

Annexe 3

Sources et méthodes

En raison de sa richesse d'informations, l'échantillon interrégimes de retraités (EIR) est la base de données de référence pour l'étude des pensions de retraite. Il couvre en effet quasiment tous les régimes et fournit des informations individuelles extrêmement détaillées. Cette édition exploite la vague 2016 de l'échantillon. Les estimations pour les années 2017 à 2020 sont issues du modèle de microsimulation ANCETRE (Actualisation annuelle par calage pour l'estimation tous régimes des effectifs de retraités), construit à partir de l'EIR 2016 et calé sur les données de l'enquête annuelle auprès des caisses de retraite (EACR). Les résultats de cette enquête sont également exploités directement dans de nombreuses fiches de cet ouvrage. Enfin, l'échantillon interrégimes de cotisants (EIC) permet d'étudier les parcours de carrières des futurs retraités et des retraités actuels. Cette édition exploite la vague 2017 de l'échantillon.

L'échantillon interrégimes de retraités

Dans le cadre de la collecte de l'échantillon interrégimes de retraités¹ (EIR), les organismes de retraite renseignent les caractéristiques individuelles d'un échantillon de retraités : nature et montant des prestations qui leur sont versées, conditions de liquidation des droits à la retraite (âge de liquidation, taux de liquidation, durée de carrière validée). Le rapprochement, individu par individu, des informations en provenance des différents régimes est indispensable pour calculer le nombre de retraités et pour reconstituer la pension globale de chacun. En effet, un même retraité peut recevoir une pension de plus d'un régime de retraite à la fois : régime de base et régime(s) complémentaire(s) correspondant(s) le cas échéant, mais aussi d'autres régimes de base et complémentaires s'il a changé de régime au cours de sa carrière (retraité dit « polypensionné »). La somme des effectifs de chaque régime mesure donc le nombre total des pensions servies et non le nombre global de retraités couverts par ces régimes. Il n'est ainsi pas possible d'additionner le nombre de pensions versées par chaque régime pour calculer le nombre total de retraités : une telle méthode conduirait à des doubles comptes. L'EIR, qui

couvre les principaux régimes, permet de calculer le nombre de retraités et de reconstituer le montant de la retraite globale des personnes ainsi que ses éléments constitutifs.

La quasi-totalité des organismes de retraite obligatoire (régimes de base et régimes complémentaires) sont interrogés, soit plus de 70 régimes au total pour l'EIR 2016. En revanche, les retraites issues des régimes supplémentaires non obligatoires sont exclues du champ de l'échantillon, tout comme les revenus provenant de l'épargne individuelle. Le champ de l'EIR comprend donc tous les individus de l'échantillon qui perçoivent une pension de droit direct ou de droit dérivé d'une caisse de retraite légalement obligatoire. L'EIR 2016 a été élargi aux pensions d'invalidité versées par le régime général, de manière à couvrir la totalité des bénéficiaires de pensions d'invalidité.

L'EIR est un panel : les personnes appartenant à l'échantillon initial sont sélectionnées de nouveau à chaque vague d'enquête (excepté les personnes décédées entre deux vagues, celles des générations les plus âgées [les centenaires] et celles ajoutées lors du millésime précédent, afin de surreprésenter certaines générations). L'EIR portant sur la situation au 31 décembre 2016 est la huitième vague du panel, la première ayant

¹. Sa création est prévue par l'article 1^{er} de la loi n° 84-575 du 9 juillet 1984 et elle est encadrée par les articles R. 161-59 à R. 161-69 du Code de la Sécurité sociale.

eu lieu en 1988. Depuis 1988, l'opération a été reconduite tous les quatre ans. L'échantillon a été complété à chaque vague par de nouvelles générations et a été étendu aux personnes nées dans les DROM (depuis la vague de 2001), puis aux personnes nées à l'étranger (depuis la vague de 2004), pour une meilleure prise en compte de la population des retraités. Les personnes âgées de 34 à 54 ans ont également été ajoutées dans le champ de l'EIR 2008, afin de mieux couvrir l'ensemble des retraités. S'agissant de la nouvelle vague disponible, la taille de l'EIR a fortement augmenté : l'EIR 2016 compte environ 650 000 assurés. Il est apparié au panel Tous Salariés de l'Insee afin d'obtenir des informations complémentaires sur les salaires des dernières années de carrière. Il est également apparié à des données historiques de Pôle emploi et aux données fiscales anonymisées. La prochaine vague de l'EIR portera sur la situation au 31 décembre 2020.

Chaque retraité de l'EIR est pondéré, pour tenir compte du fait qu'un retraité retenu dans l'échantillon « représente » plusieurs autres retraités non échantillonnés. Ces pondérations sont calées, par la méthode statistique du calage sur marges, sur un très grand nombre de variables auxiliaires démographiques et relatives aux bénéficiaires des caisses (effectifs de retraités et montants moyens).

L'enquête annuelle auprès des caisses de retraite

L'enquête annuelle auprès des caisses de retraite² (EACR) porte sur les principaux régimes de retraite de base en matière de nombre de pensions servies (19 régimes en 2020, dont les principaux régimes spéciaux) et de retraite complémentaire (6 régimes en 2020) [tableau 1], ainsi que sur les données relatives à l'invalidité de la Caisse nationale d'assurance maladie (CNAM). Par le biais de cette enquête, la DREES collecte annuellement auprès de ces caisses de retraite des informations agrégées sur les bénéficiaires d'un droit direct ou d'un droit dérivé, vivants au 31 décembre de l'année.

2. En avril 2021, l'enquête a obtenu le label d'intérêt général et de qualité statistique, avec le caractère obligatoire (obligation de réponse). Depuis décembre 2021, les données de l'EACR sont disponibles sur l'espace Open Data de la DREES, rubrique Retraite : <https://data.drees.solidarites-sante.gouv.fr>.

Ces informations concernent principalement les montants moyens de pensions (avantage principal de droit direct et de droit dérivé), et les effectifs correspondant au champ de l'ensemble des bénéficiaires et au champ des nouveaux bénéficiaires au cours de l'année. Ces indicateurs sont détaillés selon le sexe, la génération et le lieu de naissance (en France ou à l'étranger). L'enquête recueille aussi des informations sur les nouveaux retraités de l'année écoulée qui perçoivent une pension avec une décote ou une surcote, et sur les motifs de la liquidation de leurs droits (handicap, carrière longue, etc.) par année de départ à la retraite et par génération. Des données relatives aux bénéficiaires de pensions d'invalidité, de rentes d'incapacité permanente ou d'un minimum de pension sont également collectées.

Les données de l'enquête annuelle réalisée par la DREES peuvent différer légèrement de celles publiées par les régimes de retraite dans leurs bilans statistiques annuels. Des concepts statistiques homogènes ont notamment été définis, afin d'obtenir des statistiques comparables entre les différents régimes. Les invalides de la fonction publique sont ainsi comptabilisés parmi les retraités dès lors qu'ils ont dépassé l'âge d'ouverture des droits à la retraite (voir fiche 23).

Estimation des grandeurs annuelles tous régimes : le modèle ANCETRE

En 2010, la DREES a conçu le modèle de micro-simulation statique ANCETRE pour estimer annuellement les grandeurs tous régimes à partir des données de l'EIR et des EACR (encadré 1). L'EIR 2016 et les données des EACR 2017, 2018, 2019 et 2020 ont permis de réaliser les vagues ANCETRE 2017, 2018, 2019 et 2020. L'EIR 2012 et les données des EACR 2013 à 2016 ont permis de réaliser les vagues ANCETRE 2013 à 2016. ANCETRE est notamment utilisé pour estimer sans double compte le nombre total de retraités et de nouveaux retraités, les montants de pension de droit direct et de droit dérivé, ainsi que les âges de liquidation.

Tableau 1 La diversité des régimes de retraite

	Régimes de base	Régimes complémentaires
Salariés du secteur privé	Caisse nationale d'assurance vieillesse (CNAV), régime général. Au 1 ^{er} janvier 2020, le régime général gère également les indépendants (anciens assurés de la SSI).	Arrco (salariés d'employeurs privés), Agirc (cadres), Ircantec (agents non titulaires de l'État et des collectivités publiques). Les régimes Agirc et Arrco ont fusionné au 1 ^{er} janvier 2019 en un seul régime, le régime Agirc-Arrco ¹ .
Salariés agricoles	Mutualité sociale agricole (MSA) aligné sur le régime général.	
Artisans Industriels et commerçants	La Sécurité sociale des indépendants (SSI, issue du RSI) est le régime obligatoire de Sécurité sociale qui assure la protection sociale des artisans, commerçants, industriels, et de certaines professions libérales. Au 1 ^{er} janvier 2020, le régime général (CNAV) gère également ces deux populations d'indépendants.	RSI complémentaire. Avant 2013, il y avait un régime complémentaire pour les artisans et un autre pour les commerçants.
Professions libérales	Régime de base (10 sections professionnelles regroupées au sein de la Caisse nationale d'assurance vieillesse des professions libérales [CNAVPL]). Caisse nationale des barreaux français (CNBF).	Régimes complémentaires obligatoires selon la section professionnelle. Avantages supplémentaires pour les praticiens et auxiliaires médicaux conventionnés.
Non-salariés agricoles	Mutualité sociale agricole (MSA).	Régime complémentaire obligatoire.
Fonctionnaires civils et militaires de l'État	Régime spécial géré par le Service des retraites de l'État (SRE).	Retraite additionnelle de la fonction publique (RAFP).
Fonctionnaire des hôpitaux et des collectivités locales	Régime spécial géré par la CNRACL.	
Salariés des entreprises publiques et autres régimes spéciaux	Régimes spéciaux de la SNCF, de la RATP, des industries électriques et gazières (CNIÉG), des clercs et employés de notaires (CRPCEN), de la Banque de France, des mines, etc.	
Salariés des cultes	Caisse d'assurance vieillesse, invalidité et maladie des cultes (Cavimac).	

1. Les cotisations de retraite complémentaire sont calculées sur la base des éléments de rémunérations brutes perçues par les salariés, par référence à l'assiette « Sécurité sociale ».

Dans cet ouvrage, l'EACR 2020 est utilisée pour les statistiques par régime en 2020, ANCETRE 2020 pour les statistiques tous régimes en 2020 et l'EIR 2016 pour les statistiques tous régimes en 2016. Les données relatives à des millésimes de l'EIR (2004, 2008, 2012 et 2016) sont calculées à partir de l'EIR et non du modèle ANCETRE. Ainsi, les évolutions entre 2015 et 2016 sont à prendre avec précaution, car les données 2015 s'appuient sur ANCETRE 2015, lui-même constitué à partir de l'EIR 2012, alors que les données 2016 se fondent sur l'EIR 2016. Pour les mêmes raisons, les évolutions de 2011 à 2012 et de 2007 à 2008 doivent également être analysées avec prudence.

L'échantillon interrégimes de cotisants (EIC 2017)

Afin d'étudier les retraites futures des personnes d'âge actif et les parcours de carrière des personnes déjà retraitées, la DREES, en collaboration avec l'ensemble des régimes de retraite, a mis en place un dispositif statistique interrégimes : l'échantillon interrégimes de cotisants (EIC). Environ 30 organismes, recouvrant une soixantaine de régimes, sont questionnés tous les quatre ans sur les droits à la retraite acquis par un échantillon anonyme de plus de 700 000 personnes de différentes générations. L'EIC est issu du rapprochement personne par personne d'informations

Encadré 1 Le modèle ANCETRE

Description et méthodologie

Le modèle Actualisation annuelle par calage pour l'estimation tous régimes des effectifs de retraités (ANCETRE) combine les sources statistiques pour conserver le caractère individuel des données de l'échantillon interrégimes de retraités (EIR) et le caractère actualisé des données de l'enquête annuelle auprès des caisses de retraite (EACR). La méthodologie est adaptée chaque année afin d'intégrer les changements réglementaires. En 2020, le modèle intègre la fusion CNAV-SSI ainsi que la mise en place d'une revalorisation différenciée selon le montant de pension totale. ANCETRE 2020 est un pseudo EIR 2020, construit à partir de l'EIR 2016 et de l'EACR 2020. Cette nouvelle base de données est utilisée pour fournir des statistiques au 31 décembre 2020.

Pour réaliser la version 2020 d'ANCETRE, dans un premier temps, l'échantillon des retraités fin 2019, issu de la version 2019 d'ANCETRE, est vieilli. Pour cela, les pondérations de la base sont réduites afin de prendre en compte les décès ayant eu lieu entre 2019 et 2020. Un modèle détermine ensuite, parmi les bénéficiaires d'une pension en 2019, ceux qui vont liquider de nouveaux droits en 2020 : des liquidations sont ajoutées à ces individus en 2020. Enfin, de nouveaux retraités liquidant leurs droits en 2020 sont également imputés. Finalement, les pondérations et les pensions sont calées, par la méthode du calage sur marges, de façon à assurer la cohérence entre les estimations d'ANCETRE et les données de l'EACR 2020. Pour l'édition 2022 du panorama, le modèle ANCETRE a été revu pour être également calé de façon plus précise sur les prélèvements sociaux concernant les pensions de retraites (la CSG notamment).

La prise en compte de la liquidation unique dans les régimes alignés

La mise en place de la liquidation unique des régimes alignés (Lura) à partir du 1^{er} juillet 2017 complexifie la réalisation d'ANCETRE 2017, 2018, 2019 et 2020. En effet, le modèle repose sur l'hypothèse que les données individuelles à partir de 2017 ont peu évolué par rapport à l'EIR 2016. Or la Lura modifie largement les conditions de liquidation des personnes polyaffiliées aux régimes alignés¹ (régime général, SSI et MSA) : leur pension est calculée comme si elles n'avaient appartenu qu'à un seul régime, déterminé en fonction de leur carrière. La Lura nécessite donc d'adapter l'étape de création de nouveaux liquidants dans ANCETRE.

Pour tenir compte de la réduction du nombre de liquidations par régime induite par la Lura, un régime verseur est attribué à chaque personne concernée par la Lura à partir des variables disponibles dans l'EIR (durées dans chaque régime, dates de première et dernière validations, existence de droits non-salariés agricoles). De même, le modèle prend en compte le fait que la Lura oblige l'assuré à liquider tous ses droits en même temps avec une pension calculée en tenant compte de l'ensemble des droits acquis au régime général et au régime des indépendants, puis revalorisée.

L'impact de la Lura sur les montants de pension des retraités concernés n'est pas directement modélisé, en revanche, dans les données individuelles. Cet impact est pris en compte par le biais du calage sur les données de référence des régimes depuis 2017. Pour les trois régimes concernés par la Lura, l'EACR fournit en effet des éléments (effectifs, masses) sur le champ spécifique des retraités Lura afin d'assurer ce calage.

1. Elle a également un impact sur le montant de retraite des personnes monoaffiliées à la SSI ou à la MSA, qui subissent le passage à un mode de calcul annualisé du salaire ou revenu d'activité annuel moyen (SAM).

provenant de ces régimes de retraite. Il recense tous les éléments utiles au calcul de leur pension et permet de retracer leur parcours de validation de trimestres. C'est le seul outil permettant une connaissance globale des droits à la retraite acquis par les actifs, sans reconstitution à partir de données déclaratives. Les informations collectées

dans l'EIC sur les durées d'assurance validées (voir annexe 1) ont des particularités (*encadré 2*). Des données sont également recueillies pour toutes les personnes auprès de Pôle emploi. Elles permettent de disposer de certaines périodes assimilées ouvrant des droits à retraite (chômage, pré-retraite). Enfin, le panel Tous actifs de l'Insee permet

de disposer notamment des revenus annuels d'une partie des personnes composant l'EIC.

La cinquième vague de l'EIC (EIC 2017) a été conçue pour représenter les personnes ayant entre 23 et 71 ans au 31 décembre 2017. L'échantillon, extrait du répertoire national d'identification des personnes physiques (RNIPP), est constitué de personnes nées en janvier, avril, juillet ou octobre, une génération paire sur deux entre la génération née en 1946 et celle née en 1954, puis toutes les générations paires jusqu'à celle née en 1994. Le taux de sondage est compris entre 2,2 % et 4,4 % selon la génération. Le champ de l'EIC comprend toutes les personnes de l'échantillon qui sont ou ont été affiliées (y compris les personnes déjà parties à la retraite) pour des droits directs à pension, au moins une fois au cours de leur carrière, à l'un des régimes de retraite partenaires de l'opération. Il inclut donc des personnes ayant travaillé durant de courtes périodes en France, qui peuvent avoir par ailleurs effectué une partie de leur carrière dans leur pays d'origine, ou des personnes ayant travaillé et validé des droits à retraite à l'étranger.

Parmi les personnes inscrites au RNIPP, 6 à 8 % de celles nées entre 1946 et 1988, 9 % de celles nées en 1990, et 11 % et 14 % de celles nées en 1992 et 1994 répondent aux critères de l'échantillon mais ne sont affiliées à aucune des caisses de retraite. Il peut s'agir soit de personnes n'ayant jamais cotisé en France (du moins à la date de référence de l'échantillon), soit de personnes ayant cotisé dans les régimes de base non couverts par l'EIC (qui représentent environ 0,01 % des cotisants).

Le prochain EIC portera sur les droits acquis au 31 décembre 2021.

Améliorations méthodologiques et révisions des données en entrée

Les données présentées dans ce panorama comportent des révisions sur les années antérieures à 2020, liées à des ajustements méthodologiques et à l'intégration de données modifiées. C'est le cas des données issues de l'EIR 2016, des modèles ANCETRE 2017 à 2019, et de l'enquête Retraite supplémentaire 2019. ■

Encadré 2 La mesure de la durée validée pour la retraite dans l'EIC

Les majorations de durée d'assurance pour enfants, les périodes validées à l'étranger et une partie des trimestres validés au titre du chômage non indemnisé, ainsi que certaines périodes d'apprentissage, ne sont pas intégrées dans le calcul des durées d'assurance validées analysées dans l'EIC. Ces informations ne sont souvent disponibles dans les systèmes d'information des caisses de retraite qu'au moment de la liquidation de la pension (et non au cours de la carrière).

Pour déterminer la nature des trimestres validés, il a été décidé de conserver par convention en priorité les trimestres cotisés au titre de l'activité professionnelle, puis les validations au titre de l'Assurance vieillesse des parents au foyer (AVPF), puis les trimestres validés au titre du service national, puis les trimestres assimilés de type chômage, formation, reconversion, préretraite, et enfin ceux de type maladie, maternité, invalidité, accident du travail.

Néanmoins, certains régimes ne distinguent pas les périodes assimilées des périodes d'emploi (comme ceux de la fonction publique, et la plupart des régimes spéciaux et des professions libérales). Elles sont donc comptabilisées dans les périodes d'emploi par ces régimes.

Pour en savoir plus

- > Présentation des sources sur le site de la DREES : <https://drees.solidarites-sante.gouv.fr>, rubrique Sources, outils et enquêtes, sous-rubrique Retraites.
- > **Aubert, P., Ducoudré, B.** (2012, décembre). Un nouveau modèle pour estimer annuellement des indicateurs « tous régimes de retraite confondus » : ANCETRE. *Retraite et société*, 1/2013 (64), p. 152-169.
- > **Aubert, P., Ducoudré, B.** (2011, septembre). Le modèle ANCETRE : Actualisation annuelle par Calage pour l'Estimation Tous Régimes des Effectifs de retraités. DREES, *Document de travail*, série Sources et Méthodes, 24.