
MODE DE COLLECTE ET QUESTIONNAIRE, QUELS IMPACTS SUR LES INDICATEURS DE L'ENQUÊTE EMPLOI ?

Klara Vinceneux ()*

() INSEE, Direction des statistiques démographiques et sociales, Département de l'emploi et des revenus d'activité, Programme d'évolution de l'enquête Emploi*

`klara.vinceneux@insee.fr`

Mots-clés. Collecte Internet, effet de mode, effet questionnaire, méthodologie d'enquête, multi-mode

Résumé

Dans le cadre du futur règlement européen sur les enquêtes sociales (IESS), le questionnaire de l'enquête Emploi va devoir évoluer. Les modifications attendues auront probablement un impact sur la mesure des principaux indicateurs relatifs au marché du travail. À la suite d'un appel à candidature européen, l'Insee a mis en œuvre une expérimentation dont l'objectif principal était de mesurer un « effet questionnaire » sur le taux d'emploi et le taux de chômage en particulier, mais aussi d'analyser le lien entre la modification des questions et les estimations obtenues.

L'Insee a choisi de mener ce test en ligne, sans aucune intervention d'enquêteurs. 40 000 ménages ont été exclusivement interrogés sur Internet : la moitié recevant le questionnaire actuel de l'enquête Emploi ; l'autre moitié une version proposée par Eurostat. La comparaison des réponses apportées à chacune des deux versions du questionnaire a permis de mettre en exergue un « effet questionnaire » sur les principaux indicateurs du marché du travail.

Une fois l'effet questionnaire analysé, les données collectées par Internet ont ensuite été confrontées à celles collectées au même moment en face-à-face dans le cadre de l'enquête Emploi. Cet effet de mode a été décomposé en un effet de sélection et un effet de mesure, comme le préconise la littérature. L'analyse du profil des répondants a permis de mettre en évidence un effet de sélection sur Internet. Ainsi, les ménages ayant des revenus élevés ont une plus grande propension à répondre sur Internet. Il en est de même pour les ménages propriétaires ou résidant dans des communes urbaines. Les personnes âgées de 65 ans ou plus sont, quant à elles, moins représentées au sein des répondants à l'enquête en ligne que les plus jeunes. Une fois cet effet de sélection corrigé, l'écart est de deux points sur le taux de chômage : il s'élève à 12 % lorsqu'il est calculé à partir des données collectées en ligne, contre 10 % en face-à-face. La comparaison des données post correction de l'effet de sélection a également permis de révéler une sur-représentation des chômeurs au sens BIT parmi les répondants sur Internet.

Abstract

The French Labour Force Survey will have to evolve, in particular following the adoption of the future European framework Regulation on social surveys. In this context, the French National Institute of Statistics and Economic Studies (INSEE) launched an experimental household online survey, whose main goal is to evaluate a “questionnaire effect” on the employment rate. This survey, led on 40,000 households in France in the second quarter of 2016, allowed to obtain first relatively reliable and innovative estimates of what could be “questionnaire effects” and “mode effects” on the main indicators of the French Labour Force Survey.

Introduction

Le questionnaire de l'enquête Emploi va devoir évoluer, notamment à la suite de l'adoption du futur règlement européen sur les enquêtes sociales (Integrated European Social Statistics - IESS) dans lequel s'inscrit la *Labour Force Survey* (LFS), dont l'enquête Emploi est la version française. L'harmonisation plus stricte de la partie du questionnaire permettant de qualifier le statut d'emploi au sens du Bureau international du travail (BIT), risque d'entraîner une rupture de série sur les principaux indicateurs du marché du travail, aux premiers rangs desquels le taux d'emploi et le taux de chômage. Eurostat a lancé en 2015 un appel à candidatures pour mener un test pilote portant sur les questions du module relatif à la situation sur le marché du travail, l'objectif étant notamment d'arriver à mieux prendre en compte les différentes formes d'emploi. Participant à la task force élaborant le modèle de questionnaire européen, l'Insee a proposé à Eurostat de tester cette partie du questionnaire.

À la suite de cet engagement, l'Insee a mis en œuvre en 2016, l'expérimentation dite de « Module A ». Son principal objectif était de mesurer un « effet questionnaire » sur le taux d'emploi et le taux de chômage. L'expérimentation devait aussi permettre d'analyser l'effet de chaque modification introduite sur ces indicateurs.

Pour des questions d'opportunité, ce test a été réalisé entièrement sur des questionnaires proposés sur Internet, sans l'intervention d'enquêteur. En effet, ce test entraine dans le cadre du projet d'expérimentation Multimode sur l'enquête Emploi (Muse), préparatoire à la refonte de l'enquête Emploi. 40 000 ménages ont donc été exclusivement interrogés sur Internet : la moitié recevant le questionnaire actuel de l'enquête Emploi ; l'autre moitié une version proposée par Eurostat.

Ce document présente les principaux résultats obtenus à partir des données collectées dans le cadre de l'expérimentation « Module A ». Il propose dans un premier temps une estimation de l'effet questionnaire sur les indicateurs d'emploi, de chômage et d'inactivité au travers d'une comparaison des résultats obtenus sur chacun des deux échantillons (questionnaire européen et questionnaire actuel sur Internet). En présence avérée d'un effet questionnaire sur les principaux indicateurs du marché du travail, il s'attelle ensuite à repérer l'existence d'un effet de mode induit par la collecte par Internet sur les principaux indicateurs du marché du travail en confrontant les données collectées par Internet à celles collectées au même moment en face-à-face dans le cadre de l'enquête Emploi. Cet effet de mode est analysé et décomposé en un effet de sélection et un effet de mesure, comme le préconise la littérature. Enfin, la question de l'agrégation des données en présence d'un effet de mode sur les indicateurs est posée en annexe. Faut-il corriger les données collectées sur Internet ? Si oui, comment procéder pour obtenir des indicateurs corrigés de l'effet de mesure ? Peut-on obtenir des indicateurs limitant les ruptures de série en considérant le face-à-face comme le mode de collecte de référence ? Est-ce légitime ?

1 L'expérimentation module A : champ retenu et méthodologie d'enquête

1.1 Le champ

L'expérimentation a été menée auprès d'un échantillon de ménages de France métropolitaine, dans lesquels la personne de référence (ou son conjoint le cas échéant) était âgée de moins de 71 ans. L'enquête Emploi portant sur les logements ordinaires, les communautés ont été exclues de l'enquête. Les personnes ont été interrogées dans leur logement habituel et seules les personnes âgées de 15 ans ou plus ont eu à remplir un questionnaire.

1.2 Méthodologie et protocole

Le tirage a été effectué dans le Répertoire statistique des logements, construit à partir des fichiers fiscaux relatifs à l'année 2015. Compte tenu des objectifs de l'enquête, le plan de sondage a été établi de façon à minimiser la variance des estimateurs des variables d'intérêt (à savoir le taux de chômage et le taux d'emploi).

Une expérimentation Internet dite « de grande envergure » réalisée au cours du 1^{er} trimestre 2016 et portant également sur la première interrogation de l'enquête Emploi a permis de construire des groupes de réponse homogènes en fonction des caractéristiques des ménages présentes dans les données de la taxe d'habitation. Le tirage de l'échantillon « Module A » a été stratifié selon quatre variables déterminées par une analyse des facteurs expliquant la non-réponse observée dans le cadre de l'expérimentation de grande envergure : l'âge de la personne de référence, le revenu du foyer, le statut d'occupation et la présence d'une adresse mail dans les fichiers fiscaux. Le plan de sondage a ainsi été construit de façon à améliorer la précision des estimations réalisées à partir des réponses des sous-populations les plus concernées par la non-réponse.

La taille totale de l'échantillon est de 40 000 logements. L'allocation déterminée pour chaque strate est proportionnelle au nombre de logements par strate en tenant compte du taux de réponse au test de grande envergure dans la strate, ainsi que d'une sur-représentation des moins de 26 ans d'un rapport de 3 et des 26-30 ans d'un rapport de 2.

Ainsi :

$$n_h = \frac{R_h}{\sum_{h=1}^{17} R_h} \cdot 40000 \qquad R_h = \frac{r}{t_h} \cdot N_h$$

avec n_h l'allocation de la strate, N_h la taille de la strate dans la base de sondage, r le rapport de sur-représentation de la strate (en fonction de l'âge) et t_h le taux de collecte de la strate dans le test précédent.

Un tirage aléatoire a permis de déterminer le type de questionnaire proposé à chaque ménage enquêté. Ainsi, la moitié de l'échantillon avait accès au questionnaire en ligne de l'enquête Emploi actuelle ; l'autre moitié à la version en ligne du questionnaire proposé par Eurostat. Compte tenu des objectifs, le questionnaire a été restreint à la partie relative au logement, au module A-BIT, à partir duquel les principaux indicateurs de l'enquête peuvent être estimés, ainsi qu'au module B, permettant de qualifier les emplois occupés. Le questionnaire a porté sur les semaines de référence du 16 au 22 mai, du 23 au 29 mai ou du 30 mai au 5 juin, en fonction du sous-échantillon auquel les ménages appartenaient. Les données ont été collectées du 23 mai au 26 juin 2016.

2 Résultats généraux de l'expérimentation

2.1 Le taux de réponse

Le taux de réponse global, défini comme la part de ménages ayant validé au moins un questionnaire individuel, s'élève à 24 % (30 %, une fois prises en compte les pondérations initiales - Figure 1).

FIGURE 1 – Résultat de la collecte et taux de réponse

| | Avant pondération | | | Après prise en compte des poids de sondage | | |
|--|-------------------|---------------|---------------|--|-------------------|-------------------|
| | Eurostat | EEC | Ensemble | Eurostat | EEC | Ensemble |
| | Part (en %) | Part (en %) | Part (en %) | Part (en %) | Part (en %) | Part (en %) |
| Répondant | 24 | 23 | 24 | 31 | 30 | 30 |
| Ayant validé l'enquête et au moins un questionnaire individuel | 23 | 22 | 22 | 29 | 28 | 29 |
| En cours avec au moins un questionnaire individuel validé | 1 | 1 | 1 | 2 | 2 | 2 |
| Non répondant | 75 | 76 | 75 | 69 | 70 | 69 |
| Ayant validé l'enquête mais aucun questionnaire individuel | 0 | 0 | 0 | 0 | 0 | 0 |
| N'ayant ni validé l'enquête ni validé aucun questionnaire validé | 3 | 3 | 3 | 4 | 4 | 4 |
| Aucune information reçue | 39 | 40 | 39 | 39 | 40 | 40 |
| Renvoi d'un coupon-réponse | 17 | 17 | 17 | 15 | 15 | 15 |
| N'habite pas à l'adresse indiquée (pli non distribué) | 16 | 16 | 16 | 10 | 10 | 10 |
| Déchet | 0 | 0 | 0 | 0 | 0 | 0 |
| Ensemble | 100 | 100 | 100 | 100 | 100 | 100 |
| Effectif | 19 945 | 19 952 | 40 000 | 11 045 100 | 11 054 800 | 22 099 900 |

Champ : Ménages enquêtés dans une résidence principale.

Source : Expérimentation « Module A ».

Certaines caractéristiques semblent liées au taux de réponse. Ainsi, les ménages ayant déclaré un revenu supérieur à 40 000 euros, propriétaires de leur résidence principale, résidant dans une maison, dont la personne de référence a entre 30 et 59 ans ou avec un enfant, sont sur-représentés parmi les répondants. Il en est de même pour les ménages « connectés », c'est-à-dire ayant transmis une adresse mail au fisc lors du paiement de la taxe d'habitation.

Inversement, d'autres caractéristiques semblent plutôt sur-représentées au sein des ménages ayant envoyé un coupon-réponse signalant leur impossibilité ou leur refus de participer à l'enquête en ligne. Ainsi, les ménages ayant déclaré des revenus inférieurs à 40 000 euros, non connectés et âgés de 50 ans ou plus (particulièrement les 60 ans ou plus) ou sans enfants sont sur-représentés parmi les ménages dans l'impossibilité ou ayant refusé de participer à l'enquête en ligne (qu'ils aient ou non démarré la saisie d'un ou plusieurs questionnaires).

La participation globale des ménages (réponses complètes ou partielles), dépend très peu du questionnaire proposé. Ceci s'explique notamment par le fait que le protocole ait été identique quel que soit le questionnaire proposé. La part de répondants, définie comme la part de ménages ayant validé au moins un questionnaire individuel, est à peine plus élevée parmi les ménages ayant reçu la version d'Eurostat (31 %) que parmi les autres (30 %).

En excluant les ménages impossibles à joindre (pli non distribuable), le taux de réponse global s'élève à 27 % (34 % une fois pris en compte les poids de sondage) (Figure 2). Ce taux est nettement inférieur au taux de réponse de l'EEC (79 %, cf infra). Il s'élève à 34 % parmi les ménages ayant été sélectionnés pour répondre à la version Eurostat du questionnaire et à 33 % parmi ceux qui ont reçu l'enquête Emploi actuelle.

FIGURE 2 – Résultat de la collecte et taux de réponse parmi les ménages contactés

| | Avant pondération | | | Après prise en compte des poids de sondage | | |
|--|-------------------|-------------|-------------|--|-------------|-------------|
| | Eurostat | EEC | Ensemble | Eurostat | EEC | Ensemble |
| | Part (en %) | Part (en %) | Part (en %) | Part (en %) | Part (en %) | Part (en %) |
| Répondant | 29 | 28 | 28 | 34 | 33 | 34 |
| Ayant validé l'enquête et au moins un questionnaire individuel | 28 | 26 | 27 | 33 | 31 | 32 |
| En cours avec au moins un questionnaire individuel validé | 1 | 2 | 2 | 2 | 2 | 2 |
| Non répondant | 71 | 72 | 72 | 66 | 67 | 66 |
| Ayant validé l'enquête mais aucun questionnaire individuel | 0 | 0 | 0 | 0 | 0 | 0 |
| N'ayant ni validé l'enquête ni validé aucun questionnaire validé | 4 | 4 | 4 | 4 | 4 | 4 |
| Aucune information reçue | 47 | 47 | 47 | 44 | 45 | 45 |
| Renvoi d'un coupon-réponse | 20,0 | 20,3 | 20,1 | 16,9 | 16,8 | 16,9 |
| Ensemble | 100 | 100 | 100 | 100 | 100 | 100 |
| Effectif | 16 708 | 16 749 | 33 457 | 9 864 876 | 9 904 565 | 19 769 441 |

Champ : Ménages enquêtés dans une résidence principale et contactés.

Source : Expérimentation « Module A ».

Le taux de réponse global parmi les ménages contactés varie de la même façon que le taux de réponse global, en fonction des caractéristiques du ménage. Le léger écart observé précédemment sur le taux de réponse global en faveur des ménages ayant reçu la version du questionnaire proposée par Eurostat s'accroît sur des populations particulières. C'est le cas des ménages ayant déclaré un revenu supérieur à 40 000 euros, qui sont 45 % à avoir validé au moins un questionnaire lorsqu'ils ont reçu la version Eurostat, contre 42 % de ceux qui ont reçu la version actuelle. De la même façon, l'écart est plus important parmi les ménages les plus jeunes ou les plus âgés. Les ménages « non connectés » ont proportionnellement davantage répondu à la version Eurostat du questionnaire, obtenant un taux de réponse de 24 % contre 22 % parmi les ménages ayant reçu la version actuelle. En revanche, les ménages composés de trois enfants ou plus ont davantage répondu à la version actuelle du questionnaire qu'à la version d'Eurostat.

2.2 La part d'interruptions et d'abandons

Le **taux d'abandon**, défini comme la proportion de ménages ayant débuté la saisie d'un questionnaire logement et/ou individuel au moins, mais n'ayant pas validé l'enquête ou ayant finalement renvoyé un coupon-réponse, parmi les ménages contactés (hors plis non distribuables), s'élève dans cette expérimentation à 14 % (une fois pris en compte les poids de sondage -Figure 3). La majeure partie des abandons concernent des ménages ayant renseigné le questionnaire logement et débuté la saisie d'au moins un questionnaire individuel (9 %) ; les autres n'ont commencé la saisie d'aucun questionnaire individuel (4,5 %) ou ont envoyé un coupon-réponse (2,2 %).

FIGURE 3 – Taux d'abandon parmi les ménages contactés

| | Avant pondération | | | Après prise en compte des poids de sondage | | |
|----------------------------------|-------------------|-------------|-------------|--|-------------|-------------|
| | Eurostat | EEC | Ensemble | Eurostat | EEC | Ensemble |
| Ensemble | 11,9 | 13,3 | 12,6 | 12,9 | 14,1 | 13,5 |
| Saisie du QL débutée | 4,6 | 4,4 | 4,5 | 4,6 | 4,5 | 4,5 |
| Saisies des QL et QI débutées | 7,3 | 8,9 | 8,1 | 8,3 | 9,6 | 9,0 |
| Ensemble | 11,9 | 13,3 | 12,6 | 12,9 | 14,1 | 13,5 |
| Enquête non validée | 9,9 | 11,1 | 10,4 | 10,8 | 11,8 | 11,3 |
| Saisie du QL débutée | 3,6 | 3,4 | 3,5 | 3,6 | 3,5 | 3,6 |
| Saisies des QL et QI débutées | 6,3 | 7,7 | 6,9 | 7,2 | 6,9 | 7,7 |
| Envoi d'un coupon-réponse | 2,0 | 2,3 | 2,2 | 2,1 | 2,3 | 2,2 |
| Saisie du QL débutée | 1,0 | 1,1 | 1,0 | 1,0 | 1,0 | 1,0 |
| Saisies des QL et QI débutées | 1,0 | 1,3 | 1,1 | 1,4 | 1,1 | 1,3 |

Champ : Ménages enquêtés dans une résidence principale et contactés.
Source : Expérimentation « Module A ».

Le taux d'abandon est un peu plus élevé parmi les ménages ayant reçu le questionnaire de l'enquête Emploi actuelle que parmi ceux interrogés via le questionnaire proposé par Eurostat. Ce constat peut s'expliquer par le fait que le questionnaire individuel proposé par Eurostat ait été allégé (au travers du module A) comparativement à celui de l'enquête Emploi actuelle.

Pour des raisons vraisemblablement similaires, **la réponse a un peu moins souvent été interrompue au cours du remplissage du questionnaire d'Eurostat : 41 % des questionnaires individuels validés ont connu une interruption de saisie d'une durée supérieure ou égale à dix minutes, contre 42 % de ceux de l'enquête Emploi (Figure 4). Finalement, 36 % des ménages interrogés se sont connectés au site pour démarrer la saisie. Parmi eux, 81 % ont validé l'enquête, 16 % l'ont débutée mais n'ont pas terminé la saisie et 3 % ont finalement envoyé un coupon-réponse (Figure 5).**

FIGURE 4 – Part d'interruptions de dix minutes ou plus parmi les répondants

| Nombre de QI à saisir | Avant pondération | | | Après prise en compte des poids de sondage | | |
|-----------------------|-------------------|-------------|-------------|--|-------------|-------------|
| | Eurostat | EEC | Ensemble | Eurostat | EEC | Ensemble |
| Un | 37,8 | 39,5 | 38,6 | 33,7 | 36,0 | 34,8 |
| Deux | 39,4 | 41,4 | 40,4 | 38,4 | 40,3 | 39,4 |
| Trois ou plus | 51,0 | 50,6 | 50,8 | 47,8 | 48,2 | 48,0 |
| Ensemble | 42,2 | 43,6 | 42,9 | 40,5 | 42,3 | 41,4 |

Champ : Ensemble des questionnaires individuels validés.
Source : Expérimentation « Module A ».

FIGURE 5 – Avancement de l'enquête parmi les ménages qui se sont connectés au site de collecte

| | Avant pondération | | | Après prise en compte des poids de sondage | | |
|-----------------------------|-------------------|-----|----------|--|-----|----------|
| | Eurostat | EEC | Ensemble | Eurostat | EEC | Ensemble |
| Connexion enregistrée | 100 | 100 | 100 | 100 | 100 | 100 |
| Enquête validée | 82 | 79 | 80 | 82 | 80 | 81 |
| Enquête débutée non validée | 15 | 17 | 16 | 15 | 16 | 16 |
| Envoi d'un coupon réponse | 3 | 4 | 3 | 3 | 3 | 3 |

Champ : Ménages enquêtés dans une résidence principale, contactés, pour lesquels une connexion a été enregistrée.
 Source : Expérimentation « Module A ».

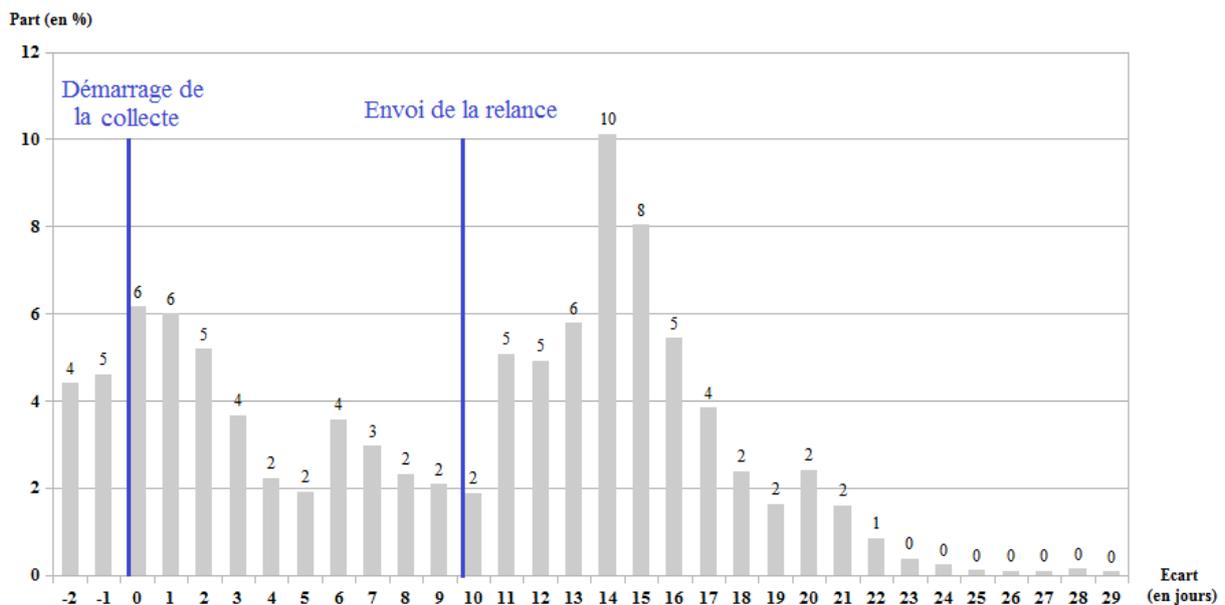
2.3 Validation des questionnaires individuels

La date de validation d'un questionnaire individuel apparaît liée à celle de la réception d'un courrier et/ou d'un mail, qu'il s'agisse de la lettre avis ou d'une relance (envoyée dix jours après le début de la collecte – Figure 6).

Les questionnaires individuels ont davantage été validés en début ou en fin de semaine, près de 30 % d'entre eux le week-end (Figure 7).

Enfin, la majorité des répondants a choisi de valider son questionnaire dans l'après-midi ou la soirée (58 %) ; 20 % ont procédé à la validation en matinée ; 18 % entre 11h30 et 14h30 et 4 % durant la nuit (Figure 8).

FIGURE 6 – Distribution du nombre de jours écoulés entre la date de début de collecte et la date de validation du questionnaire individuel



Champ : Questionnaires individuels validés dans un ménage répondant.

Source : Expérimentation Module A

Note : Compte tenu de l'ouverture précoce du site, certains questionnaires ont pu être saisis avant la date officielle de début de collecte.

FIGURE 7 – Distribution des dates de validation des questionnaires individuels par jour

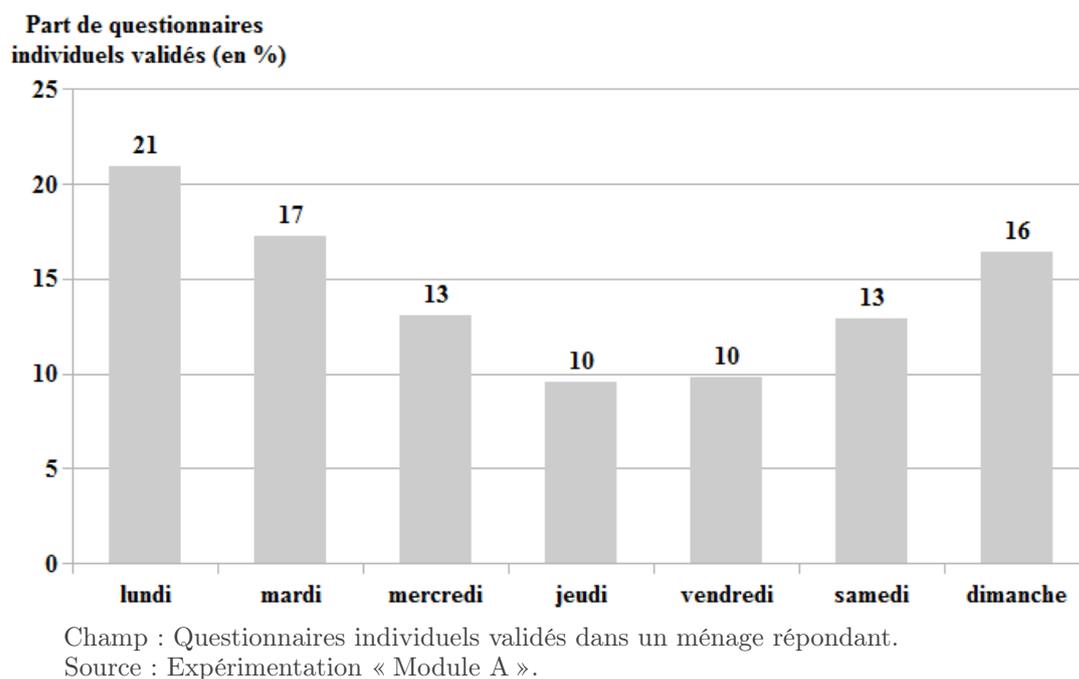
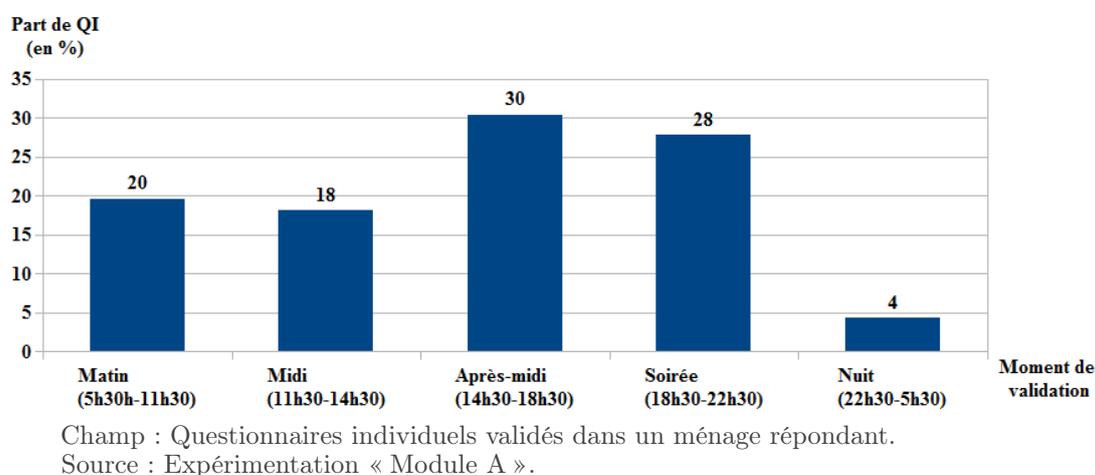


FIGURE 8 – Répartition des questionnaires individuels en fonction du moment de validation



2.4 Le taux d'exhaustivité

Le **taux d'exhaustivité**, c'est-à-dire la part de ménages ayant validé l'intégralité de leurs questionnaires individuels parmi l'ensemble des ménages répondants, **est à peine plus élevé parmi les ménages ayant reçu le questionnaire proposé par Eurostat que parmi ceux ayant reçu la version actuelle** (93 % contre 92 %, une fois pris en compte les poids de sondage). En se restreignant aux ménages ayant au moins deux questionnaires individuels à saisir, le constat reste identique mais les taux d'exhaustivité diminuent un peu pour atteindre respectivement 91 % et 90 % (Figure 9). Comme précédemment, ce constat s'explique vraisemblablement par le fait que la version proposée par Eurostat soit plus courte et plus intuitive que la version actuelle du questionnaire.

FIGURE 9 – Taux d'exhaustivité

| Taux d'exhaustivité | Avant pondération | | | Après prise en compte des poids de sondage | | |
|------------------------|-------------------|------|----------|--|------|----------|
| | Eurostat | EEC | Ensemble | Eurostat | EEC | Ensemble |
| Au moins 2 QI à saisir | 90,1 | 88,8 | 89,5 | 90,5 | 89,6 | 90,1 |
| Ensemble | 93,6 | 92,5 | 93,1 | 93,0 | 92,1 | 92,6 |

Champ : Questionnaires individuels validés dans un ménage ayant validé a minima un questionnaire individuel.

Source : Expérimentation « Module A ».

2.5 Le taux de proxy

Le **taux de proxy corrigé**, c'est-à-dire le nombre de questionnaires individuels saisis par un proche de la personne à laquelle le questionnaire était initialement destiné, rapporté au nombre de questionnaires individuels moins le nombre de ménages ayant validé au moins un questionnaire¹, **s'élève à 16,3 %**, une fois prises en comptes les pondérations initiales (Figure 10). **Les questionnaires saisis par un proxy concernent le plus fréquemment des jeunes âgés de moins de 29 ans et/ou inactifs** (ni en emploi, ni chômeur au sens du BIT). Ainsi 54 % d'inactifs sont décomptés à partir des questionnaires saisis par proxy, contre 34 % dans l'ensemble de la population interrogée. Les questionnaires concernant les femmes sont également plus souvent saisis par un proxy que les autres (55 % de questionnaires contre 45 % pour les hommes). **Après pondération, le taux de proxy corrigé est à peine plus élevé parmi les répondants au questionnaire proposé par Eurostat que parmi ceux ayant répondu au questionnaire actuel (16,6 % contre 15,9 %)**. À titre de comparaison, le **taux de proxy corrigé de l'enquête Emploi en continu s'est établi, au T2 2016, à 48 % en face-à-face et 74 % par téléphone**.

1. Nombre de QI validés par un proxy / (nombre de QI validés - nombre de ménages ayant validé au moins un QI).

FIGURE 10 – Taux de proxy corrigé

| | Avant pondération | | | Après prise en compte des poids de sondage | | |
|-------------------------------|-------------------|------|----------|--|------|----------|
| | Eurostat | EEC | Ensemble | Eurostat | EEC | Ensemble |
| Taux de proxy corrigé* | 15,9 | 15,7 | 15,8 | 16,6 | 15,9 | 16,3 |

Champ : Questionnaires individuels validés dans un ménage ayant validé a minima un questionnaire individuel.

Source : Expérimentation « Module A ».

* Nombre de QI validés par un proxy/(nombre de QI validés - nombre de ménages ayant validé au moins un QI).

3 L'effet questionnaire : aspects méthodologiques

Dans le cadre de l'expérimentation « Module A », la détection d'un potentiel impact de la modification du questionnaire sur les estimations des principaux indicateurs de l'enquête Emploi repose sur la comparaison des résultats obtenus en exploitant les réponses apportées d'une part par les ménages ayant reçu la version actuelle du questionnaire de l'enquête Emploi, et d'autre part par les ménages ayant reçu la version proposée par Eurostat. L'effet questionnaire mesuré par comparaison est alors un effet observé dans le cadre d'une expérimentation Internet. Le résultat obtenu ne peut donc pas nécessairement être transposé au cadre d'une modification du questionnaire intervenue sous le protocole classique d'une première interrogation en face-à-face, en l'absence de travaux complémentaires. Une expérimentation similaire réalisée dans les mêmes conditions que l'enquête Emploi actuelle, c'est-à-dire notamment en face-à-face avec un enquêteur et en première interrogation, pourrait permettre d'estimer quel serait l'impact réel d'un changement de questionnaire sur les principaux indicateurs du marché du travail en première interrogation.

3.1 Méthodologie mise en œuvre pour corriger la non-réponse

Pour observer un potentiel impact du questionnaire sur les principaux indicateurs calculés à partir des données de l'enquête Emploi, il est possible de comparer directement les indicateurs calculés à partir des données non corrigées de la non-réponse². Cependant, compte tenu du faible taux de réponse obtenu dans le cadre de cette expérimentation, et en vue d'obtenir des indicateurs comparables à ceux de l'enquête Emploi en continu actuellement sur le terrain, il est préférable de corriger la non-réponse.

La méthode mise en œuvre consiste, dans un premier temps, à analyser la non-réponse totale, en considérant, comme précédemment, tout ménage ayant validé au moins un questionnaire individuel et n'ayant pas retourné de coupon-réponse, comme un ménage répondant. Cette analyse permettra, par l'intermédiaire d'une procédure de repondération, d'ajuster à la hausse les poids des unités répondantes, et de compenser ainsi l'impact des ménages non-répondants sur l'estimation. En vue de réaliser une estimation fiabilisée des principaux indicateurs calculés à partir de l'enquête Emploi, il est nécessaire de procéder à des corrections des poids initiaux en tenant compte de la non-réponse totale. La non-réponse partielle (cas d'un ménage n'ayant pas validé l'intégralité de ses questionnaires individuels) sera par la suite corrigée en ajustant à la hausse le poids des individus ayant répondu, de façon à compenser l'absence de réponses des individus non répondants au sein du ménage.

2. Les résultats présentés par la suite portent sur les données corrigées de la non réponse. Cependant, les analyses menées sur les données non corrigées de la non-réponse confirment les résultats obtenus après traitement de la non-réponse

La procédure suivie est la suivante :

- **Étape 1** : Sélection des variables auxiliaires liées au fait d'avoir répondu, mais aussi à la variable permettant de déterminer le statut de l'enquêté au regard de l'emploi (actif en emploi, chômeur BIT, inactif) au sein des ménages ayant répondu au moins partiellement à l'enquête. En effet, si l'une des variables retenues est liée à la probabilité de réponse, mais n'est pas liée à la variable d'intérêt, alors il n'est pas souhaitable de l'utiliser, car elle n'aura aucun impact sur la réduction du biais de non-réponse mais pourrait contribuer à faire augmenter la variance des estimateurs.
- **Étape 2** : Modélisation de la variable indicatrice de réponse r_i à l'aide des variables retenues à l'étape précédente, et à partir des questionnaires individuels validés :

La variable indicatrice r_i pour un individu i elle que :

$$r_i = \begin{cases} 1 & \text{si le ménage est répondant} \\ 0 & \text{sinon} \end{cases}$$

- **Étape 3** : Constitution de classes homogènes fondées sur les probabilités de réponse estimées à partir des ménages appartenant à l'échantillon. La probabilité de réponse estimée à l'étape précédente pourrait éventuellement être directement exploitée en vue de calculer les poids corrigés. Il s'agirait alors d'utiliser les probabilités estimées par régression logistique à l'étape précédente, en divisant le poids de chaque ménage répondant par l'inverse de la probabilité estimée. Cependant un tel choix pourrait conduire à des estimateurs biaisés, si la forme de la fonction (ici, logistique) était mal spécifiée. La constitution de classes homogènes de répondants, de nature non paramétrique, assure une protection contre l'éventuelle mauvaise spécification de la forme de la fonction logistique. Pour définir les classes homogènes de ménages du point de vue de la probabilité de réponse, quatre méthodes seront utilisées :
 - la classification ascendante hiérarchique avec constitution de 3 classes (CAH) ;
 - la méthode des quantiles égaux avec constitution de 5 classes (QE5) ;
 - la méthode des quantiles égaux avec constitution de 10 classes (QE10) ;
 - la méthode Haziza et Beaumont (H-B) avec constitution de 11 classes.
- **Étape 4** : Calcul des poids corrigés de la non-réponse, à partir du taux de réponse observé dans chacune des classes construites à l'étape précédente. Une fois les classes homogènes de ménages constituées, il est possible de calculer le taux de réponse propre à chacune des classes, et ce, quelle que soit la méthode utilisée pour former ces dernières. Pour chaque classe constituée, la probabilité de réponse d'un ménage appartenant à la classe est estimée par :

$$\hat{p}_g = \frac{n_{rg}}{n_g}$$

où n_{rg} désigne le nombre de ménages répondants dans la classe g et n_g désigne le nombre de ménages dans la classe g . Ainsi, les probabilités de réponse estimées figurant dans le tableau ci-dessous (Figure 11) sont obtenues.

FIGURE 11 – Calcul des probabilités de réponse estimées par classe

| | CAH | QE5 | QE10 | H-B |
|-----------|------|------|------|------|
| Classe 1 | 0,13 | 0,11 | 0,10 | 0,10 |
| Classe 2 | 0,28 | 0,15 | 0,11 | 0,12 |
| Classe 3 | 0,45 | 0,22 | 0,14 | 0,16 |
| Classe 4 | | 0,30 | 0,17 | 0,20 |
| Classe 5 | | 0,41 | 0,20 | 0,25 |
| Classe 6 | | | 0,25 | 0,29 |
| Classe 7 | | | 0,28 | 0,32 |
| Classe 8 | | | 0,32 | 0,35 |
| Classe 9 | | | 0,37 | 0,41 |
| Classe 10 | | | 0,46 | 0,44 |
| Classe 11 | | | 0,49 | |

Note de lecture : Dans la classe 1, construite à partir de la méthode de classification ascendante hiérarchique (CAH), la probabilité de réponse estimée est de 13% ; dans la classe 1, construite à partir de la méthode des 5 quantiles égaux (QE5), la probabilité de réponse estimée est de 11% ; dans la classe 1, construite à partir de la méthode des 10 quantiles égaux (QE10), la probabilité de réponse estimée est de 10% ; dans la classe 1, construite à partir de la méthode Haziza-Beaumont (HB), la probabilité de réponse estimée est de 10%.

Le poids corrigé de la non-réponse d'un ménage répondant dans la classe est alors calculé, pour chaque méthode, de la façon suivante :

$$w_i^* = w_i \frac{1}{\hat{p}}$$

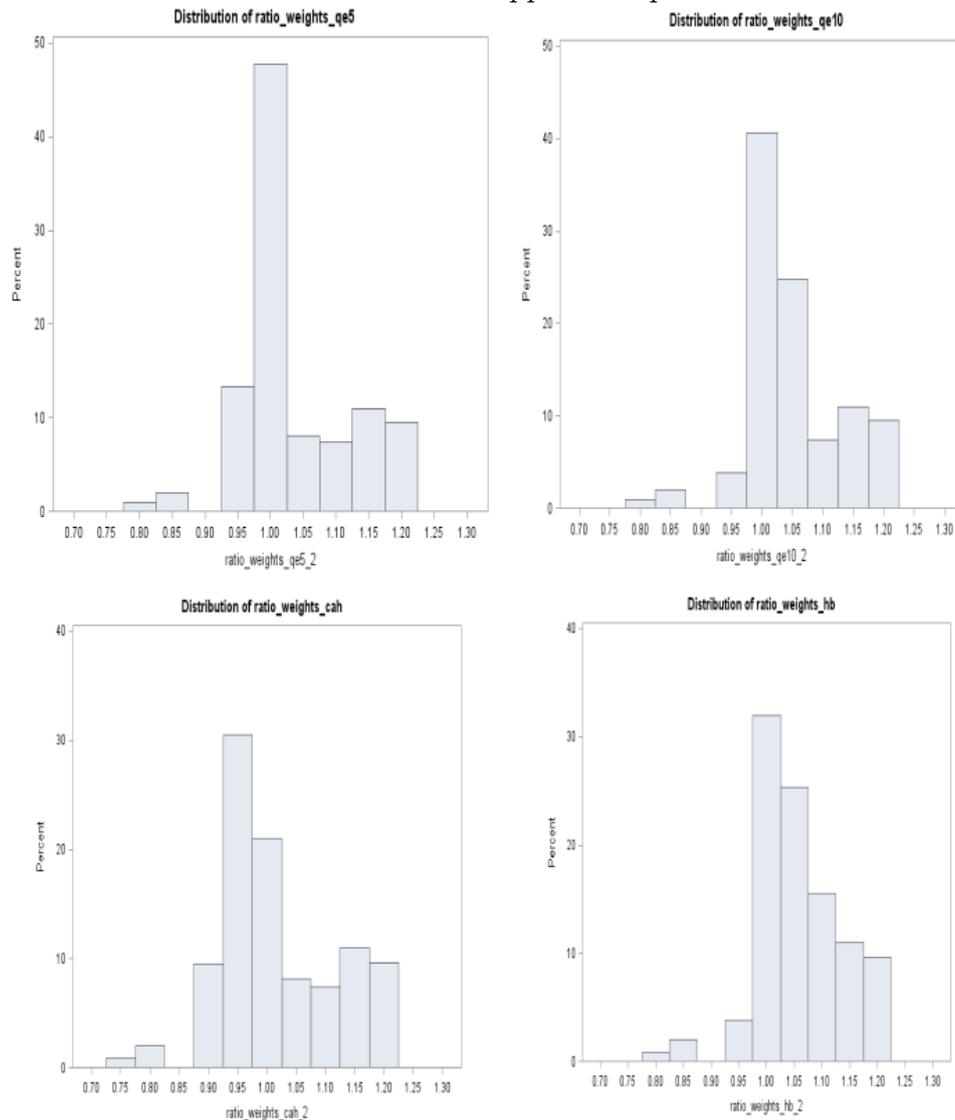
où w_i^* désigne le poids corrigé de la non-réponse et w_i , le poids de sondage initial, tenant compte des probabilités de réponse observés dans le cadre de l'expérimentation de grande envergure.

- **Étape 5** : Correction de la non-réponse partielle et calcul des poids corrigés individuels, chaque individu du ménage se voyant attribuer dans un premier temps le poids corrigé du ménage. Dans un second temps, les poids des individus n'ayant pas validé leur questionnaire individuel sont équitablement répartis sur les poids des répondants.
- **Étape 6** : Correction des poids corrigés de la non-réponse, par calage sur marges dans chacun des deux échantillons. Le calage sur marges permet de compléter le traitement préalable de la non-réponse totale et vise à mettre en cohérence les estimations obtenues avec certaines variables de contexte. Les variables de calage retenues sont le sexe et l'âge.

Au terme de cette étape, les poids corrigés de la non-réponse et calés permettent de calculer une estimation des principaux indicateurs calculés à partir de l'enquête Emploi, à commencer par le taux d'emploi.

Les rapports de poids obtenus entre les poids corrigés de la non-réponse et les poids finaux obtenus après calage apparaissent sur la figure 12.

FIGURE 12 – Rapports de poids



3.2 Les concepts du Bureau International du Travail (BIT)

Par la suite, les différents concepts définis par le Bureau international du travail (BIT) seront utilisés pour définir les indicateurs analysés.

Une personne est en emploi au sens du BIT (ou active occupée) si :

- elle est âgée de 15 ans ou plus et a effectué au moins une heure de travail rémunéré au cours de la semaine de référence ;
- ou si :
- elle a été temporairement absente durant la semaine de référence mais a gardé un lien formel avec l'emploi (maladie de moins d'un an, congés, vacances, grève, congé formation, congé maternité ou parental de moins de trois mois, chômage technique, etc.).

Une personne est au chômage au sens du BIT si :

- elle est âgée de 15 ans ou plus et n'est pas en emploi au cours de la semaine de référence ;
- et disponible pour travailler dans les deux semaines ;

- et a entrepris des démarches actives de recherche d'emploi dans le mois précédent, ou a trouvé un emploi qui commence dans les trois mois.

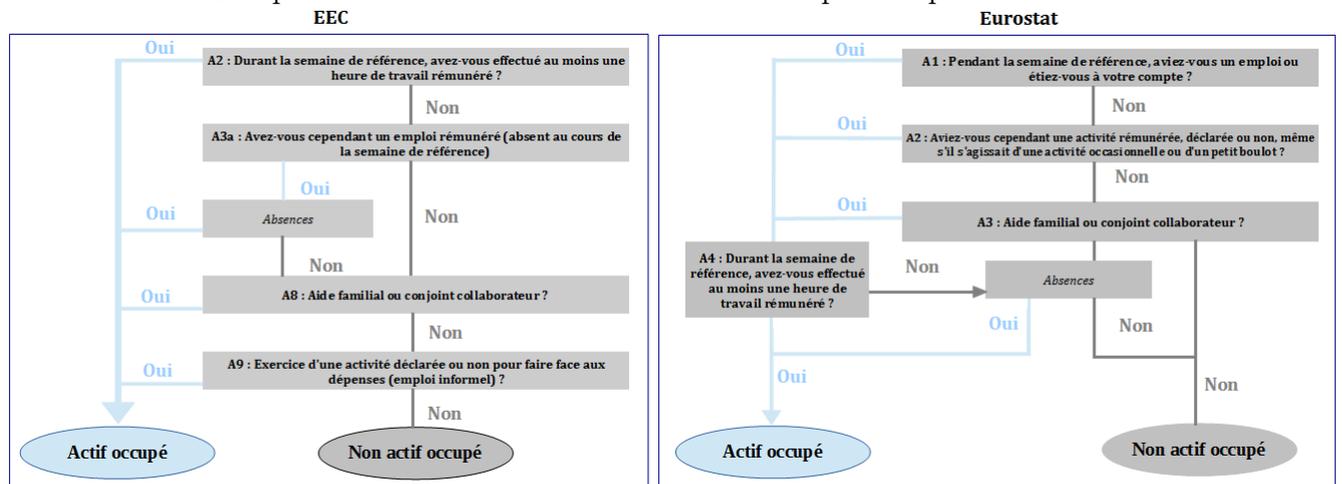
Une personne est inactive au sens du BIT si elle est âgée de 15 ans ou plus et n'est ni en emploi, ni au chômage au sens du BIT au cours de la semaine de référence.

3.3 Principales différences entre les deux versions du Module A-BIT

3.3.1 Concept d'emploi au sens du BIT

Du point de vue du concept d'emploi au sens du BIT, les questionnaires présentent deux différences majeures (Figure 13) :

FIGURE 13 – Comparaison des flowcharts relatifs au concept d'emploi au sens du BIT



- Une différence de concept « At work » vs « One hour », qui conduit à considérer davantage de personnes en emploi via le questionnaire actuel que via la version proposée par Eurostat. Alors que le questionnaire de l'enquête Emploi actuelle pose directement la question permettant de déterminer si l'individu a travaillé au moins une heure pendant la semaine de référence, la version proposée par Eurostat pose prioritairement les questions relatives au fait d'être en emploi ou à son compte, d'avoir une activité rémunérée ou d'être aide familial ou conjoint collaborateur. Ainsi, dans le questionnaire actuel de l'enquête Emploi, le simple fait d'être aide familial ou conjoint collaborateur, ou de déclarer avoir exercé une activité (déclarée ou non), suffit à placer le répondant dans la catégorie des actifs occupés.
- À l'inverse, une restriction de la liste des raisons d'absence associée à des regroupements différents, qui conduisent à considérer davantage de personnes en emploi via le questionnaire proposé par Eurostat que via le questionnaire français.

Voici la liste détaillée des différences observées (Figure 14) :

FIGURE 14 – Comparaison des modalités d'absence dans les deux versions proposées du questionnaire

| Module A – EEC | | Module A – Eurostat | |
|--|--|---|--|
| <i>Pourquoi n'avez-vous pas travaillé cette semaine là ?</i> | En emploi ? | <i>Pour quelle raison principale n'avez-vous pas travaillé la semaine du ... au ... ?</i> | En emploi ? |
| Congés rémunérés (yc RTT, repos compensateur) | Oui | Congés/vacances | Oui |
| Congé maladie (yc enfants malades) ou accident du travail | Si l'interruption dure un an ou moins | Congé maladie | Oui |
| Congé de maternité/paternité | Oui | Congé de maternité/paternité | Oui |
| Temps partiel | Oui | Aménagement du temps de travail (horaires flexibles, repos compensateur, période précédant un nouvel emploi) | Oui |
| Congé parental | Oui si l'interruption dure 3 mois ou moins | Congé parental | Si l'interruption dure 3 mois ou moins ou rémunéré |
| Autres types de congés rémunérés | Oui si l'interruption dure 3 mois ou moins | Autre raison | Si l'interruption dure 3 mois ou moins |
| Formation rémunérée par l'employeur ou dans le cadre d'un contrat en alternance ou en apprentissage | Oui | | |
| Chômage partiel | Oui | | |
| Grève | Oui | | |
| Intempéries | Oui | | |
| Mise à pied, période de fin d'emploi | Non | | |
| Période de morte-saison dans le cadre d'une activité de saisonnier, ou période précédant le début d'emploi | Non | Période de morte-saison | Si réalisation de tâches ou de travaux en lien avec l'activité |

- Dans l'enquête Emploi actuelle, une personne est considérée comme en emploi en cas de congé maladie, sous condition que ce dernier ait une durée totale estimée inférieure à un an. Dans la version proposée par Eurostat, toute personne ayant déclaré être en congé maladie est considérée comme en emploi, quelle que soit la durée du congé.
- La version proposée par Eurostat conduit à considérer une personne ayant déclaré un congé parental rémunéré comme en emploi, contrairement à l'enquête Emploi actuelle ; ceci conduirait par exemple à considérer en emploi une personne en congé parental bénéficiant d'un CLCA « PREPARE ».
- La version proposée par Eurostat conduit à classer « en emploi » une personne ayant déclaré une autre raison que celles proposées, dès lors que la durée totale de l'interruption est estimée à trois mois ou moins. La modalité « autre raison » n'est pas proposée dans la version du questionnaire de l'enquête Emploi actuelle.
- Le questionnaire de l'enquête Emploi actuel conduit à considérer une personne absente dans le cadre d'une période précédant la fin d'un emploi comme non active occupée. À l'inverse, la version d'Eurostat conduit à considérer que la personne est alors en emploi.
- Enfin, alors qu'une personne déclarant être en période dite de « morte saison » n'est pas considérée comme en emploi dans l'enquête Emploi actuelle, elle l'est dans la version proposée par Eurostat, dès lors qu'elle déclare avoir effectué une tâche en lien avec l'activité saisonnière.

Au total, les deux principales différences de questionnaire ont théoriquement un impact de sens contraire sur le fait de considérer une personne comme en emploi au sens du BIT.

Eurostat. Cela peut donc l'amener à se déclarer plus facilement indisponible dans le questionnaire européen et donc à être plus souvent considéré comme inactif.

- La question relative au souhait de travailler est différemment positionnée dans chacun des deux questionnaires. Alors qu'elle figure après la question relative à la recherche d'emploi mais avant la question portant sur les démarches de recherche d'emploi dans le questionnaire proposé par Eurostat (et n'est posée qu'aux personnes ayant déclaré ne pas avoir recherché d'emploi), elle se situe avant l'ensemble des questions relatives à la recherche d'emploi dans le questionnaire français.
- Enfin, la liste des démarches considérées comme étant des démarches actives de recherche d'emploi est réduite dans la version proposée par Eurostat, ce qui conduit à considérer plus souvent une personne comme inactive (Figure 16). En particulier, les modalités « Concours pour entrer dans la fonction publique, un établissement public ou une collectivité territoriale » et « Déplacement dans un salon professionnel, une bourse d'emploi ou un forum des métiers » sont absentes du questionnaire européen.

Contrairement aux deux premières différences décrites précédemment entre les deux questionnaires, ces dernières conduiraient plutôt à considérer davantage de personnes inactives à partir du questionnaire proposé par Eurostat que du questionnaire actuel de l'enquête Emploi.

FIGURE 16 – Comparaison des modalités de recherche d'emploi dans les deux versions proposées

| Module A – EEC | Module A – Eurostat |
|--|--|
| Prise de contact avec Pôle emploi, l'Apec, un opérateur de placement, la chambre de commerce et d'industrie ou un autre organisme public | Avez-vous pris contact avec Pôle emploi, l'Apec, la chambre des métiers ou un autre organisme public ? |
| Prise de contact avec une agence d'intérim ou un opérateur de placement | Avez-vous contacté une agence d'intérim ou un cabinet de placement ? |
| Relations personnelles, amis ou famille | Vous êtes-vous adressé(e) à votre entourage (famille, amis, connaissance) ? |
| Relations professionnelles, anciens collègues, syndicat | |
| Diffusion du profil professionnel sur les réseaux sociaux | Avez-vous mis ou actualisé votre CV sur Internet ? |
| Test ou entretien d'embauche | Test ou entretien d'embauche |
| Concours pour entrer dans la fonction publique, un établissement public ou une collectivité locale | X |
| Dépôt d'une candidature spontanée (salon professionnel ou entreprise) | Avez-vous fait une candidature spontanée auprès d'un employeur ? |
| Envoi d'une candidature spontanée par courrier postal ou électronique ou sur le site Internet de l'entreprise | |
| Déplacement dans un salon professionnel, une bourse d'emploi ou un forum des métiers | X |
| Démarches pour reprendre une entreprise, une exploitation ou un cabinet | Avez-vous fait des démarches pour vous installer à votre compte ? |
| Recherche d'un terrain, de locaux ou de matériel | |
| Démarches pour obtenir des ressources financières | |
| Demande de permis, de licence ou d'autorisation pour s'installer | |
| Dépôt d'une annonce pour trouver un emploi | Avez-vous passé une annonce ou répondu à une offre d'emploi (y compris sur Internet?) |
| Réponse à une annonce d'offre d'emploi | |
| Étude d'annonces d'offre d'emploi | Avez-vous cependant étudié des offres d'emploi (y compris sur Internet ?) |

4 L'effet questionnaire : résultats

4.1 L'effet questionnaire sur les principaux indicateurs

Dans le cadre de cette expérimentation, l'ensemble des indicateurs sont calculés sur la population restreinte aux ménages dont la personne de référence (ou son conjoint, le cas échéant) est âgée de moins de 71 ans. Compte tenu des différences entre les deux versions du questionnaire présentées ci-dessus, les principaux indicateurs sont susceptibles de diverger d'une population à l'autre, sans que l'on puisse déterminer a priori dans quel sens³.

Le taux d'emploi, défini comme le rapport entre le nombre de personnes en emploi et le nombre de personnes âgées de 15 ans ou plus, **a quasiment la même valeur dans les deux populations et s'élève à 57 % environ** (Figure 17).

FIGURE 17 – Principaux indicateurs calculés après correction de la non-réponse

| Indicateurs | | CAH | QE5 | QE10 | HB |
|------------------|---------------------------|-------------|-------------|-------------|-------------|
| Taux d'emploi | Eurostat | 57,5 | 57,2 | 57,1 | 57,0 |
| | EEC | 57,7 | 57,2 | 57,0 | 57,0 |
| | Ecart EEC-Eurostat | 0,1 | -0,1 | -0,1 | 0,0 |
| Taux d'activité | Eurostat | 64,5 | 64,2 | 64,2 | 64,1 |
| | EEC | 65,4 | 65,1 | 65,1 | 65,0 |
| | Ecart EEC-Eurostat | 0,9 | 0,9 | 0,9 | 1,0 |
| Taux de chômage | Eurostat | 10,8 | 10,9 | 11,1 | 11,1 |
| | EEC | 11,9 | 12,2 | 12,4 | 12,4 |
| | Ecart EEC-Eurostat | 1,1 | 1,3 | 1,3 | 1,3 |
| Part de chômeurs | Eurostat | 7,0 | 7,0 | 7,1 | 7,1 |
| | EEC | 7,8 | 7,9 | 8,1 | 8,0 |
| | Ecart EEC-Eurostat | 0,8 | 0,9 | 1,0 | 0,9 |
| Part d'inactifs | Eurostat | 35,5 | 35,8 | 35,8 | 35,9 |
| | EEC | 34,6 | 34,9 | 34,9 | 35,0 |
| | Ecart EEC-Eurostat | -0,9 | -0,9 | -0,9 | -1,0 |
| Effectifs | Eurostat | 44 765 308 | 44 765 308 | 44 765 308 | 44 765 308 |
| | EEC | 44 765 308 | 44 765 308 | 44 765 308 | 44 765 308 |

Note : CAH=Classification ascendante hiérarchique ; QE5=5 quantiles égaux ; QE10=10 quantiles égaux ; HB=Haziza-Beaumont.

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins). Données corrigées de la non-réponse.

Source : Expérimentation « Module A ».

Le taux d'activité, défini comme le rapport entre le nombre d'actifs, qu'ils soient occupés ou non (actifs occupés et chômeurs au sens du BIT), et le nombre de personnes âgées de 15 ans ou plus, **est légèrement plus élevé avec la version actuelle du questionnaire (65 % contre 64 %)**.

3. les résultats présentés dans le tableau 17 sont calculés après traitement de la non-réponse. Le comportement de réponse étant globalement identique quel que soit le type de questionnaire renseigné (Eurostat ou EEC), les écarts entre indicateurs peuvent être calculés sans traiter la non-réponse. Les écarts observés sont alors identiques aux écarts calculés après traitement de la non-réponse.

Il en est de même pour le taux de chômage, défini comme le rapport entre le nombre de chômeurs au sens du BIT et le nombre de personnes actives, occupées ou non, (12 % contre 11 %), et pour la part de chômeurs définie comme le rapport entre le nombre de chômeurs BIT et le nombre de personnes âgées de 15 ans ou plus (8 % contre 7 %).

Inversement et par complémentarité, **la part d'inactifs**, définie comme le rapport entre le nombre de personnes inactives (ni actives occupées, ni chômeuses BIT) et le nombre de personnes âgées de 15 ans ou plus **est plus élevée parmi les répondants au questionnaire proposé par Eurostat**.

Compte tenu des faibles écarts observés entre les méthodes de repondération sélectionnées, les résultats exposés par la suite seront les résultats obtenus avec la méthode « QE10 ».

Qu'advient-il des effets questionnaires sur les principaux indicateurs, une fois prises en compte les caractéristiques des ménages ?

4.2 L'effet questionnaire sur les principaux indicateurs par caractéristiques

L'écart sur le taux d'emploi observé entre les estimations obtenues à partir des répondants au questionnaire actuel de l'enquête Emploi (57,0 %) d'une part et des répondants à la version proposée par Eurostat (57,1 %) d'autre part est très faible (0,1 point – Figure 18). Bien que l'écart soit également relativement faible une fois prises en compte les caractéristiques des personnes enquêtées et des ménages auxquels elles appartiennent, il existe des catégories de personnes pour lesquelles l'effet « questionnaire » a eu un impact plus marqué sur le taux d'emploi. En particulier, le taux d'emploi est plus élevé lorsqu'il est calculé à partir des répondants au questionnaire proposé par Eurostat parmi les répondants âgés de moins de 20 ans ou parmi ceux âgés de 50 à 59 ans ; parmi les ménages dont le revenu déclaré est inférieur à 10 000 euros, compris entre 20 000 et 30 000 euros ou supérieur à 40 000 euros ; ou encore parmi les ménages connectés. Pour certaines caractéristiques, l'effet questionnaire joue en sens inverse : au sein des jeunes de 20 à 29 ans, le taux d'emploi est plus élevé parmi les répondants au questionnaire actuel de l'enquête Emploi que parmi ceux qui ont répondu au questionnaire proposé par Eurostat (58,3 %). De la même façon, les individus dont le questionnaire a été rempli par un proche ont un taux d'emploi de 40,0 % lorsqu'ils ont répondu au questionnaire actuel de l'enquête Emploi, contre 38,5 % lorsqu'ils ont répondu au questionnaire proposé par Eurostat.

L'écart sur le taux de chômage observé entre les estimations obtenues à partir des répondants au questionnaire actuel de l'enquête Emploi (12,4 %) d'une part et des répondants à la version proposée par Eurostat (11,1 %) d'autre part est plus marqué (1,3 point). La version proposée par Eurostat pourrait donc conduire à sous-estimer la proportion de chômeurs BIT, comparativement au questionnaire actuel de l'enquête Emploi. Sauf exception, l'écart observé semble relativement homogène, quelles que soient les caractéristiques des personnes enquêtées ou des ménages auxquels elles appartiennent. Toutefois, un écart important existe pour les moins de 20 ans : l'estimation fondée sur la version proposée par Eurostat conduit à un taux de chômage de 25 %, contre 45 % avec le questionnaire de l'enquête Emploi. Les modifications réalisées sur le questionnaire auraient donc eu un impact particulier sur les plus jeunes (et donc potentiellement sur les étudiants). Dans une moindre mesure, l'écart sur le taux de chômage est également plus marqué parmi les ménages déclarant des revenus inférieurs à 10 000 euros ou compris entre 20 000 et 30 000 euros.

FIGURE 18 – Principaux indicateurs par caractéristiques

| | Taux d'emploi (en %) | | Taux d'activité (en %) | | Taux de chômage (en %) | | Part de chômeurs (en %) | | Part d'inactifs (en %) | |
|--|-------------------------|-------------|---------------------------|-------------|---------------------------|-------------|----------------------------|------------|---------------------------|-------------|
| | Eurostat | EEC | Eurostat | EEC | Eurostat | EEC | Eurostat | EEC | Eurostat | EEC |
| Ensemble | 57,1 | 57,0 | 64,2 | 65,1 | 11,1 | 12,4 | 7,1 | 8,1 | 35,8 | 34,9 |
| Sexe | | | | | | | | | | |
| Femme | 54,7 | 55,0 | 61,6 | 62,8 | 11,3 | 12,5 | 6,9 | 7,8 | 38,4 | 37,2 |
| Homme | 59,4 | 58,9 | 66,6 | 67,2 | 10,9 | 12,4 | 7,3 | 8,3 | 33,4 | 32,8 |
| Age du répondant | | | | | | | | | | |
| Moins de 20 ans | 8,4 | 7,1 | 11,3 | 13,0 | 25,3 | 45,3 | 2,9 | 5,9 | 88,7 | 87,0 |
| Entre 20 et 29 ans | 58,3 | 60,9 | 71,1 | 75,7 | 18,0 | 19,5 | 12,8 | 14,8 | 28,9 | 24,3 |
| Entre 30 et 39 ans | 78,4 | 78,6 | 88,0 | 88,6 | 10,9 | 11,2 | 9,6 | 9,9 | 12,0 | 11,4 |
| Entre 40 et 49 ans | 82,1 | 82,8 | 89,8 | 90,6 | 8,6 | 8,7 | 7,7 | 7,9 | 10,2 | 9,4 |
| Entre 50 et 59 ans | 77,2 | 74,1 | 84,4 | 82,5 | 8,5 | 10,2 | 7,1 | 8,4 | 15,6 | 17,5 |
| Entre 60 et 69 ans | 17,2 | 17,2 | 18,9 | 19,3 | 8,9 | 11,1 | 1,7 | 2,1 | 81,1 | 80,7 |
| 70 ans ou plus | 0,8 | 2,9 | 0,8 | 2,9 | 0,0 | 0,0 | 0,0 | 0,0 | 99,2 | 97,1 |
| Revenu 2015 déclaré | | | | | | | | | | |
| Inférieur à 10 000 € | 38,3 | 35,8 | 57,8 | 58,2 | 33,8 | 38,4 | 19,6 | 22,4 | 42,2 | 41,8 |
| Entre 10 000 et 20 000 € | 55,1 | 56,4 | 64,9 | 67,4 | 15,1 | 16,3 | 9,8 | 11,0 | 35,1 | 32,6 |
| Entre 20 000 et 30 000 € | 60,6 | 59,2 | 65,7 | 66,8 | 7,7 | 11,3 | 5,1 | 7,6 | 34,3 | 33,2 |
| Entre 30 000 et 40 000 € | 59,7 | 62,4 | 63,8 | 66,7 | 6,4 | 6,5 | 4,1 | 4,3 | 36,2 | 33,3 |
| Supérieur à 40 000 € | 62,1 | 61,2 | 65,6 | 64,7 | 5,4 | 5,4 | 3,5 | 3,5 | 34,4 | 35,3 |
| Statut d'occupation du logement | | | | | | | | | | |
| Locataire | 56,7 | 55,6 | 68,2 | 67,8 | 16,9 | 18,1 | 11,5 | 12,3 | 31,8 | 32,2 |
| Propriétaire | 57,0 | 57,7 | 61,1 | 62,8 | 6,6 | 8,2 | 4,1 | 5,2 | 38,9 | 37,2 |
| Type de logement | | | | | | | | | | |
| Appartement | 57,6 | 57,3 | 67,5 | 68,5 | 14,6 | 16,4 | 9,9 | 11,2 | 32,5 | 31,5 |
| Maison | 56,7 | 56,8 | 61,9 | 62,8 | 8,5 | 9,6 | 5,2 | 6,0 | 38,1 | 37,2 |
| Réponse par proxy | | | | | | | | | | |
| Non | 58,5 | 58,3 | 65,8 | 66,6 | 11,0 | 12,4 | 7,3 | 8,3 | 34,2 | 33,4 |
| Oui | 38,5 | 40,0 | 43,7 | 45,7 | 11,9 | 12,4 | 5,2 | 5,6 | 56,3 | 54,3 |
| Ménage connecté | | | | | | | | | | |
| Oui | 62,6 | 61,9 | 68,7 | 68,8 | 8,9 | 10,0 | 6,1 | 6,9 | 31,3 | 31,2 |
| Non | 50,0 | 50,2 | 58,3 | 60,0 | 14,4 | 16,2 | 8,4 | 9,7 | 41,7 | 40,0 |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins). Données corrigées de la non-réponse.
Source : Expérimentation « Module A ».

La part d'inactifs est estimée à 35,8 % à partir du questionnaire proposé par Eurostat contre 34,9 % avec le questionnaire actuel de l'enquête Emploi. L'écart entre les deux questionnaires est donc estimé en moyenne à 0,9 point. L'écart observé est à nouveau relativement homogène selon les caractéristiques des personnes enquêtées ou des ménages auxquels elles appartiennent. Toutefois, des écarts importants existent pour certaines caractéristiques déjà repérées dans le cadre de l'analyse du taux d'emploi. Ainsi, l'écart est de près de 5 points parmi les 20–29 ans. Il est également davantage marqué parmi les ménages dont le revenu déclaré est compris entre 10 000 et 20 000 euros ou entre 30 000 et 40 000 euros ; les ménages non connectés ou encore parmi les personnes dont le questionnaire a été rempli par un proche (proxy). En revanche, de la même façon que pour le taux d'emploi et de façon symétrique, l'effet questionnaire joue parmi les personnes âgées de 50 à 59 ans : la part d'inactifs est moins élevée parmi les répondants au questionnaire proposé par Eurostat (15,6 %), que parmi ceux qui ont répondu au questionnaire de l'enquête Emploi (17,5 %).

L'analyse par caractéristiques socio-démographiques montre donc que l'écart est principalement porté par les moins de 30 ans et dans une moindre mesure par les 50-71 ans. Ces populations sont moins souvent au chômage au sens du BIT dans le questionnaire proposé par Eurostat que dans le questionnaire actuel.

4.3 Décomposition de l'effet questionnaire

La détermination du statut au regard de l'emploi (en particulier, les distinctions entre un actif occupé et un inactif, un chômeur BIT et un inactif) dépend des réponses successives apportées à différentes questions.

Dans la suite de l'analyse, les indicateurs et les répartitions présentés seront calculés sur le champ des 15-74 ans de façon à disposer d'un champ comparable quel que soit le questionnaire étudié (Eurostat ou actuel). Par ailleurs, comme précédemment, les résultats présentés seront calculés à l'aide des poids corrigés de la non-réponse par la méthode « QE10 » (Figure 19).

FIGURE 19 – Répartition des répondants en fonction du statut au regard de l'emploi

| Statut au regard de l'emploi (15-74 ans) | | QE10 | |
|--|----------|------------|-------------|
| | | Effectif | Part (en %) |
| En emploi BIT | Eurostat | 25 539 806 | 57,4 |
| | EEC | 25 509 288 | 57,3 |
| Chômeur BIT | Eurostat | 3 180 220 | 7,2 |
| | EEC | 3 617 024 | 8,1 |
| Inactif | Eurostat | 15 750 317 | 35,4 |
| | EEC | 15 369 070 | 34,5 |
| Ensemble | Eurostat | 44 470 343 | 100,0 |
| | EEC | 44 495 383 | 100,0 |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins). Données corrigées de la non-réponse (méthode « QE10 »).
Source : Expérimentation « Module A ».

4.3.1 Décomposition des effets sur l'emploi

En se restreignant au champ des 15-74 ans, le taux d'emploi s'établit à 57,3 % lorsqu'il est calculé à partir des répondants au questionnaire actuel et à 57,4 % lorsqu'il est calculé à partir des répondants au questionnaire européen. Après décomposition de la provenance des actifs occupés (en fonction des questions ayant permis de les considérer comme en emploi au sens du BIT), il semble que le changement de structure du questionnaire n'ait pas eu d'impact sur le taux d'emploi. Cependant, les actifs occupés ayant déclaré avoir travaillé au moins une heure dans la semaine de référence représentent 94,1 % de l'ensemble de la population active occupée parmi les répondants au questionnaire de l'enquête Emploi, contre 95,1 % parmi les répondants au questionnaire proposé par Eurostat (Figures 21 et 23). La compensation se fait, de façon légère, au travers des autres critères. Ainsi, les personnes ayant déclaré avoir été absentes pour des motifs liés à l'emploi, avoir été aide-familiaux ou avoir eu une activité rémunérée (déclarée ou non) au cours de la semaine de référence sont plus nombreuses parmi les répondants au questionnaire

actuel de l'enquête Emploi que parmi ceux du questionnaire proposé par Eurostat.

Les répondants au questionnaire proposé par Eurostat ne semblent pas davantage déclarer d'activité temporaire ou d'appoint. En revanche, ils sont légèrement plus nombreux à déclarer exercer plusieurs activités (6 % dans le questionnaire proposé par Eurostat contre 5 % dans le questionnaire actuel).

L'analyse plus fine de la décomposition des absences confirme que le changement de questionnaire a un impact sur la part de personnes en emploi (Figure 24).

Plusieurs facteurs expliquent ce résultat :

- Les personnes absentes en raison d'un congé maladie ont intégralement été considérées comme en emploi dans le questionnaire proposé par Eurostat, alors que seules 90,5 % d'entre elles l'ont été dans le questionnaire actuel de l'enquête Emploi ;
- Les personnes absentes en raison d'un congé parental ont été considérées dans le questionnaire actuel comme en emploi uniquement lorsque la durée prévisionnelle du congé était inférieure à trois mois ; à l'inverse, elles l'ont été dans le questionnaire européen dès lors que le congé était rémunéré ou que sa durée prévisionnelle était inférieure à trois mois. Ainsi, 85 % des personnes ayant déclaré être en congé parental ont été considérées comme en emploi avec le questionnaire européen, contre seulement 10 % avec le questionnaire français (Figure 25).
- Les saisonniers ont été considérés comme en emploi dans le questionnaire proposé par Eurostat, dès lors qu'ils ont déclaré avoir réalisé des tâches ou des travaux en lien avec l'activité, alors qu'ils ne le sont pas dans le questionnaire actuel.
- Enfin, le regroupement d'un certain nombre de modalités d'absence dans une catégorie « autre raison » dans le questionnaire proposé par Eurostat aurait quant à lui eu un effet à la baisse sur le taux d'emploi. En effet, la durée d'absence estimée pour ces autres raisons est généralement supérieure à trois mois ou non renseignée.

En résumé, chacun des deux questionnaires permet, a priori, de capter l'emploi de façon similaire. Les deux différences principales semblent n'avoir aucun effet sur le taux d'emploi. Ainsi, l'ordre des questions permettant de capter l'emploi importe vraisemblablement peu. La question « Avez-vous effectué au moins une heure de travail rémunéré durant la semaine de référence ? » reste la plus discriminante, quelle que soit sa position. La faire apparaître plus tôt permettrait, par conséquent, d'alléger la charge de réponse.

Enfin, la restriction de la liste des raisons d'absence et la modification des regroupements des motifs conduisent à considérer davantage de personnes en emploi via le questionnaire proposé par Eurostat que via le questionnaire de l'enquête Emploi. Par ailleurs, les modifications effectuées ont un impact sur la définition même de la notion d'emploi : à titre d'exemple un saisonnier est considéré, hors saison, comme en emploi avec le questionnaire proposé par Eurostat dès lors qu'une tâche ou un travail en lien avec son activité a été réalisé pendant sa période d'absence (ce qui n'est pas le cas avec le questionnaire actuel).

FIGURE 20 – Arbre de décision relatif à l'emploi – Questionnaire actuel

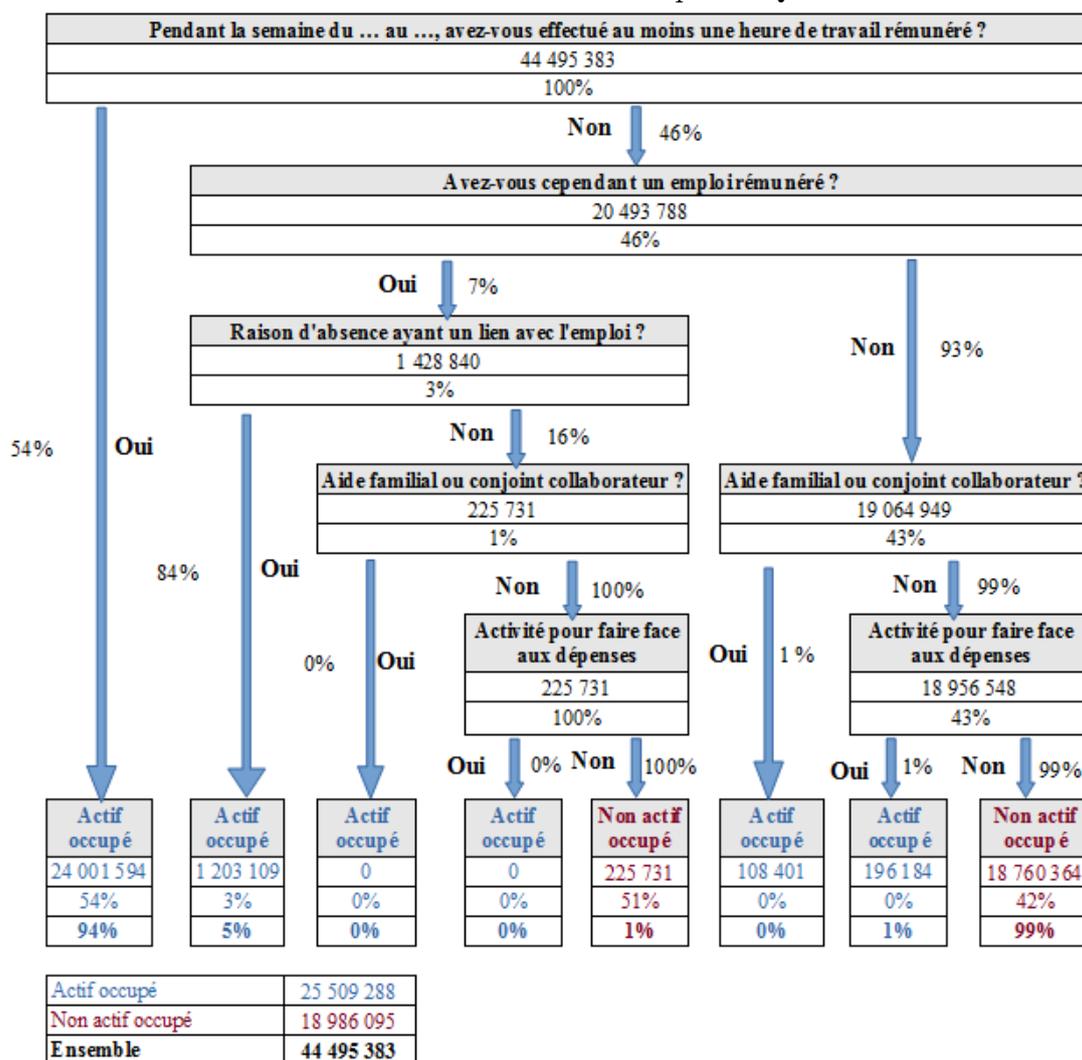


FIGURE 21 – Répartition des actifs occupés en fonction de l'identification de l'emploi – Questionnaire actuel

| | EEC | Effectif | Part | Taux d'emploi |
|---------------------------------|---|-------------------|------------|---------------|
| TRAREF | Pendant la semaine de référence, avez-vous effectué au moins une heure de travail rémunéré ? | 24 001 594 | 94,1 | 53,9 |
| PASTRA + RABS | Avez-vous cependant un emploi rémunéré ? (+ prise en compte des motifs d'absence) | 1 203 109 | 4,7 | 2,7 |
| AIDFAM | Etes-vous aide-familial ou conjoint collaborateur ? | 108 401 | 0,4 | 0,2 |
| INFORM | Toujours pendant la semaine du... au ..., avez-vous effectué une ou plusieurs activités, déclarée(s) ou non, pour faire face aux dépenses ? | 196 184 | 0,8 | 0,4 |
| En emploi au sens du BIT | | 25 509 288 | 100 | 57,3 |
| Ensemble | | 44 495 383 | - | - |

FIGURE 22 – Arbre de décision relatif à l'emploi - Questionnaire proposé par Eurostat

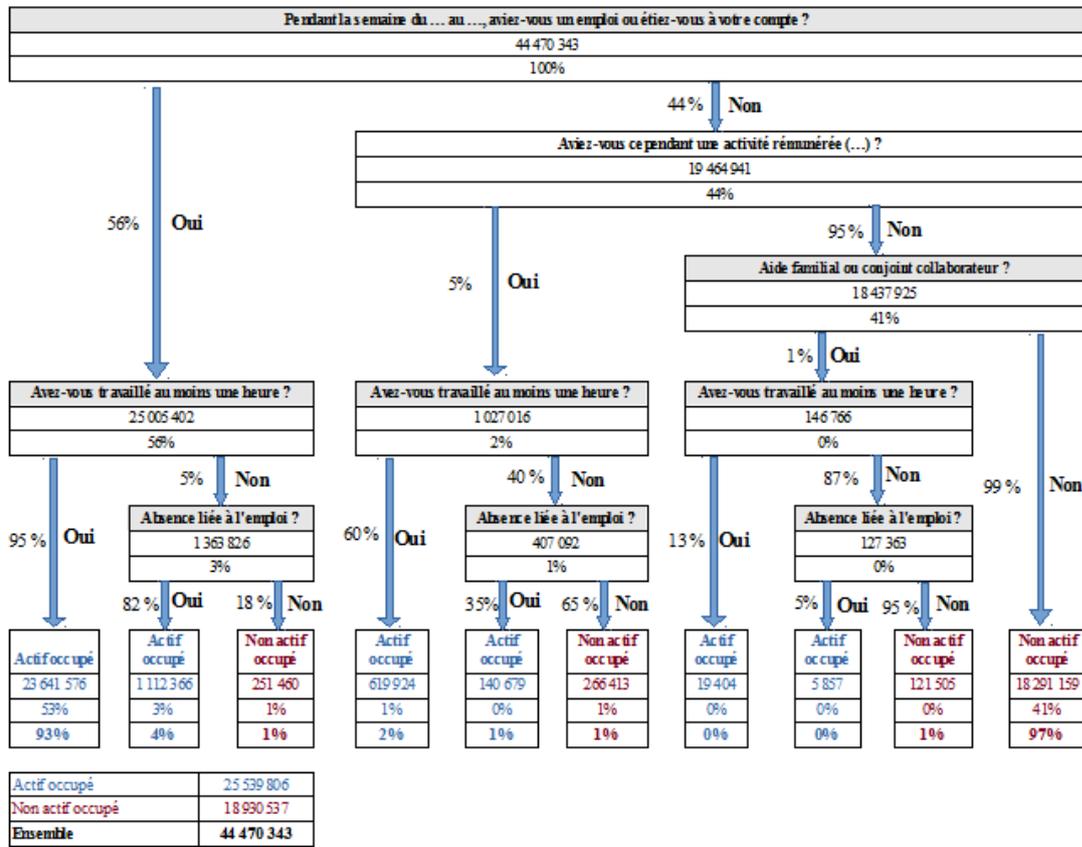


FIGURE 23 – Répartition des actifs occupés en fonction de l'identification de l'emploi - Questionnaire proposé par Eurostat

| | Eurostat | Effectif | Part | Taux d'emploi |
|---------------------------------|--|-------------------|--------------|---------------|
| TRAREF | Pendant la semaine du ... au ... , avez-vous travaillé au moins une heure ? | 24 280 904 | 95,1 | 54,6 |
| PASTRA + RABS | Avez-vous un emploi ou étiez-vous à votre compte ? (+ prise en compte des motifs d'absence) | 1 112 366 | 4,4 | 2,5 |
| AIDFAM + RABS | Pendant la semaine du ... au ... , avez-vous travaillé, sans être rémunéré(e), comme aide-familial ou conjoint collaborateur ? (+ prise en compte des motifs d'absence) | 5 857 | 0,0 | 0,0 |
| INFORM + RABS | Avez-vous cependant une activité rémunérée, déclarée ou non, même s'il s'agissait d'une activité occasionnelle ou d'un petit boulot ? (+ prise en compte des motifs d'absence) | 140 679 | 0,6 | 0,3 |
| En emploi au sens du BIT | | 25 539 805 | 100,0 | 57,3 |
| Ensemble | | 44 470 343 | | 100 |

FIGURE 24 – Part de personnes absentes de l'emploi en fonction de la modalité d'absence sélectionnée

| Module A – EEC | | | | Module A – Eurostat | | | |
|--|--------------|--|----------------|--|--------------|--|----------------|
| Pourquoi n'avez-vous pas travaillé cette semaine là ? | Part en % | En emploi ? | Part en emploi | Pour quelle raison principale n'avez-vous pas travaillé la semaine du... au... ? | Part en % | En emploi ? | Part en emploi |
| Congés rémunérés (yc RTT, repos compensateur) | 22,3 | Oui | 100,0 | Congés/vacances | 23,4 | Oui | 100,0 |
| Congé maladie (yc enfants malades) ou accident du travail | 42,0 | Si l'interruption dure un an ou moins | 90,5 | Congé maladie | 25,3 | Oui | 100,0 |
| Congé de maternité/paternité | 12,4 | Oui | 100,0 | Congé de maternité/paternité | 8,0 | Oui | 100,0 |
| Temps partiel | 4,0 | Oui | 100,0 | Aménagement du temps de travail (horaires flexibles, repos compensateur, période précédant un nouvel emploi) | 1,2 | Oui | 100,0 |
| Congé parental | 7,1 | Oui si l'interruption dure 3 mois ou moins | 10,1 | Congé parental | 4,0 | Si l'interruption dure 3 mois ou moins ou rémunéré | 84,8 |
| Autres types de congés cumulés | 10,6 | | 64,0 | | | | |
| Autres types de congés rémunérés | 4,4 | Oui si l'interruption dure 3 mois ou moins | 53,1 | Autre raison | 35,0 | Si l'interruption dure 3 mois ou moins | 11,9 |
| Formation rémunérée par l'employeur ou dans le cadre d'un contrat en alternance ou en apprentissage | 2,2 | Oui | 100,0 | | | | |
| Chômage partiel | 1,9 | Oui | 100,0 | | | | |
| Grève | 0,4 | Oui | 100,0 | | | | |
| Intempéries | 0,0 | Oui | 0,0 | | | | |
| Mise à pied, période de fin d'emploi | 1,8 | Non | 0,0 | | | | |
| Période de morte-saison dans le cadre d'une activité de saisonnier, ou période précédant le début d'emploi | 1,6 | Non | 0,0 | Période de morte-saison | 3,1 | Si réalisation de tâches ou de travaux en lien avec l'activité | 27,9 |
| Ensemble | 100,0 | | 18,1 | Ensemble | 100,0 | | 66,3 |

FIGURE 25 – Répartition des personnes ayant choisi le congé parental comme motif d'absence dans le questionnaire proposé par Eurostat

| | |
|--|---------------|
| Nombre de personnes absentes pour congé parental | 75 078 |
| Part percevant un salaire ou une indemnité versé(e) au moins partiellement par l'employeur | 84,8 |
| non versé par l'employeur | 17,2 |
| Part ne percevant pas de salaire ou d'indemnité | 67,7 |
| | 15,2 |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins) par une personne ayant déclaré un congé parental.

Données corrigées de la non-réponse (méthode « QE5 »).

Source : Expérimentation « Module A ».

4.3.2 Décomposition des effets sur le chômage

En se restreignant au champ des 15-74 ans, la part de chômeurs s'établit à 8,1 % lorsqu'elle est calculée à partir des répondants au questionnaire actuel et à 7,2 % lorsqu'elle est calculée à partir des répondants au questionnaire proposé par Eurostat (Figures 27 et 29). Cet écart résulte d'une combinaison de facteurs.

Tout d'abord, les personnes déclarant avoir trouvé un emploi débutant dans trois mois ou moins sont moins nombreuses parmi les répondants au questionnaire proposé par Eurostat. Cela s'explique assez simplement par un changement de positionnement de la question. Alors que l'intégralité des non actifs occupés sont interrogés sur le fait d'avoir trouvé un emploi débutant dans les trois mois dans l'enquête Emploi, seules les personnes déclarant ne pas avoir recherché

d'emploi dans les quatre semaines se terminant par la semaine de référence le sont dans la version proposée par Eurostat.

Par ailleurs, les personnes ayant effectué au moins une démarche active de recherche d'emploi sont moins nombreuses parmi les répondants au questionnaire proposé par Eurostat.

Cela peut s'expliquer par la suppression de deux items dans la liste des démarches actives de recherche d'emploi dans le questionnaire européen comparativement au questionnaire français : l'inscription à un concours public d'une part et le déplacement dans un salon professionnel, un forum ou une bourse à l'emploi d'autre part.

Enfin, la complexité de la question relative à la disponibilité dans le questionnaire proposé par Eurostat a sans doute été à l'origine d'une sous-estimation du nombre de personnes disponibles pour travailler dans les deux semaines. Alors que le questionnaire actuel de l'enquête Emploi interroge sur une disponibilité dans les deux semaines suivant l'interrogation et porte sur un emploi susceptible de convenir à l'enquêté « Si vous trouviez un travail qui vous convienne, seriez-vous disponible pour le commencer dans les deux semaines ? » ; la version d'Eurostat oblige l'enquêté à une projection rétrospective « Si on vous avait proposé un emploi pendant la semaine de référence, auriez-vous été disponible pour commencer à travailler dans les deux semaines ? ».

Au total, une combinaison de trois facteurs explique que l'estimation du chômage soit moins élevée dans le questionnaire proposé par Eurostat : l'ordre inversé des questions relatives au fait d'avoir trouvé un emploi débutant plus tard et d'avoir effectué des démarches de recherche d'emploi ; la réduction du nombre de démarches considérées comme des démarches actives ; la complexification de la question relative au fait d'être disponible pour travailler dans les deux semaines.

Il est important de souligner que l'ensemble des résultats présentés ici est difficilement extrapolable à une modification du questionnaire administré par un enquêteur. En effet, les écarts concernent un échantillon de 40 000 ménages interrogés sur Internet, en première interrogation, sans aucune intervention d'un enquêteur (pour introduire l'enquête notamment ou pour apporter des explications sur certaines questions complexes). Par ailleurs, les estimations sont obtenues par comparaison de répondants sur Internet, qui ne représentent sans doute pas parfaitement l'ensemble des ménages interrogés dans le cadre de l'enquête Emploi. Cette analyse fera l'objet de la seconde partie.

FIGURE 26 – Arbre de décision relatif au chômage BIT - Questionnaire actuel

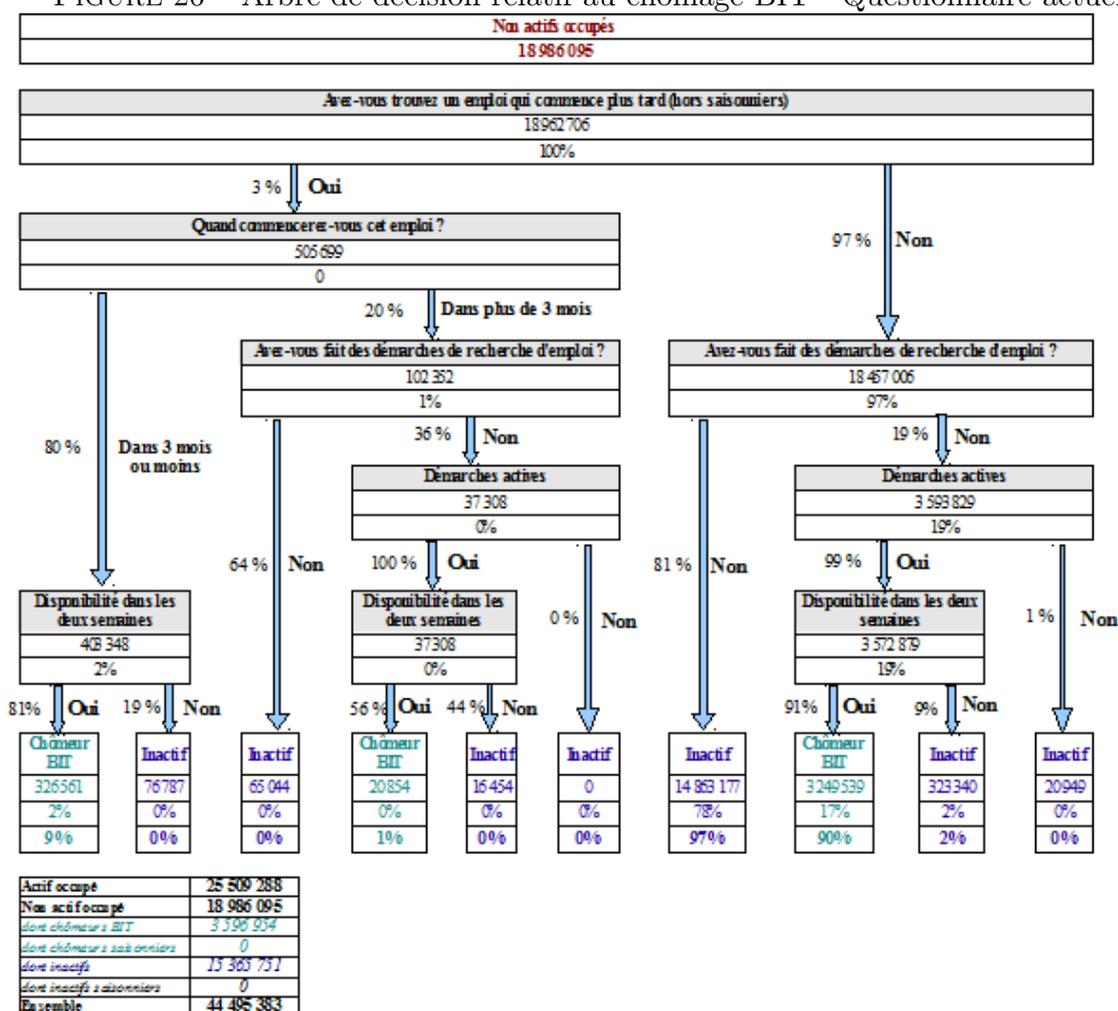


FIGURE 27 – Répartition des chômeurs BIT en fonction de l'identification du chômage - Questionnaire actuel

| | EEC | Effectif | Part | Part de chômage |
|----------------------------------|---|-------------------|--------------|-----------------|
| PASTRB | Avez-vous trouvé un emploi qui commence plus tard (moins de 3 mois) ? | 326 561 | 9,1 | 0,7 |
| PASTRB + MREC | Avez-vous trouvé un emploi qui commence plus tard (plus de 3 mois et démarches actives de recherche d'emploi) ? | 20 854 | 0,6 | 0,0 |
| MREC | Pendant la période ..., avez-vous fait des démarches actives de recherche d'emploi ? (+prise en compte des démarches actives) | 3 249 539 | 90,3 | 7,3 |
| Au chômage au sens du BIT | | 3 596 954 | 100,0 | 8,1 |
| Ensemble | | 44 495 383 | 100,0 | 100,0 |

FIGURE 28 – Arbre de décision relatif au chômage BIT - Questionnaire proposé par Eurostat

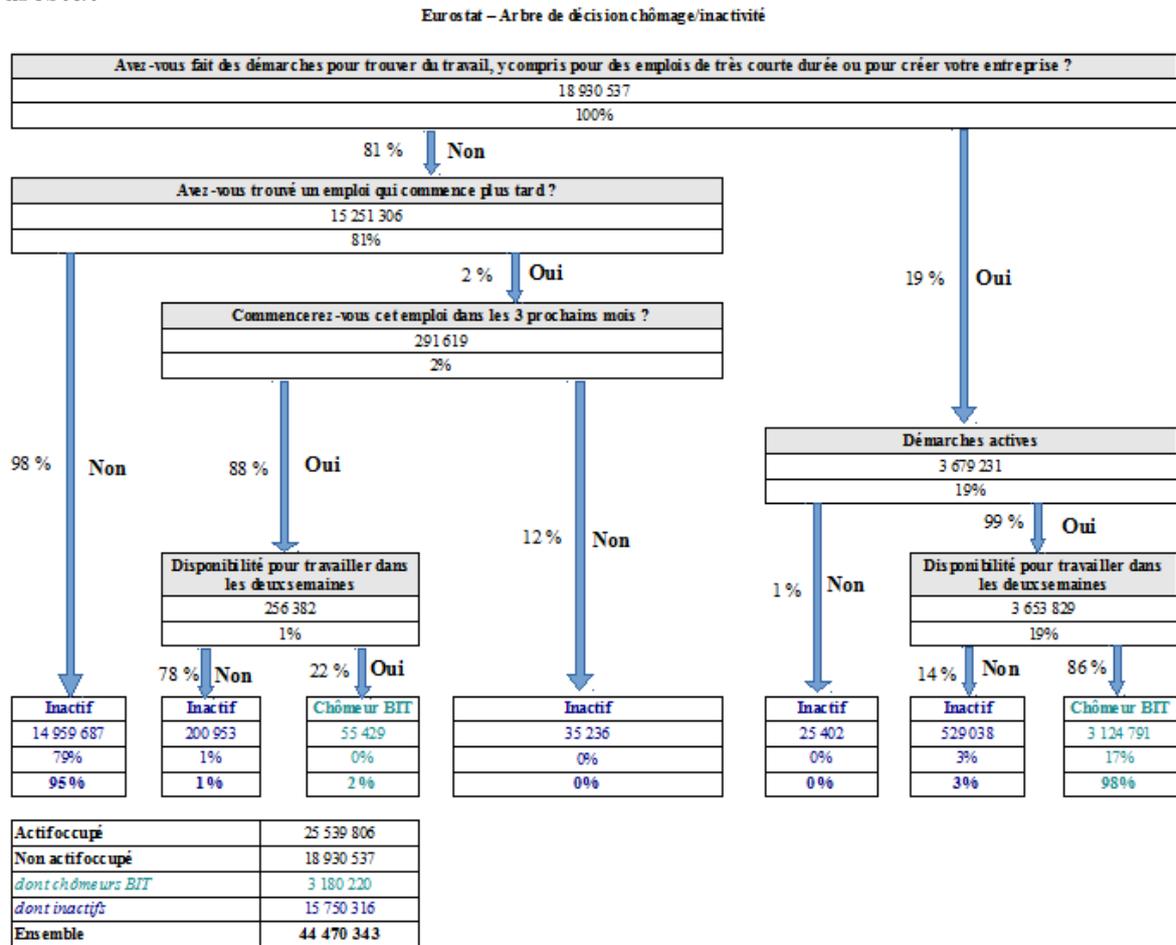


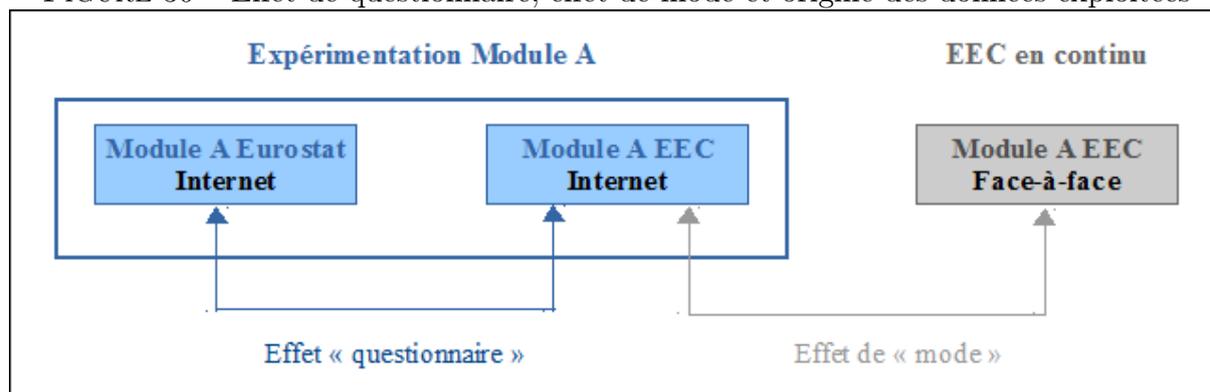
FIGURE 29 – Répartition des chômeurs BIT en fonction de l'identification du chômage - Questionnaire proposé par Eurostat

| Eurostat | | Effectif | Part | Part de chômage |
|----------------------------------|---|-------------------|------------|-----------------|
| PASTRB | Avez-vous trouvé un emploi qui commence plus tard (moins de 3 mois et sans déclarer avoir recherché un emploi) ? | 55 429 | 1,7 | 0,1 |
| PASTRB + MREC | Avez-vous trouvé un emploi qui commence plus tard (plus de 3 mois et démarches actives de recherche d'emploi) ? | 0 | 0,0 | 0,0 |
| MREC | Pendant la période ..., avez-vous fait des démarches actives de recherche d'emploi ? (+prise en compte des démarches actives) | 3 124 791 | 98,3 | 7,0 |
| Au chômage au sens du BIT | | 3 180 220 | 100 | 7,2 |
| Ensemble | | 44 470 343 | ' | ' |

5 L'effet de mode : effet de sélection et effet de mesure

Une fois étudié l'effet « questionnaire » par comparaison des données obtenues à partir des répondants au questionnaire de l'enquête Emploi d'une part et à partir des répondants au questionnaire proposé par Eurostat d'autre part, il est possible d'estimer un potentiel effet de mode en comparant cette fois les données collectées par Internet sur le questionnaire de l'enquête Emploi, à celles collectées, au même moment, sur le terrain en face-à-face par les enquêteurs (EEC) (Figure 30).

FIGURE 30 – Effet de questionnaire, effet de mode et origine des données exploitées



5.1 Sélection d'un champ comparable

Avant de procéder à toute analyse, il est nécessaire de sélectionner un champ comparable au sein des données collectées en face-à-face, à celui des données collectées sur Internet dans le cadre du test de « Module A ». Dans un premier temps, seules les données collectées par l'enquête Emploi en continu en première interrogation durant le deuxième trimestre 2016 ont été conservées, de façon à disposer d'un périmètre temporel commun. Puis, dans un second temps, seuls les ménages interrogés dans leur résidence principale et dont la personne de référence (ou son conjoint le cas échéant) était âgée de 70 ans ou moins ont été conservés dans la base de données. Une fois cette sélection opérée (c'est-à-dire un champ comparable sélectionné), les taux de réponse obtenus pour chacun des deux échantillons peuvent être comparés.

Le taux de réponse global est bien plus élevé lorsque l'enquête a été réalisée en face-à-face par un enquêteur (79 %), que lorsqu'elle a été proposée sur Internet dans le cadre de l'expérimentation « Module A » (26 % - Figure 31).

5.2 Décomposition de l'effet de mode

L'effet de mode inclut deux effets distincts :

- **l'effet de sélection** : les répondants sur Internet ont-ils les mêmes caractéristiques que les répondants en face-à-face ?
- **l'effet de mesure** : à caractéristiques socio-démographiques équivalentes, les enquêtés répondent-ils de la même façon en face-à-face et sur Internet ?

FIGURE 31 – Taux de réponse par mode

| | Face-à-face | | Internet | |
|----------------------|--------------------|-------------|--------------------|-------------|
| | Nombre d'individus | Part (en %) | Nombre d'individus | Part (en %) |
| Dans le champ | 9 898 | 100 | 16 749 | 100 |
| Répondant | 7 827 | 79 | 4 355 | 26 |
| Non répondant | 2 071 | 21 | 12 394 | 74 |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins).

Sources : Expérimentation « Module A » ; Enquête Emploi en continu, T2 2016.

L'effet de sélection peut être défini comme l'effet provenant des différences socio-démographiques existant entre les personnes ayant répondu sur Internet et les personnes enquêtées en face-à-face. Les deux enquêtes (expérimentation « Module A » sur Internet et Enquête Emploi en continu) étant obligatoires, l'effet de sélection devrait, dans ce cadre, être moins fort que ce qu'il serait si les répondants avaient eu le choix de répondre sur Internet ou en face-à-face avec un enquêteur. **L'effet de mesure**, quant à lui, peut être défini, à caractéristiques socio-démographiques égales, comme l'écart induit par les différences observées dans les réponses apportées au questionnaire web, relativement à celles que l'enquêteur saisit lors de la collecte.

5.3 L'effet de sélection

Dans un premier temps, les caractéristiques des répondants à l'enquête Emploi, interrogés en face-à-face par un enquêteur en première interrogation dans leur résidence principale au cours du deuxième trimestre 2016, sont comparées à celles des répondants sur Internet au questionnaire actuel de l'enquête Emploi, dans le cadre de l'expérimentation « Module A ».

Les répondants sur Internet n'ont pas le même profil que ceux qui ont répondu aux enquêteurs. En particulier, les ménages déclarant des revenus supérieurs à 40 000 euros représentent 43 % des répondants sur Internet, alors qu'ils ne représentent que 30 % des répondants en face-à-face (Figure 32). De la même façon, les ménages propriétaires sont sur-représentés parmi les répondants sur Internet. Enfin, les habitants de la région parisienne, et plus généralement les personnes résidant en milieu urbain, sont sur-représentées parmi les répondants sur Internet.

Du point de vue des caractéristiques individuelles (Figure 33), les hommes sont plus enclins à répondre sur Internet que les femmes. Les plus jeunes sont sur-représentés parmi les répondants sur Internet, tandis que les 65 ans ou plus sont plus nombreux à répondre en face-à-face.

L'analyse des caractéristiques des répondants dans chacun des deux échantillons a donc permis de mettre en évidence l'existence d'un effet de sélection non négligeable entre les modes de collecte « face-à-face » et « Internet ».

FIGURE 32 – Caractéristiques des ménages répondants en fonction du mode de collecte

| | Mode de collecte | |
|--|------------------|------------|
| | Face-à-face | Internet |
| Caractéristiques du ménage | | |
| Revenu déclaré dans la taxe d'habitation | | |
| Inférieur à 15 000 € | 20 | 12 |
| Entre 15 000 € et 40 000 € | 50 | 46 |
| Supérieur à 40 000 € | 30 | 43 |
| Statut d'occupation | | |
| Locataire | 35 | 30 |
| Propriétaire | 64 | 69 |
| Non renseigné ou autres | 2 | 1 |
| Type de commune | | |
| Rurale | 29 | 24 |
| Urbaine | 71 | 76 |
| Type de voisinage | | |
| Maisons dispersées, hors agglomération | 17 | 21 |
| Maison en lotissement, quartier pavillonnaire ou ville | 48 | 41 |
| Immeubles en ville | 20 | 18 |
| Immeuble en cité ou grand ensemble | 9 | 5 |
| Habitat mixte : immeuble et maisons | 5 | 10 |
| Non renseigné | 0 | 5 |
| Région | | |
| Ile-de-France | 14 | 19 |
| Autres régions | 86 | 81 |
| Ensemble | 100 | 100 |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins).

Sources : Expérimentation « Module A » ; Enquête Emploi en continu, T2 2016.

FIGURE 33 – Caractéristiques des ménages répondants en fonction du mode de collecte

| | Mode de collecte | |
|---|------------------|------------|
| | Face-à-face | Internet |
| Caractéristiques des individus | | |
| Sexe | | |
| Homme | 47 | 50 |
| Femme | 53 | 50 |
| Age de l'individu au dernier jour de la semaine de référence | | |
| 15-29 ans | 21 | 22 |
| 30-39 ans | 15 | 19 |
| 40-49 ans | 18 | 19 |
| 50-64 ans | 27 | 29 |
| 65 ans ou plus | 19 | 11 |
| Ensemble | 100 | 100 |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins).

Sources : Expérimentation « Module A » ; Enquête Emploi en continu, T2 2016.

5.4 L'effet de mesure

5.4.1 Qu'est-ce que l'effet de mesure ?

L'effet de mesure induit par le recours à des modes de collecte différents correspond, à caractéristiques équivalentes, à l'écart créé par un comportement de réponse variant d'un mode à l'autre.

S'il existe, il peut être le fruit de différents éléments :

- la **désirabilité sociale** qui pourrait par exemple conduire un enquêté à ne pas avouer avoir renoncé à la recherche d'emploi ; ou encore à déclarer davantage de démarches de recherche d'emploi face un enquêteur ;
- - les **reclassements opérés par l'enquêteur** dans les cas où il détecterait des incohérences dans les réponses apportées ;
- - l'**apport d'exemples par l'enquêteur** aiguillant l'enquêté vers certaines modalités de réponse plus souvent que vers d'autres ;
- plus généralement, les **difficultés de compréhension** des concepts ou des modalités de réponse ;
- - le manque d'intérêt initial, ou accru par la longueur du questionnaire, qui conduirait l'enquêté à sélectionner prioritairement les premières modalités proposées sans lire les suivantes, de façon à atteindre plus rapidement la complétion du questionnaire (« satisficing »).

5.4.2 Comment isoler l'effet de mesure ?

Afin de tenter d'isoler ce qui constituerait l'effet de mesure au sein de l'effet de mode, il est nécessaire de corriger l'effet de sélection préalablement mis en exergue. Pour cela, les deux échantillons (face-à-face et Internet) doivent être rendus comparables au moyen d'un traitement de la non-réponse totale. Le traitement de la non-réponse totale est effectué, dans chacun des deux échantillons, en deux temps, comme précédemment : la non-réponse totale est d'abord corrigée par repondération en ajustant à la hausse les poids des unités répondantes au sein de groupes de réponse homogènes constituées sur la base des informations disponibles dans les bases de sondage exploitées pour la construction des échantillons ; dans un second temps, un calage est effectué sur quelques variables supplémentaires en vue de mettre en cohérence la population échantillonnée avec la population française sur le champ concerné par l'enquête.

La méthodologie appliquée est donc comparable, en totalité, à celle mise préalablement en œuvre dans le cadre de l'analyse de l'effet questionnaire. L'étape de constitution des groupes de réponse homogène dans l'échantillon interrogé en face-à-face conduit à l'obtention des probabilités de réponse ci-dessus qui varient de 60 % à 87 % selon les méthodes de constitution des groupes de réponse exploitées et selon les classes considérées (Figure 34).

Une fois la non-réponse totale corrigée, il est procédé, comme précédemment, au traitement de la non-réponse partielle et au calcul des poids corrigés individuels, chaque individu du ménage se voyant attribuer dans un premier temps le poids corrigé du ménage. Dans un second temps, les poids des individus n'ayant pas validé leur questionnaire individuel sont équitablement répartis sur les poids des répondants.

Les poids corrigés de la non-réponse calculés sur l'échantillon interrogé en face-à-face sont ensuite corrigés à nouveau au moyen d'un calage sur marges. Ce dernier permet de compléter

FIGURE 34 – Calcul des probabilités de réponse estimées par classe

| | CAH | QE5 | QE10 | H-B |
|------------------|------------|------------|-------------|------------|
| Classe 1 | 0,68 | 0,70 | 0,65 | 0,60 |
| Classe 2 | 0,77 | 0,77 | 0,73 | 0,65 |
| Classe 3 | 0,85 | 0,78 | 0,75 | 0,74 |
| Classe 4 | | 0,85 | 0,77 | 0,77 |
| Classe 5 | | 0,86 | 0,79 | 0,78 |
| Classe 6 | | | 0,80 | 0,85 |
| Classe 7 | | | 0,84 | |
| Classe 8 | | | 0,86 | |
| Classe 9 | | | 0,86 | |
| Classe 10 | | 0,87 | | |

Note de lecture : Dans la classe 1, construite à partir de la méthode de classification ascendante hiérarchique (CAH), la probabilité de réponse estimée est de 68 % ; dans la classe 1, construite à partir de la méthode des 5 quantiles égaux (QE5), la probabilité de réponse estimée est de 70 % ; dans la classe 1, construite à partir de la méthode des 10 quantiles égaux (QE10), la probabilité de réponse estimée est de 65 % ; dans la classe 1, construite à partir de la méthode Haziza-Beaumont (HB), la probabilité de réponse estimée est de 60 %.

le traitement préalable de la non-réponse totale de façon à mettre en cohérence les estimations obtenues avec certaines variables de contexte. Les variables de calage retenues sont, comme précédemment, le sexe et l'âge.

Au terme de cette étape, les deux échantillons sont désormais similaires en termes de caractéristiques socio-démographiques. Les poids corrigés de la non-réponse et calés permettent donc de calculer une estimation des principaux indicateurs à partir de l'échantillon enquêté en face-à-face. Il sera donc possible de les comparer aux estimations obtenues à partir de l'échantillon interrogé directement sur Internet et d'évaluer ce qui a été préalablement appelé l'effet de mesure.

5.4.3 L'effet de mesure sur les principaux indicateurs

Abstraction faite de la méthode « classification ascendante hiérarchique », l'écart entre le taux d'emploi calculé à partir de l'échantillon interrogé sur Internet dans le cadre de l'expérimentation « Module A » et celui calculé à partir des données collectées en face-à-face dans le cadre de l'enquête Emploi est faible (Figure 35). Le mode de collecte ne semble donc pas avoir eu d'impact fort sur le taux d'emploi.

FIGURE 35 – Principaux indicateurs calculés à partir du Module A-BIT

| | Indicateurs | CAH | QE5 | QE10 | HB | Moyenne |
|------------------|-----------------------------------|-------------|-------------|-------------|-------------|-------------|
| Taux d'emploi | EEC face-à-face | 56,7 | 56,7 | 56,7 | 56,7 | 56,7 |
| | EEC Internet | 57,7 | 57,2 | 57,0 | 57,0 | 57,2 |
| | Ecart Face-à-face/Internet | -1,0 | -0,5 | -0,3 | -0,3 | -0,5 |
| Taux d'activité | EEC face-à-face | 62,9 | 62,8 | 62,9 | 62,9 | 62,9 |
| | EEC Internet | 65,4 | 65,1 | 65,1 | 65,0 | 65,2 |
| | Ecart Face-à-face/Internet | -2,6 | -2,3 | -2,2 | -2,2 | -2,3 |
| Taux de chômage | EEC face-à-face | 9,8 | 9,8 | 9,8 | 9,8 | 9,8 |
| | EEC Internet | 11,9 | 12,2 | 12,4 | 12,4 | 12,2 |
| | Ecart Face-à-face/Internet | -2,1 | -2,4 | -2,6 | -2,6 | -2,4 |
| Part de chômeurs | EEC face-à-face | 6,1 | 6,2 | 6,1 | 6,2 | 6,1 |
| | EEC Internet | 7,8 | 7,9 | 8,1 | 8,0 | 8,0 |
| | Ecart Face-à-face/Internet | -1,6 | -1,8 | -1,9 | -1,9 | -1,8 |
| Part d'inactifs | EEC face-à-face | 37,1 | 37,2 | 37,1 | 37,1 | 37,1 |
| | EEC Internet | 34,6 | 34,9 | 34,9 | 35,0 | 9,0 |
| | Ecart Face-à-face/Internet | 2,6 | 2,3 | 2,2 | 2,2 | 2,3 |
| Effectifs | EEC face-à-face | 44 765 308 | 44 765 308 | 44 765 308 | 44 765 308 | ' |
| | EEC Internet | 44 765 308 | 44 765 308 | 44 765 308 | 44 765 308 | ' |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins). Données corrigées de la non-réponse.

Sources : Expérimentation « Module A » ; Enquête Emploi en continu, T2 2016.

En revanche, l'effet de mesure est observable et non négligeable sur l'ensemble des autres indicateurs, pour lesquels il varie de 1,6 point à 2,6 points en fonction des indicateurs et des méthodes considérées.

Cependant, il est important de garder en mémoire le faible taux de réponse sur Internet, potentiellement à l'origine d'une imprécision plus importante sur les estimations réalisées à partir des données collectées sur Internet, qu'à partir de celles collectées en face-à-face.

5.5 Limites de l'estimation de l'effet de mesure

5.5.1 Un effet de sélection inobservable ?

Compte tenu de l'ampleur des écarts observés (notamment sur le taux de chômage), il est possible que l'effet de mesure, défini comme l'effet induit par le fait de répondre sur Internet plutôt qu'en face-à-face, à caractéristiques sociodémographiques semblables, ne soit pas un « pur » effet de mesure. En effet, la notion de chômage BIT étant construite à partir d'un certain nombre de questions claires et successives, il peut sembler surprenant qu'un effet d'une telle ampleur soit observé. En particulier, il se pourrait que l'effet de sélection n'ait pu être totalement corrigé au moyen des variables socio-démographiques disponibles. En particulier, il ne semble pas déraisonnable de penser que les personnes considérées comme au chômage BIT à l'aide des réponses apportées puissent être davantage enclines à répondre sur Internet qu'en face-à-face.

L'analyse montre que les chômeurs sont, en proportion, bien plus présents dans l'échantillon des questionnaires individuels renseignés sur Internet que dans celui des questionnaires recueillis en face-à-face. En effet, 13 % des ménages ayant répondu sur Internet comprennent au moins un chômeur BIT, contre 9 % des ménages ayant répondu à un enquêteur. En outre, cette observation

FIGURE 36 – Présence de chômeurs BIT en fonction de la composition du ménage

| Composition du ménage | | Mode de collecte | |
|--------------------------------|-------------------------------|------------------|-------------|
| | | Face-à-face | Internet |
| Une personne | Aucun chômeur | 94,3 | 90,0 |
| | Un chômeur | 5,7 | 10,0 |
| Deux personnes | Aucun chômeur | 91,2 | 87,4 |
| | Un chômeur | 8,3 | 11,1 |
| | Deux chômeurs | 0,5 | 1,5 |
| Trois personnes ou plus | Aucun | 80,9 | 79,1 |
| | Un chômeur | 15,9 | 17,9 |
| | Deux chômeurs | 2,8 | 3,0 |
| | Trois chômeurs ou plus | 0,4 | 0,0 |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins). Données corrigées de la non-réponse (méthode « QE5 »). Sources : Données issues de l'exploitation « Module A » ; Enquête Emploi en continu, T2 2016.

est la même, quelle que soit la composition du ménage (Figure 36).

Les données semblent donc plutôt confirmer l'hypothèse selon laquelle les chômeurs BIT sont plus enclins à répondre sur Internet qu'en face-à-face. Ceci peut s'expliquer de différentes façons. En particulier, un sentiment de culpabilité lié à leur situation pourrait conduire les chômeurs à redouter l'interrogation d'un enquêteur sur le sujet de l'emploi. En outre, le nom court de l'enquête, « Enquête Emploi » (pour « Enquête sur l'emploi, le chômage et l'inactivité »), pourrait désinciter les chômeurs BIT à participer à l'enquête, dans la mesure où leur situation n'est pas une situation d'emploi. Sur Internet, la possibilité de se connecter, sans contrainte, peut avoir un impact positif sur le taux de réponse des chômeurs. En effet, ce mode de collecte leur offre la possibilité de prendre conscience rapidement que leur situation est également prise en compte dans le questionnaire. Enfin, les chômeurs BIT, pour partie demandeurs d'emploi (c'est-à-dire inscrits à Pôle emploi), doivent nécessairement se connecter sur Internet de façon à actualiser leur profil chaque mois. Ils sont également tenus de consulter leur messagerie afin de s'assurer qu'une offre d'emploi ne leur a pas été transmise par Pôle emploi. Il s'agit donc, au moins partiellement, d'une population particulièrement connectée, peut-être plus encline à répondre sur Internet.

L'effet de mesure observé pourrait ainsi constituer un effet de mode résiduel dans lequel serait inclus un effet de sélection lié à la variable d'intérêt, non pris en compte par les variables socio-démographiques disponibles pour traiter la non-réponse totale au sein des deux échantillons. Cependant, le faible taux de réponse obtenu sur Internet rend l'hypothèse de la sur-représentativité des chômeurs BIT au sein des répondants Internet, forte, même à caractéristiques observables données.

D'autres limites pourraient potentiellement fragiliser l'estimation de l'effet de mesure. En particulier, le mode de traitement de non-réponse partielle au sein d'un ménage.

5.5.2 Un nombre de chômeurs sur Internet artificiellement accru par le traitement de la non-réponse partielle.

Les analyses précédentes ont montré que la non-réponse partielle était plus fréquente au sein des ménages ayant répondu sur Internet qu'au sein des ménages ayant répondu en face-à-face. Par ailleurs, les ménages « incomplets » ayant répondu sur Internet sont plus souvent constitués d'au moins un chômeur BIT que les ménages « incomplets » ayant répondu en face-à-face. Or, la méthodologie choisie pour traiter la non-réponse partielle au sein des ménages « incomplets » consiste à redistribuer le poids des non-répondants du ménage sur celui des répondants. Ainsi, l'application de cette méthode induit automatiquement un renforcement plus important du poids des chômeurs parmi les personnes ayant répondu sur Internet que parmi celles qui ont répondu en face-à-face.

Afin de quantifier l'impact de cette méthode sur les écarts observés, les indicateurs sont recalculés sans qu'aucun traitement de la non-réponse partielle ne soit appliqué. Autrement dit, le poids des individus non-répondants au sein d'un ménage n'est pas redistribué sur celui des répondants (sa valeur est automatiquement réduite à 0) et seul le poids des individus répondants est pris en compte dans le calcul des indicateurs.

FIGURE 37 – Comparaison des indicateurs en fonction de la méthodologie choisie pour traiter la non-réponse partielle

| Traitement de non réponse partielle | Mode de collecte | Taux d'emploi (%) | Taux d'activité (%) | Taux de chômage (%) | Part de chômage (%) | Part d'inactifs (%) |
|-------------------------------------|------------------|-------------------|---------------------|---------------------|---------------------|---------------------|
| Avec | Face-à-face | 56,7 | 62,8 | 9,8 | 6,2 | 37,2 |
| | Internet | 57,2 | 65,1 | 12,2 | 7,9 | 34,9 |
| <i>Ecart Face-à-face/Internet</i> | | <i>-0,5</i> | <i>-2,3</i> | <i>-2,4</i> | <i>-1,7</i> | <i>2,3</i> |
| Sans | Face-à-face | 56,7 | 62,8 | 9,8 | 6,1 | 37,2 |
| | Internet | 56,9 | 64,8 | 12,2 | 7,9 | 35,2 |
| <i>Ecart Face-à-face/Internet</i> | | <i>-0,2</i> | <i>-1,9</i> | <i>-2,4</i> | <i>-1,8</i> | <i>1,9</i> |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins). Données corrigées de la non-réponse (méthode « QE5 »). Sources : Données issues de l'exploitation « Module A » ; Enquête Emploi en continu, T2 2016.

Sans traitement de la non-réponse partielle, les écarts observés restent globalement semblables à ceux observés préalablement, avec traitement de la non-réponse partielle (Figure 37). Cela démontre que la méthodologie envisagée pour corriger la non-réponse partielle n'a pas d'impact majeur sur les estimations précédentes de l'effet de mode résiduel, une fois l'effet de sélection, non lié à la variable d'intérêt, évincé.

Les écarts observés entre les indicateurs calculés à partir de chacun des deux échantillons pourraient également refléter l'impact des réponses apportées par des proxys, beaucoup plus nombreux en face-à-face que sur Internet.

5.5.3 Un plus grand nombre de proxy en face-à-face à l'origine d'un nombre moindre de chômeurs en face-à-face ?

Il pourrait exister un potentiel effet « proxy » sur les parts de chômeurs obtenues dans l'un ou l'autre des échantillons.

Le taux de proxy est en effet bien plus élevé en face-à-face que sur Internet : 23 % contre 6 %. Par ailleurs, la probabilité que l'enquêté soit inactif est plus élevée lorsque le questionnaire a été renseigné par un proxy, quel que soit le mode de collecte. L'effet proxy pourrait donc automatiquement conduire à une surestimation du nombre d'inactifs en face-à-face. Cependant, l'effet proxy est amplifié sur Internet. Un questionnaire rempli par un proxy sur Internet aboutit bien plus souvent, en proportion, à un statut d'inactif qu'un questionnaire rempli par un proxy en face-à-face. Un questionnaire rempli par un proxy sur Internet a 2,4 fois plus de chances d'aboutir à une catégorisation en inactif qu'un questionnaire rempli par la personne concernée. En revanche, un questionnaire rempli par un proxy en face-à-face a 1,2 fois plus de chances d'aboutir à une catégorisation en inactif qu'un questionnaire rempli par la personne concernée.

D'un point de vue théorique, les deux effets se compensent donc de telle sorte que l'effet proxy dans sa globalité ne puisse expliquer, de façon simple, l'écart observé sur les principaux indicateurs de l'enquête Emploi.

Enfin, des différences entre les méthodes d'échantillonnage utilisées peuvent constituer une dernière limite identifiable aux estimations proposées.

5.5.4 Un biais d'estimation potentiellement induit par des méthodes d'échantillonnage légèrement différentes

Bien que les champs aient été rendus comparables, quelques différences entre les deux échantillons subsistent tout de même :

- Les bases de sondage diffèrent : l'échantillon de l'EEC est tiré dans la taxe d'habitation, alors que celui de l'expérimentation « Module A » a été tiré dans le Répertoire statistique des logements (RSL).
- Le plan de sondage diffère : il s'agit, pour l'EEC, d'un sondage à 2 degrés, stratifié par région en métropole, équilibré sur un certain nombre de caractéristiques présentes dans la base de sondage et uniformément réparti dans le temps ; alors que l'échantillon « Module A » a été stratifié selon quatre variables déterminées par une analyse des facteurs expliquant la non-réponse dans le cadre de l'expérimentation de grande envergure (préalablement effectué dans des conditions similaires - Annexe B).
- L'échantillon de l'EEC contient des logements vacants ou des résidences secondaires qui ne sont enquêtés que s'ils sont devenus des résidences principales, cela n'a pas été le cas de l'échantillon « Module A ».
- Les semaines de référence ne sont pas exactement identiques : les données issues de l'expérimentation « Module A » correspondent aux semaines de référence comprises entre le 16 mai et le 5 juin 2016 ; alors que l'ensemble des données relatives aux semaines de référence du deuxième trimestre 2016 de l'EEC ont été conservées de façon à accroître la taille de l'échantillon tout en conservant une période de collecte cohérente. La restriction des données collectées en face-à-face aux 3 semaines de référence correspondant à celles de l'expérimentation « Module A » aboutit à des conclusions similaires, les écarts entre indicateurs étant légèrement amplifiés.

Ces différences peuvent également induire un biais sur l'effet de mode résiduel estimé, sans que celui-ci ne puisse être mesuré.

En conclusion, la principale limite de l'estimation de l'effet de mesure mise en exergue dans les analyses précédentes semble être celle liée à l'impossibilité de corriger parfaitement l'effet de sélection au moyen des seules variables explicatives disponibles dans les bases de sondage. Il serait donc possible que l'effet de mesure estimé englobe un effet de sélection non corrigé et directement lié à la variable d'intérêt. Cependant, le faible taux de réponse obtenu sur Internet rend l'hypothèse de la sur-représentativité des chômeurs BIT au sein des répondants Internet, forte, même à caractéristiques observables données.

6 Conclusion

L'expérimentation « Module A » a permis de mettre en exergue les risques encourus liés à l'introduction d'un nouveau questionnaire sur l'estimation des principaux indicateurs calculés à partir de l'enquête Emploi. En particulier, elle a révélé qu'une modification des questions, aussi légère soit-elle, pourrait être à l'origine d'une rupture de série, expliquée par l'inversion de l'ordre des questions, la suppression de certaines d'entre elles ou encore des reformulations induites par de nouvelles directives européennes. La refonte du questionnaire est actuellement en cours et chaque modification devra donc faire l'objet d'une attention particulière lors du test pilote envisagé dans le cadre du projet, de façon à anticiper au maximum les ruptures de séries.

Par ailleurs, cette expérimentation tout Internet a permis de révéler, pour la première fois, un effet de mode potentiel non négligeable sur l'estimation de certains des principaux indicateurs et notamment du taux de chômage (environ deux points). Une analyse plus fine de l'effet de mode a permis de mettre en évidence un effet de sélection lié aux caractéristiques socio-démographiques, à l'origine d'une sur-représentation des ménages ayant les revenus les plus élevés, des hommes, des personnes âgées de 30 à 39 ans ainsi que des résidents d'Île-de-France au sein des répondants sur Internet, comparativement à la population répondant habituellement en face-à-face. Au-delà de cette conclusion, les travaux menés ont conduit à envisager, puis à confirmer, l'hypothèse selon laquelle les chômeurs au sens du BIT auraient davantage tendance à répondre à l'enquête Emploi sur Internet qu'ils ne le feraient en face-à-face, soit l'hypothèse d'un effet de sélection directement lié à la variable d'intérêt.

Sous réserve que cette hypothèse soit avérée, l'introduction de la collecte par Internet⁴ dans l'enquête Emploi aurait pour effet bénéfique de traiter un potentiel défaut de couverture des chômeurs BIT dans l'enquête Emploi actuelle et permettrait d'améliorer les estimations réalisées. Partant de ce point de vue, une correction des données collectées sur Internet en amont d'une agrégation avec les données collectées en face-à-face pourrait être à l'origine d'un biais et aboutir à des estimations moins fiables que celles qui auraient été obtenues sans traitement particulier de l'effet de mode. Là encore, l'introduction d'Internet lors du test pilote de 2020 devra donc être l'occasion d'apporter une attention particulière aux effets de mode induits en réinterrogation.

4. Le projet de refonte de l'enquête Emploi prévoit d'introduire la collecte par Internet lors des réinterrogations, la première interrogation restant en face-à-face.

Annexe A - Repérer les individus porteurs de l'effet de mesure en vue d'une possible correction des données collectées sur Internet avant agrégation

Corriger les réponses collectées auprès des internautes ?

Compte tenu de l'ampleur de l'effet de mesure estimé dans ce document, il peut s'avérer pertinent de s'interroger sur l'étape future d'agrégation des données. En présence de données provenant à la fois d'une collecte en face-à-face et du web, est-il préférable, ou non, de corriger les données avant de les agréger ? Si oui, lesquelles et comment ? Rappelons ici, que cette situation reste théorique, la future enquête Emploi n'ayant pas vocation à proposer simultanément une collecte en face-à-face et par Internet en première interrogation, mais plutôt une collecte par téléphone et par Internet en réinterrogation.

Supposons donc que les données collectées sur Internet et en face-à-face doivent permettre de calculer des indicateurs les plus fiables possible et les plus représentatifs de la situation sur le marché du travail.

Deux options se présentent :

- Agréger les données collectées sans opérer de correction sur l'un ou l'autre des échantillons (face-à-face ou Internet). Cela revient à considérer que le biais de mesure est ici limité par le fait que les questionnaires soient extrêmement proches dans un mode et dans l'autre et que les réponses apportées par les populations ayant répondu dans l'un et l'autre des modes soient de qualités comparables.
- Corriger les données provenant de l'un ou l'autre des échantillons en considérant que l'un des modes est le mode de « référence ». Ce mode de collecte peut être un mode historique ayant permis la construction de séries longues. Il peut également s'agir d'un mode considéré comme plus fiable (la qualité des réponses dans l'autre mode étant, quant à elle, supposée moindre).

Ici, le mode de référence du point de vue historique serait le face-à-face. En revanche, compte tenu des analyses précédentes, rien ne permet de prouver qu'il serait bénéfique, du point de vue de la qualité des estimateurs, de corriger les données collectées sur Internet. L'agrégation des données sans correction pourrait même être à l'origine d'une réduction du biais de couverture des chômeurs dans l'enquête Emploi s'il s'avérait que ces derniers soient réellement plus prompts à répondre sur Internet qu'à un enquêteur.

En vue d'affiner l'analyse relative à l'effet de mesure et, en particulier, de comprendre les écarts persistant sur le taux de chômage, il est possible de repérer les individus susceptibles de porter l'effet. Pour cela, une méthode conçue par Stéphane Legleye⁵ [1] et visant à repérer des individus ayant des caractéristiques sociodémographiques communes mais un comportement de réponse différent dans chacun des deux modes est appliquée. Elle permet, lorsque cela s'avère nécessaire, de corriger les réponses apportées par les individus repérés comme porteurs de l'effet de mesure.

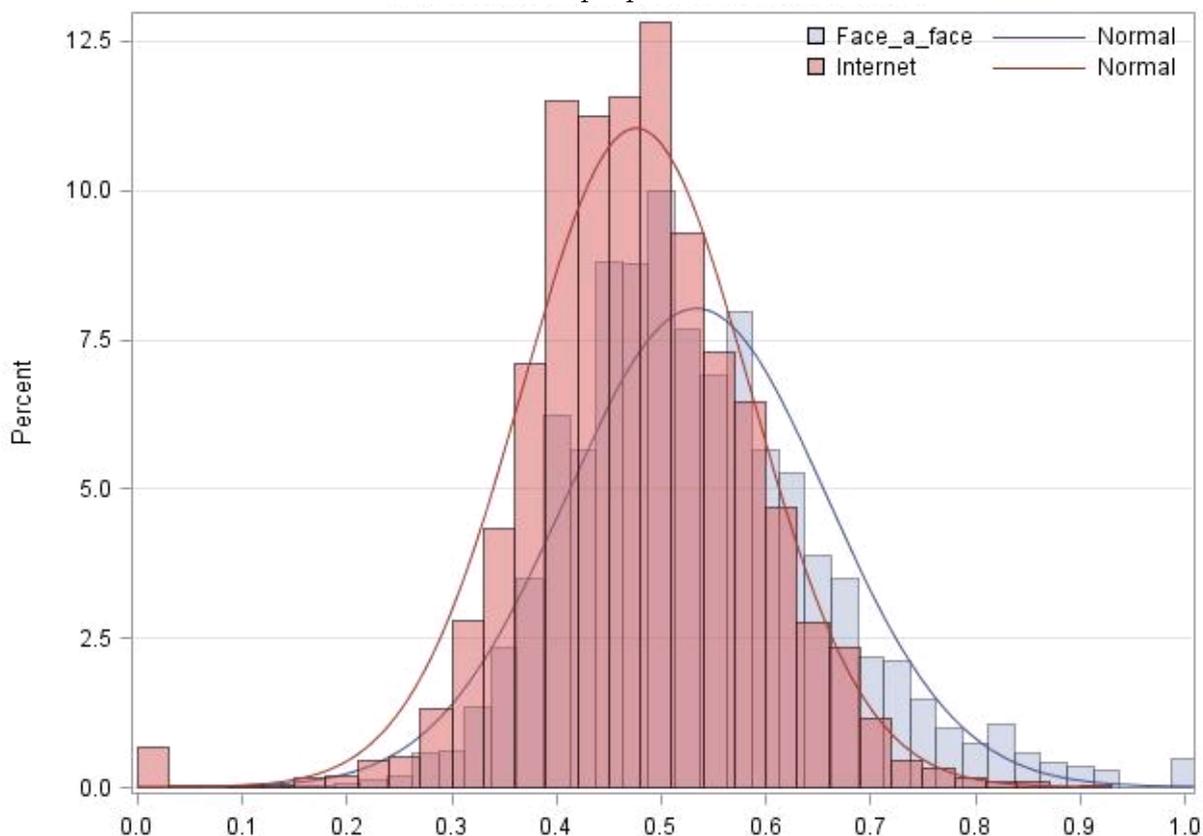
5. Insee, Direction de la méthodologie et de la coordination statistique et internationale, Division Recueil et traitement de l'information.

L'approche « Support commun »

L'approche « Support commun » comprend plusieurs étapes :

- Dans un premier temps, une probabilité de répondre sur Internet plutôt qu'en face-à-face (ou score de propension) est calculée pour chaque individu de l'échantillon total (répondants en face-à-face et sur Internet). La variable correspondant au mode de collecte (face-à-face/Internet) est modélisée par régression logistique à partir des variables socio-démographiques disponibles et de leurs interactions. Le modèle ainsi construit permet de calculer, pour chaque individu, une probabilité prédite de répondre en face-à-face (Figure 38).

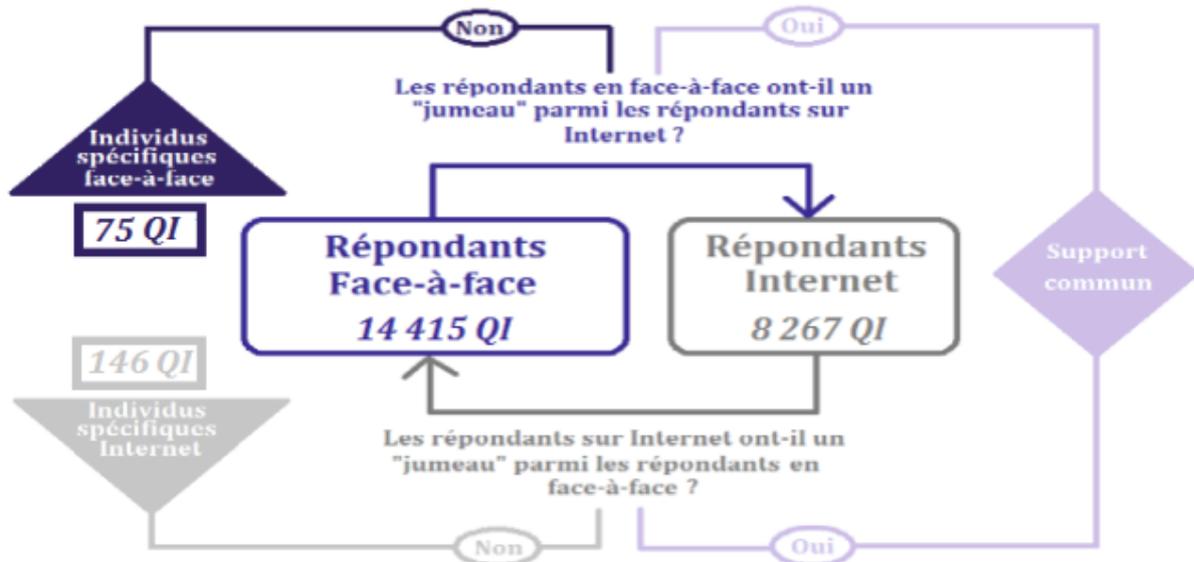
FIGURE 38 – Distribution des scores de propension en fonction du mode de collecte



- Dans un second temps, pour chaque individu répondant sur Internet ayant un score de propension $sp_{Internet}$, il est recherché un individu répondant en face-à-face ayant un score de propension $sp_{Face-a-face}$ tiré aléatoirement dans l'intervalle $[sp_{Internet} - 0,25 \cdot \sigma_{sp_{Internet}}; sp_{Internet} + 0,25 \cdot \sigma_{sp_{Internet}}]$. Un appariement est alors réalisé dès lors qu'un individu a été sélectionné aléatoirement dans l'intervalle défini. Une opération similaire est réalisée pour chaque individu répondant en face-à-face.
- L'ensemble des individus pour lesquels un « jumeau » au sens du score de propension a été trouvé forme « le support commun ». Les individus ayant répondu sur Internet pour lesquels aucun jumeau n'a été trouvé forment un ensemble d'individus « spécifiques de la collecte sur Internet » ; symétriquement, les individus ayant répondu en face-à-face pour lesquels aucun jumeau n'a été trouvé au sein des répondants sur Internet, forment un

ensemble d'individus « spécifiques de la collecte en face-à-face » (Figure 39). En supposant l'effet de sélection corrigé du point de vue des variables socio-démographiques, les individus susceptibles de porter l'effet de mode observé se trouvent potentiellement parmi les répondants sur Internet du support commun ayant un statut d'emploi différent de celui de l'individu ayant répondu en face-à-face auquel ils sont appariés.

FIGURE 39 – Appariement sur score de propension



Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins).

Sources : Données issues de l'exploitation « Module A » ; Enquête Emploi en continu, T2 2016.

FIGURE 40 – Répartition des répondants en fonction de leur statut d'emploi et de leur appartenance au support commun

| | Actif occupé | Chômeur BIT | Inactif | Total |
|--|--------------|----------------|---------|--------|
| Support commun | 12 158 | 1 439 | 8 864 | 22 461 |
| Individus spécifiques Internet | 77 | 8 | 61 | 146 |
| Individus spécifiques Face-à-face | 23 | 12 | 40 | 75 |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins). Données corrigées de la non-réponse (méthode « QE5 »).

Sources : Données issues de l'exploitation « Module A » ; Enquête Emploi en continu, T2 2016.

L'appariement de chacun des individus ayant répondu sur Internet à un individu ayant répondu en face-à-face, tel que défini précédemment, est réalisé pour 8 121 individus⁶. Ces individus appariés, associés à l'ensemble des individus ayant répondu en face-à-face constituent le support commun (8 121 + 14 415 = 22 536 individus). Au sein de ce support commun, certains individus ayant répondu sur Internet ont été appariés à des individus ayant répondu en face-à-face et n'ayant pas le même statut d'emploi. Compte tenu des résultats observés sur les principaux indicateurs, les individus susceptibles de porter l'effet de mesure seraient des répondants sur Internet ayant un statut de chômeur BIT, appariés à des répondants en face-à-face ayant un statut d'inactif. Cette catégorie représente 3 % (soit 243 individus) de l'ensemble du support commun (Figure 41). Bien qu'il ne soit pas possible de le mesurer ou de l'observer par cette méthode, il est envisageable que certains individus n'appartenant pas au support commun (ayant donc des caractéristiques particulières), soient également porteurs d'une partie de l'effet de mesure.

FIGURE 41 – Répartition des individus ayant répondu sur Internet en fonction du statut d'emploi de l'individu apparié

| Type d'appariement | Statut d'emploi de l'individu répondant sur Internet | Statut d'emploi de l'individu apparié répondant en Face-à-face | Effectif non pondéré | Effectif pondéré | Part (en %) |
|--|--|--|----------------------|-------------------|-------------|
| Convergent | Actif occupé | Actif occupé | 2 969 | 19 096 629 | 44 |
| | Chômeur BIT | Chômeur BIT | 62 | 1 007 354 | 2 |
| | Inactif | Inactif | 1 397 | 11 015 450 | 25 |
| Divergent | Actif occupé | Chômeur BIT | 280 | 1 295 887 | 3 |
| | Actif occupé | Inactif | 1 512 | 4 718 771 | 11 |
| | Chômeur BIT | Actif occupé | 364 | 1 290 876 | 3 |
| | Chômeur BIT | Inactif | 243 | 1 130 808 | 3 |
| | Inactif | Actif occupé | 1 158 | 3 639 408 | 8 |
| | Inactif | Chômeur BIT | 136 | 648 177 | 1 |
| Ensemble des individus Internet appartenant au support commun | | | 8 121 | 43 843 360 | 100 |

Champ : Questionnaires individuels validés sur Internet par un individu du support commun dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins). Données corrigées de la non-réponse (méthode « QE5 »).

Sources : Données issues de l'exploitation « Module A » ; Enquête Emploi en continu, T2 2016.

Supposer que l'ensemble de ces individus sont porteurs de l'effet de mesure résiduel (une fois l'effet de sélection lié aux variables socio-démographiques évincé) revient à émettre l'hypothèse selon laquelle l'intégralité d'entre eux ont déclaré un statut d'emploi erroné comparativement à ce qu'ils auraient répondu en face-à-face. Si tel est le cas, une correction de leur statut d'emploi par un statut d'inactif en lieu et place de leur statut de chômeur BIT devrait conduire à annuler l'effet de mesure.

6. Le modèle retenu pour estimer les scores de propension est un modèle logistique dans lequel toutes les interactions bivariées ont été prises en compte. L'ensemble des variables socio-démographiques disponibles ne présentant pas un nombre trop important de valeurs manquantes dans l'un ou l'autre des modes de collecte ont été conservées. Par ailleurs, que la variable cible soit intégrée ou non au modèle, que les poids corrigés de la non-réponse soient pris en compte ou non, que les interactions soient prises en compte ou non, le nombre d'individus considérés comme étant spécifiques à l'un des deux modes reste très faible (au maximum 173 questionnaires individuels Internet hors du support). Quel que soit le modèle choisi, cette méthode ne permet pas d'expliquer l'effet de mode résiduel, compte tenu de l'information auxiliaire disponible.

Or, la correction du statut d'emploi de l'ensemble des individus considérés comme éventuels porteurs de l'effet de mode conduit à une réduction du taux de chômage sur Internet tel, qu'il en devient inférieur au taux de chômage calculé à partir des données collectées en face-à-face dans le cadre de l'enquête Emploi actuelle (Figure 42).

Fort de ce résultat, il peut s'avérer intéressant de tenter de recourir à un mode de correction des statuts d'emploi moins systématique.

FIGURE 42 – Indicateurs du marché du travail après correction sur appariement

| Mode de collecte | | Taux d'activité | Taux d'emploi | Taux de chômage | Part de chômeurs |
|------------------|------------------|-----------------|---------------|-----------------|------------------|
| Face-à-face | | 62,8 | 56,7 | 9,8 | 6,2 |
| Internet | avant correction | 65,1 | 57,0 | 12,4 | 8,1 |
| | après correction | 62,2 | 57,0 | 8,4 | 5,7 |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins). Données corrigées de la non-réponse (méthode « QE5 »).

Sources : Données issues de l'exploitation « Module A » ; Enquête Emploi en continu, T2 2016.

Une méthode de correction avant agrégation des données moins systématique

L'idée consiste à prédire le statut d'emploi BIT des individus enquêtés sur Internet comme chômeurs BIT et appariés à des individus inactifs en face-à-face, à partir de l'information collectée sur les personnes ayant répondu en face-à-face uniquement, qu'elles appartiennent, ou non, au support commun préalablement défini.

La méthode consiste à :

- **Étape 1** : Prédire le fait de répondre sur Internet plutôt qu'en face-à-face à l'aide des variables sociodémographiques disponibles afin d'obtenir des probabilités de réponse par Internet ;
- **Étape 2** : Remplacer le statut d'emploi des chômeurs BIT ayant répondu par Internet, préalablement appariés à des inactifs en face-à-face par des valeurs manquantes ;
- **Étape 3** : Pour chacun de ces individus (chômeurs sur Internet apparié à un inactif en face-à-face), imputer 10 statuts d'emploi BIT prédits à partir du modèle suivant exclusivement construit à partir des données collectées en face-à-face :

$$\text{Statut d'emploi prédit} = \text{Probabilité prédite de répondre sur Internet}$$

Les résultats obtenus, très similaires aux résultats observés préalablement figurent sur la figure 43.

FIGURE 43 – Indicateurs du marché du travail après correction par imputations

| Mode de collecte | | Taux d'activité | Taux d'emploi | Taux de chômage | Part de chômeurs |
|------------------|-------------------|-----------------|---------------|-----------------|------------------|
| Face-à-face | | 62,8 | 56,7 | 9,8 | 6,2 |
| Internet | avant imputations | 65,1 | 57,0 | 12,4 | 8,1 |
| | après imputations | 64,0 | 58,5 | 8,5 | 5,4 |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins). Données corrigées de la non-réponse (méthode « QE5 »).

Sources : Données issues de l'exploitation « Module A » ; Enquête Emploi en continu, T2 2016.

L'effet de mesure est ici corrigé. Deux conclusions découlent de cette analyse : si l'effet de mesure est porté par les individus repérés précédemment, tous n'ont probablement pas déclaré un statut d'emploi différent de celui qu'ils auraient déclaré en face-à-face auprès d'un enquêteur ; l'effet de mode résiduel, une fois écarté l'effet de sélection porté par les variables socio-démographiques disponibles, n'est probablement pas un pur effet de mesure. Il pourrait s'agir, comme le suggéraient certaines des analyses précédentes, au moins partiellement d'un effet de sélection indépendant de l'information auxiliaire disponible, voire directement lié à la variable cible. Cela revient à dire qu'un certain nombre de répondants, ayant un profil particulier, auraient une propension plus forte à répondre sur Internet qu'en face-à-face, dès lors que leur statut serait celui d'un chômeur BIT.

Pour aller plus loin : une méthode de correction plus souple avant agrégation

L'analyse précédente conduit à émettre l'hypothèse selon laquelle seuls quelques individus au sein du support commun pourraient être à l'origine d'un effet de mesure inclus dans l'effet de mode résiduel. Pour tenter de repérer les individus les plus susceptibles de porter l'effet de mesure, il s'avère nécessaire d'affiner l'analyse au-delà d'un seul appariement sur score de propension, en se focalisant davantage sur la population de chômeurs BIT, sur-représentée au sein des répondants sur Internet. Plus précisément, l'idéal serait de déterminer quels individus ont le risque le plus important d'être chômeurs BIT en répondant sur Internet, et inactifs en face-à-face. La méthode envisagée ici pour atteindre cet objectif consiste à construire un modèle prédictif à partir des données collectées en face-à-face, permettant d'estimer pour tout individu non actif occupé ayant participé à l'enquête, sa probabilité d'être inactif plutôt que chômeur BIT. Afin d'améliorer la qualité du modèle, la variable « Inscription à Pôle emploi », liée à la variable cible (chômeur BIT/Inactif) est ajoutée à l'ensemble des variables socio-démographiques utilisées jusqu'ici. Cet ajout implique cependant une restriction du champ aux moins de 65 ans.

Le modèle prédictif une fois établi est appliqué à l'ensemble des individus ayant répondu sur Internet. Chacun se voit alors attribuer, par ce biais, une probabilité prédite d'être inactif plutôt que chômeur BIT. En ordonnant ces probabilités, les individus déclarés chômeurs BIT sur Internet les plus enclins à se déclarer inactifs en face-à-face sont identifiés. Si l'effet résiduel est intégralement considéré comme un effet de mesure et que le face-à-face est considéré comme le mode de collecte de référence, il est alors possible de procéder à la correction du statut d'une partie des individus repérés comme étant les plus enclins à se déclarer inactifs en face-à-face. À titre d'exemple, une correction des individus appartenant au quartile des chômeurs BIT ayant

la plus forte probabilité de se déclarer inactifs en face-à-face, conduit à un taux de chômage de 9,7 % contre 9,8 % en face-à-face (Figure 44).

FIGURE 44 – Indicateurs du marché du travail après correction par quartiles

| Mode de collecte | | Taux d'activité | Taux d'emploi | Taux de chômage | Part de chômeurs |
|--------------------|---|-----------------|---------------|-----------------|------------------|
| Face-à-face | | 62,8 | 56,7 | 9,8 | 6,2 |
| Internet | Sans correction du statut d'emploi | 65,1 | 57,0 | 12,4 | 8,1 |
| | Avec correction du statut d'emploi des individus ayant une probabilité d'être inactif supérieure à | 0,50 | 63,1 | 57,0 | 9,7 |
| | Q1 = 0,87 | 63,5 | 57,0 | 10,3 | 6,5 |
| | Q2 = 0,94 | 64,5 | 57,0 | 11,6 | 7,5 |
| | Q3 = 0,96 | 64,9 | 57,0 | 12,2 | 7,9 |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins). Données corrigées de la non-réponse (méthode « QE5 »).

Sources : Données issues de l'exploitation « Module A » ; Enquête Emploi en continu, T2 2016.

La probabilité prédite pour un chômeur sur Internet d'être inactif en face-à-face est, en moyenne, plus élevée hors du support commun (Figure 46). Cela confirme l'hypothèse selon laquelle les individus n'appartenant pas au support commun ont des caractéristiques particulières, sans qu'il soit possible de déterminer quelles sont celles qui sont à l'origine de cette « spécificité Internet ».

FIGURE 45 – Probabilités prédites d'être inactif en face-à-face selon l'appartenance au support commun

| | Support commun | Hors support commun |
|--------------------------------|----------------|---------------------|
| Effectif | 2733 | 56 |
| Moyenne | 0,78 | 0,82 |
| Minimum | 0,11 | 0,23 |
| 1^{er} quartile | 0,49 | 0,84 |
| Médiane | 0,93 | 0,96 |
| 3ème quartile | 0,96 | 0,99 |
| Maximum | 1 | 1 |

Champ : Questionnaires individuels validés sur Internet dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins), relatifs à des non actifs occupés de moins de 65 ans. Données corrigées de la non-réponse (méthode « QE5 »).

Source : Données issues de l'exploitation « Module A » ; Enquête Emploi en continu, T2 2016

Annexe B - L'expérimentation « de grande envergure », des résultats qui confirment l'analyse

Qu'est-ce que l'expérimentation de grande envergure ?

Deux expérimentations quantitatives ont été menées en 2016 dans le cadre du projet Muse : l'expérimentation dite « de grande envergure » (EGE), comprenant trois interrogations espacées d'un trimestre, débutant en mars, juin et septembre 2016, et l'expérimentation « module A ». L'objectif de l'expérimentation de grande envergure était d'étudier, à grande échelle, la passation du questionnaire de l'enquête Emploi sur Internet et de tester plusieurs variantes, en vue de concevoir le protocole le plus incitatif possible. L'échantillon était constitué de 40 000 logements tirés dans les fichiers consolidés de la taxe d'habitation 2015. Comme l'expérimentation « Module A », l'EGE a été menée exclusivement sur Internet.

Des résultats qui confirment l'effet de mode observé à partir de l'expérimentation « Module A »

Après sélection d'un champ comparable à celui de l'expérimentation « Module A », et application d'un traitement de la non-réponse équivalent à ceux réalisés dans ce document, il est possible de comparer les indicateurs calculés à partir des données collectées lors de la première interrogation de l'EGE, à ceux calculés à partir de l'enquête Emploi en continu en face-à-face.

L'effet de mode observé après correction de la non-réponse (et donc traitement de l'effet de sélection lié aux variables auxiliaires disponibles dans les bases de sondage) est comparable à celui observé sur les données issues de l'expérimentation « Module A », quel que soit l'indicateur considéré, bien que l'ampleur des écarts soit légèrement différente (Figure 46). Ces derniers sont moins importants sur les indicateurs de chômage et d'inactivité, que ce qu'ils n'étaient lorsqu'ils étaient calculés à partir de l'expérimentation « Module A ». À titre d'exemple, le taux de chômage calculé à partir des données collectées en vague 1 dans le cadre de l'EGE est, en moyenne selon la méthode de traitement de la non-réponse considérée, de 11,2 %, contre 9,8 % en face-à-face et 12,2 % à partir des données collectées dans le cadre de l'expérimentation « Module A ». À l'inverse, les écarts sont amplifiés sur les indicateurs d'emploi et d'activité. Ainsi, le taux d'emploi calculé à partir des données de l'EGE est de 58,3 % en moyenne, de 56,7 % en face-à-face et de 57,2 % à partir des données collectées en vague 1 dans le cadre de l'expérimentation « Module A ».

FIGURE 46 – Indicateurs du marché du travail calculés à partir des données collectées dans le cadre de l'EGE

| | | CAH | QE5 | QE10 | HB | Moyenne |
|--------------------------|---|-------------|-------------|-------------|-------------|-------------|
| Taux d'emploi | EEC Face-à-face | 56,7 | 56,7 | 56,7 | 56,7 | 56,7 |
| | EEC EGE Internet | 58,6 | 59,0 | 57,9 | 57,7 | 58,3 |
| | EEC Module A Internet | 57,7 | 57,2 | 57,0 | 57,0 | 57,2 |
| | <i>Ecart Face-à-face/Internet Module A</i> | -1,0 | -0,5 | -0,3 | -0,3 | -0,5 |
| | <i>Ecart Face-à-face/Internet EGE</i> | -1,9 | -2,3 | -1,2 | -1,0 | -1,6 |
| Taux d'activité | EEC Face-à-face | 62,9 | 62,8 | 62,9 | 62,9 | 62,9 |
| | EEC EGE Internet | 65,9 | 66,1 | 65,5 | 65,3 | 65,7 |
| | EEC Module A Internet | 65,4 | 65,1 | 65,1 | 65,0 | 65,2 |
| | <i>Ecart Face-à-face/Internet Module A</i> | -2,5 | -2,3 | -2,2 | -2,1 | -2,3 |
| | <i>Ecart Face-à-face/Internet EGE</i> | -3,0 | -3,3 | -2,6 | -2,4 | -2,8 |
| Taux de chômage | EEC Face-à-face | 9,8 | 9,8 | 9,8 | 9,8 | 9,8 |
| | EEC EGE Internet | 11,0 | 10,6 | 11,6 | 11,6 | 11,2 |
| | EEC Module A Internet | 11,9 | 12,2 | 12,4 | 12,4 | 12,2 |
| | <i>Ecart Face-à-face/Internet Module A</i> | -2,1 | -2,4 | -2,6 | -2,6 | -2,4 |
| | <i>Ecart Face-à-face/Internet EGE</i> | -1,2 | -0,8 | -1,8 | -1,8 | -1,4 |
| Part de chômage | EEC Face-à-face | 6,1 | 6,2 | 6,1 | 6,2 | 6,2 |
| | EEC EGE Internet | 7,3 | 7,0 | 7,6 | 7,6 | 7,4 |
| | EEC Module A Internet | 7,8 | 7,9 | 8,1 | 8,0 | 8,0 |
| | <i>Ecart Face-à-face/Internet Module A</i> | -1,7 | -1,7 | -2,0 | -1,8 | -1,8 |
| | <i>Ecart Face-à-face/Internet EGE</i> | -1,2 | -0,8 | -1,5 | -1,4 | -1,2 |
| Taux d'inactivité | EEC Face-à-face | 37,1 | 37,2 | 37,1 | 37,1 | 37,1 |
| | EEC EGE Internet | 34,1 | 33,9 | 34,5 | 34,7 | 34,3 |
| | EEC Module A Internet | 34,6 | 34,9 | 34,9 | 35,0 | 34,9 |
| | <i>Ecart Face-à-face/Internet Module A</i> | 2,5 | 2,3 | 2,2 | 2,1 | 2,3 |
| | <i>Ecart Face-à-face/Internet EGE</i> | 3,0 | 3,3 | 2,6 | 2,4 | 2,8 |

Champ : Questionnaires individuels validés dans un ménage ayant validé au moins un questionnaire individuel (ménages dont la personne de référence (ou son conjoint le cas échéant) est âgée de 70 ans ou moins). Données corrigées de la non-réponse.

Sources : Données issues de l'EGE et de l'expérimentation « Module A » ; Enquête Emploi en continu, T2 2016.

Les écarts observés, un peu plus marqués dans le cadre de l'expérimentation de grande envergure, confortent l'analyse menée dans ce document. Les différences observées sur les estimations réalisées à partir de l'une ou l'autre des expérimentations étudiées peuvent notamment s'expliquer par une meilleure précision des indicateurs calculés dans le cadre de l'expérimentation « Module A » induite par un plan de sondage tenant compte des taux de réponse observés par profil dans le cadre de l'EGE. En effet, l'expérimentation de grande envergure a permis de mettre en évidence des variables soio-démographiques liées à la non-réponse sur Internet. À partir de ces variables, des strates ont été construites et ont permis de surpondérer les ménages les moins susceptibles de répondre sur Internet dans l'échantillon de l'expérimentation « Module A ».

Références

- [1] S. LEGLEYE, T. RAZAFINDRANOVONA et G. DE PERETTI. *Agréger les échantillons d'une enquête multimode en limitant l'effet de mesure : une proposition d'imputation raisonnable et pragmatique*, Journées de la méthodologie statistique de l'Insee, 2018.
- [2] S. LEGLEYE. *Agréger des données multimodes : état des lieux et propositions de traitement*, Document de travail, Insee, à paraître.
- [3] S. BECK, B. BRENDLER et M. TÔ. *Static on the line : How does the survey mode affect the measurement of ILO indicators ?*, à paraître.
- [4] J. VANNIEUWENHUYZE and G. LOOSVELDT. *Evaluating relative mode effects in mixed-mode survey data using covariate adjustment models*, Journal of Official Statistics, Vol. 30, N° 1, pp.1-21, 2014.
- [5] T. KÖRNER, A. LIERSCH. *Case study on mode effects in the German labour force survey*, Technical report, Destatis, 2014.
- [6] T. RAZAFINDRANOVONA. *Exploitation de l'enquête expérimentale Vols, violence et sécurité*. Document de travail M2016/08, Insee, 2016.
- [7] C. ROBERTS. *Mixing modes of data collection in surveys : A methodological review*. Technical report, ESRC National Center for Research Methods, 2007.
- [8] T. RAZAFINDRANOVONA. *Exploitation de l'enquête expérimentale Logement Internet/papier*. Document de travail M2016/08, Insee, 2016.
- [9] C. BARRET, C. DZIKOWSKI. *Comparaison des modes de collecte par téléphone et par Internet dans le cadre de l'enquête Génération du Cereq. Analyse des abandons en corus et tentative pour distinguer les effets de sélection et de mesure*, Cereq, 2016.
- [10] D. A. DILLMAN, G. PHELP, R. TORTORA, K. SWIFT, J. KOHRELL, J. BERCK, et al.. *Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet*, 2009.
- [11] E. D. LEEUW. *To Mix or Not to Mix Data Collection Modes in Surveys*. Journal of Official Statistics, 2005.
- [12] G. CHARRANCE, C. HAMEL, S. LEGLEYE. *Utiliser des enquêtes ouvertes sur Internet pour améliorer le redressement d'une enquête téléphonique probabiliste ? Le cas de Virage et de la population LGBT*, Ined.
- [13] J. BAULNE, R. COURTEMANCHE. *Comment intégrer le multimode tout en conservant la comparabilité dans une enquête transversale répétée sur la santé ?* Institut de la statistique du Québec.
- [14] M. BOUCHER, E. GAGNON. *Effet du mode de collecte dans l'enquête sur les pratiques de formation en emploi au Québec en 2014*, Institut de la statistique du Québec.

[15] M. WU, F. BRISEBOIS, L. HUNTER. *Mise à l'essai des stratégies de collecte pour les enquêtes en ligne fondées sur l'autodéclaration, Recueil du Symposium 2014 de Statistique Canada, Au-delà des méthodes traditionnelles d'enquêtes : l'adaptation à un monde en évolution.*

[16] G. DE PERETTI, T. RAZAFINDRANOVONA. *Les enquêtes multimode au prisme de l'erreur d'enquête totale, Recueil du Symposium 2014 de Statistique Canada, Au-delà des méthodes traditionnelles d'enquêtes : l'adaptation à un monde en évolution.*

[17] R. BIGOT, P. CROUTTE, F. RECOURS. *Enquêtes en ligne : peut-on extrapoler les comportements et les opinions des internautes à la population générale ? Cahier de recherche N° 273, décembre 2010, Crédoc.*