

Sources et méthodes

Enquête Emploi en continu

L'enquête Emploi est l'une des pièces centrales du dispositif statistique de connaissance de l'emploi et du chômage. L'enquête apporte de nombreuses informations sur l'état du marché du travail et son évolution : le volume de la population active, le nombre de chômeurs et le taux de chômage, les caractéristiques des personnes présentes sur le marché du travail, de celles qui arrivent ou le quittent, etc. Elle fournit aussi des données originales sur la structure par professions, sur l'activité des femmes ou des jeunes, sur la durée du travail et sur les emplois temporaires. Elle contribue ainsi au débat social sur les politiques de l'emploi et de l'éducation et sert de support à de nombreux travaux de recherche dans le domaine économique et social.

L'enquête Emploi est en outre l'unique source d'information qui permet de mettre en œuvre la mesure de l'activité suivant les concepts du Bureau international du Travail (BIT), définis et adoptés lors des conférences internationales des statisticiens du travail (1954 et 1982). Elle s'inscrit dans le cadre des enquêtes sur les forces de travail réalisées dans tous les pays de l'Union européenne (« Labour Force Survey »). La France peut ainsi comparer le niveau et l'évolution de son taux de chômage avec ceux des autres pays européens, mais aussi avec ceux des pays membres de l'organisation internationale du Travail (OIT).

L'enquête Emploi en continu

Objectifs de l'enquête

L'enquête Emploi est un élément central de l'observation structurelle et conjoncturelle du marché du travail et de la situation des individus vis-à-vis de ce marché. Elle est la seule source donnant directement une mesure du chômage au sens du BIT.

L'enquête permet :

- de mesurer directement les grandes catégories d'activité retenues par le BIT (chômage, population active, sous emploi...), ce qui permet notamment de comparer le niveau et l'évolution des taux d'activité et de chômage de la France à ceux des autres pays européens ainsi que ceux des pays membres de l'organisation internationale du travail ;
- de préciser les caractéristiques (durée du travail, temps partiel, multi-activité...) et la structure des emplois ;
- d'étudier des trajectoires individuelles ;
- de constituer une base de données permettant la réalisation d'études approfondies sur les différentes approches de l'emploi.

Les données de l'enquête Emploi sont souvent utilisées par les autres enquêtes de l'Insee ou par les partenaires du système statistique public. La plupart des instituts de sondage et organismes d'études utilisent pour leurs travaux des données de cadrage issues de l'enquête Emploi.

Principales caractéristiques de l'enquête actuelle

L'enquête Emploi est produite selon un calendrier trimestriel et sa collecte se déroule en continu tout au long de chaque trimestre. L'administration de l'enquête « en continu » (c'est-à-dire tout au long de l'année) permet notamment la mesure des évolutions conjoncturelles de l'emploi et du chômage grâce à des moyennes trimestrielles.

Dans cette enquête, ce sont les logements qui sont tirés au sort. Une fois qu'un logement a été tiré, ses occupants seront enquêtés. Et ce, pendant six trimestres consécutifs. Ainsi, puisque les enquêtés sont interrogés jusqu'à six fois, pendant six trimestres consécutifs, l'enquête peut fournir des éléments sur les parcours professionnels individuels à moyen terme (à condition de tenir compte du fait que les personnes suivies six trimestres ne sont pas représentatives de l'ensemble de la population).

A partir du 2^{ème} trimestre 2010, l'échantillon trimestriel de l'enquête est constitué d'environ 67 000 résidences principales, correspondant après collecte à environ 108 000 personnes répondantes.

Une enquête qui a évolué

Depuis sa création en 1954, l'enquête Emploi a connu plusieurs mues importantes. L'outil a dû évoluer et s'adapter aux exigences et aux moyens techniques de chaque époque. Nous donnons ici quelques jalons particulièrement notables de cette longue histoire.

En 2007, suite au débat sur les statistiques du chômage et au rapport de l'IGF-Igas¹, l'Insee a décidé d'augmenter de 50% la taille de l'échantillon. Cette augmentation a été progressive et initiée début 2009. Elle précède le renouvellement total de l'échantillon initial de l'enquête à partir du 3ème trimestre 2010. Par ailleurs, une refonte est programmée pour 2013, elle porte sur une rénovation du dispositif informatique encadrant l'enquête et du questionnaire lui-même.

L'enquête Emploi ne s'est pas toujours déroulée tout au long de l'année, comme actuellement : jusqu'en 2002, elle avait lieu une fois par an, généralement en mars. Outre ce changement de périodicité, l'enquête Emploi a connu d'autres évolutions notables à ce moment de son histoire. Le questionnaire a été modifié, à la fois dans un souci d'harmonisation européenne (pour une meilleure comparabilité avec les pays voisins) et dans le sens d'un enrichissement de la connaissance relative à l'emploi, à la durée annuelle du travail, à la pluriactivité ou encore à la formation continue.

Le questionnaire remanié s'est en particulier attaché à fournir une mesure harmonisée au niveau européen des concepts d'emploi, de chômage et d'inactivité, respectant ainsi l'approche européenne privilégiant les mesures les plus factuelles et objectives possibles. Le questionnaire offre également des informations plus riches que le précédent sur le sous-emploi, sur la durée du travail et permet de mieux cerner la frontière entre emploi, chômage et inactivité.

1950	Première enquête Emploi
1960, 1968, 1975, 1982, 1990	Refontes de l'enquête Emploi
1992 - 1995	Expérimentation d'une enquête Emploi trimestrielle
1998	Début du dispositif léger, visant à anticiper les problèmes organisationnels et à mettre au point les méthodes d'exploitation d'une enquête en continu
2001	Juillet : début de la collecte de la nouvelle enquête Emploi
2002	Mars : dernière enquête annuelle
2007	Novembre : changements de définition de l'activité au sens du BIT, du champ de l'enquête et de la méthode de pondération appliqués sur l'ensemble des enquêtes depuis 2003 ; Décembre : publication trimestrielle des taux d'emploi et de chômage dans l'Informations Rapides
2009	janvier : début de la phase d'augmentation de 50% de la taille de l'échantillon
2010	Juillet : début de la phase de remplacement de l'échantillon initial
2011	Octobre : A partir de cette date, l'échantillon est intégralement issu de la Taxe d'Habitation
2013	Révision de l'application informatique et du questionnaire

¹ « Rapport sur les statistiques d'estimation du chômage », Inspection Générale des Finances, Inspection Générale des Affaires Sociales.

Les concepts d'activité, de chômage et de sous-emploi

La population active regroupe les personnes ayant un emploi (ou population active occupée) et les chômeurs. Les personnes n'appartenant pas à la population active sont les inactifs, c'est-à-dire notamment les retraités, les étudiants, les femmes au foyer. Dans les publications, les concepts retenus sont ceux du BIT (bureau international du travail), définis lors de la XIII^{ème} conférence internationale des statisticiens du travail d'octobre 1982 (nouvelle version du texte adopté lors de la VII^{ème} conférence de 1954). Pour mesurer les effectifs de ces populations, l'Insee interprète les recommandations émises dans le texte et les traduit en questions dans l'enquête Emploi. Les définitions du BIT n'ont pas changé depuis 1982, mais le questionnaire de l'enquête a été modifié en 2003 afin que les statistiques produites soient comparables avec celles des autres pays européens. Le nouveau questionnaire est ainsi conforme au règlement européen n° 1897/2000 sur la définition européenne du chômage. Il diffère du questionnaire de l'enquête annuelle qui débutait par une question invitant les enquêtés à se classer spontanément dans les catégories « en emploi », « au chômage », « inactif ». Cette question est déplacée plus loin dans le nouveau questionnaire, dans le calendrier mensuel rétrospectif d'activité posé à chaque interrogation de l'enquête emploi. Malgré cette harmonisation des questionnaires européens, la France a conservé une interprétation spécifique des définitions BIT.

En 2007, l'Insee a finalement entièrement adopté l'interprétation communautaire, celle de l'office statistique des communautés européennes (Eurostat). Ce rapprochement a été opéré de façon à harmoniser au niveau européen les statistiques sur le marché européen, ce qui facilite les comparaisons internationales. La méthode présentée ici et utilisée dans les publications de l'Insee depuis cette date correspond à cette nouvelle interprétation.

La population active occupée

L'interprétation européenne de la définition BIT de la population active occupée ou « personnes pourvues d'un emploi », fournie dans le paragraphe 9 de la résolution BIT (cf. encadré 1), est la suivante :

- les personnes âgées d'au moins 15 ans et qui ont effectué au moins une heure de travail rémunéré durant la semaine de référence (semaine du lundi au dimanche, dont les dates sont explicitement mentionnées lors de l'interrogation et qui précède de peu la date d'interrogation, de 1 à 16 jours dans la plupart des cas) d'une part ;
- et les personnes âgées d'au moins 75 ans et qui déclarent occuper un emploi actuellement d'autre part, constituent l'essentiel de la population active occupée.

Toutefois, parmi les personnes de 15 ans et plus n'ayant pas effectué au moins une heure de travail rémunéré la semaine de référence, sont incluses dans la population active occupée les personnes :

- qui étaient en congé rémunéré, en congé de maternité ou de paternité, à temps partiel, en formation rémunérée par l'employeur ou dans le cadre d'un contrat en alternance, au chômage partiel (chômage technique), en grève ;
- ou qui étaient en congé non rémunéré par l'employeur (congé parental, congé sans solde, congé individuel de formation) d'une durée (prévue) inférieure ou égale à 91 jours ;

- ou qui étaient en congé maladie ou accident du travail d'une durée inférieure ou égale à 365 jours ;
- ou qui aidaient un membre de la famille (sans être salarié), ou qui travaillaient pour l'entreprise, l'exploitation ou le cabinet familial (aides familiaux).

Ainsi, selon cette définition, une personne ayant pris sa retraite, mais effectuant quelques heures de travail, est classée parmi les actifs occupés ; de même un étudiant qui exerce une activité occasionnelle (job d'été par exemple) sera également classé parmi les actifs occupés s'il est interrogé pendant cette période.

Le chômage

Selon le BIT, les « chômeurs » (cf. paragraphe 10 de l'encadré 1) sont des personnes :

- (a) « sans travail » ;
- (b) « disponibles pour travailler » ;
- et (c) « à la recherche d'un travail ».

Le critère (a) ne pose pas de problème d'interprétation : « sans travail » signifie « n'est pas actif occupé » au sens de la définition exposée au paragraphe précédent. Sont donc exclues de la population des chômeurs les personnes ayant effectué ne fût-ce qu'une heure de travail effectif au cours de la semaine de référence, même s'il s'agissait d'un travail purement occasionnel.

La « disponibilité pour travailler » est comprise dans le sens de « disponibilité pour commencer une activité en tant que salarié ou non-salarié dans un délai de deux semaines suivant la semaine de référence ». Les personnes qui ne sont pas disponibles pour cause de maladie temporaire (moins d'un mois) sont considérées comme disponibles.

Le critère de « recherche d'un travail », lui, est observé sur la période de 4 semaines se terminant à la fin de la semaine de référence. Une personne est considérée comme « à la recherche d'un travail » si, au cours de cette période, elle a entrepris au moins une démarche spécifique pour en trouver.

Sont considérées comme des démarches spécifiques de recherche de travail² :

- rechercher un terrain, des locaux, du matériel, des clients ou faire une étude de marché ;
- rechercher des ressources financières ;
- demander un permis, une licence ou une autorisation pour s'installer ;
- prendre des conseils auprès d'un organisme public pour trouver un emploi ;
- être contacté par Pôle Emploi, l'Apéc ou un autre organisme public ;
- avoir un contact avec une agence d'intérim ;

² Depuis 2007, le simple renouvellement de l'inscription à Pôle Emploi n'est plus considéré comme une démarche spécifique de recherche d'emploi. Il l'est s'il s'accompagne par une des démarches ci-dessus.

- effectuer des démarches directes auprès d'employeurs ;
- se rendre à un salon professionnel ;
- s'adresser à des relations personnelles ou professionnelles pour s'installer ;
- passer un test, un entretien, un concours de recrutement, un concours administratif ;
- passer ou afficher une annonce d'offre d'emploi ;
- répondre à une offre d'emploi ;
- étudier les annonces d'offres d'emploi ;
- rechercher une entreprise, une exploitation ou un cabinet à racheter ou à prendre en location-gérance ;
- effectuer des démarches autres que celles ci-dessus.

En plus de cette catégorie de personnes, la population des chômeurs comprend également les personnes qui sont sans travail (a), disponibles pour travailler (b), et qui ont trouvé un emploi qui commencera plus tard, ou pour lesquelles la période considérée est une morte saison (saisonniers) ou précède le début d'un emploi. Seules celles dont l'emploi ultérieur débute dans un délai inférieur à trois mois sont considérées comme des chômeurs.

Le sous-emploi

A lui seul, l'indicateur du chômage ou du taux de chômage est un indicateur frustré de la fluidité du marché du travail. Des indicateurs complémentaires fournissent d'autres éléments de diagnostic. Le sous-emploi est l'un de ces indicateurs. Les personnes en sous-emploi sont celles qui ont un emploi, mais qui souhaiteraient travailler davantage. Une définition précise de ce concept a été établie lors de la XVI^{ème} conférence internationale des statisticiens du travail d'octobre 1998 (cf. encadré 2).

La population en sous-emploi est constituée de deux groupes de personnes :

- celles qui travaillent à temps partiel (moins de 35 heures), qui souhaitent travailler plus, et qui recherchent un emploi et/ou qui sont disponibles pour travailler plus ;
- celles qui ont involontairement travaillé moins que d'habitude, pour cause de chômage partiel ou d'intempéries par exemple, qu'elles travaillent à temps plein ou à temps partiel.

ENCADRÉ 1

**EXTRAIT DE LA RÉOLUTION CONCERNANT LES STATISTIQUES
DE LA POPULATION ACTIVE, DE L'EMPLOI, DU CHÔMAGE ET DU SOUS EMPLOI**
adoptée par la conférence internationale des statisticiens du travail
(Genève, octobre 1982).

La main-d'œuvre (la population active du moment)

8. La « main-d'œuvre » ou « population active du moment » comprend toutes les personnes qui remplissent les conditions requises pour être incluses parmi les personnes pourvues d'un emploi ou les chômeurs, comme défini aux paragraphes 9 et 10 ci-après.

Emploi

9.1. Les "personnes pourvues d'un emploi" comprennent toutes les personnes ayant dépassé un âge spécifié qui se trouvaient, durant une brève période de référence spécifiée telle qu'une semaine ou un jour, dans les catégories suivantes :

a) « emploi salarié »

a1) « personnes au travail » : personnes qui, durant la période de référence, ont effectué un travail moyennant un salaire ou un traitement en espèces ou en nature ;

a2) « personnes qui ont un emploi mais ne sont pas au travail » : personnes qui, ayant déjà travaillé dans leur emploi actuel, en étaient absentes durant la période de référence et avaient un lien formel avec leur emploi :

i) le service ininterrompu du salaire ou du traitement ;

ii) une assurance de retour au travail à la fin de la situation d'exception ou un accord sur la date de retour ;

iii) la durée de l'absence du travail qui, le cas échéant, peut être la durée pendant laquelle les travailleurs peuvent recevoir une indemnisation sans obligation d'accepter d'autres emplois qui leur seraient éventuellement proposés.

b) « emploi non salarié »

b1) « personnes au travail » : personnes qui, durant la période de référence, ont effectué un travail en vue d'un bénéfice ou d'un gain familial, en espèces ou en nature ;

b2) « personnes ayant une entreprise mais n'étant pas au travail » : personnes qui, durant la période de référence, avaient une entreprise industrielle, un commerce, une exploitation agricole ou une entreprise de prestations de services, mais n'étaient temporairement pas au travail pour toute raison spécifique.

9.2. Dans la pratique, on peut interpréter la notion de « travail effectué au cours de la période de référence » comme étant un travail d'une durée d'une heure au moins.

9.3. Les personnes temporairement absentes de leur travail pour raison de maladie ou d'accident, de congé ou de vacances, de conflit du travail ou de grève, de congé-éducation ou formation, de congé-maternité ou parental, de mauvaise conjoncture économique ou de suspension temporaire du travail due à des causes telles que : conditions météorologiques défavorables, incidents mécaniques ou électriques, pénurie de matières premières ou de combustibles, ou toute autre cause d'absence temporaire avec ou sans autorisation, devraient être considérées comme pourvues d'un emploi salarié, à condition qu'elles aient un lien formel avec leur emploi.

9.4. Les employeurs, les personnes travaillant à leur propre compte et les membres des coopératives de producteurs devraient être considérés comme travailleurs non salariés et classés comme « étant au travail » ou « n'étant pas au travail », selon les cas.

9.5. Les travailleurs familiaux non rémunérés devraient être considérés comme travailleurs non salariés indépendamment du nombre d'heures de travail effectué durant la période de référence. Les pays qui, pour des raisons particulières, préféreraient choisir comme critère une durée minimale de temps de travail pour inclure les travailleurs familiaux non rémunérés parmi les personnes pourvues d'un emploi devraient identifier et classer séparément les personnes de cette catégorie qui ont travaillé moins que le temps prescrit

9.6. Les personnes engagées dans la production de biens et services pour leur propre consommation ou celle du ménage devraient être considérées comme travailleurs non salariés si une telle production apporte une importance à la consommation totale du ménage.

9.7. Les apprentis qui ont reçu une rétribution en espèces ou en nature devraient être considérés comme personnes pourvues d'un emploi salarié et classés comme « étant au travail » ou « n'étant pas au travail » sur la même base que les autres catégories de personnes pourvues d'un emploi salarié.

9.8. Les étudiants, les personnes s'occupant du foyer et autres personnes principalement engagées dans des activités non économiques durant la période de référence et qui étaient en même temps pourvues d'un emploi salarié ou non salarié comme défini au sous-paragraphe 1 ci-dessus devraient être considérés comme ayant un emploi, sur la même base que les autres catégories de personnes ayant un emploi, et être identifiés séparément lorsque cela est possible.

9.9. Les membres des forces armées devraient être inclus parmi les personnes pourvues d'un emploi salarié. Les forces armées devraient comprendre aussi bien les membres permanents que les membres temporaires, comme spécifié dans la récente révision de la Classification internationale type des professions (CITP).

Chômage

10.1. Les « chômeurs » comprennent toutes les personnes ayant dépassé un âge spécifié qui, au cours de la période de référence, étaient :

a) « sans travail », c'est-à-dire qui n'étaient pourvues ni d'un emploi salarié ni d'un emploi non salarié ;

b) « disponibles pour travailler » dans un emploi salarié ou non salarié durant la période de référence ;

c) « à la recherche d'un travail », c'est-à-dire qui avaient pris des dispositions spécifiques au cours d'une période récente spécifiée pour chercher un emploi salarié ou un emploi non salarié. Ces dispositions spécifiques peuvent inclure : l'inscription à un bureau de placement public ou privé ; la candidature auprès d'employeurs ; les démarches sur les lieux de travail, dans les fermes ou à la porte des usines, sur les marchés ou dans les autres endroits où sont traditionnellement recrutés les travailleurs ; l'insertion ou la réponse à des annonces dans les journaux ; les recherches par relations personnelles ; la recherche de terrain, d'immeubles, de machines ou d'équipement pour créer une entreprise personnelle ; les démarches pour obtenir des ressources financières, des permis et licences, etc.

10.2. Dans les situations où les moyens conventionnels de recherche de travail sont peu appropriés, où le marché du travail est largement inorganisé ou d'une portée limitée, où l'absorption de l'offre de travail est, au moment considéré, insuffisante, où la proportion de main-d'œuvre non salariée est importante, la définition standard du chômage donnée au sous-paragraphe 1 ci-dessus peut être appliquée en renonçant au critère de la recherche de travail.

10.3. Pour appliquer le critère de la disponibilité pour le travail, spécialement dans les situations couvertes par le sous-paragraphe 2 ci-dessus, des méthodes appropriées devraient être mises au point pour tenir compte des circonstances nationales. De telles méthodes pourraient être fondées sur des notions comme l'actuelle envie de travailler et le fait d'avoir déjà travaillé, la volonté de prendre un emploi salarié sur la base des conditions locales ou le désir d'entreprendre une activité indépendante si les ressources et les facilités nécessaires sont accordées.

10.4. En dépit du critère de recherche de travail incorporé dans la définition standard du chômage, les personnes sans travail et disponibles pour travailler, qui ont pris des dispositions pour prendre un emploi salarié ou pour entreprendre une activité indépendante à une date ultérieure à la période de référence, devraient être considérées comme chômeurs.

10.5. Les personnes temporairement absentes de leur travail sans lien formel avec leur emploi, qui étaient disponibles pour travailler et à la recherche d'un travail, devraient être considérées comme chômeurs conformément à la définition standard du chômage. Les pays peuvent, cependant, en fonction des situations et politiques nationales, préférer renoncer au critère de la recherche d'un travail dans le cas des personnes temporairement mises à pied. Dans de tels cas, les personnes temporairement mises à pied qui n'étaient pas à la recherche d'un travail mais qui étaient néanmoins classées comme chômeurs devraient être identifiées et former une sous catégorie à part.

10.6. Les étudiants, les personnes s'occupant du foyer et les autres personnes principalement engagées dans des activités non économiques durant la période de référence et qui satisfont aux critères exposés aux sous-paragraphes 1 et 2 ci-dessus devraient être considérés comme chômeurs au même titre que les autres catégories de chômeurs et être identifiés séparément lorsque cela est possible.

ENCADRE 2

EXTRAIT DE LA RÉSOLUTION CONCERNANT LES STATISTIQUES DE LA MESURE DU SOUS-EMPLOI ET DES SITUATIONS D'EMPLOI INADÉQUAT

adoptée par la conférence internationale des statisticiens du travail

(Genève, octobre 1998).

Mesures du sous-emploi lié à la durée du travail

7. Le sous-emploi lié à la durée du travail existe quand la durée du travail d'une personne employée est insuffisante par rapport à une autre situation d'emploi possible que cette personne est disposée à occuper et disponible pour le faire.

8.1) Les personnes en sous-emploi lié à la durée du travail comprennent toutes les personnes pourvues d'un emploi -telles qu'elles sont définies dans les directives internationales en vigueur concernant les statistiques de l'emploi- qui répondent aux trois critères suivants pendant la période de référence utilisée pour définir l'emploi :

a) « disposées à faire davantage d'heures », c'est-à-dire souhaitant prendre un autre (ou plusieurs autres) emploi(s) en plus de leur(s) emploi(s) actuel(s) afin d'effectuer davantage d'heures de travail ; de remplacer tel ou tel de leurs emplois actuels par un autre (ou plusieurs autres) emploi(s) assorti(s) d'une durée de travail supérieure ; d'effectuer davantage d'heures de travail dans tel ou tel de leurs emplois actuels ; ou une combinaison de ces différents éléments. Dans la perspective de montrer comment la « disposition à effectuer plus d'heures de travail » est significative en termes d'action selon les circonstances nationales, il doit y avoir distinction entre ceux qui ont activement cherché à travailler plus et les autres. La recherche active d'heures de travail complémentaires doit être définie selon les critères utilisés dans la définition de recherche d'un emploi tels qu'utilisés dans la définition de la population active, tout en tenant compte également des activités nécessaires en vue d'augmenter de nombre d'heures de travail dans l'emploi occupé ;

b) « disponibles pour faire davantage d'heures » c'est-à-dire prêtes, pendant une période ultérieure spécifiée, à faire davantage d'heures, si la possibilité leur en était offerte. La période ultérieure à spécifier lorsque l'on détermine la disponibilité des travailleurs pour faire davantage d'heures devrait être choisie en fonction des circonstances nationales et inclure la période dont ont également besoin les travailleurs pour quitter un emploi et en commencer un autre ;

c) « ayant travaillé moins qu'un seuil relatif à la durée du travail », c'est-à-dire les personnes dont les heures de travail « réellement effectuées » dans tous les emplois confondus pendant la période de référence, telles que définies dans les directives internationales en vigueur concernant les statistiques du temps de travail, étaient inférieures à un seuil à choisir parmi les circonstances nationales. Ce seuil pourrait être défini, par exemple, par rapport à la distinction entre emploi à temps plein et emploi à temps partiel, aux valeurs médianes, moyennes, ou aux normes relatives aux heures de travail telles que spécifiées par la législation pertinente, les conventions collectives, les accords d'aménagement du temps de travail, ou les habitudes de travail selon les pays.

8. 2) Afin de donner une certaine souplesse analytique à l'élaboration et à l'évaluation des politiques, ainsi qu'à des fins de comparabilité nationale, les pays devraient s'efforcer de recenser tous les travailleurs qui, pendant la période de référence, étaient disposés et disponibles pour faire davantage d'heures, quel que soit le nombre d'heures qu'ils ont réellement effectuées pendant la période de référence.

La méthodologie de l'enquête

Le champ de l'enquête : personnes interrogées et représentativité

Les personnes sont enquêtées dans leur résidence principale, c'est-à-dire la résidence où elles vivent habituellement.

Les personnes interrogées dans l'enquête sont celles vivant dans des logements dits « ordinaires » : ainsi, les populations vivant en communauté ne font pas partie du champ de l'enquête. Il s'agit des personnels des établissements hospitaliers, scolaires et hôteliers vivant en collectivité et des membres des communautés religieuses. La population des habitations mobiles et les marinières ne sont pas non plus interrogés. Échappant à cette règle, les élèves internes sont intégrées au champ de l'enquête.

En outre, les élèves ou étudiants habitant dans un logement indépendant (hors cité universitaire ou foyer d'étudiants) font partie du champ de l'enquête : ils sont interrogés dans leur logement s'ils le considèrent comme leur logement principal, ou dans le logement de leurs parents s'ils y retournent fréquemment notamment pendant les vacances scolaires. De même, une personne ayant un logement occasionnel mais retournant le week-end dans son foyer principal est interrogée dans ce dernier.

Toutes les personnes des logements appartenant au champ de l'enquête sont enregistrées lors du passage de l'enquêteur, mais seules les personnes âgées d'au moins 15 ans révolus dans l'année civile sont effectivement interrogées.

Le champ géographique de l'enquête est la France métropolitaine. Une enquête Emploi indépendante a également lieu dans les départements d'outre mer (Dom). Elle fait l'objet de publications indépendantes.

L'échantillon de l'enquête

Les caractéristiques de l'échantillon

L'échantillon de l'enquête Emploi est constitué de logements. Pour bien mesurer les évolutions trimestrielles d'emploi et de chômage, une part importante de l'échantillon est conservée d'un trimestre sur l'autre : seul 1/6ème de l'échantillon est renouvelé chaque trimestre. Ainsi, chaque logement de l'échantillon est interrogé pendant 6 trimestres consécutifs³.

L'échantillon trimestriel de l'enquête est constitué d'environ :

- 67 000 logements identifiés comme résidences principales ;

³ Il est important de noter ici que ce sont bien les logements qui constituent les unités enquêtées : lorsqu'un ménage quitte son logement au cours des 6 trimestres pendant lesquels se déroule l'enquête, c'est le ménage qui s'installera à sa suite dans le logement qui continuera d'être enquêté jusqu'au terme de cette période. Les ménages qui quittent leur logement ne sont donc pas suivis dans leur nouvelle destination.

- 57 000 logements répondants (c'est-à-dire comptant au moins une personne répondante⁴) ;
- 108 000 personnes de 15 ans et plus répondantes.

Sur une année, et compte tenu des renouvellements partiels, l'échantillon compte environ :

- 268 000 logements (résidences principales), dont 100 500 distincts ;
- 228 000 logements répondants, dont 85 000 distincts ;
- 432 000 personnes de 15 ans et plus répondantes, dont 162 000 distinctes.

Une base de sondage élaborée à partir des fichiers de la taxe d'habitation

A partir du 3^{ème} trimestre 2010, la base de sondage dans laquelle sont tirés les nouveaux logements entrants dans l'échantillon de l'enquête Emploi est construite à partir des informations disponibles dans les fichiers de la taxe d'habitation 2006. Cette base de sondage⁵ est mise à jour chaque année avec les données de la taxe d'habitation les plus récentes. En effet, la disponibilité d'un identifiant pérenne des logements autorise, par appariement des fichiers de la taxe d'habitation relatifs à différentes années, l'identification des logements nouveaux et leur suivi par un échantillonnage adapté.

La qualité et la fraîcheur des informations de repérage (nom, adresses foncière et postale,...) apportées par la base de sondage sont autant d'atouts pour la collecte. Un atout supplémentaire de cette base est qu'elle fournit les références cadastrales des 32 millions de logements concernés par la taxe d'habitation. Ces références permettent de localiser précisément les logements et de les regrouper selon leur proximité géographique. En effet, depuis longtemps, l'enquête Emploi se caractérise par une stratégie d'échantillonnage qualifiée d'« aréolaire » : l'échantillon de cette enquête n'est pas constitué d'un tirage de logements, mais d'un tirage de groupes de logements contigus.

Les références cadastrales permettent aisément de procéder à ces regroupements. Elles sont constituées par la section cadastrale (un quartier) et, à l'intérieur de la section, de la parcelle cadastrale (plus petite unité de propriété foncière). Sections et parcelles cadastrales sont codées de telle sorte qu'elles permettent de constituer par une procédure automatique des aires géographiques, appelées « grappes » et couvrant une vingtaine de logements principaux.

La procédure de construction des grappes au sein de chaque section cadastrale peut se résumer ainsi :

- A partir du nombre de résidences principales de la section, on détermine le nombre de grappes à créer, et leur taille, qui doit être la plus proche possible de 20 résidences principales ;

⁴ Parce qu'ils ne constituent pas le cœur de l'enquête et que l'Insee ne souhaite les solliciter qu'à minima, les ménages d'inactifs de 65 ans ou plus ne sont enquêtés qu'en 1^{ère} et dernière interrogations. Ce sont les réponses données lors de la 1^{ère} interrogation qui alimentent les interrogations intermédiaires pour lesquelles ces ménages sont donc considérés comme répondants.

⁵ La méthode de construction de l'échantillon est détaillée dans « La construction du nouvel échantillon de l'enquête emploi en continu à partir des fichiers de la taxe d'habitation », Vincent Loonis, Les actes des journées de méthodologie statistique, 2009.

- On trie les logements par numéro de parcelle ;
- On lit séquentiellement le fichier en comptant les résidences principales, jusqu'à en obtenir le nombre déterminé précédemment : tous les logements lus, principaux ou non principaux, constituent une grappe.

Deux parcelles contigües, dont l'une a subi une modification, peuvent ne pas avoir des numéros de parcelles consécutifs, et donc se retrouver dans deux grappes différentes. On corrige donc, dans la mesure du possible, le numéro de parcelle à partir du nom de la voie et du numéro de façon à recréer de la contiguïté.

Dès le départ, le choix d'un échantillon aréolaire a été dicté par les raisons suivantes :

- la plus grande concentration géographique des interviews permet un gain de temps et de coût pour la collecte, condition nécessaire à la réalisation d'une enquête de grande ampleur ;
- l'obtention d'un meilleur taux de réponse que dans les enquêtes où les logements ne sont pas tirés par aires géographiques. En effet, on observe en général un effet d'entraînement positif entre les ménages d'une même aire. L'observation empirique montre que les effets d'entraînement négatifs sont beaucoup plus rares.

En contrepartie, l'échantillon aréolaire présente un inconvénient principal : à taille identique, les résultats obtenus auprès d'un échantillon de grappes de logements sont moins précis que ceux provenant d'un échantillon de logements dispersés. En effet, dans une même grappe (souvent un même quartier), les ménages ont souvent des caractéristiques socio-économiques proches. C'est ce que l'on appelle l'« effet de grappe ». Cet effet est particulièrement important pour les répartitions par nationalité, par activité économique et relativement important pour certaines catégories socioprofessionnelles. C'est pour minimiser cet effet que des grappes de taille réduite ont été constituées.

Le tirage de l'échantillon

Après avoir été enquêtée pendant 6 trimestres consécutifs, une grappe (dite « sortante ») doit être remplacée par une nouvelle grappe (« entrante »). Si aucune contrainte n'est donnée, il n'y a pas de raison que cette grappe entrante soit proche géographiquement de la grappe sortante. Ce fonctionnement pourrait avoir des conséquences négatives sur l'organisation du travail des enquêteurs. C'est pour assurer la proximité géographique des grappes entrantes et sortantes qu'a été introduite la notion de « secteur ». Un secteur est un regroupement de grappes proches. Quand une grappe est sortante, on la remplace par une grappe du même secteur.

- La taille du secteur, en nombre de grappes, dépend de la durée de vie souhaitée de l'échantillon. Celle-ci a été fixée à 9 ans, soit 36 trimestres. Chaque grappe étant enquêtée 6 trimestres, il faut que les secteurs aient 6 grappes.

Pour parvenir à un échantillon de 67 000 logements par trimestre, la taille moyenne d'une grappe étant légèrement supérieure à 20 résidences principales, le nombre de secteurs a été fixé à 3 200.

Le nombre de secteurs à tirer dans chaque région a été déterminé sur la base d'une répartition proportionnelle de l'échantillon par région. Pour tenir compte des attentes d'Eurostat en matière de

précision régionale, cette répartition a été adaptée de façon à ce que pour chaque région, le nombre de secteurs à tirer soit au moins égal au nombre de secteurs présents dans l'ancien échantillon⁶.

Le tirage des secteurs par région s'est fait de manière équilibrée, proportionnellement au nombre de résidences principales. Les variables d'équilibrage retenues sont, par ordre décroissant d'importance :

- La répartition des résidences principales par type d'espace selon le Zonage en Aires Urbaines : pôles urbains, couronnes périurbaines, communes multi polarisées, communes rurales ;
- La répartition des résidences principales par quintile de revenus ;
- Le nombre de locataires ;
- La répartition selon la date d'achèvement, essentiellement pour les logements récents ;
- Le nombre de logements sociaux ;
- Le nombre de logements collectifs ;
- Le nombre de résidences principales dont le chef de ménage a plus de 55 ans ;
- Le nombre de résidences principales ;
- La probabilité d'inclusion pour assurer un sondage de taille fixe.

Dans chaque secteur, les 6 grappes ont été numérotées de manière aléatoire. Le numéro attribué constitue le numéro de série. Il détermine le rang d'introduction de la grappe dans l'échantillon : les grappes de la première série seront interrogées en 1er, etc. ...

Sélection des logements non principaux

Par construction, tous les logements, y compris non principaux, sont automatiquement rattachés à une grappe. L'échantillon de l'enquête Emploi doit nécessairement intégrer ces logements, pour deux raisons :

- Certains logements identifiés comme vacants ou résidences secondaires au moment du tirage de l'échantillon peuvent avoir changé de statut au moment de l'enquête et être devenus des logements principaux ;
- La notion de logements principaux au sens de la taxe d'habitation (qui répond à une règle fiscale) ne correspond pas strictement à celle utilisée dans l'enquête Emploi.

Il peut alors arriver qu'une grappe ait un nombre très important de logements non principaux. Dans ce cas, pour limiter la charge de travail des enquêteurs, une procédure d'échantillonnage a été mise en place sur les logements non principaux, selon les règles suivantes :

⁶ Dans l'ancien échantillon, la répartition par région avait été déterminée par un calcul d'allocation optimale.

- si la grappe contient 10 logements non principaux ou moins au total, tous ces logements sont inclus dans l'échantillon ;
- si la grappe contient de 11 à 40 logements non principaux au total, seuls 10 logements non principaux sont inclus aléatoirement dans l'échantillon;
- si la grappe contient de 41 à 100 logements non principaux au total, un quart de ces logements non principaux sont inclus aléatoirement dans l'échantillon ;
- si la grappe contient plus de 100 logements non principaux au total, 25 logements non principaux sont inclus aléatoirement dans l'échantillon.

Procédures de mise à jour de l'échantillon

Les fichiers annuels de la taxe d'habitation permettent de mettre à jour chaque année les informations permettant le repérage des logements sur le terrain, et de recueillir l'information auxiliaire la plus fraîche possible pour procéder aux opérations de correction de la non-réponse.

Il permettent par ailleurs, en amont de la collecte, d'enrichir l'échantillon par les logements nouveaux. L'existence d'un identifiant pérenne des logements permet de repérer les logements nouveaux, en sélectionnant les identifiants apparus dans le dernier fichier de la taxe d'habitation disponible, mais absents du fichier initial. Après avoir trié ce fichier par section et parcelle cadastrale, on rattache chaque logement nouveau repéré au premier logement non nouveau qui le précède. Si ce premier logement est présent dans l'échantillon, alors le logement nouveau est échantillonné et prend les caractéristiques de ce logement auquel il a été associé (numéro de secteur, numéro de grappe et pondération).

Après les mises à jours successives, si le nombre de logements nouveaux sélectionnés dans une grappe entrante est trop important, on recourt à une procédure d'échantillonnage identique à celle utilisée pour sélectionner les logements non principaux :

- si la grappe contient 10 logements nouveaux ou moins au total, tous ces logements sont inclus dans l'échantillon ;
- si la grappe contient de 11 à 40 logements nouveaux au total, seuls 10 logements neufs sont inclus aléatoirement dans l'échantillon;
- si la grappe contient de 41 à 100 logements nouveaux au total, un quart de ces logements neufs sont inclus aléatoirement dans l'échantillon ;
- si la grappe contient plus de 100 logements nouveaux au total, 25 logements neufs sont inclus aléatoirement dans l'échantillon.

Le déroulement de l'enquête

L'organisation de la collecte

La méthode d'enquête a été conçue de manière à diminuer, dans la mesure du possible, les erreurs d'observation susceptibles d'être commises sur le terrain.

Lorsqu'ils se rendent sur le terrain, les enquêteurs disposent d'une série d'informations présentes dans la base de sondage et qui leur permettent de procéder à un repérage du terrain : il s'agit du nom des occupants (i.e. des dernières personnes ayant été soumises à la taxe d'habitation dans ces logements), de leurs adresses foncière et postale, mais aussi des références cadastrales. Ces dernières informations leur permettent de faire un repérage du bâti, lorsque les adresses fournies sont ambiguës.

Le premier travail de l'enquêteur consiste à reconnaître les lieux à l'aide de ces informations et de renseigner les logements qu'il identifie comme ayant été détruits dans la grappe à enquêter.

L'enquêteur dispose sur son micro portable d'une « feuille de local » pour chaque logement de l'échantillon. Sur cette « feuille de local », l'enquêteur mentionne si le logement est bien à usage d'habitation ou non. Lorsque c'est bien le cas, il renseigne la catégorie du logement (résidence principale, logement vacant, résidence secondaire,...) ainsi que le résultat de l'interview.

Dans le cas où l'enquête est acceptée, il indique en outre la liste des occupants, leur situation dans le ménage et leur attribue à chacun un numéro d'ordre. Cette « feuille de local » permet d'attribuer le même numéro de logement à chaque logement d'un trimestre sur l'autre, et le même numéro individuel à chaque personne d'un trimestre sur l'autre. Cette numérotation rigoureuse et identique des personnes et des logements a un double intérêt : d'une part, elle permet au moment des enquêtes ultérieures de tenir compte des réponses précédentes et d'adapter le questionnaire à ces réponses passées. Ce repérage permet d'autre part, une fois la collecte terminée, et à des fins d'exploitation, de comparer les réponses d'une même personne au fil du temps et de mesurer les changements de sa situation.

Dans un deuxième temps, l'enquêteur procède aux interviews proprement dites. Dans le cas où l'enquête n'a pas à se faire (logement vacant, résidence secondaire) ou ne peut pas se faire (refus ou absence pendant toute la durée de l'enquête), l'enquêteur l'indique sur la « feuille de local ». Sinon il enquête la totalité des membres du logement âgés de 15 ans ou plus.

L'enquêteur n'est pas tenu d'interroger personnellement chacun des membres du ménage : l'une des personnes présentes dans le logement au moment de sa visite peut répondre pour les autres. Ainsi, en 2010, sur 100 questionnaires remplis, environ 30 ne l'ont pas été par la personne concernée, mais par un autre membre du logement, soit parce que la personne concernée était absente au moment de la visite ou de l'appel téléphonique de l'enquêteur, soit parce qu'elle ne pouvait pas ou ne souhaitait pas répondre. Si cette procédure rend l'enquête plus rapide et moins coûteuse, elle peut limiter en revanche sur certains points la précision des informations recueillies, et les enquêteurs doivent donc tenter autant que possible d'obtenir des réponses directes.

Le mode d'interrogation et la répartition des interviews dans le temps

Les logements sont interrogés six trimestres consécutifs. L'interrogation se déroule en vis-à-vis avec l'enquêté dans son logement lors de la première et de la sixième vagues d'enquête. Les deuxième, troisième, quatrième et cinquième vagues (ou « vagues intermédiaires ») se déroulent par téléphone, pour des raisons principalement budgétaires. Les enquêtes téléphoniques sont moins coûteuses, car elles économisent le déplacement physique de l'enquêteur. Si l'enquêteur n'a pas pu réaliser l'interview par téléphone, il doit effectuer une visite dans le logement à enquêter. L'interrogation par

visite est également préconisée dans le cas des logements comportant un nombre important de personnes, pour lesquels une interrogation téléphonique serait inadaptée, ainsi que dans les logements non munis de téléphone, ou dont les habitants ne souhaitent pas communiquer leur numéro. L'interrogation se fait également par visite en vague intermédiaire lorsqu'il s'agit de la première interrogation du ménage, soit parce que le ménage n'a pas répondu en première vague, soit parce que le ménage résidant dans le logement a changé.

Une lettre-avis prévient les ménages qu'ils vont être interrogés prochainement dans le cadre de l'enquête Emploi. En moyenne, d'après les tests réalisés en 2000, l'interrogation d'une personne dure 19 minutes lors de la première interrogation. Certaines questions n'étant posées que la première fois, les enquêtes suivantes sont en moyenne plus courtes : 7 minutes.

L'enquête se déroule tout au long des 52 semaines de l'année. Chaque interrogation est rattachée à une semaine de référence (du lundi au dimanche), sur laquelle portent les principales questions, notamment celles qui sont destinées à classer les personnes comme actifs, chômeurs ou inactifs. Pour une semaine de référence donnée, la collecte commence le lundi qui suit la semaine de référence, et se termine deux semaines et deux jours après. Les enquêteurs ont donc deux semaines et deux jours pour interroger tous les logements qui sont associés à cette semaine de référence. Ils ont pour consigne de réaliser le plus grand nombre d'entretiens possible en début de période de collecte, d'une part parce que les personnes enquêtées peuvent fournir des réponses plus précises si on les interroge sur des événements récents, d'autre part parce que l'enquêteur se donne ainsi les moyens de parvenir à joindre le maximum de personnes.

Pour un logement donné, les six interrogations successives sont attachées à des semaines de référence espacées d'un trimestre exactement (13 semaines⁷). Les exploitations trimestrielles se font par agrégation des réponses obtenues au cours des 13 semaines du trimestre. Les résultats annuels publiés sont une moyenne des résultats trimestriels.

La qualité des résultats issus de l'enquête

La qualité des résultats de l'enquête dépend d'abord du bon déroulement de la collecte, mais aussi des traitements effectués sur les données recueillies.

La qualité de la collecte

Au stade de la collecte, le succès de l'enquête suppose un échantillonnage correct, un questionnaire d'enquête adapté au sujet traité et au public interrogé et un réseau d'enquêteurs qualifiés. Dans l'ensemble, ces conditions sont remplies dans l'enquête Emploi.

Le questionnaire a été rénové à plusieurs occasions, afin d'améliorer la formulation des questions et l'adéquation du questionnaire aux besoins d'informations sur le marché du travail. Les enquêteurs sont formés chaque année, et nombre d'entre eux ont une longue expérience de l'enquête Emploi.

⁷ Une année comptant 52 semaines plus un ou deux jours, un décalage se crée progressivement entre les trimestres de référence de l'enquête et le calendrier. Une semaine supplémentaire est donc ajoutée à certains trimestres.

De plus, les directions régionales de l'Insee assurent l'apurement des anomalies détectées, avec visualisation éventuelle des questionnaires sur ordinateur et retour à l'enquêteur si cela est nécessaire.

Malgré tout cela, les résultats de l'enquête restent imparfaits : ils comportent une certaine part d'aléa, due au fait qu'ils sont obtenus à partir d'un échantillon. En outre, certains ménages refusent de répondre ou sont impossibles à joindre ; le dénombrement des logements ou de leurs occupants peut comporter des erreurs ou des omissions ; enfin, certaines questions (le salaire, la profession, l'adresse de l'employeur, le diplôme, etc.) peuvent être sujettes à des erreurs, des imprécisions ou même au refus de répondre des personnes interrogées.

L'analyse des résultats de l'enquête sous-échantillon par sous-échantillon (sixième entrant, sous-échantillon en deuxième interrogation, etc.) ne donne pas les mêmes résultats sur chacun d'entre eux. Ce résultat est connu depuis longtemps. Dès 1975, Barbara Bailar (« The effects of rotation group bias on estimates from panel surveys », *Journal of American Statistical Association*) identifiait ces différences sur le Current Population Survey (CPS), l'équivalent américain de l'enquête Emploi. Dans l'enquête Emploi annuelle, on avait noté et étudié ces différences dès le début des années 1980. L'enquête Emploi trimestrielle présente le même type de limites. Selon les vagues, les résultats peuvent être différents, que ce soit dû à des fluctuations d'échantillonnage (le hasard ne donne pas forcément la même chose sur plusieurs échantillons, les risques de divergence étant d'autant plus grands que leur taille est petite) ou que ce soit dû au fait que le comportement de réponse des ménages (acceptation de l'enquête, façon d'y répondre, ...) est susceptible de dépendre du fait d'avoir déjà été interrogé ou pas. Des calages et redressements sont effectués pour corriger ces biais (cf. infra) mais ils peuvent en laisser subsister une partie. Ce phénomène est identifié dans la littérature académique sous le nom de « biais de rotation ».

Les redressements

Comme pour toutes les enquêtes réalisées auprès des ménages, le redressement des résultats de l'enquête a deux objectifs principaux : compenser le biais introduit par la non-réponse et réduire autant que possible les fluctuations d'échantillonnage⁸.

Les redressements sont réalisés par la méthode de « calage sur marges ». Cette méthode consiste à modifier les poids des personnes répondantes de telle sorte que le total de certaines variables (les marges) estimé à partir des répondants de l'échantillon soit égal au vrai total connu par ailleurs.

La correction de la non-réponse

Dans l'enquête en continu, le taux de sondage⁹ est d'environ 1/380e, c'est-à-dire que pour un trimestre donné, un logement de l'échantillon représenterait 380 logements de la population totale si tous les logements pouvaient effectivement être enquêtés.

⁸ Une description détaillée de la démarche d'élaboration de cette méthode est exposée dans « Repondération dans la nouvelle enquête Emploi en continu », Philippe Février et Pauline Givord, Journées de méthodologie statistique, 2002 (accessible sur : <http://jms.insee.fr/>), dans « Travaux sur la repondération de l'enquête Emploi en continu », Philippe Février, note interne n° 056/DG75-F410, et enfin dans la note interne « Repondération de l'enquête emploi en continu », Dominique Place, N° 3141/DG75-F410.

De fait, il n'est pas possible d'obtenir une réponse dans tous les logements. Corriger de la non-réponse consiste à effectuer une modification des poids initiaux des logements, pour prendre en compte le fait que tous ne répondent pas. Dans la pratique, les non-répondants ne se répartissent pas aléatoirement dans l'échantillon : le taux de non-réponse des personnes vivant seules est plus élevé que celui des logements comportant plusieurs personnes, il est également plus élevé dans les grandes villes que dans les campagnes, dans les immeubles que dans les maisons individuelles. Le redressement sera d'autant meilleur qu'il transformera les poids initiaux en sorte que les ménages défaillants soient représentés par des ménages possédant autant que possible les mêmes caractéristiques. Ces caractéristiques doivent donc être connues à la fois sur les répondants et sur les non-répondants. Pour ce qui concerne l'enquête Emploi, les variables qui ont été sélectionnées sont celles qui caractérisent le logement dans la taxe d'habitation :

- la situation du logement (résidence principale, secondaire, vacante) ;
- le statut d'occupation du logement (locataire, propriétaire) ;
- le type de logement (individuel, collectif) ;
- le nombre de pièces du logements ;
- le statut HLM ou non du logement;
- le décile de revenu du ménage ;
- le zonage en aires urbaines.

Les marges correspondantes sont mises à jour annuellement, dès la réception des nouveaux fichiers de la taxe d'habitation.

Depuis le 1er trimestre 2007, une enquête par courrier postal est menée chaque trimestre auprès des logements non-répondants : le questionnaire envoyé correspond à une version simplifiée de l'enquête Emploi, centré principalement sur les questions sur le statut d'activité. Parmi les répondants à cette enquête¹⁰, on constate une surreprésentation des actifs occupés. Les variables de calage mentionnées plus haut ne permettent qu'imparfaitement de corriger le biais dû à la non-réponse. Il a donc été décidé de tenir compte des informations sur le statut d'activité apportées par l'enquête non-répondants dans la procédure de calage¹¹.

Le redressement pour fluctuations d'échantillonnage

Lorsque l'on tire au sort un échantillon pour une enquête, il arrive fréquemment que l'échantillon obtenu ne soit pas exactement représentatif de la population de référence, au regard de certains critères importants pour le sujet de l'enquête. Par exemple, on sait qu'en 2010, il y avait 51,6 % de

⁹ L'inverse du taux de sondage est appelé « poids initial » ou pondération initiale. Ainsi on peut dire que le poids initial des logements de l'échantillon est d'environ 380.

¹⁰ Durant l'année 2009, le taux de réponse à cette enquête postale était de 22,4 %.

¹¹ La méthodologie de l'enquête non-répondants et son intégration dans l'enquête Emploi sont décrites dans « L'enquête auprès des non-répondants à l'enquête Emploi », Sébastien Durier et Corinne Prost, Les actes des journées de méthodologie statistique, 2009.

femmes dans la population métropolitaine française vivant en ménages ordinaires. Quand on calcule la proportion de femmes ayant répondu à l'enquête, on en trouve 52 %. Cette différence entre la vraie valeur et la valeur fournie par l'enquête s'appelle une fluctuation d'échantillonnage.

Ces fluctuations d'échantillonnage peuvent être redressées par une modification des poids des individus de l'échantillon. Corriger les fluctuations d'échantillonnage suppose d'avoir des informations issues d'une autre source que l'enquête et de meilleure qualité, ou mieux encore de la « vraie » valeur. Les possibilités de correction se trouvent de ce fait limitées. Les critères sur lesquels s'effectuent la repondération dans l'enquête Emploi sont :

-au niveau individuel, l'âge en tranches d'âge quinquennal, le sexe et la région ;

-au niveau logement, le nombre total de logements, le nombre de résidences principales, le nombre de logements neufs.

Ainsi, le redressement modifie le poids de chaque répondant en sorte que :

-le nombre de personnes de même sexe et de même âge quinquennal estimé au niveau régional par l'enquête soit égal au nombre de personnes de même sexe et de même âge quinquennal connu par ailleurs pour chacune des 21 régions de la France métropolitaine (informations démographiques issues de l'actualisation par les données de l'état civil des données du recensement de la population). Dans l'enquête Emploi, c'est plus précisément la population des logements ordinaires détaillée par sexe, âge quinquennal au niveau régional qui est utilisée comme information de calage ;

-le nombre total de logements et les nombres de résidences principales et de logements neufs estimés par l'enquête soient égaux au nombre total de logements et aux nombres de résidences principales et de logements neufs donnés par le compte du logement.

La combinaison des deux types de redressement

Les deux types de repondération, pour correction de la non-réponse et pour fluctuations d'échantillonnage, peuvent se faire soit successivement (transformation des poids pour correction de la non-réponse, puis modification des poids obtenus pour correction des erreurs aléatoires), soit simultanément. Lorsqu'on procède en deux étapes, comme c'était le cas auparavant dans l'enquête annuelle, le deuxième calage déforme le jeu de pondérations obtenu à la première étape et « détruit » au moins partiellement le premier calage.

Dans l'enquête actuelle, c'est un calage en une seule étape qui est réalisé, mais au prix d'une certaine opacité puisqu'il n'est plus possible de dissocier les modifications de poids dues à la non-réponse et aux fluctuations d'échantillonnage. Par ailleurs, afin d'assurer la cohérence entre les estimations au niveau logement et les estimations au niveau individu, l'unité statistique sur laquelle a été effectué le calage est le logement. Toutes les personnes de 15 ans ou plus d'un même logement ont donc le même poids.

Numériquement, alors que les pondérations initiales des logements étaient en moyenne de 380, les pondérations après redressement (redressement pour la non-réponse et redressement pour fluctuations d'échantillonnage) sont en moyenne de 500.

La repondération est réalisée séparément pour chacun des six sous-échantillons. En effet, le comportement de réponse semble dépendre du rang d'interrogation de l'enquête (les taux de réponse sont plus élevés lors des interrogations intermédiaires de la grappe qu'en première et dernière interrogation). Il est donc préférable de traiter séparément les sous-échantillons. Les estimations finales des variables sur la totalité de l'échantillon du trimestre sont obtenues par la moyenne des estimations obtenues sur chacun des six sous-échantillons.

L'actualisation des pondérations

L'enquête Emploi étant exploitée trimestriellement, le redressement se fait chaque trimestre sur la répartition de la population régionale par sexe et âge quinquennal au milieu du trimestre. Ces informations sont issues du bilan démographique produit chaque année à partir du recensement de la population. A chaque nouvelle estimation du recensement de la population, le bilan démographique est révisé rétrospectivement sur les trois années précédentes. Les calages des enquêtes emploi sont alors refaits à chaque révision annuelle de population, ce qui peut expliquer des différences marginales entre des chiffres publiés avec des pondérations différentes.

Évaluation des erreurs aléatoires de l'enquête

Les données de l'enquête Emploi sont, comme toutes les enquêtes par sondage, affectées d'erreurs qui peuvent être mesurées par l'écart-type de chacune des estimations. L'estimation de ces erreurs aléatoires repose sur deux éléments :

- d'une part, le mode d'élaboration des données, c'est-à-dire le type d'échantillonnage (tirage stratifié aréolaire) décrit précédemment ;
- d'autre part, le fait que l'enquête est recalée sur une information qui lui est exogène et qui est considérée comme certaine (redressements). L'échantillon utilisé pour faire cette enquête est l'un des multiples échantillons qu'il aurait été possible de constituer avec le plan de sondage utilisé. Chacun de ces échantillons aurait donné des chiffres plus ou moins différents. La technique d'échantillonnage utilisée est sans biais, c'est-à-dire que pour un chiffre donné, la moyenne des estimations de chacun des échantillons que l'on aurait pu constituer par cette technique est égale à ce qu'on obtiendrait en menant l'enquête auprès de l'ensemble de la population.

Il reste donc à déterminer dans quelle mesure l'estimation de ce chiffre par notre échantillon particulier risque de s'écarter de la moyenne. C'est ce que mesure précisément l'écart-type de l'estimation à partir duquel est établi l'intervalle de confiance, cet écart-type étant lui-même estimé.

Si on menait l'enquête sur la multitude des échantillons qu'il est possible de constituer par la technique utilisée et si, pour chacun d'eux, on procédait au calcul de l'effectif d'une population donnée et à celui de l'écart-type de cette estimation, alors à peu près 95 % des intervalles de deux écarts-types de part et d'autre des estimations contiendraient la moyenne de toutes les estimations possibles (qui est égale, rappelons-le, dans un échantillon sans biais comme celui de l'enquête Emploi, à la vraie valeur).

Pour un échantillon particulier, cet intervalle a donc une probabilité de 0,95 (95 chances sur 100) de contenir l'effectif que l'on obtiendrait en menant l'enquête sur toute la population. De même, un intervalle de plus ou moins un écart-type a 2 chances sur 3 de contenir cet effectif. La proportion atteint 99 % pour un intervalle de 3 écarts-types de part et d'autre de l'estimation. Certains de ces intervalles sont fournis dans certaines publications tirées de l'enquête emploi.

Historique des évolutions de l'enquête

L'enquête Emploi a connu plusieurs évolutions notables. Cette partie détaille les épisodes les plus récents de cette histoire.

Pas encore « en continu », l'enquête Emploi annuelle (jusqu'en 2002)

L'enquête Emploi ne s'est pas toujours déroulée, comme actuellement, tout au long de l'année. Avant le 1^{er} juillet 2001, elle était organisée au cours du mois de mars. L'enquête était donc annuelle et ne permettait donc pas d'investiguer des évolutions infra-annuelles.

De plus, les ménages étaient interrogés trois années consécutives (contre six trimestres consécutifs dans l'enquête actuelle). Enfin, la période au cours de laquelle peut s'effectuer la collecte était plus longue dans l'enquête annuelle, puisque les enquêteurs disposaient de quatre à cinq semaines pour interroger un nombre de logements sensiblement équivalent à ceux qui leur sont attribués actuellement¹² (contre deux semaines et deux jours dans l'enquête Emploi actuelle).

Ce changement de méthodologie a introduit une rupture importante dans l'enquête. La proportion d'enquêtes acceptées a notamment baissé avec le passage à l'enquête Emploi en continu. Plusieurs raisons expliquent la légère détérioration de cet indicateur. Tout d'abord, le raccourcissement des périodes autorisées pour la collecte, ce qui a mécaniquement accru la difficulté pour les enquêteurs de joindre les ménages. Ensuite, la collecte étant désormais étalée sur toute l'année, elle couvre des périodes où le taux d'absence des ménages est élevé (vacances scolaires par exemple), alors que l'enquête annuelle se déroulait à un moment de l'année caractérisé par un fort taux de présence. Enfin, le nombre et la fréquence des ré-interrogations dans l'enquête en continu accroît le risque de lassitude des ménages.

L'enquête Emploi annuelle portait sur environ 75 000 logements répondants, comptant 150 000 personnes de 15 ans et plus.

Le passage à la méthodologie actuelle « en continu » date du 1^{er} juillet 2001. Toutefois, pour comparer les résultats, les deux enquêtes ont été menées de front en 2002 : une enquête annuelle en mars, et une enquête collectée en continu tout au long de l'année.

¹² Dans l'enquête annuelle, le nombre de logements par aire était de 40 en zone rurale et de 20 en zone urbaine.

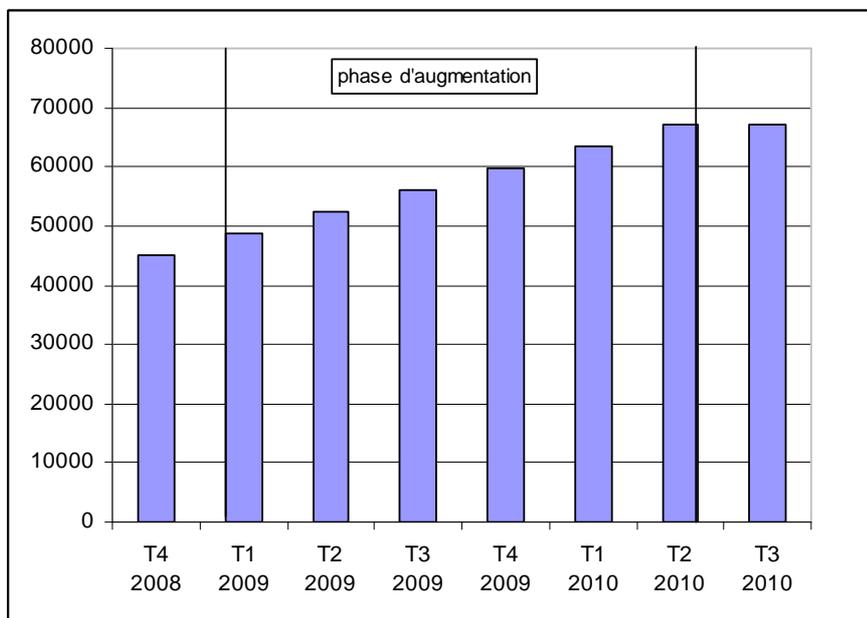
Une phase d'agrandissement de l'échantillon (du T1 2009 au T2 2010)

Suite au débat sur les statistiques du chômage de 2007, l'IGF-Igas a expertisé l'enquête Emploi et a formulé une liste de recommandations visant à améliorer la qualité des résultats issus de l'enquête¹³.

L'une des recommandations portées par le rapport de l'IGF-Igas concernait la taille de l'échantillon : pour améliorer la précision des estimateurs, il allait falloir augmenter le nombre d'enquêtés. L'Insee s'est ainsi engagé dans une opération permettant d'augmenter progressivement de 50% de la taille de l'échantillon de l'enquête.

Cette augmentation, initiée début 2009, a été réalisée par l'introduction dans l'échantillon d'un nombre de ménages supérieur à ceux qui en sortent. Ce surnombre a été calibré afin de permettre à l'échantillon d'atteindre son objectif de taille définitive au T2 2010.

Graphique : Nombre de logements à enquêter (résidences principales)



A titre de comparaison, l'échantillon trimestriel de l'enquête a évolué de la façon suivante :

- Au lieu de 45 000 résidences principales à enquêter avant le T1 2009, elles sont 67 000 à partir du T2 2010 ;
- Au lieu de 38 000 logements répondants, l'échantillon en compte désormais 57 000 ;
- Au lieu de 72 000 personnes de 15 ans et plus répondantes, 108 000 personnes.

¹³ « Rapport sur les statistiques d'estimation du chômage », Inspection Générale des Finances, Inspection Générale des Affaires Sociales.

Une nouvelle base de sondage à partir de 2009

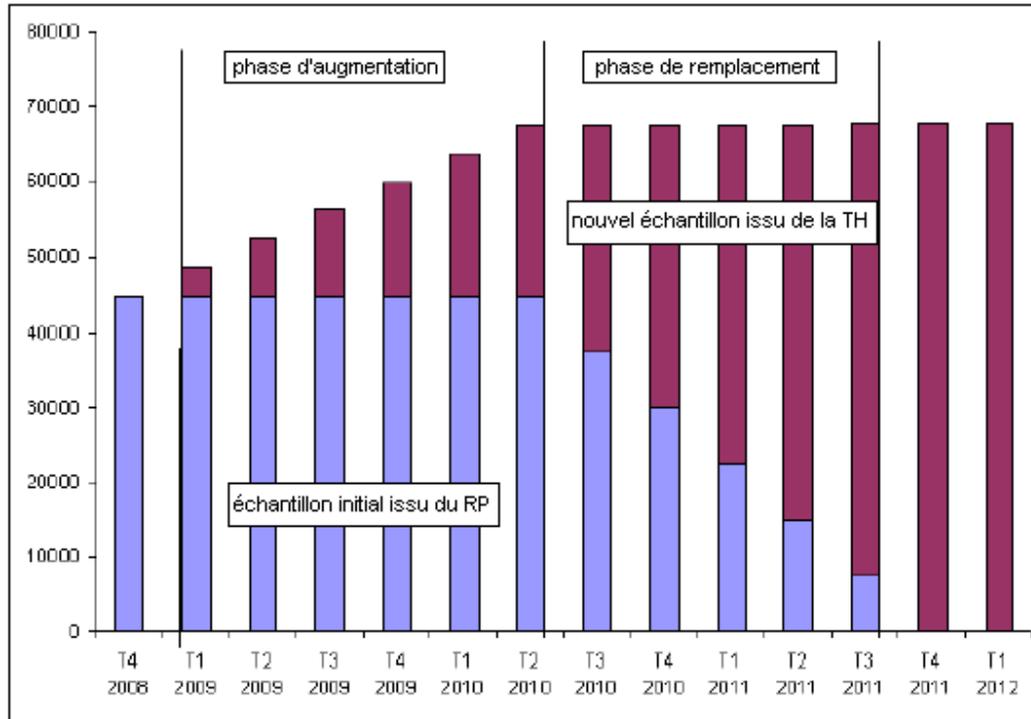
L'opération d'augmentation de la taille de l'échantillon a coïncidé avec une autre évolution majeure de l'enquête : l'introduction progressive d'une nouvelle base de sondage pour le tirage de l'échantillon.

Avant 2009, l'échantillon de l'enquête était construit à partir des informations issues du recensement de la population de 1999, et avait été calibré pour pouvoir être utilisé jusqu'à mi-2010. La durée de vie totale de l'échantillon était donc de 36 trimestres ou 9 ans¹⁴.

Ce sont les données de la taxe d'habitation qui ont servi de base de sondage à partir de 2009. A partir du 1er trimestre 2009, en même temps que l'augmentation de la taille de l'échantillon, a débuté l'opération de remplacement de l'échantillon initial issu du recensement de 1999, par celui de la taxe d'habitation.

Ainsi, la phase d'augmentation de l'échantillon, du 1er trimestre 2009 au 2ème trimestre 2010, se caractérise par le maintien de l'échantillon initial dans son rythme de croisière, auquel on ajoute 1/3 du nouvel échantillon, par sixième. A partir du 3ème trimestre 2010, l'échantillon initial commence à décliner par sixième et est remplacé par les 2/3 restants du nouvel échantillon. Au 4ème trimestre 2010, l'échantillon de l'enquête n'est plus constitué que de l'échantillon tiré de la taxe d'habitation.

¹⁴ La procédure de construction de l'échantillon initial issu du recensement est expliquée de façon détaillée dans « La construction de l'échantillon de la future enquête Emploi en continu à partir du recensement de 1999 », Marc Christine, Les actes des journées de méthodologie statistique, 4 et 5 décembre 2000, Insee Méthodes n°100, novembre 2002.

Graphique : répartition du nombre de résidences principales selon le type d'échantillon

Dans le nouvel échantillon, issu de la taxe d'habitation, les grappes et secteurs sont construits automatiquement selon une logique ascendante (construction des grappes, puis de secteurs). Les aires et secteurs de l'échantillon issu du recensement étaient eux construits manuellement selon une logique descendante. Dans un premier temps, l'intégralité du territoire avait été découpé en unités primaires, puis un tirage avait été effectué parmi ces unités primaires. Seules les unités tirées ont été découpées en secteurs, puis en aires. Secteurs et aires avaient ainsi été conçus pour constituer des regroupements de logements contigus.

Le schéma global de sélection du nouvel échantillon est le suivant :

- Sélection de l'ensemble des secteurs par région ;
- Répartition des secteurs en 6 groupes de même taille ;
- Scission de chaque groupe en 2 sous-échantillons, le premier comportant 1/3 des secteurs le second les 2/3.

Les secteurs issus de la taxe d'habitation ont été répartis en 6 groupes, chacun déterminant la date d'entrée des secteurs dans l'échantillon pour chacune des deux phases d'augmentation et de remplacement : les secteurs du groupe 1 entrent dans l'échantillon au 1er trimestre de la phase d'augmentation de l'échantillon (1er trimestre 2009) ou au 1er trimestre de la phase de renouvellement de l'échantillon (3ème trimestre 2010),... Enfin, deux sous-échantillons ont été construits à partir de chaque groupe, l'un correspondant aux secteurs mobilisés en phase

d'augmentation de l'échantillon (et contenant donc 1/3 des secteurs), l'autre aux secteurs mobilisés en phase de remplacement de l'échantillon correspondant (et contenant les 2/3 des secteurs).

On dispose ainsi des 12 sous-échantillons, dont le rythme d'entrée est donné par le tableau 2.

Tableau 2 : nombre de secteurs entrants par trimestre (sous-échantillon) et par région de gestion

Région de gestion	Phase 1 : augmentation de la taille de l'échantillon						Phase 2 : remplacement de l'échantillon						Ensemble
	Sous-échantillon de secteurs (et groupe)												
	1 (gr 1)	2 (gr 2)	3 (gr 3)	4 (gr 4)	5 (gr 5)	6 (gr 6)	7 (gr 1)	8 (gr 2)	9 (gr 3)	10 (gr 4)	11 (gr 5)	12 (gr 6)	
	Trimestre d'entrée												
	T1 2009	T2 2009	T3 2009	T4 2009	T1 2010	T2 2010	T3 2010	T4 2010	T1 2011	T2 2011	T3 2011	T4 2011	
Île-de-France	22	20	21	25	26	22	49	47	49	40	42	48	411
Champagne-Ardenne	8	12	14	8	8	8	18	10	18	22	23	10	159
Picardie	0	2	6	7	5	6	8	13	11	10	13	9	90
Haute-Normandie	8	10	5	8	10	7	14	14	17	18	13	16	140
Centre	9	12	7	3	14	11	19	22	18	22	21	22	180
Basse-Normandie	9	6	6	6	4	4	5	9	9	8	10	10	86
Bourgogne	4	2	6	7	5	5	13	11	9	7	10	12	91
Nord-Pas-de-Calais	11	9	10	9	11	12	21	21	23	24	20	20	191
Lorraine	8	5	4	6	6	11	15	12	15	17	9	8	116
Alsace	8	3	5	6	6	3	8	10	8	8	14	13	92
Franche-Comté	6	7	4	4	5	1	10	8	4	12	9	9	79
Pays de la Loire	12	11	12	9	9	10	19	18	17	22	17	19	175
Bretagne	8	8	8	10	11	9	17	20	17	17	20	17	162
Poitou-Charentes	8	6	4	2	4	6	9	12	10	10	8	12	91
Aquitaine	9	8	10	12	10	5	18	19	17	13	20	20	161
Midi-Pyrénées	6	7	10	9	8	6	19	15	16	17	15	15	143
Limousin	3	5	5	2	3	6	9	9	9	8	8	10	77
Rhône-Alpes	17	20	18	18	13	17	34	31	34	34	34	33	303
Auvergne	4	4	4	4	4	5	9	9	5	8	8	9	73
Languedoc-Roussillon	9	7	7	6	5	10	13	15	17	15	12	14	130
PACA	11	15	13	14	12	14	30	27	30	25	30	30	251
Corse	0	0	2	2	2	1	1	4	0	0	0	4	16
Ensemble	180	179	181	177	181	179	358	356	353	357	356	360	3217

Du 1er trimestre 2009 et jusqu'au 3ème trimestre 2011, l'échantillon initial issu du recensement de 1999 et le nouvel échantillon issu de la taxe d'habitation coexistent. Le redressement doit donc en tenir compte et s'effectue en utilisant des caractéristiques propres aux deux échantillons. Pour l'échantillon issu du recensement 1999, il s'agit de :

- la taille de l'unité urbaine dans laquelle est localisé le logement en 5 groupes ;
- le type de logement (maison individuelle, ferme ou immeuble) ;
- la nature du logement (résidence principale, secondaire ou logement vacant)
- le nombre de pièces du logement ;
- le type socio-économique du quartier au sens de la typologie Tabard¹⁵.

Pour les logements de l'échantillon initial construits après le recensement de 1999, on ne dispose pas par définition des variables de calage caractérisant le logement au moment du recensement 1999. Les logements neufs sont donc considérés séparément. On cale le nombre de logements neufs trouvés dans l'enquête Emploi sur le nombre total de constructions neuves (en distinguant selon la tranche d'unité urbaine), calculé par l'Insee à partir des sources du ministère chargé de l'équipement.

L'échantillonnage à partir du recensement présentait plusieurs particularités par rapport à la base de sondage actuelle :

- L'échantillon ne pouvait pas être mis à jour, comme il l'est actuellement, avec les nouveaux logements construits après 1999 (millésime du recensement mobilisé pour la constitution de la base de sondage). Les enquêteurs devaient donc s'acquitter d'un lourd travail de repérage visant à identifier toutes les modifications éventuelles du parc de logements (fusion, éclatement de logement, changement d'affectation, construction, démolition, etc.), par rapport aux informations qui leur étaient fournies dans un plan détaillé de l'aire d'enquête et un dossier d'aire.
- Les aires de l'échantillon issu du recensement ont été construites manuellement de façon à obtenir des regroupements de logements contigus. Au contraire, la construction automatique des grappes a créé des ruptures de contiguïté, qu'on a néanmoins essayé de corriger dans la mesure du possible.

¹⁵ Cf. Document de travail Insee n°F0208: « Représentation socio-économique du territoire », Nicole Tabard et Géraldine Martin-Houssart, 2002

Table des matières

L'enquête Emploi en continu	2
Les concepts d'activité, de chômage et de sous-emploi	5
La méthodologie de l'enquête	12
Le champ de l'enquête : personnes interrogées et représentativité.....	12
L'échantillon de l'enquête	12
Le déroulement de l'enquête	16
La qualité des résultats issus de l'enquête	18
Historique des évolutions de l'enquête	25
Pas encore « en continu », l'enquête Emploi annuelle (jusqu'en 2002).....	25
Une phase d'agrandissement de l'échantillon (du T1 2009 au T2 2010).....	26
Une nouvelle base de sondage à partir de 2009.....	27
Table des matières	31