

L'ENQUETE AUPRES DES NON-REONDANTS A L'ENQUETE EMPLOI

Sébastien DURIER (), Corinne PROST (*)*

() Insee, Division Emploi*

L'enquête non-répondants (ENR) est une enquête postale auprès des non-répondants à l'enquête Emploi, qui reprend de façon succincte le module servant à déterminer le statut d'activité au sens du bureau international du travail. Cette enquête étend la période de collecte de l'enquête Emploi : 2 semaines et 2 jours pour l'enquête principale, étendue à 5 semaines avec le questionnaire papier de l'ENR.

L'enquête Emploi étant une enquête en panel, se pose la question de l'impact de l'ENR sur les comportements de réponses des ménages au fur et à mesure des interrogations. Le protocole de mise en place de l'ENR ne permet pas une évaluation directe de son effet sur le comportement des ménages car l'ENR a été appliqué directement à l'ensemble du territoire. A la lecture des simples taux de réponse, il ne semble pas y avoir d'effet « d'éviction » qui conduirait les ménages ayant répondu à l'ENR à ne plus vouloir répondre à l'enquêteur lors des interrogations suivantes. Par ailleurs, l'ENR aurait eu un effet positif sur l'organisation de la collecte (obligation pour les équipes régionales de mieux respecter le calendrier de collecte, obligation pour les enquêteurs de retourner sur le terrain pour récupérer les noms des non-répondants).

Les résultats à l'ENR montrent que les actifs occupés sont sur-représentés parmi les répondants de l'ENR, y compris à tranche d'âge donnée (66% d'actifs occupés contre 55% dans l'enquête Emploi en 2007). Toutefois, le taux de réponse à l'ENR est faible (autour de 25%) et les non-répondants finaux représentent encore 10% de l'échantillon total au T1 2007. Faut-il intégrer les résultats de l'ENR à ceux de l'enquête principale, sachant qu'on n'observe qu'un quart des non-répondants ? Deux difficultés se posent alors. Tout d'abord, les réponses ne sont peut-être pas complètement comparables car il peut y avoir un effet de mode de collecte : papier pour l'ENR contre face-à-face et téléphone pour l'enquête Emploi. Malheureusement, le protocole de l'ENR ne permet pas d'évaluer un possible effet de mode. Par ailleurs, l'intégration des résultats d'une enquête non-répondants ne conduit pas forcément à réduire les biais liés à la non-réponse. Une façon empirique d'évaluer la pertinence de l'intégration d'une telle enquête repose sur une mesure de la représentativité des réponses. A partir des variables disponibles, on peut montrer que les répondants à l'ENR permettent bien d'améliorer la représentativité de l'ensemble des répondants.

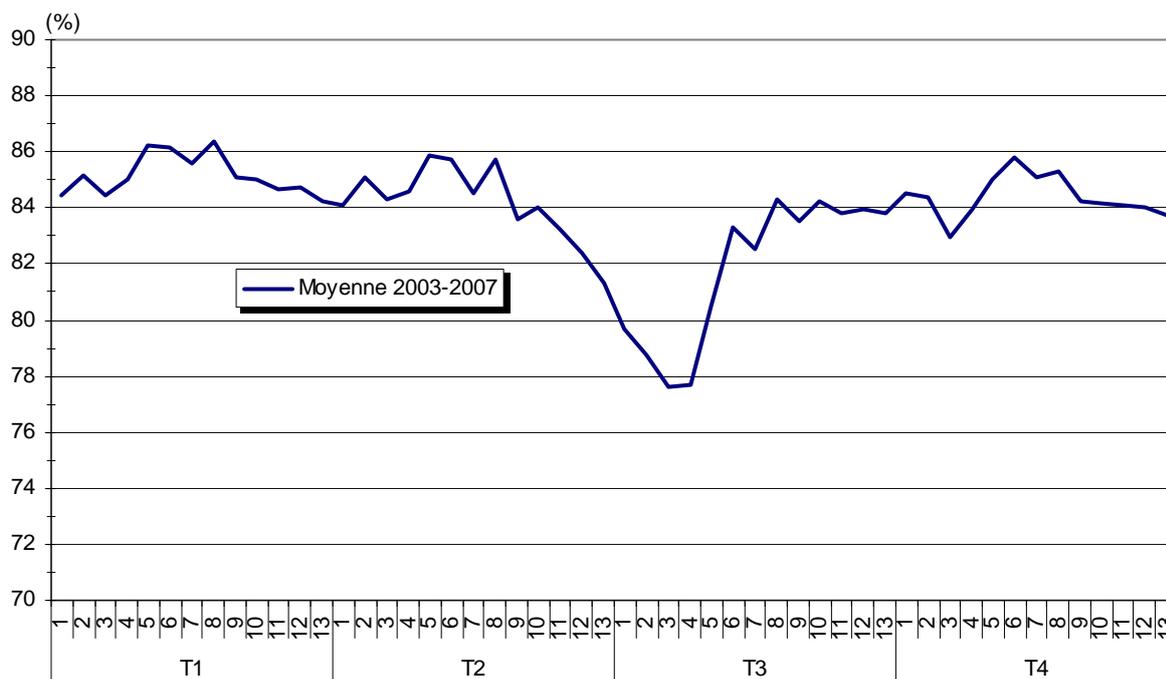
Les résultats de l'ENR sont maintenant intégrés à l'enquête Emploi, à partir de marges d'activité calculées sur les répondants à l'enquête Emploi et à l'ENR. Cela conduit à diminuer le taux de chômage global d'environ 0,1 point.

1. Un dispositif méthodologique

1.1. Les taux de réponse à l'enquête Emploi

Le taux de réponse de l'enquête Emploi se situe en moyenne autour de 84%. Il est très saisonnier, avec un creux d'environ 6 points aux troisièmes trimestres, du fait des vacances estivales (cf. graphique 1).

Graphique 1 : Taux de réponse par semaine de référence

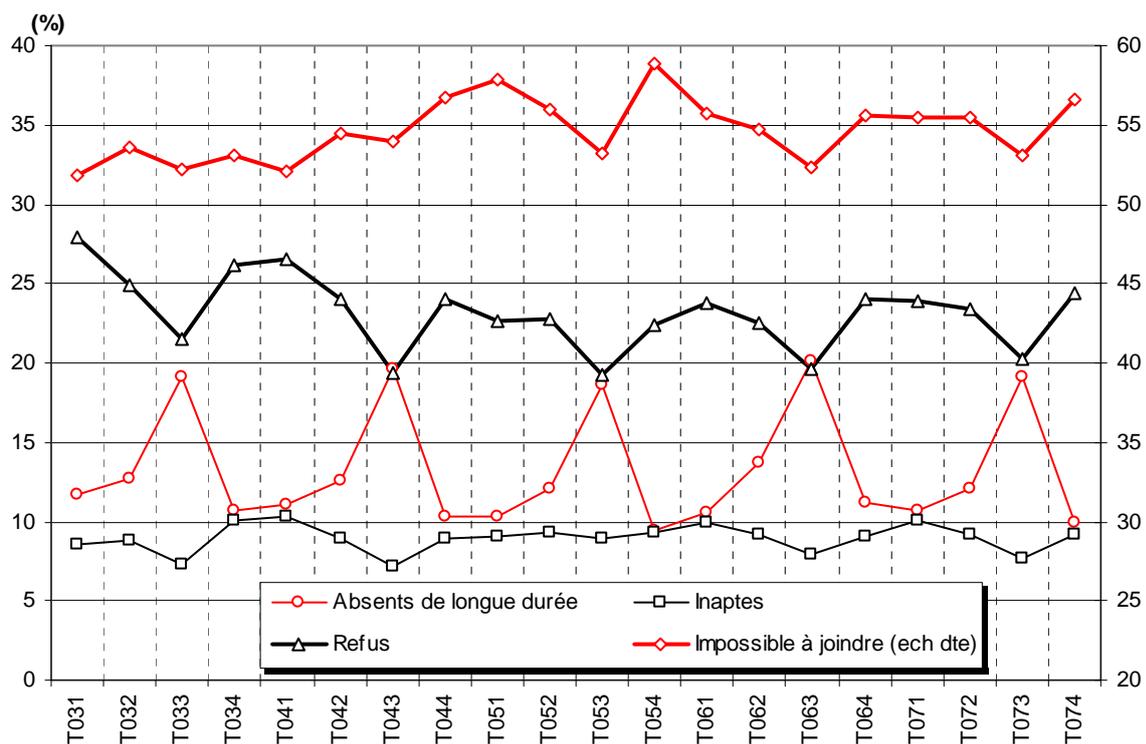


Lecture : pour chaque semaine de l'année, environ 3400 résidences principales sont interrogées. Au mois de juillet, soit les 4 premières semaines du 3^{ème} trimestre, moins de 80 % des logements répondent.

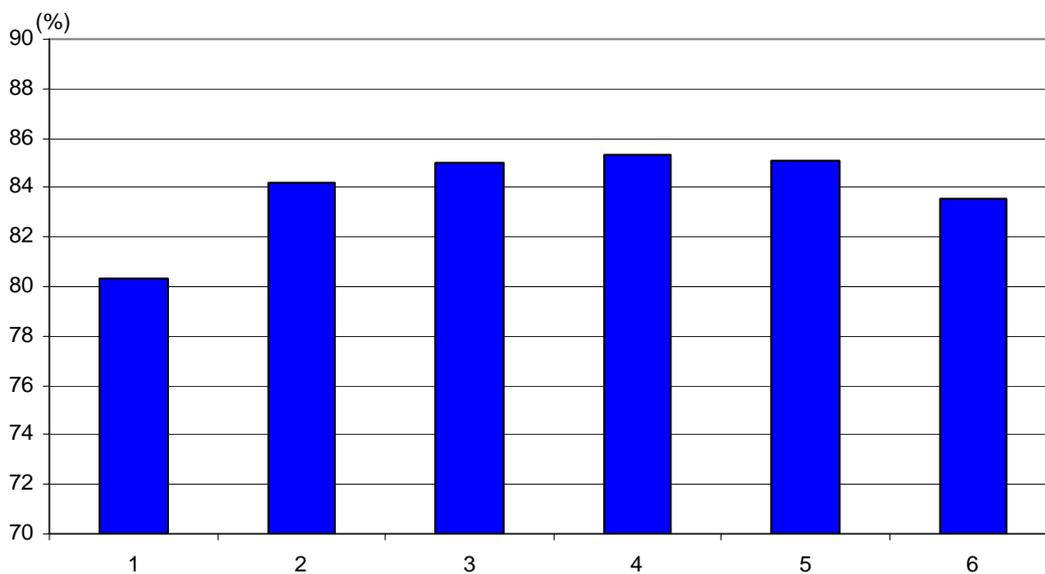
Les refus représentent moins d'un quart des raisons de la non-réponse. Celle-ci semble le plus souvent due à la difficulté de contacter les ménages. Les deux motifs correspondants recouvrent une majorité de non-réponses : impossibilité à joindre (autour de 55%) et absence de longue durée (autour de 15%). La répartition des motifs de non-réponse apparaît très stable dans le temps (cf. graphique 2).

Le taux de réponse varie également en fonction du rang d'interrogation (cf. graphique 3). Il est le plus faible pour les aires entrantes où seulement 80 % des logements sont interrogés, puis progresse pour les interrogations suivantes avant de baisser à nouveau en dernière interrogation. On a donc un phénomène de récupération des non-répondants au fil des interrogations contrebalancée par une lassitude progressive de l'enquête en panel. Il faut noter toutefois que le champ diffère selon le rang d'interrogation : les ménages d'inactifs de + de 65 ans ne sont pas interrogés aux vagues 2 à 5 ; le statut des résidences secondaires n'est pas réévalué à ces vagues intermédiaires ; les logements neufs sont intégrés seulement en vague 6. Toutes ces modifications de champ rendent plus difficile la comparaison des taux de réponse d'une vague à l'autre.

Graphique 2 : Répartition de la non-réponse selon le motif donné par l'enquêteur



Graphique 3 : Taux de réponse selon le rang d'interrogation (moyenne 2003-2007)



L'échantillon actuel de l'enquête Emploi contient assez peu de variables permettant de caractériser les non-répondants. Les seules variables disponibles dans les données sont des variables du recensement de 1999, qui fournissent des caractéristiques des logements et de leur localisation : nombre de pièce, type de logement, tranche d'unité urbaine, région, ZUS, année de construction, typologie Tabard. La probabilité de répondre semble tout d'abord étroitement corrélée au nombre de pièces, au type de logement (maison individuelle contre immeuble collectif) et à son année de construction. A cela s'ajoute un fort effet géographique, à la fois lié au degré d'urbanisation et à la composition socio-économique du quartier (cf. annexe 1 pour les résultats de la régression logistique sur les variables explicatives de la réponse).

1.2. Mise en place d'une enquête auprès des non-répondants

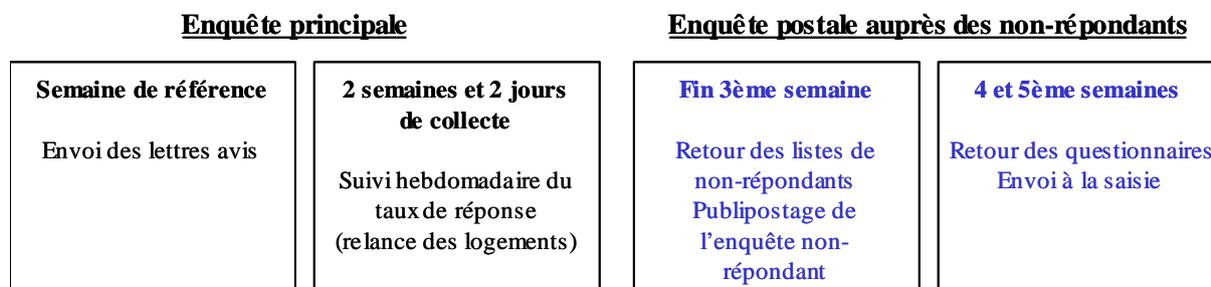
L'enquête auprès des non-répondants (ENR) a été mise en place en tant qu'enquête méthodologique, visant à mieux connaître la non-réponse. C'est une enquête postale auprès des non-répondants à l'enquête Emploi, qui reprend de façon succincte le module servant à déterminer le statut d'activité au sens du bureau international du travail. Elle permet donc de reconstituer une variable d'activité.

L'un des objectifs de l'enquête méthodologique est de demander directement au ménage pourquoi il n'a pas répondu à l'enquêteur ; et vérifier ainsi que les motifs de non-réponse renseignés par l'enquêteur ne sont pas biaisés par une sous-déclaration des refus ou une sur-déclaration des absences de longue durée.

L'autre objectif est plus directement lié à ce qu'on cherche à mesurer : est-ce que la non-réponse est corrélée au statut d'activité ? Des soupçons de biais de réponse sont nés notamment du constat d'un « biais de rotation » dans l'enquête Emploi. En effet, dans l'enquête, les ménages sont interrogés six trimestres d'affilée, s'ils ne déménagent pas. Or les résultats, et notamment le taux de chômage, présentent des différences systématiques selon la vague d'interrogation considérée. Ce biais de rotation a de nombreuses causes possibles. L'une de ces causes pourrait être une non-réponse spécifique selon la vague d'interrogation et le statut d'activité (par exemple une propension plus forte des chômeurs à refuser de répondre plusieurs fois de suite).

L'enquête a d'abord été mise en place de façon partielle aux T1 et T2 2005 puis au T4 2006. Puis elle a été généralisée à l'ensemble des non-répondants et l'ensemble des régions à partir du T1 2007. A la fin de la troisième semaine après la semaine de référence, c'est-à-dire deux jours après la fin de la collecte de l'enquête Emploi, les informations sur les noms et les adresses des non-répondants sont rassemblées et transmises pour publipostage. Les courriers contenant l'enquête papier sont envoyés ; les ménages ont alors deux semaines pour envoyer leurs réponses.

Schéma 1 : Protocole de collecte de l'enquête Emploi



L'enquête papier est composée d'une feuille logement (sexe, âge des habitants) et de 3 questionnaires individuels pour les individus de 15 ans ou plus (hormis les inactifs de 65 ans et plus). Le questionnaire individuel est restreint aux questions permettant d'établir le statut d'activité au sens du BIT. Il reprend les questions et l'ordre de l'enquête principale à quelques aménagements près, certains filtres ne pouvant être reproduits (cf. questionnaire en annexe 2).

Tous les logements non-répondants sont théoriquement concernés par l'enquête non-répondant. Toutefois, certains noms et adresses s'avèrent inutilisables ou ne sont pas transmis à temps. Par ailleurs, les directions régionales peuvent décider d' « exempter » certains ménages principalement lorsqu'il s'agit de personnes inaptes qui, de façon certaine, ne répondront pas au questionnaire. Finalement, le taux de couverture des non-répondants se situe aux alentours de 86 % (cf. tableau 1).

Tableau 1 : Taux de couverture et taux de réponse de l'enquête non-répondant

En %	Taux de couverture des non-répondants (1)	Taux de retour de l'ENR	Taux de retour pour les aires entrantes	Taux de réponse des non-répondants (2)
T1	85,4	29,9	30,4	25,5
T2	85,4	23,8	27,8	20,3
T3	87,3	24,8	27,2	21,7
T4	87,9	22,8	26,7	20,0
Ens 2007	86,6	25,3	28,0	21,9

(1) Proportion de logements non-répondants auxquels l'ENR est envoyée

(2) Ce taux cumule le taux de couverture et le taux de retour : il s'agit du nombre de réponses à l'ENR rapporté à la population totale des non-répondants.

En moyenne sur l'année 2007, le taux de retour a été de 25 %, soit environ 1400 logements et 2100 individus de 15 ans ou plus par trimestre.

1.3. Qui sont les répondants à l'enquête auprès des non-répondants ?

1.3.1. Les motifs de non-réponse renseignés par les enquêteurs semblent corroborés par l'enquête auprès des non-répondants

L'un des objectifs de l'enquête auprès des non-répondants est d'obtenir des informations sur la raison de la non-réponse, et de comparer ces réponses avec les informations déjà fournies par les enquêteurs. Les résultats de l'ENR ne peuvent pas être entièrement conclusifs puisqu'ils ne couvrent que moins d'un quart des non-répondants. Ils donnent toutefois des éléments qualitatifs sur la qualité des motifs renseignés par les enquêteurs.

Au vu des résultats (cf. tableau 2), il semble que la variable renseignée par l'enquêteur est un bon indicateur des motifs de non-réponse. Ainsi, 2/3 des IAJ (impossible à joindre) et des ALD (absence de longue durée) déclarent l'absence comme motif de non-réponse. De même, les logements classés comme refus par l'enquêteur sont ceux qui déclarent le plus souvent des motifs de refus. Enfin, les « inaptés » déclarent massivement d'autres raisons que celles proposées, qui manifestement ne correspondent pas à leur situation. Au-delà de ces concordances entre motifs déclarés par les ménages et ceux inscrits par les enquêteurs, l'ENR décèle également quelques incohérences. Ainsi, 20 % des IAJ déclarent un manque de temps, motif qui est très souvent cité par le « refus ». Autrement dit, derrière certains IAJ se cachent probablement des refus.

Tableau 2 : Motif de non-réponse selon l'enquêteur et le logement

Motifs de non-réponse fournis par le logement (plusieurs réponses possibles)	Ensemble	Impossible à joindre	Absence de longue durée	Inaptes	Refus
Vous étiez absent	53	63	69	25	26
Vous n'aviez pas le temps	20	20	9	9	28
Vous êtes contre les enquêtes	12	8	4	10	28
Vous ne voulez pas qu'on entre chez vous	14	11	5	14	26
Vous ne voulez pas être interrogé par téléphone	15	11	9	19	28
Vous ne vous sentez pas concerné par le sujet	12	8	8	15	25
Vous ne voulez pas être interrogé plusieurs fois	13	10	7	15	24
L'enquête est trop longue	7	6	4	6	12
Autres raisons	32	29	29	56	35

Champ : logements répondant à l'ENR, année 2007

Lecture : 53 % des logements déclarent l'absence du domicile comme motif de non-réponse.

Les motifs de non-réponse du logement sont corrélés à l'âge des habitants du logement (voir tableau 3). Les motifs d'indisponibilité ou d'absence diminuent avec l'âge et à l'inverse les motifs de refus croissent avec l'âge.

Tableau 3 : Motifs de non-réponse des logements selon l'âge

Motifs de non-réponse du logement	Ens	15-29	30-49	50-64	65+
Vous étiez absent	53	66	58	55	43
Vous n'aviez pas le temps	20	25	26	19	8
Vous êtes contre les enquêtes	12	5	9	15	15
Vous ne voulez pas qu'on entre chez vous	14	10	11	13	19
Vous ne voulez pas être interrogé par téléphone	15	7	10	16	24
Vous ne vous sentez pas concerné par le sujet	12	4	5	11	23
Vous ne voulez pas être interrogé plusieurs fois	13	7	11	14	16
L'enquête est trop longue	7	4	7	8	6
Autres raisons	32	25	29	30	38

Champ : logements répondant à l'ENR, année 2007. Age de l'habitant le plus âgé du logement.

1.3.2. Les répondants à l'ENR sont plus souvent des personnes habitant seules et actives.

Les ménages répondant à l'ENR sont plus petits : 45 % sont des personnes seules contre un tiers dans l'enquête principale (cf. tableau 4). La composition par âge des logements montre deux particularités : d'une part il s'agit moins souvent de ménages avec enfant, et de surcroît avec un nombre d'enfant moindre. D'autre part, on constate dans l'ENR une sous-représentation des ménages d'inactifs de 65 ans et plus, ce qui est cohérent avec le protocole particulier qui leur est réservé dans l'enquête Emploi¹. On remarque cependant que la présence de ce type de ménages est plus importante en 6^{ème} interrogation lorsqu'ils réintègrent l'enquête.

¹ Lorsqu'un logement enquêté est composé uniquement d'inactifs âgés de 65 ans ou plus, le logement n'est pas interrogé lors des vagues suivantes à l'exception de la dernière.

Tableau 4 : Types de ménages répondant à l'enquête Emploi et ceux répondant à l'ENR en 2007

Nombre d'habitants du logement	ENR	Enquête Emploi
1	45	32
2	33	34
3	12	14
4	7	13
5+	3	7
Ensemble	100	100
Proportion de logements avec enfant de 14 ans et - dont proportion avec un seul enfant	15 57	24 50
Proportion de ménages d'inactifs de 65 ans et + dont proportion avec un seul habitant	17 65	24 56

Quant aux informations apportées par l'ENR sur le biais lié à la non-réponse, l'ENR semble indiquer que la non-réponse à l'enquête Emploi n'est pas indépendante du statut d'activité puisque les actifs occupés sont sur-représentés parmi les répondants de l'ENR, y compris à tranche d'âge donnée (66% d'actifs occupés contre 55% dans l'enquête Emploi en 2007, cf. tableau 5).

Tableau 5 : Statut d'activité des répondants à l'enquête Emploi et des répondants à l'ENR en 2007

Groupe d'âge		Actifs occupés	Chômeurs	Inactifs	Total
15-24	EE	27	7	66	100
	ENR	46	5	49	100
25-34	EE	77	9	14	100
	ENR	83	6	11	100
35-44	EE	82	6	12	100
	ENR	86	5	9	100
45-54	EE	81	5	14	100
	ENR	83	4	13	100
55-64	EE	41	2	57	100
	ENR	58	2	41	100
65-74	EE	3	0	97	100
	ENR	5	0	95	100
Ensemble	EE	55	5	40	100
	ENR	66	4	30	100

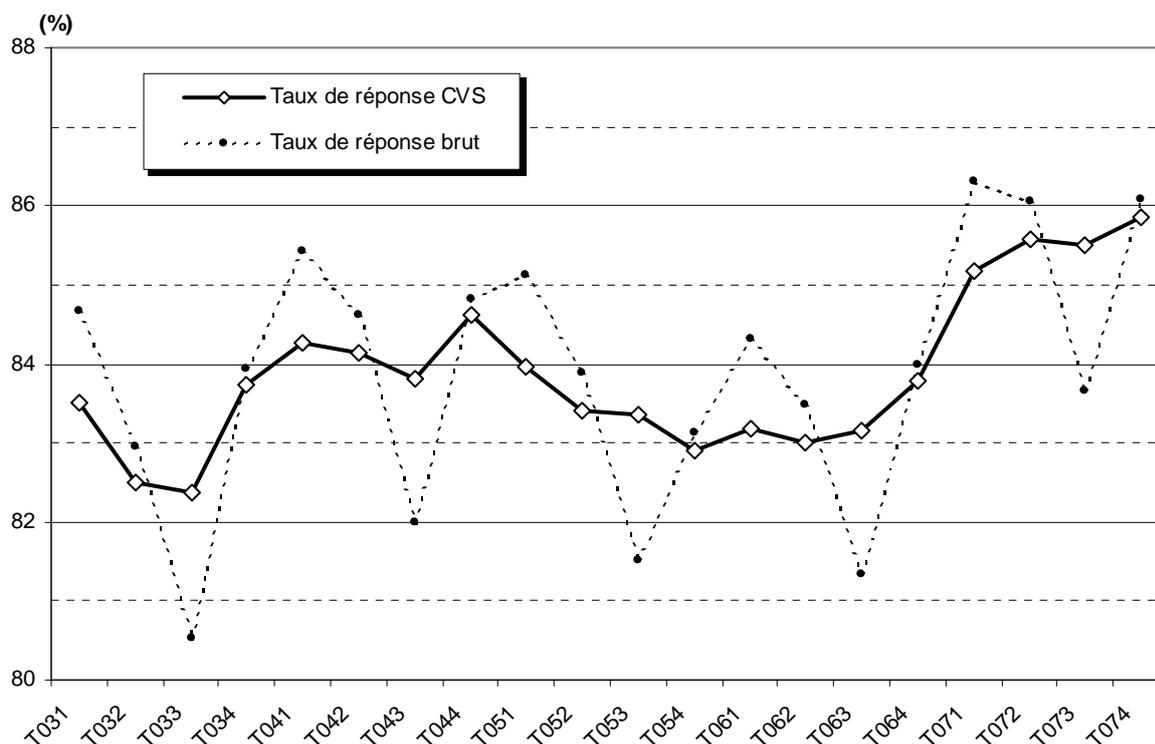
1.4. L'enquête auprès des non-répondants a-t-elle une influence sur la réponse en panel à l'enquête Emploi ?

Lorsque l'enquête principale est une enquête longitudinale, un des risques lors de la mise en place d'une enquête auprès de non-répondants, enquête légère et auto-administrée, est d'influer sur les comportements de réponse lors des interrogations suivantes. Il serait évidemment dommageable pour la qualité de l'enquête Emploi que les ménages répondant à l'ENR décident ensuite de ne pas répondre à l'enquêteur et privilégient l'ENR. Il est toutefois impossible d'évaluer directement l'effet de l'ENR sur les comportements de réponse. En effet, le protocole de l'ENR ne permet pas d'avoir un groupe de contrôle pour une telle évaluation². Plusieurs indicateurs semblent toutefois indiquer que l'ENR n'a pas significativement conduit certains ménages à ne plus répondre à l'enquête principale.

² Les enquêtes méthodologiques des T1 et T2 2005 ont été appliquées sur des champs partiels (certains motifs de non-réponse combinés avec le fait d'avoir été répondant en vague 1). En outre, des problèmes pratiques de

Tout d'abord, l'ENR ne semble pas avoir pesé sur le taux de réponse à l'enquête principale. Même si ce n'est pas conclusif, l'évolution du taux de réponse à l'enquête Emploi peut donner des indices sur la façon dont l'ENR a impacté ce taux. On peut constater qu'après une baisse atteignant presque 2 points en 2005 et 2006, le taux de réponse désaisonnalisé a nettement augmenté ensuite pour rester supérieur à 85% tout au long de 2007, alors que c'est l'année où l'ENR est généralisée (cf. graphique 4). En outre, d'après les questionnaires, l'ENR aurait eu un effet positif sur l'organisation de la collecte (obligation pour les équipes régionales de mieux respecter le calendrier de collecte, obligation pour les enquêteurs de retourner sur le terrain pour récupérer les noms des non-répondants).

Graphique 4 : Taux de réponse trimestriel à l'enquête Emploi



Par ailleurs, parmi les ménages ayant répondu une fois à l'ENR, la proportion d'entre eux qui continuent à répondre à l'ENR le trimestre suivant n'est pas très importante (22%). Le suivi des répondants à l'ENR le trimestre suivant permet en effet d'identifier trois sous-populations (cf. tableau 6). Il y a d'abord les logements qui retournent à l'enquête principale. On pourrait les qualifier de non-répondants temporaires dans le sens où leur non-réponse ne résulte pas d'un refus de l'enquête mais d'une impossibilité passagère. Ils donnent d'ailleurs plus souvent l'absence comme motif de non-réponse et sont nettement moins souvent des personnes âgées. Ils représentent un peu plus de 40 % des non-répondants à l'ENR, mais sont un peu plus nombreux au troisième trimestre en raison des vacances. Viennent ensuite un gros tiers de logements qui ne répondent qu'une fois à l'ENR, et ne répondent ni à l'enquête principale ni à l'ENR au trimestre suivant. Enfin, seuls 22 % des logements répondent à nouveau à l'ENR le trimestre suivant.

Il faut noter que ces deux dernières sous-population sont très proches en terme d'âges et de motifs de non-réponse ; les modalités de motifs de non-réponse ont été déterminées par rapport à l'enquête principale, et ne permettent pas de faire la part entre les personnes qui acceptent de répondre plusieurs fois à l'enquête postale et celles qui s'y refusent. On peut juste constater une légère surreprésentation des personnes déclarant ne pas vouloir être interrogées plusieurs fois et par téléphone parmi les personnes répondant deux fois à l'ENR.

collecte ont affecté l'enquête du T4 2006. De fait, ces enquêtes antérieures à 2007 sont difficiles à exploiter pour obtenir un groupe de contrôle.

Tableau 6 : Motifs de non-réponse et situation le trimestre suivant

Motifs de non-réponse fournis par le logement le trimestre T	Retour à l'enquête en T+1	Maintien dans l'ENR en T+1	Non-réponse à l'ENR en T+1	Ensemble
Vous étiez absent	58	49	50	53
Vous n'aviez pas le temps	16	23	22	20
Vous êtes contre les enquêtes	7	17	16	12
Vous ne voulez pas qu'on entre chez vous	7	21	17	14
Vous ne voulez pas être interrogé par téléphone	10	22	18	15
Vous ne vous sentez pas concerné par le sujet	7	16	15	12
Vous ne voulez pas être interrogé plusieurs fois	7	19	15	13
L'enquête est trop longue	5	9	8	7
Autres raisons	41	21	38	32
Répartition selon la situation en T+1	42	22	36	100

Lecture : 42 % des logements sont retournés à l'enquête en T+1. Parmi ceux-ci, 58 % avaient donné l'absence comme motif.

2. Une enquête intégrée aux résultats de l'enquête Emploi

Hormis le fait que l'enquête est limitée aux questions portant sur le statut d'activité, l'enquête non-répondants pourrait être considérée comme une simple extension de la période de collecte de l'enquête Emploi : les personnes qui n'ont pas pu être contactées pendant la courte période de collecte ont la possibilité de répondre à l'ENR, étendant la période de collecte à 5 semaines. Dans ce cas, intégrer les résultats de l'ENR à ceux de l'enquête Emploi ne semble pas poser de difficultés méthodologiques.

Pourtant, au moins deux arguments peuvent être opposés à l'intégration des résultats de l'ENR. Tout d'abord, les modes de collecte ne sont pas les mêmes ; les résultats sont-ils donc totalement comparables ? En outre, le processus de réponse à l'ENR suit une logique probablement différente à celle de l'enquête Emploi, créant par elle-même des possibilités de biais de réponse ; rajouter les répondants de l'ENR ne garantit pas de corriger le biais possible de réponse de l'enquête Emploi, cela pourrait même aggraver ce biais. Nous verrons que cela ne semble pas être le cas avec l'ENR.

2.1. Les résultats de l'ENR : un effet de mode possible

Le mode de collecte de l'ENR est différent de celui de l'enquête principale : une enquête papier d'un côté ; une enquête en face-à-face ou par téléphone de l'autre côté. Il est possible que les différences observées entre l'enquête principale et l'ENR soient en partie dues au fait que l'ENR est un questionnaire papier auto administré. L'analyse détaillée des questionnaires montre par exemple que les filtres ne sont pas toujours respectés, et que le risque d'erreur de mesure est donc plus important que dans l'enquête principale.

Toutefois, le statut d'actif occupé dépend principalement de la première question (cf. questionnaire en annexe 2) ; le risque d'erreur ou d'effet de mode sur cette mesure est probablement assez limité. Il reste toutefois l'éloignement de la semaine de référence (au maximum 4) qui peut entraîner des biais de mémoire. La distinction entre chômage et inactivité est potentiellement plus difficile à établir avec un questionnaire auto-administré (plus de questions mobilisées et plus de filtres à franchir).

Le protocole de mise en place de l'ENR ne permet pas d'effectuer l'évaluation d'un possible effet de mode. Une étude méthodologique sur la non-réponse à l'enquête Emploi hollandaise a été réalisée en 2005, visant notamment à évaluer des effets de mode (cf. Cobben, 2007). Les logements du groupe de contrôle étaient contactés par téléphone, le groupe de traitement recevait une enquête assez proche de l'ENR. Aucun effet de mode concernant le statut d'activité n'a pu être mis en évidence.

2.2. Faut-il intégrer les résultats d'une enquête non-répondants ?

Les motifs de non-réponse à l'enquête Emploi renseignés le plus fréquemment par les ménages dans l'ENR sont l'absence, le manque de temps puis l'opposition à être interrogé par téléphone. L'ENR permet donc probablement de récupérer certains types de ménages, en étendant la période de collecte et en offrant un mode alternatif de réponse. Grâce à cette enquête, le statut d'activité de 3 % des logements de l'échantillon de l'enquête Emploi peut être mesuré. L'ENR permet donc d'augmenter le taux de réponse de l'enquête Emploi de 3 points, en ce qui concerne la seule mesure de l'activité et de quelques caractéristiques des ménages.

Pour autant, augmenter partiellement le taux de réponse ne permet pas forcément de diminuer les potentiels biais de non-réponse. En effet, les ménages récupérés au moyen de l'ENR peuvent être très spécifiques, différents à la fois des répondants et de ceux qui finalement ne répondent ni à l'enquête Emploi ni à l'ENR, population que nous dénommerons « non-répondants finaux » par la suite. Supposons que les répondants à l'ENR soient beaucoup plus souvent en emploi que les répondants à l'enquête Emploi parce que ce sont des actifs que justement les enquêteurs ont du mal à joindre pendant la courte période de collecte. Supposons que les non-répondants finaux soient au contraire plutôt des inactifs, qui refusent de répondre à l'enquête parce qu'ils ne se sentent pas concernés par le sujet. Comme les non-répondants finaux sont beaucoup plus nombreux que les répondants à l'ENR, supposons que le taux d'activité de l'enquête Emploi soit de fait biaisé vers le haut, c'est-à-dire que l'ensemble des non-répondants soit plus inactifs que les répondants. Dans ce cas, intégrer les répondants à l'ENR biaiserait encore davantage la mesure du taux d'activité.

Il est impossible de mesurer l'évolution du biais de non-réponse lorsqu'on intègre une enquête non-répondants puisque par nature, nous ne connaissons pas le statut d'activité des non-répondants finaux. Mais il est possible de constater sur les variables observables pour tous comment se positionnent les répondants à l'ENR face aux deux autres catégories de personnes. Une façon synthétique de poser cette question est de l'envisager sous l'angle de la représentativité. Est-ce que l'ajout des répondants à l'ENR augmente la représentativité de l'échantillon par rapport à l'échantillon global, cette représentativité étant de fait mesurée sur les variables internes à l'échantillon ?

De fait, les répondants à l'ENR semblent avoir des caractéristiques intermédiaires entre celles des répondants à l'enquête Emploi et celles des non-répondants finaux. Ainsi, ils habitent des logements plus petits que les répondants à l'enquête Emploi, et ceci est encore plus vrai pour les non-répondants finaux (cf. tableau 7). Est-ce que rajouter les répondants à l'ENR permet alors de se rapprocher de la distribution observée sur l'ensemble de l'échantillon ? Le test du Chi2 comparant la distribution des tailles de logements des répondants à l'enquête Emploi et celle de l'ensemble de l'échantillon montre que les répondants à l'enquête Emploi sont significativement différents de l'ensemble en ce qui concerne le nombre de pièces de leurs logements. Rajouter les répondants à l'ENR ne change pas ce diagnostic, mais baisse légèrement la statistique du Chi2, améliorant très légèrement la représentativité des répondants.

Tableau 7 : Répartition des répondants selon le nombre de pièces du logement

en %	Non-répondants finaux	Répondants ENR	Répondants enquête Emploi	Echantillon
1	11,7	8,0	5,5	6,4
2	20,0	16,4	12,2	13,4
3	25,2	24,2	21,4	22,0
4	22,9	25,3	27,4	26,7
5 et +	20,2	26,1	33,5	31,5
Ensemble	100,0	100,0	100,0	100,0
Chi2 répondants EE versus échantillon			495	<0,0001
Chi2 répondants EE+ENR versus échantillon			380	<0,0001

Les résultats sont un peu moins homogènes concernant la tranche d'unité urbaine (cf. tableau 8). Globalement, les répondants à l'ENR vivent plus souvent dans des villes, et notamment Paris, que les répondants à l'enquête Emploi, et c'est encore plus net pour les non-répondants finaux. Toutefois, pour les unités urbaines de plus de 20 000 habitants, hors Paris, les répondants à l'ENR semblent différents des non-répondants finaux, vivant plus souvent dans ces zones que les répondants à l'enquête Emploi, alors que c'est l'inverse pour les non-répondants finaux. Toutefois, ces différences sont assez faibles, et la comparaison des tests du Chi2 indique une fois de plus que rajouter les répondants à l'ENR améliore légèrement la représentativité.

Tableau 8 : Répartition des répondants selon la tranche d'unité urbaine

en %	Non-répondants finaux	Répondants ENR	Répondants enquête Emploi	Echantillon
Commune rurale	12,6	14,6	20,6	19,4
UU moins de 20 000	13,7	14,2	15,6	15,3
UU 20 000-199 999	23,5	25,1	24,6	24,5
UU plus de 200 000	24,0	25,1	24,5	24,5
UU de Paris	26,2	21,0	14,7	16,4
Ensemble	100,0	100,0	100,0	100,0
Chi2 répondants EE versus échantillon			350	<0,0001
Chi2 répondants EE+ENR versus échantillon			254	<0,0001

Le tableau 9 montre que sur toutes les variables disponibles, sauf la variable ZUS, l'ensemble des répondants à l'enquête Emploi et à l'ENR est plus représentatif de l'échantillon que les seuls répondants à l'enquête Emploi, puisque les statistiques du Chi2 diminuent lorsqu'on rajoute les répondants à l'ENR. Les effets sont toutefois très faibles.

Tableau 9 : Récapitulatif des Chi2 pour les variables de l'échantillon

Variable	DL	EE seule		EE+ENR	
		Chi2	p-value	Chi2	p-value
Trimestre	3	29	<0,0001	21	<0,0001
Rang d'interrogation	5	14	0,0136	4	0,5938
Nombre de pièces	5	495	<0,0001	380	<0,0001
ZUS	1	5	0,0218	8	0,0050
Type de logement	5	411	<0,0001	297	<0,0001
Tranche d'unité urbaine (9 modalités)	8	356	<0,0001	259	<0,0001
Région	21	462	<0,0001	322	<0,0001
Année de construction du logement	8	85	<0,0001	68	<0,0001
Typologie Tabard des quartiers	27	463	<0,0001	322	<0,0001

Cobben et Schouten (2007) propose de synthétiser cette représentativité à partir de la variance d'une estimation de la probabilité de réponse. Cette probabilité de répondre peut être mesurée par une régression logistique de l'indicatrice répondant / non-répondant sur les variables disponibles sur l'ensemble de l'échantillon. Plus cette probabilité estimée a de variabilité, plus les variables observées permettent de discriminer entre la réponse et la non-réponse, ce qui signifie que l'ensemble des répondants est éloigné d'une bonne représentativité de l'échantillon.

Plusieurs mesures reposant sur cette idée de variabilité sont proposées. Nous utiliserons dans la suite le R-index, qui s'écrit ainsi (cf. Schouten, 2007) :

$$R = 1 - \frac{1}{(n-1)\bar{\rho}(1-\bar{\rho})} \sum_{i=1}^n (\rho_i - \bar{\rho})^2$$

où ρ_i est la probabilité estimée de réponse du logement i , et $\bar{\rho}$ la moyenne des probabilités estimées de réponse.

Le R-index varie entre 0 et 1 (0 à 100 dans le tableau), il vaut 0 lorsque les probabilités de réponses sont les plus dispersées, et il vaut 1 lorsque toutes les probabilités individuelles sont égales à la probabilité moyenne. Dans ce cas, il n'y a aucune sélection de non-réponse observée sur les variables disponibles. Entre les deux, cela signifie qu'on a une sur ou sous-représentation de certains groupes.

Cet indicateur permet notamment de mesurer la variation de représentativité apportée par l'apport de répondants supplémentaires, en comparant le R-index mesuré sur les répondants à l'enquête Emploi avec celui mesuré en rajoutant les répondants à l'ENR. Le tableau 10 montre que l'ENR améliore très légèrement la représentativité. L'effet est plus important au T1, probablement en partie du fait d'un taux de réponse à l'ENR plus important.

Tableau 10 : R-index

	EE seule	EE +ENR
T1	95,5	96,0
T2	95,7	96,0
T3	95,8	96,1
T4	96,9	97,1
Ens 2007	96,1	96,4

Note : La probabilité de réponse utilisée pour le calcul du R-index est estimée à partir d'un modèle logistique. Les résultats de ce modèle sont en annexe 1.

2.3. Méthode d'intégration des résultats de l'ENR

Depuis novembre 2007, il a été décidé d'intégrer les résultats de l'ENR à ceux de l'enquête Emploi (cf. Durier, Gonzalez, Macario-Rat, Thélot, 2007). Néanmoins, intégrer les résultats de l'ENR pose plusieurs difficultés. Tout d'abord, faut-il les rajouter simplement aux répondants à l'enquête Emploi ou peut-on considérer qu'ils sont représentatifs des non-répondants finaux et extrapoler leurs réponses à l'ensemble des non-répondants ? Faire cette hypothèse revient grossièrement à extrapoler les répondants de l'ENR en multipliant leur poids par 4. Cependant, avec 25 % en moyenne sur l'année 2007, le taux de réponse à l'ENR est faible et les non-répondants finaux représentent encore plus de 10 % de la population totale. Il semble donc plus prudent de supposer que les non-répondants finaux ont en moyenne le même profil que la population des répondants à l'enquête principale et à l'ENR.

Par ailleurs, l'ENR se limite aux questions sur le statut d'activité. Rajouter les données individuelles à celles de l'enquête Emploi créerait donc beaucoup de non-réponse partielle. L'ENR est alors uniquement utilisée pour redresser le biais sur les variables d'activité.

Pour cela, un premier calage par sous-échantillon est effectué avec les répondants à l'ENR pour établir des marges d'activité sur l'ensemble des 15 ans ou plus. Puis, on refait un calage sans les non-répondants mais en ajoutant les marges précédemment calculées. On fait donc comme si on disposait de données externes concernant l'emploi et le chômage.

Tableau 11 : Effet de l'ENR sur les taux d'activité et de chômage

	Sur le taux de chômage					Sur le taux d'activité				
	T071	T072	T073	T074	Ens	T071	T072	T073	T074	Ens
Ensemble	-0,1	-0,1	-0,1	-0,1	-0,1	0,2	0,1	0,0	0,1	0,1
Par âge : 15-24 ans	-0,3	-0,2	-0,2	-0,1	-0,2	0,2	0,2	0,1	0,1	0,1
25-49 ans	-0,1	-0,1	-0,1	-0,1	-0,1	0,2	0,1	0,0	0,1	0,1
50-64 ans	-0,1	-0,1	-0,1	0,0	-0,1	0,3	0,2	0,1	0,1	0,2
65 ans et +	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

Lecture : Au 1^{er} trimestre (resp. en 2007) l'ENR a fait diminuer le taux de chômage de 0,1 point (resp. 0,1 point) ; de 0,3 point (resp. 0,2 point) pour les jeunes de 15 à 24 ans.

Un des objectifs de l'enquête non-répondant était de mesurer l'impact de la non-réponse sur le biais de rotation. Malheureusement, le biais de rotation observé sur l'année 2007 est peu marqué, peut-être du fait de taux de réponse particulièrement élevés (cf. tableau 12). La conséquence est que l'intégration des résultats de l'ENR ne permet pas de conclure sur une éventuelle correction du biais. Ainsi, l'impact de l'ENR est une diminution du chômage quelque soit le rang d'interrogation. On observe bien quelques variations dans l'ampleur de l'effet, mais elles restent faibles et pourraient tout à fait être le fait des fluctuations d'échantillonnage. De façon générale, vu les faibles effectifs de l'ENR, on ne pourra probablement pas savoir si l'ENR permet de corriger en partie le biais de rotation avant quelques années d'observation.

Tableau 12 : Effet de l'ENR selon le rang d'interrogation (T1-T4 2007)

Rang d'interrogation	Taux de chômage			Taux d'activité		
	sans ENR	avec ENR	Variation	sans ENR	avec ENR	Variation
1	8,2	8,1	-0,14	56,2	56,4	0,17
2	8,0	7,9	-0,11	56,2	56,2	0,04
3	7,9	7,8	-0,04	56,1	56,2	0,08
4	8,0	7,9	-0,09	56,0	56,1	0,14
5	8,1	8,0	-0,06	56,2	56,3	0,10
6	8,1	8,0	-0,08	56,7	56,8	0,08

Bibliographie :

Cobben F., Schouten B., « R-indexes for the comparison of different fieldwork strategies and data collection modes », *Discussion Paper*, n°07002, 2007, Statistics Netherlands

Cobben F., « Mode Effects in a basic question approach for the Dutch LFS », *Discussion Paper*, n°07003, 2007, Statistics Netherlands

Schouten B., « A Follow-up of Non-response in the Dutch Labour Force Survey », n°07004, 2007, Statistics Netherlands

Durier S., Gonzalez L., Macario-Rat I., Thélot H., «Le chômage baisse depuis 2006», *Insee Première*, n° 1164, novembre 2007, INSEE

Annexe 1: probabilité de réponse estimée par un modèle logistique

Variable	Modalité	EE seule		EE+ENR	
		Param.	P-value	Param.	P-value
Intercept		1,96	<0,001	2,30	<0,001
Trimestre	T1	ref		ref	
	T2	-0,02	ns	-0,10	<0,001
	T3	-0,22	<0,001	-0,27	<0,001
	T4	-0,03	ns	-0,11	<0,001
Rang d'interrogation	1	-0,38	<0,001	-0,22	<0,001
	2	-0,06	0,023	0,02	ns
	3	0,02	ns	0,05	ns
	4	ref		ref	
	5	0,00	ns	-0,01	ns
	6	-0,19	<0,001	-0,16	<0,001
Type de logement	Ferme	0,56	ns	0,74	ns
	Hôtel, etc ...	-0,54	0,002	-0,72	<0,001
	Construction provisoire	0,15	ns	0,03	ns
	Maison individuelle	ref		ref	
	Immeuble d'habitation	-0,27	<0,001	-0,26	<0,001
	Logement de fonction	0,13	ns	0,11	ns
Tranche d'unité urbaine	Comune rurale	0,17	<0,001	0,18	<0,001
	UU <5 000	0,09	0,045	0,12	0,019
	UU <10 000	-0,04	ns	-0,05	ns
	UU <20 000	0,06	ns	0,04	ns
	UU <50 000	ref		ref	
	UU <100 000	0,07	ns	0,10	0,010
	UU <200 000	0,07	ns	0,09	0,044
	UU <2 000 000	0,12	<0,001	0,15	<0,001
	UU de Paris	0,07	ns	0,05	ns
Nombre de pièces du logement	1	-0,47	<0,001	-0,51	<0,001
	2	-0,22	<0,001	-0,24	<0,001
	3	ref		ref	
	4	0,17	<0,001	0,19	<0,001
	5 et +	0,30	<0,001	0,35	<0,001
Année de construction	Avant 1975	ref		ref	
	De 1975 à 1981	0,10	<0,001	0,13	<0,001
	De 1982 à 1989	0,15	<0,001	0,17	<0,001
	De 1990 à 1998	0,07	0,003	0,10	<0,001
	Après 1998	0,02	ns	0,06	ns
ZUS	Non	ref		ref	
	Oui	-0,04	ns	-0,08	0,016
Région	Île-de-France	-0,40	<0,001	-0,45	<0,001
	Champagne-Ardenne	-0,05	ns	-0,11	ns
	Picardie	ref		ref	
	Haute-Normandie	0,05	ns	0,05	ns
	Centre	-0,22	<0,001	-0,14	0,021
	Basse-Normandie	-0,13	0,033	-0,16	0,018
	Bourgogne	0,22	<0,001	0,27	<0,001
	Nord-Pas-de-Calais	0,37	<0,001	0,31	<0,001
	Lorraine	-0,07	ns	-0,06	ns
	Alsace	0,10	ns	0,05	ns
	Franche-Comté	0,19	0,002	0,16	0,027
	Pays de la Loire	0,51	<0,001	0,43	<0,001

	Bretagne	0,28	<0,001	0,25	<0,001
	Poitou-Charentes	0,14	0,028	0,11	ns
	Aquitaine	-0,15	0,006	-0,24	<0,001
	Midi-Pyrénées	-0,31	<0,001	-0,37	<0,001
	Limousin	0,04	ns	0,08	ns
	Rhône-Alpes	-0,18	<0,001	-0,21	<0,001
	Auvergne	0,05	ns	-0,02	ns
	Languedoc-Roussillon	0,01	ns	-0,07	ns
	Provence-Alpes-Côte d'Azur	-0,13	0,014	-0,15	0,007
	Corse	-0,65	<0,001	-0,75	<0,001
Typologie Tabard	ADPUB1	ref		ref	
	ADPUB3	-0,15	<0,001	-0,18	<0,001
	AGRI12	0,02	ns	0,00	ns
	AGRI13	0,35	<0,001	0,28	0,001
	AGRI21	-0,22	<0,001	-0,17	0,020
	AGRI22	0,18	0,002	0,09	ns
	AGRI31	0,24	<0,001	0,16	0,022
	CHOMA1	-0,06	ns	-0,15	0,008
	CHOMA2	0,14	0,003	0,12	0,022
	CHOMA3	0,31	<0,001	0,24	<0,001
	CHOMA4	-0,04	ns	-0,15	ns
	DIR1	-0,17	0,001	-0,17	0,002
	DIR3	-0,08	ns	-0,07	ns
	DIR4	0,10	0,031	0,02	ns
	DIR5	-0,20	<0,001	-0,15	0,003
	INDOUV1	0,08	ns	0,00	ns
	INDOUV3	0,14	0,036	0,10	ns
	INDOUV4	0,22	0,002	0,19	0,013
	INDOUV5	-0,01	ns	-0,07	ns
	INDQ2	-0,02	ns	-0,09	0,037
	INDQ3	0,04	ns	0,04	ns
	INDQ4	0,08	0,043	0,06	ns
	INDQ5	0,19	<0,001	0,12	0,038
	SEMAG2	-0,12	0,010	-0,16	0,002
	SEMAG3	0,15	0,001	0,10	ns
	TEC2	0,06	ns	0,03	ns
	TEC3	0,09	0,031	0,08	ns

Annexe 2: questionnaires de l'enquête auprès des non-répondants

Questionnaire Logement



ENQUETE SUR L'EMPLOI

Feuille de logement

Numéro de la DR : _ _	Numéro de l'enquêteur : _ _
Date de début de la semaine de référence : _ _ _ _ _ _ _ _	jour mois année
Numéro d'aire : _ _ _ _ _ _ _ _ _ _	
Numéro d'immeuble : _ _ _	Numéro de local : _ _

1 Inscrivez l'ensemble des personnes habitant dans ce logement, même une partie de la semaine, y compris celles qui sont temporairement absentes.

	Prénom	Sexe	Date de naissance Jour/mois/année	
1		M <input type="checkbox"/> F <input type="checkbox"/>	_ _ _ _ _ _ _ _	<input type="checkbox"/>
	Si la personne a 65 ans et plus et ne travaille plus, ne pas remplir de bulletin individuel, cochez la case ci-contre →			<input type="checkbox"/>
2		M <input type="checkbox"/> F <input type="checkbox"/>	_ _ _ _ _ _ _ _	<input type="checkbox"/>
	Si la personne a 65 ans et plus et ne travaille plus, ne pas remplir de bulletin individuel, cochez la case ci-contre →			<input type="checkbox"/>
3		M <input type="checkbox"/> F <input type="checkbox"/>	_ _ _ _ _ _ _ _	<input type="checkbox"/>
	Si la personne a 65 ans et plus et ne travaille plus, ne pas remplir de bulletin individuel, cochez la case ci-contre →			<input type="checkbox"/>
4		M <input type="checkbox"/> F <input type="checkbox"/>	_ _ _ _ _ _ _ _	<input type="checkbox"/>
	Si la personne a 65 ans et plus et ne travaille plus, ne pas remplir de bulletin individuel, cochez la case ci-contre →			<input type="checkbox"/>
5		M <input type="checkbox"/> F <input type="checkbox"/>	_ _ _ _ _ _ _ _	<input type="checkbox"/>
	Si la personne a 65 ans et plus et ne travaille plus, ne pas remplir de bulletin individuel, cochez la case ci-contre →			<input type="checkbox"/>
6		M <input type="checkbox"/> F <input type="checkbox"/>	_ _ _ _ _ _ _ _	<input type="checkbox"/>
	Si la personne a 65 ans et plus et ne travaille plus, ne pas remplir de bulletin individuel, cochez la case ci-contre →			<input type="checkbox"/>
7		M <input type="checkbox"/> F <input type="checkbox"/>	_ _ _ _ _ _ _ _	<input type="checkbox"/>
	Si la personne a 65 ans et plus et ne travaille plus, ne pas remplir de bulletin individuel, cochez la case ci-contre →			<input type="checkbox"/>
8		M <input type="checkbox"/> F <input type="checkbox"/>	_ _ _ _ _ _ _ _	<input type="checkbox"/>
	Si la personne a 65 ans et plus et ne travaille plus, ne pas remplir de bulletin individuel, cochez la case ci-contre →			<input type="checkbox"/>

Merci de nous répondre par retour du courrier et au plus tard dans la prochaine quinzaine

Ce questionnaire, une fois rempli, doit être replié de façon à ce que l'adresse de l'Insee (en bas à droite) apparaisse dans la fenêtre de l'enveloppe-retour T ci-jointe

INSEE - DR du Nord-Pas-de-Calais
POLE ENQUETE EMPLOI
AUTORISATION 91404
59049 LILLE CEDEX

