

ENQUÊTES EN LIGNE : COMPARAISON DE MODES DE QUESTIONNEMENT SUR LE THÈME DU LOGEMENT

Marie-Hélène AMIEL (), Thomas DENOYELLE (**)*

() Insee, Inspection générale*

*(**) Insee, Direction des statistiques démographiques et sociales*

Introduction

Dans le cadre des réflexions sur le moyen terme « Insee Ambition 2015 », il a été proposé par le groupe « grands chantiers statistiques », d'expérimenter auprès des ménages des enquêtes combinant plusieurs modes de collecte (face-à-face, téléphone, internet...), appelées dès lors « enquêtes multimodales ».

La direction des statistiques démographiques et sociales (DSDS) a repris cette thématique dans son plan d'action de moyen terme, en soulignant qu'elle contribue aux ambitions « mettre l'excellence technique et l'innovation au cœur de nos travaux » et « dégager des marges de manœuvre pour mieux remplir nos missions et répondre à de nouvelles demandes ».

Un processus de modernisation des statistiques sociales européennes a par ailleurs été engagé par Eurostat. Cette réingénierie devrait s'appuyer notamment sur les enquêtes par internet et la collecte multimodale. Il est donc essentiel pour l'Insee de pouvoir prendre position par rapport à ces projets en pleine connaissance de cause.

L'expérimentation est donc indispensable car il s'agit de porter une appréciation objective sur les nouveaux modes de collecte (internet, applications smartphone...) en vue de décider de l'opportunité de leur utilisation lors d'enquêtes auprès des ménages et des individus.

Parallèlement à cela, il a été envisagé, pour la future enquête nationale sur le logement, de pouvoir disposer de modes de collecte alternatifs au face-à-face pour une partie de l'enquête. Il a donc paru judicieux d'expérimenter une enquête par internet sur le thème du logement et de la mobilité résidentielle.

Avant cela, deux enquêtes par internet auprès des ménages avaient été réalisées : l'enquête auprès des non-répondants de l'enquête emploi et une opération méthodologique en parallèle avec l'enquête TIC 2010.

Par ailleurs, pour tirer enseignement des enquêtes en ligne déjà expérimentées par d'autres services statistiques ou organismes, un groupe de réflexion a été mis en place à partir du printemps 2010. Il est animé par l'Insee et le Credoc. Les participants viennent de plusieurs organismes : Insee, Credoc, Ined, Inserm, Inpes, Drees, Irdes, Cevipof. Les expériences passées ou en cours y sont discutées et des monographies sur les enquêtes en ligne y sont présentées. Les réunions du groupe ont aussi permis des échanges avec des instituts privés. Les auteurs remercient le groupe qui, à plusieurs stades de cette expérimentation, a accepté d'en discuter les modalités et a ainsi contribué à enrichir le protocole.

Les auteurs tiennent aussi à remercier Daniel Verger pour ses conseils et sa relecture attentive.

1. Quelques notions

Qu'entend-on par « mode » ?

On entend par « mode » une façon de faire. Dans les enquêtes statistiques, la notion de « mode » est employée dans des sens différents selon le stade de l'opération.

Le « **mode de contact** » indique les moyens employés pour proposer à un ménage ou à un individu de participer à l'enquête. Exemples : lettre-avis ou coup de téléphone.

Le « **mode de collecte** » ou « mode de questionnement » indique le moyen par lequel les questions sont posées. On distingue les modes de collecte par entretien (collecte avec un enquêteur réalisée en face-à-face ou par téléphone) et les modes de collecte « auto-administrés » (réponse sur papier ou par internet).

Le « **mode de suivi** », en particulier pour les enquêtes avec ré-interrogations, se rapporte aux relations post-enquête avec les répondants (envoi de premiers résultats, suivi d'adresse).

Diverses combinaisons sont possibles et varient effectivement suivant les enquêtes. Par exemple, un ménage peut recevoir une lettre-avis lui demandant de participer à une enquête par internet et recevoir par la suite un coup de fil de remerciement et des résultats par e-mail. Ou bien, un enquêteur peut déposer une lettre-avis puis téléphoner à un ménage pour l'avertir d'une enquête en face-à-face dont quelques résultats seront ensuite transmis par courrier.

Le questionnement lui-même peut comporter des étapes en face-à-face suivies d'une phase auto-administrée¹.

Les expérimentations qui vont être présentées dans la suite de cette étude concernaient principalement le mode de questionnement et, dans une moindre mesure, le mode de contact, mais n'ont pas porté sur le mode de suivi. On s'est intéressé surtout au questionnement par internet.

Qu'est-ce qu'un « effet de mode » ?

La conception d'une enquête statistique, dans le cadre d'un sondage probabiliste, vise à minimiser l'erreur totale définie comme la somme de l'erreur d'échantillonnage, de l'erreur d'observation et de l'erreur de couverture et de non-réponse.

Chaque mode de collecte comporte des effets propres au regard des différents types d'erreur.

L'erreur d'échantillonnage :

Le mode de collecte peut contraindre le tirage de l'échantillon : lorsqu'un enquêteur se déplace auprès des enquêtés, il ne peut le faire que sur des zones limitées, ce qui entraîne une perte d'efficacité de l'échantillon. Par téléphone, seule la population présente dans l'annuaire peut être couverte.

L'erreur d'observation :

Lorsqu'un enquêteur interroge un enquêté, celui-ci peut avoir tendance à répondre en fonction de ce qu'il croit être la réponse attendue par l'enquêteur, au détriment de la réalité. Il s'agit de l'effet dit de « désirabilité sociale », qui déforme le discours sur la réalité pour la rapprocher d'une norme réelle ou fictive. D'autre part, si l'enquêteur cite à l'enquêté une longue liste de modalités, ce dernier peut avoir tendance à choisir la première pouvant à peu près convenir. C'est ce qu'on appelle, en anglais, l'effet de « satisficing », mot-valise formé de « satisfy » et de « suffice » décrivant la tendance selon laquelle

¹ Par exemple, dans l'enquête Santé et itinéraire professionnel, en face-à-face, quelques questions « délicates », en particulier le module sur la consommation d'alcool, faisaient l'objet d'un questionnaire auto-administré, sur papier en 2006 et sous casque audio en 2010.

la réponse choisie n'est peut-être pas la meilleure mais la plus commode à indiquer tout en restant honnête.²

Ces effets (« désirabilité sociale » et « satisficing ») ne sont attendus que pour les enquêtes en face-à-face ou par téléphone. Mais à l'inverse, pour une enquête auto-administrée sur papier, l'enquêté peut commettre des erreurs dans les filtres (cela étant bien évidemment impossible par internet). L'enquêté peut aussi être tenté de répondre très vite, par la modalité « ne sait pas » par exemple. Il peut surtout être particulièrement difficile pour l'enquêté de reconstituer seul un calendrier ou encore de recalculer des éléments financiers à partir de nombreux documents, tous points pour lesquels un enquêteur peut jouer un rôle, pour motiver, expliquer, clarifier... Enfin, bien que la recherche en ce domaine n'en soit qu'à ses débuts en psychologie et en sociologie, il semble qu'internet, comme il offre la possibilité de se construire une identité que d'aucuns iraient jusqu'à qualifier de « revendiquée », conduise à l'expression d'opinions plus tranchées dans le cas de questionnements subjectifs.

L'erreur de couverture et la non-réponse :

L'intervention d'un enquêteur permet de s'assurer que le répondant est bien dans le champ de l'enquête. Jusqu'à aujourd'hui, elle a permis de parvenir à des taux de réponse significativement supérieurs aux modes auto-administrés : outre la question de l'équipement internet du ménage, la force de persuasion de l'enquêteur est supérieure à celle des lettres-avis et des lettres de relance.³

Par ailleurs, pour les enquêtes bimodales internet-téléphone ou internet-face-à-face, les caractéristiques sociodémographiques des répondants sont potentiellement différentes d'un mode à l'autre, au moins tant que l'accès à internet n'est pas généralisé. C'est une des raisons pour lesquelles il reste nécessaire d'offrir la possibilité d'une réponse sur papier aux enquêtes auto-administrées, par exemple au moment de la relance.

De tout ce qui précède, il ressort que les modes par entretien et auto-administrés ont leurs avantages et leurs inconvénients au regard des différents types d'erreur exposés.

2. Le déroulement de l'expérimentation de l'automne 2010 sur le thème du logement

2.1. Le protocole

Afin d'évaluer le processus de collecte d'informations sur le logement et la mobilité résidentielle par internet, une enquête expérimentale a été réalisée auprès de 10 000 ménages à l'automne 2010, du 15 octobre au 31 décembre.

L'objectif de cette opération était de tester l'organisation à mettre en place en vue d'obtenir des taux de réponse satisfaisants par internet et d'évaluer la qualité des données recueillies par ce mode de collecte en les comparant à celles issues de l'Enquête Nationale Logement 2006 (ENL 2006).

Deux sous-échantillons de 5 000 ménages ont été constitués afin de tester deux questionnaires de longueur différente : un module court (15 à 20 minutes) et un module plus long (25 minutes), cela afin d'apprécier l'influence de la durée d'interrogation sur le taux de réponse.

Un échantillon témoin de 500 logements a été enquêté en face-à-face en décembre 2010, en Lorraine et en Midi-Pyrénées, et ce pour essayer de pallier d'éventuels problèmes venant du décalage temporel entre l'enquête Logement 2006 et le test par internet. Un plus gros échantillon n'a pas été envisageable pour raisons budgétaires.

² Cela peut se corriger en partie par rotation aléatoire des modalités dans les enquêtes sous CAPI.

³ On a parfois observé des enquêtés qui répondaient par respect pour le travail de l'enquêteur (on l'aide à gagner sa vie).

Par ailleurs, en complément du test annuel du dispositif statistique sur les revenus et les conditions de vie (SRCV), il a été proposé aux enquêtés de répondre sur le web à l'enquête expérimentale. Il s'agissait de tester le dépôt de l'invitation à l'enquête web par un enquêteur, en complément d'une autre enquête réalisée en face-à-face⁴.

Enfin, en décembre 2010, une enquête téléphonique s'est déroulée auprès de 1 000 non-répondants de l'enquête auto administrée afin de cerner les causes de la non réponse.

La base de sondage :

Le tirage a été effectué dans le fichier de la taxe d'habitation. En effet, il convient de disposer d'une base de noms, prénoms et adresses sous forme de fichier pour le bon déroulement de l'opération⁵.

Ce tirage en France métropolitaine n'a pas comporté de contrainte de localisation géographique⁶. Par ailleurs, certaines variables présentes dans le fichier de la taxe d'habitation ont été récupérées pour les traitements aval.

Protocole du test par internet :

Une lettre faisant avis de l'enquête par internet a été adressée, début octobre, à chaque ménage⁷.

Deux semaines après (le 21 octobre) a été envoyée une lettre de relance indiquant à nouveau aux ménages qu'ils pouvaient répondre à l'enquête par internet, ainsi qu'un coupon-réponse pour obtenir un questionnaire papier (sans enveloppe préaffranchie).

Une seconde relance par courrier a eu lieu le 20 novembre.

L'application a été fermée en début d'année 2011.

Le questionnaire :

Le questionnaire a été élaboré sur la base de modules de l'Enquête Nationale Logement 2006.

Les questions sont reprises telles quelles pour la plupart. Lorsque des définitions étaient données par l'enquêteur dans l'ENL 2006, elles apparaissent à l'écran sous forme de « bulle » et un lien hypertexte conduit alors à une page de définitions. Dans quelques cas, certaines modalités très peu fréquentes ont été regroupées dans une modalité « autre ».

C'est un questionnaire pouvant s'adresser a priori à l'ensemble des ménages et des situations de logement ordinaire. Deux thèmes sont abordés : la qualité du logement et la mobilité résidentielle.

La structure du questionnaire est la suivante :

- description du logement
- module thématique sur la qualité du logement
- module thématique sur la mobilité résidentielle
- caractéristiques du ménage et du répondant

⁴ Le taux de réponse dans le test couplé à SRCV a été faible. Il est vrai que le thème du logement est déjà très largement abordé dans SRCV et l'enquête par internet a pu paraître redondante. Avant de conclure que, pour les enquêtes en ligne, le mode de contact par courrier fonctionne mieux que le mode de contact en face-à-face, de nouvelles expérimentations seraient nécessaires.

⁵ À ce stade, Octopusse, échantillon maître reposant sur le recensement de la population, n'offre pas cette possibilité (les coordonnées devraient être ressaisies à partir des images des bulletins de recensement).

⁶ Comme la probabilité de recoupement avec d'autres enquêtes est ainsi particulièrement faible, aucun travail spécifique de disjonction n'a été réalisé.

⁷ Cette lettre est malencontreusement partie avant l'ouverture de l'application, la relance a donc été adressée à tout l'échantillon.

Le module contenant les questions sociodémographiques nécessaires pour connaître les caractéristiques du ménage et du répondant a été placé à la fin en vue de ne pas décourager le répondant. Ce module comporte une question sur les revenus, essentiellement dans le but d'évaluer la capacité à répondre à ce type de question via internet.

La version longue du questionnaire comporte plus de questions sur la qualité du logement et sur les causes de mobilité résidentielle. La durée réelle du questionnement a souvent été plus brève que prévue, notamment du fait des filtres, une partie du module sur la mobilité résidentielle ayant concerné peu de répondants.

Une page de résultats sur le logement en France et en Europe, construite avec le département de l'offre éditoriale, était affichée en fin d'enquête, après validation définitive de la réponse. Cette page a été voulue comme un cadeau, remplaçant le cadeau matériel traditionnellement offert dans certaines enquêtes en face-à-face⁸.

Le développement applicatif et l'hébergement du site de l'enquête ont été réalisés par l'Insee, il en a été de même pour les publipostages, le suivi de la collecte, la réponse et l'assistance aux enquêtés.

2.2. Éléments d'appréciation du protocole

Le mode d'organisation est satisfaisant :

L'enquête s'est insérée dans le mode de conception habituel des enquêtes auprès des ménages et l'organisation des travaux s'est fortement inspirée de celle de l'enquête annuelle sur l'utilisation des technologies de l'information et de la communication (TIC) auto-administrée qui a lieu depuis avril 2010.

Le professionnalisme des équipes en charge de l'enquête et leur montée en compétences sur la gestion des enquêtes en ligne sont particulièrement appréciables. Il s'agit désormais de capitaliser et de standardiser les processus, en particulier pour le calendrier de l'envoi des courriers.

L'outil informatique est adapté :

Le développement de l'application de collecte a été très satisfaisant. L'utilisation de Blaise permet une vérification du modèle de données dans un cadre connu de nombreux statisticiens, et il en va de même pour le poste de gestion.

En outre, il est possible de réutiliser tout ou partie de modèles de données existants pour la collecte par enquêteur, c'est ce qui a été pratiqué pour TIC, en assurant ainsi à moindre coût la cohérence entre les différents modes en ce qui concerne le questionnement et le contrôle des réponses.

Les taux de réponse sont encore bas :

Le taux de réponse par internet est de 19,5 %⁹, et le taux de réponse au test est de 24,1 % au total si l'on tient compte des réponses sur papier.

Les taux de réponse pour des enquêtes par internet réalisées par des instituts privés sont quelquefois meilleurs, mais pas toujours, par exemple dans le domaine des enquêtes de satisfaction.

Le taux de réponse est équivalent, qu'il s'agisse du questionnaire court ou du questionnaire long. Il est même très légèrement supérieur pour le sous-échantillon du questionnaire long.

L'enquête téléphonique auprès de 1 000 non-répondants a porté sur les raisons de la non réponse.

⁸ Il n'y a pas eu de décompte du temps passé par les enquêtés à consulter cette page. De même, la consultation des définitions n'a pas été observée dans ce test. Ces décomptes pourraient cependant être prévus lors de nouvelles enquêtes expérimentales.

⁹ Hors adresses erronées (NPAI).

Lors de cette enquête, qui a eu 84 % de répondants, si 29 % ont déclaré ne pas avoir répondu car ils n'avaient pas accès à internet, 11 % parce qu'ils n'avaient pas reçu la lettre-avis, 27 % ont dit ne pas avoir eu le temps de répondre et 15 % ont affirmé ne pas répondre aux enquêtes en général.

Depuis cette expérimentation, une enquête auto-administrée a été conduite au printemps 2011, l'enquête expérimentale sur la qualité de la vie. Son protocole était le suivant :

- une lettre-avis adressée à un échantillon de 9968 individus tiré dans le fichier d'imposition des personnes ; cette lettre invitait à répondre à l'enquête par internet
- une première relance invitant à répondre par internet ou papier (questionnaire joint)
- une deuxième relance invitant à répondre par internet ou papier (en réclamant le questionnaire à l'aide d'un coupon réponse)

Le taux de réponse de cette enquête s'est élevé à 38,4 %¹⁰. Les réponses par internet ont représenté 46 % de l'ensemble.

3. Les résultats de l'expérimentation¹¹

3.1. La non-réponse

Le taux de réponse à l'enquête par internet est bas. La non-réponse ne peut donc pas être ignorée.

De plus, en raison d'un dysfonctionnement lors de la récupération des données de collecte, 1 624 réponses seulement, sur les 1 742 recueillies, ont pu faire l'objet de l'exploitation statistique¹².

Malgré le faible taux de réponse, puisqu'il s'agissait d'une opération purement expérimentale, dont les résultats n'ont pas vocation à être diffusés en tant que tels, il a été décidé de poursuivre l'analyse et de les comparer à ceux de l'Enquête Nationale Logement 2006.

Un des modules n'a cependant pas fait l'objet de cette comparaison : le module sur la mobilité résidentielle passée. En effet, très peu de réponses ont été recueillies. Moins de 20 % des répondants ayant déclaré avoir changé de logement depuis le 1^{er} octobre 2006, cela conduit à des effectifs très faibles pour chaque modalité des différentes questions ce qui rend la comparaison avec l'ENL 2006 peu significative.

De même, certaines questions relatives aux intentions de mobilité (choix de la région et demande de logement HLM) , situées en amont du véritable module sur la mobilité, n'ont pas pu être prises en compte en raison des très faibles effectifs de répondants.

Les taux de réponse en fonction du questionnaire sont les suivants :

Logement Web (questionnaire court)	15,8 %	soit : 788 logements
Logement Web (questionnaire long)	16,7 %	836 logements
Logement face-à-face 2010 « test mini-enquête »	64,8 %	324 logements

On constate une différence d'un point entre le taux de réponse au questionnaire court et le taux de réponse au questionnaire long, différence apparaissant contre-intuitive et qui est sans doute due au plus complet hasard, à moins que la mention, dans la lettre-avis, du fait que le questionnaire dure plus ou moins longtemps donne un aspect plus ou moins « sérieux », « important » à l'enquête (plus long étant perçu comme « plus sérieux »). De telles hypothèses psychologiques ne sont pas extravagantes : on peut observer, par exemple que beaucoup de spectateurs abordent avec davantage de patience, s'appropriant à maintenir une plus longue attention, un film qui dure plus de deux heures qu'un film qui ne fait qu'une heure et demie.

¹⁰ Hors adresses erronées (NPAI).

¹¹ Les auteurs tiennent à remercier particulièrement Octave Falco (Insee-DSDS) qui a bien voulu réaliser les premiers traitements statistiques et la mise en forme des résultats.

¹² Les réponses sur papier n'ont pas été prises en compte dans les traitements, faute de moyens.

3.2. Les caractéristiques socio-démographiques

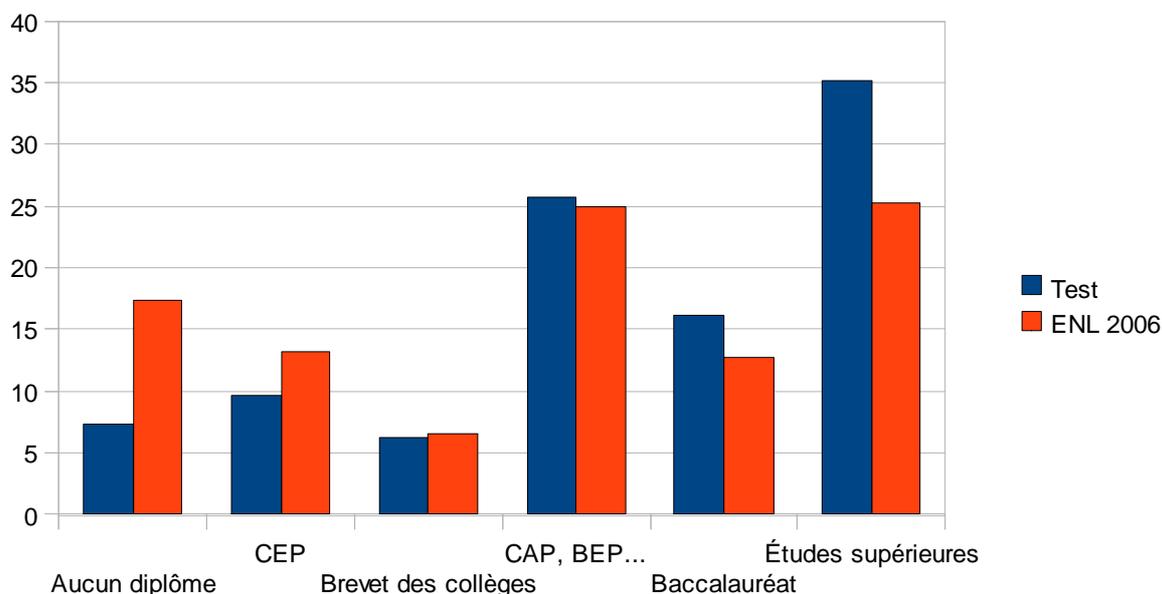
Dans neuf cas sur dix, la personne ayant répondu au test est celle à qui la lettre-avis était adressée, c'est-à-dire celle dont le nom figure dans le fichier de la taxe d'habitation. Ceci nous a permis de faire l'approximation que le répondant était la « personne de référence » du ménage (au sens des enquêtes auprès des ménages réalisées par l'Insee) et ainsi de pouvoir comparer les caractéristiques socio-démographiques déclarées par le répondant avec celles de la personne de référence dans l'Enquête Nationale Logement 2006¹³.

L'âge moyen des répondants au test était de 54 ans, alors que l'âge moyen de la personne de référence dans l'ENL 2006 était de 49 ans. Il s'agissait d'hommes dans 60 % des cas, légèrement moins que dans l'ENL 2006.

Les répondants du test sont plutôt des diplômés. Cela confirme le biais de sélection des enquêtes par internet, ou même auto-administrées sur papier, qui supposent la maîtrise de l'outil informatique et à tout le moins de l'écrit¹⁴.

F2. Quel est le diplôme le plus élevé que vous ayez obtenu ?

DIPLOME	Test Brut	ENL 2006 Pondéré
1. aucun diplôme	117 7.30	17.38
2. un CEP (certificat d'études primaires) ou diplôme étranger de même niveau	154 9.61	13.15
3. le brevet des collèges, BEPC, brevet élémentaire ou diplôme étranger de même niveau	100 6.24	6.58
4. CAP, BEP ou diplôme de ce niveau	411 25.64	24.87
5. baccalauréat, capacité en droit, diplôme d'éducateur spécialisé ou diplôme étranger de même niveau	259 16.16	12.81
6. diplôme d'études supérieures	562 35.06	25.21



¹³ Il aurait été possible de donner aux enquêtés la définition de la personne de référence et d'indiquer que c'est à elle de répondre au questionnaire. Mais d'une part le mode auto-administré ne nous permet pas de vérifier que la personne répondant au questionnaire est bien celle qui était censée répondre, d'autre part cela aurait sans doute encore augmenté le taux de non-réponse.

¹⁴ Au contraire, le mode auto-administré sous casque audio peut permettre de toucher plus de répondants, en particulier les personnes ayant des difficultés pour la lecture.

En ce qui concerne la situation vis-à-vis de l'emploi, la proportion relativement élevée d' « autres situations » tend à montrer que certains répondants au test, sans l'aide d'un enquêteur, ont eu du mal à se classer. Quoi qu'il en soit, le déficit d'actifs est à souligner.

F3. Quelle est votre situation principale vis-à vis du travail ?

SITUA	Test Brut		ENL 2006 Pondéré
1. vous occupez un emploi	839	51.85	57.69
2. vous êtes apprenti	1	0.06	0.16
3. vous êtes étudiant	8	0.49	1.70
4. vous êtes chômeur	70	4.33	4.58
5. vous êtes retraité	565	34.92	32.10
6. vous êtes homme ou femme au foyer	31	1.92	1.75
7. vous êtes dans une autre situation	104	6.43	2.02

Le « nombre de personnes vivant dans le logement » est une notion difficile à définir. En auto-administré, on peut avoir du mal à obtenir de l'enquêté qu'il adopte une définition proche des concepts des statisticiens¹⁵.

Quoi qu'il en soit, le nombre de personnes seules est beaucoup plus faible dans le test que dans l'Enquête Nationale Logement 2006.

F7. Combien de personnes vivent-elles dans ce logement ?

NHAB	Test Brut		ENL 2006 Pondéré
1	390	24.36	33.46
2	606	37.85	32.32
3	236	14.74	14.85
4	258	16.11	13.14
5 et plus	111	6.93	6.23

De même, les répondants au test appartiennent à des ménages où les moins de 14 ans sont plus nombreux.

F8. Parmi ces personnes, combien ont moins de 14 ans ?

NENF14	Test Brut		ENL 2006 Pondéré
0	1133	73.33	76.08
1	202	13.07	12.33
2	165	10.68	8.64
3 et plus	40	2.59	2.94

Notons enfin que le taux de réponse varie en fonction du statut d'occupation du logement tel qu'il apparaît dans la base de sondage : il est de 21 % pour les propriétaires et de 15 % pour les locataires.

3.3. Les questions générales sur le logement

¹⁵ Par exemple : dans le tronc commun des enquêtes auprès des ménages, les habitants du logement sont les personnes qui habitent au moins un mois dans l'année (en cumulé) dans le logement.

Les résultats du test font apparaître moins de résidents en maison individuelle et plus en appartement en immeuble collectif que les résultats de l'Enquête Nationale Logement 2006. Les répondants du test, sans l'aide d'un enquêteur, ont eu plus de mal à se classer, comme le montrent les 5,6 % de « autre ».

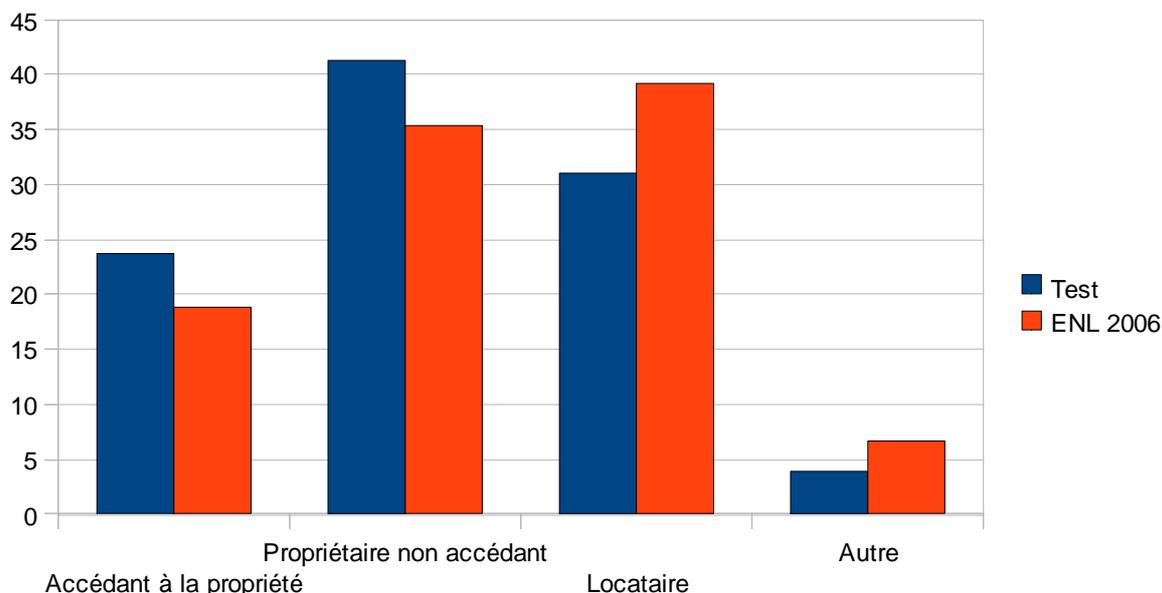
Q1. Votre logement est-il ?

HTCL	Test Brut		ENL 2006 Pondéré
1. une maison individuelle	960	59.19	56.12
2. un appartement dans un immeuble collectif	604	37.24	42.04
3. autre	58	3.58	1.83

Le déficit de locataires est cohérent avec le profil des répondants dégagé jusqu'à présent : des personnes plus âgées et plus diplômées qu'en moyenne dans la population. En ce qui concerne la répartition entre propriétaires accédants et non accédants, elle peut aussi provenir du mode de questionnement et d'une mauvaise compréhension des définitions, sans aide de l'enquêteur.

Q2. Occupez-vous ce logement comme ?

STOC	Test Brut		ENL 2006 Pondéré
1. accédant à la propriété (vous remboursez des emprunts pour l'achat de ce logement)	452	28.04	18.79
2. propriétaire non accédant (vous ne remboursez pas ou plus d'emprunt pour l'achat de ce logement)	653	40.51	35.37
3. locataire	450	27.92	39.18
4. autre	57	3.54	6.66



D'après les réponses des locataires, les propriétaires institutionnels sont sous-représentés dans le test. On peut penser que l'enquêteur, dans l'ENL 2006, propose au ménage de faire référence au bail ou aux quittances de loyer, ce qui permet de vérifier si le propriétaire est un particulier ou non. En mode auto-administré, une telle vérification n'est pas possible.

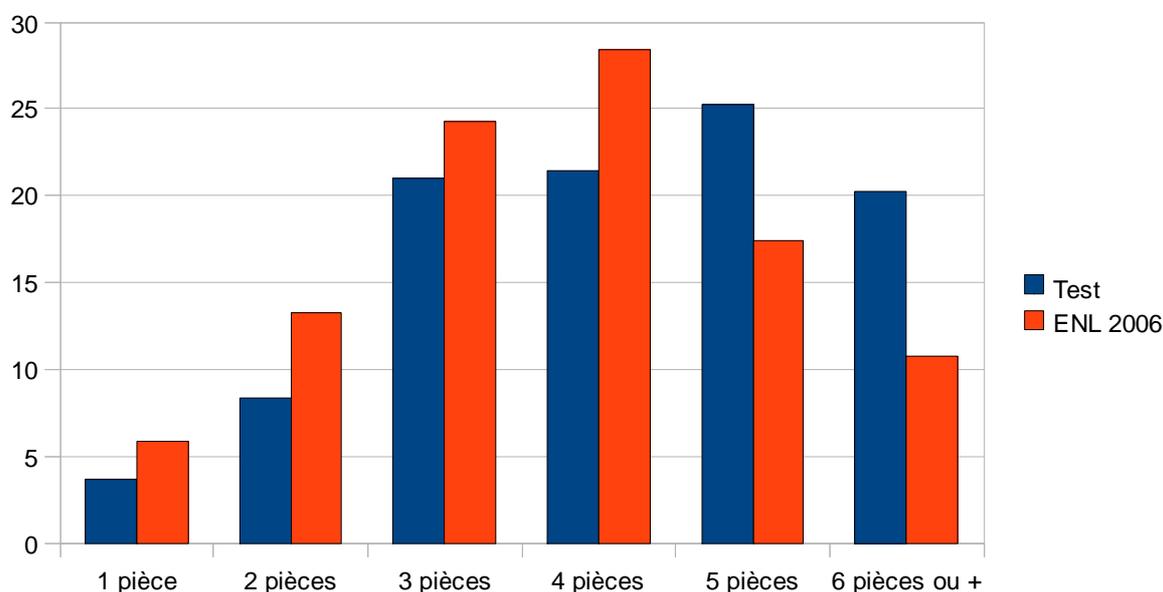
Q3. Qui est propriétaire de ce logement ? (question posée au locataires)

PROPRI	Test Brut		ENL 2006 Pondéré
1. un particulier	219	43.45	47.68
2. un organisme HLM ou assimilé (OPAC, offices, sociétés, fondations)	224	44.44	37.57
3. une administration, une banque, une compagnie d'assurances	25	4.96	8.52
4. un membre de votre famille	36	7.14	6.23

La distribution du nombre de pièces diverge beaucoup entre le test web et l'ENL 2006, notamment en ce qui concerne les 5 pièces et 6 pièces et plus. Une bulle sur le questionnaire internet précisait ce qu'est une pièce : « Comptez les pièces d'habitation y compris la cuisine si elle fait plus de 12 m² et les annexes utilisées par le titulaire du logement. Ne comptez pas les pièces à usage exclusivement professionnel. » Il est possible qu'elle n'ait pas été lue par certains ou mal comprise. Il est également possible que cela soit en cohérence avec la surreprésentation des diplômés et des revenus mensuels moyens élevés.

Q6. Combien de pièces compte votre logement ?

HPH	Test Brut		ENL 2006 Pondéré
1. 1 pièce	43	2.65	5.83
2. 2 pièces	129	7.96	13.30
3. 3 pièces	303	18.69	24.28
4. 4 pièces	371	22.89	28.39
5. 5 pièces	411	25.35	17.41
6. 6 pièces ou plus	364	22.46	10.79



Les données recueillies sur la surface sont très proches, qu'elles portent sur la surface déclarée spontanément (question Q7) ou par tranche de surface. Là encore une bulle indiquait la définition reprise des instructions aux enquêteurs de l'ENL 2006 : « Hors véranda, et pièces à usage exclusivement professionnel, y compris les pièces annexes ».

Q7. Quelle est la surface de votre logement ?

La distribution des surfaces par tranche révèle une sur-représentation dans l'échantillon des logements avec surfaces habitables de plus de 100 m².

HSHT	Test Brut	ENL 2006 Pondéré
1. de 1 à 19 m ²	0.63	0.80
2. de 20 à 39 m ²	4.25	6.33
3. de 40 à 59 m ²	11.51	12.81
4. de 60 à 79 m ²	17.01	22.32
5. de 80 à 99 m ²	21.20	21.96
6. de 100 à 119 m ²	17.26	14.29
7. de 120 et plus	28.14	21.49

Comme le tirage de l'échantillon s'est fait dans la taxe d'habitation, il a été possible de comparer les réponses à l'enquête aux données administratives correspondantes¹⁶. Les résultats de ces comparaisons se trouvent en annexe 4. Ils révèlent quelques divergences entre les données issues de la taxe d'habitation et les données collectées, mais cette observation ne semble pas liée au mode de collecte.

3.4. La pondération des données collectées : le calage sur marges

Les tableaux précédents permettent de constater que **la structure des répondants au test, pour ce qui concerne les caractéristiques socio-démographiques et pour leurs conditions matérielles de logement, est très différente de celle de l'enquête Logement 2006.**

Pour confirmer ces résultats, des tests de Kolmogoroff-Smirnoff ont été réalisés sur les deux sous-populations des répondants et des non-répondants, sur les variables recueillies à partir de la base de sondage (âge, statut d'occupation du logement, surface du logement dans la taxe d'habitation). Ils ont conclu, compte tenu de leur niveau d'exigence élevé, **à la non-identité de structure deux sous-populations.**

Un calage sur marges était donc tout à fait indispensable.

Un premier calage a été effectué grâce à la procédure CALMAR, utilisée à partir de six variables :

- l'âge du répondant (qu'on a supposé être la personne de référence),
- son sexe,
- son niveau de diplôme,
- le nombre de personnes habitant le logement
- le nombre de pièces du logement
- et la tranche d'unité urbaine du logement

Les marges proviennent de l'Enquête Emploi en Continu 2010 (EEC 2010). Les pondérations ont été obtenues en sortie du calage par la méthode raking-ratio (modèle M2).

L'objectif de ce calage était de rendre plus comparables les résultats obtenus aux autres questions.¹⁷

¹⁶ La déclaration de l'enquête à la CNIL avait explicitement envisagé cette possibilité.

¹⁷ Afin de rendre encore plus comparables les données du test et celle de l'enquête Logement 2006 (sous l'hypothèse forte que le profil sociodémographique de la population et les conditions de logement n'ont pas changé entre 2006 et 2010), nous avons calé l'enquête Logement 2006 sur les mêmes marges que celles utilisées pour le test. Les résultats de ce calage étant très similaires aux données pondérées de cette enquête calculées à partir des marges de 2006, nous avons décidé de conserver les résultats pondérés officiels de l'époque, qui ont le mérite de représenter une réalité concrète : l'état du logement en 2006.

Cependant, les résultats de ce calage n'ont pas été très satisfaisants, car les variables non calées conservaient des valeurs éloignées de celles de l'Enquête Nationale Logement 2006. Le nombre d'itérations nécessaires au calage (18) indiquait déjà que la structure des répondants et celle de la population totale sont très différentes. Par conséquent, nous avons décidé de procéder à un second calage, sur davantage de variables. Les marges proviennent cette fois de l'échantillon, donc de la taxe d'habitation (TH). En plus des variables précédentes, nous avons donc calé sur :

- le revenu fiscal avant abattement (en déciles)
- la surface du logement (en quintiles)
- la nature du logement (maison / appartement)
- le statut d'occupation du logement (propriétaire / locataire)

Ce second calage donne des résultats plus satisfaisants concernant les revenus, dont nous allons maintenant parler.

3.5. Une variable difficile à observer : les revenus

Un des enjeux du test était de savoir si les enquêtés répondraient à la question sur leur revenu mensuel moyen. En effet, on a souvent constaté dans les enquêtes auprès des ménages que cette question dérangeait certaines personnes, et on pouvait s'attendre à davantage de refus de répondre sur internet qu'en face-à-face, sans la présence d'un enquêteur pour les convaincre.

La question était ainsi formulée : « **F9. Quel est au total le revenu mensuel moyen de votre ménage ?** »

Il était demandé d'indiquer en clair le montant du revenu mensuel moyen, mais en cas de non-réponse, des tranches de revenus étaient proposées. En cas de nouvelle non-réponse, la case « je ne souhaite pas répondre » s'affichait et le ménage pouvait la sélectionner.

Enquête	Répondants	Rép. question	% rép. quest.	Rép. en clair	% rép. clair	Rép. en tranche	% rép. tranche
Web (courte)	788	737	93,5	663	84,1	74	9,4
Web (longue)	836	796	95,2	710	84,9	86	10,3
Face-à-face	324	234	72,2	234	72,2	0	0,0

Au total, seulement 6,5 % des répondants n'ont pas répondu à cette question, ce qui est inférieur à la non-réponse habituellement constatée dans le tronc commun des enquêtes auprès des ménages (TCM) en face-à-face (environ 25 % dans les enquêtes sans ré-interrogation, dont les neuf dixièmes en clair).

Les revenus bruts obtenus par internet (moyenne : 3 318 € / mois, médiane : 2 500 €¹⁸) sont très supérieurs aux revenus (bruts et pondérés) obtenus dans les enquêtes ménages récentes (moyenne entre 2 400 et 2 500 € et médiane entre 2 000 et 2 050 €).

Nous avons d'abord tenté de corriger ces revenus via un programme d'apurement qui avait développé pour l'enquête HSM 2008 et dont la possibilité d'utilisation a ensuite été étendue à toutes les enquêtes utilisant le tronc commun des enquêtes auprès des ménages (TCM). Ce programme fonctionne en trois étapes :

- une régression permettant de repérer et de supprimer les valeurs aberrantes
- une imputation des revenus en non-réponse, sur la base des caractéristiques socio-démographiques des ménages (en cas de réponse en tranche, le revenu imputé est dans la tranche déclarée)

¹⁸ Ici nous n'avons pris en compte que les réponses en clair, car les revenus déclarés en tranche seront imputés durant le programme d'apurement (voir ci-après). Cependant, si à ce stade on impute la moyenne de tranche, on obtient une moyenne des revenus de 3 294 € / mois, très proche donc de la valeur précédente.

- un lissage par moyenne mobile¹⁹.

Après apurement, le revenu moyen obtenu était de 2 876 € / mois et le revenu médian de 2 489 € / mois. (La différence par rapport à la moyenne brute est donc de 450 €.) Il est à noter toutefois que le nombre de valeurs jugés aberrantes durant la première phase du programme d'apurement peut varier beaucoup selon le seuil de rejet que l'on retient (ici on a rejeté les observations pour lesquelles le résidu était supérieur à 4).

Après le premier calage (sur les marges de l'EEC), le revenu pondéré moyen était de 2 506 € /mois et le revenu médian de 2 053 € / mois. On peut constater l'effet très bénéfique du calage pour la vraisemblance des résultats.

Cependant, le second calage (sur les marges de l'EEC et sur celles de l'échantillon issu de la TH) donne de meilleurs résultats encore, et c'est ceux que nous retiendrons : une moyenne de revenus de 2 489 € / mois et une médiane de 2 028 € / mois. Ces valeurs s'approchent, tout en leur restant quelque peu supérieures, des valeurs observées dans l'enquête CVS 2010 : une moyenne de 2 410 € / mois et une médiane de 2 012 € / mois (après apurement)²⁰.

Dans l'enquête en face-à-face, le revenu mensuel brut est de 2 353 € / mois et le revenu après apurement et calage de 2 245 € / mois. Il s'agit d'un calage très osé puisque effectué sur les seuls 324 logements répondants et ne pouvant donc, vu ce faible effectif, servir qu'à titre indicatif pour comparaison (les variables de calage étaient d'ailleurs un peu moins nombreuses). Notons que cette valeur est inférieure à celle trouvée dans les enquêtes ménages comme CVS 2010.

Enfin, à titre indicatif, dans l'échantillon issu du fichier de la taxe d'habitation, le revenu du ménage avant abattement est de 2 606 € / mois²¹. On constate donc une différence de plus de 270 € entre la valeur TH et la valeur déclarée dans l'enquête logement web.

Voici un tableau récapitulatif :

Enquête	Revenu brut moyen	Revenu brut médian	Revenu pondéré moyen	Revenu pondéré médian
Web ²²	3 318	2 500	2 489	2 028
Face-à-face	2 353	2 000	2 245	1 923
CVS 2010	2 444	2 000	2 410	2 000
Échantillon TH			2 606	2 012

Nous concluons de ces résultats que les résultats bruts obtenus par le test web à la question sur les revenus sont assez mauvais (personne ne pouvant contrôler la réponse du ménage), et qu'un apurement et un calage poussé sont nécessaires pour pouvoir comparer les revenus issus de déclarations faites sur internet avec les revenus issus de collectes en face-à-face.

* *
*

Dans les paragraphes suivants, nous présentons les résultats bruts et pondérés par calage obtenus à la plupart des questions du test par internet en les comparant aux résultats pondérés obtenus lors de

¹⁹ Pour de plus amples détails sur le fonctionnement de ce programme et pour l'utiliser, se rapport à la note enregistrée N° 1838/DG75-F420.

²⁰ Les revenus après apurement obtenus dans CVS 2011 ont presque les mêmes moyenne et médiane qu'en 2010. Ce résultat nous intéresse dans la mesure où le test web 2010 a été réalisé quelques semaines avant le début de CVS 2011.

²¹ Cette valeur a été calculée en divisant par douze le revenu du ménage avant abattement issu du fichier de la taxe d'habitation. Mais cette grandeur n'est pas vraiment comparable avec le revenu déclaré par les ménages lors d'enquêtes car elle ne tient pas compte d'éventuelles prestations sociales.

²² Ici encore il ne s'agit que des revenus déclarés en clair.

l'Enquête Nationale Logement 2006. Des graphiques accompagnent les tableaux lorsque ceux-ci révèlent des différences notables.

3.6. Les questions d'opinion sur le logement

Les deux premières questions qualitatives sur les conditions de logement ont été posées à l'ensemble des répondants. Elles conduisent à des réponses proches de celles de l'enquête Logement 2006.

Q9. Estimez-vous que vos conditions de logement sont... ?

OLA	Test Brut		Test Pondéré		ENL 2006 Pondéré
1. très satisfaisantes	476	29.51	24,79		28.07
2. satisfaisantes	795	49.29	48,64		45.93
3. acceptables	251	15.56	19,86		19.32
4. insuffisantes	63	3.91	5,24		5.05
5. très insuffisantes	28	1.74	1,48		1.63

Q10. Vous plaisez-vous dans votre quartier ?

OQA	Test Brut		Test Pondéré		ENL 2006 Pondéré
1. oui	1465	92.43	92,44		92.51
2. non	120	7.57	7,56		7.49

Les questions suivantes ont été posées seulement aux répondants du questionnaire long²³.

Certaines questions conduisent à des réponses foncièrement différentes, **d'où la suspicion de réels effets de mode** sur cette partie du module qualitatif. Ces effets tiennent sans doute à la fois à la présence ou non d'un enquêteur mais également au biais de sélection de l'enquête par internet dont les répondants sont probablement les personnes qui ont le plus à se plaindre de leur logement. En effet, les réponses sont presque toujours plus « négatives » dans le test web que dans l'enquête logement 2006 et dans le test en face-à-face. Il convient ici de souligner que le calage sur marges a permis de corriger les biais de sélection pour les variables observables alors que ces biais n'ont pas été corrigés pour les variables inobservables.

C'est le cas en ce qui concerne la perception de la température : alors que 82,6 % n'ont pas eu de problème au cours du dernier hiver pour l'enquête Logement 2006, c'est seulement 64,8 % qui sont dans ce cas avec le test²⁴. Il est possible que cela corresponde à un biais de désirabilité sociale, si l'on considère qu'il peut être difficile de reconnaître une telle privation de confort en face-à-face. Ce phénomène est encore plus remarquable si on prend en considération le biais en faveur d'une population en haut de la hiérarchie sociale (davantage diplômés, habitant un logement plus grand, etc.).

Les causes déclarées de la mauvaise température sont également très différentes. La mauvaise isolation et les raisons financières sont citées beaucoup plus fréquemment.

²³ Les données pondérées suivantes proviennent d'un calage effectué uniquement sur les données du questionnaire long. Pour ces questions, il y avait en effet de légères différences entre les données brutes provenant des deux questionnaires court et long, mais rien de très important.

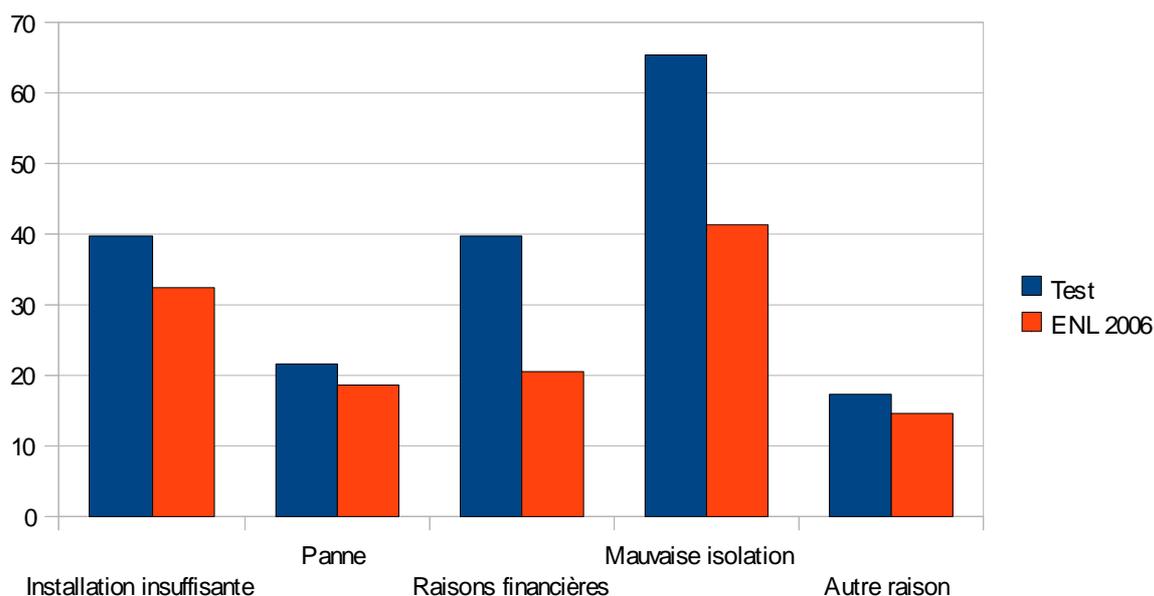
²⁴ Ces impressions semblent contredire les données de Météo France. En effet, l'hiver 2005-2006 a été le plus froid de ces vingt-cinq dernières années, avec une moyenne de température hivernale à -1,6 °C (pour le s mois de décembre-janvier-février). Tandis que l'hiver 2009-2010 a été un peu moins froid, avec une moyenne à -1,2 °C.

Q11. Au cours de l'hiver dernier, dans votre logement, votre ménage a-t-il souffert, pendant au moins 24 heures... ?

GCHAUF	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. du froid	125	21.04	21,42	14.53
2. d'un excès de chauffage	11	1.85	1,58	2.10
3. du froid et d'un excès de chauffage	35	5.89	7,6	0.78
4. d'aucun problème de température dans le logement	423	71.21	69,39	82.59

Q12. Votre ménage a souffert du froid, était-ce dû à ... ? (plusieurs réponses possibles)

GCHAUFC	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. l'installation qui est insuffisante ?	31.25		39.57	32.40
2. une panne de votre installation (ou du chauffage collectif) ?	16.25		21.58	18.59
3. des raisons financières : vous limitez votre chauffage en raison de son coût ?	34.38		39.56	20.60
4. une mauvaise isolation de votre logement ?	59.38		65.22	41.23
5. d'autres raisons (intempéries, mise en route tardive du chauffage collectif, ...) ?	15.63		17.27	14.44



En ce qui concerne la perception du bruit, les résultats sur la fréquence sont très proches selon les deux modes de questionnement. En revanche, la cause des bruits invoquée est différente. Le mode auto-administré incrimine nettement plus le voisinage (établissements environnants et voisins) que le face-à-face.

Q13. Pendant la journée, les bruits que vous entendez dans votre logement (fenêtres fermées) sont-ils... ?

KBJ	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. rares ou inexistants	435	73.11	72,34	69.50
2. assez fréquents	110	18.49	18,89	20.60
3. très fréquents	50	8.40	8,78	9.89

Q14. Leur origine principale est-elle... ?

KBJO	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. la circulation (voitures, trains, avions, ...)	82	51.25	45,95	61.90
2. les établissements environnants (usines, commerces, écoles, ...)	7	4.38	2,44	4.42
3. les installations techniques du logement ou de l'immeuble (ascenseur, chaufferie, ventilation, ...)	5	3.13	3,63	1.13
4. Les bruits de voisinage (bricolage, musique, enfants, chiens, ...)	66	41.25	47,98	32.55

Q15. Durant la nuit, les bruits que vous entendez dans votre logement (fenêtres fermées) sont-ils... ?

KBN	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. rares ou inexistants	504	85.57	83,99	83.11
2. assez fréquents	73	12.39	13,58	12.62
3. très fréquents	12	2.04	2,43	4.27

Q16. Leur origine principale est-elle... ?

KBNO	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. la circulation (voitures, trains, avions, ...)	39	46.43	40,33	49.60
2. les établissements environnants (usines, commerces, écoles, ...)	7	8.33	6,69	6.79
3. les installations techniques du logement ou de l'immeuble (ascenseur, chaufferie, ventilation, ...)	4	4.76	5,82	2.19
4. Les bruits de voisinage (bricolage, musique, enfants, chiens, ...)	34	40.48	47,16	41.41

Les questions suivantes font également apparaître d'importantes divergences entre les deux enquêtes. Elles concernent l'exposition du logement, le voisinage, l'accès au logement, la qualité de l'air, la sécurité et l'accessibilité du logement.

Q17. Trouvez-vous que l'exposition de votre logement (clarté, ensoleillement) est... ?

KEXPO	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. bonne	444	75.64	71,28	81.70
2. moyenne	112	19.08	23,33	14.85
3. mauvaise	31	5.28	5,38	3.45

Q18. Trouvez-vous que vos relations avec le voisinage sont... ?

KVOIS	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. bonnes	388	65.32	63,49	77.08
2. moyennes	142	23.91	23,59	10.35
3. mauvaises	9	1.52	1,45	1.43
4. je n'ai pas de relations avec le voisinage	55	9.26	11,46	11.14

Q19. Trouvez-vous que l'accès à votre logement (porte d'entrée, couloirs, escaliers, ascenseurs) est... ?

KACCESS	Test Brut		Test Pondéré	ENL 2006 Pondéré
----------------	------------------	--	---------------------	-------------------------

1. facile	314	53.77	50.53	87.56
2. normal	227	38.87	42.44	9.85
3. moyen	43	7.36	7.03	2.60

Q20. Trouvez-vous que la qualité de l'air de votre quartier (poussières, pollution) est... ?

KAIR	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. bonne	388	65.87	65,45	69.04
2. moyenne	172	29.20	29,46	23.30
3. mauvaise	29	4.92	5,09	7.66

Q21. Trouvez-vous que la sécurité de votre quartier (risques de vol, d'agression) est...

KSECU	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. bonne	366	61.93	60,41	73.74
2. moyenne	184	31.13	31,52	21.59
3. mauvaise	41	6.94	8,07	4.67

Q 24. Trouvez-vous que l'accessibilité de votre logement par les voitures particulières (stationnement dans le quartier, embouteillages) est...

KVOITU	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. bonne	352	59.36	57,28	70.26
2. moyenne	161	27.15	28,03	16.82
3. mauvaise	80	13.49	14,68	12.92

Il en va de même pour l'entretien des rues et la proximité des commerces, pour laquelle le biais de sélection a dû jouer, les répondants du test, vivant plus souvent en immeuble collectif, disposant de plus de commerce, en zone urbaine. Il semble en aller de même pour les transports en commun.

Q 25. Trouvez-vous que l'entretien des rues ou de la voirie et des espaces publics est ?

KRUE	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. bon	304	52.69	50,39	66.40
2. moyen	199	34.49	35,35	24.13
3. mauvais	74	12.82	14,26	9.47

Q 26. Trouvez-vous que la proximité des commerces est ?

KCOM	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. bonne	337	56.54	53,76	63.07
2. moyenne	159	26.68	29,45	17.49
3. mauvaise	43	7.21	7,22	5.63
4. il n'y a pas de commerces	57	9.56	9,58	13.82

Q23. Trouvez-vous que l'accessibilité des transports en commun est ?

KTRANS	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. bonne	284	47.89	46,05	54.85
2. moyenne	142	23.95	21,73	14.60
3. mauvaise	76	12.82	14,72	8.78
4. pas de transports en commun	91	15.35	17,5	21.77

La question sur les espaces verts est un cas particulier : il semble que le terme « espaces verts » n'a pas été compris de la même manière si l'on en juge la modalité « il n'y a pas d'espace vert » dans l'enquête Logement 2006. On s'attendrait en effet à ce que ce soient les répondants au test, plus urbains, qui constatent l'absence d'espace verts ou de campagnes proches.

Cela semble révéler le caractère ambigu de la notion « d'espace vert », dont le contenu peut correspondre à des réalités très diverses.

Q22. Trouvez-vous que la présence et la qualité des espaces verts dans votre quartier (ou des campagnes lorsqu'elles sont proches) sont ?

KVERT	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. bonne	375	64.77	62,06	66.11
2. moyenne	151	26.08	26,77	15.92
3. mauvaise	24	4.15	3,72	4.10
4. il n'y a pas d'espaces verts	29	5.01	7,46	13.87

3.7. Les prévisions de mobilité

Le test fait ressortir des prévisions de mobilité nettement plus élevées que dans l'enquête Logement 2006, soit que les habitants souhaitent déménager, soit qu'ils pensent y être bientôt contraints.

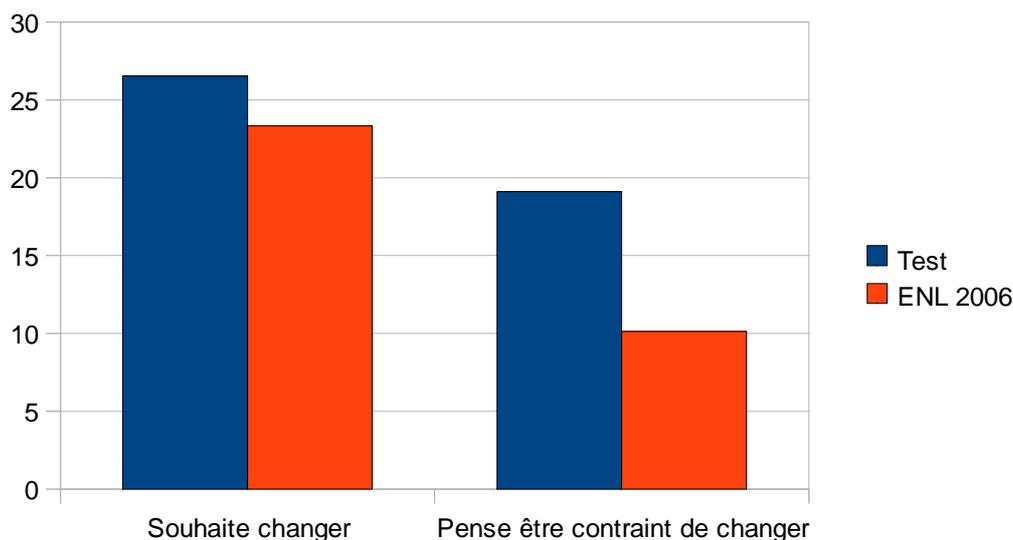
L'effet de mode de questionnement peut jouer. Cependant, durant les quatre années qui séparent les deux opérations statistiques, le marché du logement a connu des tensions accrues, et il serait hâtif de conclure au seul effet de mode.

P1. Souhaitez-vous changer de logement ?

ODL	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. oui	417	25.79	26.44	23.32
2. non	1200	74.21	73.56	76.68

P2. Pensez-vous être contraint de quitter votre logement dans les trois ans à venir ?

ODF	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. oui	289	18.02	19.05	10.03
2. non	1315	81.98	80.95	89.97



Si les raisons professionnelles ou familiales prédominent pour le test comme pour l'enquête Logement 2006, les « autres raisons » sont nettement plus nombreuses dans le test, ce qui correspond bien aux difficultés des répondants à classer leur situation sans l'aide d'un enquêteur.

P3. Pour quelle raison pensez-vous y être contraint ? (question posée aux personnes ayant répondu « oui » à P2)

ODFRA	Test		ENL 2006 Pondéré
	Brut	Pondéré	
1. raisons professionnelles : mutation, études	44	15.44	17.62
2. votre propriétaire veut reprendre le logement	7	2.46	1.65
3. votre logement vous est prêté ou loué provisoirement	4	1.40	1.25
4. raisons financières	45	15.79	18.90
5. raisons familiales : modification de la taille de la famille	68	23.86	21.61
6. raisons de santé	22	7.72	8.97
7. retraite, fin d'activité	20	7.02	6.73
8. autre raison	75	26.32	23.27

Les prévisions de mobilité par zone géographique ont donné lieu à peu de réponses lors du test et ne peuvent donc pas être comparées à l'enquête Logement 2006.

Pour les caractéristiques et le statut d'occupation du futur logement, ici encore l'indétermination des « ne sait pas » est plus forte dans le test, même si certaines modalités ont des résultats très comparables.

P8. Envisagez-vous d'occuper ...

DTL	Test		ENL 2006 Pondéré
	Brut	Pondéré	
1. une maison individuelle	214	45.73	46.88
2. un appartement	182	38.89	35.33

3. une chambre ou un logement dans une structure collective (foyer, résidence pour personnes âgées, ...)	10	2.14	2.82	1.83
4. autre	7	1.50	0.98	0.71
5. ne sait pas	55	11.75	13.99	4.05

Le résultat obtenu à la question P9 corrobore ceux indiquant qu'il y avait une proportion inhabituellement grande, parmi les répondants, de personnes âgées, inactives.

P9. Par rapport à votre logement actuel, envisagez-vous d'occuper un logement ...

DTA	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. plus grand (surface ou nombre de pièces)	237	50.43	48.34	55.88
2. plus petit (surface ou nombre de pièces)	90	19.15	19.62	16.10
3. de taille équivalente	110	23.40	23.77	26.68
4. ne sait pas	33	7.02	8.28	1.35

P10. Dans votre futur logement, envisagez-vous d'être ...

DST	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. propriétaire	273	58.21	54.52	48.87
2. locataire	132	28.14	28.82	45.44
3. autre	9	1.92	2.59	3.27
4. ne sait pas	55	11.73	14.07	2.42

Étonnamment, ce sont les questions plus précises sur les démarches qui offrent les résultats les plus proches, avec peu d'indétermination pour le test. Leur formulation est bien adaptée à un mode auto-administré, en particulier la modalité « vous n'avez encore rien fait ».

P11. Où en êtes-vous de vos démarches (achat) ?

DDEMP	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. vous n'avez encore rien fait	153	56.25	51.91	56.27
2. vous avez débuté vos recherches	63	23.16	22.91	24.68
3. vous êtes en phase de réservation, de compromis de vente ou vous avez signé la vente	43	15.81	19.79	15.05
4. ne sait pas	13	4.78	5.39	4.00

P12. Où en êtes-vous de vos démarches (location)?

DDEML	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. vous n'avez encore rien fait	66	50.77	48.28	55.41
2. vous avez débuté vos recherches	51	39.23	39.88	34.01
3. vous êtes en phase de réservation, de compromis de vente ou vous avez signé la vente	9	6.92	5.97	6.98
4. ne sait pas	4	3.08	5.88	3.60

Concernant la date du futur départ, les réponses sont très comparables pour les intentions à moins de deux ans ; au-delà le mode auto-administré conduit à beaucoup plus d'indétermination qu'avec le face-à-face.

P13. Quand envisagez-vous de quitter votre logement ?

DDAT	Test Brut		Test Pondéré	ENL 2006 Pondéré
1. dans moins de 6 mois	60	12.90	14.37	17.79
2. d'ici 6 mois à moins d'un an	67	14.41	13.53	16.52
3. d'ici un à moins de deux ans	76	16.34	15.43	21.07
4. dans deux ans ou plus	95	20.43	17.28	26.66
5. ne sait pas	167	35.91	39.39	17.96

3.8. Les différences observées viennent-elles du mode de collecte ? un modèle LOGIT

Afin d'observer l'influence de la source (test web ou enquête Logement 2006) et des variables socio-démographiques sur les réponses aux questions, nous avons conçu un modèle de régression logistique utilisable pour expliquer des variables dichotomiques ou polytomiques ordonnées.

Modèle retenu :

Les variables explicatives sont les suivantes :

Source	s1	test web
	s2	Logement 2006
Âge du répondant	age1	18-24 ans
	age2	25-44 ans
	age3	45-64 ans
	age4	64 ans et +
Sexe du répondant	sex1	homme
	sex2	femme
Revenu mensuel moyen du ménage	r1	1er quintile de revenu
	r2	2e quintile de revenu
	r3	3e quintile de revenu
	r4	4e quintile de revenu
	r5	5e quintile de revenu
Diplôme	d1	aucun diplôme
	d2	CEP
	d3	BEPC
	d4	CAP, BEP
	d5	BAC, capacité en droit
	d6	études supérieures
Nombre d'habitants du logement	nh1	1
	nh2	2
	nh3	3 ou 4
	nh4	5 et +
Situation du répondant vis-à-vis de l'emploi	sit1	emploi
	sit2	étudiant, apprenti
	sit3	chômeur
	sit4	retraité, au foyer, invalide
Tranche d'unité urbaine	t1	rural (< 2 000 hbts)
	t2	ville < 50 000 hbts
	t3	ville > 50 000 hbts
	t4	Paris + agglo

Nous n'avons pas pu comparer directement les données issues du test web avec celles issues de la mini-enquête en face-à-face, et cela pour deux raisons : 1° le faible nombre d'observations obtenues à cette mini-enquête et 2° la mini-enquête comportait seulement les questions du questionnaire web court, et pas les questions d'opinion présentes seulement dans le questionnaire long ; or ce sont surtout les réponses à ces questions qui nous intéressent ici.

Cependant, pour les questions présentes dans les deux questionnaires (questions « objectives » et quelques questions d'opinion), le même modèle logit a été testé sur les données test web + mini-

enquête face-à-face et les résultats obtenus sont très similaires à ceux obtenus sur les données test web + ENL 2006. La seule différence est que certaines variables, explicatives dans le modèle web + ENL, n'apparaissent plus aussi significatives dans le modèle web + face-à-face. Ceci nous semble dû au faible nombre d'observations faites en face-à-face.

Par conséquent, nous avons conclu que dans le modèle logit web + ENL, c'est le mode de collecte, et non pas l'année d'enquête (2006 ou 2010), qui est décisif pour expliquer les importantes différences de résultats constatées pour les questions d'opinion.

* *
*

Voici les variables à expliquer dans le modèle test web + ENL 2006 :

Variable	Libellé	Modalités	Nombre de modalités
HPH	Nombre de pièces	ordre croissant	3
HSHT	Surface du logement	ordre croissant	7
OLA	Satisfaction des conditions de logement	de très insuffisantes à très satisfaisantes	5
OQA	Quartier plaisant	non / oui	2
KBJ	Bruit pendant la journée,	fréquent à rare	3
KBN	Bruit pendant la nuit	fréquent à rare	3
KEXPO	Exposition du logement	mauvaise à bonne	3
KVOIS	Relations avec le voisinage	mauvaises à bonnes	3
KACCES	Accès au logement	moyen à bon	3
KAIR	Qualité de l'air du quartier	mauvaise à bonne	3
KSECU	Sécurité du quartier	mauvaise à bonne	3
KVOITU	Accessibilité en voiture	mauvaise à bonne	3
KRUE	Entretien des rues et espaces publics	mauvais à bon	3
KCOM	Proximité des commerces	aucune à bonne	4
KTRANS	Accessibilité des transports en commun	aucune à bonne	4
KVERT	Présence et qualité des espaces verts	aucune à bonne	4
ODL	Souhait de changer de logement	oui / non	2
ODF	Contrainte de changer de logement	oui / non	2

Robustesse du modèle :

Pour mesurer le degré de robustesse du modèle, nous avons testé deux variantes établies à partir de celui-ci mais en prenant des tranches d'âge et d'unité urbaine plus fines. Les résultats sont très proches, comme on peut le constater dans le tableau ci-dessous.

Log 2006 vs. Test web	Modèle de base		Modèle âge fin		Modèle âge fin + tranche d'unité urbaine fine	
	Coeff	Khi ²	Coeff	Khi ²	Coeff	Khi ²
HPH	-0,29	158	-0,29	152	-0,29	147
OLA		1		1		1
KBJ	0,83	493	0,73	493	0,73	498
KVOITU	0,93	773	0,93	774	0,95	795
ODL	0,09	8	0,10	9	0,10	10
ODF	0,28	110	0,39	114	0,39	115

Nous avons aussi testé le modèle en séparant les données selon la taille de l'unité urbaine où se trouve le logement. Pour cela nous avons séparé l'habitat rural (communes de moins de 2 000 habitants, les « petites » villes (de 2 000 à 50 000 habitants), les grandes villes (de 50 000 à 2 millions d'habitants) et l'agglomération parisienne. Les résultats sont proches d'un cas à l'autre.

Log 2006 vs. Test web	Modèle rural		Modèle petit ville (<50M)		Modèle grande ville		Modèle Paris + aglo	
	Coeff	Khi ²	Coeff	Khi ²	Coeff	Khi ²	Coeff	Khi ²
HPH	-0,43	70	-0,35	33	-0,31	80		0
OLA		2	0,20	10		1		1
KBJ	0,70	86	0,70	60	0,78	262	0,77	101
KVOITU	1,00	167	1,03	137	1,04	408	0,78	94
ODL		18		3		0	0,23	11
ODF	0,21	5	0,46	25	0,32	38	0,54	54

Enfin, nous avons testé un dernier modèle en remplaçant les revenus déclarés par les revenus bruts avant abattement issus du fichier de la taxe d'habitation.

Ces résultats nous conduisent à considérer le modèle comme suffisamment robuste.

Résultats du modèle LOGIT :

Le résultat important à tirer de ces tableaux est que la source est significative pour presque toutes les questions (sauf celle sur la satisfaction à propos des conditions de logement). Si l'on fait l'hypothèse que malgré la diachronie de l'enquête Logement 2006 et du test effectué en 2010, les résultats sont comparables (hypothèse étayée par le fait que le modèle logit test web + mini-enquête donne des résultats similaires au modèle test web + ENL 2006), alors on peut conclure à un effet de mode pour toutes les questions d'opinion (de ressenti), le mode de collecte par internet donnant alors des résultats plus « négatifs ».

Le mode de collecte est même la variable la plus explicative pour les questions sur :

- le bruit (dans la journée et la nuit)
- l'exposition du logement
- les relations avec le voisinage
- l'accès au logement
- la sécurité du quartier
- l'entretien des rues
- la proximité des commerces.

Conclusion

Le protocole d'expérimentation avait certes ses limites, notamment dans la mesure où l'enquête de référence, l'Enquête Nationale Logement 2006, avait quatre d'ans d'antériorité par rapport au test sur le web, de surcroît l'enquête de 2006 était une enquête d'intérêt général à caractère obligatoire.

Cependant, la très forte non-réponse rencontrée dans le test et les grandes différences de structure socio-démographiques des répondants sont particulièrement flagrantes et ne peuvent être expliquées par les différences entre les deux opérations mentionnées ci-dessus.

Un calage très poussé a été nécessaire pour rendre les données du test web représentatives de la population française, et ainsi pour rendre comparables les résultats aux questions d'opinion du test web avec ceux de l'ENL 2006. En plus du calage, l'apurement des revenus a été indispensable pour

les étudiant, et finalement ceux-ci se sont révélés de qualité similaire à celle obtenue dans les enquêtes ménages en face-à-face.

Une fois les données rendues comparables, on a constaté de très grandes différences entre elles : les données issues du web ont révélé une insatisfaction beaucoup plus grande relative au logement. Ces constatations ont été corroborées par un modèle logit montrant que le mode de collecte explique très significativement les résultats aux questions d'opinion sur le logement²⁵. Pour plusieurs questions, il s'agit même de l'effet le plus significatif.

Au vu de ces résultats nous pouvons conclure à un effet de mode dans l'enquête sur le logement, pour ce qui concerne les questions sur le ressenti des personnes. Deux explications possibles, qui peuvent se combiner, nous sont apparues.

D'un côté, le fait de répondre seul, face à son écran d'ordinateur, sans la présence d'un enquêteur, peut neutraliser l'effet de « désirabilité sociale » et rendre les réponses plus sincères et plus tranchées.

De l'autre, au vu du taux de réponse, il est possible que ce soit surtout les personnes ayant à se plaindre de leurs conditions de logement et espérant peut-être que répondre à un questionnaire sur ce thème permettra une évolution positive de ces conditions, qui aient fait la démarche de se connecter au site et de répondre à l'enquête.

La question demeure, de savoir si cet effet de mode qui nous est apparu lors de cette enquête sur le logement apparaîtrait aussi lors d'enquêtes portant sur d'autres thèmes. Un premier élément de réponse peut nous être fourni par l'enquête Qualité de la Vie. Si les résultats de celle-ci ne peuvent être comparées avec des données issues d'une collecte en face-à-face, on peut cependant constater que ce ne sont pas seulement ni majoritairement les personnes ayant à se plaindre de leur vie ou les personnes pessimistes qui y ont répondu.

²⁵ Au delà des questions d'opinion, la comparaison à l'ENL 2006 sur certaines questions factuelles (« qui est le propriétaire de ce logement ? », ou encore « pour quelle raison pensez-vous être contraint à quitter votre logement ? ») montre qu'il est plus difficile pour l'enquêté de répondre de façon précise en mode auto-administré dès qu'il doit se référer à des documents, ici le bail, ou qu'il doit se classer face à des modalités complexes.

Bibliographie

- [1] Ardilly P., « Les techniques de sondage », Éditions Technip, 2006
- [2] Dillman D.A., Smyth J.D., Christian L.M., « Internet, mail and mixed-mode surveys, the tailored design method”, third edition, Wiley, 2009
- [3] De Leeuw E.D., Hox J.J., Dillman D.A., « International handbook of survey methodology”, European association of survey methodology, 2008
- [4] Jäckle A., Roberts C., Lynn P., “Assessing effects of data collection mode on measurement”, ISER, 2008

Annexe 1 : Quelques images de l'application

Voici quelques captures d'écran de l'application d'enquête en ligne à laquelle les enquêtés ont pu se connecter.

Page d'accueil:

 <p>INSEE INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES</p>	<h3>Le logement et la mobilité résidentielle</h3> <h4>Enquête 2010</h4>	
<p>Pourquoi cette enquête</p> <hr/> <p>Votre contact à l'Insee</p> <hr/> <p>Consignes de remplissage</p>	<p>Bonjour, nous vous avons envoyé un courrier annonçant la réalisation d'une enquête sur le logement et la mobilité résidentielle.</p> <p>Ce questionnaire est à remplir par une seule personne de votre ménage : vous-même ou toute personne de plus de 18 ans habitant le logement.</p> <p>Pour accéder au questionnaire de l'enquête, vous aurez besoin de l'identifiant et du mot de passe indiqués sur la lettre que vous avez reçue (en haut et à droite).</p> <p>Nous vous remercions de bien vouloir y consacrer un peu de temps.</p>	<p>La loi n° 78-17 du 6 janvier 1978 modifiée, relative à l'informatique, aux fichiers et aux libertés, s'applique aux réponses faites à la présente enquête. Elle garantit aux personnes concernées un droit d'accès et de rectification pour les données les concernant. Ce droit peut être exercé auprès de l'Insee. Les réponses à ce questionnaire sont protégées par le secret statistique et destinées à l'Institut national de la statistique et des études économiques (Insee).</p>
<p>Démarrer l'enquête</p> <p>Pour toute difficulté d'accès à l'application résultant d'un handicap visuel ou autre, cliquez sur "Votre contact à l'Insee"</p>		

Début du questionnaire :

 <hr/> Sommaire: 1- Description de votre logement 2- Opinion sur votre logement 3- Vos projets de mobilité 4- Les mobilités précédentes 5- Votre situation personnelle © INSEE 2010	LOGEMENT 2010 Description de votre logement
	<p>Tout d'abord, voilà quelques questions pour décrire vos conditions actuelles de logement.</p> <p>Votre logement principal est-il ?</p> <ul style="list-style-type: none"><input type="radio"/> une maison individuelle<input type="radio"/> un appartement dans un immeuble collectif<input type="radio"/> autre <p>Occupez-vous ce logement comme ?</p> <ul style="list-style-type: none"><input type="radio"/> accédant à la propriété (vous remboursez des emprunts pour l'achat de ce logement)<input type="radio"/> propriétaire non accédant (vous ne remboursez pas ou plus d'emprunt pour l'achat de ce logement)<input type="radio"/> locataire<input type="radio"/> autre

Questions sur les prévisions de mobilité :

 <hr/> Sommaire: 1- Description de votre logement 2- Opinion sur votre logement 3- Vos projets de mobilité 4- Les mobilités précédentes 5- Votre situation personnelle © INSEE 2010	LOGEMENT 2010 Vos projets de mobilités
	<p>Nous allons vous interroger maintenant sur vos intentions éventuelles de déménager.</p> <p>Souhaitez-vous changer de logement ?</p> <ul style="list-style-type: none"><input checked="" type="radio"/> oui<input type="radio"/> non <p>Pensez-vous être contraint de quitter votre logement dans les trois ans à venir ?</p> <ul style="list-style-type: none"><input checked="" type="radio"/> oui<input type="radio"/> non <p>Pour quelle raison pensez-vous y être contraint ?</p> <ul style="list-style-type: none"><input type="radio"/> raisons professionnelles : mutation, études<input type="radio"/> votre propriétaire veut reprendre le logement<input checked="" type="radio"/> votre logement vous est prêté ou loué provisoirement<input type="radio"/> raisons financières<input type="radio"/> raisons familiales : modification de la taille de la famille<input type="radio"/> raisons de santé<input type="radio"/> retraite, fin d'activité<input type="radio"/> autre raison

Question sur le revenu mensuel, proposition de tranches :

 <hr/> Sommaire: 1- Description de votre logement 2- Opinion sur votre logement 3- Vos projets de mobilité 4- Les mobilités précédentes 5- Votre situation personnelle	LOGEMENT 2010	Quelques informations sur votre situation personnelle
	<p>Pouvez-vous néanmoins situer le <u>revenu mensuel moyen</u></p> <p> Vous n'avez pas répondu à cette question. Si c'est un oubli cochez ou renseignez cette zone, si vous ne souhaitez pas répondre cochez la case "sans réponse"</p> <ul style="list-style-type: none"><input type="radio"/> moins de 400 euros<input type="radio"/> de 400 à moins de 600 euros<input type="radio"/> de 600 à moins de 800 euros<input type="radio"/> de 800 à moins de 1000 euros<input type="radio"/> de 1 000 à moins de 1 200 euros<input type="radio"/> de 1 200 à moins de 1 500 euros<input type="radio"/> de 1 500 à moins de 1 800 euros<input type="radio"/> de 1 800 à moins de 2 000 euros<input type="radio"/> de 2 000 à moins de 2 500 euros<input type="radio"/> de 2 500 à moins de 3 000 euros<input type="radio"/> de 3 000 à moins de 4 000 euros<input type="radio"/> de 4 000 à moins de 6 000 euros<input type="radio"/> de 6 000 à moins de 10 000 euros<input type="radio"/> 10 000 euros ou plus<input type="radio"/> sans réponse	
<small>© INSEE 2010</small>		

Possibilité de laisser un commentaire en fin d'enquête :

 <hr/> Sommaire: 1- Description de votre logement 2- Opinion sur votre logement 3- Vos projets de mobilité 4- Les mobilités précédentes 5- Votre situation personnelle	LOGEMENT 2010	Quelques informations sur votre situation personnelle
	<p>A la fin de cette enquête, avez-vous un commentaire ? Vous pouvez par exemple indiquer les questions qui vous ont paru difficiles et pourquoi. Vous pouvez également nous informer des difficultés techniques que vous avez pu rencontrer.</p> <p>Je répondrai volontiers à d'autres enquêtes !</p>	
<small>© INSEE 2010</small>		

Remerciements :



INSEE

LOGEMENT 2010

Remerciements

Votre questionnaire est maintenant complet, nous vous remercions pour le temps que vous avez passé à répondre.

Pour valider et envoyer vos réponses à l'Insee, veuillez cliquer sur Valider, vous quitterez alors le site de l'enquête et vous pourrez accéder à des informations statistiques sur le logement.

Sommaire:

- 1- Description de votre logement
- 2- Opinion sur votre logement
- 3- Vos projets de mobilité
- 4- Les mobilités précédentes
- 5- Votre situation personnelle

© INSEE 2010

Annexe 2 : Ce que les enquêtés ont pensé de l'enquête par internet

Les dernières questions permettaient de connaître l'avis des enquêtés sur la présente enquête en ligne.

Plus de 60 % des répondants ont trouvé que le questionnaire avait une bonne longueur. En détail, le questionnaire a été trouvé :

	Logement Web (court)	Logement Web (long)
Trop long	4 %	8 %
Trop court	4 %	2 %
Comme il faut	68 %	63 %
Non-réponse	24 %	27 %

Comme on pouvait s'y attendre (objectivement parlant), il y a eu davantage de répondants qui ont trouvé le questionnaire « trop long » parmi ceux qui ont dû répondre au questionnaire long que parmi ceux qui ont dû répondre au questionnaire court. Et inversement, les premiers ont été moins nombreux que les seconds à trouver le questionnaire « trop court ». [La question n'a pas été posée en face-à-face.]

Il y a eu autant de commentaires de fin d'enquête après les questionnaires longs et après les questionnaires courts, 450 au total.

45 % de ces commentaires signifiaient que l'enquêté n'avait eu aucun problème pour répondre et n'avait rien à ajouter, 16,7 % étaient des conseils pour améliorer le questionnaire et 13,1 % des précisions apportées aux réponses déjà données.

9 % des commentaires sont des critiques du mode de collecte par internet : soit des indications à propos de problèmes techniques de connexion ou de déconnexion, soit des recommandations pour rendre l'accès au questionnaire plus facile, soit un rejet des enquêtes en ligne pour raisons ergonomiques (certains ne disposant pas de ce média à domicile et devant remplir le questionnaire chez quelqu'un d'autre).

Enfin, quelques enquêtés pensent qu'il n'est pas très sûr de révéler le montant de ses revenus sur internet, et d'autres trouvent l'ensemble du questionnaire indiscret, sans dire si cette remarque est liée au mode de collecte.

Annexe 3 : Résultats du modèle logit

Le tableau suivant contient les résultats des régressions logistiques avec, pour chaque variable à expliquer et chaque variable explicative, la statistique du Khi^2 de Wald (qui est significative lorsqu'elle est supérieure à 4) et le coefficient correspondant dans les cas significatifs. On a retenu ici les résultats non pondérés (la pondération ne pratiquement change rien).

Par exemple, la source est significative pour le nombre de pièces car le Khi^2 vaut 158. Le coefficient est alors de -0,29. On peut en conclure que le nombre de pièces était plus petit dans l'enquête Logement Web que dans le test.

Pour faciliter l'interprétation, les variables à expliquer sont classés par ordre de valeur la plus négative à la plus positive selon ce qui nous a paru être l'opinion courante. Par exemple :

- La surface du logement, de petite à grande.
- Le bruit pendant la journée, de fréquent à rare.
- Les relations avec le voisinage, de mauvaises à bonnes.
- Le souhait de changer de logement, de oui à non.

	HPH		HSHT		OLA		OQA		KBJ		KBN	
	Coeff	Khi ²										
s1	ref	ref										
s2	-0,29	158	-0,06	7		1	0,15	12	0,73	493	0,94	749
age1	-0,38	234	-0,39	253		2	-0,08	4		1		0
age2	ref	ref										
age3	0,46	1421	0,44	1288	0,15	155	0,11	24	0,08	33	0,10	33
age4	0,64	950	0,60	832	0,32	225	0,44	126	0,25	102	0,30	104
sex1	ref	ref										
sex2	0,08	55	0,04	13		0	-0,07	13	-0,05	18	-0,09	35
r1	-0,36	499	-0,36	524	-0,37	536	-0,18	43	-0,12	46	-0,17	64
r2	-0,17	133	-0,16	120	-0,22	228	-0,11	17	-0,06	12	-0,08	18
r3	ref	ref										
r4	0,23	244	0,26	307	0,25	282	0,13	20	0,15	70	0,19	72
r5	0,61	1433	0,72	1982	0,55	1098	0,30	81	0,29	218	0,35	196
d1	-0,19	108	-0,23	162	-0,30	262	-0,23	54	-0,10	25	-0,19	61
d2	-0,08	17	-0,12	33	-0,16	57		1		0		1
d3		2	-0,05	4	-0,14	37	-0,10	6		1	-0,08	6
d4		0		2	-0,07	20	-0,12	15	-0,05	8	-0,05	6
d5	ref	ref										
d6	0,09	31	0,09	35	0,09	29		3		0		3
nh1	-0,54	1585	-0,45	1175	0,17	152	0,05	4	0,05	12	0,08	20
nh2	ref	ref	ref	ref	ref	ref	ref					
nh3	0,44	1182	0,35	749	-0,14	109	-0,09	15		0		1
nh4	1,02	2824	0,73	1443	-0,21	119	-0,18	31		0	-0,06	6
sit1	ref	ref										
sit2	-0,62	220	-0,63	250	0,11	8	0,15	5		0		0
sit3	-0,14	49	-0,17	71	-0,25	149	-0,17	33	-0,10	23	-0,10	16
sit4	0,18	122	0,16	90	0,04	5	-0,07	5		1		1
t1	1,01	3716	1,32	5982	0,68	1690	0,66	316	0,59	812	0,80	820
t2	0,86	2521	1,07	3754	0,57	1101	0,44	167	0,36	316	0,50	402
t3	0,56	1978	0,69	2929	0,38	922	0,14	48	0,14	108	0,23	217
t4	ref	ref										

	KEXPO		KVOIS		KACCES		KAIR		KSECU		KVOITU	
	Coeff	Khi ²	Coeff	Khi ²	Coeff	Khi ²	Coeff	Khi ²	Coeff	Khi ²	Coeff	Khi ²
s1	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
s2	1,07	1025	1,07	947	1,54	2073	0,83	623	1,01	922	0,93	773
age1		2	-0,16	34		0	0,08	8		1		3
age2	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
age3	0,10	34	0,16	96	0,13	44		2	-0,05	10	0,08	31
age4	0,29	93	0,32	139	0,19	33	0,13	27	0,10	16	0,07	7
sex1	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
sex2		2	-0,04	12	-0,04	7	-0,06	24	-0,05	15	-0,06	23
r1	-0,16	57	-0,07	13	-0,17	52	-0,06	11	-0,07	15	-0,06	9
r2	-0,10	26	-0,04	4	-0,12	29	-0,05	9	-0,04	6		2
r3	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
r4	0,17	64	0,08	17	0,19	61	0,11	41	0,12	43	0,10	31
r5	0,27	129	0,16	53	0,36	162	0,22	127	0,23	131	0,16	68
d1	-0,09	14	-0,12	28	-0,12	20	-0,11	28	-0,17	60		2
d2		0		0		0		0	-0,06	5	0,06	7
d3		1		2		0	-0,05	4	-0,07	7		0
d4		0		2		0	-0,06	9	-0,10	28		1
d5	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
d6		0		0		2		1	0,07	14		2
nh1		1	-0,05	10	0,05	5	0,07	21	0,06	15		0
nh2			ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
nh3		4	0,05	11	-0,06	10		0		4	0,06	17
nh4		0		0	-0,10	13		2	-0,11	24	0,05	5
sit1	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
sit2	-0,11	6		3		1		0	0,15	10		0
sit3	-0,11	20	-0,09	14	-0,13	25	-0,14	44	-0,10	20	-0,07	9
sit4		0		0		1		0	-0,04	5	0,05	6
t1	0,28	154	0,29	177	0,44	258	1,01	1975	0,87	1393	0,99	1847
t2	0,22	86	0,20	83	0,29	118	0,74	1206	0,65	850	0,64	965
t3	0,17	108	0,09	33	0,14	62	0,31	526	0,13	92	0,35	670
t4	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref

	KRUE		KCOM		KTRANS		KVERT		ODL		ODF	
	Coeff	Khi²	Coeff	Khi²	Coeff	Khi²	Coeff	Khi²	Coeff	Khi²	Coeff	Khi²
s1	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
s2	0,95	859	0,76	550	0,73	509	0,75	516	0,09	8	0,38	110
age1	0,18	34	0,11	11	0,07	6	-0,09	11		3	-0,23	52
age2	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
age3	-0,05	15	-0,13	78	-0,04	9	0,08	36	0,40	644	0,31	201
age4	-0,06	6	-0,25	111	-0,15	39	0,13	30	0,92	897	0,62	210
sex1	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
sex2		1	0,03	5		3		0	-0,03	6		3
r1		0		1		2	-0,11	38	-0,25	143	-0,20	61
r2		0		2		0	-0,05	9	-0,13	45	-0,12	22
r3	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
r4		0	0,04	4		0	0,09	29	0,22	124	0,14	24
r5		3	-0,04	4		3	0,21	123	0,37	286	0,24	60
d1		1	0,04	4		3	-0,15	60	-0,08	11	0,06	4
d2		1		0		1	-0,06	6	0,08	7	0,14	11
d3		1		3		4		1		0	0,09	5
d4		0		1		3		1	0,07	10	0,10	13
d5	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
d6		0		2		0	0,05	6		1	-0,07	7
nh1	0,08	31		2		0		0	0,19	111		1
nh2	ref	ref	ref	ref								
nh3		0		2		2		0	-0,07	19		0
nh4		1	-0,06	8		0	-0,08	13		2		2
sit1	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
sit2		1		3		0		1	0,32	48	-0,44	94
sit3		2	0,05	4	0,06	6	-0,08	13	-0,28	131	-0,24	75
sit4		2	-0,09	20		1		3	0,05	4		1
t1	0,06	12	-0,66	1427	-1,26	4563	0,80	1301	0,80	1162	0,27	72
t2	0,04	5		0	-0,94	2426	0,33	287	0,53	571	0,12	15
t3	0,03	4	0,07	23	-0,12	61	-0,05	17	0,31	413		2
t4	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref

Annexe 4 : Comparaison des données collectées avec les données de la taxe d'habitation

La surface diffère grandement (de plus de 10 %), selon que la donnée provient du fichier de taxe d'habitation (TH) ou du questionnaire web. En revanche, l'adéquation est beaucoup plus forte pour le statut d'occupation et de la nature du logement.

Surface	50,0 % d'adéquation (marge d'erreur de 10 %)
	70,2 % (marge d'erreur de 20 %)
	88,9 % (marge d'erreur de 50 %)
Statut d'occupation	93,3 % d'adéquation
Nature du logement	93,4 % d'adéquation

Des tests d'indépendance du khi-deux portant sur le statut d'occupation et la nature du logement indiquent que les données collectées et les données provenant de la taxe d'habitation ne sont pas indépendantes les unes des autres (p-value inférieure à 0,05).

Un test de corrélation indique aussi que la surface du logement collectée et la surface du logement issue de la TH sont corrélées (le coefficient est de 0,64) ; mais un test de Student de la différence entre la valeur collectée et la valeur issue de la TH rejette l'égalité des moyennes.

La surface déclarée dans le questionnaire web est supérieure à la surface figurant dans le fichier TH dans 70 % des cas ; et elle est inférieure dans 30 % des cas. On pourrait nommer le premier phénomène « biais de prétention » et le second « biais d'humilité », et un redressement pourrait être envisagé si l'on disposait des profils psychologiques !

Pour les trois variables étudiées (surface, statut d'occupation et nature du logement), le mode de collecte (questionnaires web vs. mini-enquête papier) ne semble pas avoir d'influence sur l'adéquation ou l'inadéquation entre les réponses collectées et les données issues de la taxe d'habitation. C'est ce qu'indiquent un test d'égalité des moyennes d'une part, des tests du khi-deux d'autre part. Cependant, pour la surface, le test d'égalité des variances indique que celles-ci ne sont pas égales ; en effet, on remarque des divergences notables des variances de la variable « différence entre la surface déclarée et la surface issue de la TH » selon le mode de collecte (en distinguant enquête Web courte, enquête Web longue et mini-enquête en face-à-face) alors que les moyennes sont assez proches les unes des autres.