

LE RECRUTEMENT DU PANEL ELIPSS

Bilan du projet réalisé avec le soutien financier de l'INPES
dans le cadre de la subvention n°060/12-DAS

Anne Cornilleau, Anne-Sophie Cousteaux, Stéphane Legleye,

Nirintsoa Razakamanana et l'équipe ELIPSS



Juin 2014

Introduction.....	3
Un panel internet pour la recherche.....	5
Les enquêtes par internet en population générale.....	5
Les expériences étrangères.....	6
Le dispositif ELIPSS.....	8
Le recrutement du pilote.....	11
L'échantillonnage.....	11
Le déroulement du terrain.....	13
Des actions pour améliorer le taux de recrutement.....	15
L'entrée des panélistes.....	17
Le taux de recrutement.....	19
La représentativité du panel.....	22
Quelques éléments sur la participation aux enquêtes.....	29
Bilan et perspectives.....	35
Bibliographie.....	38

INTRODUCTION

Internet est très utilisé par les sondeurs pour produire des enquêtes rapides et à bas coût. Pour cela, les instituts ont mis en place de larges panels d'internautes volontaires, qui peuvent s'y inscrire par exemple par des bannières postées sur des sites populaires. Ces panels sont donc constitués sans base de sondage ni procédure probabiliste, alors même que plus de 20% des ménages n'ont pas accès à Internet en 2012 (Gombault, 2013). Le projet d'*étude longitudinale par internet pour les sciences sociales* (ELIPSS) a vu le jour pour permettre aux chercheurs de mener des enquêtes probabilistes sur des thèmes peu traités par la statistique publique française et de conduire des enquêtes longitudinales de grande ampleur, ordinairement hors d'atteinte de la recherche académique en sciences humaines et sociales. ELIPSS est un projet unique en France combinant le recrutement d'un panel probabiliste (représentatif) et son interrogation via un outil mobile connecté à Internet, une tablette tactile.

ELIPSS s'insère dans le cadre de l'équipement d'excellence DIME-SHS (Données, infrastructures et méthodes en sciences humaines et sociales), lauréat de l'appel à projet 2010 des "investissements d'avenir" et financé par l'Agence nationale de la recherche (ANR). Cet équipement, porté par Sciences Po, en partenariat avec l'Institut national d'études démographiques (INED), le Groupe des écoles nationales d'économie et statistiques (GENES), l'Université Paris Descartes, Telecom Paris Tech, la branche recherche et développement d'Electricité de France (EDF R&D) et l'Ecole des hautes études en sciences sociales (EHESS) vise à collecter, enrichir et diffuser des données utilisables par la recherche en SHS, autant quantitative que qualitative (<http://www.sciencespo.fr/dime-shs/>). ELIPSS représente le volet « quantitatif » de cet équipement, les deux autres volets étant relatifs à la collecte des données sur le web et à l'archivage et la mise à disposition d'enquêtes qualitatives. Ce dispositif d'enquêtes par internet est destiné à la communauté scientifique afin de combler l'absence de moyens d'enquête par questionnaire propres aux chercheurs français en SHS. Il s'agit d'un panel internet représentatif de la population résidant en France métropolitaine. Une tablette tactile et un abonnement internet mobile sont fournis aux panélistes sélectionnés aléatoirement à partir du recensement afin qu'ils participent à des enquêtes mensuelles élaborées par des chercheurs.

De 2012 à 2014, les deux premières années du projet correspondent à l'étude pilote, nécessaire à la mise en place de la méthodologie et des outils informatiques, à l'expérimentation et à la validation du dispositif. Le recrutement du pilote s'est déroulé en deux temps. L'exploitation de l'échantillon principal (3500 adresses) a eu lieu de juin 2012 à février 2013. Trois modes de contact ont été mis en œuvre. Une invitation à participer a été envoyée par courrier postal. Ce premier contact a fait l'objet d'une expérimentation : la moitié des courriers a été envoyée avec un bon cadeau de 10 euros et l'autre moitié sans bon cadeau. Des relances postales et téléphoniques ont ensuite été réalisées. À l'issue d'un appel d'offres, l'institut de sondage TNS Sofres a été retenu pour exploiter en face-à-face les adresses pour lesquelles il n'y avait pas eu de contact et pour tenter de convaincre certains des ménages qui avaient refusé de participer. De janvier à avril 2013, les 1000 adresses de réserve ont été exploitées par téléphone et en face-à-face par TNS Sofres. Au terme de cette procédure, 1026 personnes ont accepté de participer au panel.

La première partie du rapport est consacrée à la présentation générale du dispositif, notamment des contextes français et étrangers en matière d'enquêtes par internet en population générale. La deuxième partie présente le recrutement du pilote du panel ELIPSS, elle décrit le déroulement du terrain et la représentativité du panel. C'est dans cette partie que sera analysé l'effet des bons cadeaux et des différents modes de contact en termes de taux de recrutement et de profil sociodémographique des répondants. La troisième partie donne des éléments sur la production des enquêtes, notamment sur la participation mensuelle des panélistes et sur l'évolution de la représentativité du panel au fil des enquêtes. Enfin, le bilan du pilote proposé en conclusion vise notamment à mettre au point la méthodologie de recrutement la plus efficace en termes de taux de participation, de biais sociodémographique et de coût financier pour obtenir un échantillon de 5000 panélistes en 2015.

UN PANEL INTERNET POUR LA RECHERCHE

Les enquêtes par internet en population générale

Si les enquêtes par internet présentent des avantages certains, leur utilisation en population générale revêt plusieurs difficultés. D'une part, on retrouve les avantages des enquêtes par questionnaires auto-administrés. Les coûts de collecte sont réduits, essentiellement par l'absence d'enquêteurs ; les enquêtés peuvent répondre au moment qui leur convient le mieux ; l'absence d'enquêteurs permet également d'aborder des questions plus personnelles (santé, sexualité...). D'autre part, le mode de passation par internet présente des avantages spécifiques. Il permet d'envisager de nouveaux modes de questionnement qui intègrent des vidéos, du son ou des applications interactives. Par ailleurs, la période de collecte peut être réduite puisqu'il n'y a pas (ou presque) de limite au nombre de personnes interrogées simultanément. De plus, les réponses sont sauvegardées au fur et à mesure de leur collecte.

Pour autant, l'utilisation d'internet pour enquêter en population générale se heurte à plusieurs difficultés qui mettent en cause la représentativité de l'échantillon et l'extrapolation des résultats. Il n'existe pas de liste exhaustive des internautes comme il existe une liste de logements ; et il n'existe pas de moyens simples pour s'adresser en particulier à un internaute via internet. C'est cette absence de base de sondage qui pose problème pour la constitution d'échantillon. Les enquêtes par internet sont ainsi généralement réalisées à partir d'échantillons de personnes recrutées dans un cadre non probabiliste qui ne satisfait pas aux exigences de la recherche. Les opérateurs commerciaux proposent en effet deux types d'enquête par internet. Des enquêtes ponctuelles sont proposées sur des sites (de journaux, des portails généralistes, etc.) à l'aide de bannières invitant à participer une seule fois sur un thème donné à un moment donné. Les panels sont au contraire des dispositifs d'enquête pérennes : le principe est de recruter un grand nombre de personnes via des moyens divers telles que des invitations affichées sur des sites internet et auxquelles il est proposé de participer régulièrement à des enquêtes en échange de compensations financières (généralement des bons d'achat). En France, les principaux panels des instituts de sondage comprennent plusieurs centaines de milliers de membres. Toutefois, il existe très peu de

documentation sur les modes de recrutement de ces panels et leur qualité statistique est sujette à caution.

Par ailleurs, outre cette absence de base de sondage, on note trois problèmes majeurs :

- Un biais de couverture : les personnes n'ayant pas accès à internet sont de fait exclues. Or en 2012, une personne sur cinq résidant en France n'avait pas d'accès internet à son domicile (Gombault, 2013).
- Un biais d'exposition : seules des personnes ayant une pratique suffisamment régulière d'internet peuvent être sollicitées pour participer à des enquêtes.
- Un biais d'auto-sélection : seules les personnes volontaires acceptent. Cet aspect est particulièrement important pour l'entrée dans des panels pour lesquels il est clairement stipulé que la personne sera sollicitée régulièrement.

Autrement dit, internet est un excellent support pour administrer des questionnaires mais n'est pas un univers dans lequel il est possible de réaliser un échantillon représentatif. La plupart des enquêtes et des panels sur internet sont non probabilistes pour les raisons exposées précédemment. Afin de dépasser ces biais, une possibilité est de dissocier les deux fonctions d'internet (comme univers et comme support de questionnaire) et de construire un panel internet à partir d'un échantillon aléatoire de la population. Ceci suppose de recourir à un mode de recrutement hors-ligne, d'inclure les personnes qui n'ont pas accès à internet et de les équiper d'une connexion le cas échéant (Das, Ester and Kaczmirek, 2011). Dans le cas du panel ELIPSS, ceci est réalisé par le recours à un échantillon aléatoire d'adresses tiré par l'INSEE et par la mise à disposition de tablettes tactiles connectées en 3G à tous les panélistes.

Les expériences étrangères

Deux expériences étrangères ont inspiré le projet de panel ELIPSS et deux initiatives allemandes sont développées en même temps que le pilote d'ELIPSS. Des discussions sont également en cours dans plusieurs pays pour mettre en place des projets similaires, notamment en Norvège, au Royaume-Uni, ou encore en Europe du Sud (un consortium regroupant Chypre, l'Espagne, l'Italie, la Grèce, le Portugal et la Turquie pourrait être construit dans les années à venir).

Le Longitudinal Internet Studies for the Social Sciences (LISS Panel) de l'institut de recherche néerlandais CentERdata (Université de Tilburg)

Ce panel internet représentatif des ménages néerlandais a été constitué en 2007 selon un plan de sondage probabiliste réalisé en collaboration avec le Centraal Bureau voor de Statistiek (CBS, Institut national de statistique des Pays-Bas). Il repose sur la mise à disposition d'un ordinateur simplifié et d'une connexion à internet aux ménages qui en sont dépourvus. Le LISS Panel est constitué de 5 000 ménages, soit 8 000 individus de 16 ans et plus. Ce dispositif est gratuit et exclusivement dédié à des opérations de recherche.

Le KnowledgePanel de Knowledge Networks aux Etats-Unis

Créé en 1999 par deux universitaires américains, ce dispositif repose sur un échantillonnage aléatoire et sur l'installation d'une connexion à internet chez les répondants qui n'en disposent pas à leur domicile. L'échantillon, tiré aléatoirement à partir d'une base d'adresses de logements, est composé de 55 000 personnes de 18 ans et plus. Contrairement au LISS Panel, ce dispositif est également ouvert aux études commerciales. En 2011, le KnowledgePanel a été racheté par l'institut de sondage GfK. Dans le cadre du programme TESS, l'accès au panel reste gratuit pour les chercheurs en sciences sociales, le programme prenant en charge le coût financier pour les projets d'enquête sélectionnés.

Le German Internet Panel (GIP) de l'Université de Mannheim

Ce panel représentatif de la population âgée de 16 à 75 ans résidant en Allemagne repose sur un échantillonnage aléatoire de 1 500 personnes. Les membres du panel sont interrogés tous les 2 mois. À l'instar du LISS Panel, un ordinateur simplifié et une connexion internet sont mis à disposition des personnes qui en sont dépourvues.

Ce dispositif est réservé aux chercheurs de l'Université de Mannheim et les questionnaires sont centrés sur les réformes politiques.

Le GESIS Panel du GESIS Leibniz Institute for the Social Sciences de Mannheim

Ce dispositif se caractérise par le recours à deux modes d'interrogation auto-administrés. Les personnes disposant d'un accès à internet depuis leur domicile peuvent répondre sur le web, les personnes ne disposant pas d'un accès à internet ou ne

souhaitant pas répondre sur le web peuvent répondre sur papier aux questionnaires envoyés par voie postale.

L'échantillon tiré aléatoirement à partir des registres municipaux est constitué de 4 000 individus âgés de 18 à 70 ans et résidant en Allemagne. Les enquêtes bimestrielles sont proposées par des équipes de recherche en sciences sociales dans le cadre d'appels à projets et ne peuvent aucunement servir un intérêt commercial.

Le dispositif ELIPSS

Le contexte français

En France, il n'existe pas de service de production d'enquêtes par questionnaire à vocation nationale dédié à la recherche en sciences humaines et sociales. Pour pallier cette absence, les chercheurs français ont deux solutions. Ils peuvent produire leurs enquêtes par le recours à un institut de sondage, mais cette solution est onéreuse, d'autant plus lorsqu'il s'agit de réaliser des enquêtes à partir d'un échantillonnage aléatoire. Ils peuvent également réutiliser des enquêtes statistiquement représentatives, par exemple les enquêtes de la statistique publique, mais celles-ci n'abordent pas l'ensemble des thèmes et des questions susceptibles d'intéresser les chercheurs.

Par ailleurs, on constate une diminution des taux de participation aux enquêtes, liée essentiellement aux refus de répondre et aux difficultés croissantes pour joindre les personnes à interroger. Or ceci fragilise la qualité statistique des enquêtes.

Ainsi, la mise en place du panel ELIPSS répond à deux objectifs principaux :

- permettre aux chercheurs de mener des enquêtes sur des thèmes qui ne sont pas traités par la statistique publique française,
- affranchir la recherche publique des intermédiaires du privé pour la réalisation d'enquêtes par questionnaire à partir d'un échantillonnage aléatoire tout en diminuant les coûts et le temps de collecte (par la passation auto-administrée sur internet).

La collecte par internet mobile

Le panel ELIPSS se distingue des dispositifs similaires à l'étranger par le choix de l'internet mobile comme mode de collecte principal. Une tablette tactile et un

abonnement 3G sont fournis à l'ensemble des panélistes en échange de leur participation. Ainsi, ils peuvent répondre aux questionnaires même s'ils ne disposent pas de connexion internet.

Le choix de cette nouvelle technologie s'explique par des raisons à la fois méthodologiques et pratiques. Il s'agit de tirer parti des possibilités offertes par internet (images, vidéos...) et par la mobilité pour renouveler certaines techniques d'enquête (étude des déplacements, carnets d'enquêtes budget-temps, etc.). Les tablettes offrent de nombreux avantages par rapport aux enquêtes sur ordinateur. Parce que leurs interfaces sont plus intuitives, elles offrent un accès internet simplifié aux personnes peu familières des nouvelles technologies. L'accès web mobile donne également aux panélistes plus de souplesse pour répondre aux enquêtes (choix du moment et du lieu). L'outil de collecte est le même pour tous les panélistes, via une application spécifique installée sur la tablette.

Outre les avantages de la tablette comme mode de collecte, ce choix a également été guidé par l'effet incitatif attendu par la mise à disposition d'un tel équipement dans le contexte actuel de développement de l'internet mobile et de la diffusion croissante de l'usage des outils tactiles connectés. Au moment du recrutement du pilote, le taux de pénétration des tablettes était faible en France (9% des ménages étaient équipés en 2012¹). Il faut toutefois noter la forte progression en seulement deux ans puisque près d'un ménage sur trois était équipé à la fin de l'année 2013. En effet, ELIPSS ne rémunère pas ses panélistes, au contraire des panels probabilistes existant à l'étranger (bien que cette rémunération soit modique) et des panels commerciaux nationaux. La tablette tactile est prêtée gratuitement et dispose d'une connexion illimitée gratuite à internet via les réseaux 3G de téléphonie mobile (c'est-à-dire indépendamment d'un équipement wifi). Le panéliste peut l'utiliser à sa guise en échange de sa participation aux enquêtes mensuelles. C'est donc la tablette qui constitue l'incitation matérielle à participer dans le cadre d'ELIPSS.

¹ Enquête GfK / Médiamétrie – Référence des Equipements Multimédia

Le calendrier et l'équipe

L'étude pilote a débuté en 2012 afin de définir la procédure de recrutement, d'affiner la méthodologie, de mettre au point les procédures de gestion de panel et de production d'enquêtes et de développer les outils informatiques (tableau 1). Pendant cette phase de test, l'utilisation du service de production d'enquêtes d'ELIPSS est réservée aux équipes de recherche du consortium DIME-SHS répondant à des appels à projets.

Tableau 1 : Calendrier du pilote

		Recrutement des panélistes	Opérateur téléphonique	Outils informatiques	Enquêtes
2012	mars-12	Tirage de l'échantillon par l'INSEE		Début développement des applications panélistes, gestion du panel et des enquêtes	
	avr-12		Appel d'offres opérateurs		
	mai-12	Préparation des documents panélistes		Ouverture du site elipss.fr et de l'application de recrutement	Formation Blaise
	juin-12	Envoi des courriers d'invitation et de relance Appel d'offres instituts de sondage	Choix de l'opérateur		Groupe de travail Enquête annuelle ELIPSS
	juil-12	Choix de TNS Sofres Relances téléphoniques			
	août-12		Négociation du contrat		
	sept-12	Préparation du terrain par TNS Sofres		1e version de l'application panélistes	1e CST DIME Quanti
	oct-12	Envoi des conventions de participation Recrutement par TNS SOFRES	Signature du contrat	1e version de l'application gestion du panel et des enquêtes	Lancement de l'appel à projets 2012
	nov-12		Envoi des tablettes Activation des abonnements	Programmation, design et tests de la 1e enquête	
	déc-12		Formation téléphonique		1e enquête Pratiques numériques
2013	janv-13				
	févr-13				
	mars-13				

À partir de 2015, le panel ELIPSS devrait être constitué de 5000 individus et les appels à projets d'enquête seront ouverts à l'ensemble de la communauté scientifique.

La construction de ce dispositif repose sur une équipe aux compétences variées qui s'est étoffée au fur et mesure du pilote (tableau 2). Cette équipe est localisée dans deux institutions. Le Centre de données socio-politiques de Sciences Po assure la coordination du projet, l'ensemble de la production et la diffusion des enquêtes ainsi que les développements et l'infrastructure informatiques. L'INED a la charge de la gestion du panel ainsi que du suivi de la qualité statistique du panel. Ainsi, huit personnes travaillent à Sciences Po (auxquelles s'ajouteront bientôt trois personnes dont le recrutement est en cours) et cinq à l'INED.

Tableau 2: l'équipe ELIPSS et son évolution depuis 2012

Rôle	Institution	2012				2013				2014			
		T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
Coordination	Sciences Po	Anne Cornilleau											
		Anne-Sophie Cousteaux								Recrutement en cours			
Gestion du panel	INED	Carmen Calandra											
		Gabrielle Bouchet				Marc Sigaud				Patricia Sossa			
						Elodie Pétorin (40%)				Kevin Boudelle			
Production des enquêtes	Sciences Po	Emmanuelle Duwez											
		Matthieu Olivier								Alexandre Mairot			
										Recrutement en cours			
Statistiques	INED									Nirintsoa Razakamanana			
Informatique	Sciences Po	Adrien Ferreira											
		*Daniele Guido											
		*Geneviève Michaud											
		*Jérémy Richard (50%)								Recrutement en cours			

*: Personnes travaillant pour les trois volets de l'équipement DIME-SHS

LE RECRUTEMENT DU PILOTE

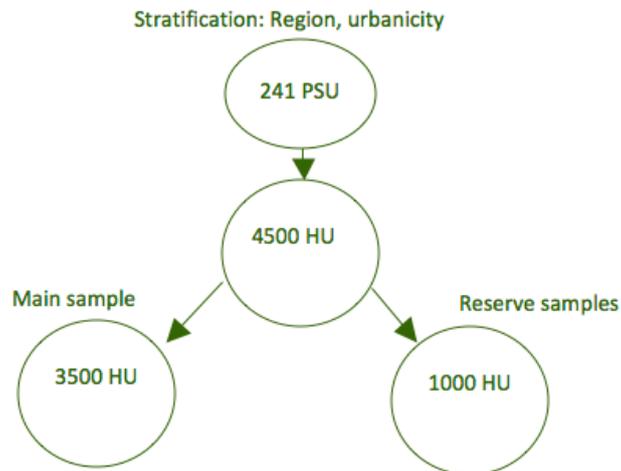
L'échantillonnage

La base de sondage

Le panel ELIPSS est un panel aléatoire d'individus résidant en ménages ordinaires au sens de l'Institut national des statistiques et des études économiques (INSEE). Sont donc

exclues les personnes sans domicile ou vivant dans des habitations précaires, les personnes résidant dans des collectivités (prisons, maisons de retraite, résidences étudiantes, etc.), et également celles ne maîtrisant pas suffisamment la langue française pour répondre aux questionnaires auto-administrés.

Figure 1 : L'échantillonnage réalisé par l'INSEE



La base de sondage est constituée des logements recensés en 2011². A partir de cette base de sondage, un échantillon de 4500 logements a été tiré par l'INSEE³ par sondage à deux degrés stratifié par région et type de commune (urbain/rural). Le premier degré est un tirage d'unités primaires (PSU) correspondant à des communes ou des groupes de communes, avec un taux de sondage de 1/3⁴. Le second degré est un tirage de logements au sein des 241 PSU. Les 4500 adresses ont été divisées en trois sous-échantillons : un échantillon principal de 3500 adresses et deux échantillons de réserve de 300 et 700 adresses à utiliser au cas où l'objectif des 1500 panélistes ne serait pas atteint.

Au sein de chaque logement sélectionné par l'INSEE, une seule personne est ensuite tirée au sort.

² 3% des logements ont été recensés en 2009 ou en 2010.

³ Cet échantillon a été fourni gracieusement par l'INSEE à ELIPSS à des fins expérimentales.

⁴ L'Échantillon-Maître de l'INSEE est constitué de 567 PSU. La taille effective de l'échantillon du pilote ELIPSS n'était pas compatible avec la mobilisation de l'ensemble des PSU définies par l'INSEE.

La préparation de la base

La phase préliminaire du travail a consisté à saisir les fiches adresses fournies par l'INSEE et sur lesquelles les noms et adresses des chefs de ménage étaient manuscrites. Une vérification de la cohérence et de la vraisemblance des adresses saisies a été faite manuellement par internet par les gestionnaires de panel de l'INED. Elle a été complétée par l'institut de sondage pour les adresses des échantillons de réserve. Le format des adresses a également été vérifié automatiquement par la Poste et la base a été enrichie par la recherche des numéros de téléphone.

Au total, 76 adresses dans l'échantillon principal et 13 dans les échantillons de réserve ont été considérées comme inexploitable. Par ailleurs, environ 10% des lettres sont revenues avec la mention NPAI (n'habite pas à l'adresse indiquée), ces adresses ont directement reçu la visite d'un enquêteur.

Des lettres d'information à propos du panel ELIPSS ont été expédiées aux mairies des communes de l'échantillon INSEE afin d'assurer une publicité autour du projet.

Le déroulement du terrain

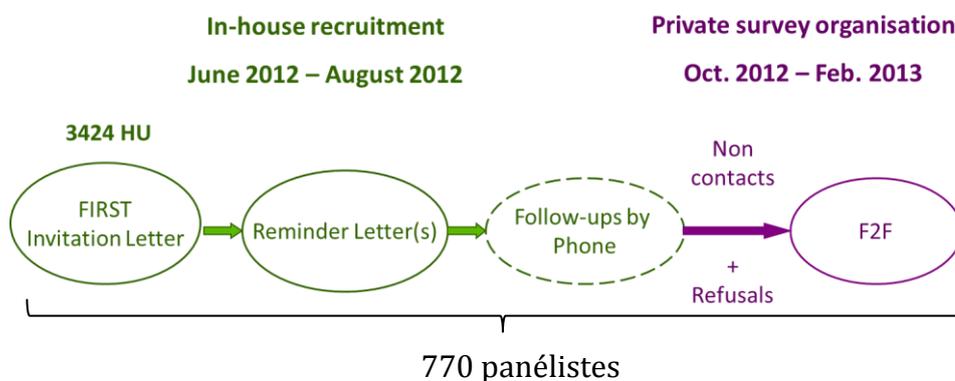
Afin de préparer le recrutement de 2015, la procédure définie pour le pilote a consisté à tester différents modes de contact ordonnés par ordre croissant du coût : courrier puis téléphone et enfin face-à-face. Deux procédures différentes d'exploitation des adresses de l'échantillon principal et des échantillons de réserve ont été mises en œuvre.

L'échantillon principal

Pour l'échantillon principal contenant 3424 adresses, la procédure de recrutement s'est faite à l'aide de trois modes de contact successifs. Une première invitation à participer a été envoyée par courrier postal en juin 2012 par l'équipe ELIPSS. Ce courrier présentait le dispositif d'enquête et demandait aux personnes de nous communiquer la composition de leur foyer afin de pouvoir procéder au tirage d'un individu dans le logement. La lettre d'invitation était suivie d'une lettre de relance 15 jours plus tard. Jusqu'à mi-juillet 2012, des relances ont été réalisées par téléphone pour les ménages dont le numéro a pu être retrouvé (c'est-à-dire pour 45% des adresses de l'échantillon principal). Le recrutement en face à face s'est déroulé d'octobre 2012 à février 2013 auprès de ménages pour lesquels aucun contact n'avait encore eu lieu (non-répondants

et NPAI) ainsi qu'auprès de la plupart des ménages qui avaient refusé de participer (ces refus représentent 17,5% des adresses qui ont été réexploitées).

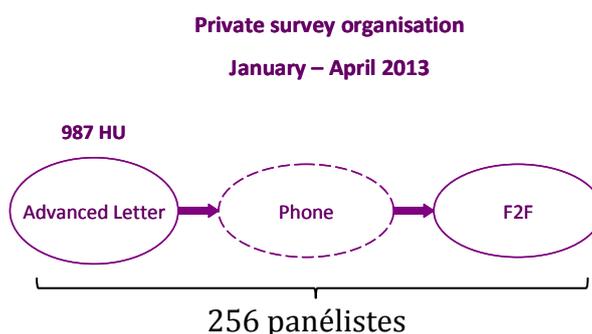
Figure 2 : Procédure de recrutement - Echantillon principal



Les échantillons de réserve

A l'issue de l'exploitation de l'échantillon principal, 770 personnes ont été recrutées. Afin de s'approcher de l'objectif visé de 1500 panélistes, il a été décidé en janvier 2013 de confier les 987 adresses de réserve à l'institut de sondage. Lorsque cela était possible, les premières tentatives de contact ont eu lieu par téléphone⁵. Si les 12 tentatives d'appel restaient infructueuses ou si une personne du ménage avait refusé, un enquêteur se présentait à l'adresse. Par ailleurs, les logements sans numéro de téléphone ont directement fait l'objet d'un recrutement en face à face. L'exploitation des échantillons de réserve a permis de recruter 256 panélistes supplémentaires.

Figure 3 : Procédure de recrutement - Echantillons de réserve



⁵ La mobilisation de plusieurs annuaires téléphoniques a permis à TNS-Sofres d'obtenir un numéro de téléphone pour plus de 60% des logements des échantillons de réserve.

Le tirage de l'individu

Au terme de l'une ou l'autre de ces phases de recrutement, la composition des ménages répondants était connue. Au sein de chaque logement, un individu était sélectionné par tirage aléatoire simple. Les critères d'éligibilité étaient les suivants : être âgé de 18 à 75 ans, habiter en France métropolitaine (hors Corse) et parler français.

Finalement, 1026 individus ont accepté de participer au panel ELIPSS en signant la convention (cf. infra). Plus de la moitié des panélistes (51%) ont été recrutés en face-à-face, 31% par courrier et 18% par téléphone.

Des actions pour améliorer le taux de recrutement

Plusieurs actions ont été mises en place en vue d'améliorer le taux de recrutement : varier les modes de réponse à la lettre d'invitation, offrir des bons cadeaux, former les enquêteurs et réexploiter les refus.

Les différents modes de réponse à l'invitation à participer

Les lettres d'invitation offraient trois modes de réponse aux ménages sélectionnés. Ainsi, 421 personnes ont renvoyé le coupon-réponse par courrier postal, 248 ont utilisé l'accès qui leur était réservé sur le site internet elipss.fr et 28 ont téléphoné au numéro indiqué sur la lettre. La moitié des personnes qui ont répondu par courrier postal ont accepté de participer alors que c'est le cas de la quasi-totalité de celles qui ont répondu par internet (95%).

L'expérimentation des bons cadeaux

Le recrutement du pilote a été l'occasion de mener une expérimentation en joignant 2000 bons cadeaux à la première lettre d'invitation envoyée aux logements issus de l'échantillon principal. L'objectif était de tester si une incitation financière indépendante de la participation (dite aussi inconditionnelle) pouvait favoriser le recrutement des panélistes. Cette expérimentation est la première de ce genre en France, où les incitations financières dans les enquêtes scientifiques en population générale sont très rares. En s'inspirant des résultats de la littérature internationale et notamment de l'expérimentation réalisée par le LISS panel (Scherpenzeel, 2009), les incitations financières d'une valeur de 10 euros ont été distribuées de manière aléatoire et inconditionnelle lors de l'invitation à participer.

Les ménages ayant reçu un bon cadeau ont plus souvent répondu à l'invitation (pour accepter ou pour refuser de participer) et ont également plus de chances d'avoir accepté de participer (tableau 3).

Tableau 3: Réponse à l'invitation au panel selon l'envoi de bon cadeau

	Non réponse	Réponse			Tous	Total
		Refus	Acceptation	Inéligible		
Sans bon cadeau (n =1424)	67,8	18,9	11,2	2,1	32,2	100
Avec bon cadeau (n = 2000)	62,1	16,6	18,8	2,5	37,9	100
Total (n = 3424)	64,5	17,6	15,7	2,3	35,5	100

Note : le test du khi-2 est significatif (p-value <0.001)

L'effet positif des incitations financières sur la réponse et sur l'acceptation méritait d'être confirmé en France où cette pratique n'est pas courante dans le domaine des enquêtes scientifiques. En outre, le fait d'avoir reçu un bon cadeau multiplie par deux les chances d'accepter de participer, toutes choses égales par ailleurs. Ce résultat est en accord avec une méta-analyse d'expérimentations menées dans le monde anglo-saxon. Cette étude montre que les incitations financières inconditionnelles sont très efficaces (l'Odds-ratio s'élève à 2) et qu'elles fonctionnent mieux qu'une incitation conditionnée à la réponse dans les enquêtes postales ou électroniques (Edwards, Roberts, Clarke et al., 2009).

La modélisation de l'acceptation à participer au panel montre qu'après l'envoi d'un bon cadeau, la disponibilité d'un numéro de téléphone est la deuxième variable la plus déterminante. L'Odds-ratio associé au fait d'avoir retrouvé un numéro pour le foyer s'élève à 1,7. Ceci peut refléter le fait que les ménages acceptant de faire figurer leur numéro dans l'annuaire sont plus coopérants que ceux sur liste rouge. Il est à noter que les effets du bon cadeau et de la publicité du numéro se multiplient. Conjointement au numéro de téléphone, le bon-cadeau multiplie par plus de trois la participation du ménage.

La formation des enquêteurs

Les 110 enquêteurs ont suivi une demi-journée de formation approfondie, dont le contenu et le déroulement ont été préparés conjointement par l'équipe ELIPSS et l'institut de sondage. Les formations, auxquelles ont pris part des membres de l'équipe ELIPSS, portaient notamment sur la présentation du panel ELIPSS, sur les arguments et les techniques pour convaincre les personnes de participer, sur l'explication des termes de la convention que les personnes doivent signer (cf. infra), sur le remplissage des

fiches-contact et sur la prise en main des tablettes numériques par les enquêteurs. En effet, sur le terrain, les enquêteurs étaient équipés d'une tablette afin de pouvoir en faire une démonstration aux personnes sélectionnées.

La conversion des refus

Parmi les refus obtenus par courrier ou téléphone, 542 ont été réexploités par l'institut de sondage. Les raisons sous-tendant ces premiers refus sont relativement classiques dans les enquêtes par questionnaire. Il s'agit le plus souvent d'un manque d'intérêt ou d'un manque de temps. Des raisons spécifiques à ELIPSS ont également été données, certaines personnes ont déclaré ne pas être à l'aise avec les nouvelles technologies, d'autres ne souhaitent pas avoir accès à internet ou une tablette. En revanche, les refus liés à des questions de confidentialité ou d'anonymat n'ont pas été confiés à l'institut de sondage. A l'issue de la réexploitation, 49 refus (9%) ont donné lieu à une acceptation à participer.

L'entrée des panélistes

La convention

Pour faire partie du panel ELIPSS, les personnes sélectionnées ont signé une convention qui régit les conditions de participation aux enquêtes et d'utilisation de la tablette.

En signant cette convention, les panélistes prennent trois engagements. Ils s'engagent à répondre personnellement et régulièrement aux enquêtes. Ils s'engagent aussi à prendre soin de la tablette et informer l'équipe ELIPSS en cas de casse, de perte ou de vol. Ils s'engagent enfin à restituer la tablette à la fin de leur participation au panel. En échange, ils peuvent librement utiliser la tablette et internet dans le respect de la législation en vigueur. A tout moment, ils peuvent suspendre ou mettre fin à leur participation.

En cas de non-réponse prolongée aux questionnaires ou d'utilisation frauduleuse de la tablette, les panélistes peuvent être exclus du panel.

Outre la mise à disposition gratuite de la tablette et de l'abonnement internet 3G, la convention interdit l'utilisation commerciale des enquêtes et garantit l'anonymat des réponses aux questionnaires.

La prise en main de la tablette

Une fois la convention signée et renvoyée, les panélistes reçoivent une tablette à leur domicile. En page d'accueil, est notamment installée l'application qui permet de remplir les questionnaires.

Tous les panélistes se sont vus proposer une formation téléphonique pour prendre en main la tablette, découvrir l'application ELIPSS et, le cas échéant, paramétrer le wifi. Ces formations ont été assurées par un prestataire extérieur à partir d'un script défini par l'équipe ELIPSS. Deux tiers des panélistes ont été formés, 30% n'ayant pu être joints par téléphone⁶ et 5% ayant refusé la formation. A l'issue de la formation, les formateurs devaient évaluer le niveau d'aisance avec les nouvelles technologies. Ils ont estimé que 20 % des panélistes formés n'étaient pas à l'aise avec la tablette.

La première enquête

Lancée en décembre 2012, la première enquête⁷ comportait deux parties. Le « didacticiel » devait permettre aux panélistes de se familiariser avec différents types de question et le design des enquêtes ELIPSS. Le module « Enquêtes et internet » visait à mesurer l'accès à internet, les pratiques numériques et la participation aux enquêtes avant l'entrée dans le panel ELIPSS. Administrée jusqu'à la fin du recrutement en avril 2013, 90% des panélistes⁸ ont répondu à cette première enquête.

D'après cette enquête, 91% des panélistes disposent d'un accès internet à domicile. Ils sont 79% à utiliser internet tous les jours et 6% à se connecter moins d'une fois par semaine. Cette enquête nous apprend également que 13% des panélistes affirment avoir toujours refusé de participer à une enquête avant le panel ELIPSS. Elle confirme enfin que la mise à disposition d'une tablette tactile a été la principale motivation pour

⁶ Dans la majorité des cas, ces personnes n'étaient pas relancées pour suivre la formation dans le cas où elles avaient répondu à la première enquête.

⁷ Les données ne sont pas pondérées

⁸ Seuls 943 panélistes ont été invités à participer à cette enquête car les dernières personnes recrutées sont entrées dans le panel au mois d'avril 2013 au moment de l'enquête annuelle 2013. En considérant les seules personnes invitées à répondre à « Enquêtes et internet », le taux de complétion est de 99% (tableau 9).

accepter de participer au panel ELIPSS (citée par 62% des panélistes), suivie de la confiance dans les institutions impliquées dans le projet (pour 46% des panélistes). Viennent ensuite l'originalité du projet (37%), l'intérêt pour la recherche (32%) et la mise à disposition d'un abonnement internet (13%).

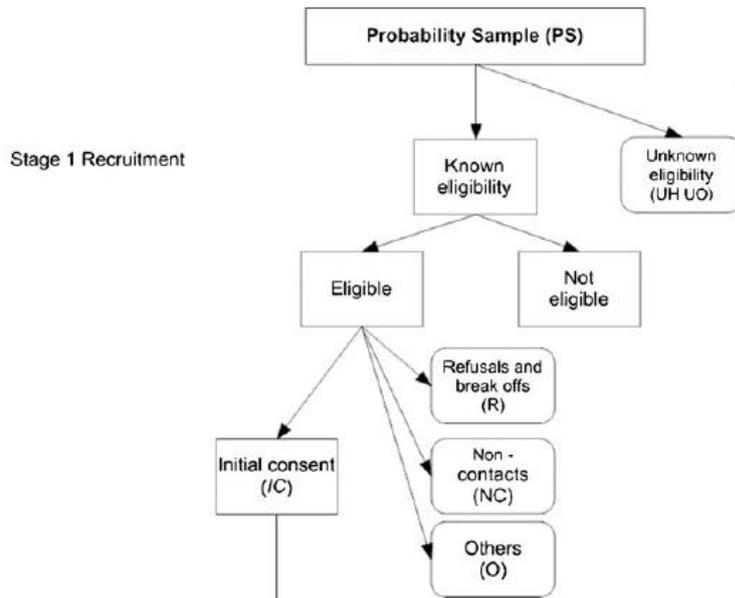
Le taux de recrutement

Résultats globaux

Rappelons qu'au terme de l'exploitation des 4500 adresses de l'échantillon fourni par l'INSEE, 1026 individus ont accepté de faire partie du panel et ont signé la convention. Le calcul du taux de recrutement est un premier élément d'évaluation du pilote. On peut distinguer ici deux types d'accord à participer (*initial consent - IC*) tel que défini par Callegaro et DiSogra (2008). En effet, dans la procédure spécifique aux adresses de l'échantillon principal, une première étape a consisté à demander l'accord du ménage pour décrire la composition des membres du logement⁹, puis le second accord intervient au niveau individuel au moment de la signature de la convention de participation au panel. A la différence de la procédure décrite par Callegaro et DiSogra (2008), il n'y a pas eu de *profil/connection stage*, ou plus précisément, elle est confondue avec l'étape dite d'*initial consent* au niveau de l'individu.

⁹ Ceci est particulièrement vrai dans le cas de la première phase de recrutement de l'échantillon principal puisque les ménages devaient avoir une démarche volontaire pour renvoyer le coupon-réponse, téléphoner aux gestionnaires de panel ou remplir le formulaire en ligne pour décrire des habitants du logement.

Figure 4 : Stratégie de recrutement d'après Callegaro et DiSogra, 2008



Les résultats finaux de la procédure de recrutement du panel ELIPSS sont reproduits dans le tableau 4 ci-dessous. Ils ont été calculés à partir des formules de l'AAPOR, en considérant, comme le proposent Callegaro et DiSogra (2008, p.1018), le *RR3* comme taux de recrutement (*RECR*), tel que définit ci-dessous :

$$\text{Recruitment rate (RECR)} = \frac{IC}{IC + (R + NC + O) + e(UH + UO)}$$

where

IC = initial consent

R = cases directly and actively refusing

NC = noncontacts

O = other cases

UH = unknown if household is occupied

UO = unknown other

e = estimated proportion of cases of unknown eligibility that are eligible.

Notons que l'éligibilité est parfois délicate voire impossible à déterminer quand aucun contact avec le ménage n'a pu avoir lieu. Pour tous ces cas, nous avons appliqué le taux d'éligibilité (*e*) observé sur les unités pour lesquelles l'éligibilité est connue.

Ainsi on obtient un taux de recrutement de 27% au niveau de l'individu (*RECR*), avec un taux de réponse de 36% (*RR3*) au niveau ménage. Par ailleurs, l'exploitation des échantillons de réserve a permis d'atteindre un taux de recrutement légèrement plus élevé que celle de l'échantillon principal. Cette différence tient notamment à un plus grand nombre de ménages pour lesquels l'éligibilité n'a pas pu être déterminée dans l'échantillon principal, les taux de refus étant similaires (45% dans l'échantillon principal et 46% dans les échantillons de réserve). Ceci s'explique en grande partie par

le travail plus approfondi de nettoyage et de vérification des adresses mis en œuvre pour les échantillons de réserve, ce qui a permis de diminuer les non-contacts.

Tableau 4: Taux de réponse des ménages et taux de recrutement des individus

		Ménages	Individus	<i>Echantillon principal (individus)</i>	<i>Echantillons de réserve (individus)</i>
Eligible	Accord (IC)	1352	1026	770	256
	Refus (R)	1481	1695	1318	377
	Non contacts (NC)	0	2	2	0
	Autres cas (O)	65	175	163	12
Inéligible		570	570	426	144
Eligibilité inconnue	Adresse non localisée (UH)	172	172	89	83
	Ménage injoignable (UO)	860	860	732	128
	e	0,836	0,836	0,841	0,817
RR3 = RECR		0,36	0,27	0,26	0,31

Résultats selon le mode de contact

Malgré les lettres de relance, différents modes de réponse possibles et les bons-cadeaux offerts à la majeure partie des ménages de l'échantillon principal, nous n'avons recueilli que 697 réponses à la lettre d'invitation, soit 20% de réponse. Ces réponses étaient certes majoritairement positives mais il a été indispensable de réexploiter en face-à-face 2217 non-contacts et 450 refus exprimés principalement lors des relances téléphoniques. Le calcul consistant à privilégier le mode de contact le plus économique, à savoir le courrier postal, s'avère finalement coûteux dans la mesure où les enquêteurs en face-à-face se sont vus contraints de réexploiter les adresses les plus délicates (soit parce que le ménage était difficilement joignable, soit parce qu'il avait déjà refusé), et ce, pour 78% des adresses de l'échantillon principal.

La comparaison des taux de recrutement dans les échantillons de réserve permet d'évaluer l'efficacité des deux procédures mises en œuvre : téléphone puis face-à-face versus face-à-face seulement. En d'autres termes, peut-on avoir au recours au téléphone pour limiter le coût financier du recrutement en face-à-face des panélistes ? En considérant seulement les ménages répondants, on obtient 54% d'accord et 46% de refus en utilisant uniquement le face-à-face contre 42% d'accord et 58% de refus en utilisant d'abord le téléphone. Le face-à-face donne également de meilleurs résultats pour transformer une acceptation au niveau ménage en une acceptation au niveau individu. C'est le cas pour 85% des ménages directement sollicités par un enquêteur à domicile contre 80% des ménages joints d'abord par téléphone.

Contrairement à l'expérimentation des bons cadeaux, la façon dont le ménage a été approché, par téléphone ou en face-à-face, n'a pas été randomisée. L'allocation entre les deux procédures de contact dépendait de la disponibilité d'un numéro de téléphone. En outre, la procédure de recrutement reposant uniquement sur les enquêteurs en face-à-face est donc d'autant plus efficace qu'elle concernait des ménages plus difficiles à joindre.

La représentativité du panel

Il n'importe pas seulement de savoir combien d'individus ont finalement accepté de faire partie du panel ELIPSS et participent régulièrement aux enquêtes, il importe aussi de savoir qui sont ces individus, quelles sont leurs caractéristiques. En d'autres termes, dans quelle mesure le panel est-il représentatif de la population générale ? Dans cette partie, nous donnerons des éléments de comparaison entre le profil du panel issu de l'enquête annuelle et les statistiques nationales et nous poursuivrons l'analyse des différents modes de contact en décrivant les profils sociodémographiques des panélistes recrutés par courrier, téléphone et face-à-face.

L'enquête annuelle

L'enquête annuelle ELIPSS a pour objectif de disposer de nombreuses variables socio-démographiques (module signalétique¹⁰) ainsi que de variables de croisement et d'indicateurs fréquemment utilisés en sciences humaines et sociales (module barométrique¹¹). Le questionnaire a été construit en collaboration avec plusieurs chercheurs spécialistes des thèmes abordés. En outre, les questions retenues sont très largement issues d'enquêtes existantes, nationales et internationales. Une majorité des variables issues du module signalétique est systématiquement appariée aux fichiers de données diffusés.

¹⁰ Ce module traite des questions suivantes : état civil, travail et formation, description socio-démographique du ménage, logement et quartier, revenus et patrimoine.

¹¹ Ce module s'intéresse aux liens sociaux, aux loisirs et pratiques culturelles, aux croyances et pratiques religieuses, aux comportements et opinions politiques, à l'état de santé, aux comportements de santé et aux habitudes de vie.

La première enquête annuelle ELIPSS a été administrée en 2013 à l'issue du recrutement du panel, en avril pour le module signalétique et en mai pour le module barométrique. En mars 2014, les deux modules ont été administrés ensemble : le module signalétique a été répliqué dans sa totalité (les questions ont été adaptées de façon à mesurer l'évolution depuis la première interrogation) et une sélection de questions du module barométrique a été reposée, à l'exception de la partie sur les comportements et opinions politiques¹².

Comparaison du profil du panel avec les statistiques nationales

Un taux de recrutement relativement bas ne signifie pas nécessairement que l'échantillon est de mauvaise qualité. En effet, le risque que les estimateurs soient biaisés dépend surtout de la différence en termes de comportements de réponse entre ceux qui ont accepté de participer et ceux qui n'ont pas accepté. Ne connaissant pas les réponses de ceux qui n'ont pas accepté, il est impossible d'évaluer ce risque pleinement. La représentativité du panel peut cependant s'apprécier par la distorsion de l'échantillon de répondants relativement à celui de la population cible. Pour chacune des 4500 adresses, nous disposons des informations renseignées par le ménage qui occupait le logement à la date du recensement.

En première approche, nous avons comparé la distribution des variables géographiques et des principales variables décrivant le logement entre l'échantillon tiré initialement par l'INSEE et l'échantillon de répondants à l'enquête annuelle (tableau 5). Relativement au recensement de la population, les marges de l'échantillon de répondants à l'enquête annuelle sont distordues en faveur des communes rurales et des grands logements. La distorsion est faible concernant la grande région (ZEAT – Zone d'études et d'aménagement du territoire) et n'est pas significative pour le caractère social du logement (HLM) et la résidence en zone urbaine sensible (ZUS).

¹² Ces questions sont posées régulièrement au panel dans le cadre du projet longitudinal sur les dynamiques de mobilisation.

Tableau 5 : Comparaison des distributions des variables géographiques et des variables décrivant le logement

Variable	Modalité	Recensement INSEE (2010) en %	Enquête annuelle ELIPSS (avril 2013) en %	
ZEAT	Région parisienne	19.4	17.0	*
	Bassin parisien	18.0	19.2	
	Nord	6.6	5.2	
	Est	8.2	10.5	
	Ouest	13.2	14.6	
	Sud-Ouest	10.6	9.9	
	Centre Est	12.2	12.2	
	Méditerranée	11.8	11.4	
Taille de l'unité urbaine	Rural	21.5	25.9	*
	De 2 000 à 4 999 habitants	2.6	2.0	
	De 5 000 à 9 999 habitants	4.6	5.3	
	De 10 000 à 19 999 habitants	6.8	6.2	
	De 20 000 à 49 999 habitants	7.7	6.7	
	De 50 000 à 99 999 habitants	8.1	6.5	
	De 100 000 à 199 999 habitants	6.5	6.4	
	De 200 000 à 1 999 999 habitants	24.1	25.3	
	Agglomération de Paris	18.1	15.7	
Nombre de pièces dans le logement	1	5.7	5.2	***
	2	12.8	10.5	
	3	21.3	17.1	
	4	25.3	23.3	
	5	19.2	22.7	
	6 et plus	15.7	21.2	
HLM	Le logement appartient à un organisme HLM	16.2	15.0	ns
	Le logement n'appartient pas à un organisme HLM	83.8	85.0	
ZUS	Le logement n'est pas situé en ZUS	93.3	94.9	ns
	Le logement est situé en ZUS	6.7	5.1	

*, **, ***, ns : valeur-test du test du χ^2 de Pearson respectivement <0.05, <0.01, <0.001 et non significative

Si la mise à disposition de tablettes tactiles aux membres du panel ELIPSS a pour objectif de faire face au problème de couverture lié aux enquêtes par internet, on constate néanmoins des différences entre les caractéristiques individuelles des panélistes et celles de la population cible. Le tableau 6 compare la distribution de plusieurs variables sociodémographiques dans le panel (à partir de l'enquête annuelle 2013) et dans la population des 18-75 ans résidant en France métropolitaine (à partir des données du recensement de 2010).

Tableau 6: Comparaison des distributions sociodémographiques

Variable	Modalité	Recensement INSEE (2010) en %	Enquête annuelle ELIPSS (avril 2013) en %			
		18-75 ans	Non Pondéré	Pondéré ¹³		
Sexe	Hommes	48.6	48.0	ns	48.6	ns
	Femmes	51.4	52.0	ns	51.4	ns
Age	18-24 ans	12.0	8.2	***	12.0	ns
	25-34 ans	17.7	19.1		17.7	
	35-44 ans	20.0	25.7		20.0	
	45-54 ans	19.6	21.8		19.6	
	55-64 ans	18.1	16.5		18.1	
	65-75 ans	12.6	8.8		12.6	
Taille ménage	1	16.7	25.9	***	19.5	***
	2	34.9	25.8		25.9	
	3	20.1	17.9		22.3	
	4	18.0	20.6		23.7	
	5 et plus	10.4	9.9		8.7	
Statut matrimonial	Célibataire	37.9	40.1	***	38.2	***
	Marié	50.2	44.9		49.6	
	Divorcé	3.7	12.3		9.6	
	Veuf	8.3	2.7		2.6	
Nationalité	Français de naissance	88.2	90.9	***	88.2	ns
	Français par acquisition	5.4	5.2		5.4	
	Etranger	6.4	3.9		6.4	
Statut d'activité professionnelle	Occupe un emploi	59.7	64.9	***	55.8	***
	Etudiant(e) ou en stage	4.6	6.0		8.7	
	Chômeur	7.8	7.5		8.3	
	Retraité(e)	20.1	15.2		19.6	
	Autre situation	8.0	6.5		7.6	
Statut d'occupation du logement	Propriétaire	60.5	58.5	***	59.0	***
	Locataire ou sous-locataire	37.2	35.3		32.6	
	Occupant à titre gratuit	2.3	4.8		6.5	
	NSP ou NVPR	0	1.4		1.8	
Niveau de diplôme	Aucun/CEP/BEPC	28.4	17.6	***	28.4	ns
	CAP/BEP	24.9	20.9		24.9	
	Bac à bac+2	32.3	36.5		32.3	
	Bac+3 et plus	14.4	24.9		14.4	

*, **, ***, ns : valeur-test du test du χ^2 de Pearson respectivement <0.05, <0.01, <0.001 et non significative

On observe des distorsions similaires à celles constatées à l'issue du recrutement du LISS panel (Leenheer and Scherpenzeel, 2013 ; Knoef and de Vos, 2009) en termes d'âge

¹³ L'âge, le sexe, le niveau de diplôme, la nationalité et la ZEAT ont été utilisées pour le calage sur marge. Les marges ont été calculées à partir du recensement de la population 2010. Au final, le poids obtenu par calage (méthode du raking ratio) a pour moyenne 4,79, écart-type de 2,58, pour un rapport max/min de 13.5 (max=18,9, min=1,40), et coefficient de variation de 53,99.

et de niveau de diplôme. Les personnes âgées de 55 à 75 ans sont sous-représentées comme attendu, ainsi que les 18-25 ans, groupe toujours difficile à atteindre dans les enquêtes en France. A l'inverse, les ménages d'une seule personne sont surreprésentés, ainsi que les personnes très diplômées et les personnes occupant un emploi. La répartition hommes-femmes est en revanche très proche de celle de la population. Des analyses multivariées sont en cours pour compléter ces premiers résultats.

Une question importante dans la mise en place d'un projet comme Elipss est le fait de fournir un accès à internet à ceux qui n'y avaient pas accès avant leur participation au panel¹⁴. En comparant les données de l'enquête sur les technologies de l'information et de la communication et le commerce électronique réalisée en 2012 par l'INSEE (tableau 7), on observe que les panélistes vivant dans des ménages déjà équipés d'un accès internet à domicile sont surreprésentés. Le panel ELIPSS comprend 9% d'*offliners* alors que la population des 18-75 ans en compte 17%. Cet écart est similaire à ceux constatés dans le LISS Panel et le German Internet Panel (Blom, Gathmann, Krieger, 2013 ; Leenheer and Scherpenzeel, 2013).

Tableau 7: Comparaison de l'accès à internet à domicile

	Enquête TIC INSEE, 2012 (18-75 ans)	ELIPSS
Equipé d'un accès internet à domicile avant ELIPSS	83%	91%
Sans accès internet à domicile avant ELIPSS	17%	9%

Comparaison du profil des personnes recrutées selon le mode de contact

Le tableau 8 montre que globalement, les différents modes de recrutement ont sélectionné des populations différentes sur le plan sociodémographique. Les femmes sont surreprésentées dans les trois modes de collecte, sans distinction significative entre les trois ($P_c=0.721$). Par ailleurs, quel que soit le mode, la répartition hommes/femmes ne se différencie pas significativement de celle du recensement rénové de la population (RRP) 2010 ($P_b>0.05$).

¹⁴ Des analyses sur le profil des *offliners* sont en train d'être menées en collaboration avec Mélanie Revilla (Université Pompeu Fabra, Barcelona).

Tableau 8 : Caractéristiques des panélistes par mode de contact (% colonne, distance du Khi² normalisée et p-value du test par rapport au RRP pour chaque mode, et p-value pour la comparaison des trois modes)

	RRP 2010	Courrier (n=286)			Face-à-face (n=475)			Téléphone (n=162)			P^c
Sexe	(%)	(%)	Khi² ^a	P^b	(%)	Khi² ^a	P^b	(%)	Khi² ^a	P^b	
Hommes	48.6	49.7			46.7			48.8			0.721
Femmes	51.4	50.4	0.0	0.722	53.3	0.13	0.417	51.2	0.00	0.966	
Age											
18-24 ans	12.0	6.3			9.1			9.3			0.045
25-34 ans	17.7	20.6			18.7			17.3			
35-44 ans	20.0	29.7			26.1			17.3			
45-54 ans	19.6	21.0			22.3			21.6			
55-64 ans	18.1	14.3			14.7			25.3			
65-75 ans	12.6	8.0	10.5	<.0001	9.1	4.6	0.001	9.3	4.9	0.154	
Diplôme											
Aucun diplôme_CEP_BEPC	28.4	12.9			21.1			16.1			0.002
CAP_BEP	24.9	17.5			23.6			19.1			
Bac	19	17.8			19.6			20.4			
Bac+2	13.3	19.6			16.4			16.1			
Bac+3 et plus	14.5	32.2	35.2	<.0001	19.4	4.4	0.000	28.4	20.7	<.0001	
ZEAT											
Région parisienne	19.1	21.7			14.1			17.3			0.027
Bassin parisien	16.9	16.1			21.9			16.7			
Nord	6.4	7.7			4.6			2.5			
Est	8.6	11.9			9.5			11.1			
Ouest	13.5	9.8			16.4			17.9			
Sud-Ouest	11.0	9.1			10.3			9.9			
Centre Est	12.0	12.6			11.2			14.8			
Méditerranée	12.5	11.2	3.4	0.201	12.0	4.1	0.007	9.9	6.0	0.197	
Taille d'unité urbaine											
Rural	21.5 ^d	23.1			28.4			23.5			0.002
2 000 à 4 999 habitants	2.6	2.8			0.8			3.7			
5 000 à 9 999 habitants	4.6	5.9			3.6			9.3			
10 000 à 19 999 habitants	6.8	5.9			6.1			6.8			
20 000 à 49 999 habitants	7.7	6.3			7.2			6.2			
50 000 à 99 999 habitants	8.1	5.2			7.2			6.8			
100 000 à 199 999 habitants	6.5	2.5			8.6			6.8			
200 000 à 1 999 999 habitants	24.1	28.0			25.1			21.6			
Agglomération de Paris	18.1	20.3	5.3	0.056	13.1	6.0	0.000	15.4	6.5	0.225	

^a : Distance du Khi² normalisée, c'est-à-dire calculée pour une taille arbitraire de 100 individus pour l'échantillon et la population cible. Ce calcul permet la comparaison des écarts de distribution entre les différents modes sans considération de la taille des différents sous-échantillons. Cette distance n'est pas celle du test du Khi² qui tient compte de la taille des échantillons comparés.

^b : valeur-p classique du test de Khi² de Pearson comparant la distribution dans le sous-échantillon du mode de collecte à celle du recensement, avec les tailles réelles observées des échantillons.

^c : Test du Khi² de Pearson pour la comparaison des trois modes de collecte.

^d : Marge provenant de l'échantillon des logements INSEE (n=4500).

Pour l'âge, on observe globalement que les jeunes de 18-24 ans et les personnes âgées de 65 à 75 ans sont sous-représentés par les trois modes de collecte. Pour le téléphone, les 55-64 ans sont surreprésentés même si la distribution observée pour ce mode de contact n'est pas significativement différente de celle du RRP ($P_b > 0.05$). Par ailleurs, c'est le recrutement par courrier qui présente la distribution la plus éloignée de celle de la population cible du RRP (la distance du χ^2 normalisée étant la plus élevée), devant le téléphone puis le face-à-face. On note cependant que les trois modes diffèrent faiblement entre eux ($P_c = 0.045$).

D'un point de vue global, selon la distance du χ^2 normalisée, c'est le face-à-face qui présente la meilleure distribution des niveaux de diplôme, devant le téléphone et le courrier. Les trois distributions, qui diffèrent effectivement entre elles ($P_c = 0.002$), sont toutefois significativement différentes de celles du recensement. Le niveau de diplôme des personnes recrutées varie toutefois fortement suivant le mode de collecte. Par courrier, les peu ou pas diplômés sont très sous-représentés ; à l'inverse, les très diplômés sont largement surreprésentés. Par téléphone, la tendance est identique mais les écarts sont moins prononcés. Enfin, pour le face-à-face, les sans diplôme et les bac +3 sont moins surreprésentés que par les deux autres modes de contact.

Pour ce qui est des variables géographiques, les écarts entre les trois modes de collecte sont plus réduits, bien que les trois modes diffèrent également entre eux ($P_c = 0.027$). C'est le courrier qui est le plus proche de la distribution du RRP, devant le face-à-face et le téléphone. Les tests montrent que seul le face-à-face se distingue de la distribution cible du RRP. La Région parisienne est surreprésentée dans l'échantillon recruté par courrier mais largement sous-représentée en face-à-face, l'Ouest est surreprésenté en face-à-face et par téléphone mais sous-représenté par courrier, le Bassin parisien est surreprésenté en face-à-face tandis que le Nord est très sous-représenté par téléphone. Enfin, concernant la taille de l'unité urbaine, le recrutement par courrier présente la répartition la plus proche de la population. Viennent ensuite le face-à-face et le téléphone. Les communes rurales sont surreprésentées par les trois modes, mais c'est particulièrement le cas en face-à-face.

Pour résumer, les trois modes de contact nous ont permis d'atteindre des populations qui diffèrent significativement selon l'âge, le niveau de diplôme, la région de résidence et la taille de l'unité urbaine. Parmi les personnes recrutées par courrier, sont

surreprésentés les adultes de 25 à 44 ans, les très diplômés, les personnes vivant dans de grandes agglomérations, en région parisienne, dans le Nord ou l'Est. Parmi les personnes qui ont accepté de participer au panel ELIPSS suite à la visite d'un enquêteur à leur domicile, sont surreprésentés les adultes de 35 à 44 ans, peu diplômées, habitant dans des communes rurales, dans le Bassin parisien ou l'Ouest. Enfin, parmi les personnes recrutées par téléphone, sont surreprésentés les individus âgés de 45 à 64 ans, plus diplômées que celles recrutées en face-à-face, les habitants des petites communes de l'Ouest.

Même si le face-à-face donne les meilleurs taux de recrutement, la diversité des profils des personnes recrutées milite aussi en faveur d'une procédure mobilisant plusieurs modes de contact susceptibles d'atteindre ces différentes populations. En effet, au regard des résultats précédents, il pourrait être intéressant de privilégier le face-à-face dans les communes rurales, le téléphone dans les petites villes et le courrier dans les grandes agglomérations, notamment en région parisienne.

QUELQUES ELEMENTS SUR LA PARTICIPATION AUX ENQUETES

Les enquêtes sélectionnées par le comité scientifique et technique

Depuis le premier appel à projet en 2011, 20 propositions d'enquête ont été reçues, parmi lesquelles 11 ont été acceptées par le comité scientifique et technique DIME Quanti. Ce comité sélectionne les projets d'enquête déposés par les chercheurs lors des appels à projets et examine les demandes d'appariement de données. Il compte 15 membres dont au moins la moitié est affiliée à des institutions extérieures au consortium DIME-SHS. Il est composé de chercheurs issus de disciplines variées en sciences sociales (sociologie, sciences politiques, démographie, épidémiologie, etc.) spécialisés en méthodes d'enquête par questionnaire. Il comprend un représentant de l'INSEE, en raison du tirage de l'échantillon par l'institut et de la validation des enquêtes par le Conseil national de l'information statistique (CNIS).

Au total, de décembre 2012 à mai 2014, 17 enquêtes ont été administrées au panel ELIPSS. Outre l'enquête sur les pratiques numériques lors de l'entrée dans le panel et l'enquête annuelle ELIPSS, les enquêtes ont abordé des thèmes variés tels que les pratiques culturelles, la contraception, les valeurs et les opinions politiques,

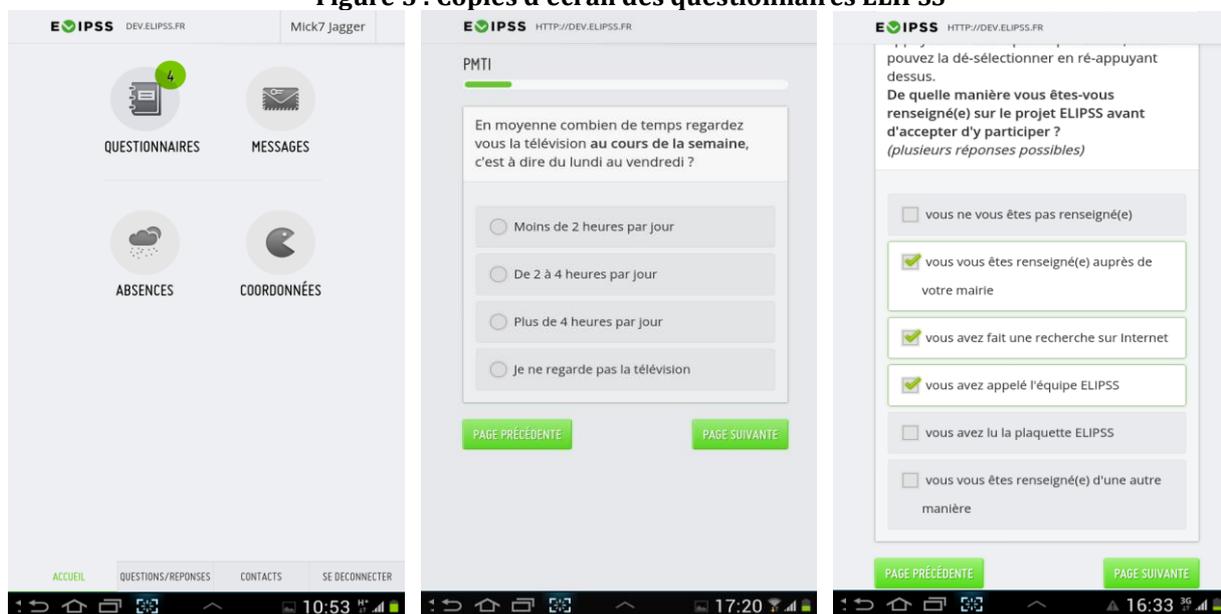
l'environnement, le couple, la famille et les relations intergénérationnelles, la santé et les expositions professionnelles (tableau 9).

Deux projets comparatifs ont été acceptés par le comité scientifique et technique en dehors du cadre des appels à projets : un projet porté par Jon Krosnick qui a pour objet de répliquer des expérimentations classiques aux Etats-Unis sur plusieurs panels internet à travers le monde et un projet comparatif entre les quatre panels internet probabilistes en Europe (le GESIS panel, le GIP, ELIPSS et le LISS panel).

La tablette comme mode collecte

Les panélistes répondent aux questionnaires via une application développée en interne et pré-installée en page d'accueil de la tablette (figure 5)¹⁵. La conception des questionnaires et la collecte des données en ligne reposent sur le logiciel Blaise. Développée par Statistics Netherlands, cette solution est destinée à la réalisation d'enquêtes de la statistique publique et est utilisée par la plupart des instituts nationaux de statistiques, y compris l'INSEE.

Figure 5 : Copies d'écran des questionnaires ELIPSS



¹⁵ Cette application permet également aux panélistes d'envoyer des messages aux gestionnaires de panel, d'indiquer leurs périodes d'indisponibilité, de modifier leurs coordonnées et d'avoir les réponses aux questions les plus fréquentes.

La participation mensuelle et l'attrition

Le tableau 9 décrit la participation aux enquêtes administrées depuis l'inclusion dans le panel.

Tableau 9 : Liste des enquêtes administrées au panel ELIPSS depuis décembre 2012

Terrain	Enquête	Taux de réponse	Nombre d'invités	Attrition
déc.2012-mars 2013	Enquêtes et internet	99%	943	0%
avril 2013	Enquête annuelle 2013 (TC1) (module signalétique)	91%	1012	0%
mai 2013	Enquête annuelle 2013 (TC2) (module barométrique)	87%	1011	0%
juin 2013	Pratiques culturelles, médias et technologies de l'information	88%	1011	0%
juil.2013-août 2013	Fécondité, contraception, dysfonctions sexuelles	87%	1005	2%
septembre 2013	Dynamique de mobilisation - vague 1	87%	996	2%
octobre 2013	Enquête sur les valeurs, l'environnement et l'énergie	85%	997	3%
novembre 2013	Situation de couple, intentions de fécondité et opinions sur la famille	88%	983	3%
déc.2013-janv.2014	Santé, travail et environnement / Dynamique de mobilisation - vague 2	83%	993	3%
février 2014	Les relations entre générations au prisme des normes de solidarité et de justice sociale	77%	984	4%
mars 2014 (2 semaines)	Dynamique de mobilisation - vague 3 (pré-municipales)	64%	945	8%
mars 2014	Enquête annuelle 2014 (TC)	90%	945	8%
avril 2014 (2 semaines)	Dynamique de mobilisation - vague 4 (post-municipales)	83%	945	8%
avril 2014	Expérimentations dans plusieurs pays	85%	945	8%
mai 2014 (2 semaines)	Dynamique de mobilisation - vague 5 (pré-européennes)	77%	936	9%
mai 2014	Enquête des panels européens	85%	936	9%

Jusqu'en novembre 2013, le taux de participation (ce taux est le COMR décrit par Callegaro, Disogra, 2008) est au-dessus de 85% et commence ensuite à décliner pour atteindre un taux de 77% à l'enquête de février 2014¹⁶. Si l'on constate une chute importante de la participation à cette enquête, il convient de noter que cette enquête est

¹⁶ La vague pré-municipale de Dynamob de mars 2014 connaît un taux encore plus bas de 64% ce qui s'explique surtout par un terrain de deux semaines seulement, ce qui était inédit pour les panélistes, comme pour l'équipe ELIPSS.

en réalité la seule à avoir connu un terrain de quatre semaines, sans délai supplémentaire et avec un problème technique qui a empêché l'envoi de mails de relance. L'enquête de mars 2014 connaît l'un des taux de réponse les plus élevés avec 90%, ceci après une campagne de relances personnalisées et un terrain rallongé de quelques jours. Des efforts importants ont en effet été mis en œuvre car il s'agit de l'enquête annuelle dans laquelle les caractéristiques socio-démographiques des panélistes sont mises à jour. Le taux de réponse se maintient depuis autour de 85%.

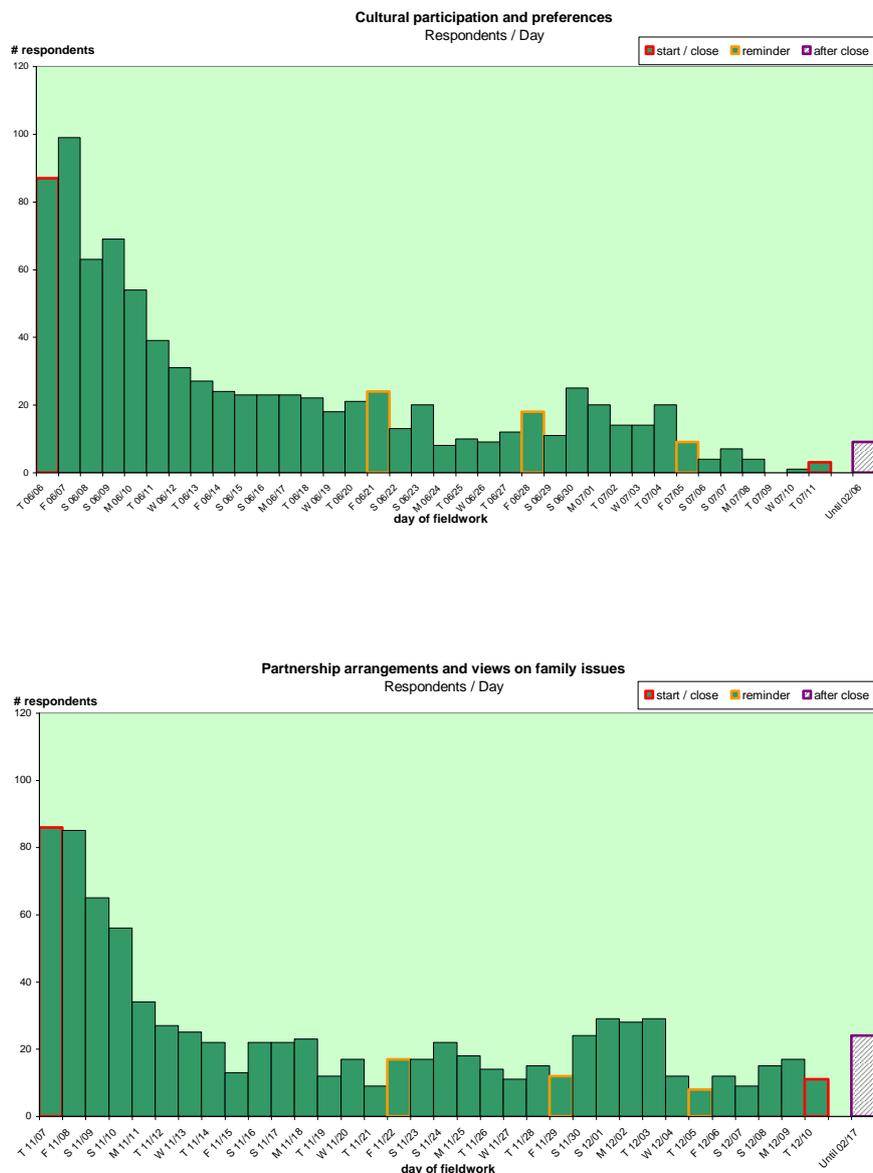
Le taux d'attrition commence à augmenter à la fin de l'année 2013 et atteint 9% en mai 2014. Ceci correspond aux premières désactivations des panélistes « invisibles » qui ont eu lieu à la suite de la systématisation des relances auprès de ce groupe.

Maintien de la participation

Les relances des panélistes pour leur participation mensuelle aux enquêtes reposent sur plusieurs dispositifs. Après avoir été invités au panel le premier jour de terrain par un message sur la tablette (et par mail lorsque l'adresse électronique est disponible), des messages automatiques de relance sont envoyés sur les tablettes des non-répondants les vendredis à partir de la seconde semaine de terrain. Les graphiques ci-dessous montrent l'évolution du nombre de répondants par jour à l'enquête de mai 2013 (première enquête avec des relances systématisées) et celle de novembre 2013. On constate une première vague de réponse après l'invitation pour les deux enquêtes, puis pour la seconde enquête, des pics plus visibles les week-ends. Dans ce dernier cas, il est difficile de savoir pour l'instant si l'on peut associer ce phénomène à un effet des relances qui ont lieu le vendredi.

La participation aux enquêtes dépend également des contacts réguliers des panélistes avec les gestionnaires du panel. Les premiers mois, les contacts consistaient principalement à traiter des problèmes techniques liés aux tablettes et la connexion 3G, des cas de casse, de perte et de vol de tablettes. Ceci a permis de définir les différentes procédures pour gérer ces situations. Des procédures de relances ont été élaborées pour suivre les panélistes ne répondant pas à plusieurs enquêtes. Ainsi, les « somnolents » (non-répondants à l'enquête en cours et à l'enquête précédente) et les « invisibles » (non répondants à au moins deux enquêtes) font l'objet de relances personnalisées par téléphone et par courrier postal chaque semaine.

Figure 6: Participation quotidienne aux enquêtes de mai et novembre 2013



Les non-répondants à l'enquête en cours ou les répondants l'ayant commencée sans la terminer¹⁷ sont relancés en fonction de la charge de travail des gestionnaires de panel et des taux de non réponse de ces deux groupes.

Les raisons de la non-réponse des panélistes évoquées lors de ces relances personnalisées sont souvent des problèmes techniques ou des cas de vols ou pertes de

¹⁷ Les panélistes qui ont commencé le questionnaire bénéficient d'un délai supplémentaire d'un mois pour le terminer.

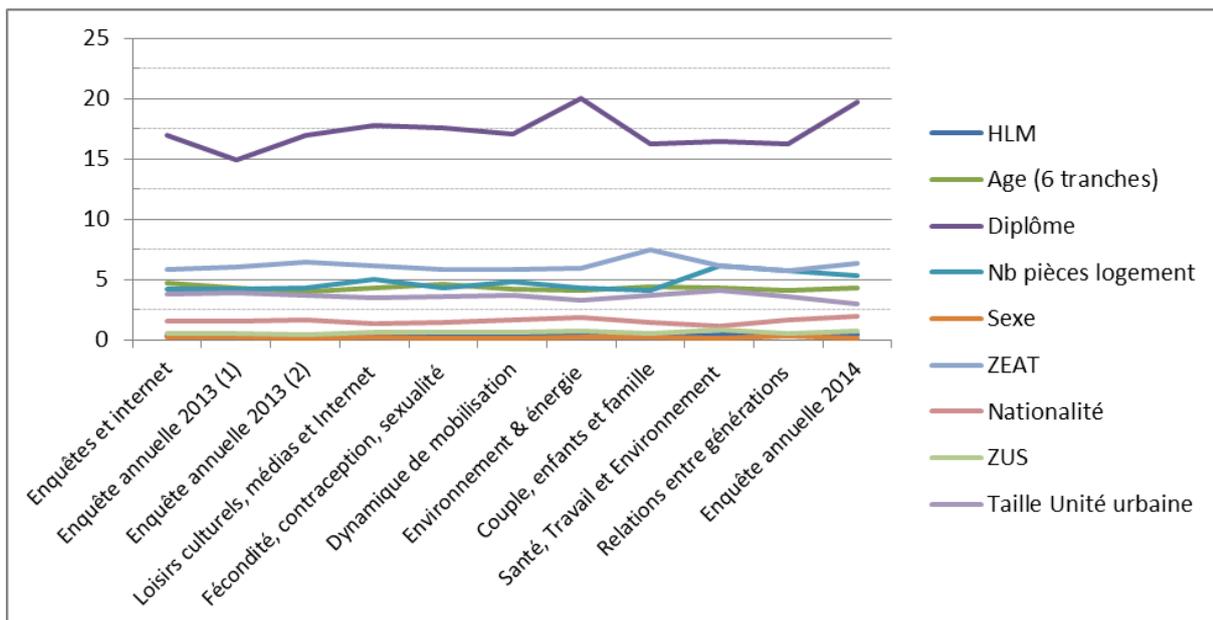
tablette. Dans le cas où il s'avère que les panélistes ne peuvent techniquement pas répondre aux enquêtes, ils sont suspendus du panel.

Afin de faciliter le suivi des panélistes, un outil en ligne de gestion du panel a été développé par le développeur de l'équipe ELIPSS¹⁸. Tous les échanges avec les panélistes sont ainsi répertoriés au quotidien par les gestionnaires du panel.

L'évolution de la représentativité

La représentativité du panel peut aussi s'apprécier au fil des enquêtes. La figure 7 montre l'évolution de la distance entre l'échantillon de répondants à chaque enquête (avant calage mais redressé par le poids de sondage) et l'échantillon INSEE. Cette distance est celle du khi deux normalisée afin de tenir compte du fait que le nombre de répondants varie d'une enquête à l'autre.

Figure 7 : Evolution des marges de l'échantillon de répondants relativement à celles de l'échantillon Insee au cours des douze premières enquêtes du panel



Les distorsions apparaissent très importantes s'agissant du niveau de diplôme, notamment pour l'enquête sur l'environnement et l'énergie qui avait pour particularité de comporter de nombreuses questions ouvertes. Les écarts à l'échantillon de l'INSEE sont élevés pour la grande région (ZEAT), le nombre de pièces du logement, l'âge et la

¹⁸ Cet outil a bénéficié de l'expérience de l'institut de recherche néerlandais CentERdata, responsable du LISS panel, qui a mis à notre disposition leur propre système de gestion des panélistes.

taille de l'unité urbaine. Elles sont en revanche faibles pour la nationalité, le caractère social du logement, la résidence en zone urbaine sensible et le sexe. Les écarts à l'échantillon INSEE sont d'abord restés plutôt stables mais ils augmentent légèrement pour le niveau de diplôme dans l'enquête annuelle 2014. Il conviendra de poursuivre cette analyse pour les prochaines enquêtes afin de déterminer dans quelle mesure la représentativité du panel se dégrade lorsque l'attrition augmente.

BILAN ET PERSPECTIVES

La construction d'un dispositif tel que le panel ELIPSS recouvre des activités très diverses qui correspondent aux différentes étapes du cycle de vie des données, allant du projet d'enquête à la mise à disposition des données produites. En particulier, le pilote, destiné à tester la faisabilité d'un tel panel en France, a nécessité d'inventer l'inexistant. En premier lieu, le recrutement des panélistes a fait l'objet d'une procédure relativement complexe afin de déterminer la meilleure approche pour le développement du panel en 2015. Par ailleurs, l'originalité d'ELIPSS de fournir une tablette tactile et un abonnement 3G à chaque panéliste a nécessité la contractualisation avec un opérateur de téléphonie mobile, ce qui impliquait des aspects techniques, opérationnels, juridiques et financiers. Aussi, il a fallu développer en interne les outils informatiques puisqu'aucun outil existant ne remplissait l'ensemble des besoins du projet. Plusieurs de ces opérations, notamment le contrat avec l'opérateur de téléphonie, étaient des préalables indispensables à la construction du panel et sont maintenant définis pour le passage à 5000 panélistes.

Le pilote du panel ELIPSS a mis au jour des difficultés à prendre en compte et des pistes d'amélioration en vue du développement du panel en 2015. En premier lieu, l'utilisation d'une base d'adresses pour recruter les panélistes présente des inconvénients. Telle que fournie par l'INSEE, elle nécessite un important travail de nettoyage et de vérification des adresses. Elle implique aussi de décrire préalablement tous les membres du logement avant de sélectionner une personne pour l'inviter à participer au panel. Si ces contraintes sont communes à toutes les enquêtes menées à partir d'un échantillon d'adresses, le recrutement du pilote a rencontré des difficultés spécifiques à la procédure retenue et au timing du projet : le terrain qui s'est en partie déroulé pendant l'été ainsi que le retard pris du fait de la contractualisation avec l'opérateur ; l'invitation par courrier et les relances par téléphone qui ne se sont pas avérées aussi

efficaces qu'escompté ; la signature de la convention de participation qui ajoute une étape pour convaincre les personnes sélectionnées et formalise l'engagement à participer. En revanche, la couverture réseau a été meilleure qu'attendue (aucune personne éligible n'a été exclue pour cette raison) et la tablette a eu l'effet prévu en incitant les personnes à participer. Elle présente aussi l'immense avantage de fournir un seul et même outil aux panélistes pour répondre à des questionnaires dont l'aspect visuel est entièrement contrôlé.

Quels enseignements tirer du pilote pour continuer à construire et faire vivre le panel ELIPSS avec 5 000 individus ? Par souci méthodologique et financier, nous avons testé différents modes de contact, mais force est de constater que le face-à-face est la stratégie la plus efficace pour convaincre les ménages puis recruter des panélistes. Il semble donc difficile de faire l'économie du recours à un institut de sondage pour le recrutement prévu au début de l'année 2015. Il en va de la taille et de la qualité de l'échantillon. Nous pouvons cependant envisager un protocole en deux phases afin de bénéficier de la diversité des profils recrutés par courrier et face-à-face : une lettre proposant aux personnes intéressées de s'inscrire directement sur le site web dédié et annonçant sinon la visite d'un enquêteur, puis un recrutement en face-à-face des non-répondants. En effet, la possibilité de s'inscrire par internet avait permis de recueillir facilement presque 15% des réponses positives au niveau du ménage. Il serait également envisageable d'autoriser les enquêteurs à utiliser le téléphone pour tenter de joindre un ménage dont le logement est inaccessible (par exemple dans les grandes agglomérations). Pour des raisons méthodologiques et techniques, nous comptons continuer à équiper les panélistes de tablettes ayant toutes le même format, fonctionnant avec le même système d'exploitation et sur lesquelles l'application ELIPSS sera pré-installée. Cependant, ces choix en matière de recrutement et d'équipement internet ne sont pas encore entièrement financés et ont fait l'objet d'une demande de financement complémentaire auprès de la région Ile de France.

Par ailleurs, le recrutement des nouveaux panélistes en 2015 mérite d'être mieux préparé. Nous bénéficierons évidemment de l'expérience acquise lors du pilote. La procédure sera également plus simple et prise en charge par l'institut de sondage. De plus, nous avons prévu de renforcer l'équipe en embauchant une personne chargée précisément de préparer et de suivre le terrain en face-à-face.

En dernier lieu, cette période sera une occasion unique de mener des expérimentations et d'étudier par exemple l'éventuelle professionnalisation en comparant les comportements de réponse des anciens et des nouveaux panélistes. Pendant le pilote, les travaux méthodologiques ont dû généralement passer au second plan par rapport à la mise en place du panel et à la production des enquêtes mensuelles car le projet avait largement été sous-dimensionné en termes de force de travail. Compte tenu du renforcement prévu de l'équipe, la recherche méthodologique, indispensable pour suivre et maintenir la qualité du dispositif, devrait occuper une plus grande place dans le projet.

BIBLIOGRAPHIE

- Blom A., Gathmann C., Krieger U., The German Internet Panel: Method and Results, 2013
- Callegaro M., 2010 – Do You Know Which Device Your Respondent Has Used to Take Your Online Survey?, *Survey Practice*, vol.3, n°6
(<http://www.surveypractice.org/index.php/SurveyPractice/article/view/250/html>)
- Callegaro M., DiSogra C., 2008 – Computing Response Metrics for Online Panels, *Public Opinion Quarterly*, vol.75, n°5, p.1008-1032
- Couper M., 2008 – *Designing effective web surveys*, New York: Cambridge University Press
- Das M., Ester P., Kaczmirek L.(eds.), 2011 – *Social and Behavioral Research and the Internet: Advances in Applied Methods and Research Strategies*, Boca Raton: Taylor & Francis
- Edwards P.J., Roberts I., Clarke, M.J., DiGuseppi C., Wentz R., Kwan I., Cooper R., Felix L.M., Pratap S., 2009 – Methods to increase response to postal and electronic questionnaires, *Cochrane database of systematic review*, n°02/2009
- Gombault V., 2013 – L'internet de plus en plus prisé, l'internaute de plus en plus mobile *INSEE première*, n°1452, juin
- Knoef M., de Vos K., 2009 – The representativeness of the LISS, an online probability panel, Tilburg, CentERdata, 29p.
(http://www.lisspanel.nl/assets/uploaded/representativeness_LISS_panel.pdf)
- Leenheer J., Scherpenzeel A.C., 2013 – Does it pay off to include non-internet households in an internet panel?, *International Journal of Interent Science*, vol. 8, n°1, p.17-29
- Scherpenzeel A., 2009 – Start of the LISS panel: Sample and recruitment of a probability-based internet panel, Tilburg, CentERdata, 9 p.
(http://www.lissdata.nl/assets/uploaded/Sample_and_Recruitment.pdf)