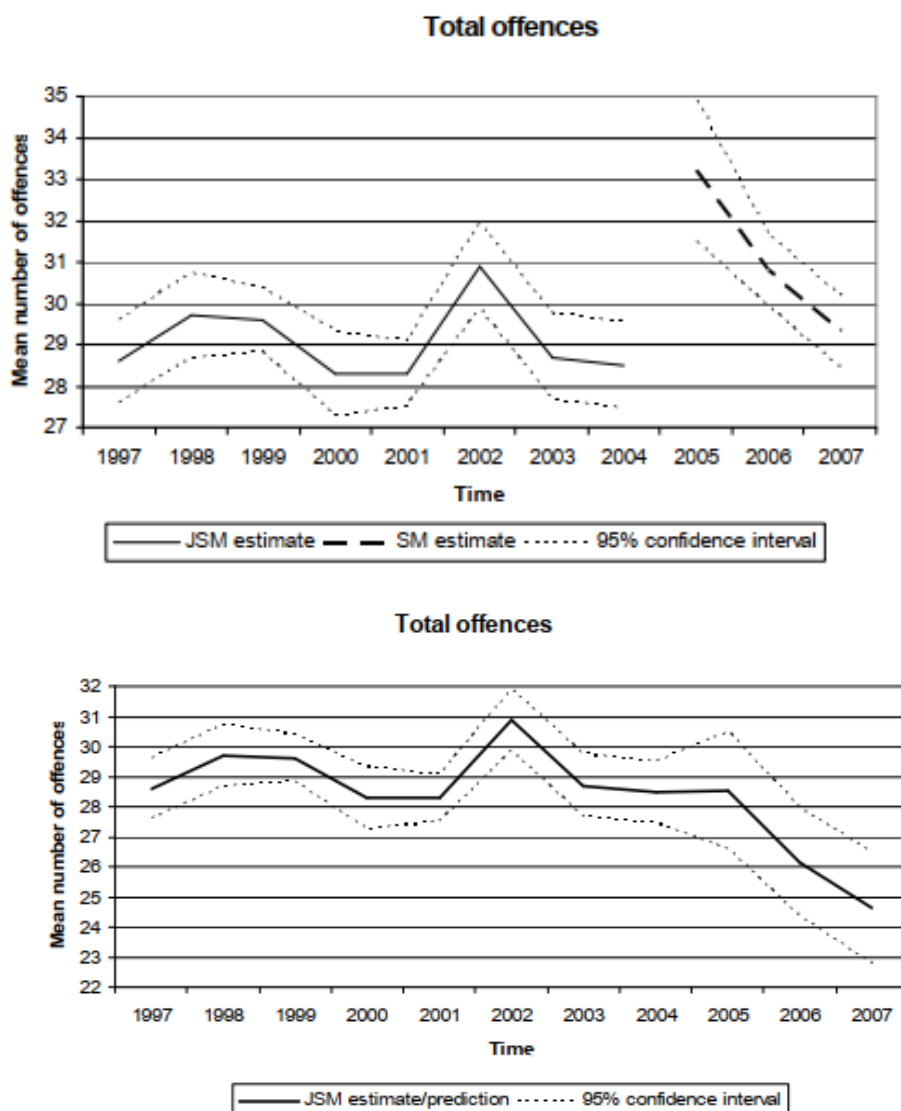
Source Insee, 2021

It emerges from this work that breaks are frequently different depending on the characteristics of the population. This is why it is important to measure and understand the breaks' determinants (sex and age in particular) and then backcast by determinant at sub-series level rather than at aggregate level.

6.1.2 The backcasting of series derived from victimisation surveys carried out in the Netherlands

In the Netherlands, until 2004, victimisation and opinions concerning insecurity and the activities of the police and the justice system were measured using the permanent survey on living conditions' *Justice and Security Module* (JSM). The collection was entirely conducted face-to-face (CAPI). For budgetary reasons, its various questions were gathered from 2005, in a special victimisation survey (*Crime Victimisation Survey*, CVS) that used CAPI and the telephone (CATI) as its collection modes. This overhaul was not the subject of the parallel processes JSM and CVS. The overhaul has had an effect on incidents of victimisation (in particular on damage to property) and the series generated by CVS were adjusted to the previous JSM levels (figure 6.3) using a purely time series approach, and using structural models (van den Brakel *et al.*, 2008). This choice (rather than using backcasting) was made to meet the needs of users at that time. The communication difficulties this may have posed at the time led the Dutch to provide for a system of parallel surveys for subsequent overhauls.

Figure 6.3: Total number of offences (per 1,000 residents) in the Netherlands, break and adjustment



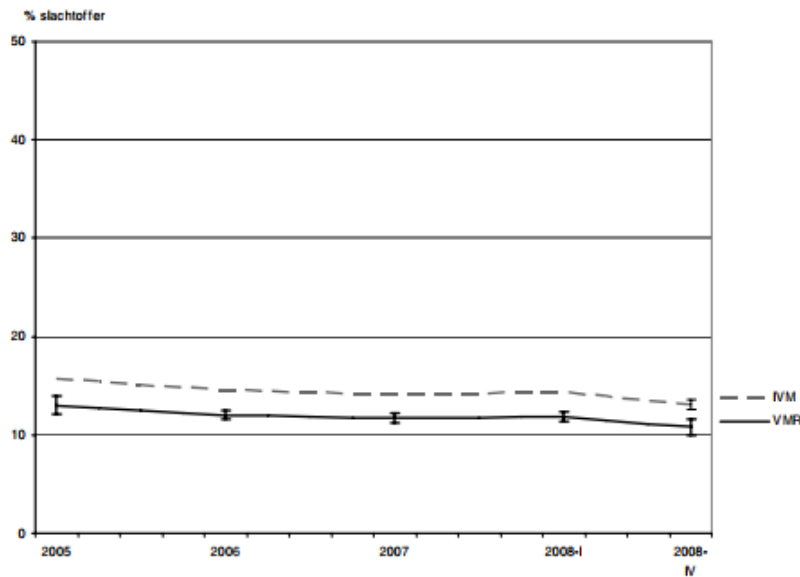
Source: van den Brakel et al. 2008

The first graph provides the series for the total number of offences assessed using JSM in 1997-2004 and the updated SM/CVS protocol for 2005-2007; the second graph gives the series for the total number of offences in 1997-2007, using JSM for 1997-2004 and SM/CVS break-corrected assessments for 2005-2007.

A significant update of the collection process took place in 2008: this mainly consisted in switching from the CAPI/CATI multimode protocol to the Internet sequential multimode protocol (CAWI) / paper (PAPI) / CATI / CAPI, with a few adjustments to the questionnaire. As in previous years, the CAPI/CATI survey was carried out under the usual conditions in the first quarter of the year 2008. And in the autumn the two protocols were conducted in parallel. This parallel operation of the two processes in 2008 enabled the series generated by the old process to be backcast using synthetic correction methods.

In the annual report *Integrale Veiligheidsmonitor 2008*, two series were provided for each indicator (figure 6.4): the VMR gives assessments directly derived from the old protocol, and the IVM series (updated protocol) gives, for the years 2005, 2006 and 2007, backcasted adjustments for IVM using VMR points 2005, 2006 and 2007 and the correction factor generated by the surveys conducted in parallel in 2008 (Cbs, 2009).

Figure 6.4: Victims of acts of vandalism (in %) in the Netherlands during the period 2005-2008



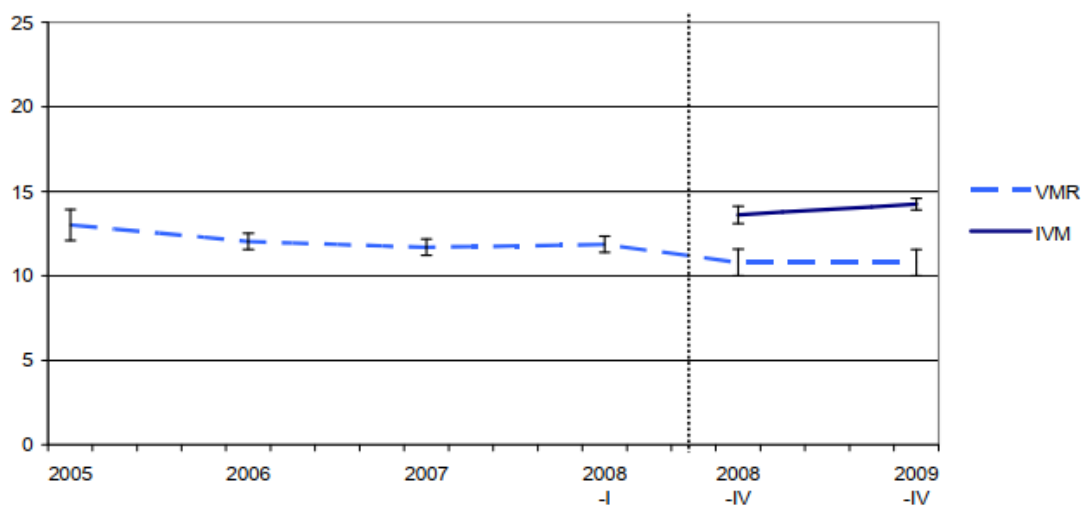
Source: *Integrale Veiligheidsmonitor 2008*, Cbs (2009)

VMR describes the old protocol, IVM the updated protocol.

Dotted line: IVM series for 2005, 2006, 2007 backcast using VMR points from 2005, 2006 and 2007 and the correction factor generated by the parallel operation of the two processes in the autumn of 2008.

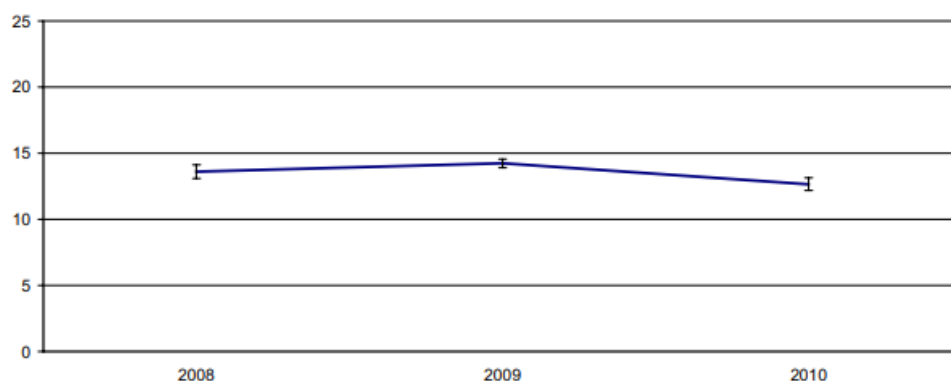
In 2009 and 2010, the old protocol was again carried out in parallel with the updated protocol to consolidate the discontinuity assessments. In the *Integrale Veiligheidsmonitor 2009* report, the indicators were presented once again via 2 series (figure 6.5): the two points directly derived from the updated IVM protocol in 2008 and 2009, and the VMR series giving the historic assessments for 2005-2008 and two points (IV) in 2008 and 2009 generated by experimental parallel surveys. By contrast, in the *Integrale Veiligheidsmonitor 2010* report, a single series is presented (figure 6.6), with solely the points 2008, 2009 and 2010, and there is no longer any reference to the names of the protocols (Cbs 2010; Cbs 2011).

Figure 6.5: Victims of acts of vandalism (in %) in the Netherlands during the period 2005-2009



Source: *Integrale Veiligheidsmonitor 2009*, Cbs (2010)

Figure 6.6: Victims of acts of vandalism (in %) in the Netherlands during the period 2008-2010



Source: *Integrale Veiligheidsmonitor 2010, Cbs (2011)*

In addition, indicators were also provided at sub-national level on the basis that the observed break at national level could be applied at local level, but caution was called for in its interpretation. Assessment work at a more local level was subsequently supplemented by discontinuity measurement approaches with small-area assessment techniques.

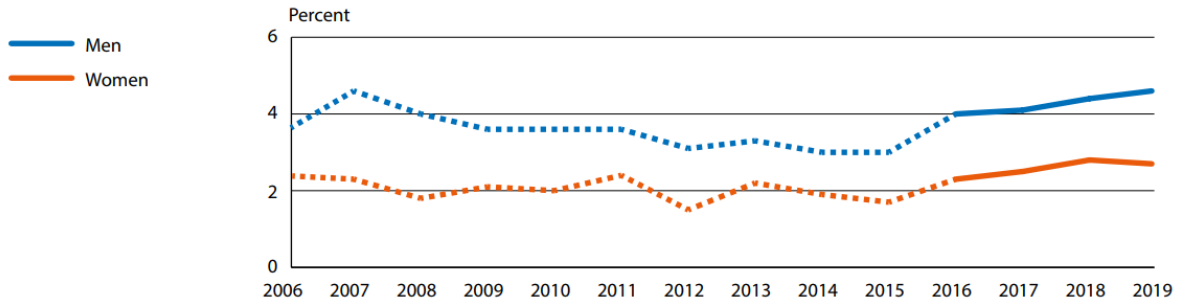
6.1.3 The backcasting of series derived from victimisation surveys carried out in Sweden

A major overhaul of the Swedish victimisation and insecurity survey system took place in 2017. It consisted in switching from the CATI protocol, in which 20,000 individuals were interviewed, to the CAWI/PAPI protocol in the initial phase, then the CATI in the second phase (victimisation monitoring module, running over 3 years), with almost 200,000 individuals interviewed in the initial phase. This evolution in the protocol was accompanied by modifications to the questionnaire: adding and deleting questions, reformulations, etc. The main reasons that led to this overhaul were the upward trend in the total non-response rate and the rising costs of the survey.

To best prepare for the transition, the two protocols were conducted in parallel in 2017. This framework thus allows the backcasting of the various series of indicators according to the standard synthetic correction method with a multiplicative approach. It therefore consists in reassessing the series generated by the old protocol by multiplying them by a binding factor that generates assessments derived from the two surveys conducted in parallel in 2017 (Brå, 2019, annexe 16). Binding factors may be specific: calculated for a question and a sub-population of interest. Hence in the 2020 report on delinquency (Brå, 2020), the disseminated series (figure 6.7) were backcast for the period 2006-2015 (the N-survey makes it possible to assess victimisation in N-1).

Figure 6.7: Victims of gender-based assault (in %) in Sweden over the period 2006-2019

Figure 3. Self-reported victimisation due to assault. Percentage for each gender, 2006–2019.^a



Source: Swedish Crime Survey 2020, English summary of Brå report (2020)

Expressed as a formula, for a given question and sub-group of interest (d), the binding factor is a prevalence ratio (p) for parallel surveys. Hence, where R is the old protocol and L the new protocol:

$$f_{d,2016} = \frac{p_{dL,2016}}{p_{dR,2016}}$$

The recalculated prevalence for N-victimisation is equal to the binding factor multiplied by the original N-prevalence.

$$p_{dR,N}^l = f_{d,2016} p_{dR,N}$$

Another possible solution was to apply calibration methods (Särndal & Lundström, 2005).

Theoretically, for an “old protocol” survey (R) from the year N (on victimisation in N-1), in order to connect it to the “new protocol” prevalence (L), a set of weights must be found:

- respecting the usual calibration totals relating to the year N survey;
- respecting the various totals of the recalculated victimisation variables (question q, modality j)

$$\hat{t}_{q=j,R,N}^l = p_{q=j,L,2016} \hat{N}_{R,N}$$

The advantage of such a method is that a single calibrated weight allows the various constraints to be integrated and may be used directly in micro-data analysis. But this requires the integration of the recalculated totals for all the variables of interest and potentially all the crosses. This may potentially create a very large number of constraints and calibration problems. Finally, this apparently theoretically attractive solution was ultimately not used.

PART III – CVS/Genese comparisons: initial elements for analysis

This part proposes a few elements for comparison between, on the one hand, assessments derived from “traditional” CVS surveys carried out between 2007 and 2019 that have already been published and, on the other, the initial raw results derived from data from the Genese survey and the CVS survey carried out in 2021. Post-collection statistical processing is currently underway for Genese and for CVS 2021; also the assessments provided in this document are purely indicative. They are liable to be modified once this processing has been completed (potential correction of mode effects, non-response correction, and margin calibration). This analysis should therefore be repeated at a later stage, on the release of the final data.

This part examines, on the one hand, the victimisation indicators for offences that figure in both surveys (prevalence, rate of reporting incidents to the police or *gendarmerie* and consistent multi-victimisation) and the main indicators for opinions and perceptions concerning the security of individuals in the 18–74 age range who reside in mainland France (“Genese” coverage). A is offered in Appendix III.

7. Victimisation indicators

Prevalence

Table 7.1 gives the annual prevalence of victimisation derived from the 2017–2019 editions of CVS as well as the raw data from Genese and CVS 2021. The CVS 2019 assessments for the year 2018 are accompanied by their accuracy at the 95 % threshold, for a greater appreciation of the significance of observed differences.

Overall, compared to the prevalence in the previous three published editions of CVS (2017-2019) covering the years 2016–2018, prevalence measured in Genese for the year 2020 appears to be fairly consistent overall and even relatively similar (within the 95 % confidence interval) for offences targeting homes (burglary, attempted burglary, theft without breaking and entering, and acts of vandalism), theft of items in or on cars, bike theft and theft without violence (figure 7.1). On the other hand, in Genese gross prevalence is much higher for car theft, theft and attempted theft with violence or threats and attempted theft without violence or threats. Conversely, gross prevalence in Genese 2020 for assaults on individuals - physical violence and sexual violence - is lower. Nevertheless, within sexual violence, the measurement of the prevalence of rape and attempted rape is identical.

The examination of gross prevalence within CVS 2021 covering the year 2020 reveals more significant discrepancies when compared to the Genese assessments (except for assaults on individuals), since all prevalence was in clear decline in 2020, the year in which the population was in lockdown for 14 weeks due to the health crisis. The gathering of incidents of victimisation in Genese did not provide for the identification of victimisation in 2019 (N-2) but made it possible to distinguish three periods in 2020: the initial lockdown (17 March–10 May), the second lockdown (29 October–14 December) and “outside lockdown”. The annual extrapolation of the prevalence of observed acts during the “outside lockdown” period in Genese is systematically higher than the annual prevalence for the whole of the year 2020, except for the theft of two-wheeled motor vehicles and scams and bank fraud. The “lockdown” effect on assessments derived from Genese over the year 2020 therefore seem fairly consistent with what is observed in CVS.

Figure 7.1: Comparison of annual prevalence measured in “traditional” CVS, CVS 2021 and in Genese

	CVS "classique"				CVS 2021 sur l'année		GENESE sur l'année 2020	
	2016	2017	2018	précision 2018 à 95%	2019	2020	Ensemble	hors confinement annualisé *
Cambriolages	1,07%	1,07%	0,83%	±0,19 pt	1,01%	0,51%	0,89%	0,97%
Tentatives de cambriolage	1,04%	0,93%	1,16%	±0,22 pt	0,99%	0,55%	1,02%	1,06%
Vols sans effraction	0,84%	0,81%	0,87%	±0,19 pt	0,58%	0,46%	0,89%	0,93%
Vandalisme contre le logement	2,46%	2,29%	2,09%	±0,30 pt	1,59%	1,91%	1,85%	1,94%
Vols de voiture (hors tentative)	0,19%	0,23%	0,15%	±0,08 pt	0,21%	0,19%	0,34%	0,36%
Vols d'objet dans ou sur la voiture (hors tentative)	2,24%	2,53%	2,17%	±0,30 pt	1,78%	1,66%	1,97%	2,03%
Vols de deux-roues à moteur (hors tentative)	0,26%	0,23%	0,14%	±0,08 pt	0,12%	0,06%	0,22%	0,20%
Vols de vélos (hors tentatives)	1,52%	1,29%	1,26%	±0,23 pt	1,56%	0,68%	1,24%	1,30%
Vols avec violences ou menace	0,17%	0,21%	0,13%	±0,07 pt	0,22%	0,19%	0,32%	0,35%
Tentatives de vol avec violences ou menaces	0,20%	0,19%	0,18%	±0,09 pt	0,18%	0,08%	0,36%	0,39%
Vols sans violences ni menaces	1,16%	1,63%	1,24%	±0,23 pt	1,02%	0,72%	1,24%	1,39%
Tentatives de vols sans violences ni menaces	0,30%	0,46%	0,64%	±0,17 pt	0,52%	0,29%	0,86%	0,96%
Arnaques ou escroqueries bancaires	ND	7,48%	8,40%	±0,58 pt	ND	ND	7,36%	6,33%
Violences physiques	1,76%	2,00%	2,15%	±0,30 pt	ND	ND	1,21%	1,24%
dont violences physiques hors ménage	1,25%	1,29%	1,43%	±0,25 pt	1,50%	0,92%	ND	ND
Violences sexuelles	0,66%	0,83%	0,49%	±0,15 pt	ND	ND	0,30%	0,36%
dont viols et tentatives de viols	0,29%	0,35%	0,22%	±0,10 pt	ND	ND	0,22%	0,26%

* prévalence qui serait observée en rapportant la période hors confinement de 38 semaines à 52 semaines.

ND : non disponible.

Champ : individus âgés de 18 à 74 ans vivant en France métropolitaine.

Lecture : d'après CVS 2021, 0,50 % des personnes âgées de 18 à 74 ans ont subi un cambriolage en 2020. Dans l'enquête Genese cette proportion atteint 0,89 %.

Source : enquêtes CVS 2017-2019 ; données brutes CVS 2021 ; données brutes Genese 2021 ; traitements SSMSI.

Reporting incidents to the police/gendarmerie

Questioning concerning the referral of incidents to the authorities varies between CVS and Genese. In CVS, for each type of offence studied, the context of the latest incident is examined and victims are initially asked whether they went to a police station or *gendarmerie* to make a statement and, if so, they are asked to specify whether they made a formal complaint, reported an incident or else abandoned the process once they had arrived. In Genese, victims must answer a more “general” question that does not specifically concern the latest incident that occurred during the year but at least one of the incidents that occurred during the year (*“Concerning this burglary (at least one of the burglaries) that occurred in 2020, did you make a statement at a police station or a gendarmerie? Yes / No”*).

A comparison of the “reporting rates” measured in the two surveys shows fairly consistent findings for the majority of offences, but there are nevertheless significant discrepancies: much higher reporting rates in Genese for acts of vandalism targeting homes and attempted theft with or without violence and, conversely, lower rates for car theft, theft of two-wheeled motor vehicles and violent robbery (figure 7.2). These discrepancies should be investigated further on the basis of the final data.

Figure 7.2: Comparison of reporting rates measured in “traditional” CVS, CVS 2021 and in Genese

	CVS "classique"			GENESE sur l'année 2020
	2016	2017	2018	
Cambriolages	79,5%	79,8%	70,0%	74,4%
Tentatives de cambriolage	30,6%	41,3%	35,4%	37,8%
Vols sans effraction	29,1% ^(b)			40,2%
Vandalisme contre le logement	16,2%	13,5%	12,9%	31,6%
Vols de voiture (hors tentative)	95,6% ^(a)			70,0%
Vols d'objet dans ou sur la voiture (hors tentative)	40,2%	39,8%	35,9%	43,3%
Vols de deux-roues à moteur (hors tentative)	88,3% ^(b)			62,9%
Vols de vélos (hors tentatives)	23,8%	26,8%	17,4%	29,2%
Vols avec violences ou menace	73,8% ^(b)			56,5%
Tentatives de vol avec violences ou menaces	11,4% ^(a)			28,9%
Vols sans violences ni menaces	46,5%	54,0%	51,2%	50,5%
Tentatives de vols sans violences ni menaces	4,4% ^(a)			15,7%
Arnaques ou escroqueries bancaires	ND	22,8%	19,0%	19,2%
Violences physiques	31,3%	25,2%	17,5%	34,2%
Violences sexuelles	16,7% ^(a)			12,7%

* taux de déclaration pour violences physiques hors ménage uniquement ; ND : non disponible

(a) moyenne sur la période 2012-2018

(b) moyenne sur la période 2016-2018

Note : pour CVS le taux de déclaration agrège le dépôt de plainte formel et la déclaration à la main courante.

Champ : individus âgés de 18 à 74 ans vivant en France métropolitaine.

Lecture : la proportion de victimes de cambriolage ayant fait une déclaration en commissariat ou en gendarmerie est de 74,4 % dans Genese 2021 contre 70,0% en 2018 d'après CVS 2019.

Source : enquêtes CVS 2017-2019 ; données brutes CVS 2021 ; données brutes Genese 2021 ; traitements SSMSI

Consistent multi-victimisation

In CVS as in Genese, for each type of offence, N-1 victims are asked to specify the number of offences of the same type that they had experienced during the year N-1. Collection is an open field in CVS. In Genese, 5 response modalities are offered: 1, 2, 3, 4, 5 or more. The proportion of victims who give 2 or more as their response to this question corresponds to the so-called consistent multi-victimisation indicator.

Figure 7.3: Comparison of consistent multi-victimisation measured in “traditional” CVS, CVS 2021 and in Genese

	CVS "classique"			GENESE sur l'année 2020
	2016	2017	2018	
Cambriolages	7,5% ^(a)			9,4%
Tentatives de cambriolage	4,1% ^(b)			17,0%
Vols sans effraction	7,7% ^(a)			13,6%
Vandalisme contre le logement	20,9%	18,8%	20,2%	28,4%
Vols de voiture (hors tentative)	2,0% ^(a)			5,7%
Vols d'objet dans ou sur la voiture (hors tentative)	7,8% ^(a)			18,3%
Vols de deux-roues à moteur (hors tentative)	2,9% ^(a)			13,7%
Vols de vélos (hors tentatives)	9,0% ^(a)			10,3%
Vols avec violences ou menace	2,9% ^(a)			20,5%
Tentatives de vol avec violences ou menaces	12,1% ^(a)			25,2%
Vols sans violences ni menaces	7,2% ^(a)			16,4%
Tentatives de vols sans violences ni menaces	11,6% ^(a)			24,0%
Arnaques ou escroqueries bancaires	21,1%*	21,17%*	20,8%*	25,6%
Violences physiques*	37,2%*	40,1%*	31,9%*	34,0%
Violences sexuelles*	46,7% ^{*(b)}			39,1%

* pour les violences physiques et sexuelles dans CVS la multivictimisation homogène est calculée sur les 24 derniers mois.

(a) moyenne sur la période 2012-2018

(b) moyenne sur la période 2016-2018

Note : la multivictimisation homogène désigne ici le fait d'avoir subi plusieurs fois des faits de même nature sur la période de référence.

Champ : individus âgés de 18 à 74 ans vivant en France métropolitaine.

Lecture : la proportion de victimes de cambriolages ayant subi deux cambriolages ou plus la même année est de 7,5% en moyenne sur la période 2016-2018 dans CVS et de 9,4% en 2020 dans Genese.

Source : enquêtes CVS 2017-2019 ; données brutes CVS 2021 ; données brutes Genese 2021 ; traitements SSMSI

Initial insights

The discrepancies between CVS and Genese in the victimisation indicators (prevalence, reporting rates and consistent multi-victimisation) suggest two sorts of problems that need to be addressed in the future survey: failure to comply with instructions on opting out and its corollary, double counting. Indeed, in Genese, the identification of vehicle thefts excluded attempted theft. The greater prevalence observed in CVS, the lower reporting rates and greater consistent multi-victimisation appear to indicate that respondents have possibly tended to include the attempted theft of vehicles in their responses despite instructions to the contrary. Concerning theft and attempted theft with violence or threats, the greater prevalence observed in Genese might be explained by a greater propensity amongst self-administering respondents to report offences already identified, in particular theft from homes and theft targeting vehicles that may have involved violence or threats (context for which gathering was not planned in Genese for theft from homes and theft targeting vehicles).

8. Indicators for perceptions and opinions concerning security

Figure 8.1: Comparison of consistent multi-victimisation measured in “traditional” CVS, and in Genese

	CVS "classique"				CVS 2021	GENESE 2021
	2017	2018	2019	précision 2019 à 95%		
Sentiment d'insécurité dans le quartier ou le village	11,3%	12,4%	11,8%	±0,67 pt	11,1%	20,5%
Sentiment d'insécurité au domicile	8,0%	8,5%	7,6%	±0,55 pt	7,1%	13,1%
Satisfaction à l'égard de l'action de la Justice	20,5%	20,3%	20,3%	±0,84 pt	21,6%	21,4%
Présence suffisante de la police et de la gendarmerie dans le quartier ou le village	49,9%	50,5%	48,7%	±1,04 pt	21,9% ^(a)	58,5%
Efficacité de la police et de la gendarmerie dans le quartier ou le village	48,8%	49,4%	47,8%	±1,04 pt	36,4%	47,4%

(a) 17,3% des individus ont répondu qu'elle est excessive contre 1 à 3% en moyenne chaque année

Champ : individus âgés de 18 à 74 ans vivant en France métropolitaine.

Lecture : la proportion de personnes âgées de 18 à 74 ans se sentant souvent ou de temps en temps en insécurité dans leur quartier ou leur village en 2020 est de 11,1 % d'après CVS 2021 et de 20,5 % d'après Genese.

Source : enquêtes CVS 2017-2019 ; données brutes CVS 2021 ; données brutes Genese 2021 ; traitements SSMSI.

The two surveys' key common indicators for perceptions and opinions concerning security are: feeling unsafe within the home and within the district or village, satisfaction regarding the operation of the justice system and opinions on the presence and effectiveness of the police or *gendarmerie* within the district or village. Overall, CVS and Genese assessments differ, with the exception of satisfaction regarding the operation of the justice system. The sense of insecurity, whether in the home or in the district or village, is twice as great in Genese. These questions, examined in the analysis of mode effects in Genese (cf. Chapter 3) demonstrate discrepancies between responses given by telephone and those given via Internet. Nevertheless assessments of the sense of insecurity obtained from amongst Genese telephone respondents remain higher than assessments in CVS.

PART IV Summary and outlook

The collection of survey data, largely multimode via Internet, addresses various challenges confronting public statistics. It may be a solution, at a sustainable cost, to the downward trend in participation via the traditional modes and be more in line with respondents' expectations. Internet also makes it possible to envisage collecting statistical information from much larger samples, for example to obtain usable data at localised levels or to improve the assessment of evolutions in short-term key indicators. This approach, the preferred method for overhauling the victimisation survey system, nevertheless raises a few crucial methodological issues to which this report attempts to provide some completely transparent answers, just a few months after the end of the Genese survey's collection. In this part, a summary review of the key insights developed in the first parts of the report will include some aspects of the outlook and road map for the future victimisation survey system.

9. Key insights

Mode effects that are generally limited as regards victimisation but more marked in questions regarding opinions and perceptions concerning security.

The use of multi-mode data collection raises fundamental issues, one of which relates to mode effects intrinsically linked to measurement: to what extent do respondents respond in the same way to a question posed in two different collection modes? In this regard, relevant quasi-experimental protocols and the availability of control variables are key to ensuring the credibility of the hypotheses assuring the comparability of the responses according to the modes. The Genese survey's system was therefore designed in advance with this in mind. It was particularly important to compare responses via Internet, the protocol's self-administered and most widely favoured mode, with telephone responses.

The questioning in phase 1 of the Genese survey thus includes a certain number of questions on victimisation that are similar to those in the historical CVS survey and destined to be repeated in the future system. For these different indicators, the findings are fairly reassuring: mode effects linked to measurement are generally limited. Thus, amongst all the indicators tested, only two have statistically significant robust effects on the various selection control methods (by regression and by *matching*): acts of vandalism (under-reporting by telephone compared to Internet) and physical violence (over-reporting by telephone). The common feature shared by these indicators is a relative ambiguity in the wording of the questions; this represents an opportunity to improve their formulation in order to limit such effects.

On the other hand, in questions concerning opinions concerning security, mode effects are more marked. These findings, which are consistent with the literature, are linked to social desirability: social interaction with an interviewer produces a certain conformity to normative expectations. Hence, telephone responses attest to greater satisfaction with the activities of the police/*gendarmérie* and the operation of the justice system, whereas a sense of insecurity is reported less frequently. Mode effects are also significant for some of the questions concerning the observation of delinquent phenomena, but are weaker.

The prevalence of victimisation derived from Genese is indeed consistent with that of the historical CVS series

It has not definitely been stated that the quantification of incidents of victimisation, using an innovative survey system that doesn't include face-to-face, is consistent with historical CVS assessments. Concerning this issue, the challenge seems to have been met: victimisation assessments derived from Genese are pretty consistent with previous years' CVS findings. These findings appear to attest to the quality of the questionnaires' design and the respondents' understanding of the concepts, whatever the response mode. These findings are also in line with conclusions regarding mode effects: although mode effects, quantified using Genese, between the telephone and Internet are weak, the traditional dichotomy (self-administration vs via the intermediary of an interviewer) in the literature between the various collection modes suggests that the face-to-face method presents no clear difference either. In addition, the success in terms of respondent participation in Genese system surveys arguably contributes to these findings. Such massive participation probably limits problems associated with non-ignorable selection that may be significant for protocols that are only able to select a certain fraction of the sampled population.

These findings can be consolidated once Genese's final processing is completed, in particular the total non-response processing and external margin calibration. Discrepancies will also need to be assessed within the current context since the recorded delinquency statistics show atypical variations in 2020 that may undermine a direct comparison. Finally, assessments derived from CVS 2021 data, even if it was not collected in the same way as usual, may also provide additional information for understanding and explaining discrepancies.

On the other hand, in questions regarding opinions concerning security, again mirroring the findings on mode effects, assessments derived from Genese differ significantly from CVS assessments.

Documenting overhauls: indispensable at the present time and in the near future, probably very useful even in the very distant future

During a major overhaul of a survey process, documentation is particularly crucial: it enables challenges to be thoroughly met, impacts to be assessed and the right decisions to be made. Certain choices that may seem obvious at the time may not necessarily be so a few years later: for the common good it is important to set aside the time and resources required. This is all the more true for systems involving surveys that are repeated over time with statistical methodology issues in terms of temporal comparability. This report thus lays the foundations for documentation on the overhaul of the victimisation survey system, and this effort will need to be maintained hereafter.

10. Outlook and road map for the future victimisation survey system

Consolidating the mode effect findings

Initial editions of the future survey system should make it possible to consolidate assessments in terms of mode effects on victimisation incidents derived from Genese data. It will be a question of renewing a protocol offering the same quasi-experimental framework for improving the assessment of effects intrinsically linked to measurement (telephone vs Internet) and therefore well purged of selection differences. This would initially be to consolidate the victimisation findings included in the Genese questioning: this additional data will be particularly appreciated as regards the rarer incidents of victimisation. This will also make it possible to quantify the mode effects for certain incidents of victimisation that were not covered in Genese: threats, abuse, acts of vandalism on cars and attempted theft from within vehicles.

Offences concerning which mode effects are estimated to be statistically significant (mainly acts of vandalism to housing and physical violence) using Genese data will require special attention during the

design phases of the future system's questionnaire. It will be a question of removing, as far as possible, any ambiguity which could result in differences in response behaviour depending on the mode.

Also, devoting part of the sample to purely face-to-face collection could provide additional information on mode effects that may explain a few differences in the assessments derived from the updated protocol and those derived from CVS.

Concerning questions regarding opinions and perceptions concerning security, the measured effects appear to be fairly robust and in line with the expectations of the literature. Additional work will help improve explanations of the underlying mechanisms of the most fundamental questions but it seems unrealistic to wish to limit these intrinsic effects at all costs.

The question of comparability over time

The major interest of a survey repeated over time is its temporal comparison dimension, including in the long term. Within the context of a major overhaul, this question of comparability over time is where several issues converge. Should we favour "correct" measurement over a given period of time, or comparability with the past? Do we need to guard against future comparability issues?

To support the major overhaul of a survey process, it is crucial to document, understand, explain and interpret potential breaks observed in indicators monitored historically over time, in advance of work on the potential connection of series. The work initiated in this report will need to be continued using **Genese** data, CVS 2021 data and data from the initial editions of the future victimisation system. This work is essential in order to avoid the communication difficulties that may arise within this context.

To ensure initial comparability over time, the dissemination of indicators generated by the future victimisation survey system's initial edition could be accompanied by historical CVS series, both raw and "temporarily" backcast, using both Genese and CVS 2021 data: CVS 2021 as the last (roughly comparable) point in the series before the break (temporal series approach), Genese for a comparison with CVS 2021 (pseudo-parallel surveys) in order to assess the break effect or as the first point in the new series (temporal series approach). Backcasting can be adjusted following each new release of each new year's data. The atypical nature of the year 2020 in terms of delinquency and the absence of a true point for 2019 (CVS 2020 was not carried out due to the health context) are liable to cause difficulties in assessing the break effect over this particular period; a certain amount of distance will no doubt be required before backcast historical series that are considered final and reliable can be delivered.

Moreover, the greater the magnitude of the break, the more the issue of the validity of the corrections and the credibility of the underlying assumptions arises. It may be necessary to identify, from amongst the questions on opinions concerning security, those which must absolutely be retained for long-term historical analysis.

Also, for an improved understanding of the challenges posed by such breaks and in order not to be limited to the sole objective of disseminating a series of indicators, work could be carried out on the preservation of correlations between key variables of interest before/after the overhaul or on breaks in the prevalence of particular sub-populations.

Correcting and tailoring mode effects

In the above, the connection/backcasting context considers the updated protocol as a whole, without specific distinctions as to collection modes.

There is no consensus in the literature on what needs to be done once the mode effects intrinsically linked to measurement have been quantified. The first findings using Genese data initially provide insights into improving questioning so as to minimise them.

If a corrective approach is preferred, the initial issue is that of defining a potential reference mode. This mode may be defined as the mode that gives the most “correct” measurement of the parameter of interest, and could theoretically vary according to the type of questions. A more pragmatic approach would be to consider the protocol’s majority mode as the reference mode. Such an approach would therefore be preferable in the case of mode effect corrections by imputations on individual data (Legleye *et al.*, 2018), on the basis of an updated protocol similar to Genese’s: since intrinsic matching can only be carried out on comparable Internet and telephone responses, thus provided in the quasi-experimental framework, this requires corrections to be introduced solely to telephone responses. A more aggregated approach to correcting mode effects would probably allow such constraints to be relaxed, but at the price of more debatable underlying hypotheses.

Another way of proceeding is to avoid the correction of mode effects as such but to ensure comparability over time for future editions of the system: the idea is to monitor changes in the proportion of respondents across the various modes in order to prevent the indicators from drifting over time for this reason and not solely due to the intrinsic evolution of the phenomena being studied (Buelens & van den Brakel, 2015 ; Buelens & van den Brakel, 2017). A natural recommendation for limiting the magnitude of such corrections over time would be to avoid substantially modifying the protocol over the course of time. Nor should the difficulties associated with such a potential drift be overestimated: if the mode effects are limited and the protocol is relatively stable overall, this is a fairly minor problem, especially if the proportions across modes are unbalanced (with one mode strongly in the majority). The relevance of the introduction of such monitoring may be anticipated on the basis of simulations using quantified mode effects.

Bibliography

Abadie, A. & Imbens, G. W. (2006). « Large sample properties of matching estimators for average treatment effect ». *Econometrica*, Vol 74, n°1, 235-2.

Abadie, A. & Imbens, G. W. (2011) « Bias-corrected Matching Estimators for Average Treatment Effect ». *American Statistical Association Journal of Business Economic Statistics*, Vol 29.

Algava, E. (2015). « Les yeux dans les yeux » Téléphone ou face à face : quel impact sur les réponses ? L'expérience de l'enquête Changement Organisationnel et Informatisation 2006 ». *Document d'études* n°196, Dares, décembre.

Amiel, M.-H. & Denoyelle, T. (2012). Enquêtes en ligne : comparaison de modes de questionnement sur le thème du logement. *XIe Journées de Méthodologie Statistique*, Insee, Paris, 24-26 janvier. http://jms.insee.fr/files/documents/2012/945_4-JMS2012_S26-3_AMIEL-ACTE.PDF.

Andersen, R., Kasper, J. D. & Frankel, M. R. (1979). *Total survey error*. Jossey-Bass.

Barret, C. & Cissé, M. (2018). Agrégation de données multimode : impact sur la modalisation des variables présentant un effet de mesure. *XIIIe Journées de Méthodologie Statistique*, Insee, Paris, 12-14 juin. http://www.jms-insee.fr/2018/S22_3_ACTEv2_BARRET_JMS2018.pdf

Beck, F. & de Peretti, G. (2017). La collecte multimode à l'Insee : d'un effet de mode à un effet sur nos travaux. *Séminaire de Méthodologie Statistique de l'Insee*.

Beque, M., Kingsada, A. & Mauroux, A. (2019). Contraintes physiques et intensité du travail. Dares, *Synthèse Stat'*, n°25. https://dares.travail-emploi.gouv.fr/sites/default/files/pdf/synthese_stat_no24.pdf

Buelens, B. & van den Brakel, J. A. (2015). Measurement error calibration in mixed-mode sample surveys. *Sociological Methods & Research*, 44(3), 391-426.

Buelens, B. & van den Brakel, J. A. (2017). Comparing two inferential approaches to handling measurement error in mixed-mode surveys. *Journal of official statistics*, 33(2), 513-531.

Burricand, C., Houdré, C. & Vallet L.-A. (2014). « Introduction : EU-SILC / SRCV, ambition et apport d'une approche longitudinale et transversale des revenus et des conditions de vie, en France et en Europe ». Insee, *Economie et Statistique*, n°469-470. <https://www.insee.fr/fr/statistiques/1377972?sommaire=1377989>

Burtless, G. (1995). «The Case for Randomized Field Trials in Economic and Policy Research ». *Journal of Economic Perspectives*, Vol 9, n°2, pp. 63-84. <https://www.aeaweb.org/articles?id=10.1257/jep.9.2.63>

Brå (2019), *Nationella trygghetsundersökningen 2018*, Teknisk rapport, Rapport 2019:2.

Brå (2020), *Swedish Crime Survey 2020*, English summary of Brå report 2020:8.

Castell, L., Clerc, M. & Legleye, S. (2021). Victimations déclarées et effets de mode : l'expérimentation CVS Panel. *11^e colloque francophone Sondages 2021, Session 19 : Méthodes et modes de collecte/ Enquêtes multi-mode*, Bruxelles, 6-8 octobre.

Castell, L. & Sillard, P. (2021). « Le traitement du biais de sélection endogène dans les enquêtes auprès des ménages par modèle de Heckman ». *Document de travail*, Insee.

Cbs (2009), *Integrale Veiligheidsmonitor 2008*.

Cbs (2010), *Integrale Veiligheidsmonitor 2009*.

Cbs (2011), *Integrale Veiligheidsmonitor 2010*.

Couper, M. P. (2011). The future of modes of data collection. *Public Opinion Quarterly*, 75(5), 889-908.

De Leeuw, E. D. (2018). Mixed-mode: Past, present, and future. In *Survey Research Methods* (Vol. 12, No. 2, pp. 75-89).

de Peretti, G. & Razafindranovona, T. (2014), « Les enquêtes multimode : attention aux effets de mode », *Statistique et Société*, vol.2 No 2.

Destandau, S. & Le Minez, S. (2015). Introduction d'internet dans les modes de collecte de l'enquête Emploi : enjeux, questions et premiers éléments de réponse. *XIIe Journées de Méthodologie Statistique*, Insee, Paris, 31 mars-2 avril. http://www.jms-insee.fr/2015/S24_2_ACTE_DESTANDAU_JMS2015.PDF

Dillman, D. A., Smyth, J. D. & Christian, L. M. (2014). Internet, phone, mail, and mixed-mode surveys: the tailored design method. John Wiley & Sons.

Dillman, D. & Edwards, M. (2016). Designing a mixed mode survey. *The sage handbook of survey methodology*, 255-268.

Durbin, J. & Koopman, S. J. (2012). *Time series analysis by state space methods*. Oxford university press.

Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of applied Psychology*, 37(2), 90.

Estival, A. & Filatriau, O. (2019). La mesure statistique de la délinquance, AJ Pénal, n°224, Dalloz.

Garnero, M. (2019). « Le projet Muse : 5 ans d'expérimentations pour préparer l'introduction d'Internet dans l'enquête Emploi ». *Insee, Documents de travail*, n°F1907. <file:///C:/Users/0354044/AppData/Local/Temp/F1907.pdf>

Givord, P. (2014). Méthodes économétriques pour l'évaluation de politiques publiques. *Economie prévision*, (1), 1-28.

Givord, P. (2003). « Une nouvelle Enquête Emploi ». *Economie et statistique*, n°362. <https://www.insee.fr/fr/statistiques/fichier/1376196/es362d.pdf>

Gottfredson, M. R. & Hindelang, M. J. (1977). A consideration of telescoping and memory decay biases in victimization surveys. *Journal of Criminal Justice*, 5(3), 205-216.

Goux, D. (2003). « Une histoire de l'Enquête Emploi ». *Economie et statistique*, n°362. <https://www.insee.fr/fr/statistiques/fichier/1376194/es362c.pdf>

Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E. & Tourangeau, R. (2011). *Survey methodology* (Vol. 561). John Wiley & Sons.

Guillaumat-Tailliet, F. & Tavan, C. (2021). « Une nouvelle enquête Emploi en 2021. Entre impératif européen et volonté de modernisation ». *Courrier des statistiques*, n° 6. <https://www.insee.fr/fr/statistiques/fichier/5398695/courstat-6.pdf>

Haziza, D. & Lesage, É. (2016). A discussion of weighting procedures for unit nonresponse. *Journal of Official Statistics*, 32(1), 129.

IAURIF (2020). *Victimation et sentiment d'insécurité en Île-de-France. Rapport final de l'enquête 2019*. https://www.institutparisregion.fr/fileadmin/NewEtudes/000pack2/Etude_2466/rapport_v12.pdf

Imbens, G. W. & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Insee (2008), « Correction des ruptures de série suite au passage en 2003 de l'enquête Emploi annuelle à l'enquête Emploi en continu ».

Insee (2015), « Rétropolation de séries de l'Insee Résultats « Marché du Travail - Séries longues » suite au changement de questionnaire en 2013 ».

Insee (2019). « L'usage des technologies de l'information et de la communication par les ménages entre 2009 et 2018 ». Enquêtes sur les TIC auprès des ménages, *Insee Résultats*, mars. <https://www.insee.fr/fr/statistiques/3743327?sommaire=3743360>

Insee, Département de l'emploi et des revenus d'activité (2021). « L'enquête Emploi se rénove en 2021 : des raisons de sa refonte aux impacts sur la mesure de l'emploi et du chômage ». *Insee Analyses*, n°65. Juin.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied cognitive psychology*, 5(3), 213-236.

Lagrange, H., Pottier, M.-L., Zauberman, R. & Robert, P. (2004). Enquêtes de victimation et statistiques de police : les difficultés d'une comparaison, *Déviance et Société*, 28, 3, 285-316.

Legleye, S., Razafindranovona, T. & de Peretti G. (2018). Agréger les échantillons d'une enquête multimode en limitant l'effet de mesure : une proposition d'imputation raisonnable et pragmatique. *10^e colloque francophone Sondages 2018, Session 19 : Multi-modes et Collecte*, Lyon, 24-26 octobre. http://sondages2018.sfds.asso.fr/assets/Documents/Session19_MultiModes_P1_Legleye.pdf

Miceli, L., Névanen, S., Robert, Ph. & Zauberman, R. (2009). De l'instantané au long métrage. L'enquête Cadre de Vie et Sécurité dans la série des données sur la victimation, *Economie & Statistique*, 426, 3-28.

Neumayr, J., Schork, J. & Osier, G. (2018). Les effets du mode de collecte des données sur la mesure de l'emploi : une comparaison entre le web et le téléphone. *10^e colloque francophone Sondages 2018, Session 19 : Multi-modes et Collecte*, Lyon, 24-26 octobre. http://sondages2018.sfds.asso.fr/assets/Documents/Session19_MultiModes_P2_Osier.pptx

Olson, K., Wagner, J. & Anderson, R. (2020). Survey Costs: Where are We and What is The Way Forward ?, *Journal of Survey Statistics and Methodology*.

Rao, J. N. & Molina, I. (2015). *Small area estimation*. John Wiley & Sons.

Razafindranovona, T. (2015). La collecte multimode et le paradigme de l'erreur d'enquête totale. *Insee, Document de travail de la DMCSI, M2015-01*, http://www.insee.fr/fr/publications-et-services/docs/doc_travail/la%20collecte%20multimode-b.pdf

Razafindranovona, T. (2016a). Exploitation de l'enquête expérimentale Vols, violence et sécurité. *Insee, Document de travail de la DMCSI, M2016-03*.

Razafindranovona, T. (2016b). Exploitation de l'enquête expérimentale Logement internet/papier. *Insee, Document de travail de la DMCSI, M2016-08*.

Razafindranovona, T. (2017). Exploitation de l'enquête expérimentale Qualité de vie au travail. *Insee, Document de travail de la DMCSI, M2017-01*.

Robert, Ph., Zauberman, R., Pottier, M. L. & Lagrange, H. (1999). Mesurer le crime. Entre statistiques de police et enquêtes de victimation (1985-1995), *Revue française de sociologie*, XL, 2, 255-294.

Robert, P., Zauberman, R., Nevanen S. & Didier E. (2006). Enquêtes locales 2005 sur la victimation et l'insécurité. *Cesdip, Études et données pénales*, n°103.

Robert, P. & Zauberman, R. (2011a). Mesurer la délinquance, Paris, *Presses de Sciences Po*.

Robert, P., Zauberman, R., Névanen S. & Didier E. (2008). L'évolution de la délinquance d'après les enquêtes de victimation, France 1984-2005, *Déviance et Société*, 32, 4, 335-371.

Robert, P. & Zauberman, R. (2011b). Le développement et les usages des enquêtes de victimation en France. *Économie et Statistique*, 448-449, 79-83.

Rosenbaum, P. & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects, *Biometrika*, pp41-55.

- Särndal, C. E. & Lundström, S. (2005). *Estimation in surveys with nonresponse*. John Wiley & Sons.
- Schork, J., Riillo, C. A. & Neumayr, J. (2021). Survey Mode Effects on Objective and Subjective Questions: Evidence from the Labour Force Survey. *Journal of Official Statistics (JOS)*, 37(1).
- Schouten, B., van den Brakel, J., Buelens, B., Giesen, D., Luiten, A. & Meertens, V. (2021). *Mixed-Mode Official Surveys: Design and Analysis*. CRC Press.
- Senaux, M. (2018). Evolution de la joignabilité dans les enquêtes téléphoniques. 10^e colloque francophone Sondages 2018, Session 19 : Multi-modes et Collecte. http://sondages2018.sfds.asso.fr/assets/Documents/Session19_MultiModes_P4_Senaux.pdf
- Sillard, P., Faivre, S., Paliod, N. & Vincent, L. (2020) Pour les enquêtes auprès des ménages, l'Insee rénove ses échantillons. *Courrier des Statistiques*, N4 2020.
- SSMSI (2021). Insécurité et délinquance en 2020: bilan statistique.
- Sudman, S. & Bradburn, N. M. (1973). Effects of time and memory factors on response in surveys. *Journal of the American Statistical Association*, 68(344), 805-815
- Tourangeau, R., Conrad, F. G. & Couper, M. P. (2013). *The science of web surveys*. Oxford University Press.
- Tourangeau, R. (2017). Mixing modes. Tradeoffs among coverage, nonresponse and measurement error. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, LE Lyberg, NC Tucker, & BT West (Eds.), *Total Survey Error in Practice*. New York : Wiley.
- Tourangeau, R., Maitland, A., Rivero, G., Sun, H., Williams, D. & Yan, T. (2017). Web surveys by smartphone and tablets: Effects on survey responses. *Public Opinion Quarterly*, 81(4), 896-929.
- van den Brakel, J., Smith, P. A. & Compton, S. (2008). Quality procedures for survey transitions-experiments, time series and discontinuities. In *Survey Research Methods* (Vol. 2, No. 3, pp. 123-141).
- van den Brakel, J. A. & Krieg, S. (2016). Small area estimation with state space common factor models for rotating panels. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 763-791.
- van den Brakel, J. A., Zhang, X. & Tam, S. (2017). Measuring discontinuities due to survey process redesigns. *Discussion Paper Cbs*.
- van den Brakel, J., Zhang, X. & Tam, S. M. (2020). Measuring discontinuities in time series obtained with repeated sample surveys. *International Statistical Review*, 88(1), 155-175.
- van den Brakel, J. A. & Boonstra, H. J. (2021). Estimation of domain discontinuities using Hierarchical Bayesian Fay-Herriot models. *Survey Methodology*, 47(1), 151-190.
- Vannieuwenhuyze, J. T., Loosveldt, G. & Molenberghs, G. (2014). Evaluating mode effects in mixed-mode survey data using covariate adjustment models. *Journal of Official Statistics*, 30(1), 1-21.
- Zauberman, R., Robert, P., Névanen, S. & Didier, E. (2009). L'acteur et la mesure : le comptage de la délinquance. *Revue française de sociologie*, 50(1), pp31-62.
- Zauberman, R. (2015). Les enquêtes de victimation. *Idées économiques et sociales*, (3), 8-21.
- Zauberman, R., Robert Ph., Perez-Diaz, C. & Lévy, R. (1990). Les victimes, comportements et attitudes. Enquêtes nationales de victimation. Déviance et contrôle social, Paris, CESDIP, 2 volumes.
- Zauberman, R. & Robert, Ph. (1995). *Du côté des victimes : un autre regard sur la délinquance*, Paris, L'Harmattan.

APPENDICES

APPENDIX I. Presentation of the Living Environment and Security survey

The “Living Environment and Security” (CVS) survey, the so-called “victimisation” survey, has been conducted each year since 2007 by the French National Institute of Statistics and Economic Studies (Insee), a project it jointly managed with the French National Observatory of Delinquency and Penal Responses (ONDRP, shut down in 2020) and the (French Ministerial Statistics Department for Internal Security, SSMSI, created in 2014). The survey’s latest edition was conducted in 2021.

The survey’s purpose is to identify incidents of delinquency of which households and individuals may have been the victims during the two years preceding the survey, whether or not they resulted in statements being made at police stations or *gendarmeries*, but also to gather the general population’s (victims and non-victims) opinions concerning their living environment and security, and analyse the sense of insecurity and the levels of satisfaction with the operation of the justice system and the security forces.

Each of the survey’s editions benefited from the compulsory General Interest and Statistical Quality Label issued by the CNIS (French National Council for Statistical Information).

Until 2014, the survey was conducted over the 3 months between mid-January and mid-April. From 2015–2018, the collection was carried out between early February and late April/early May. In 2019, the collection was carried out over the 2.5 months between early February and mid-April. In 2020, the collection was scheduled for 20 April–11 July in order to leave room at the beginning of the year for the collection of the Statistics on Income and Living Conditions (SRCV) survey, and in order to comply with the European regulations governing this operation. In 2021, the latest edition of the survey was conducted between 19 April and 26 June 2021.

Selected households are contacted by means of a notification letter presenting the survey and the modalities of its administration, accompanied by a leaflet briefly presenting the survey and the Institute’s missions. The CVS survey’s collection takes the form of a computerised (“*capied*”) questionnaire completed face-to-face with one of the Insee network’s interviewers. In their notification letters, households are asked to book an appointment for the interviewer’s visit, but in fact it is mainly up to the interviewers to remind households to actually book a date for their visit.

The survey is based on 4 questionnaires which are administered in succession. Their structure and content have a common basis throughout the editions. However, certain modules could be added or deleted between 2007 and 2020 (public transport, scams, corruption, discrimination, etc.). In 2019, the CVS survey’s questionnaire was structured in the following way:

1. the “Common Household Core” (TCM), administered in the vast majority of Insee’s face-to-face household surveys and more generally within public statistics, provides a wealth of socio-demographic information on households and their members. Preferably, the household’s reference individual responds to the TCM (average time taken to administer the TCM: 15 mins).
2. the household questionnaire (QM) to which the household’s reference individual responds face-to-face. This allows the gathering of incidents of victimisation that targeted household property. (Average time taken to administer the QM: 8 mins);
3. the individual questionnaire (QI), administered face-to-face, is intended for one individual per household drawn from amongst the so-called eligible individuals aged 15 + on 31 December of the year in which the survey is collected. The individual concerned (the “kish”) is automatically designated by computer after the table listing the dwelling’s residents is entered (the individual selected is the person whose birthday falls closest to 1st January of the year in which the survey is carried out). This individual is obliged to respond: proxies are not allowed but a second visit is possible should he/she not be present at the time of the draw. This questionnaire allows the gathering of incidents of “personal” victimisation outside of “sensitive” violence, i.e. within the

meaning of the survey, assaults on the “kish” individual selected from within the household. (Average time taken to administer the QI: 15 mins).

4. the individual questionnaire self-administered through headphones (QAA) for so-called “kish” individuals interviewed face-to-face (average time taken to administer the QAA: 9 mins). It covers so-called sensitive violence (sexual and domestic). It is intended for individuals drawn for the “individual” face-to-face questionnaire, on condition that such individuals are aged at least 18 on the day of the survey or 75 or under on 1st January of the year in which the survey is carried out. The QAA is available in French and in 4 other languages (Arabic, Turkish, Portuguese, English and German).

The response time obviously varies according to the events experienced by the respondents: if there has been no victimisation, the administration of the questionnaire takes 40 minutes on average. In cases where there have been multiple incidents of victimisation, this may take over an hour.

APPENDIX II. Presentation of the Genese survey

The Genese (Gender and Security) survey is a new European-initiative victimisation survey. The French Ministerial Statistics Department for Internal Security (SSMSI) was very much in demand on the subject of violence against women; it applied for, and in November 2019 obtained, EU funding to conduct the national Genese survey on gender-based violence in 2021, which was based on the *Gender-Based Violence* questionnaire developed by Eurostat.

The survey obtained the compulsory General Interest and Statistical Quality Label issued by the French National Council for Statistical Information (CNIS).

This survey was designed to fulfil two aspirations: i/ inform public debate at European level with new data on the trajectories of gender-based and sexual violence; ii/ carry out a full-scale multimode experiment to measure victimisation and perceptions concerning security with a view to overhauling the “Living Environment and Security” survey by 2022.

Measuring victimisation posed two major challenges associated with: i/ the low incidence of measured offences, in particular the most serious violence, which requires significant sample sizes; ii/ the potentially sensitive nature of the subjects covered, which makes balancing the advantages and disadvantages of the various collection modes particularly complicated.

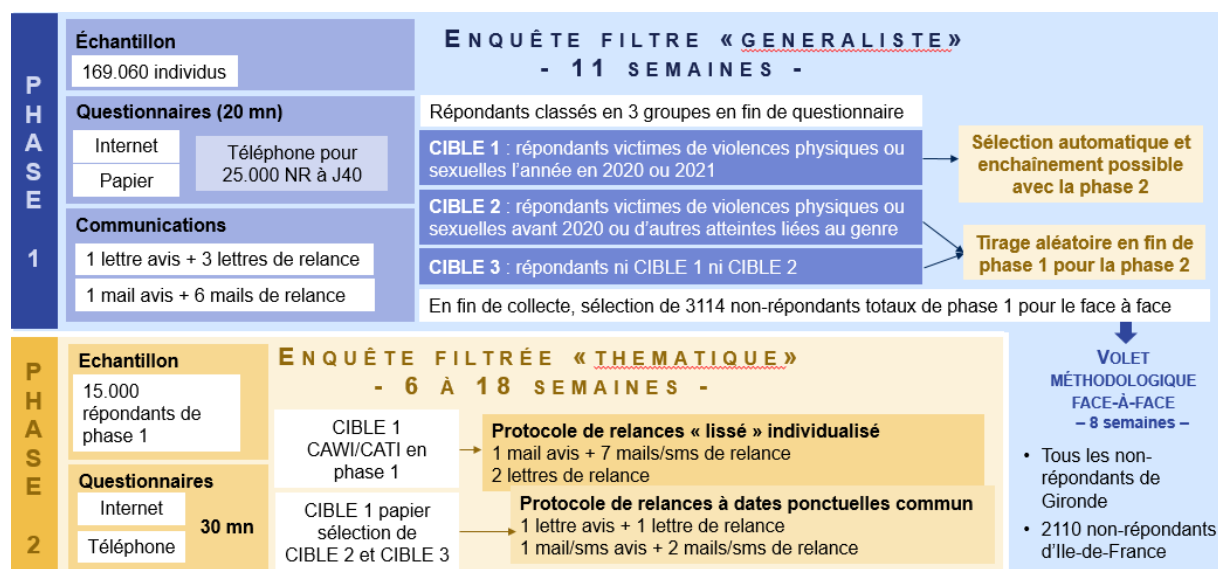
Drawing on existing surveys such as the *Vie quotidienne et santé* (Daily Life and Health, VQS) survey and in conjunction with Insee, the SSMSI has designed a strictly nominative two-phase multimode protocol.

The initial phase consisted of a short “general-purpose” victimisation survey (20 mins on average) based on gathering socio-demographic characteristics, and questions identifying offences and perceptions and opinions concerning security, modelled on what features in the *Living Environment and Security* survey. It interviewed a sample of 169,060 individuals who were aged 18–74 on 1st January 2021, living in mainland France and residing in ordinary housing. These individuals were drawn from fiscal files (Fidéli) by Insee. This was deferred concurrent collection, carried out by a polling institute: it began via Internet then provided telephone follow-up amongst a random selection of 25,000 non-respondents and sent paper self-administered questionnaires at various stages of the collection, depending on the contact details available for each respondent. The initial phase of the survey was conducted over the 11 weeks between 1st March and 16 May 2021. In total, almost 109,000 individuals (64 %) responded to the initial phase: 51 % via Internet, 3 % by telephone and 10 % on paper.

The second phase consisted of a longer thematic follow-up survey (30 mins on average) based on the questionnaire designed by Eurostat, which allows very detailed information to be gathered on lifelong gender-based and sexual violence. It was offered via Internet and telephone to a sample of 15,000 Phase 1 respondents in which victims of physical and sexual violence were over-represented. The collection's protocol was designed to target the "most qualified" Internet and telephone respondents (Target 1) so that they would all be automatically selected for the second phase, be able to complete the two questionnaires one after the other and be followed up according to a smooth and personalised follow-up protocol before the start of the collection amongst the other "less qualified" respondents (Target 2 and Target 3) who, at the end of the first phase of the collection, were drawn at random to participate in the survey's second phase. The second phase of the survey was conducted between 1st March and 10 July 2021 for Target 1 respondents and between 27 May and 10 July 2021 for Target 2 and 3 respondents. Over 10,000 individuals (68 %) responded: 53 % via Internet and 15 % by telephone.

Finally, at the end of Phase 1, a face-to-face methodological component was carried out between 1st June and 24 July 2021 amongst over 3,000 individuals: the entirety of the total Gironde non-respondents and a random selection from the total Île-de-France non-respondents. During this component, only the questionnaire from the initial phase was administered. It enabled over 1,000 questionnaires to be collected.

Genese survey protocol: two multimode phases of the screening survey/ follow-up type survey completed face-to-face



APPENDIX III. Table comparing the formulation of questions on offences in CVS and Genese

N-1 victimisation and perceptions / opinions indicators	CVS 2021	GENESE 2021
Burglary	Combination of 2 questions: <i>In [N-2] or [N-1], were you the victim of a burglary or attempted burglary in your home? Yes / No</i> + <i>How many times (excluding attempted burglaries) was your main residence burgled in N-1?</i>	<i>Within your lifetime, have you been subjected to one of the following offences in your home?</i> - <i>A burglary?</i> Yes in N / Yes in N-1 / Yes before N-1 / No, never
Attempted burglary	Combination of 2 questions: <i>In [N-2] or [N-1], were you the victim of a burglary or attempted burglary in your home? Yes / No</i> + <i>How many times have you experienced attempted burglaries at your main residence in N-1?</i>	<i>Within your lifetime, have you experienced one of the following offences in your home?</i> - <i>An attempted burglary?</i> Yes in N / Yes in N-1 / Yes before N-1 / No, never
Theft without breaking and entering	Combination of 2 questions: <i>In [N-2] or [N-1], were you the victim of a theft without breaking and entering in your home (including outbuildings)? Yes / No</i> + <i>How many times have you experienced theft without breaking and entering at your main residence in N-1?</i>	<i>Within your lifetime, have you experienced one of the following offences in your home?</i> - <i>theft without breaking and entering?</i> Yes in N / Yes in N-1 / Yes before N-1 / No, never
Acts of vandalism to your home	Combination of 2 questions: <i>In [N-2] or [N-1], apart from the thefts already covered, were you the victim of acts of wilful destruction or damage to your main residence? Yes/ No</i> + <i>How many acts of wilful destruction or damage to your home occurred in N-1?</i>	<i>Within your lifetime, have you experienced one of the following offences in your home?</i> - <i>an act of vandalism?</i> Yes in N / Yes in N-1 / Yes before N-1 / No, never
Car theft	Combination of 2 questions: <i>In [N-2] or [N-1], were you the victim of car theft or attempted car theft? Yes / No</i> + <i>How many times has a car been stolen from you (excluding attempted car theft) in N-1?</i>	Combination of 2 questions: <i>Within your lifetime, have vehicles belonging to your household or items within or on your vehicle ever been stolen (cars, two-wheeled motor vehicles or bikes)? Yes / No / You have never owned a car or a two-wheeled vehicle</i> + <i>Which of the following types of theft have you or someone within your household experienced:</i> - <i>car theft?</i> Yes in N / Yes in N-1 / Yes before N-1 / No, never
Theft from within/the exterior of cars	Combination of 2 questions: <i>In [N-2] or [N-1], were you the victim of the theft or attempted theft of items, accessories, or parts located within or on the exterior of your car? Yes / No</i> + <i>How many times have items, accessories or parts within or on the exterior of your car (excluding attempted thefts), been stolen from you in N-1?</i>	Combination of 2 questions: <i>Within your lifetime, have vehicles belonging to your household or items within or on the exterior of your vehicle ever been stolen (cars, two-wheeled motor vehicles or bikes)?</i> Yes / No / You have never owned a car or a two-wheeled vehicle + <i>Which of the following types of theft have you or someone within your household experienced:</i> - <i>theft of items or accessories within or on the exterior of your car?</i> Yes in N / Yes in N-1 / Yes before N-1 / No, never
Theft of two-wheeled motor vehicles	Combination of 2 questions: <i>In [N-2] or [N-1], were you the victim of the theft or attempted theft of a two-wheeled motor vehicle (motorbike, scooter, moped) belonging to yourself? Yes/ No</i> +	Combination of 2 questions <i>Within your lifetime, have vehicles belonging to your household or items within or on the exterior of your vehicles ever been stolen (cars, two-wheeled motor vehicles or bikes)?</i> Yes / No / You have never owned a car or a two-wheeled vehicle

	How many times have two-wheeled motor vehicles been stolen from you (excluding attempted thefts) in N-1?	+ Which of the following types of theft have you or someone within your household experienced: - theft of two-wheeled motor vehicles? Yes / No / You have never owned a car or a two-wheeled vehicle
Bike theft	Combination of 2 questions: In [N-2] or [N-1], were you the victim of the theft or attempted theft of your bike? Yes/ No + How many times has a bike been stolen from you (excluding attempted thefts) in N-1?	Combination of 2 questions: Within your lifetime, have vehicles belonging to your household or items within or on the exterior of your vehicle ever been stolen (cars, two-wheeled motor vehicles or bikes)? Yes / No / You have never owned a car or a two-wheeled vehicle + Which of the following types of theft have you or someone within your household experienced: - bike theft? Yes in N / Yes in N-1 / Yes before N-1 / No, never
Theft with physical violence or threats	Combination of 2 questions: In [N-2] or [N-1], were you personally the victim of theft or attempted theft with physical violence or threats? Yes / No + How many thefts of this type have you experienced in N-1?	Within your lifetime, apart from the thefts already covered, have you personally experienced one of the following offences: - theft with physical violence or threats? Yes in N / Yes in N-1 / Yes before N-1 / No, never
Attempted theft with physical violence or threats	Combination of 2 questions: In [N-2] or [N-1], were you personally the victim of theft or attempted theft with physical violence or threats? Yes / No + How many attempted thefts have you experienced in N-1?	Within your lifetime, apart from the thefts already covered, have you personally experienced one of the following offences: - attempted theft with physical violence or threats? Yes in N / Yes in N-1 / Yes before N-1 / No, never
Theft without physical violence or threats	Combination of 2 questions: In [N-2] or [N-1] again, were you personally the victim of theft or attempted theft without physical violence or threats? Yes / No + How many thefts of this type have you experienced in N-1?	Within your lifetime, apart from the thefts already covered, have you personally experienced one of the following offences: - theft without physical violence or threats? Yes in N / Yes in N-1 / Yes before N-1 / No, never
Attempted theft without physical violence or threats	Combination of 2 questions: In [N-2] or [N-1] again, were you personally the victim of theft or attempted theft without physical violence or threats? Yes / No + How many attempted thefts have you experienced in N-1?	Within your lifetime, apart from the thefts already covered, have you personally experienced one of the following offences: - attempted theft without physical violence or threats? Yes in N / Yes in N-1 / Yes before N-1 / No, never
Scams or fraud	Combination of X questions: In [N-2] or in [N-1] was money ever fraudulently debited from one of your bank accounts? Yes / No + How many times was money fraudulently debited in N-1? + In [N-2] or [N-1], as a consumer or user, were you the victim of a scam or an attempted scam (attempted consumer fraud, fraud)? Yes and at least one scam cost me money / I was the victim of at least one attempted scam but it didn't cost me anything / No + How many times were you the victim of a scam that cost you money in N-1?	Within your lifetime, apart from the thefts already covered, have you personally experienced one of the following offences: - bank fraud or a scam aimed at stealing money from you? Yes in N / Yes in N-1 / Yes before N-1 / No, never
Physical violence	Combination of 4 questions: In [N-2] or [N-1] again, were you personally the victim of physical violence on the part of an	Within your lifetime, has anyone ever: - slapped you, hit you, beat you or intentionally subjected you to any

	<p>individual who does not currently live within the same accommodation as you? (incl. on the part of an ex-spouse/partner or spouse/partner who no longer lives with you, excluding sexual violence) Yes / No</p> <p>+</p> <p>How many times were you the victim of acts of physical violence in N-1?</p> <p>+</p> <p>In [N-2] or [N-1], apart from sexual violence, did anyone who currently lives with you ever slap you, hit you, beat you or subject you to any other form of physical violence? Yes / No</p> <p>+</p> <p>Did this incident (at least one of these incidents) take place in [N-1]? Yes / No</p>	<p>other form of physical violence, apart from acts of a sexual nature?</p> <p>Yes in N / Yes in N-1 / Yes before N-1 / No, never</p>
Sexual violence	<p>Combination of 4 questions:</p> <p>In [N-2] or in [N-1], [apart from individuals who currently live with you if NHAB>1], has anyone ever subjected you to fondling or unwanted sexual intercourse, or attempted to do so using violence, threats, coercion or surprise? Yes / No</p> <p>+</p> <p>Did this assault take place in N-1? Yes / No</p> <p>+</p> <p>Outside these violent episodes, in [N-2] or [N-1] did anyone who currently lives with you ever subject you to fondling or unwanted sexual intercourse or attempt to do so using violence, threats, coercion or surprise? Yes / No</p> <p>+</p> <p>Did this incident (at least one of these incidents) take place in [N-1]? Yes / No</p>	<p>Within your lifetime, has anyone ever:</p> <ul style="list-style-type: none"> - forced you to have sexual intercourse by using violence, threats, coercion or surprise? - attempted to force you to have sexual intercourse by using violence, threats, coercion or surprise? - subjected you to, or forced you to carry out, the fondling of genitals? <p>Yes in N / Yes in N-1 / Yes before N-1 / No, never</p>
Feeling unsafe at home	<p>Do you ever personally feel unsafe in your home? Often / From time to time / Rarely / Never</p>	<p>Do you ever personally feel unsafe:</p> <ul style="list-style-type: none"> - in your home? <p>Often / From time to time / Rarely / Never</p>
Feelings unsafe in your district/village	<p>Do you ever personally feel unsafe in your district or in your village? Often / From time to time / Rarely / Never</p>	<p>Do you ever personally feel unsafe:</p> <ul style="list-style-type: none"> - in your district or your village? <p>Often / From time to time / Rarely / Never</p>
Satisfaction as regards the operation of the justice system	<p>Would you say that the general operation of the justice system and the courts in the handling of delinquency is:</p> <p>very satisfactory / satisfactory / not very satisfactory / unsatisfactory</p>	<p>Would you say that the general operation of the justice system and the courts in the handling of delinquency is:</p> <p>very satisfactory / satisfactory / not very satisfactory / unsatisfactory</p>
Sufficient police and gendarmerie presence in your district or your village	<p>What would you say concerning the presence of the police or gendarmerie in your district (or your village)? (responses displayed in random order)</p> <p>That it is sufficient / That it is excessive / That it is insufficient / That it is non-existent despite being necessary / That it is non-existent and unnecessary</p>	<p>Would you say that the presence of the police or gendarmerie in your district (or your village) is:</p> <p>excessive / sufficient / insufficient</p>
Performance of the police and gendarmerie in your district or your village	<p>What would you say concerning the activities of the police or gendarmerie in combating delinquency in your district (or your village)?</p> <p>That they are very successful / That they are successful / That they aren't very successful / That they are not at all successful</p>	<p>Would you say that the activities of the police or gendarmerie in combating delinquency in your district (or your village) are:</p> <p>very successful / successful / not very successful / not at all successful</p>