

Appariement entre l'enquête Emploi et le fichier Historique de Pôle emploi sur la période 2012-2017 Méthode et premiers résultats

**Alexis HAMEAU, Sylvain LARRIEU, Anis MARRAKCHI, Alexis MONTAUT,
Christophe DIXTE, Yohan CODER et Sophie HAMMAN**

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES

Série des Documents de Travail

de la

DIRECTION DES STATISTIQUES DÉMOGRAPHIQUES ET SOCIALES

N°F1904

Appariement entre l'enquête Emploi et le fichier Historique de Pôle emploi sur la période 2012-2017

Méthode et premiers résultats

ALEXIS HAMEAU, SYLVAIN LARRIEU, ANIS MARRAKCHI, ALEXIS MONTAUT
(Insee, Division synthèse et conjoncture du marché du travail)

CHRISTOPHE DIXTE
(Dares, Département marché du travail)

YOHAN CODER, SOPHIE HAMMAN
(Pôle emploi, Département marché du travail)

Document de travail

Juillet 2019

Les auteurs tiennent à remercier Boris Guannel, Yves Jauneau, Nicolas Vanni, Anne-Juliette Bessonne, Emmanuel Chion et Vladimir Passeron dont les relectures attentives ont fait progresser la qualité de ce document de travail. Ils remercient aussi Pierre Faury pour avoir constitué les bases du Fichier Historique. Ils remercient enfin la Division Statistiques et Analyses Urbaines, dont le procédé de géolocalisation a permis d'améliorer sensiblement la qualité de l'appariement.

Ce travail a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du programme Investissements d'avenir portant la référence ANR-10-EQPX-17 (Centre d'accès sécurisé aux données – CASD).

Ces documents de travail ne reflètent pas la position de l'INSEE et n'engagent que leurs auteurs.
Working-papers do not reflect the position of INSEE but only their authors'views.

Appariement entre l'enquête Emploi et le fichier Historique de Pôle emploi sur la période 2012-2017

Méthode et premiers résultats

Alexis Hameau, Sylvain Larrieu, Anis Marrakchi, Alexis Montaut (Insee)
Christophe Dixte (Dares)
Yohan Coder, Sophie Hamman (Pôle emploi)

2 juillet 2019

Résumé

En moyenne en 2017 et en France métropolitaine, 2,6 millions de personnes de 15 à 64 ans sont au chômage au sens du Bureau international du travail (BIT) et 3,4 millions sont inscrites à Pôle emploi en catégorie A. Entre 2013 et 2017, l'écart entre ces deux mesures de référence s'est accentué, passant de 0,3 à 0,8 million. Deux populations sont susceptibles d'expliquer ce phénomène : les inscrits en catégorie A non chômeurs au sens du BIT et les chômeurs BIT non inscrits en catégorie A. L'appariement entre les enquêtes Emploi de l'Insee et le Fichier Historique de Pôle emploi sur cette période permet de les identifier et de les dénombrer toutes les deux. En pratique, on cherche à retrouver les personnes de l'enquête Emploi dans le Fichier Historique à partir de leur prénom, leur sexe, leur date de naissance et leur(s) adresse(s) sur la période 2012-2017.

En 2017, parmi les 2,7 millions d'inscrits en catégorie A ayant pu être reconstitués *via* l'appariement, 56 % sont au chômage au sens du BIT, 20 % sont dans le halo autour du chômage, 16 % sont inactifs hors du halo et 9 % sont en emploi. Inversement, selon l'appariement, 22 % des chômeurs au sens du BIT en 2017 ne sont pas inscrits à Pôle emploi et 11 % sont inscrits en catégorie B ou D. Les jeunes sont surreprésentés parmi les chômeurs BIT non inscrits à Pôle emploi, tandis que les seniors sont plus nombreux parmi les inactifs (hors halo) inscrits en catégorie A.

Entre 2013 et 2017, la hausse du nombre d'inscrits en catégorie A qui sont inactifs au sens du BIT (+0,3 million de personnes) est celle qui a le plus contribué à accroître l'écart entre chômage BIT et inscrits en catégorie A. La part des seniors inactifs au sens du BIT et inscrits en catégorie A a augmenté, dans un contexte de fin des dispenses de recherche d'emploi et de hausse de l'âge légal de départ à la retraite. Enfin, les jeunes, plus souvent chômeurs au sens du BIT et non inscrits, ont davantage bénéficié de l'embellie économique en 2017.

Mots-clés : chômage, halo, enquête Emploi, Pôle emploi, appariement

Matching between the Labour Force Surveys and the Historical File from Pôle emploi over the period 2012-2017

Method and first results

Alexis Hameau, Sylvain Larrieu, Anis Marrakchi, Alexis Montaut (Insee)
Christophe Dixte (Dares)
Yohan Coder, Sophie Hamman (Pôle emploi)

2 july 2019

Abstract

On average, in 2017 and in metropolitan France, 2.6 million people aged 15 to 64 are unemployed according to the International Labour Office (ILO) definition whereas 3.4 million are registered with the public employment service ("Pôle emploi") and classified as "category A" job seekers. Between 2013 and 2017, the gap between these two benchmark measures widened from 0.3 million to 0.8 million. Two populations are likely to explain it: jobseekers registered in category A who are not ILO-unemployed and ILO-unemployed persons who are not registered in category A. The matching between Labour Force Surveys and the Historical File of Pôle emploi over this period allows to identify and count them. In practice, we try to find people from the Labour Force Surveys in the Historical File based on their first name, sex, date of birth and address(es) over the period 2012-2017.

In 2017, of the 2.7 million category A jobseekers reconstituted through matching, 56% are unemployed according to the ILO definition, 20% are in the "halo of unemployment", 16% are "economically inactive" (excluding halo) and 9% are employed. Conversely, 22% of the unemployed according to the ILO are not registered with Pôle emploi in 2017 at the time of the survey and 11 % are registered, in category B or D. Young people are over-represented among the ILO-based unemployed not registered at Pôle emploi, while seniors are more numerous among the economically inactive (excluding halo) registered in category A.

Between 2013 and 2017, the number of people registered in category A and who are economically inactive, according to the ILO definition, increased (+0.3 million people) and strongly contributed to widen the gap between ILO-based unemployment and jobseekers registered in category A. The share of economically inactive seniors according to the ILO and registered in category A has increased, in a context of the end of job-search exemptions and of the increase in the legal retirement age. Lastly, young people, who are more often ILO-based unemployed and not registered, benefited more from the economic upturn in 2017.

Keywords : Unemployment, halo of unemployment, Labour Force Survey, public employment service, matching

Sommaire

Principaux résultats	5
Main results	7
Introduction	9
Quelques définitions	9
Les données	11
Chapitre 1- La méthode d'appariement.....	12
1-1 Algorithme général de l'appariement	12
1-2 Comment comparer les prénoms ?	13
1-3 Comment comparer les adresses ?	15
1-4 Règle de décision pour accepter ou rejeter un appariement	16
1-5 Autres critères pour apparier.....	18
1-6 Taux de couverture et biais d'appariement.....	20
Chapitre 2 – Les principaux résultats en coupe	24
2-1 Recoupements des statuts BIT et des catégories de demandeur d'emploi.....	24
2-2 Les demandeurs d'emploi en catégorie A non chômeurs au sens du BIT	30
2-3 Les chômeurs BIT non inscrits en catégorie A.....	35
Chapitre 3 – Évolutions 2013-2017	39
3.1 - Davantage d'inactifs inscrits en catégorie A et davantage de personnes en emploi et inscrites en catégories B et C	39
3.2 – Le recul de l'âge de départ à la retraite et la fin des dispenses de recherche d'emploi ont accru le nombre de seniors inactifs au sens du BIT et inscrits en catégorie A.....	40
3.3 – Les jeunes ont davantage profité de l'embellie économique en 2017, ce qui a réduit le nombre de chômeurs non inscrits.....	44
3.4 - Le chômage, un état plus instable en 2017 qu'en 2013	45
Bibliographie	47
Annexe 1 – précisions méthodologiques.....	48
1.1 Fonctions « Comped » et « Soundex » sous SAS	48
1.2 Un nombre de prénoms qui peut être différent entre le Fichier Historique et l'enquête Emploi	48
1.3 Géolocalisation	49
1.4 Probabilité théorique de « faux-positifs »	50
1.5 Une méthode d'appariement robuste aux faux-positifs	52
1.6 Traitement des libellés en clair	53
Annexe 2 – tableaux et graphiques complémentaires	55

Principaux résultats

En moyenne en 2017, en France métropolitaine, 2,6 millions de personnes de 15 à 64 ans sont au chômage au sens du Bureau international du travail (BIT) et 3,4 millions sont inscrites à Pôle emploi et classées comme demandeurs d'emplois de catégorie A. L'écart entre ces deux mesures de référence, qui ne reposent pas sur les mêmes concepts, s'est accentué sur la période 2012-2017, passant de 0,3 million en 2012 à 0,8 million en 2017.

Les enquêtes Emploi en continu de 2012 à 2017 (hors DROM) permettent de dénombrer les chômeurs au sens du BIT. Elles ont été appariées au fichier Historique des demandeurs d'emploi de Pôle emploi, qui compile l'ensemble des personnes inscrites à Pôle emploi sur la même période. . Apparier les deux sources consiste à retrouver les personnes de l'enquête Emploi dans le fichier Historique à partir de leur prénom, leur sexe, leur date de naissance et leur(s) adresse(s) sur la période 2012-2017 (appariement dit « statistique »). Ceci permet alors d'enrichir l'enquête Emploi avec des variables du fichier Historique : dates d'inscription à Pôle emploi, catégorie de demandeur d'emploi (A, B, C, D, E), nombre d'heures travaillées en cas d'activité réduite (pour les personnes inscrites en catégorie B ou C), etc. Dans la mesure où l'enquête Emploi est représentative de la population et si l'appariement était techniquement parfait, l'enquête appariée devrait dénombrer, en 2017, 3,4 millions de demandeurs d'emplois de catégorie A en France métropolitaine, comme dans le fichier Historique. En pratique, elle ne permet d'en dénombrer que 2,7 millions, soit un « taux de couverture » de l'appariement estimé à 81 % en catégorie A et à 84 % sur l'ensemble des catégories de demandeurs d'emploi.

Malgré ses imperfections, l'appariement permet tout d'abord d'établir qu'en 2017, au moins **1,5 million de personnes sont à la fois au chômage au sens du BIT et inscrites à Pôle emploi en catégorie A**, tandis qu'au plus **1,1 million de personnes sont chômeuses au sens du BIT sans être inscrites à Pôle emploi en catégorie A** et qu'au moins **1,2 million de personnes inscrites à Pôle emploi en catégorie A sont considérées comme inactives ou en emploi au sens du BIT**.

L'appariement permet d'étudier de façon détaillée les deux populations causant l'écart entre les deux agrégats : **les demandeurs d'emploi de catégorie A non classés comme chômeurs au sens du BIT** d'une part, et **les chômeurs au sens du BIT qui ne sont pas inscrits en catégorie A à Pôle emploi** d'autre part.

- En 2017, au moins **1,2 million de personnes sont ainsi inscrites en catégorie A sans être au chômage au sens du BIT** : 0,5 million, soit 20 % des inscrits en catégorie A, sont dans le halo autour du chômage, 0,4 million, soit 16 % des inscrits, sont inactifs hors halo et enfin 0,2 million, soit 9 % des inscrits, sont en emploi.
 - Les 0,5 million de personnes dans le halo autour du chômage et inscrites en catégorie A sont le plus souvent des personnes ne respectant pas les critères de recherche ou de disponibilité pour être classées au chômage au sens du BIT pour des raisons ponctuelles (vacances, déménagement, courte maladie, découragement passager).
 - Les 0,4 million d'inactifs hors halo inscrits en catégorie A incluent notamment un grand nombre de seniors proches de la retraite et souvent éloignés du marché du travail pour raisons de santé.
 - Quant aux 0,2 million de personnes en emploi classées en catégorie A, il s'agit principalement de situations transitoires qui pourraient s'expliquer d'une part par des manières différentes de définir le fait d'être en emploi entre les deux sources, et d'autre part par des retards ou erreurs de déclarations, dans l'enquête ou à Pôle emploi.
- Inversement, au plus **1,1 million de chômeurs au sens du BIT ne sont pas inscrits en catégorie A** : 0,8 million de chômeurs ne sont pas inscrits à Pôle emploi au moment de l'enquête et 0,3 million sont inscrits mais dans d'autres catégories, principalement B ou D.
 - Sur les 0,8 million de chômeurs BIT non retrouvés parmi les inscrits à Pôle emploi, 0,4 million, soit 14 % des chômeurs, ont été inscrits au moins un mois sur la période 2012-2017 et sont appariés, mais ne sont pas inscrits au moment où ils sont interrogés dans l'enquête. Il ne semble pas s'agir de demandeurs d'emploi ponctuellement désinscrits puisque seuls un quart d'entre eux étaient inscrits le mois juste avant ou le mois juste après. Par ailleurs, 0,2 million, soit 8 % des chômeurs BIT, semblent être réellement non connus de Pôle emploi, souvent des étudiants cherchant un emploi à temps partiel en parallèle de leurs études. Finalement, seuls 0,2 million correspondraient à des échecs d'appariement.
 - Les 0,3 million de chômeurs inscrits en catégorie B et D s'expliqueraient quant à eux par des périodes de référence différentes entre Pôle emploi, qui observe l'activité sur l'ensemble d'un mois, et l'enquête Emploi, qui l'observe sur une semaine de référence au sein de ce mois.

L'appariement permet enfin d'analyser l'évolution de l'écart entre les deux mesures, le chômage au sens du BIT mesuré par l'enquête Emploi, et le nombre d'inscrits à Pôle emploi en catégorie A, qui a crû de 0,4 million de personnes entre 2013 et 2017.

- Sur cette période, le nombre **d'inactifs au sens du BIT inscrits en catégorie A** a augmenté de 0,3 million. La hausse de cette population, pour moitié dans le halo et pour moitié hors halo, est celle qui a le plus contribué à accroître l'écart entre chômage BIT et demandeurs d'emploi en catégorie A.
 - En ce qui concerne les **inactifs BIT hors halo inscrits en catégorie A**, une grande partie de la hausse de cette population s'expliquerait par le recul de l'âge de départ à la retraite et la fin du dispositif de dispense de recherche d'emploi. La part des seniors (50-64 ans) dans cette population a en effet fortement augmenté, de 38 % en 2013 à 46 % en 2017.
 - **Dans le halo, les inactifs BIT et inscrits en catégorie A** en 2017 ont des profils sociodémographiques très proches de ceux observés en 2013. Ils semblent toutefois un peu plus éloignés du marché du travail, puisque plus souvent non disponibles et n'ayant pas fait de recherches actives d'emploi.
 - Sur la période 2013-2017, Pôle emploi a simplifié les processus d'inscriptions (inscription sur internet, actualisation par SMS). Ces facilités pour s'inscrire et s'actualiser pourraient avoir contribué à accroître le nombre d'inscrits à Pôle Emploi, en bénéficiant à la fois à des chômeurs et à des personnes un peu plus éloignées du marché du travail. De plus, sur cette même période, le nombre d'embauches en CDD a fortement augmenté, ce qui a pu conduire davantage de personnes à demeurer inscrites à Pôle Emploi pour des périodes prolongées, qu'elles soient en emploi ou entre deux périodes d'emploi.

- Par ailleurs, entre 2013 et 2017, le nombre de **chômeurs BIT non inscrits à Pôle emploi** a reculé (-0,1 million, surtout concentrés entre 2016 et 2017), contribuant également à accroître l'écart entre chômage BIT et inscrits en catégorie A. La croissance économique et les créations d'emploi de 2017 ont en effet beaucoup profité aux jeunes, lesquels sont plus souvent chômeurs non inscrits.

Tableau : part des statuts BIT en fonction de la catégorie d'inscription, en 2017

	Inscrits en cat. A	Inscrits en cat. B, C ou E	Inscrits en cat. D
Chômage BIT	56%	10%	23%
Emploi BIT	9%	87%	15%
Halo	20%	2%	30%
Inactivité hors halo	16%	1%	33%
Ensemble	100%	100%	100%

Lecture : en 2017, parmi les personnes de l'enquête emploi qui ont pu être appariées avec le Fichier Historique et qui sont inscrites en catégorie A au moment de l'enquête, 56 % sont au chômage au sens du BIT, 9 % sont en emploi au sens du BIT, etc.

Champ : France métropolitaine, 15-64 ans, ménages ordinaires, personnes de l'enquête Emploi appariées avec le Fichier Historique et inscrites à Pôle emploi au moment de l'enquête.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Main results

On average, in 2017, in metropolitan France, 2.6 million people aged 15 to 64 are unemployed according to the International Labour Organization (ILO) definition whereas 3.4 million are registered with the public employment service ("Pôle emploi") and classified as "category A" job seekers. The gap between these two benchmark measures, which are not based on the same concepts, widened over the period 2012-2017, from 0.3 million in 2012 to 0.8 million in 2017.

The Labour Force Surveys from 2012 to 2017 (excluding overseas departments) make it possible to count the ILO-based unemployment. They have been matched to the Pôle emploi jobseekers' Historical File, which compiles registered unemployment over the same period. Matching the two sources is about finding people from the Labour Force Survey in the Historical File based on their first name, sex, date of birth and address(es) over the period 2012-2017 (so-called "statistical" matching). This then makes it possible to enrich the Labour Force Surveys with variables from the Historical File: dates of registration, category of jobseeker (A, B, C, D, E), number of hours worked in case of reduced activity (for persons registered in category B or C), etc. Insofar as the Labour Force Survey is representative of the population and if the matching was technically perfect, the matched survey should allow to count, in 2017, 3.4 million A category jobseekers in metropolitan France, as in the Historical File. In practice, it only allows for a count of 2.7 million, i.e. a "coverage rate" of the matching estimated to 81% in category A and 84% for all categories of job seekers.

Despite its shortcomings, the matching first shows that in 2017, at least **1.5 million people are simultaneously unemployed according to the ILO definition and category A jobseekers**, while at most **1.1 million people are unemployed according to the ILO without being category A jobseekers** and at least **1.2 million category A jobseekers are not unemployed according to the ILO** (economically inactive or employed).

The matching makes it possible to study precisely the two populations that cause the gap between the two aggregates: **jobseekers in category A who are not classified as unemployed according to the ILO** on the one hand, and **unemployed persons according to the ILO who are not registered in category A** on the other hand.

- In 2017, at least **1.2 million category A jobseekers are not unemployed according to the ILO**: 0.5 million, or 20% of people registered in category A, are in the "halo of unemployment", 0.4 million, or 16% of people registered, are economically inactive (excluding halo) and finally 0.2 million, or 9% of people registered, are employed.
 - The 0.5 million people in the halo of unemployment and registered in category A are most often people who do not meet the search or availability criteria to be classified as ILO-based unemployed for short-term reasons (holidays, moving, short illness, temporary discouragement).
 - The 0.4 million other inactive people registered in category A include a large number of seniors who are close to retirement and often away from the labour market for health reasons.
 - As for the 0.2 million people in employment classified in category A, these are mainly transitional situations that could be explained on the one hand by different ways of defining the status of employment between the two sources, and on the other hand by delays or errors in declarations, in the survey or at Pôle emploi.
- Conversely, at most **1.1 million unemployed persons according to the ILO are not registered in category A**: 0.8 million unemployed persons are not registered in Pôle emploi at the time of the survey and 0.3 million are registered but in other categories, mainly B or D.
 - Of the 0.8 million unemployed people not registered, 0.4 million were registered at least one month over the 2012-2017 period and are matched, but are not registered at the time they are surveyed. These do not appear to be occasional unsubscribed jobseekers since only a quarter of them were registered in the month just before or the month just after. In addition, 0.2 million seem to be ILO-based unemployed people who are really absent from Pôle emploi registers, particularly students seeking part-time work in parallel with their studies. Finally, the remaining 0.2 million would correspond to mismatches.
 - The 0.3 million unemployed registered in categories B and D are explained by different reference periods between Pôle emploi, which observes activity over the whole month, and the Labour Force Surveys, which observes it over a reference week within that month.

Finally, the matching allows to understand the evolution of the gap between the two measures, the ILO-based unemployment and the number of category A jobseekers, which increased by 0.4 million people between 2013 and 2017.

- Over this period, the number of **economically inactive persons according to the ILO definition and registered in category A** increased by 0.3 million. The increase in this population, half in and half out of the halo, is the one that has contributed most to widening the gap between ILO-based unemployment and category A jobseekers.
 - Regarding the **other inactive persons (non-halo) registered in category A**, a large part of the increase in this population is explained by the increase in the retirement age and the end of the job-search exemption scheme. The share of older people (50-64 years old) in this population has indeed increased significantly, from 38% in 2013 to 46% in 2017.
 - **In the halo, inactive registered in category A** in 2017 have socio-demographic profiles very similar to those observed in 2013. However, they seem to be a little further away from the labour market, as they are more often unavailable and have not actively sought employment.
 - Over the 2013-2017 period, Pôle emploi has simplified the registration processes (online registration, SMS updates). These facilities for registering and updating may have contributed to increasing the number of people registered at Pôle Emploi, benefiting both the unemployed and people who are a little further away from the labour market. In addition, over the same period, the number of hires on fixed-term contracts increased significantly, which may have led more people to remain registered at Pôle Emploi for extended periods, whether they are in employment or between two periods of employment.
- Furthermore, between 2013 and 2017, the number of **ILO-based unemployed not registered in Pôle emploi** fell (-0.1 million, mainly concentrated between 2016 and 2017), also contributing to an increase in the gap between ILO-based unemployment and registered jobseekers in category A. Economic growth and job creation in 2017 have greatly benefited young people, who are more often unregistered unemployed.

Table: Share of ILO statutes by registration category of public employment services, 2017

	Registered in cat. A	Registered in cat. B, C or E	Registered in cat. D
Unemployment (ILO)	56%	10%	23%
Employment (ILO)	9%	87%	15%
Halo of unemployment	20%	2%	30%
Other economically inactive	16%	1%	33%
Total	100%	100%	100%

Reading: in 2017, among the persons in the Labour Force Surveys who may have been matched with the "Historical File" of public employment service (Pôle emploi), and who are registered in category A at the time of the survey, 56% are unemployed according to the ILO, 9% are in employment according to the ILO, etc.

Scope: Metropolitan France, people aged 15 to 64, ordinary households, persons in the Labour Force Surveys matched with the Historical File and registered in the public employment service at the time of the survey.

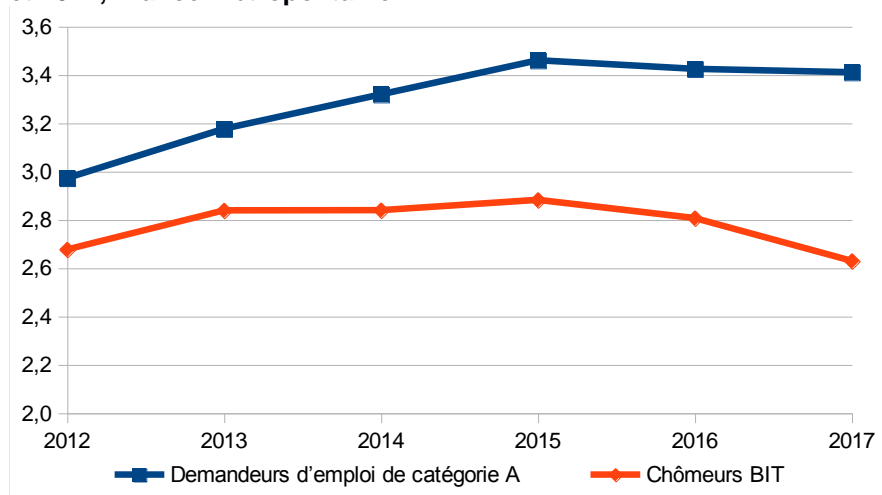
Source: Dares, INSEE, Pôle emploi, matching between the Historical File from Pôle emploi and the Labour Force Surveys.

Introduction

Quelques définitions

Un **chômeur au sens du Bureau international du travail (BIT)** est une personne de 15 ans ou plus, n'ayant pas eu d'activité rémunérée pendant la semaine de référence sur laquelle porte l'enquête, qui serait disponible pour occuper un emploi dans les 15 jours et qui a recherché activement un emploi au cours des quatre semaines précédant et incluant la semaine de référence (ou en a trouvé un qui commence dans moins de trois mois). Le chômage au sens du BIT est mesuré chaque trimestre par l'enquête Emploi en continu (EEC), qui interroge un échantillon représentatif de la population. En 2017, en moyenne sur l'année, on dénombre, en France métropolitaine, 2,6 millions de chômeurs au sens du BIT (cf. graphique 0.1).

Graphique 0.1 : Nombre moyen de chômeurs BIT et de demandeurs d'emploi de catégorie A entre 2012 et 2017, France métropolitaine.



En millions de personnes, en moyenne sur l'année

Champ : France métropolitaine, 15-64 ans.

Source : Fichier Historique 2012-2017 pour le nombre de demandeurs d'emploi en catégorie A, enquêtes Emploi en continu 2012 à 2017 pour le nombre de chômeurs BIT.

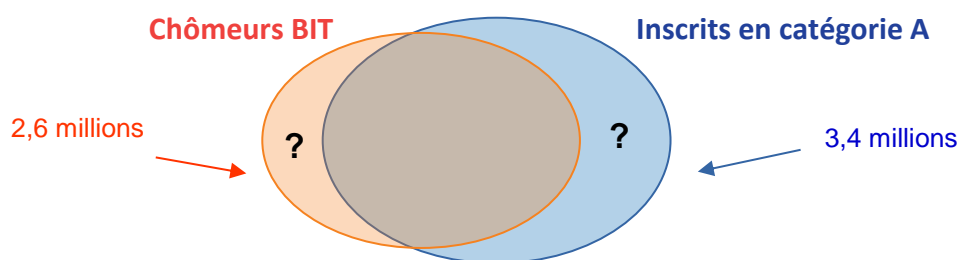
L'enquête permet d'identifier aussi les personnes en emploi au sens du BIT (personnes ayant eu au moins une heure d'activité rémunérée au cours de la semaine de référence, ou absentes de leur emploi pour un certain nombre de raisons répertoriées, congés, etc.), et par solde, les personnes inactives (ni en emploi, ni au chômage). Parmi les inactifs, on distingue notamment le halo autour du chômage : il s'agit de personnes inactives et qui souhaitent travailler, mais ne remplissent pas les critères pour être considérées comme au chômage au sens du BIT (non disponibles, ou n'ayant pas fait de recherches actives).

L'inscription sur les listes de Pôle emploi (PE) résulte quant à elle d'une démarche administrative. Le nombre mensuel de **demandeurs d'emploi (DE)** inscrits à Pôle emploi est issu des fichiers de gestion produits par Pôle emploi et la Dares, le service statistique du ministère du travail. Les inscrits en catégorie A regroupent les personnes sans emploi tenues d'effectuer des actes positifs de recherche. C'est la catégorie conceptuellement la plus proche de celle du chômage au sens du BIT et elle est particulièrement observée pour suivre les évolutions conjoncturelles du nombre de chômeurs. En 2017, en France métropolitaine, on dénombre 3,4 millions de demandeurs d'emplois inscrits en catégorie A en moyenne sur l'année et dont l'âge est compris entre 15 et 64 ans.

Les demandeurs d'emploi peuvent aussi être inscrits à Pôle emploi dans d'autres catégories que la A. Les catégories B et C correspondent aux demandeurs d'emploi ayant exercé une activité réduite au cours du mois (moins de 78 heures de travail pour la catégorie B et plus de 78 heures pour la catégorie C). La catégorie D regroupe les demandeurs d'emploi sans emploi non tenus de faire des recherches actives (en raison d'un stage, d'une formation, d'un arrêt maladie...). Enfin, la catégorie E correspond à des demandeurs d'emploi non tenus de faire des actes positifs de recherche, car déjà en emploi (contrats aidés par exemple). Par analogie avec les statuts BIT, les catégories B, C et E correspondraient à des personnes en emploi, tandis que la catégorie D correspondrait plutôt à des inactifs (personnes en formation, en arrêt maladie...).

Les deux concepts de chômeur au sens du BIT et d'inscrits en catégorie A sont proches, mais ne se recoupent pas totalement. Ainsi, un demandeur d'emploi inscrit en catégorie A peut ne pas être classé comme chômeur au sens du BIT dans l'enquête, par exemple s'il n'est pas disponible immédiatement à la date de l'enquête pour travailler, en raison de vacances, d'une maladie, d'un déménagement, etc. Inversement, un chômeur au sens du BIT n'a aucune obligation de s'inscrire à Pôle emploi. Notamment, les jeunes qui viennent de finir leurs études et se mettent à chercher un emploi ne s'inscrivent pas toujours à Pôle emploi, lorsqu'ils n'ont pas encore ouvert de droits à indemnisation. L'appariement des données individuelles de ces deux sources va permettre de bien identifier ces différentes populations et de les quantifier (cf. graphique 0.2). Il permettra ainsi de mieux comprendre l'écart entre les deux concepts de chômage, lequel a parfois fluctué au cours des années passées (cf. graphique 0.3).

Graphique 0.2 : représentation schématique des chômeurs BIT et des inscrits en catégorie A en 2017

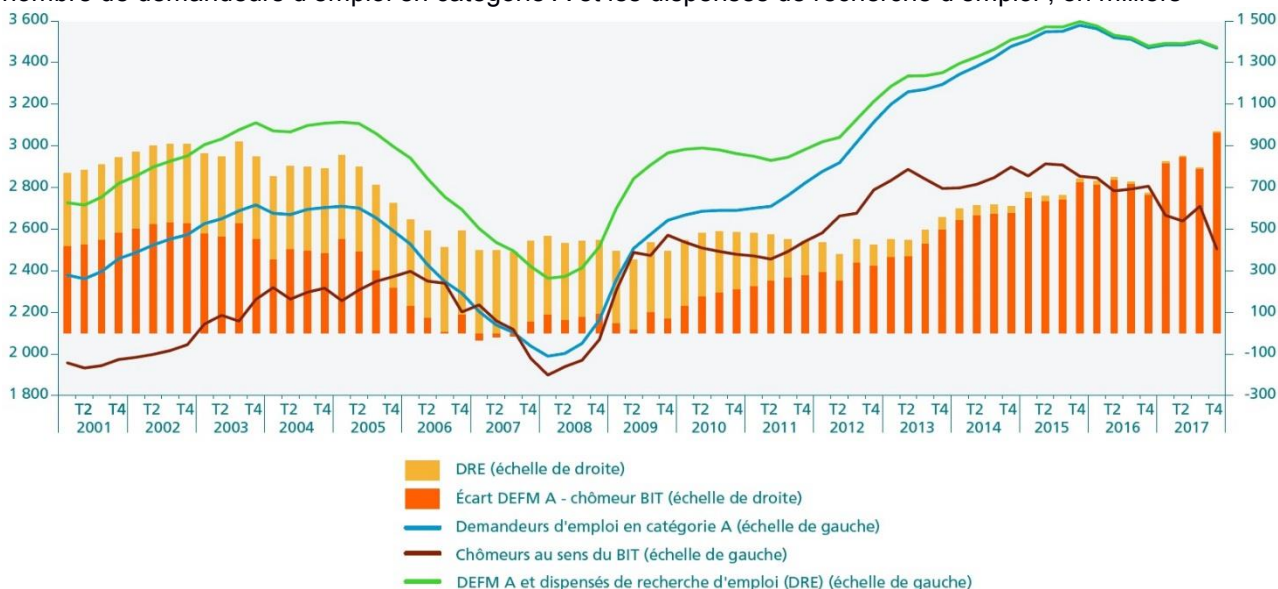


Champ : France métropolitaine, 15-64 ans, année 2017

Source : Fichier Historique pour le nombre d'inscrits en catégorie A, enquêtes Emploi en continu pour le nombre de chômeurs BIT.

Graphique 0.3 : demandeurs d'emploi en catégorie A, dispensés de recherche d'emploi et chômeurs au sens du BIT

Données en moyenne trimestrielle CVS pour le nombre de chômeurs au sens du BIT et CVS-CJO pour le nombre de demandeurs d'emploi en catégorie A et les dispensés de recherche d'emploi ; en milliers



Lecture : au 1^{er} trimestre 2010, on comptait 2 537 000 chômeurs au sens du BIT, 2 668 000 demandeurs d'emploi inscrits à Pôle emploi en catégorie A et 315 000 dispensés de recherche d'emploi. L'écart entre les inscrits (respectivement les inscrits ou dispensés de recherche d'emploi) et les chômeurs était de 131 000 (respectivement 446 000).

Note : ce graphique est issu du Dares Analyses n°031, « Emploi, chômage, population active en 2017 : nouvelle accélération de l'emploi salarié privé et amplification de la baisse du chômage », juillet 2018.

Champ : population des ménages (chômeurs au sens du BIT) ; ensemble de la population y compris les personnes vivant en collectivité (inscrits en catégorie A et dispensés de recherche d'emploi) ; France métropolitaine.

Source : Pôle emploi-Dares, STMT ; Insee, enquêtes Emploi.

Les données

Les enquêtes Emploi 2012 à 2017 sont mobilisées. L'enquête Emploi est réalisée par l'Insee depuis 1950 annuellement d'abord, et en continu en France métropolitaine depuis 2003. Les enquêtes 2012 à 2017 sont ici restreintes au champ de la France métropolitaine, car les départements-régions d'Outre-mer (DROM) n'ont été inclus dans l'enquête en continu qu'à partir de 2014. Pour l'appariement, le champ est également restreint aux personnes de 15 à 64 ans. Enfin, l'enquête est réalisée auprès des ménages ordinaires. Les chômeurs qui vivent hors des ménages ordinaires (en communautés, habitations mobiles, sans abris...) semblent toutefois marginaux : en 2015, hors ménages ordinaires, 45 000 personnes seulement se déclarent spontanément au chômage d'après le recensement de la population.

L'enquête Emploi est un panel rotatif : les logements enquêtés le sont 6 trimestres consécutifs. Cette dimension panel permet ainsi d'illustrer les transitions dans les parcours des personnes. Au global, en France métropolitaine et pour les personnes âgées de 15 à 64 ans, les Enquête emploi en continu 2012 à 2017 contiennent 1,8 million d'observations, correspondant à 0,45 million d'individus distincts (un individu répond donc en moyenne 4 fois à l'enquête, et non pas 6 fois, en raison de non-réponses à certaines ou interrogations, ou de déménagements).

Le fichier Historique de Pôle emploi est une base de données administratives recensant toutes les personnes ayant été inscrites à un moment ou un autre à Pôle emploi sur la période 2012-2017. Il est constitué de deux bases : la base des individus et la base des demandes. La base des individus recense l'ensemble des individus inscrits avec leur catégorie de demandeur d'emploi au mois le mois de 2012 à 2017. 16,9 millions d'individus distincts ont été inscrits à Pôle emploi sur cette période. La base des demandes contient, quant à elle, une ligne par période d'inscription à Pôle emploi de ces 16,9 millions d'individus. Au total, cette base contient 70 millions d'observations, avec pour chaque période d'inscription la date de début et de fin d'inscription, et l'adresse de l'individu lors de cette période.

Un certain nombre de variables en commun entre les deux fichiers permettent de rapprocher les individus. Certaines variables sont considérées comme « les plus robustes » pour l'appariement au sens où elles sont définies et codées à peu près de la même manière dans l'enquête et le Fichier Historique et qu'elles sont peu sujettes à des erreurs. Il s'agit du sexe, de la date de naissance (jour, mois, année) et de la commune de résidence.

Les deux autres variables disponibles pour rapprocher les individus, l'adresse et le prénom, sont plus difficiles à comparer entre l'enquête et le Fichier Historique, car elles ne sont pas forcément renseignées de la même manière et ne répondent pas à la même logique dans les deux bases. Dans le Fichier Historique, l'adresse est l'adresse postale telle qu'elle est déclarée par le demandeur d'emploi, tandis que dans l'enquête Emploi, il s'agit de l'adresse fiscale, issue de la taxe d'habitation dans laquelle l'échantillon de l'enquête a été tiré (par exemple, une adresse est identifiée par un lieu-dit dans le Fichier, tandis qu'elle est identifiée avec un nom de voie dans la taxe d'habitation, donc dans l'enquête Emploi).

Quant au prénom, dans le Fichier Historique, le prénom officiel de l'individu est renseigné (prénom(s) de l'état civil). Dans l'enquête Emploi, on collecte plutôt un prénom d'usage, afin d'identifier la personne au sein du foyer, pour poser plus facilement le questionnaire et identifier simplement les liens entre les individus qui composent le ménage. Les enquêtés déclarent donc parfois des surnoms ou des diminutifs. De plus, des erreurs ou des variantes dans la façon d'orthographier les prénoms par les enquêteurs peuvent rendre plus difficile l'exercice de comparaison.

Chapitre 1- La méthode d'appariement

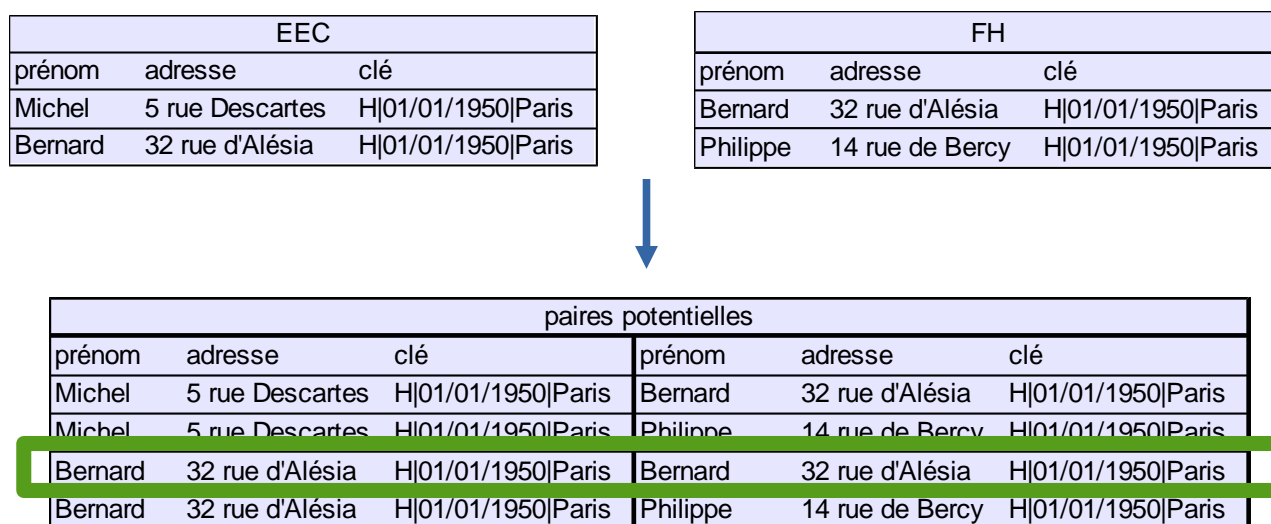
Le Fichier Historique est un fichier exhaustif, tandis que l'enquête Emploi est une base d'enquête. Ici, appairer les deux bases de données signifie donc enrichir l'enquête Emploi par de nouvelles variables issues du Fichier Historique (dates d'inscription à Pôle emploi, catégorie, etc.).

1-1 Algorithme général de l'appariement

Tout d'abord, les variables communes et les plus robustes (sexe, date de naissance, commune de résidence) sont utilisées pour créer une « clé de hachage ». Cela signifie que ne peuvent être appariés que deux personnes (l'une issue du Fichier Historique et l'autre de l'enquête Emploi) qui partagent le même sexe, la même date de naissance et la même commune de résidence.

Toutes les « paires potentielles » (paires ayant en commun la clé de hachage) sont ainsi créées (cf. graphique 1.1). La comparaison des prénoms et adresses permet enfin d'accepter ou de rejeter les paires potentielles.

Graphique 1.1 : représentation schématique de l'algorithme de l'appariement



Comme l'objectif de l'appariement est d'étudier les individus pour lesquels le statut au sens du BIT diverge de celui sur les listes de Pôle emploi (ex : inscrits en catégorie A non classés comme chômeurs au sens du BIT), on cherche ici à limiter au maximum les « faux-positifs », c'est-à-dire les individus appariés à tort. En effet, pour un individu apparié à tort, les statuts BIT et Pôle emploi sont fortement susceptibles de diverger puisqu'il ne s'agit pas de la même personne dans les deux bases. Un appariement à tort parasiterait donc les résultats sur les populations présentant le plus d'intérêt pour l'étude.

1-2 Comment comparer les prénoms ?

Un travail préliminaire de normalisation des prénoms a tout d'abord été effectué, dans les deux bases : les caractères spéciaux sont retirés, les accents ainsi que les lettres en double sont supprimées, de façon à uniformiser l'écriture (exemple : Rémi → Remi, Matthieu → Mathieu, etc.).

Les difficultés rencontrées dans la comparaison des prénoms entre les deux bases

- Certains prénoms peuvent s'écrire de différentes façons, et l'enquêteur choisit d'orthographier selon une de ces déclinaisons lorsqu'il saisit le prénom dans son outil de collecte. Exemple : Kristel (FH) orthographiée Christelle (EEC), Mostefa (FH) orthographié Mustapha (EEC), Laurène (FH) orthographiée Lorraine (EEC)...
- Certaines personnes déclarent à Pôle emploi un prénom « francisé » différent du prénom d'usage qu'ils déclarent à l'enquêteur de l'EEC (ou inversement). Exemple : Joseph (FH) ↔ Giuseppe (EEC), Baudoin (FH) ↔ Boudewijn (EEC)...
- Certains prénoms sont rares et inconnus de l'enquêteur qui se débrouille alors pour les orthographier tels qu'il les a entendus. Ex : Symphorien (FH) ↔ SaintFlorian (EEC)...
- Enfin, les enquêtés peuvent déclarer dans l'EEC des surnoms, diminutifs ou contractions de prénoms : ex : Catherine (FH) ↔ Cathy (EEC), Maria Isabelle (FH) ↔ Maribel (EEC)...

Des définitions de distance pour comparer les prénoms

Distance « scripturale » : comparer la façon dont les prénoms s'écrivent

Pour comparer les prénoms, on utilise la distance dite « de Levenshtein ». Il s'agit du nombre minimum d'opérations (insertion, suppression, remplacement...) qui permettent de passer du prénom du Fichier Historique au prénom enquête Emploi.

Par exemple : Stéphan (FH) ↔ Stéphane (EEC) : distance Levenshtein = 1 (ajout d'une lettre « E »). La distance de Levenshtein est implémentée sous SAS (fonction « Compged », voir annexe 1.1).

Cette distance est ensuite pondérée par la longueur du prénom : $\text{Longueur}(\text{prénom}) = \min(\text{longueur}_{\text{prénomFH}} ; \text{longueur}_{\text{prénomEEC}})$. L'idée sous-jacente est que plus un prénom est long, plus une erreur d'orthographe est susceptible de se produire.

Par exemple : Stéphan ↔ Stéphane: distance scripturale = 1/7

La distance scripturale fournit ainsi une sorte de « pourcentage de dissemblance » entre les écritures des deux prénoms (1/7, soit 14 % dans l'exemple précédent)

Distance « anagramme » : comparer les lettres qui composent les prénoms

Chaque prénom est converti en vecteur des lettres qui le composent, sans tenir compte de l'ordre des lettres. Ainsi, le prénom Jeanne par exemple s'écrit avec une fois la lettre A, 0 fois les lettres B, C et D, 2 fois la lettre E, etc. Jeanne → (1 0 0 0 2 ...) x (A B C D E...). Puis les vecteurs de lettres sont comparés.

Cette fonction permet d'associer une distance nulle à deux prénoms qui sont anagrammes l'un de l'autre, ce qui est utile dans le cas par exemple où plusieurs prénoms sont renseignés dans le Fichier Historique et l'enquête, mais dans un ordre différent. Exemple : Paule Marie (FH) ↔ Marie Paule (EEC). La distance anagramme est ici nulle alors que la distance scripturale serait élevée.

La distance anagramme est elle aussi pondérée par la longueur des prénoms ($\text{Longueur}(\text{prénom}) = \min(\text{longueur}_{\text{prénomFH}} ; \text{longueur}_{\text{prénomEEC}})$) de façon à s'interpréter là aussi comme un pourcentage de dissemblance.

Distance « phonétique » : comparer la façon dont les prénoms se prononcent

Les prénoms du Fichier Historique et de l'enquête Emploi sont préalablement convertis en un code phonétique correspondant à la façon dont ils se prononcent. Ici, la fonction SAS « Soundex » a été utilisée (cf. annexe 1.1). Cette fonction renvoie un code phonétique basé sur les consonnes qui composent le mot.

Puis les codes phonétiques sont comparés avec la distance scripturale. À nouveau, on pondère la distance obtenue par la longueur des prénoms. Cette approche phonétique permet de relier des prénoms qui peuvent avoir été orthographiés de façon assez différente mais ont une prononciation proche (exemple : Mostefa et Mustapha). Elle permet de « rattraper » des prénoms mal orthographiés parce qu'ils sont rares ou d'origine étrangère.

Une distance synthétique sur les prénoms qui combine les trois distances précédentes

Les trois distances précédentes sont préalablement normalisées puis combinées pour constituer une unique distance synthétique de comparaison des prénoms :

$\text{Distance_prénom} = 1/3 \text{ distance scripturale} + 1/3 \text{ distance anagramme} + 1/3 \text{ distance phonétique}$

Dans le Fichier Historique, plusieurs prénoms peuvent être déclarés, tandis que dans l'enquête Emploi, en général, un seul prénom est fourni. Plusieurs distance_prénom sont alors calculées de façon à maximiser les chances d'appariement (cf. annexe 1.2).

1-3 Comment comparer les adresses ?

Difficultés pour la comparaison des adresses

La principale difficulté pour comparer les adresses de l'enquête Emploi et du Fichier Historique est la différence entre le référentiel « postal » dans lequel est fournie l'adresse à Pôle emploi et le référentiel « fiscal » pour les adresses de l'enquête. On rencontre notamment les cas suivants :

- Des immeubles, résidences ou cités qui donnent sur plusieurs voies, ou des maisons à l'angle de deux rues, pour lesquels les adresses postales et fiscales ne sont pas les mêmes (l'adresse postale est sur une voie, tandis que l'adresse fiscale est sur une autre).
- Un même logement sous deux adresses différentes (un nom de voie dans la taxe d'habitation et un lieu-dit ou un nom de maison pour l'adresse postale par exemple).

De plus, les adresses de l'enquête ayant été tirées dans la taxe d'habitation, elles ont bénéficié d'une certaine validation/certification par les services fiscaux. Leurs libellés présentent donc moins d'erreurs que les adresses du Fichier Historique, qui ont été simplement déclarées par les inscrits à Pôle emploi.

Distance entre les libellés d'adresses : comparer la façon dont les adresses s'écrivent

La première manière de comparer les adresses est de regarder si elles s'écrivent de la même façon. Comme pour la comparaison des prénoms, une distance de Levenshtein entre les libellés d'adresses (numéro, type et libellé de voie) est calculée, pondérée par le nombre de caractères du libellé d'adresse. Cette distance fournit ainsi une sorte de pourcentage de divergence entre les libellés d'adresse du Fichier Historique et de l'enquête pour une paire potentielle donnée.

Distance géographique : comparer la localisation géographique des adresses

Les adresses ont été soumises à géolocalisation (cf. annexe 1.3). Tous les libellés d'adresses et de communes, qu'ils soient issus du Fichier Historique ou de l'enquête Emploi, ont été traités par la division statistique et analyses urbaines de l'Insee qui fournit les coordonnées longitude/latitude (x,y) de l'adresse si celle-ci est retrouvée dans son référentiel géolocalisé.

Pour chaque paire potentielle, une distance géographique peut alors être calculée comme la distance euclidienne entre les coordonnées (x,y) de l'adresse du Fichier Historique et celles de l'adresse issue de l'enquête Emploi .

Distance géographique = $\sqrt{(x_{\text{eec}} - x_{\text{fh}})^2 + (y_{\text{eec}} - y_{\text{fh}})^2}$, où x_{eec} et y_{eec} sont les coordonnées de l'adresse de l'enquête Emploi et x_{fh} et y_{fh} celles de l'adresse du Fichier Historique.

Cette distance ne peut toutefois être calculée que si la géolocalisation a réussi au sein de la commune, c'est à dire ici si la voie (au minimum) a pu être identifiée (cf. annexe 1.3).

Distance synthétique sur les adresses

Les deux distances (libellés d'adresse et distance géographique) se complètent bien. En effet, la distance géographique peut venir nuancer la simple comparaison des libellés.

Exemple 1 : « 8 rue de Grenelle » et « 8 bd de Grenelle » : la distance géographique élevée (3 km) indique que ce ne sont pas les mêmes adresses, même si les libellés sont pourtant très proches

Exemple 2 : « 16 rue Victor Hugo » et « 2 rue Diderot » : la distance géographique entre les deux adresses est très faible (60 mètres) alors que les libellés ne se ressemblent pas du tout. Il peut s'agir d'un immeuble donnant sur deux voies par exemple.

Quant à la distance sur les libellés, elle a l'avantage d'être toujours disponible, même en cas d'échec de géolocalisation.

On construit ainsi une distance synthétique sur les adresses qui combine les 2 distances élémentaires décrites ci-dessus (et préalablement normalisées) :

distance adresses =

- 50 % distance libellés + 50 % distance géographique, si la géolocalisation a réussi ;
- distance libellés seulement, si échec de géolocalisation

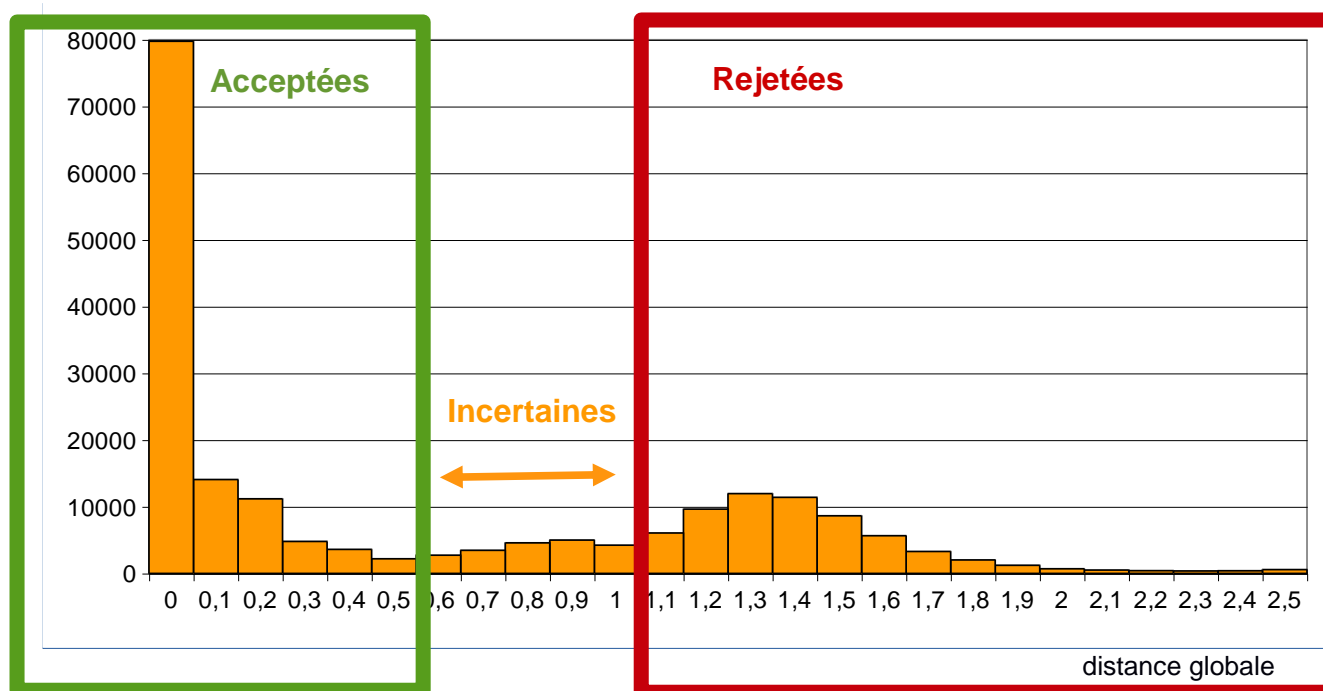
1-4 Règle de décision pour accepter ou rejeter un appariement

Une dernière distance est calculée, qui synthétise la proximité des prénoms et des adresses :
Distance globale = $\frac{2}{3}$ * distance adresse + $\frac{1}{3}$ * distance prénom

Il semble préférable de combiner les distances sur l'adresse et le prénom plutôt que de se donner un seuil pour chacune d'entre elles. En effet, lorsqu'une des deux distances élémentaires (prénom ou adresse) vaut 0, cela permet de relâcher la contrainte sur l'autre distance. Par exemple, si pour une paire potentielle les deux adresses concordent parfaitement, alors on peut s'autoriser à être un peu plus souple sur la comparaison des prénoms : le risque d'apparier à tort est déjà très faible puisque les adresses coïncident. Ici, on donne également un peu plus de poids à l'adresse qu'au prénom ($\frac{2}{3}$ vs $\frac{1}{3}$), car l'adresse est plus discriminante. En effet, dans la population d'une commune, on a en général davantage de personnes qui partagent un même prénom que de personnes qui partagent une même adresse.

Sur la période 2012-2017, 1,1 million de paires potentielles sont constituées. Pour un individu de l'enquête Emploi, seule la paire potentielle la plus « probable » est conservée (celle dont la distance globale est minimale, c'est-à-dire l'individu du Fichier Historique dont les prénoms et adresses sont les plus proches de ceux de l'individu de l'enquête Emploi). Symétriquement, on s'assure aussi qu'un individu du Fichier Historique ne peut pas être apparié avec deux personnes distinctes de l'enquête Emploi. Enfin, le fait qu'une personne a eu plusieurs adresses différentes au sein d'une même commune sur la période est aussi source de doublons (cela crée en effet autant de paires potentielles que la personne a eu d'adresses). Après élimination de l'ensemble de ces paires potentielles peu probables et doublons, il reste 200 000 paires potentielles acceptables (cf. graphique 1.2).

Graphique 1.2 : distribution des 200 000 paires potentielles acceptables selon la « distance globale » d'appariement



Note : les classes ont été construites par arrondi des valeurs de distance globale. Par exemple, la classe 0,7 correspond aux paires dont la distance globale s'arrondit à 0,7 (i.e. de 0,650 à 0,749). Les distances globales supérieures à 2,450 ont été regroupées au sein de la classe « 2,5 ».

Champ : France métropolitaine, personnes de 15 à 64 ans, ménages ordinaires.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Le 1^{er} pic de distribution à gauche (zone verte, distance globale inférieure au seuil de 0,6) correspond aux paires pour lesquelles à la fois les prénoms et les adresses concordent parfaitement ou quasiment (cf. tableau 1.3). La distance globale pour ces paires est très faible et on accepte donc ces appariements avec un risque marginal de faux-positifs. 116 000 paires acceptables sont ainsi acceptées.

Tableau 1.3 : exemple de paire potentielle acceptée

FH		EEC		dist_geoloc
prénom	adresse	prénom	adresse	
Bernard	32 rue d'Alésia	Bernard	32 rue d'Alésia	0m

Le 2^e pic de distribution à droite (zone rouge, distance globale supérieure ou égale au seuil de 1,1) correspond essentiellement à des appariements à tort, pour lesquels ni les prénoms ni les adresses ne coïncident (cf. tableau 1.4).

Tableau 1.4 : exemple de paire potentielle rejetée

FH		EEC		dist_geoloc
prénom	adresse	prénom	adresse	prénom
Léa	12 rue des peupliers	Manon	57 Bd Molière	3500m

Enfin, le plateau entre ces deux pics de distribution correspond à une zone d'incertitude (cf. tableau 1.5) avec de nombreux cas litigieux, déjà évoqués précédemment :

- prénoms qui se ressemblent phonétiquement mais avec des écritures différentes
- diminutifs, surnoms
- immeubles donnant sur plusieurs voies (adresses différentes mais géographiquement proches)
- logements ayant une adresse postale et une adresse fiscale différentes...

Mais aussi des appariements à tort :

- jumeaux vivant à la même adresse (prénoms discordants alors que les adresses sont parfaitement identiques)
- personnes distinctes partageant la même date de naissance et le même prénom et géographiquement proches l'une de l'autre en raison d'une grande densité de population au km²...

Tableau 1.5 : exemple de paires potentielles incertaines

FH		EEC		dist_geoloc
prénom	adresse	prénom	adresse	
Joseph	12 rue des peupliers	Giuseppe	12 rue des peupliers	0m
Franck	88 avenue Verdier	Etienne	88 avenue Verdier	0m
Maxime	16 rue Victor Hugo	Maxime	2 rue Diderot	60m
Laurence	5 rue des plantes	Aurélie	156 rue des plantes	900m
Catherine	Résidence Bellevue	Cathy	12 Route de Bellevue	échec geoloc

1-5 Autres critères pour appairier

Exploiter le tirage en grappes de l'enquête

L'enquête Emploi en continu est tirée par grappes : les logements sélectionnés dans l'échantillon sont volontairement proches les uns des autres afin de réduire les temps de trajet des enquêteurs et de maximiser ainsi leurs chances de réaliser tous les entretiens dans le temps imparti. De ce fait, les enquêteurs interrogent plusieurs logements d'un même immeuble ou plusieurs maisons d'une même rue, ce qui permet d'identifier les immeubles ou maisons qui possèdent deux adresses différentes au sens des référentiels postaux et fiscaux.

En effet, le fait de trouver plusieurs paires potentielles qui partagent exactement la même discordance entre adresses du Fichier Historique et de l'enquête Emploi est le signe qu'il s'agit pourtant bien de la même localisation (voir exemples en tableau 1.6). Cette méthode permet d'accepter 5 000 paires supplémentaires, dont 70 % situées dans la zone d'incertitude (distance globale comprise entre 0,6 et 1,1) et 30 % dans la zone de rejet (distance globale supérieure à 1,1).

Tableau 1.6 : p-uplet de paires potentielles partageant la même discordance entre adresses du Fichier Historique et adresses de l'enquête Emploi

FH		EEC		dist_geoloc
prénom	adresse	prénom	adresse	
Mohamed	16 rue Victor Hugo	Mohamed	2 rue Diderot	60m
Eva	16 rue Victor Hugo	Eva	2 rue Diderot	60m
Pauline	16 rue V. Hugo	Pauline	2 rue Diderot	60m
Maxime	16 rue Victor Hugo	Maxime	2 rue Diderot	60m
Evelyne	Les églantiers	Evelyne	116 route de la forêt	échec geoloc
Malik	Les églantiers	Malik	118 route de la forêt	échec geoloc

Appairier en contrôlant la probabilité théorique de faux-positifs

Pour certaines paires d'individus, le prénom est exactement identique alors que l'adresse diffère. On souhaite appairier certaines de ces paires, tout en minimisant la probabilité de réaliser des faux-positifs.

Dans une petite commune, le nombre d'habitants est réduit, si bien qu'il est relativement peu probable que deux personnes partagent la même date de naissance et le même prénom. Cela signifie que même si les adresses divergent, il est plus probable qu'il s'agisse bien de la même personne (avec une adresse fiscale différente de l'adresse postale, ou bien une erreur de déclaration/saisie de l'adresse par exemple si les noms de voies ne sont pas normalisés, etc.) plutôt que deux personnes distinctes. Dans ce cas, il semble légitime d'appairier en se référant uniquement à la date de naissance et au prénom. La question qui se pose est de savoir jusqu'à quel seuil de taille de commune ce raisonnement peut être tenu.

Pour ce faire, on calcule les probabilités théoriques qu'il y ait, pour un individu donné, une autre personne dans la commune qui ait la même date de naissance et le même prénom (voir annexe 1.4 pour le calcul de ces probabilités). On conclut que jusqu'à 4 000 habitants, on maintient une probabilité de faux-positif inférieure à 0,5 % : il semble raisonnable dans ce cas d'appairier au sein de la commune deux individus partageant la même date de naissance et le même prénom, même si les libellés d'adresses divergent.

Dans les communes de plus de 4 000 habitants, un raisonnement similaire est tenu : si dans le Fichier Historique une personne de même prénom et de même date de naissance que l'enquêté est trouvée, et que son adresse issue du Fichier Historique est géolocalisée dans un rayon de 400m autour de l'adresse de l'enquête Emploi, la paire potentielle est acceptée (cf. annexe 1.4 également pour plus de détail).

Ces deux règles basées sur des probabilités théoriques de faux-positifs permettent d'accepter respectivement 3 500 et 2 000 paires supplémentaires, situées là aussi pour 70 % d'entre elles dans la zone d'incertitude et pour 30 % dans la zone de rejet.

Reprise manuelle pour les derniers cas litigieux

Un certain nombre de paires potentielles souffrent du fait que, malgré des prénoms proches, l'une des deux adresses est mal orthographiée ce qui conduit simultanément à un échec de géolocalisation et à une distance élevée dans la comparaison des libellés d'adresses (cf. tableau 1.7). 800 paires présentant ces caractéristiques sont identifiées et traitées au cas par cas. Cette étape de reprise permet de valider environ 300 paires supplémentaires sur les 800 identifiées.

Tableau 1.7 : exemples de paires en échec de géolocalisation

FH		EEC		dist_geoloc
prénom	adresse	prénom	adresse	
Catherine	Résidence Bellevue	Cathy	6 Route de Bellevue	échec geoloc
Sophia	Av du douzième régiment	Sophia	Av du 12 ^e régiment d'infanterie	échec geoloc
Alain	Bois vert	Alain	Les combes du Bois vert	échec geoloc
Jorge Santiago	27 rue P et M Curie	Georges	27 rue Curie	échec geoloc

Trois autres tours d'appariement pour rattraper les erreurs de déclaration sur la date de naissance

Jusque-là, on a supposé les variables de date de naissance (jour, mois, année) parfaitement fiables, en ne s'autorisant ainsi à appairer que des personnes qui partagent cette « clé de hachage » entre l'enquête Emploi et le Fichier Historique. Toutefois, des erreurs dans la déclaration des dates de naissance sont possibles, aussi bien dans l'enquête Emploi que dans le Fichier Historique.

La contrainte sur la clé de hachage est donc relâchée pour chacune des trois variables à tour de rôle (année, mois, jour de naissance). En contrepartie, comme le risque de faux-positifs est plus élevé, l'exigence sur la proximité des prénoms et des adresses est rehaussée. Un seuil d'acceptation est défini pour les prénoms, de façon à n'autoriser que des erreurs d'orthographe minimales. De même, pour les adresses, les libellés doivent être très proches pour que la paire potentielle soit acceptée.

Ces trois nouveaux tours d'appariement permettent d'accepter 3 500 paires supplémentaires (dont 1 000 paires en retirant l'année de naissance de la clé de hachage, 800 paires en retirant le mois de naissance et 1 700 paires en retirant le jour de naissance).

Au final, environ 131 000 paires potentielles sont acceptées (cf. tableau 1.8). Ainsi, sur les 450 000 individus distincts des échantillons des enquêtes Emploi de 2012 à 2017, 131 000 individus ont été retrouvés dans la première table du Fichier Historique, celle contenant les dates de naissance, prénoms et adresses. De plus, 124 000 ont également été retrouvés dans la seconde table du Fichier Historique, celle qui fournit la catégorie d'inscription au mois le mois, et donc, en particulier, la catégorie de l'enquêté au moment de l'enquête. Pour ces 124 000 individus, l'appariement est ici considéré comme réussi. Enfin, le risque de faux-positifs semble contenu, en dessous de 0,6 % (cf. annexe 1.5).

Tableau 1.8 : nombre de paires potentielles acceptées en fonction du critère d'acceptation

Critère	Paires acceptées
Distance globale <0,6	116 437
Grappes	5 027
Commune de moins de 4000 habitants	3 644
Communes >4000 hbts – Rayon de 400m	2 014
Reprise manuelle	316
Clé de hachage hors année de naissance	990
Clé de hachage hors mois de naissance	841
Clé de hachage hors jour de naissance	1 664
Total retrouvés	130 933
Total avec catégorie d'inscription au mois le mois	123 971

Champ : France métropolitaine, 15-64 ans, 2012 à 2017.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

1-6 Taux de couverture et biais d'appariement

Un taux de couverture estimé à 81 % en catégorie A

Comment estimer le succès global de l'algorithme mis en œuvre ? La difficulté est que, au niveau individuel, il n'est pas possible de savoir si une personne n'est pas appariée à *raison*, parce qu'elle n'a jamais été inscrite à Pôle emploi au cours de la période 2012-2017, ou à *tort*, parce qu'elle a été inscrite, mais que l'algorithme présenté dans la première partie n'a pas permis de la retrouver dans le Fichier Historique. Les causes de tels échecs d'appariement peuvent être purement techniques (des erreurs de saisie sur la date de naissance, le prénom, la commune de résidence et l'adresse), ou liées à un écart de concept : si l'adresse transmise à Pôle emploi n'est pas la résidence habituelle, la probabilité d'enquêter la personne à cette adresse dans l'enquête Emploi est nulle.

Au niveau global cependant, un taux de couverture peut être estimé. Si l'appariement était techniquement parfait, dans la mesure où l'enquête Emploi est construite de façon à être représentative de la population française, alors l'enquête appariée devrait permettre de retrouver le nombre total d'inscrits connu via le Fichier Historique (en utilisant les pondérations de l'enquête, chaque individu apparié représentant plusieurs centaines de personnes similaires en termes d'inscription à Pôle emploi, de sexe, âge, etc.).

En procédant ainsi on retrouve, via l'appariement, 83 à 84 % de l'ensemble des demandeurs d'emploi entre 2012 et 2017 (cf. tableau 1.9). Ce taux de couverture estimé est légèrement inférieur en catégorie A, égal à 81 %, tandis qu'il avoisine 90 % pour les catégories B et C (voir encadré 1).

Tableau 1.9 : Nombres d'inscrits issus du Fichier Historique, de l'appariement et taux estimé de couverture, par catégorie et par année (2012-2017)

En millions de personnes, en % pour les taux de couverture

	Catégorie A	Catégorie B	Catégorie C	Catégorie D	Catégorie E	Ensemble
Fichier Historique (1)						
2012	3,0	0,6	0,9	0,3	0,4	5,2
2013	3,2	0,7	0,9	0,3	0,4	5,5
2014	3,3	0,7	1,0	0,3	0,4	5,8
2015	3,5	0,7	1,2	0,3	0,4	6,1
2016	3,4	0,7	1,3	0,3	0,5	6,2
2017	3,4	0,8	1,4	0,3	0,4	6,3
Enquêtes Emploi appariées (2)						
2012	2,3	0,6	0,8	0,2	0,3	4,2
2013	2,5	0,6	0,8	0,3	0,3	4,5
2014	2,7	0,6	0,9	0,3	0,3	4,8
2015	2,8	0,7	1,0	0,3	0,4	5,1
2016	2,8	0,7	1,2	0,3	0,4	5,2
2017	2,7	0,7	1,2	0,3	0,3	5,3
Taux de couverture (3)=(2)/(1)						
2012	78%	91%	89%	86%	88%	82%
2013	80%	88%	89%	89%	82%	83%
2014	81%	88%	87%	91%	83%	83%
2015	81%	90%	89%	87%	89%	84%
2016	81%	89%	91%	85%	81%	84%
2017	81%	89%	89%	88%	78%	84%

Champ : France métropolitaine, 15-64 ans, ménages ordinaires.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Un taux de couverture plus faible pour les hommes et les personnes nées à l'étranger

En plus de la catégorie et de l'année, 5 variables sociodémographiques, présentes à la fois dans le Fichier Historique et dans l'enquête Emploi appariée, sont testées : sexe, âge (en trois tranches), pays de naissance (France ou étranger), région de résidence (13 régions métropolitaines) et taille de la commune de résidence (4 tranches de taille). Ces 7 variables au total ont été croisées et pour chacun des croisements ainsi obtenus, un taux de couverture est estimé. Une régression linéaire permet alors de mesurer l'effet propre de chacune des 7 variables et d'identifier celles qui sont significativement associées à un taux de couverture plus faible (cf. tableau 1.10).

Le taux de couverture est significativement plus faible pour les hommes, les moins de 50 ans et les personnes nées à l'étranger. Il est également plus faible en Île-de-France, Occitanie et Provence-Alpes-Côte d'Azur (la Corse est un cas particulier avec un taux d'appariement nettement plus faible que dans les autres régions, en raison d'adresses moins souvent « normalisées » et de difficultés de géolocalisation). Quant à l'année, elle n'est pas significative sur le taux de couverture. Enfin, même une fois contrôlée de ces variables, la catégorie A reste significativement moins bien appariée que les autres catégories, ce qui indique que le taux de couverture plus faible de cette catégorie A ne s'explique pas seulement par le fait qu'elle concentre davantage d'hommes, de personnes de moins de 50 ans ou bien nées à l'étranger par exemple.

Plusieurs hypothèses peuvent expliquer ces résultats. Tout d'abord, on peut supposer que certaines catégories sont moins bien appariées car la qualité des variables utilisées dans l'algorithme serait insuffisante pour retrouver les personnes. Par exemple, les prénoms des personnes nées à l'étranger sont généralement orthographiés de manière moins homogène entre les sources, ce qui peut diminuer les chances d'apparier. De plus, au-delà de ces difficultés techniques, les inscrits à Pôle emploi qui résident en communautés ou qui sont sans domicile fixe ne peuvent être représentés par l'enquête Emploi, puisque cette dernière n'est réalisée qu'auprès des personnes qui résident en ménages ordinaires. Dans le recensement de la population de 2015, hors ménages ordinaires, seules 45 000 personnes se déclarent spontanément au chômage. Il se peut toutefois que les inscrits à Pôle emploi qui ne résident pas en ménages ordinaires, compte tenu de leurs situations spécifiques (précarité, handicap, etc.), ne se déclarent pas spontanément chômeurs dans le recensement, même s'ils recherchent un emploi. Le nombre d'inscrits en catégorie A résidant hors ménages ordinaires pourrait donc être plus élevé que 45 000. Enfin, le plus faible taux de couverture de certaines catégories pourrait aussi s'expliquer par des biais de l'enquête Emploi. L'enquête est en effet calée sur des marges représentant l'ensemble de la population des ménages suivant des dimensions variées (sexe, âge, région etc.), mais les redressements et calages peuvent s'avérer insuffisants pour reproduire fidèlement une sous-population précise, comme celle des inscrits à Pôle emploi.

Un appariement basé sur une clé plus « forte » ou plus « identifiante » comme par exemple le numéro de sécurité sociale (NIR) permettrait sans doute d'aboutir à de meilleurs taux de couverture, tout en limitant au maximum les risques de faux-positifs. Dans l'enquête Emploi, un tel identifiant n'est toutefois pas disponible. En effet, les personnes enquêtées ne communiquent pas leur NIR, ni même un nom de famille qui permettrait de retrouver le NIR après la collecte. Seul le prénom est demandé pour repérer les personnes composant le ménage. Par ailleurs, les informations identifiantes disponibles à partir de la base de sondage de l'enquête font référence au ménage occupant le logement au sens de la taxe d'habitation, mais pas aux occupants à la date de l'enquête, qui peuvent différer en cas de déménagement. Pour ces raisons, un éventuel appariement sur la base du NIR demanderait des évolutions dans la collecte de l'enquête Emploi ainsi qu'un accord de la CNIL.

Tableau 1.10 : Effet sur le taux de couverture du sexe, de l'âge, du lieu de naissance, de la région de résidence, de l'année et de la catégorie de demandeur d'emploi

Variable	Taux de couverture
Individu de référence	97%
Sexe	
Homme	-7% ***
Femme	ref.
Age	
15-24 ans	-11% ***
25-49 ans	-8% ***
50-64 ans	ref.
Lieu de naissance	
A l'étranger	-9% ***
En France	ref.
Région	
Île-de-France	-8% ***
Centre-Val de Loire	ns
Bourgogne-Franche-Comté	ns
Normandie	ref.
Nord-Pas-de-Calais-Picardie	ns
Alsace-Champagne-Ardenne-Lorraine	ns
Pays de la Loire	ns
Bretagne	-4% **
Aquitaine-Limousin-Poitou-Charentes	-4% ***
Languedoc-Roussillon-Midi-Pyrénées	-8% ***
Auvergne-Rhône-Alpes	-3% **
Provence-Alpes-Côte d'Azur	-10% ***
Corse	-64% ***
Taille de la commune	
Inférieure à 5 000 habitants	2% **
De 5 000 à 20 000 habitants	2% *
De 20 000 à 100 000 habitants	ns
Supérieure à 100 000 habitants	ref.
Année	
2012	ns
2013	ns
2014	ns
2015	ns
2016	ns
2017	ref.
Catégorie	
A	ref.
B	7% ***
C	7% ***
D	7% ***
E	ns

Note : *** = significatif à 1 %, ** = significatif à 5 %, * = significatif à 10 %, ns = non significatif. L'individu de référence est celui qui combine toutes les modalités de référence (femme, 50-64 ans, née en France, etc.).

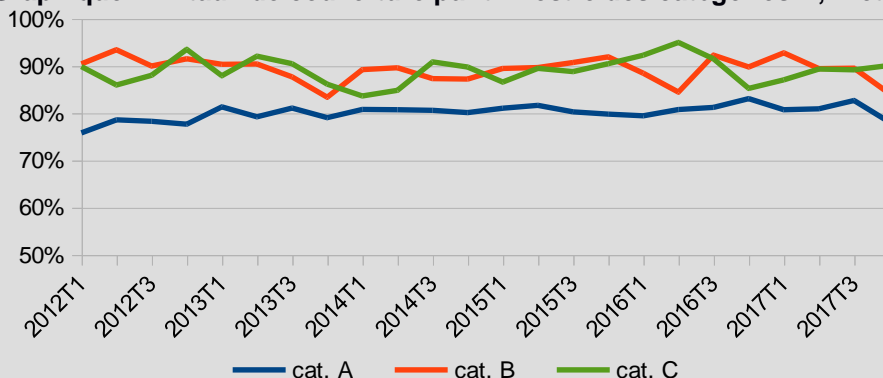
Champ : France métropolitaine, 15-64 ans, population des ménages.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Encadré 1 : Zoom sur le taux de couverture des catégories A, B et C

La catégorie A est celle dont le taux de couverture est le plus faible, tandis que les catégories B et C bénéficient des taux de couverture les plus importants. Pour les trois catégories, le taux de couverture est assez stable dans le temps, même à un niveau trimestriel (graphique E1). Pour les catégories B et C, dont les effectifs sont un peu plus faibles, le taux de couverture trimestriel est un peu plus volatil, mais il est d'environ 90 % sur l'ensemble de la période.

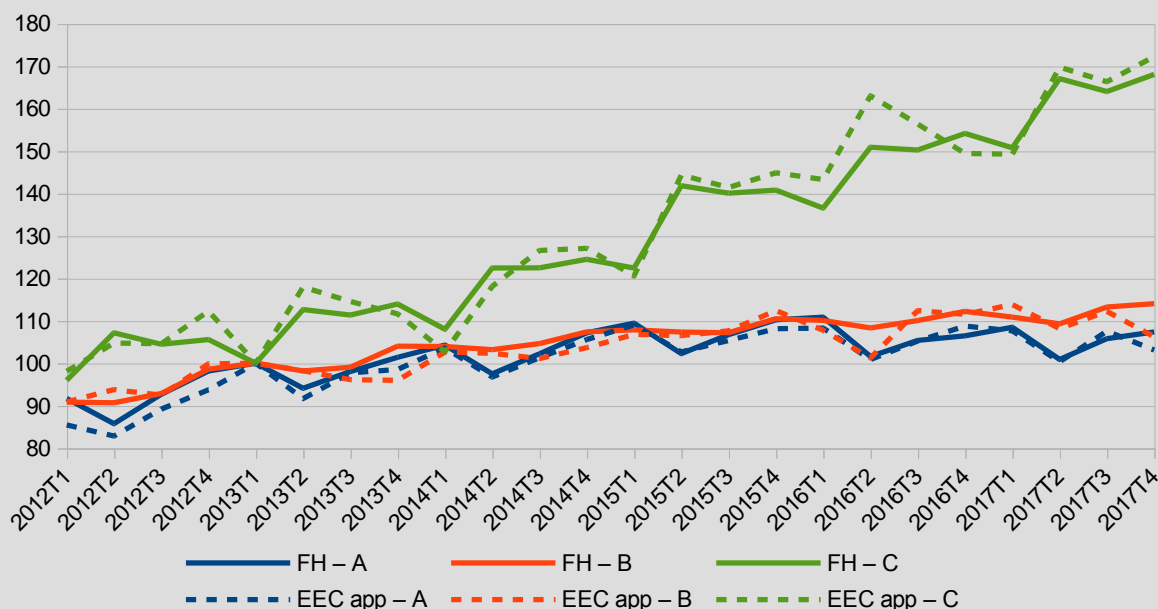
Graphique E1 : taux de couverture par trimestre des catégories A, B et C, 2012-2017



Champ : France métropolitaine, 15-64 ans, ménages ordinaires, inscrits en catégories A, B ou C
 Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Pour les trois catégories, les effectifs reconstruits à partir de l'enquête Emploi appariée évoluent sur longue période de façon très proche de ceux issus du Fichier Historique (graphique E2). Ils reflètent de plus à peu près les saisonnalités. L'appariement semble donc permettre de bien retracer les évolutions temporelles de ces trois catégories, et ce malgré les fluctuations du taux de couverture.

Graphique E2 : évolution des effectifs des catégories A, B et C dans le Fichier Historique et l'enquête Emploi appariée, 2012-2017



Champ : France métropolitaine, 15-64 ans, ménages ordinaires, inscrits en catégories A, B ou C, nombre d'inscrits en base 100 au 1^{er} trimestre 2013
 Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Chapitre 2 – Les principaux résultats en coupe

2-1 Recoupements des statuts BIT et des catégories de demandeur d'emploi

Note pour l'interprétation des résultats

Dans les résultats qui suivent, réalisés à partir de la base appariée, seuls 84 % des inscrits sont identifiés (81 % en catégorie A, cf. partie 1-6 « Taux de couverture »). Les échecs d'appariement ne peuvent pas être détectés et redressés au niveau individuel. Il n'a pas non plus été choisi de redresser les agrégats de manière globale, faute d'hypothèse satisfaisante pour réaliser une telle extrapolation. Par conséquent, dans la suite de cette partie, **le statut « non inscrit à Pôle emploi » doit en général être compris comme « non identifié parmi les inscrits à Pôle emploi », ce qui peut recouvrir des situations de non-inscription réelle (majoritairement) ou de défaut d'appariement** (minoritairement, mais plus concentrées dans certaines sous-populations, les hommes, les moins de 50 ans, etc.).

Les résultats en coupe présentés par la suite, dans ce chapitre 2, sont relatifs à l'année 2017, année la plus récente. Les évolutions présentées au chapitre 3 sont étudiées sur les périodes 2013-2017 ou 2014-2017. L'année 2012 n'est pas utilisée ici. En effet, le questionnaire de l'enquête Emploi de 2012 est sensiblement différent de celui des années qui suivent, occasionnant des ruptures de séries pour les variables d'intérêt. Le questionnaire est resté très stable sur toute la période 2013-2017, à l'exception des composantes du halo, qui ont été modifiées en cours d'année 2013 ce qui conduit, pour certains indicateurs impliquant le halo, à privilégier l'analyse des évolutions sur la période 2014-2017.

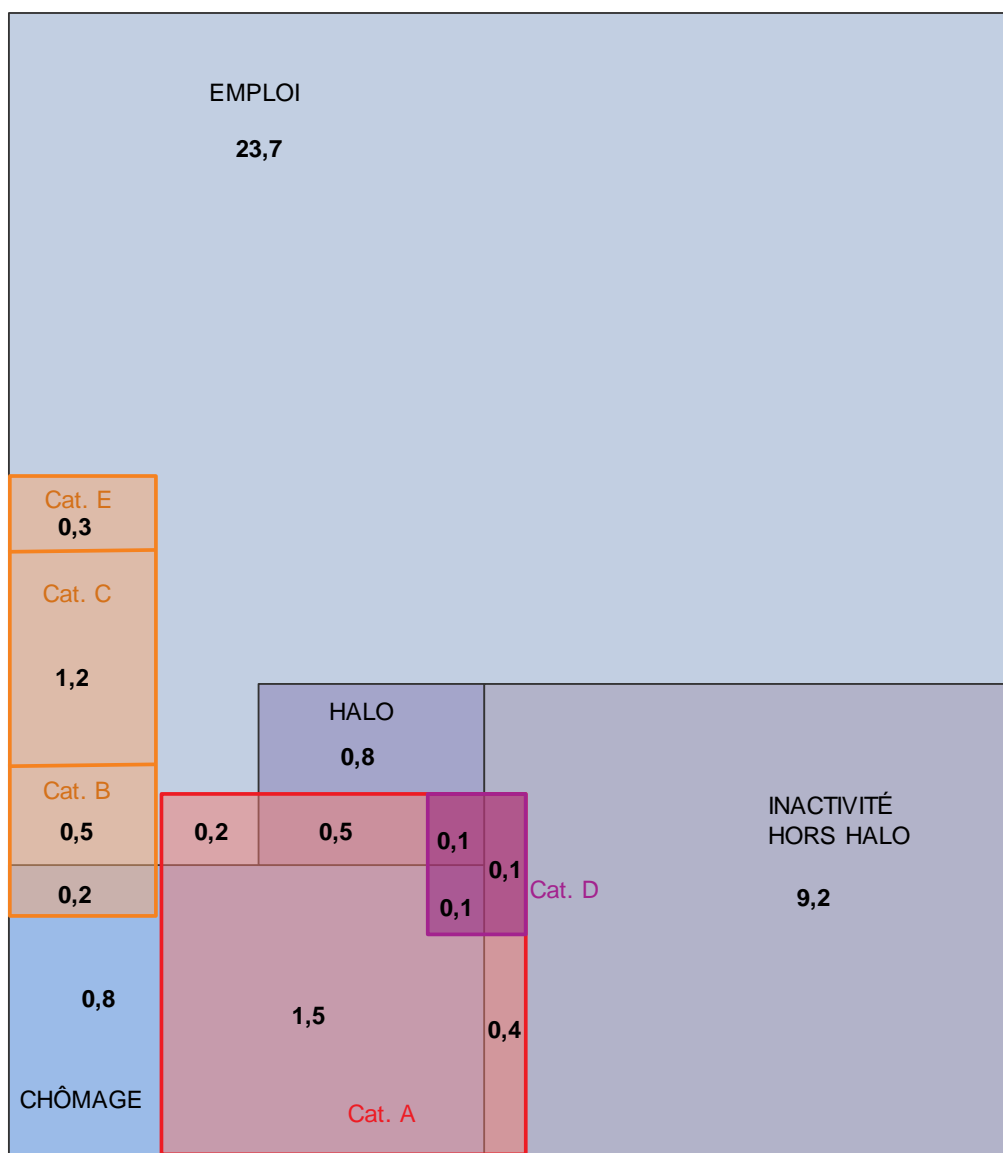
Une fois l'appariement réalisé, il est possible de voir comment les deux grandes définitions du « chômage » (statut BIT et catégorie A à Pôle emploi) se recoupent. En 2017, en moyenne sur l'année, parmi les personnes appariées :

- **1,5 million de personnes sont à la fois au chômage au sens du BIT et inscrites en catégorie A à Pôle emploi** (cf. graphique 2.1).
- **1,2 million de personnes sont inscrites catégorie A sans être au chômage au sens du BIT**, dont :
 - 1,0 inactives (dont 0,5 dans le halo et 0,4 hors halo) ;
 - 0,2 en emploi au sens du BIT
- **1,1 million de personnes sont au chômage au sens du BIT, sans être inscrites en catégorie A**, dont :
 - 0,8 million de personnes non inscrites à Pôle emploi (ce qui recouvre des situations d'échec d'appariement, cf. partie 2.3) ;
 - 0,3 million de personnes inscrites dans une catégorie autre que la A (0,2 million en catégorie B et 0,1 million en catégorie D).

Pour chaque personne, on appelle par la suite « situation » le croisement entre le statut BIT (CHO pour chômeur, EMP pour emploi, HAL pour le halo, INA pour les inactifs hors halo) et la catégorie de demandeur d'emploi (A, B, C, D, E et N pour les non-inscrits). Par exemple : la situation « CHO/A » signifie chômeur au sens du BIT et inscrit en catégorie A, la situation « EMP/N » signifie en emploi au sens du BIT et non inscrit à Pôle emploi.

Graphique 2.1 : Superposition des statuts BIT (emploi, chômage, halo et inactivité hors halo) et des catégories de Pôle emploi (A, B, C, D et E), en 2017

En millions de personnes



Note : les effectifs en millions ont été arrondis à la centaine de milliers près. Pour rendre le graphique plus lisible, seules les catégories dont les effectifs sont supérieurs à 50 000 personnes (et qui s'arrondissent donc à au moins 0,1 million) sont représentées ici. Attention : la catégorie de demandeur d'emploi n'est connue que pour les demandeurs d'emploi appariés. Dans la figure présentée ci-dessus, aucune extrapolation n'a été réalisée pour les autres personnes. Par conséquent, le nombre de demandeurs d'emploi en catégorie A (2,7 millions) dans cette figure est inférieur au total connu dans le Fichier Historique (3,4 millions). Il en est de même pour les autres catégories de demandeurs d'emploi (B, C, D et E).

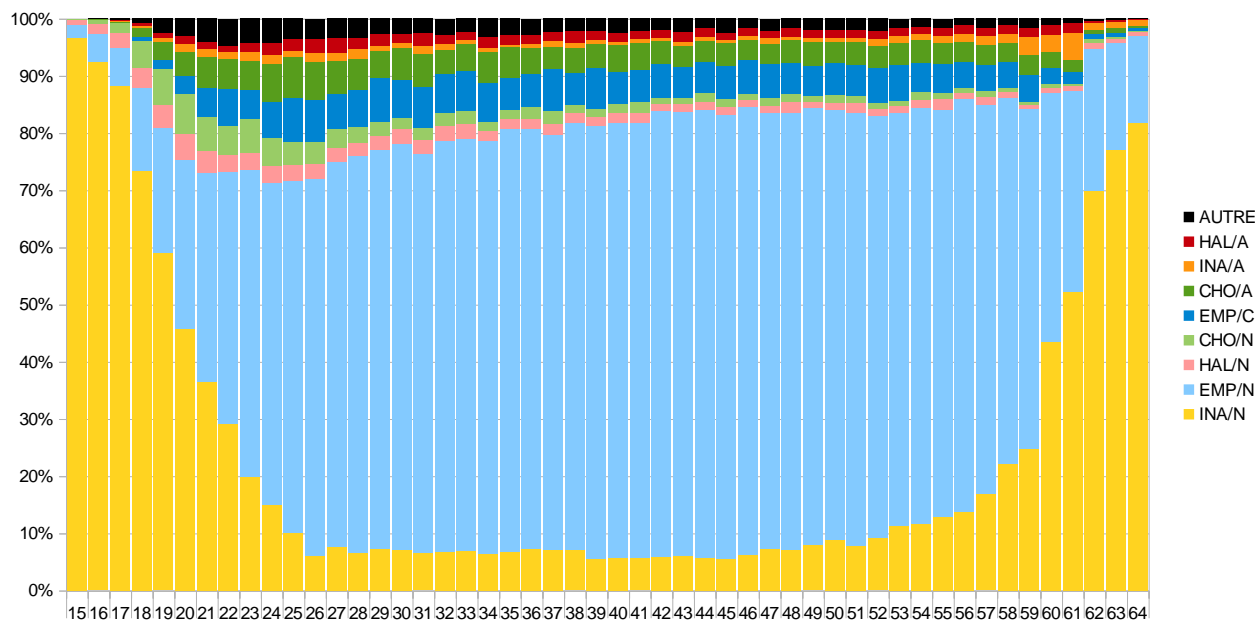
Champ : France métropolitaine, 15-64 ans, ménages ordinaires.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

L'âge est particulièrement structurant dans les situations précédemment définies (cf. graphique 2.2). Ainsi, en 2017, les chômeurs non inscrits (CHO/N ; vert pâle) et les inactifs en halo non inscrits (HAL/N ; rose) sont plutôt des jeunes tandis que les inactifs en catégorie A (INA/A ; orange) concentrent un grand nombre de seniors, notamment très proches de l'âge de départ à la retraite (entre 59 et 61 ans). En revanche, les CHO/A (vert foncé), HAL/A (rouge) et EMP/BCE (bleu foncé) proviennent quant à eux d'à peu près toutes les tranches d'âges.

Graphique 2.2 : principales situation BIT et à Pôle emploi pour chaque âge, année 2017

En %



Note : toutes les tranches d'âges sont d'effectifs proches (environ 800 000 personnes à chaque âge), ce qui permet d'interpréter le graphique aussi bien « en ligne » qu'en colonne. Ainsi, on peut à la fois conclure par exemple que les CHO/N représentent une part importante des jeunes, mais aussi que les jeunes représentent une part importante des CHO/N.

Lecture : en 2017, 98 % des jeunes de 15 ans sont INA/N, c'est-à-dire inactifs hors halo au sens du BIT et non inscrits à Pôle emploi.

Champ : France métropolitaine, 15-64 ans, ménages ordinaires.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

2-1-a Les trois populations de « référence »

Les croisements des statuts BIT et des catégories de de Pôle emploi délimitent trois populations principales, pour lesquelles les deux sources apparaissent cohérentes : les chômeurs BIT inscrits en catégorie A (1,5 million en 2017), les personnes en emploi au sens du BIT et inscrites en catégorie B, C et E (1,9 million), et enfin les inactifs au sens du BIT inscrits en catégorie D (0,2 million). Ces trois populations sont ici nommées « populations de référence » car c'est principalement en référence à ces groupes que seront décrites par la suite les populations pour lesquelles les statuts BIT et catégories de Pôle emploi semblent divergents.

Parmi les 2,7 millions d'inscrits en catégorie A et appariés en 2017, 1,5 million, soit 56 % (cf. tableau 2.3), sont aussi au chômage au sens du BIT (CHO/A). Toutes les tranches d'âge sont représentées parmi les CHO/A, bien que, chez les plus jeunes, l'inscription à Pôle Emploi soit moins fréquente (graphique 2.2). Enfin, la quasi-totalité des CHO/A se déclarent spontanément au chômage (cf. tableau 2.4). Pour une majorité d'individus, les trois « définitions » du chômage (statut BIT, catégorie de DEFM et statut spontanément déclaré) coïncident donc bien.

Tableau 2.3 : part des statuts BIT en fonction de la catégorie d'inscription, en 2017

	Inscrits en cat. A	Inscrits en cat. B, C ou E	Inscrits en cat. D
Chômage BIT	56%	10%	23%
Emploi BIT	9%	87%	15%
Halo	20%	2%	30%
Inactivité hors halo	16%	1%	33%
Ensemble	100%	100%	100%

Lecture : en 2017, parmi les personnes de l'enquête emploi qui ont pu être appariées avec le Fichier Historique et qui sont inscrites en catégorie A au moment de l'enquête, 56 % sont au chômage au sens du BIT, 9 % sont en emploi au sens du BIT, etc.

Champ : France métropolitaine, 15-64 ans, ménages ordinaires, personnes de l'enquête emploi appariées avec le Fichier Historique et inscrites à Pôle emploi au moment de l'enquête.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Tableau 2.4 : principales caractéristiques des populations de référence en 2017

	Chômeurs et cat. A	En emploi et cat. B/C/E	Inactifs et cat. D	
			Halo	Hors halo
<i>Nombre de personnes représentées et appariées (en millions)</i>	1,5	1,9	0,1	0,1
Hommes	52	42	45	32
Femmes	48	58	55	68
15-24 ans	14	10	15	38
25-49 ans	61	65	66	54
50-64 ans	24	26	19	8
Se déclare spontanément				
Chômeur	96	3	54	24
En emploi	1	97	5	5
Inactif	2	1	41	71

Champ : France métropolitaine, 15-64 ans, ménages ordinaires, CHO/A, EMP/BCE, HAL/D et INA/D.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Parmi les 2,2 millions de personnes inscrites en catégories B, C ou E et appariées en 2017, 1,9 million, soit 87 %, sont en emploi au sens du BIT. Il s'agit plus souvent de femmes, de jeunes, et d'emplois courts (CDD, intérim) localisés dans certains secteurs spécifiques : services à la personne, industrie, arts et spectacles. Les populations inscrites en catégories B, C et E, dont le nombre a nettement augmenté entre 2012 et 2017 (en particulier pour la catégorie C), sont décrites plus en détail au paragraphe 2-1-b.

Enfin, parmi les 0,3 million de personnes inscrites en catégorie D et appariées en 2017, 0,2 million, soit 63 %, sont inactives au sens du BIT, dont 0,1 million dans le halo (HAL/D) et 0,1 million hors halo (INA/D). Les inactifs inscrits en catégorie D ne représentent qu'une toute petite partie de l'ensemble des inactifs, et constituent une population très spécifique. La catégorie D est destinée à comptabiliser, dans les statistiques de Pôle emploi, des inscrits non tenus de faire des démarches actives de recherche d'emploi car en formation ou arrêt maladie. Toutefois le motif de maladie reste minoritaire¹ et les HAL/D et INA/D sont donc essentiellement des personnes en formation. Le statut auto-déclaré est lui aussi plus ambigu : les personnes dans le halo (HAL/D) se déclarent spontanément plutôt au chômage, tandis que les inactifs hors halo (INA/D) se déclarent spontanément plutôt inactifs.

¹ D'une part, les formations ou arrêts maladie courts (inférieurs à 15 jours) ne donnent pas lieu à un reclassement par Pôle emploi en catégorie D et d'autre part, pour les personnes durablement éloignées du marché du travail (maladie chronique par exemple), rester inscrit à Pôle emploi ne présente pas nécessairement d'utilité.

2-1-b Zoom sur les personnes en emploi BIT et inscrites en catégories B, C ou E : des formes d'emploi atypiques

En 2017, en moyenne, 1,9 million de personnes sont en emploi au sens du BIT et inscrites à Pôle emploi en catégories B, C ou E, dont 1,2 million en catégorie C, 0,5 million en catégorie B et 0,3 million en catégorie E. En plus des 0,5 million de personnes en emploi et inscrits en catégorie B, 0,2 million de personnes sont en catégorie B mais sont classées comme chômeuses au sens du BIT (ce qui peut s'expliquer, cf. partie 2.3, par une différence de concept entre Pôle emploi, qui mesure l'activité sur le mois, et l'enquête Emploi, qui la mesure au cours d'une semaine de référence) ce qui fait au total 0,7 million d'inscrits en catégorie B. Dans ce paragraphe, on se concentre surtout sur les personnes en catégories B et C, car elles sont tenues de faire des recherches actives d'emploi comme les catégories A. Le nombre de personnes en catégorie C a de plus particulièrement augmenté entre 2012 et 2017.

Les femmes sont surreprésentées en catégorie B et dans une moindre mesure en catégorie C (cf. tableau 2.5). Cette population plus féminine va de pair avec le type de professions occupées (cf. catégories socioprofessionnelles). Les catégories E sont plus mixtes. Les catégories B et C comportent de plus une part importante de jeunes, comparé aux personnes en emploi non inscrites. Dans cette tranche d'âge où les jeunes sont majoritairement en cours d'études, ceux qui se sont insérés dans le marché du travail se retrouvent plus souvent inscrits à Pôle emploi dans ces deux catégories.

Les professions de services à la personne sont particulièrement surreprésentées dans les catégories B (27 %) et dans une moindre mesure C (19 %, contre 6 % pour les personnes en emploi non inscrites). Les professions de l'information, des arts et des spectacles et les ouvriers non qualifiés de l'artisanat sont fréquents dans la catégorie B tandis que les non-qualifiés de type industriel sont très présents dans la catégorie C, probablement en lien avec le fait qu'un quart des personnes en catégorie C sont des intérimaires. Hormis les professions des arts et des spectacles, les professions de CSP supérieure sont sous-représentées en B et C (voir tableaux détaillés A5 en annexe pour plus de précisions sur la structure par CS des inscrits en catégories B et C). Ces quatre catégories socioprofessionnelles (professions des services à la personne et de l'information, des arts et des spectacles, ainsi que ouvriers non qualifiés de l'industrie ou de l'artisanat) regroupent à elles seules 50 % des inscrits en catégorie B et presque 40 % des inscrits en catégorie C, contre seulement 13 % des personnes en emploi non inscrites à Pôle emploi.

Comparées à l'ensemble de la population en emploi, les inscrits en catégories B et C sont plus souvent dans des formes particulières d'emploi. Les deux-tiers des catégories B et C sont en CDD ou en intérim : cette proportion est inférieure à 10 % parmi les personnes en emploi non inscrites. En contrepartie, la part des CDI est particulièrement faible chez les demandeurs d'emploi de catégorie C et B (respectivement 37 % et 29 % contre 79 % chez les personnes en emploi non inscrites). Enfin, 38 % des inscrits en catégorie E sont des non-salariés, cette catégorie regroupant notamment des bénéficiaires de dispositifs d'aides pour monter leur entreprise. Les inscrits en catégorie E salariés occupent quant à eux fréquemment des emplois aidés dans des associations ou dans la fonction publique.

Les demandeurs d'emploi de catégories B et C sont également davantage en temps partiel, et notamment en sous-emploi, et ils déclarent davantage que leur activité est temporaire ou d'appoint. Ils déclarent également plus souvent souhaiter un emploi en plus ou un emploi de remplacement. Ils ont plus souvent des horaires de travail atypiques et sont plus souvent pluriactifs (voir tableaux détaillés A6 et A7 en annexe). Par rapport aux autres personnes en emploi non inscrites, les personnes en catégorie C salariées sont plus souvent dans des entreprises privées ou des associations, tandis que celles en catégorie B travaillent très souvent pour des particuliers employeurs. Corrélativement, la proportion des personnes en catégories B et C ayant plusieurs employeurs est trois fois plus forte que pour le reste des salariés.

Enfin, lorsqu'ils sont interrogés sur leur situation principale sur le marché du travail, 8 % des demandeurs d'emploi en catégorie B se déclarent spontanément chômeurs. Cette proportion est en revanche très faible pour les inscrits en catégorie C, comparable à celle des inscrits en catégorie E ou des personnes non inscrites à Pôle emploi.

Tableau 2.5 : principales caractéristiques des personnes en emploi au sens du BIT, selon qu'elles sont ou non inscrites comme demandeur d'emploi, et selon leur catégorie en 2017

	En %			
	En emploi et cat. B	En emploi et cat. C	En emploi et cat. E	En emploi non inscrit
<i>Nombre de personnes représentées et appariées (en millions)</i>	0,5	1,2	0,3	23,7
Hommes	35	44	47	53
Femmes	65	56	53	47
15-24 ans	10	10	8	8
25-49 ans	58	66	64	62
50-64 ans	32	23	28	30
Se déclare spontanément chômeur	8	1	1	0
Se déclare spontanément en emploi	91	98	99	98
Se déclare spontanément inactif	1	0	0	1
CS35 – professions de l'information, arts et spectacles	7	2	1	1
CS56 – personnels des services directs aux particuliers	27	19	5	6
CS67 – ouvriers non qualifiés de type industriel	6	12	3	3
CS68 – ouvriers non qualifiés de type artisanal	9	4	3	3
Contrat CDD	40	42	45	7
Contrat CDI	37	29	16	79
Contrat intérimaire	15	27	1	1
Non salarié	8	3	38	11
Employeur – entreprise privée ou association	67	75	62	71
Employeur - entreprise publique	3	2	3	4
Employeur – état	4	4	16	11
Employeur – collectivités territoriales	8	5	15	8
Employeur – hôpitaux publics	1	2	3	4
Employeur – Particuliers	18	12	1	2
Emploi d'appoint/temporaire	19	14	5	2
Temps partiel	68	29	46	16
Sous-emploi	47	19	29	4
Souhaite un emploi de remplacement	26	19	15	8
Souhaite un emploi en plus	25	9	11	3
Se déclare inscrit à Pôle Emploi	77	73	50	2
Déclare percevoir une allocation chômage	45	28	25	0
Déclare percevoir le RSA	16	8	14	4

Champ : France métropolitaine, ménages ordinaire, 15-64 ans, année 2017, personnes en emploi au sens du BIT et inscrites en catégories B, C, E ou non inscrites.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

2-2 Les demandeurs d'emploi en catégorie A non chômeurs au sens du BIT

Les demandeurs d'emploi en catégorie A inactifs au sens du BIT

En 2017, dans la base appariée, 1,0 million de personnes sont inactives au sens du BIT et inscrites en catégorie A (soit un tiers des inscrits en catégorie A), dont 0,5 million dans le halo autour du chômage (HAL/A, 20 % des inscrits en catégorie A) et 0,4 million hors halo (INA/A, 16 % des inscrits en catégorie A). Les raisons d'inactivité déclarées par les personnes sont variées. Il peut s'agir de motifs temporaires (maladie de courte durée, vacances, déménagement...) ou de motifs plus durables (problème de santé, formation, s'occuper de sa famille...). En principe, la catégorie D a été conçue pour comptabiliser, dans les statistiques de Pôle emploi, les personnes répondant à certains motifs durables (formation ou problème de santé notamment).

Les caractéristiques des HAL/A sont très proches du chômage. Il s'agit même souvent de chômeurs ayant basculé ponctuellement dans l'inactivité. En effet, les motifs les plus cités par les HAL/A pour expliquer leur inactivité sont des motifs de court terme (vacances, déménagement, maladie de courte durée, etc., 37 %) ainsi que le découragement (15 %, cf. tableau 2.6).

Tableau 2.6 : Les inactifs en 2017, selon leur catégorie d'inscription à Pôle emploi et leur appartenance ou non au Halo

	En %					
	Halo Cat. A	Halo Cat. D	Halo non inscrit	Inact. Cat. A	Inact. Cat. D	Inact. non inscrit
<i>Nombre de personnes représentées et appariées (en millions)</i>	0,5	0,1	0,8	0,4	0,1	0,2
Hommes	45	45	41	39	32	42
Femmes	55	55	59	61	68	58
15-24 ans	11	15	29	14	38	45
25-49 ans	61	66	51	40	54	15
50-64 ans	28	19	20	46	8	40
Se déclare spontanément chômeur	88	54	38	52	24	2
Se déclare spontanément en emploi	2	5	6	1	5	2
Se déclare spontanément inactif	11	41	56	47	71	96
Halo1 (a recherché, mais pas disponible)	26	30	20			
Halo2 (disponible, mais n'a pas recherché)	48	18	41			
Halo3 (ni recherches, ni disponible)	26	52	39			
A déjà travaillé	90	91	67	86	77	49
Se déclare inscrit à Pôle Emploi	94	92	21	71	69	1
Déclare percevoir allocation chômage	52	59	5	46	47	0
Déclare percevoir le RSA	18	11	17	14	6	3
Motif d'inactivité						
Invalidité, problèmes de santé	12	4	18	38	4	13
S'occuper de sa famille, de son foyer	16	3	20	19	2	9
Retraite	1	0	2	7	1	25
Études, formations	12	76	21	15	89	45
Motifs de courte durée*	37	11	23	2	0	0
Découragement	15	3	11	7	0	1
Autres motifs ou motifs non renseignés	8	3	6	13	3	7

Note : la variable « motif » a été compilée à partir des différents motifs de ne pas avoir recherché d'emploi (RAISNREC), de ne pas être disponible (NONDIC), de ne pas souhaiter travailler (RAISNSOU), et par quelques modalités du statut autodéclaré (SP00 : invalidité, retraite, études/formations, congé parental) afin de limiter autant que possible la modalité « autres motifs ». L'étude des libellés en clair a aussi permis de reclasser un grand nombre de motifs « autres » dans la liste des motifs précédents (cf. annexe 1.6).

** : vacances, déménagement, attente de résultats de démarches antérieures, maladie de courte durée, a déjà trouvé un emploi qui commence ultérieurement, en cours de création d'entreprise.*

Champ : France métropolitaine, 15-64 ans, ménages ordinaires, inactifs au sens du BIT

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

De plus, en 2017, 33 % des HAL/A un trimestre donné ont basculé vers CHO/A le trimestre suivant (tableau 2.7). Pour les inscrits en catégorie A, le motif de découragement concerne plutôt des seniors, et il semble lui aussi être un motif d'assez courte durée : 36 % des inactifs inscrits en catégorie A se déclarant découragés un trimestre donné ont basculé comme chômeurs au sens du BIT au trimestre suivant, tandis que 26 % seulement sont restés classés comme inactifs et découragés (tableau 2.8). Enfin, 88 % des HAL/A se déclarent spontanément au chômage dans l'enquête Emploi. Les caractéristiques des personnes par motif sont détaillées au tableau A8 en annexe.

Tableau 2.7 : Situation des personnes au trimestre T+1 selon leur situation au trimestre T, moyenne sur l'année 2017

En % de la situation en T

Situation en T	Situation en T+1										
	CHO/A	CHO/BCE	CHO/N	EMP/A	EMP/BCE	EMP/N	HAL/A	HAL/N	INA/A	INA/N	AUTRE
CHO/A	53%	4%	4%	2%	12%	6%	10%	1%	4%	1%	3%
CHO/BCE	20%	16%	3%	2%	40%	7%	5%	1%	2%	0%	6%
CHO/N	7%	1%	41%	0%	1%	22%	1%	12%	1%	13%	1%
EMP/A	16%	3%	1%	33%	23%	16%	3%	0%	2%	1%	2%
EMP/BCE	6%	3%	0%	3%	73%	11%	2%	0%	1%	0%	2%
EMP/N	0%	0%	0%	0%	0%	96%	0%	0%	0%	1%	0%
HAL/A	33%	2%	3%	2%	7%	4%	27%	3%	13%	3%	5%
HAL/N	4%	0%	15%	0%	0%	11%	2%	28%	1%	37%	0%
INA/A	15%	1%	1%	1%	2%	1%	14%	2%	48%	11%	5%
INA/N	0%	0%	1%	0%	0%	3%	0%	3%	0%	91%	0%

Lecture : parmi les personnes au chômage et inscrites en catégorie A au trimestre T (CHO/A), 53 % restent dans la catégorie CHO/A au trimestre T+1, tandis que 12 % basculent vers la catégorie EMP/B, C ou E.

Champ : France métropolitaine, ménages ordinaires, 15-64 ans.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Tableau 2.8 : Situation au trimestre T+1 des personnes inactives et inscrites en catégorie A au trimestre T, moyenne sur l'année 2017

En % de la situation en T

HAL/A ou INA/A Situation Au trimestre T	Situation au trimestre T+1							Chômage	Emploi
	Halo ou inactivité hors halo								
	Problèmes de santé	S'occuper de sa famille	Retraite	Études ou formation	Motifs temp.*	Découragé	Autres motifs		
Problèmes de santé	63%	2%	2%	2%	3%	3%	2%	20%	3%
S'occuper de sa famille	3%	50%	0%	2%	2%	2%	7%	26%	6%
Retraite	8%	0%	43%	0%	2%	9%	24%	12%	1%
Études ou formation	1%	1%	0%	43%	5%	1%	4%	33%	11%
Motifs temporaires*	5%	4%	1%	5%	14%	4%	6%	40%	20%
Découragé	9%	5%	3%	3%	6%	26%	7%	36%	5%
Autres motifs	6%	8%	5%	6%	7%	7%	23%	24%	13%

Lecture : parmi les personnes inscrites en catégorie A au trimestre T et classées comme inactives au sens du BIT (halo ou hors halo) en raison de problèmes de santé, 63 % sont toujours inactives pour raison de santé au trimestre T+1, tandis que 20 % sont au chômage au sens du BIT au trimestre T+1.

* : vacances, déménagement, attente de résultats de démarches antérieures, maladie de courte durée...

Champ : France métropolitaine, 15-64 ans, ménages ordinaires, inactifs (halo et hors halo) inscrits en catégorie A.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Les inactifs hors halo et en catégorie A (INA/A) sont moins « proches » du chômage. Cette catégorie comporte un nombre important de seniors, inactifs pour invalidité ou problème de santé. Le fait de ne pas avoir été comptabilisé en catégorie D peut s'expliquer par des problèmes de santé ne s'accompagnant pas d'un arrêt maladie. Il peut aussi s'agir d'arrêts inférieurs à 15 jours, pour lesquels Pôle emploi n'opère pas de reclassement. Enfin, il est aussi possible que des demandeurs d'emploi bénéficient d'une prescription d'arrêt maladie supérieure à 15 jours, mais ne l'utilisent pas. En effet, pour les bénéficiaires d'une allocation, il peut être désavantageux de déclarer un arrêt, surtout s'il s'agit d'un arrêt de courte durée, car le délai de carence de trois jours représente une perte de revenu. En 2014, le délai moyen entre le début de l'arrêt maladie et la reprise du droit est d'ailleurs de 180 jours en moyenne d'après le fichier national des allocataires de Pôle

emploi, soit nettement plus que la durée moyenne des arrêts observés sur l'ensemble de la population (33 jours).

Les INA/A sont constitués de nombreux seniors très proches de l'âge de la retraite, puisqu'un tiers d'entre eux a entre 58 et 64 ans et la moitié a plus de 50 ans. À l'approche de 60 ans, la catégorie CHO/A a tendance à se réduire tandis que la catégorie INA/A augmente nettement (graphique 2.2). Ainsi, en 2017, parmi les inscrits en catégorie A âgés de 58 à 64 ans, 42 % indiquent ne pas souhaiter travailler (INA/A), 38 % sont au chômage au sens du BIT (CHO/A), 19 % sont dans le halo (HAL/A) et 1 % sont en emploi (EMP/A).

La situation INA/A n'est pas aussi transitoire que celle des personnes en HAL/A : en 2017, 48 % des INA/A un trimestre donné le sont encore au trimestre suivant et seuls 15 % d'entre eux ont basculé en CHO/A au trimestre suivant. De plus, la moitié des INA/A se déclare spontanément inactifs, ce qui pose question sur leur inscription en catégorie A.

En 2017, 12 % des HAL/A et 15 % des INA/A (soit environ 0,13 million de personnes) déclarent être en formation alors que la catégorie D a été conçue pour comptabiliser notamment ces personnes. Il peut s'agir de formations d'une durée inférieure à 15 jours, lesquelles ne donnent pas nécessairement lieu à un reclassement en catégorie D de la part de Pôle emploi. Il peut aussi s'agir de jeunes qui terminent leur formation initiale.

Les inactifs en catégorie A comportent aussi une proportion importante de personnes inactives (16 % dans le halo et 19 % hors halo) car elles s'occupent de leur famille (enfants ou personne âgée dépendante). Il s'agit essentiellement de femmes de moins de 40 ans. 50 % des personnes inscrites en catégorie A et inactives au sens du BIT (halo ou hors halo) car elles s'occupent de leur famille sont encore dans cette situation un trimestre plus tard. Il s'agit donc d'un motif plutôt durable, tout comme le motif « études ou formation » (43 %).

Enfin, le motif « retraite » est rarement invoqué par les INA/A (7 % d'entre eux en 2017) et très marginalement par les HAL/A (1 %). Parmi les inscrits en catégorie A, la retraite n'est pas un motif aussi stable qu'on pourrait attendre : seules 43 % des personnes l'invoquant un trimestre T l'invoquent encore au trimestre T+1, et une part importante bascule vers le motif « découragement ». Pour certaines personnes, le motif « retraite » pourrait donc aussi être une autre façon de se déclarer découragé² dans ses recherches d'emploi, en raison de son âge.

² Il pourrait s'agir de seniors qui indiquent à l'enquêteur avoir pris leur retraite, ou pris la décision de la prendre, même si elles n'en ont pas encore fait les démarches officielles. Le fait qu'une forte proportion de personnes qui invoquent ce motif « retraite » ne l'invoquent plus trois mois plus tard signale en effet qu'il ne s'agit peut-être pas d'une retraite « officielle ». Les personnes très proches de la retraite et sans emploi pourraient ainsi utiliser indifféremment les motifs « je suis bientôt à la retraite » et « je suis découragé » ou encore des motifs « autres » pour expliquer les raisons de leur absence de recherche ou de disponibilité.

Les inscrits en catégorie A en emploi au sens du BIT

En 2017, 0,2 million de personnes sont en emploi au sens du BIT et inscrites en catégorie A (situation « EMP/A »), soit 9 % des inscrits en catégorie A. Par rapport aux personnes en emploi non inscrites à Pôle emploi, les EMP/A sont plutôt non salariées, en CDD et en intérim (tableau 2.9). Parmi les non-salariés, les professions artistiques sont particulièrement représentées (auteurs, graphistes, journalistes, etc.). Les personnes en EMP/A occupent plus souvent un emploi d'appoint ou temporaire et à temps partiel. Les emplois d'agent de ménage, de garde d'enfants, d'employés dans des hôtels, cafés, bars ou restaurants sont également surreprésentés. Les « EMP/A » ne ressemblent pas non plus tout à fait aux personnes en emploi et inscrites en catégories B, C ou E (lesquelles sont plus souvent des femmes, et davantage des intérimaires ou personnes en CDD). Aussi, seule la moitié des EMP/A déclarent être inscrits à Pôle emploi, bien moins que la part correspondante des personnes inscrites en catégories B/C/E (70 %).

Tableau 2.9 : Comparaison des personnes en emploi inscrites en catégorie A avec les autres personnes en emploi en 2017

	En %		
	En emploi et cat. A	En emploi et cat. B, C ou E	En emploi non inscrit
<i>Nombre de personnes représentées et appariées (en millions)</i>	0,2	1,9	23,7
Hommes	54	42	53
Femmes	46	58	47
15-24 ans	13	10	8
25-49 ans	66	65	62
50-64 ans	21	26	30
Se déclare spontanément chômeur	10	3	0
Se déclare spontanément en emploi	88	97	98
Se déclare spontanément inactif	2	1	1
Contrat CDD	29	42	7
Contrat CDI	34	28	80
Contrat intérimaire	15	20	1
Non salarié	21	10	11
Si non salarié – artiste/auteur/journaliste	13	5	3
Conjoint collaborateur	2	0	0
Emploi informel	3	0	0
Aucune heure rémunérée lors de la semaine de ref.	24	9	14
Autre situation d'emploi au sens du BIT	72	91	86
Emploi d'appoint/temporaire	14	14	2
Temps partiel	35	41	16
Souhaite un emploi de remplacement	24	21	8
Souhaite un emploi en plus	15	13	3
Se déclare inscrit à Pôle Emploi	53	70	2
Déclare percevoir une allocation chômage	24	32	0
Déclare percevoir le RSA	13	11	4

Champ : France métropolitaine, ménages ordinaires, 15-64 ans, personnes en emploi au sens du BIT
Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

La situation « EMP/A » est contradictoire. Le statut BIT indique en effet que la personne est en emploi (c'est-à-dire a travaillé au moins une heure dans la semaine de référence, ou est absente de son poste pour une durée limitée : congés etc.), tandis que la catégorie de Pôle emploi indique l'inverse (aucune heure de travail dans le mois qui contient la semaine de référence). Il y a donc une impossibilité théorique à cumuler les deux statuts. Toutefois, cette situation semble souvent transitoire puisque seul un tiers des personnes en « EMP/A » le sont encore au trimestre suivant (tableau 2.7). Les deux tiers ont basculé, principalement vers l'emploi (inscrits en catégorie B/C/E ou non inscrits) ou vers le chômage (CHO/A) mettant ainsi fin à la contradiction des statuts. La situation EMP/A semble donc assez « frictionnelle ».

Elle pourrait s'expliquer en partie par les différences de concept d'emploi entre l'enquête Emploi et Pôle emploi. Les personnes absentes de leur emploi lors de la semaine de référence (en raison notamment de temps partiel,

congés non rémunérés, chômage partiel ou intempéries), les conjoints collaborateurs ou les personnes ayant eu une activité « informelle » sont par exemple considérés en emploi au sens du BIT, mais peuvent n'avoir déclaré aucune heure de travail à Pôle emploi. Ces populations sont ainsi surreprésentées parmi les EMP/A.

De plus, pour certains non-salariés (professions artistiques ou personnes en cours de création de leur entreprise par exemple) et salariés aux missions irrégulières (intérim par exemple), il n'est pas toujours aisé de comptabiliser les heures de travail effectuées dans le mois. Cette situation théoriquement contradictoire pourrait donc s'expliquer aussi par des délais ou erreurs dans la mise à jour de la situation de la personne à Pôle emploi ou par des biais de mémoire ou de déclaration de la part des enquêtés dans l'enquête Emploi³.

Enfin, les personnes en EMP/A pour au moins deux trimestres consécutifs (environ 70 000 personnes chaque trimestre) sont plutôt des professions artistiques ou des non-salariés, tandis que les employés d'hôtels, cafés, restaurants, les ouvriers de nettoyage, manutention, etc. ont plutôt tendance à basculer dans une autre situation au trimestre suivant.

³ Dans l'enquête Emploi, la semaine de référence peut être à cheval sur deux mois. Dans ce cas, elle est rattachée au mois majoritaire en nombre de jours et mise en regard de la catégorie de demandeur d'emploi de ce même mois majoritaire (par exemple, une semaine de référence comportant les 5 derniers jours de janvier et les 2 premiers de février est rattachée au mois de janvier et le statut BIT de l'enquête lors de cette semaine de référence est confronté à la catégorie de DEFM du mois de janvier). Les situations EMP/A ne semblent toutefois pas s'expliquer par des semaines de référence à cheval sur deux mois. En effet, la part des « semaines à cheval sur deux mois » est à peu près identique dans toutes les populations étudiées (environ 20 %), y compris les populations « contradictoires » de type EMP/A, ou CHO/B, etc.

2-3 Les chômeurs BIT non inscrits en catégorie A

Les chômeurs BIT non inscrits à Pôle emploi

En 2017, 0,8 million d'individus sont chômeurs au sens du BIT sans être inscrits à Pôle emploi. Parmi eux, on distingue trois populations.

Tout d'abord, 0,2 million de chômeurs BIT (soit 8 % d'entre eux) n'ont pas été retrouvés dans le Fichier Historique, mais déclarent dans l'enquête être inscrits à Pôle emploi (variable « OFFICC ») et sont donc suspectés d'être, en grande majorité, des échecs d'appariement. Leurs caractéristiques (en termes de statut spontanément déclaré, de démarches de recherche d'emploi, etc.) sont d'ailleurs assez proches de celles des chômeurs inscrits en catégorie A, ce qui conforte cette hypothèse (cf. tableau 2.10, colonne 4, à comparer à la colonne 1).

Tableau 2.10 : Principales caractéristiques des chômeurs BIT selon leur inscription à Pôle emploi (2017)
En %

	CHO et cat. A	CHO et cat. B/C/E	CHO et cat. D	CHO non appariés mais se déclarant inscrits	CHO appariés, non inscrits à la date de l'enquête	CHO non appariés et ne se déclarant pas inscrits
<i>Nombre de personnes représentées et appariées (en millions)</i>	1,5	0,2	0,1	0,2	0,4	0,2
<i>En % de l'ensemble des chômeurs BIT</i>	59%	8%	2%	8%	14%	8%
Hommes	52	52	49	54	56	51
Femmes	48	48	51	46	44	49
15-24 ans	14	20	19	26	38	52
25-49 ans	61	62	52	57	52	35
50-64 ans	24	17	30	17	10	13
Se déclare spontanément chômeur	96	84	87	94	85	52
Se déclare spontanément en emploi	1	14	2	3	5	7
Se déclare spontanément inactif	2	2	11	3	10	41
Chômeur de longue durée	21	18	16	22	24	13
Chômeur de très longue durée	25	10	11	25	26	17
A déjà travaillé	91	92	88	80	70	51
Se déclare inscrit à Pôle Emploi	97	96	97	100	58	0
Déclare percevoir allocation chômage	58	61	65	40	7	0
Déclare percevoir le RSA	19	10	14	17	23	8
Cherche un emploi à temps partiel	12	7	7	10	12	25
Démarches de recherche actives d'emploi effectuées						
Contact PE, APEC, CCI ou autre org. public	71	71	81	74	51	19
Contact agence d'intérim, Op. de placement	39	59	35	43	43	32
Recherches via des proches	57	56	54	60	60	62
Recherches via des relations professionnelles	33	41	41	30	28	25
Recherches via réseaux sociaux	33	34	35	29	31	30
A passé un test d'embauche ou un entretien	19	29	27	21	18	23
Candidature spontanée en personne	44	49	48	45	43	49
Candidature spontanée via courrier/mail	53	55	55	50	50	49
Salon prof., forum ou bourse d'emplois	13	15	17	9	9	9
A cherché à reprendre entreprise, cabinet...	13	14	30	17	12	13
A recherché terrain, locaux, matériel	17	11	26	19	12	16
A recherché emprunt, aides publiques	20	21	30	26	25	21
A demandé un permis, une licence...	5	7	7	9	6	8
A fait passer une annonce pour trouver un emploi	13	12	18	13	13	12
A répondu à une annonce d'offre d'emploi	55	56	55	53	46	44
A étudié des offres d'emploi	79	72	79	78	75	64
Nombre moyen de démarches actives	3,7	3,7	3,8	3,7	3,5	3,0

Note : nombre moyen de démarches actives différentes hors « a contacté Pôle emploi, l'APEC, un opérateur de placement, la chambre de commerce et d'industrie ou un autre organisme public ».

Champ : France métropolitaine, ménages ordinaires, 15-64 ans, personnes au chômage au sens du BIT.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

De plus, 0,4 million de chômeurs BIT (soit 14 % d'entre eux) ne sont pas inscrits à Pôle emploi au moment de l'enquête, en 2017, mais l'ont été au moins une fois sur la période 2012-2017 : ces chômeurs ont été retrouvés dans le Fichier Historique (tableau 2.10, colonne 5). Ils représentent 62 % des chômeurs BIT non inscrits (hors échecs d'appariement suspectés), ce qui signifie que la grande majorité des chômeurs de 2017 se sont tout de même retrouvés à un moment ou un autre inscrits à Pôle emploi sur la période 2012-2017. Ces chômeurs sont plutôt des jeunes et 58 % d'entre eux se pensent inscrits à Pôle emploi. Toutefois, seuls un quart d'entre eux sont inscrits à Pôle emploi le mois juste avant celui de l'enquête, ou le mois juste après⁴. Ces chômeurs non inscrits au moment de l'enquête (mais connus à un autre moment du Fichier Historique) ne semblent donc pas majoritairement désinscrits de Pôle emploi pour des raisons ponctuelles (défaut d'actualisation par exemple).

Enfin, 0,2 million de chômeurs au sens du BIT (soit 8 % d'entre eux) n'ont pas été appariés et ne déclarent pas être inscrits à Pôle emploi dans l'enquête (tableau 2.10, colonne 6). Ils se caractérisent par une recherche un peu moins intensive (cf. nombre moyen de démarches actives différentes). Au sein de cette population, deux groupes se distinguent. Le premier groupe de personnes, représentant les deux tiers des effectifs, sont des jeunes (moins de 30 ans) qui sont en train de finir leurs études mais ont commencé à chercher un emploi, ou bien cherchent un emploi en parallèle de leurs études (le plus souvent à temps partiel). Un grand nombre d'entre elles s'auto-déclarent d'ailleurs comme inactives (études/formations) ou en emploi (stage/apprentissage). Le dernier tiers correspond à des personnes plus âgées, souvent chômeurs de très longue durée. Il est toutefois possible que ces personnes aient été inscrites à Pôle emploi avant 2012.

Les chômeurs BIT inscrits à Pôle emploi dans une catégorie autre que la catégorie A

En 2017, 0,2 million d'individus sont chômeurs au sens du BIT et inscrits en catégorie B, C ou E (principalement B). Ici aussi, la situation CHO/B-C-E semble contradictoire puisque la catégorie de demandeur d'emploi indique que la personne est en emploi, tandis que le statut BIT affirme l'inverse. Comme pour les personnes en situation EMP/A (cf. paragraphe 2.2), cette situation serait principalement frictionnelle, due aux écarts de concepts entre le statut BIT (basée sur une semaine de référence) et la catégorie de Pôle emploi (basée sur un mois). Contrairement à la situation EMP/A, il n'y a pas ici d'impossibilité théorique : une personne peut par exemple être enquêtée en début de mois, lors d'une période de chômage, puis avoir une activité rémunérée par la suite et ainsi être classée en catégorie B-C-E à Pôle emploi. Ainsi, seules 16 % des personnes CHO/B-C-E le sont encore un trimestre après, tandis qu'elles basculent pour 40 % d'entre elles vers EMP/B-C-E et pour 20 % vers CHO/A (tableau 2.7).

Quant à la catégorie CHO/D, elle ne concerne que 60 000 personnes environ. Il s'agit là aussi d'une catégorie frictionnelle avec seulement un tiers des personnes qui sont dans la même situation le trimestre suivant, la majorité basculant vers CHO/A, HAL/D ou INA/D.

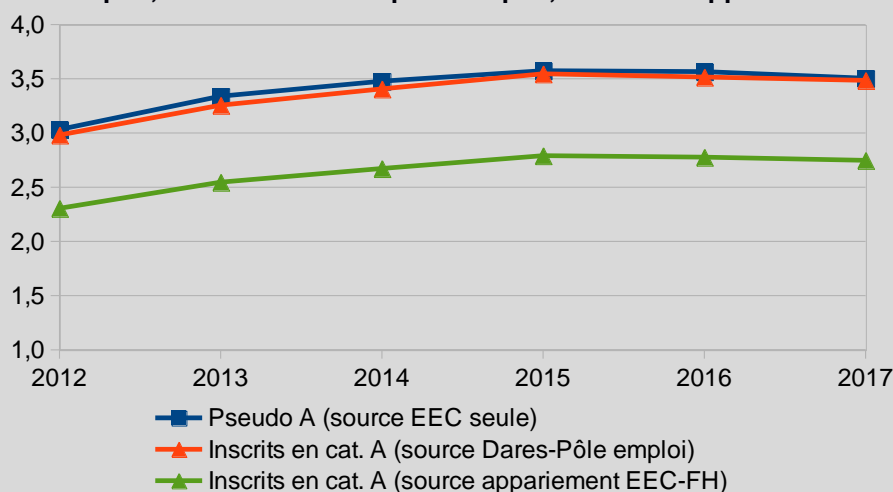
⁴ Dans le détail, en 2017, 13% sont inscrits le mois juste avant, 10 % sont inscrits le mois juste après, et 2 % sont inscrits à la fois le mois juste avant et le mois juste après.

Encadré 2 : Reconstituer un nombre d'inscrits en catégorie A déclarés dans l'enquête Emploi

Dans l'enquête Emploi, une question porte sur l'inscription à Pôle emploi (« êtes-vous actuellement inscrit comme demandeur d'emploi à Pôle emploi ou auprès d'un opérateur de placement ou d'une association d'insertion ? OUI/NON » - variable « OFFICC »). Il est donc possible d'estimer un nombre de personnes se déclarant inscrites à Pôle emploi dans l'enquête Emploi, toutes catégories confondues. En 2017, 5,4 millions de personnes se déclarent ainsi inscrites dans l'enquête Emploi, tandis que 6,3 millions sont véritablement inscrites d'après le fichier Historique de Pôle emploi.

Pour reconstituer le périmètre des seuls inscrits en catégorie A (appelés ici « pseudo A »), il faut croiser cette inscription déclarée avec d'autres informations et variables du questionnaire. En principe, les « pseudo A » sont les personnes se déclarant inscrites à Pôle emploi, mais ni en emploi, ni en formation (les inscrits en emploi devraient constituer les catégories B, C et E, et ceux en formation sont censés constituer la catégorie D). En prenant une définition assez restrictive de la formation (diplômante seulement, mais d'autres périmètres auraient été possibles), il est possible de reconstituer une série de « pseudo A » (cf. graphique E1, courbe bleue) qui coïncide très bien en niveau et en évolution avec la vraie série d'inscrits en catégorie A de source Dares-Pôle emploi (courbe rouge).

Graphique E1 : nombre de demandeurs d'emploi de catégorie A, selon les différentes sources (Dares-Pôle emploi, déclaré dans l'enquête Emploi, issue de l'appariement EEC-FH)

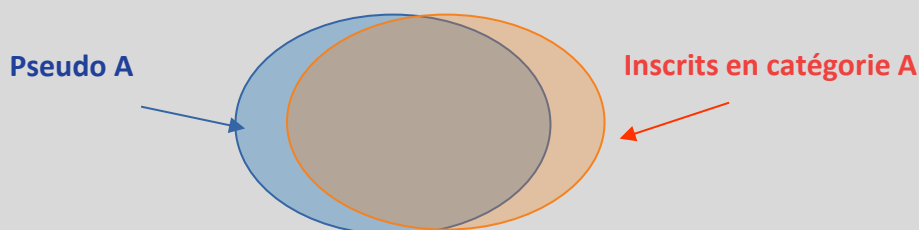


Champ : France métropolitaine, 15-64 ans.

Source : Appariement EEC-FH, EEC seule, Statistique trimestrielle Dares-Pôle emploi

Des études précédentes ont ainsi souligné la capacité de l'enquête Emploi, à travers ce concept de « pseudo A », à correctement retracer les niveaux et évolutions du vrai nombre d'inscrits (cf. Emploi, chômage, revenus du travail, édition 2018, pp. 20 à 22). L'appariement permet d'apporter un nouvel éclairage sur cet exercice. La correspondance apparemment parfaite entre « inscrits en catégorie A » selon Pôle emploi et « pseudo A » selon l'enquête Emploi est en fait le résultat d'erreurs qui se compensent (cf. schéma E2). Les inscrits en catégorie A non « pseudo A » (par exemple les 0,2 million d'inscrits en catégorie A en emploi) et les « pseudo A » non inscrits en catégorie A (par exemple, des personnes qui se déclarent inscrites à Pôle emploi alors qu'elles ne le sont pas en réalité) sont en nombre globalement équivalents, à chaque trimestre de la période considérée.

Graphique E2 : représentation schématique du recouvrement entre « pseudo A » et « inscrits en catégorie A »



Note : les « pseudo A » sont les personnes se déclarant en catégorie A d'après l'enquête Emploi seule.

Comme l'appariement EEC-FH ne permet de retrouver que 81 % des demandeurs d'emplois inscrits en catégorie A (cf. graphique E1, courbe verte, 2,7 millions de personnes en 2017 par exemple), dans la suite on s'attachera à expliquer pourquoi les courbes verte (inscrits en catégorie A appariés) et bleue (pseudo A selon l'enquête) évoluent de façon similaire sur la période.

En moyenne, sur la période 2012-2017, les inscrits en catégorie A non « pseudo A » sont constitués de 0,2 million de personnes en emploi (cf. tableau E3) et de 0,2 million de personnes qui se trompent à la question OFFICC (ne se déclarent pas inscrites alors qu'elles le sont). Les inscrits en catégorie A et en emploi ont été décrits au chapitre 2.2. Quant aux personnes qui ne se déclarent pas inscrites dans l'enquête alors qu'elles le sont, ce sont plutôt des hommes, des jeunes, des personnes peu diplômées, d'origine étrangère. Il s'agit aussi fréquemment de réponses par « proxy » (proche qui répond à la place de l'enquêté, lorsque celui-ci n'est pas disponible), ce qui peut expliquer les erreurs de déclaration. Parmi l'ensemble de ces personnes qui ne se déclarent pas inscrites alors qu'elles le sont, seules un sixième étaient effectivement inscrites le mois juste avant ou le mois juste après.

Symétriquement, en moyenne sur 2012-2017, les « pseudo A » non inscrits en catégorie A sont constitués de 0,5 million d'inscrits dans une catégorie autre que la catégorie A (principalement les chômeurs inscrits en catégorie B et D, détaillés au chapitre 2.3), et pour 0,7 million des personnes qui se disent inscrites à Pôle emploi mais ne le sont pas à la date de l'enquête d'après l'appariement (ces 0,7 million incluent des échecs d'appariement). Là aussi, les personnes qui se déclarent inscrites alors qu'elles ne le sont pas sont plutôt des jeunes, des personnes peu diplômées, d'origine étrangère et des réponses par proxy, et seules un quart d'entre elles étaient inscrites le mois juste avant ou juste après celui de l'enquête.

Tout au long de la période, les populations pour lesquelles le classement en catégorie A diverge (tableau E3, lignes (1) et (3)) sont très stables. Ainsi, les évolutions des pseudo A et des vrais inscrits en catégorie A sont très similaires à l'évolution de leur intersection (inscrits en catégorie A et pseudo A, tableau E3, ligne (2)), ce qui explique la bonne coïncidence des séries d'inscrits en catégorie A et de pseudo A, en niveau comme en évolution.

Tableau E3 : écarts entre la population des « pseudo A » et celle des « vrais » inscrits en catégorie A

	2012	2013	2014	2015	2016	2017
Inscrits en catégorie A non pseudo A (1)	0,4	0,4	0,4	0,4	0,4	0,4
- en emploi	0,2	0,2	0,2	0,2	0,2	0,2
- ne se déclarent pas inscrits dans l'enquête	0,2	0,2	0,2	0,2	0,2	0,2
Inscrits en catégorie A et pseudo A (2)	1,8	2,1	2,3	2,4	2,4	2,3
pseudo A non inscrits en catégorie A (3)	1,2	1,2	1,2	1,2	1,2	1,2
- inscrits en catégories B, C, D ou E	0,4	0,4	0,5	0,5	0,5	0,5
- non inscrits à Pôle emploi (dont échecs d'appariement)	0,8	0,8	0,7	0,7	0,7	0,7
Inscrits en catégorie A =(1)+(2)*	2,3	2,5	2,7	2,8	2,8	2,7
Pseudo A =(2)+(3)	3,0	3,3	3,5	3,6	3,6	3,5

Lecture : les « pseudo A » sont des personnes déclarant être inscrites et ni en emploi ni en formation diplômante. Dans la partie haute du tableau, parmi les inscrits en catégorie A, les « non pseudo A » sont donc soit des personnes en emploi (0,2 million), soit des personnes ne déclarant pas être inscrites (0,2 million). Les personnes en formation diplômantes étant très peu nombreuses, la ligne a été omise.

() Seules les personnes appariées sont comptées (cf. courbe verte dans le graphique E1)*

Champ : France métropolitaine, population des ménages, 15-64 ans.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Chapitre 3 – Évolutions 2013-2017

3.1 - Davantage d'inactifs inscrits en catégorie A et davantage de personnes en emploi et inscrites en catégories B et C

Comme vu au chapitre 2, tous les inscrits à Pôle emploi n'ont pu être identifiés. Mais cela n'empêche pas d'essayer de comprendre les évolutions des différentes catégories, tout d'abord car le taux de couverture est suffisamment élevé (81 % en catégorie A et 84 % pour l'ensemble des catégories de Pôle emploi) et de plus parce qu'il est stable sur la période 2013-2017. Les évolutions constatées sur le champ des inscrits appariés reflètent donc *a priori* principalement les évolutions de l'ensemble des inscrits, et non les variations du taux de couverture. Entre 2013 et 2017, l'écart entre le nombre de chômeurs au sens du BIT et le nombre d'inscrits en catégorie A issu du FH est passé de 0,3 à 0,8 million, soit +0,4 million, cette hausse pouvant être expliquée sur le sous-champ de l'enquête Emploi appariée.

Entre 2013 et 2017, la hausse des inactifs inscrits en catégorie A (0,3 million au global, dont 0,2 million hors halo et 0,1 million dans le halo, tableau 3.1) est celle qui contribue le plus à accroître l'écart entre les deux indicateurs d'inscrits en catégorie A et de chômage BIT, et explique les trois quarts de la divergence. La baisse du nombre de chômeurs non inscrits entre 2013 et 2017 (-0,1 million) explique le reste.

Tableau 3.1 : répartition des individus selon leur statut BIT et catégorie de demandeurs d'emploi, 2013 à 2017

En million de personnes

	2013	2014	2015	2016	2017	2013-2017
Chômeurs cat. A	1,6	1,7	1,7	1,6	1,5	-0,1
Chômeurs cat. B/C/E	0,2	0,2	0,2	0,2	0,2	0,0
Chômeurs cat. D	0,1	0,1	0,1	0,1	0,1	0,0
Chômeurs non inscrits, dont	0,9	0,9	0,9	0,9	0,8	-0,1
- non appariés mais se déclarant inscrits	0,2	0,2	0,2	0,2	0,2	0,0
- appariés mais non inscrits à la date de l'enquête	0,5	0,5	0,5	0,5	0,4	-0,1
- non appariés et ne se déclarant pas inscrits	0,2	0,2	0,2	0,2	0,2	0,0
En emploi cat. A	0,2	0,2	0,2	0,2	0,2	0,0
En emploi cat. B/C/E	1,5	1,6	1,8	1,9	1,9	0,5
En emploi non inscrits	23,8	23,7	23,5	23,5	23,7	-0,1
Halo cat. A	0,4	0,5	0,5	0,5	0,5	0,1
Halo cat. D	0,1	0,1	0,1	0,1	0,1	0,0
Halo non inscrits	0,8	0,8	0,8	0,8	0,8	0,0
Inactifs hors halo cat. A	0,3	0,3	0,3	0,4	0,4	0,2
Inactifs hors halo cat. D	0,1	0,1	0,1	0,1	0,1	0,0
Inactifs hors halo non inscrits	9,9	9,6	9,5	9,3	9,2	-0,6

Champ : France métropolitaine, 15-64 ans, ménages ordinaires.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Notons enfin que les personnes en emploi et inscrites en catégories B, C et E ont particulièrement augmenté sur cette période elles aussi (+0,5 million, essentiellement porté par la catégorie C). La simplification des règles pour pouvoir cumuler une indemnisation du chômage et des revenus d'activité, notamment permise dans le cadre de la convention d'assurance chômage de 2014, pourrait expliquer une partie de cette hausse. Au total, le nombre de personnes inscrites à Pôle emploi augmente de 0,7 million sur la période. Comme l'ensemble de la population des 15-64 ans est quasiment stable (-0,1 million), le nombre de personnes non inscrites diminue donc d'autant en contrepartie.

3.2 – Le recul de l'âge de départ à la retraite et la fin des dispenses de recherche d'emploi ont accru le nombre de seniors inactifs au sens du BIT et inscrits en catégorie A

Entre 2013 et 2017, les inactifs inscrits en catégorie A sont passés de 0,6 à 0,9 million, soit +0,3 million de personnes, dont +0,2 million hors halo (INA/A) et +0,1 million dans le halo (HAL/A). En ce qui concerne les inscrits en catégorie A inactifs hors halo (INA/A), les nouveaux inscrits sont plus souvent des seniors. La part des 50-64 ans est en effet passée de 38 % en 2013 à 46 % en 2017 (graphique 3.2).

Tableau 3.2 : caractéristiques des inactifs (halo et hors halo) inscrits en catégorie A, variations entre 2013 et 2017

	Halo cat. A			Inactifs hors halo cat. A		
	2013	2017	2013-2017	2013	2017	2013-2017
<i>Nombre de personnes représentées et appariées (en millions)</i>	0,4	0,5	0,1	0,3	0,4	0,2
Hommes	45	45	0	40	39	-1
Femmes	55	55	0	60	61	1
15-24 ans	13	11	-2	20	14	-7
25-49 ans	61	61	0	41	40	-1
50-64 ans	26	28	2	38	46	8
Se déclare spontanément chômeur	87	88	0	49	52	4
Se déclare spontanément en emploi	2	2	0	1	1	0
Se déclare spontanément inactif	11	11	0	50	47	-3
Halo1 (a recherché, mais pas disponible)	26*	26	0			
Halo2 (disponible, mais n'a pas recherché)	54*	48	-6			
Halo3 (ni recherches, ni disponible)	20*	26	6			
A déjà travaillé	84	90	6	76	86	11
Circonstances dans laquelle s'est terminé l'emploi antérieur (pour les personnes ayant déjà travaillé)						
- Fin de CDD	43	46	4	29	35	6
- Démission	7	8	1	8	9	1
- Licenciement	24	19	-5	27	20	-7
- Rupture conventionnelle	10	14	5	12	12	0
- Rupture pour maladie, invalidité	7	6	-1	13	15	3
- Autres raisons	10	7	-3	12	9	-4
Se déclare inscrit à Pôle Emploi	92	94	2	61	71	11
Déclare percevoir allocation chômage	53	52	-1	41	46	5
Déclare percevoir le RSA	19	18	-1	12	14	2
Motif d'inactivité						
Invalidité, problèmes de santé	12	12	0	29	38	9
S'occuper de sa famille, de son foyer	16	16	0	20	19	-1
Retraite	1	1	0	7	7	0
Études, formations	12	12	0	23	15	-9
Motifs de courte durée*	37	37	0	2	2	0
Découragement	16	15	-1	10	7	-3
Autres motifs ou motifs non renseignés	7	8	1	9	13	4

* : données 2014, car la définition des composantes du halo a été modifiée en cours d'année 2013 et n'est pas comparable à celle utilisée à partir de 2014.

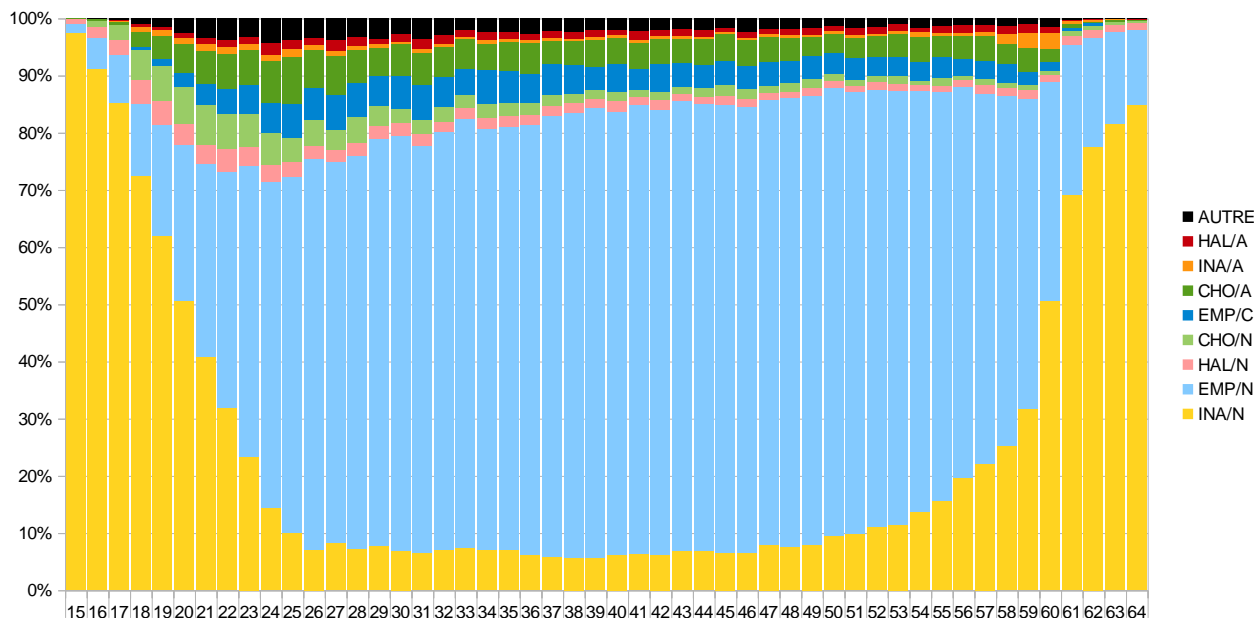
Champ : France métropolitaine, population des ménages, 15-64 ans, inactifs inscrits en catégorie A.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

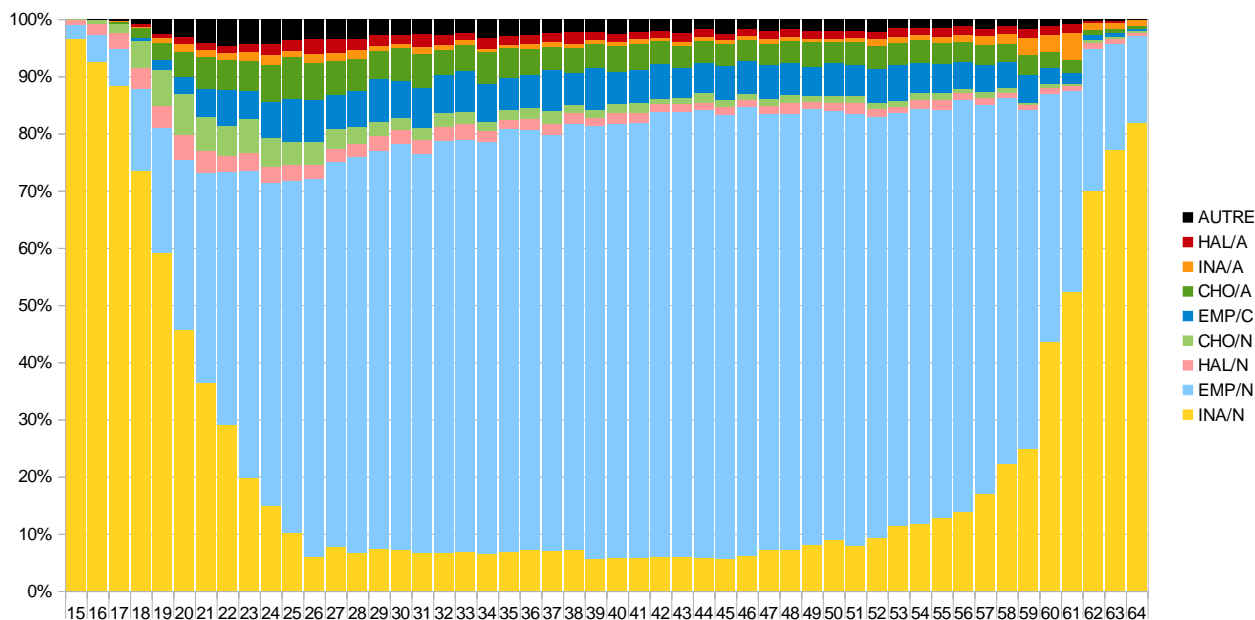
Le passage progressif de l'âge de départ à la retraite de 60 à 62 ans, ainsi que l'ensemble des réformes conduisant à augmenter l'âge individuel du départ à taux plein, peuvent expliquer une partie de la hausse de la catégorie des INA/A sur la période. En effet, ces réformes maintiennent sur le marché du travail des personnes qui sont susceptibles de rester inscrites à Pôle emploi une année supplémentaire, sans être pour autant chômeuses au sens du BIT. Ainsi, entre 2013 et 2017, le surcroît d'inactifs hors halo de catégorie A à

l'approche du passage à la retraite s'est accentué et décalé de la tranche d'âge des 58-60 ans en 2013 à la tranche d'âge des 59-61 ans en 2017 (cf. série INA/A en orange, en haut à droite sur le graphique 3.3). Entre 62 et 64 ans également, la proportion d'INA/A s'est légèrement accrue. Sur la période, ces réformes pourraient directement expliquer environ 50 000 personnes supplémentaires inscrites à Pôle emploi et inactives au sens du BIT, entre 61 et 64 ans. La hausse des inactifs inscrits en catégorie A ne résulte pas d'un effet démographique, dû à l'arrivée, aux âges élevés, d'une classe d'âge particulièrement nombreuse, mais bien d'un effet de comportement au sein de ces tranches d'âge.

Graphique 3.3 : Structure par situation BIT et Pôle emploi pour chaque âge, 2013 et 2017
Année 2013



Année 2017



Champ : France métropolitaine, 15-64 ans, ménages ordinaires.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

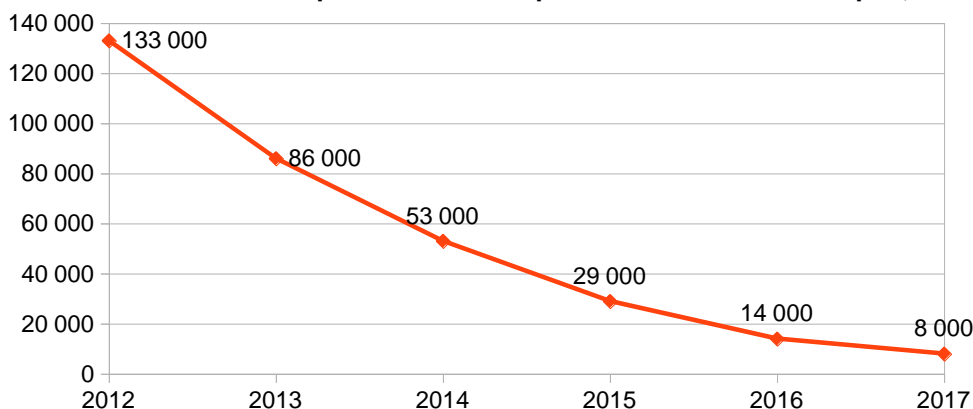
Par ailleurs, jusqu'à fin 2012, les seniors proches de l'âge de départ à la retraite pouvaient bénéficier, à leur demande, d'une dispense de recherche d'emploi (DRE). Les personnes dispensées de recherche n'apparaissent pas dans les nombres d'inscrits à Pôle emploi. L'accès à ce dispositif a été progressivement retreint à partir de 2009 et complètement supprimé en 2012. Les personnes qui étaient déjà entrées dans ce

dispositif avant sa suppression définitive ont cependant pu encore en bénéficier sur la période 2013-2017. En moyenne en 2013, 86 000 personnes étaient encore dispensées de recherche, pour seulement 8 000 en 2017 (graphique 3.4). La fin de ce dispositif aurait donc pu expliquer un maintien en catégorie A de 80 000 seniors en 2017 par rapport à 2013. On peut estimer qu'entre 30 000 et 80 000 d'entre eux se seraient retrouvés classés comme INA/A (et le reste principalement comme CHO/A), expliquant donc en partie la hausse de cette population sur la période.

L'estimation basse s'obtient en supposant que les personnes ne pouvant plus bénéficier de la DRE se seraient réparties conformément à la structure des seniors de 58 à 64 ans inscrits en catégorie A et non en emploi. En 2013, 34 % d'entre eux sont inactifs au sens du BIT (INA/A), 47 % sont au chômage au sens du BIT (CHO/A) et 19% sont dans le halo (HAL/A), soit 30 000 personnes sur les 80 000. Cette hypothèse semble cependant particulièrement forte, puisque les personnes qui auraient pu demander la DRE étaient, pour des raisons diverses (démotivation, problèmes de santé et/ou plus faible employabilité), moins susceptibles d'être chômeurs BIT. L'estimation haute (80 000) s'obtient en faisant l'hypothèse que les inscrits en catégorie A qui auraient pu demander la DRE auraient tous été classés inactifs hors halo au sens du BIT.

Sur la période 2013-2017, il est difficile de séparer de manière certaine les deux effets présentés (recul de l'âge de départ à la retraite et fin de la DRE), car ils concernent des individus similaires, en particulier les plus de 60 ans⁵. Les chiffres présentés ne peuvent donc pas être additionnés. Enfin, la hausse entre 2013 et 2017 de la proportion de personnes invoquant des problèmes de santé comme motif d'inactivité est étroitement liée à la hausse de la part des seniors. Quant à la proportion de personnes qui déclarent percevoir des allocations chômage, il est difficile d'interpréter sa hausse, car cette question n'est posée qu'aux personnes qui se déclarent inscrites à Pôle emploi, dont la proportion a également augmenté⁶.

Tableau 3.4 : Nombre de personnes en dispense de recherche d'emploi, de 2012 à 2017



Lecture : en moyenne en 2013, 86 000 personnes bénéficiaient d'une dispense de recherche d'emploi de la part de Pôle emploi.

Champ : France entière.

Source : Pôle emploi, Fichier National des Allocataires.

⁵ En effet, les personnes qui bénéficiaient encore d'une DRE en 2013 étaient pour la plupart âgées de 61 ans et plus.

⁶ La population des INA/A est, par construction, intégralement inscrite à Pôle emploi au moment de l'enquête. Or, seules 61 % se déclarent inscrites en 2013 et 71 % en 2017, loin des 100 % qu'on aurait pu attendre. La question sur l'inscription déclarée dans l'enquête est donc fragile, des erreurs de déclaration de la part des enquêtés ou de leurs proxys étant possibles. De plus, la qualité de cette question s'améliore entre 2013 et 2017, les enquêtés se trompant moins en 2017 qu'en 2013. Dans l'enquête Emploi, le fait de percevoir une indemnisation du chômage n'est demandé qu'aux personnes qui se déclarent inscrites à Pôle emploi. Le fait que le taux de personnes indemnisées progresse peut donc refléter simplement le fait que les enquêtés se sont moins trompés à la question sur leur inscription à Pôle emploi, et que la question sur le fait de toucher des indemnités chômage a ainsi été posée à davantage de personnes en 2017 qu'en 2013.

Dans le halo, les inscrits en catégorie A de 2017 ne semblent pas se distinguer nettement de ceux de 2013 (en termes de sexe, d'âge, de statut autodéclaré, de raisons de l'inactivité, cf. tableau 3.2), excepté qu'il semble s'agir un peu plus souvent de personnes souhaitant travailler, mais n'ayant pas fait de recherches actives et n'étant pas disponibles.

Sur la période 2013-2017, Pôle emploi a simplifié les processus d'inscriptions sur les listes (inscription sur internet, actualisation par SMS). Ces facilités pour s'inscrire et s'actualiser pourraient avoir contribué à accroître le nombre d'inscrits à Pôle Emploi, en bénéficiant à la fois à des chômeurs et à des personnes un peu plus éloignées du marché du travail.

De plus, sur cette même période, le nombre d'embauches en CDD a fortement augmenté [Jauneau et Vidalenc, 2019], ce qui a pu conduire davantage de personnes à demeurer inscrites à Pôle Emploi pour des périodes prolongées, qu'elles soient en emploi ou entre deux périodes d'emploi. Ainsi, parmi les inscrits en catégorie A inactifs au sens du BIT (dans le halo ou hors halo), la proportion de personnes ayant déjà travaillé a augmenté entre 2013 et 2017 (figure 3.2), tout comme la proportion de personnes dont le dernier emploi s'est terminé par une fin de CDD.

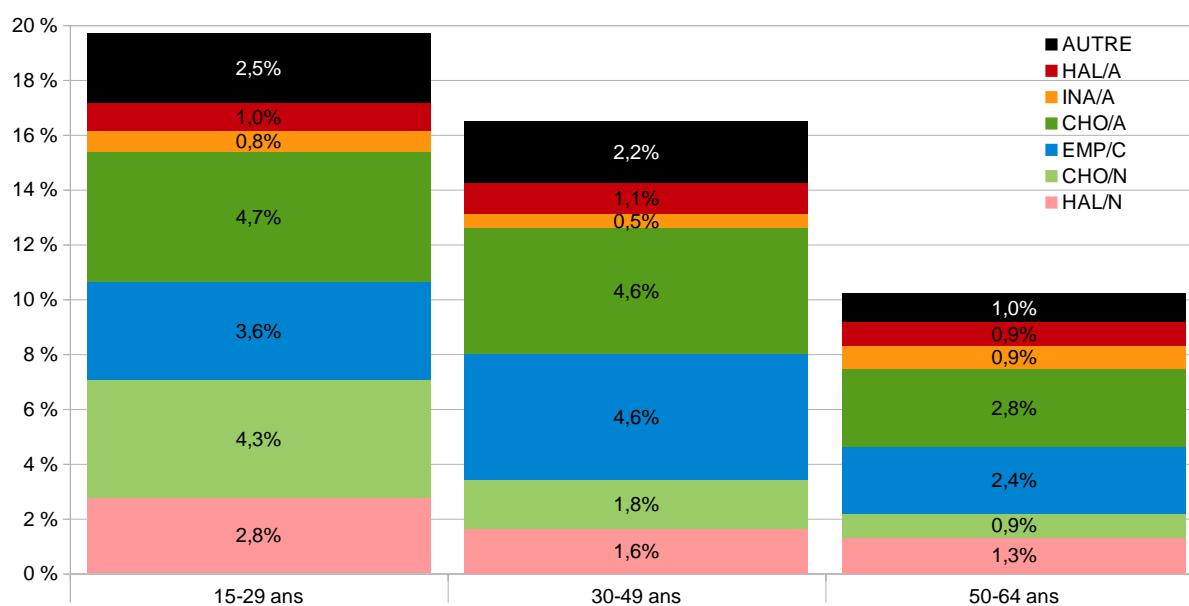
Les facilités à s'inscrire et à rester inscrits et la hausse des contrats en CDD constituent donc des pistes d'explications, mais que les données disponibles ne permettent toutefois pas de vérifier sans investigation complémentaire.

3.3 – Les jeunes ont davantage profité de l’embellie économique en 2017, ce qui a réduit le nombre de chômeurs non inscrits

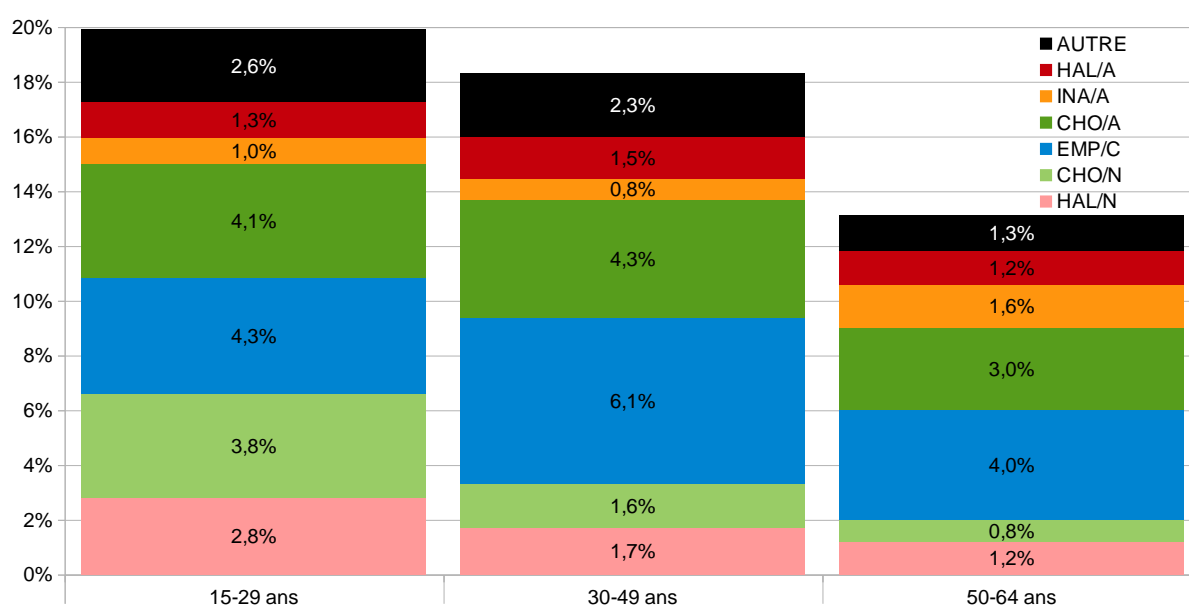
La baisse du nombre de chômeurs non inscrits à Pôle emploi a également contribué à accroître la divergence entre le nombre de chômeurs au sens du BIT et le nombre d’inscrits en catégorie A. Cette baisse a surtout été concentrée entre 2016 et 2017, lors d’une période d’embellie économique. En moyenne entre ces deux années, le nombre de personnes en emploi a augmenté de 0,2 million tandis que le nombre de chômeurs a baissé symétriquement de 0,2 million, dont -0,1 million de chômeurs inscrits en catégorie A et -0,1 million de chômeurs non inscrits (tableau 3.1). Les créations d’emploi de 2017 ont en effet beaucoup profité aux jeunes, lesquels sont plus souvent chômeurs non inscrits. La part des chômeurs non inscrits (CHO/N) a ainsi davantage reculé parmi les 15-29 ans (-0,5 %) que dans les autres tranches d’âge (graphique 3.5).

Graphique 3.5 : Structure par situation BIT et catégorie d’inscription à Pôle emploi par tranche d’âge (hors EMP/N et INA/N), 2013 et 2017

Année 2013



Année 2017



Note : Les catégories EMP/N et INA/N ne sont ici pas représentées, afin d’alléger le graphique et de mieux mettre en évidence les autres situations, moins fréquentes. Les tranches d’âge ont également été légèrement modifiées dans ce graphique, par rapport au reste du document.

Champ : France métropolitaine, 15-64 ans, ménages ordinaires.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l’enquête Emploi.

3.4 - Le chômage, un état plus instable en 2017 qu'en 2013

Le statut de demandeur d'emploi en catégorie A a plus « d'inertie » que celui de chômeur BIT. Les changements de statut BIT sont en effet plus rapides et fréquents que les changements de catégorie à Pôle emploi. En moyenne sur 2013-2017, environ 60 % des chômeurs BIT le sont encore au trimestre suivant. Pour les inscrits en catégorie A, 70 % le sont encore au trimestre suivant. En effet, être malade, en vacances ou découragé pendant quelques jours ou semaines ne fait pas forcément changer de catégorie à Pôle emploi alors que cela peut faire basculer ponctuellement un chômeur vers l'inactivité selon les critères du BIT. Cette différence explique que, en coupe un trimestre donné, on puisse mesurer davantage de personnes inscrites en catégorie A que de chômeurs BIT, toutes choses égales par ailleurs.

La question est de savoir si ce phénomène s'est amplifié au cours du temps, et contribue donc à l'écart croissant entre les deux agrégats. On constate effectivement qu'entre 2013 et 2017, les frontières entre chômage et inactivité sont devenues de plus en plus poreuses. En 2013, 18 % des chômeurs un trimestre T basculaient vers l'inactivité au trimestre T+1, contre 21 % en 2017 (cf. tableau 3.6). Réciproquement, en 2013, 22 % des chômeurs un trimestre donné étaient en provenance de l'inactivité le trimestre précédent, contre 26 % en 2017 (cf. tableau 3.7). Les transitions ont en revanche peu évolué pour les catégories de Pôle emploi (cf. tableaux 3.8 et 3.9).

Tableau 3.6 : situation au trimestre T+1 des personnes qui étaient au chômage au sens du BIT au trimestre T, moyennes sur les années 2013 à 2017

	Situation en T+1		
	CHO	EMP	HAL/INA
2013	61%	21%	18%
2014	60%	21%	19%
2015	59%	22%	19%
2016	58%	22%	19%
2017	56%	23%	21%

Lecture : en moyenne en 2013, parmi les personnes qui étaient au chômage un trimestre T, 21 % ont basculé vers l'emploi, 18 % vers l'inactivité (halo ou hors halo), et 61 % sont restées au chômage au trimestre T+1.

Champ : France métropolitaine, population des ménages, 15 à 64 ans.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Tableau 3.7 : situation au trimestre T des personnes qui sont au chômage au sens du BIT au trimestre T+1, moyennes sur les années 2013 à 2017

Situation en T	2013	2014	2015	2016	2017
CHO	60%	59%	59%	58%	57%
EMP	18%	18%	18%	18%	18%
HAL/INA	22%	23%	23%	24%	26%

Lecture : en moyenne en 2013, parmi les personnes qui étaient au chômage au trimestre T+1, 18 % étaient en emploi, 22 % étaient inactives (halo ou hors halo), et 60 % étaient déjà au chômage au trimestre T.

Champ : France métropolitaine, population des ménages, 15 à 64 ans.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Tableau 3.8 : situation au trimestre T+1 des personnes qui étaient inscrites en catégorie A au trimestre T, moyennes sur les années 2013 à 2017

	Situation en T+1		
	Cat. A	Cat. B/C/D/E	Non inscrits
2013	70%	16%	14%
2014	71%	16%	13%
2015	71%	17%	12%
2016	71%	17%	12%
2017	71%	17%	13%

Lecture : en moyenne en 2013, parmi les personnes qui étaient inscrites en catégorie A un trimestre T, 70 % sont restées inscrites en catégorie A, 16 % ont basculé vers les catégories B, C, D ou E, et 14 % ne sont plus inscrites à Pôle emploi au trimestre T+1.

Champ : France métropolitaine, population des ménages, 15 à 64 ans.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Tableau 3.9 : situation au trimestre T des personnes qui sont inscrites en catégorie A au trimestre T+1, moyennes sur les années 2013 à 2017

Situation en T	2013	2014	2015	2016	2017
Cat. A	70%	71%	71%	71%	71%
Cat. B/C/D/E	12%	12%	12%	12%	13%
Non inscrits	18%	18%	17%	17%	16%

Lecture : en moyenne en 2013, parmi les personnes qui étaient inscrites en catégorie A au trimestre T+1, 12 % provenaient d'une autre catégorie (B, C, D ou E), 18 % n'étaient pas inscrites à Pôle emploi, et 70 % étaient déjà inscrites en catégorie A au trimestre T.

Champ : France métropolitaine, population des ménages, 15 à 64 ans.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Bibliographie

Jauneau Y. et Vidalenc J. (2019), « Les salariés en contrats courts : des allers-retours plus fréquents entre emploi, chômage et inactivité », Insee première n°1736, Février.

Marrakchi A (2018), « Comparaison sur la période récente entre l'évolution du chômage au sens du BIT et celle du nombre de demandeurs d'emploi en fin de mois inscrits à Pôle Emploi », emploi chômage et revenus du travail édition 2018, Insee références

Gaignon A., Lieurade C. (2018), « Le contrôle de la recherche d'emploi : l'impact sur le parcours des demandeurs d'emploi », éclairages et synthèses, Pôle emploi, août.

De Waroquier S., Dixte C., Guillaneuf J., Minni C., Rebière S., Rey M. (2018), « Emploi, chômage, population active en 2017 : nouvelle accélération de l'emploi salarié privé et amplification de la baisse du chômage », Dares Analyses n°031, juillet.

Beck S., Vidalenc J. (2018), « une photographie du marché du travail en 2017. La baisse du chômage se confirme : -0,7 point par rapport à 2016 », Insee première n°1694, avril.

Insee (2018), « Le taux de chômage diminue de 0,7 point au quatrième trimestre 2017 », Informations rapides n°38, février.

Gonthier P., Vinceneux K. (2017), « Des chômeurs qui travaillent ? Les 7 visages de l'activité réduite », DARES document d'études n° 212, juillet 2017

Bessone A.-J., Cabannes P.-Y., Marrakchi A. (2016), « Halo autour du chômage : une population hétérogène et une situation transitoire », emploi chômage et revenus du travail édition 2016, Insee références

Marrakchi (2016), « Comparaison sur la période récente entre l'évolution du chômage au sens du BIT et celle du nombre de demandeurs d'emploi en fin de mois inscrits à Pôle emploi », note de conjoncture de l'Insee, juin .

Dares (2013) : « Emploi chômage, population active : bilan de l'année 2012 », Dares Analyses n°037, juin

Beauvoir R., Chanteloup G., Minni C., Nguyen L., Rémy V., Rey M. (2013), « Emploi chômage, population active : bilan de l'année 2012 », Dares Analyses n° 037, juin.

Ourliac B., Rochut J. (2013), « Quand les demandeurs d'emploi travaillent. Avec la crise le nombre de demandeurs d'emploi en activité atteint son plus haut niveau », Dares Analyses n°002, Dares, janvier.

Annexe 1 – précisions méthodologiques

1.1 Fonctions « Compged » et « Soundex » sous SAS

Fonction Compged

La fonction Compged est une généralisation de la distance de Levenshtein, qui mesure la dissemblance entre deux chaînes de caractères. Il s'agit du nombre minimum d'opérations permettant de passer de la chaîne de caractères 1 à la chaîne 2, les principales opérations étant insertion, suppression, remplacement d'une lettre ou d'un caractère, ajout d'un caractère (à la fin) ou troncature, etc.

L'instruction « call compost » permet d'affecter des coûts, éventuellement différenciés, aux différentes opérations. Ici, un coût de 1 a été affecté de façon identique à toutes les opérations : il n'est pas plus coûteux d'insérer une lettre, ou un espace que d'en supprimer une ou d'en intervertir deux.

Fonction Soundex

La fonction Soundex encode une chaîne de caractères en un code phonétique représentant sa prononciation. La fonction Soundex conserve la première lettre puis supprime les voyelles (A E I O U Y) ainsi que les consonnes H et W.

Exemple : Soundex(ALEX) → ALX

Les consonnes (autres que H et W) sont regroupées en 6 classes, selon la façon dont elles utilisent la langue, les dents et les lèvres pour se prononcer.

Classe 1 : B F P V

Classe 2 : C G J K Q S X Z

Classe 3 : D T

Classe 4 : L

Classe 5 : M N

Classe 6 : R

Les consonnes sont enfin converties en un code numérique, celui correspondant à leur classe.

Ex : ALX → A42

La fonction Soundex a été conçue pour la langue anglaise, mais elle s'est ici pourtant révélée plus efficace que ses variantes francophones. En français, un grand nombre de lettres ne se prononcent pas, souvent lorsqu'elles sont en fin de mot, mais pas seulement : par exemple le « r » à la fin de Roger, le « s » et le « n » dans François, le « e » dans Jeanne, etc. Dans les autres langues, en général (sauf la lettre « h »), toutes les lettres se prononcent (Frances, Jihane, Roger, etc.). C'est peut-être pour cette raison que le Soundex en anglais s'est avéré plus efficace que sa version francophone pour encoder phonétiquement les prénoms d'origine étrangère, car dans les autres langues étrangères aussi, en général, toutes les lettres se prononcent.

1.2 Un nombre de prénoms qui peut être différent entre le Fichier Historique et l'enquête Emploi

Dans le Fichier Historique, il arrive souvent que plusieurs prénoms soient donnés (cf. tableau A1). Il n'est pas toujours possible de savoir s'il s'agit de l'ensemble des prénoms de l'état civil (prénom + 2^e prénom + 3^e prénom, etc.), ou d'un prénom composé.

Exemple : Jean Pierre Paul. S'agit-il de

→ Jean-Pierre, Paul (prénom composé + 2^e prénom) ?

→ Jean, Pierre, Paul (prénom + 2^e prénom + 3^e prénom) ?

Dans l'enquête Emploi, en général, un seul prénom est fourni. Les enquêtés n'ont en effet pas besoin de décliner tout leur état civil. Si plusieurs prénoms sont fournis, c'est qu'il s'agit sauf erreur d'un prénom composé.

Tableau A1 : nombre de prénoms renseignés selon le FH ou l'EEC

nombre de prénoms renseignés	FH	EEC
1	75%	93%
2 ou +	25%	7%

Champ : 2012 à 2017, France métropolitaine, personnes de 15 à 64 ans.

Source : EEC 2012 à 2017 (Insee), FH 2012 à 2017 (Pôle emploi).

On crée donc 2 variables, pour stocker les 2 possibilités de prénom du Fichier Historique: une avec les 2 premiers prénoms (Jean Pierre, au cas où il s'agisse d'un prénom composé), appelée « prénom long », une avec le 1^{er} prénom seulement (Jean), appelée « prénom court ».

La comparaison entre les paires et les calculs des distances (scripturale, anagramme et phonétique) se fait alors sur ces deux variables de prénoms (prénom_long_FH et prenom_EEC d'une part, puis prénom_court_FH et prenom_EEC d'autre part) de façon à maximiser les chances d'appariement. La distance minimale entre ces deux comparaisons est retenue.

1.3 Géolocalisation

La Division Statistiques et analyses urbaines (DSAU) de l'Insee fournit, pour chaque adresse, son géoréférencement, c'est-à-dire sa géolocalisation (couple de coordonnées x,y) ainsi que son géocodage (attribution des codes de zonage pour les QPV, Iris, ZUS, ZFU et QVA). Avant de géoréférencer le fichier d'adresse, la DSAU effectue certains traitements : corrections ou ajout des codes communes (à partir des codes postaux et libellés de communes), nettoyage et normalisation des adresses, etc.

La géolocalisation consiste à appairer chaque adresse du fichier à géolocaliser avec une adresse du référentiel et à associer le x,y de l'adresse du référentiel à celle du fichier en cas d'appariement réussi. Lorsque la voie et le numéro sont identifiés, l'appariement est considéré comme un succès. Lorsque la voie est identifiée mais pas le numéro, les coordonnées (x,y) sont alors imputées aléatoirement dans la voie. Dans ces cas, on considère ici que les coordonnées (x,y) ont été à peu près bien positionnées et donc que la géolocalisation a réussi également. Enfin, lorsque la voie n'est pas retrouvée dans le référentiel, le couple x,y est imputé aléatoirement dans la commune. On considère ici qu'il s'agit alors d'un échec de géolocalisation, car le placement des coordonnées x,y est trop incertain compte tenu du niveau de précision ici nécessaire (chercher à appairer au sein d'une commune donnée).

Le taux de réussite de la géolocalisation est plus faible dans le Fichier Historique que dans l'enquête Emploi : dans le Fichier Historique, 5 % des adresses sont en échec de géolocalisation contre seulement 0,6 % des adresses dans l'enquête Emploi (cf. tableau A2). Les adresses fiscales de l'enquête Emploi ont en effet été davantage certifiées en amont par les services fiscaux.

Tableau A2 : taux de réussite de la géolocalisation des adresses

	FH	EEC
Voie et numéro identifiés	85%	89%
Voie identifiée seulement	10%	11%
Voie inconnue	5%	0,6%

Champ : France entière.

Source : EEC 2012 à 2017 (Insee), FH 2012 à 2017 (Pôle emploi).

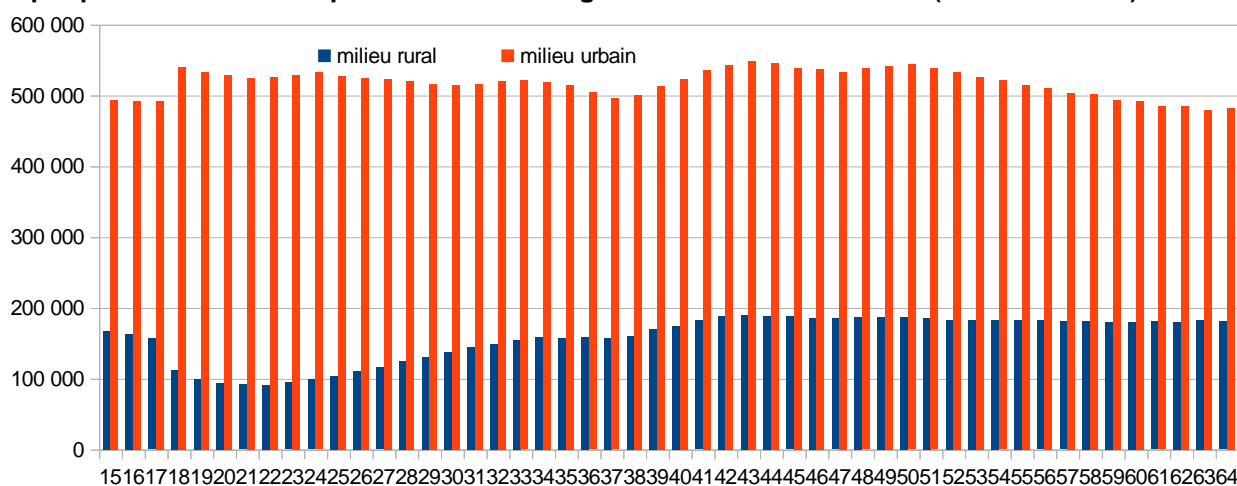
1.4 Probabilité théorique de « faux-positifs »

On calcule ici la probabilité, pour une personne donnée (« monsieur A »), qu'il y ait une autre personne dans la commune qui ait la même date de naissance et le même prénom que monsieur A.

- probabilité d'avoir le même jour et mois de naissance : 1/365. Cette probabilité est indépendante des autres facteurs.

- probabilité d'avoir la même année de naissance : on suppose ici que cette probabilité dépend du milieu, rural ou urbain. En milieu urbain, toutes les tranches d'âge sont représentées dans les mêmes proportions à peu près (cf. graphique A.3). La probabilité que deux personnes entre 15 et 64 ans aient la même année de naissance y est donc 1/50 environ. En milieu rural, tous les âges ne sont pas équi-représentés, avec un déficit de jeunes notamment. Ici, on suppose qu'en petite commune tous les résidents aient un âge compris entre 50 à 64 ans (hypothèse volontairement très défavorable de façon à obtenir un majorant de la vraie probabilité). Dans ce cas, la probabilité que deux personnes aient la même année de naissance est 1/15.

Graphique A3 : Nombre de personnes selon l'âge et le milieu de résidence (rural ou urbain)



Champ : France métropolitaine, 15-64 ans.

Source : Recensement de la population 2015, fichier INDREG.

- probabilité que deux personnes aient le même prénom : 0,0075. Cette probabilité a été estimée de façon empirique dans l'enquête Emploi et le Fichier Historique. La méthode ici mise en œuvre repose sur des tirages multiples effectués dans les deux bases de données. Un individu, nommé par exemple « Maxime », est tiré au hasard dans l'enquête Emploi, puis comparé à 50 000 individus du Fichier Historique tirés au hasard également, mais partageant la même année de naissance (car le prénom est considéré comme fortement dépendant de l'année de naissance). Le prénom Maxime est comparé aux prénoms des 50 000 personnes tirées dans le Fichier Historique sur la base de ce qui est présenté au chapitre 1.2 (distances scripturale, anagramme et phonétique) et une variable « meme_prenom » est construite, valant 1 si les prénoms concordent, 0 sinon. La moyenne de cette variable sur les 50 000 individus du Fichier Historique tirés au sort indique la probabilité pour l'individu « Maxime » qu'une personne du Fichier Historique aie le même prénom et la même année de naissance :

$\text{Proba}(\text{même prénom} \mid \text{prénom} = \text{« Maxime »}) = \text{moy}(\text{meme_prenom})$.

Cette opération est ensuite répétée 1 000 fois (pour 1 000 individus différents, de façon à faire varier le prénom de départ) et la probabilité finale empirique est la moyenne des 1 000 probabilités ainsi obtenues.

- probabilité que deux personnes aient le même sexe : 1. En effet, le prénom détermine le sexe, sauf à de rares exceptions (prénoms masculins et féminins, comme Dominique, Claude...).

On estime ainsi la probabilité, pour une personne A, qu'une autre personne B aie la même date de naissance et le même prénom :

En milieu rural, $\text{proba} = 1/365 \times 1/15 \times 0,0075 \times 1 = 1,4 \times 10^{-6}$

La proba qu'il y ait un faux-positif au sein d'une commune en milieu rural dépend enfin de la taille de la commune. Pour une commune de 1000 habitants par exemple, dont monsieur « A », la probabilité est donc $999 \times 1,4 \times 10^{-6} = 0,1 \%$ (999 autres personnes sont susceptibles d'avoir le même prénom et la même date de naissance, avec une probabilité de $1,4 \times 10^{-6}$ pour chacune).

La question est alors de savoir jusqu'à quelle taille de commune maximale on peut maintenir ce raisonnement tout en conservant une probabilité de faux-positifs très faible, par exemple inférieure à 0,5 %.
Pour 4 000 habitants, la probabilité de faux-positif s'écrit donc $3999 \times 1,4 \times 10^{-6} \approx 0,5 \%$, ce qui est le seuil recherché.

En grande commune, un raisonnement semblable est tenu.

La probabilité pour une personne A qu'une autre personne B aie la même date de naissance et le même prénom est alors $1/365 \times 1/50 \times 0,0075 \times 1 = 4,1 \times 10^{-7}$ (toutes les tranches d'âge sont à peu près équi-réparties, ici). Dans un rayon de 400 mètres autour de monsieur A (soit un cercle de superficie $\pi R^2 = 0,5 \text{ km}^2$), avec de plus l'hypothèse d'une densité de population égale à celle de Paris (20 000 habitants au km^2 , hypothèse volontairement très défavorable de façon à obtenir un majorant de la probabilité recherchée), on a donc 10 000 personnes dans ce rayon de 400 mètres (dont monsieur A), et la probabilité de faux-positif s'écrit donc $9\,999 \times 4,1 \times 10^{-7} \approx 0,5 \%$.

1.5 Une méthode d'appariement robuste aux faux-positifs

La méthode d'appariement est susceptible de générer des « faux-positifs », c'est-à-dire des individus appariés à tort. On cherche dans cette partie à quantifier leur nombre.

Pour cela, des 16,9 millions d'individus du Fichier Historique, un sous-échantillon de 0,4 million de personnes est extrait (soit de taille à peu près équivalente au nombre d'individus distincts ayant été enquêtés dans les enquêtes Emploi de 2012 à 2017). Ce sous-échantillon est alors apparié avec les 16,5 millions d'individus restant dans le Fichier Historique. L'ensemble de la chaîne de l'appariement statistique est reproduite, ainsi que les différents critères mis en place pour accepter ou rejeter les paires (distance globale $<0,6$, grappes de l'enquête, appariement sans l'adresse dans les communes de moins de 4 000 habitants, appariement dans un rayon de 400m dans les autres communes, appariement en relâchant le jour, le mois et l'année de naissance). Seule la reprise manuelle n'est ici pas reproduite.

En principe, comme les 0,4 million d'individus extraits sont différents des 16,5 millions restant, un appariement parfait ne devrait accepter aucune paire potentielle. En pratique, ici, 2 350 individus ont été appariés, vraisemblablement à tort, dont :

- 450 au titre de la distance globale $<0,6$;
- 0 au titre de l'exploitation des grappes ;
- 330 au titre des communes de moins de 4 000 habitants ;
- 170 au titre du rayon de 400m dans les autres communes ;
- et 1 400 en relâchant les contraintes sur le jour, le mois ou l'année de naissance (400 sans l'année de naissance, 100 sans le mois de naissance, 900 sans le jour de naissance).

Cela représenterait donc un taux de faux-positif de l'ordre de 0,6 %, surtout concentrés dans les étapes où la contrainte sur le jour, le mois et l'année de naissance est relâchée. À une adresse donnée, il n'est pas si rare que cohabitent deux personnes distinctes ayant le même prénom et une grande partie de leur date de naissance en commun.

Toutefois, un certain nombre de ces faux-positifs identifiés déclarent une adresse qui correspond à celle d'un centre communal d'action sociale (CCAS) ou à une communauté. Les CCAS proposent une fonction de « domiciliation administrative » : les personnes qui n'ont pas de domicile fixe peuvent déclarer l'adresse du CCAS afin d'y faire adresser leur courrier, ce qui peut concerner des inscrits à Pôle emploi. Un certain nombre d'adresses correspondent aussi à des communautés (foyers, centres d'hébergement et de réinsertion, etc.). Une grande partie des faux-positifs s'explique donc ici par le fait que de nombreuses personnes sont enregistrées à Pôle emploi à ces adresses, et il n'est donc pas si rare que deux d'entre elles partagent le même prénom et la même date de naissance (ou une partie de celle-ci). Dans l'appariement entre l'enquête Emploi et le Fichier Historique, cet écueil devrait être nettement réduit, puisque l'enquête ne porte que sur les personnes vivant en logement ordinaires, et exclue donc les CCAS et les communautés. De ce fait, le taux de faux-positif réel de l'appariement EEC-FH devrait donc être inférieur à 0,6 %.

1.6 Traitement des libellés en clair

Les personnes qui ne remplissent pas les critères pour être considérées comme au chômage au sens du BIT sont invitées à expliquer pourquoi elles n'ont pas recherché d'emploi (variable RAISNREC), ne sont pas disponibles (NONDIC), ou ne souhaitent pas travailler (RAISNSOU). Un certain nombre de motifs leurs sont proposés dans le questionnaire de l'enquête Emploi (voir détail des modalités des trois variables ci-dessous) mais les personnes qui ne se reconnaissent pas dans la liste peuvent aussi déclarer un motif « autre » et le préciser dans un libellé en clair.

RAISNREC : Pourquoi n'avez-vous pas recherché d'emploi au cours des quatre semaines du ... au ... ?

1. Car vous attendiez le résultat de démarches antérieures (concours, entretien, autorisation pour vous installer)
2. Vous faisiez des études ou suiviez une formation
3. Pour vous occuper d'enfants ou d'une personne dépendante
4. Vous aviez pendant cette période des problèmes de santé
5. Vous étiez en vacances
6. Vous pensez qu'il n'y a pas d'emploi disponible dans votre domaine de compétence
7. Vous pensez que vous ne trouverez pas d'emploi
8. Vous n'avez pas réussi à dégager du temps avec votre travail actuel
9. Autre

NONDIC : Pour quelle raison, principalement, n'êtes-vous pas disponible pour occuper un emploi dans les deux semaines ?

1. Pour achever vos études ou votre formation
2. Parce que vous ne pouvez pas quitter votre emploi actuel dans les deux semaines
3. Parce que vous gardez des enfants (y compris congé maternité) ou bien parce que vous vous occupez d'une personne dépendante
4. Pour cause de maladie temporaire (moins d'un mois)
5. Pour cause de longue maladie (ou d'invalidité)
6. Pour cause de vacances
7. Parce que vous avez ou allez déménager
8. Pour d'autres raisons

RAISNSOU : Pourquoi ne souhaitez-vous pas travailler actuellement ?

1. Vous faites des études ou suivez une formation
2. Vous vous occupez d'enfants ou d'une personne dépendante
3. Vous avez des problèmes de santé
4. Vous pensez qu'il n'y a pas de travail disponible
5. Vous avez déjà pris votre retraite
6. Autre

En raison de la proportion élevée de ces motifs « autres » (25 % des motifs évoqués par les personnes dans les situations HAL/A et INA/A par exemple), un nouveau codage de ces libellés en clair a été réalisé, permettant de les reclasser dans les modalités préexistantes. Les deux tiers environ des libellés en clair ont ainsi été codés et reclassés, principalement dans les motifs « s'occuper de sa famille, de son foyer » et dans les « motifs de courte durée » (cf. tableau A4).

Dans le détail, les personnes en cours d'inscription à une formation, ou en attente de la débiter ont été reclassées dans le motif « études/formations ». Les personnes déclarant un handicap, une blessure, une opération chirurgicale à venir ont été reclassées dans le motif « problèmes de santé ». Les personnes se déclarant « au foyer » par choix, pour s'occuper de leur conjoint ou en raison d'une activité non rémunérée (artistique, associative, sportive, politique), les personnes se déclarant enceintes, ainsi que celles invoquant un événement familial (décès du conjoint, divorce ou séparation, mariage, fête religieuse) ont été classées dans le motif « s'occupe de sa famille, de son foyer ». Les personnes se déclarant trop proches de la date officielle de leur retraite ont été reclassées dans le motif « retraite ». Les personnes se déclarant en congés sabbatique ont été classées dans le motif « vacances » et celles invoquant un départ à l'étranger ou des travaux réalisés dans leur logement ont été reclassées dans le motif « déménagement ». Les personnes se déclarant trop âgées ou trop jeunes, démotivées ou démoralisées ont été reclassées dans le motif « découragement ». Enfin, les personnes présentant des difficultés d'ordre administratif (en cours de procédure ou de litige, que ce soit au civil ou en lien avec leur emploi précédent, ou en attente d'un titre de séjour ou de son renouvellement), celles qui passent le permis de conduire et celles en attente de résultats d'un concours ou d'un agrément ou en cours de formalités de création d'entreprise ont été considérées comme

« en attente de résultats de démarches antérieures ». Enfin, les personnes ayant déjà trouvé un emploi qui commence plus tard ont également été classées en « attente de résultats de démarches antérieures ».

Tableau A4 : répartition des libellés en clair codés, en fonction du motif dans lequel ils ont été reclassés

Motif de reclassement	part de libellés en clair
S'occuper de sa famille, de son foyer	39%
Motifs de courte durée*	28%
Études, formations	11%
Problèmes de santé	10%
Retraite	7%
Découragement	5%

* : *Vacances, déménagement, en attente de résultats de démarches antérieures...*

Champ : libellés en clair des autres raisons de ne pas avoir fait de recherche, de ne pas être disponible, de ne pas souhaiter travailler.

Source : EEC 2012 à 2017.

Annexe 2 – tableaux et graphiques complémentaires

Tableau A5 : catégories socioprofessionnelles des personnes inscrites en catégorie B et C (source base appariement EEC-FH), VS ensemble des actifs occupés (source EEC), données au T4 2017

Catégories B

Données brutes

T4-2017	Enquête Emploi		Base appariée	
Total catégorie B	26 063 193	100,0%	439 661	100,0%
00 Non renseigné	57 229	0,2%	2 784	0,6%
11 Agriculteurs sur petite exploitation	137 306	0,5%	0	0,0%
12 Agriculteurs sur moyenne exploitation	56 409	0,2%	0	0,0%
13 Agriculteurs sur grande exploitation	198 298	0,8%	0	0,0%
21 Artisans	782 572	3,0%	8 240	1,9%
22 Commerçants et assimilés	671 841	2,6%	9 241	2,1%
23 Chefs d'entreprise de 10 salariés ou plus	157 959	0,6%	618	0,1%
31 Professions libérales	437 193	1,7%	4 073	0,9%
33 Cadres de la fonction publique	474 537	1,8%	1 639	0,4%
34 Professeurs, professions scientifiques	729 668	2,8%	3 406	0,8%
35 Professions de l'information, des arts et des spectacles	307 748	1,2%	32 024	7,3%
37 Cadres administratifs et commerciaux d'entreprise	1 405 354	5,4%	2 463	0,6%
38 Ingénieurs et cadres techniques d'entreprise	1 388 182	5,3%	329	0,1%
42 Professeurs des écoles, instituteurs et assimilés	933 553	3,6%	23 627	5,4%
43 Professions intermédiaires de la santé et du travail social	1 573 836	6,0%	20 862	4,7%
44 Clergé, religieux	8 941	0,0%	0	0,0%
45 Professions intermédiaires administratives de la fonction publique	448 538	1,7%	1 301	0,3%
46 Professions intermédiaires administratives et commerciales des entreprises	1 923 964	7,4%	10 620	2,4%
47 Techniciens	1 203 391	4,6%	5 767	1,3%
48 Contremaîtres, agents de maîtrise	565 888	2,2%	2 605	0,6%
52 Employés civils et agents de service de la fonction publique	2 280 878	8,8%	39 900	9,1%
53 Policiers et militaires	489 484	1,9%	7 624	1,7%
54 Employés administratifs d'entreprise	1 418 731	5,4%	18 966	4,3%
55 Employés de commerce	1 162 664	4,5%	21 190	4,8%
56 Personnels des services directs aux particuliers	1 760 191	6,8%	109 289	24,9%
62 Ouvriers qualifiés de type industriel	1 091 855	4,2%	9 451	2,1%
63 Ouvriers qualifiés de type artisanal	1 340 410	5,1%	20 070	4,6%
64 Chauffeurs	676 117	2,6%	11 124	2,5%
65 Ouvriers qualifiés de la manutention, du magasinage et du transport	486 387	1,9%	4 740	1,1%
67 Ouvriers non qualifiés de type industriel	922 777	3,5%	25 997	5,9%
68 Ouvriers non qualifiés de type artisanal	756 262	2,9%	34 077	7,8%
69 Ouvriers agricoles	215 030	0,8%	7 634	1,7%

Note : Les modalités surreprésentées sont colorées en orange et celles sous-représentées sont colorées en bleu.

Champ : France métropolitaine, ménages ordinaires, 15-64 ans, personnes en emploi au sens du BIT et inscrites en catégorie B à Pôle emploi, T4 2017.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Catégories C

Données brutes

T4-2017	Enquête Emploi		Base appariée	
Total catégorie C	26 063 193	100,0%	1 139 935	100,0%
00 Non renseigné	57 229	0,2%	2 913	0,3%
11 Agriculteurs sur petite exploitation	137 306	0,5%	238	0,0%
12 Agriculteurs sur moyenne exploitation	56 409	0,2%	855	0,1%
13 Agriculteurs sur grande exploitation	198 298	0,8%	339	0,0%
21 Artisans	782 572	3,0%	5 856	0,5%
22 Commerçants et assimilés	671 841	2,6%	12 695	1,1%
23 Chefs d'entreprise de 10 salariés ou plus	157 959	0,6%	343	0,0%
31 Professions libérales	437 193	1,7%	3 044	0,3%
33 Cadres de la fonction publique	474 537	1,8%	3 120	0,3%
34 Professeurs, professions scientifiques	729 668	2,8%	6 908	0,6%
35 Professions de l'information, des arts et des spectacles	307 748	1,2%	21 752	1,9%
37 Cadres administratifs et commerciaux d'entreprise	1 405 354	5,4%	8 810	0,8%
38 Ingénieurs et cadres techniques d'entreprise	1 388 182	5,3%	7 875	0,7%
42 Professeurs des écoles, instituteurs et assimilés	933 553	3,6%	28 104	2,5%
43 Professions intermédiaires de la santé et du travail social	1 573 836	6,0%	36 783	3,2%
44 Clergé, religieux	8 941	0,0%	574	0,1%
45 Professions intermédiaires administratives de la fonction publique	448 538	1,7%	6 677	0,6%
46 Professions intermédiaires administratives et commerciales des entreprises	1 923 964	7,4%	47 100	4,1%
47 Techniciens	1 203 391	4,6%	27 099	2,4%
48 Contremaîtres, agents de maîtrise	565 888	2,2%	7 956	0,7%
52 Employés civils et agents de service de la fonction publique	2 280 878	8,8%	126 842	11,1%
53 Policiers et militaires	489 484	1,9%	9 852	0,9%
54 Employés administratifs d'entreprise	1 418 731	5,4%	68 948	6,0%
55 Employés de commerce	1 162 664	4,5%	60 976	5,3%
56 Personnels des services directs aux particuliers	1 760 191	6,8%	210 399	18,5%
62 Ouvriers qualifiés de type industriel	1 091 855	4,2%	62 325	5,5%
63 Ouvriers qualifiés de type artisanal	1 340 410	5,1%	69 357	6,1%
64 Chauffeurs	676 117	2,6%	32 541	2,9%
65 Ouvriers qualifiés de la manutention, du magasinage et du transport	486 387	1,9%	53 853	4,7%
67 Ouvriers non qualifiés de type industriel	922 777	3,5%	144 567	12,7%
68 Ouvriers non qualifiés de type artisanal	756 262	2,9%	47 359	4,2%
69 Ouvriers agricoles	215 030	0,8%	23 875	2,1%

Note : Les modalités surreprésentées sont colorées en orange et celles sous-représentées sont colorées en bleu.

Champ : France métropolitaine, ménages ordinaires, 15-64 ans, personnes en emploi au sens du BIT et inscrites en catégorie C à Pôle emploi, T4 2017.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Tableau A6 : type de contrat de travail et type d'horaires de travail par catégorie, T4 2017

Données brutes

T4-2017	Enquête Emploi		Base appariée	
Total catégorie B	26 063 193	100,0%	439 661	100,0%
CDI	19 631 119	75,3%	167 980	38,2%
CDD	2 407 205	9,2%	181 563	41,3%
Intérim	731 777	2,8%	56 540	12,9%
Autre emploi (saisonnier, apprenti, ...)	402 419	1,5%	469	0,1%
Non salarié	2 890 673	11,1%	33 109	7,5%
Temps partiel de moins de 15 heures	706 477	2,7%	143 470	32,6%
Temps partiel de 15 à 29 heures	2 456 259	9,4%	120 228	27,3%
Temps partiel de 30 heures ou plus	1 302 284	5,0%	11 789	2,7%
Temps complet de moins de 30 heures	329 779	1,3%	11 530	2,6%
Temps complet de 30 à 34 heures	313 358	1,2%	6 105	1,4%
Temps complet de 35 à 39 heures	12 216 419	46,9%	79 255	18,0%
Temps complet de 40 heures ou plus	8 349 489	32,0%	27 928	6,4%
Pas d'horaire habituel ou horaire habituel non déclaré	389 128	1,5%	39 356	9,0%

Données brutes

T4-2017	Enquête Emploi		Base appariée	
Total catégorie C	26 063 193	100,0%	1 139 935	100,0%
CDI	19 631 119	75,3%	327 091	28,7%
CDD	2 407 205	9,2%	464 485	40,7%
Intérim	731 777	2,8%	311 552	27,3%
Autre emploi (saisonnier, apprenti, ...)	402 419	1,5%	4 156	0,4%
Non salarié	2 890 673	11,1%	32 651	2,9%
Temps partiel de moins de 15 heures	706 477	2,7%	30 245	2,7%
Temps partiel de 15 à 29 heures	2 456 259	9,4%	208 850	18,3%
Temps partiel de 30 heures ou plus	1 302 284	5,0%	63 899	5,6%
Temps complet de moins de 30 heures	329 779	1,3%	14 506	1,3%
Temps complet de 30 à 34 heures	313 358	1,2%	10 355	0,9%
Temps complet de 35 à 39 heures	12 216 419	46,9%	511 139	44,8%
Temps complet de 40 heures ou plus	8 349 489	32,0%	216 901	19,0%
Pas d'horaire habituel ou horaire habituel non déclaré	389 128	1,5%	84 040	7,4%

Note : Les modalités surreprésentées sont colorisées en orange et celles sous-représentées sont colorisées en bleu.

Champ : France métropolitaine, ménages ordinaires, 15-64 ans, personnes en emploi au sens du BIT et inscrites en catégorie B ou C à Pôle emploi, T4 2017.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Tableau A7 : statut de l'établissement employeur et nombre d'employeurs par salarié, catégories B et C confondues, T4 2017

Données brutes

T4-2017	Enquête Emploi		Base appariée	
Total catégories B et C	23 172 520	100,0%	1 513 836	100,0%
<i>Entreprise privée ou association</i>	16 420 634	70,9%	1 084 132	71,6%
<i>Entreprise publique (EDF, La Poste, SNCF, etc.)</i>	956 403	4,1%	39 669	2,6%
<i>État</i>	2 294 782	9,9%	60 008	4,0%
<i>Collectivités territoriales</i>	1 796 845	7,8%	98 873	6,5%
<i>Hôpitaux publics</i>	890 655	3,8%	18 918	1,2%
<i>Sécurité sociale</i>	33 374	0,1%	0	0,0%
<i>Particulier</i>	687 194	3,0%	201 025	13,3%
<i>Non renseigné</i>	92 633	0,4%	11 211	0,7%
<i>1 employeur principal</i>	22 616 540	97,6%	1 389 674	91,8%
<i>2 employeurs</i>	329 911	1,4%	54 418	3,6%
<i>3 employeurs ou plus</i>	169 069	0,7%	61 597	4,1%
<i>Non renseigné</i>	57 000	0,2%	8 147	0,5%

Note : Les modalités surreprésentées sont colorisées en orange et celles sous-représentées sont colorisées en bleu.

Champ : France métropolitaine, ménages ordinaires, 15-64 ans, personnes en emploi au sens du BIT et inscrites en catégorie B ou C à Pôle emploi, T4 2017.

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Tableau A8 : caractéristiques des HAL/A et INA/A, en fonction du motif d'inactivité, en 2017.

	HAL/INA inscrits en catégorie A						
	Problèmes de santé	S'occupe de sa famille, foyer	Retraite	Études, formations	Motifs de courte durée*	Découragé	Autres motifs ou non renseignés
Nombre de personnes (millions)	0,22	0,16	0,03	0,12	0,19	0,10	0,09
Hommes	49	10	54	45	50	54	45
Femmes	51	90	46	55	50	46	55
15-24 ans	4	9	0	37	12	9	14
25-49 ans	34	82	2	53	58	36	49
50-64 ans	62	9	98	10	29	55	37
Se déclare spontanément chômeur	54	67	44	53	94	94	84
Se déclare spontanément en emploi	1	0	1	2	2	1	3
Se déclare spontanément inactif	45	33	56	45	4	5	13
Halo1 (a recherché, pas disponible)	5	18	3	14	26	15	12
Halo2 (disponible, n'a pas recherché)	5	15	4	19	57	51	26
Halo3 (ni recherches, ni disponible)	19	18	3	17	13	7	6
Inactif hors halo	71	49	90	50	3	27	56
A déjà travaillé	95	84	97	73	93	88	87
Se déclare inscrit à Pôle Emploi	79	83	54	76	94	93	86
Déclare percevoir allocation chômage	53	42	40	43	57	52	51
Déclare percevoir le RSA	17	20	6	9	16	19	15

Note : le motif a été compilé à partir des différents motifs de ne pas avoir recherché d'emploi (RAISNREC), de ne pas être disponible (NONDIC), de ne pas souhaiter travailler (RAISNSOU), et par quelques modalités du statut autodéclaré (SP00 : invalidité, retraite, études/formations, congé parental) afin de limiter autant que possible la modalité « autre motifs ». L'étude des libellés en clair a aussi permis de reclasser un grand nombre de motifs « autres » dans la liste des motifs précédents.

** : vacances, déménagement, attente de résultats de démarches antérieures, maladie de courte durée, a déjà trouvé un emploi qui commence ultérieurement, en cours de création d'entreprise.*

Champ : France métropolitaine, 15-64 ans, ménages ordinaires, inactifs au sens du BIT (halo ou hors halo).

Source : Dares, Insee, Pôle emploi, appariement du Fichier Historique et de l'enquête Emploi.

Série des documents de travail de la DSDS

- F1903-G2019/04** : L'économie numérique fausse-t-elle le partage volume-prix du PIB ?
Lorraine Aeberhardt, Florian Hatier, Marie Leclair, Benoît Pentinat et Jean-Denis Zafar
- F 1902** : Les modèles de contrôle des effets de structure : comparaisons et application à une analyse de la disparité départementale des revenus non salariés moyens - Christophe Bertran et Laurianne Salembier
- F1901 - G2019/01** : Effets de moyen terme d'une hausse de TVA sur le niveau de vie et les inégalités : une approche par microsimulation - Mathias André et Anne-Lise Biotteau
- F1808** : La construction d'un indice des prix des logements neufs - Thomas Balcone
- F1807** : Estimer les effectifs de couples de personnes de même sexe au recensement : expérimentation d'une solution de validation du sexe par le prénom - Élisabeth Algava et Sébastien Hallépée
- F1806** : L'effet d'une variation du montant de certains transferts du système socio-fiscal sur le niveau de vie : résultats sur 2016 à partir du modèle de microsimulation Ines (Cahier de variantes) - Maëlle Fontaine et Michaël Sicsic
- F1805** : Family, Firms and the Gender Wage Gap in France - Elise Coudin, Sophie Maillard et Maxime Tô
- F1804** : Mode de collecte et questionnaire, quels impacts sur les indicateurs européens de l'enquête Emploi ? - Klara Vinceneux
How do the collection mode and questionnaire used affect the European indicators in the French Labour Force Survey ? - Klara Vinceneux
- F1803** : Trois versions du taux d'effort en matière de logement - Pascal Godefroy
- F1802** : Heterogeneous exposure to labor earnings risk - Pierre Pora et Lionel Wilner
- F1801** : L'espérance de vie par niveau de vie Méthode et principaux résultats - Nathalie Blanpain
- F1708** : Les données fiscales de l'EDP : une nouvelle source d'informations sur les couples et les familles ? - Vianney Costemalle
- F1707** : La situation matrimoniale dans le recensement : impact de la refonte du questionnaire de 2015. - Guillemette Buisson
- F1706** : Indices de prix à la consommation - Patrick Sillard
- F1705** : Effet d'un choc d'inflation sur le revenu disponible et ses composantes deux ans après : une approche par microsimulation - Anne-Lise Biotteau et Maëlle Fontaine
- F1704** : Scanner data and quality adjustment - Isabelle Léonard, Patrick Sillard, Gaëtan Varlet et Jean-Paul Zoyem
- F1703** : Les structures familiales en France : comparaison entre le recensement, l'enquête famille et logements et l'enquête emploi - Guillemette Buisson et Aude Lapinte
- F1702** : Projections de la population active à l'horizon 2070 - Malik Koubi et Anis Marrakchi

- F1701** : Les taux marginaux effectifs de prélèvement pour les personnes en emploi en France en 2014 - Juliette Fourcot et Michaël Sicsic
- F1606** : Projections de population 2013-2070 pour la France : méthode et principaux résultats - Nathalie Blanpain et Guillemette Buisson
- F1605** : Les durées passées en famille monoparentale - Méthode d'estimation des durées et résultats - Vianney Costemalle
- F1604** : ESeG = European Socio economic Groups - Nomenclature socio-économique européenne - Monique Meron, Michel Amar, Charline Babet, Milan Bouchet-Valat, Fanny Bugeja-Bloch, François Gleizes, Frédéric Lebaron, Cédric Huguée, Étienne Penissat et Alexis Spire
- F1603** : Catégorie sociale d'après les déclarations annuelles de données sociales et catégorie sociale d'après le recensement : quels effets sur les espérances de vie par catégorie sociale ? Comparaison entre les déclarations annuelles de données sociales et les recensements de la population. Comparaison de méthodes d'estimation des espérances de vie - Vianney Costemalle
- F1602** : L'espérance de vie par catégorie sociale et par diplôme - Méthode et principaux résultats - Nathalie Blanpain
- F1601** : Échantillonnage des agglomérations de l'IPC pour la base 2015 - Laurence Jaluzot et Patrick Sillard
- F1508** : Worker-firm matching and the family pay gap: Evidence from linked employer-employee data - Lionel Wilner
- F1507** : Effet des nouvelles mesures sociales et fiscales sur le niveau de vie des ménages : méthodologie de chiffrage avec le modèle de microsimulation Ines - Mathias André, Marie-Cécile Cazenave, Maëlle Fontaine, Juliette Fourcot et Antoine Sireyjol
- F1506** : Nowcasting du taux de pauvreté par la micro-simulation - Maëlle Fontaine et Juliette Fourcot
- F1505/376-501** : Bilan du projet EDP++ - division Camap et division Enquêtes et études démographiques
- F1504** : Contrôles des rémunérations dans les déclarations annuelles de données sociales (DADS) - Une analyse exploratoire pour améliorer la détection des points atypiques - Claire Jacod
- F1503** : Précision de l'enquête Patrimoine 2010 - Pierre Lamarche et Laurianne Salembier
- F1502** : Pourquoi l'indicateur de pauvreté en conditions de vie baisse malgré la crise économique ouverte en 2008 ? Jean-Louis Pan ké shon
- F1501** : Évolution de la population de la France entre 1981 et 2011 : contributions de la fécondité, de la mortalité, du solde migratoire et de la structure de la pyramide des âges - Catherine Beaumel et Pascale Breuil-Genier
- F1410** : "Personal network" and retirement: Is retirement bad for friendship and good for family relationships ? Anne Laferrère
- F1409** : Retraités mais pas en retrait : La retraite pousse-t-elle à de nouvelles activités ? Anne Laferrère
- F1407** : Production "aval" de l'enquête emploi en continu EEC2 2013 - 20XX - Fabien Guggemos
- F1406 bis** : La constitution de l'échantillon démographique permanent de 1968 à 2012 - Stéphane Jugnot

- F1405 (tome 1)** : Hommes et femmes vivant en couple en 2009, 1999 et aux recensements précédents - Fabienne Daguet
- F1405 (tome 2)** : Hommes et femmes vivant en couple en 2009, 1999 et aux recensements précédents - Fabienne Daguet
- F1404** : L'addition est-elle moins salée ? La réponse des prix à la baisse de TVA dans la restauration en France - Quentin Lafféter et Patrick Sillard
- F1403** : Estimer les flux d'entrées sur le territoire à partir des enquêtes annuelles de recensement - Chantal Brutel
- F1402** : Une rotation de la main d'œuvre presque quintuplée en 30 ans : plus qu'un essor des formes particulières d'emploi, un profond changement de leur usage - Claude Picart
- F1401** : Calculs statistiques de stock et de flux sur la révision électorale 2012-2013 - Christelle Rieg