

Source et Méthode

Enquête « Revenus fiscaux et sociaux »

Édition 2023



**Direction des Statistiques démographiques et sociales Département des ressources
et des conditions de vie des ménages
Division Revenus des ménages**

Avertissement

L'enquête Revenus fiscaux et sociaux (ERFS) mesure les niveaux de vie en associant aux données de l'enquête Emploi du quatrième trimestre de chaque année les revenus issus de sources administratives.

L'enquête Emploi a connu une refonte majeure en 2021 [[document du Département de l'emploi et des revenus d'activité, 2021](#)], qui a nécessité de revoir la chaîne de production de l'ERFS. L'adaptation de la production à cette nouvelle enquête Emploi a par ailleurs été l'occasion de rationaliser l'ensemble de la chaîne de production de l'ERFS et de revoir la méthode de correction de la non-réponse.

Depuis 2005, l'enquête Revenus fiscaux et sociaux (ERFS) fournit chaque année une évaluation du revenu disponible des ménages de l'année considérée, enrichie des informations sociodémographiques de l'enquête Emploi de l'Insee.

Elle s'appuie, comme les enquêtes Revenus fiscaux (ERF) de 1996 à 2005, sur les données de l'enquête Emploi en continu du 4^{ème} trimestre de l'année appariées avec les données administratives (déclarations de revenus) émanant de la direction générale des Finances publiques (DGFiP). En outre, les données sont depuis 2005 appariées avec celles des organismes sociaux afin de disposer des prestations effectivement perçues par les ménages (prestations sociales)¹. Elle évalue également les revenus générés par des produits financiers non soumis à l'impôt sur le revenu.

L'objectif de cette enquête est de mesurer un revenu économique : ce dont les ménages ont disposé au cours d'une période pour consommer ou épargner. La richesse de l'enquête réside dans les différents concepts de revenus qu'elle permet de mesurer.

¹ Les données ont été rétropolées sur 1996-2004.

Table des matières

Présentation générale de l'enquête.....	4
Champ et limites de l'enquête.....	5
Les améliorations introduites dans la série d'enquêtes Revenus fiscaux et sociaux (ERFS).....	6
Méthodologie et déroulement de l'enquête.....	9
1. L'échantillon de départ.....	9
2. L'appariement des données avec les fichiers fiscaux.....	9
3. L'appariement avec les fichiers sociaux.....	10
4. Le traitement de la non-réponse et le redressement.....	11
4.1 Traitement de la non-réponse.....	11
4.1.1 Non-réponse totale ou partielle ?.....	11
4.1.2 Le traitement de la non-réponse partielle : imputation des données manquantes.....	11
4.1.3 Les corrections de la non-réponse totale.....	14
4.2 Le redressement pour fluctuations d'échantillonnage.....	14
Glossaire.....	16

Présentation générale de l'enquête

L'enquête Revenus fiscaux et sociaux (ERFS) permet de produire des statistiques et de réaliser des études sur les revenus, le niveau de vie, la **pauvreté et les inégalités monétaires** des individus résidant dans des ménages ordinaires de France métropolitaine à l'aide d'indicateurs usuels d'analyse de la distribution des revenus (nombres, quartiles, déciles, moyenne, médiane, etc.) et d'indicateurs de structure et de composition des ménages. Elle permet ainsi l'analyse des revenus selon des critères sociodémographiques : **catégorie socioprofessionnelle** et âge des personnes composant le ménage, activité au sens du **BIT** (Bureau international du travail) de ces personnes, taille du ménage, etc. Elle permet également d'analyser comment se cumulent les différents types de revenus (salaires, chômage, pensions, revenus agricoles, industriels, commerciaux, non commerciaux, etc.) perçus par chaque membre du ménage.

L'ERFS s'appuie sur les données du 4^e trimestre de l'année N de l'enquête Emploi en continu (EEC, enquête trimestrielle depuis 2003), qui est une enquête par sondage, c'est-à-dire qu'un échantillon représentatif de l'ensemble de la population est interrogé tous les trimestres. L'ERFS ne constitue pas à proprement parler d'une enquête, puisque les données recueillies portant sur les revenus et les prestations sociales ne sont pas collectées par enquête mais par exploitation de sources administratives pour un échantillon représentatif de l'ensemble de la population.

Depuis 2021, la production de l'ERFS se déroule en deux étapes. La *première* consiste en un appariement du fichier de l'EEC avec les informations provenant des fichiers fiscaux et organismes sociaux de l'année N. Le principe de l'appariement consiste à essayer de retrouver les déclarations fiscales ou les prestations versées aux individus composant les ménages enquêtés à l'EEC au 4^e trimestre de l'année N. Concernant les prestations, la recherche s'effectue parmi les prestations versées à chaque allocataire l'année N par la caisse nationale d'allocations familiales (CNAF) et la caisse nationale de la mutualité sociale agricole (CCMSA), et les droits constatés et versés en décembre N par la caisse nationale de l'assurance vieillesse (CNAV). Cette étape était auparavant réalisée en deux temps, avec en premier lieu un appariement avec les fichiers fiscaux (qui permettait de récupérer les noms des individus de l'EEC, qui n'étaient pas disponibles dans les données d'enquête) et dans un second temps un appariement avec les fichiers sociaux. Depuis 2021, la nouvelle EEC permet de disposer du nom des individus dans l'enquête : les appariements fiscaux et sociaux peuvent être réalisés en parallèle, en utilisant tous les deux le nom des individus.

Dans un *second temps*, les informations sur les revenus des ménages absentes des sources administratives fiscales et sociales sont complétées par des estimations réalisées par l'Insee. Ces estimations concernent les revenus financiers non soumis à l'impôt, ou plus récemment, les dispositifs exceptionnels exonérés. Cette étape est indispensable pour passer du concept de **revenu déclaré** (ou revenu fiscal) à des concepts de revenus économiques (**revenu disponible**, **niveau de vie**, etc.). Ces estimations concernent également des revenus présents dans les sources administratives pour les individus qui n'ont pas pu être retrouvés dans ces fichiers (par exemple pour les salaires, pensions, revenus des indépendants, chômage ou pour des prestations). Enfin, cette enquête, comme toutes les enquêtes par sondage, fait l'objet de corrections de la non-réponse pour les individus de plus de 90 ans n'ayant pas été retrouvés dans les sources administratives² et d'un redressement afin d'être représentative au niveau de la France métropolitaine.

Ainsi, l'ERFS d'une année donnée fournit une approche du revenu disponible monétaire des ménages pour l'année considérée, enrichie des données de l'enquête Emploi collectées durant le 4^{ème} trimestre de l'année.

² Les individus de 90 ans et plus ne sont plus interrogés individuellement dans l'EEC. L'absence d'informations sur ces personnes rend l'imputation de leurs revenus trop fragile. Ces individus demeurent donc exclus du champ de l'ERFS lorsqu'ils ne sont retrouvés ni dans les fichiers fiscaux ni dans les fichiers sociaux. Le biais introduit par cette exclusion est corrigé par repondération des individus de 90 ans et plus appariés.

Champ et limites de l'enquête

Le champ couvert par l'enquête est celui des individus appartenant à des **ménages** dits « ordinaires » résidant en France métropolitaine. Sont donc exclues les personnes vivant en collectivité (maisons de retraite, communautés religieuses, cités universitaires, prisons, etc.) ainsi que les personnes sans abri.

L'enquête Revenus fiscaux et sociaux (ERFS) est représentative des ménages métropolitains (approximativement 97 % de la population totale française hors **DOM-COM**) et elle a pour unité statistique le ménage (et non l'individu ou le **foyer fiscal**). Plus précisément, jusqu'en 2020, l'échantillon de l'ERFS stricto sensu est constitué des ménages répondant à l'enquête Emploi et pour lesquels on a retrouvé au moins une déclaration fiscale ou bien un dossier social (dossier dans un des organismes sociaux). Depuis 2021, l'ensemble des ménages répondant à l'enquête Emploi sont conservés dans l'échantillon à l'exception des ménages de 90 ans et plus qui n'ont été retrouvés ni dans les fichiers fiscaux, ni dans les fichiers sociaux. Les autres ménages non retrouvés dans les données administratives sont désormais intégrés au champ de l'enquête et leurs revenus sont imputés à l'aide de modèles économétriques.

Par conception, l'enquête est tributaire de la législation fiscale qui peut faire varier le revenu déclaré d'une année sur l'autre. Les évolutions et les caractéristiques mêmes de la législation fiscale peuvent avoir une incidence sur le revenu déclaré (ou revenu fiscal). C'est par exemple le cas pour les revenus des indépendants (extension des régimes micro-entreprise, revenus des entrepreneurs individuels dont l'imposition est basée sur une évaluation forfaitaire, par exemple) ou les mises en place de dispositifs exonérés. L'enquête dépend également des modifications de la législation sociale en vigueur (modifications des caractéristiques des personnes bénéficiaires, etc.).

C'est également le cas de l'impôt sur le revenu, dont la prise en compte du paiement a évolué dans l'ERFS pour reconstituer le plus fidèlement possible le revenu disponible de l'année en cours. Jusqu'en 2012, l'ERFS de l'année N reconstituait le **revenu disponible** au titre des revenus déclarés pour l'année N et ne correspondait pas exactement au revenu disponible réellement perçu sur la période. Par exemple :

- Dans l'enquête ERFS de l'année N, le revenu disponible du ménage était égal au revenu déclaré en N, augmenté des revenus financiers non déclarés imputés l'année N et des prestations sociales perçues en N, le tout diminué de l'impôt sur le revenu de l'année N (payé en N+1). Les prestations reçues étaient relatives à la législation en vigueur l'année N, calculées par les caisses d'allocations familiales (CAF) en fonction des revenus déclarés en N-1 (ou N-2) ou des revenus perçus au cours du trimestre précédent et de la composition du ménage tout au long de l'année N (avec l'âge des enfants au cours de l'année N).

- Dans la réalité, le revenu disponible réel du ménage de l'année N était égal au revenu déclaré en N, diminué de l'impôt sur le revenu de l'année N-1 (c'est celui qui est payé en N), augmenté des prestations reçues en N.

Ce décalage entre ce qui était reconstitué dans les enquêtes ERFS et le revenu disponible réel n'avait guère d'incidence pour les ménages dont les caractéristiques (en termes de revenus, situation familiale, etc.) étaient stables au cours du temps.

A partir de l'ERFS 2013, un changement méthodologique a été introduit pour pallier ce défaut et introduire l'impôt sur le revenu de l'année N-1 payé en N dans le calcul du revenu disponible du ménage. Un appariement est réalisé avec le fichier issu de la DGFIP de l'année N-1 (pour récupérer l'impôt sur le revenu de l'année N-1) ; et pour les 8,2 % n'ayant pas pu être appariés, une imputation du montant a été effectuée (montant nul pour un jeune étudiant de 18 à 25 ans dont c'est la première déclaration d'impôts, montant de l'année N déflaté pour les autres cas - changement de composition familiale notamment -, montant imputé pour les non retrouvés). Ce changement méthodologique a également été introduit dans la série de l'ERFS 2012 rétropolée afin de pouvoir estimer des évolutions à méthodologie constante et d'évaluer de ce changement : il conduit à un impact sur le taux de pauvreté égal à 0,3 point de pourcentage.

A partir de l'ERFS 2019, avec la mise en place du prélèvement à la source, l'impôt sur le revenu est désormais principalement calculé à partir du fichier issu de la DGFIP de l'année N : l'impôt prélevé à la source (sur les salaires, chômage, retraite, etc.) ou acquitté par le déclarant sous forme d'acomptes (revenus d'indépendants, revenus fonciers, etc.) est directement remonté dans des cases de la déclaration de l'année N. L'impôt tient également compte des crédits d'impôts et de la régularisation d'impôts sur les revenus N-1 qui s'effectuent au cours de l'année N.

Les améliorations introduites dans la série d'enquêtes Revenus fiscaux et sociaux (ERFS)

L'enquête Revenus fiscaux est remplacée par une nouvelle série d'enquêtes Revenus fiscaux et sociaux (ERFS) mise en place à partir des revenus déclarés en 2005³. Les améliorations introduites concernent l'intégration des prestations sociales réellement perçues par les bénéficiaires, alors qu'auparavant ces montants étaient estimés à partir des barèmes applicables, et également une prise en compte plus complète des revenus des produits financiers.

Une prise en compte des transferts sociaux reçus par les ménages

Pour constituer le revenu disponible, le revenu que les ménages déclarent à l'administration fiscale est complété par les revenus sociaux non imposables : prestations familiales, minima sociaux et prestations logement. Jusqu'en 2005 avec l'enquête ERF, ceux-ci étaient estimés à l'aide des barèmes en vigueur, sur la base des revenus déclarés de l'année et étaient attribués pour l'année entière. Dans la nouvelle série des enquêtes Revenus fiscaux et sociaux (ERFS), les montants réellement perçus par les ménages sur l'année de revenu considérée sont collectés auprès de la caisse nationale des allocations familiales (CNAF), de la caisse nationale assurance vieillesse (CNAV) et de la caisse centrale de la mutualité sociale agricole (CCMSA).

Le recours aux fichiers des organismes gestionnaires a également permis d'intégrer, dans la nouvelle série d'enquêtes, des ménages non retrouvés dans le fichier fiscal et qui étaient auparavant exclus du champ de l'enquête (considérés comme non-répondants) alors qu'ils sont globalement plus pauvres que l'ensemble des ménages (environ 1,2 % de l'échantillon).

Une meilleure couverture des revenus du patrimoine

L'ancienne série de l'enquête Revenus fiscaux qui s'appuyait exclusivement sur la source fiscale appréhendait imparfaitement les revenus du patrimoine, principalement en raison des exonérations fiscales ou des mécanismes de **prélèvement libératoire jusqu'en 2012 et forfaitaire depuis 2013**. Cela avait pour effet de minimiser la mesure des inégalités, puisque les revenus du patrimoine sont en général d'autant plus élevés que les revenus d'activité sont importants.

Dans la nouvelle série ERFS, depuis 2005, les revenus générés par différents produits financiers non recensés par la source fiscale sont estimés à partir de modèles statistiques et ajoutés au revenu des ménages. Plus précisément, tous les ménages qui déclarent des revenus d'assurance-vie au fisc sont identifiés comme détenteurs. Pour les autres ménages, la détention est imputée ou non d'après les résultats de modèles statistiques. Pour cela, un montant des différents actifs financiers détenus par les ménages est imputé à partir de l'enquête Patrimoine réalisée par l'Insee⁴. Le patrimoine total ainsi imputé reste cependant sous-estimé. Les encours des différents actifs (hors assurance-vie) sont alors recalés, produit par produit, sur les données macroéconomiques des comptes nationaux et de la Banque de France (BDF). Les revenus générés par les actifs (hors assurance-vie) sont ensuite calculés par application de taux de rendement moyens réactualisés chaque année. Pour l'assurance-vie, le recalage se fait sur les revenus et non sur les encours. Faute d'information, le recalage des produits financiers est cependant effectué de manière uniforme : il est indépendant du revenu des ménages.

Si la prise en compte des revenus du patrimoine permet de mieux appréhender la dispersion réelle des revenus, on peut penser que l'application d'un taux de recalage uniforme n'en rend pas complètement compte : la distribution des revenus estimée serait encore moins concentrée qu'elle n'est en réalité. Par ailleurs, les revenus soumis au prélèvement libératoire ou forfaitaire, hormis les contrats d'assurance-vie - obligations, bons, OPCVM (Organismes de placements collectifs en valeurs mobilières) - restent mal appréhendés. Au total, la série ERFS permet une meilleure estimation des indicateurs d'inégalité qui se rapproche de la situation réelle tout en restant probablement encore un peu en deçà de celle-ci.

Dans les ERFS de 2005 à 2010, l'enquête Patrimoine 2004 avait été utilisée comme base pour l'imputation des revenus financiers. Dans les ERFS 2011 à 2014, la méthode d'imputation des revenus financiers

3 Pour les revenus de l'année 2005, l'enquête existe en deux versions : ERF et ERFS.

4 La méthode d'élaboration de l'**enquête Patrimoine** est accessible dans cette rubrique « Sources et méthodes », à l'emplacement suivant : Source : enquête Patrimoine.

absents de la déclaration fiscale a été revue et s'appuie sur l'enquête Patrimoine 2010⁵. Même si la méthode mise en œuvre reste globalement inchangée, il était nécessaire de prendre en compte le fait que l'enquête Patrimoine 2010 appréhende mieux que celle de 2004 le haut de la distribution des patrimoines. Aussi son intégration comme base des imputations à partir de l'ERFS 2011 avait été l'occasion d'améliorer les descripteurs pris en compte à la fois dans les modèles d'imputation de la détention et dans ceux de simulation des encours détenus pour accroître la précision des modèles sur le haut de la distribution. En outre, les masses de calage des encours avaient été revues afin d'assurer une homogénéité de champ avec celui couvert par l'ERFS (ménages dits « ordinaires » vivant en France métropolitaine).

A partir de l'ERFS 2015, la méthode d'imputation a été revue pour intégrer les améliorations de l'enquête Patrimoine 2014-2015. Les montants d'actifs déclarés par les ménages, désormais recueillis directement et non plus sous forme d'intervalles, sont plus fiables. Les revenus financiers des ERFS 2012 à 2014 ont été recalculés avec la nouvelle méthode, afin de pouvoir apprécier les évolutions depuis 2012 à méthode constante.

Une meilleure cohérence avec les concepts de la comptabilité nationale.

À partir de l'ERFS 2013, outre le changement de millésime d'impôts décrit page 3, d'autres améliorations importantes ont été introduites. La première a été la prise en compte des évolutions introduites dans l'enquête Emploi en continu de 2013 (qui constitue le socle de l'ERFS), visant notamment à mieux cerner certaines populations difficiles à appréhender, comme les étudiants.

Deux changements fiscaux importants ont été introduits à compter de 2013. D'une part, les majorations de pension de retraite pour avoir élevé 3 enfants ou plus sont désormais prises en compte dans le revenu déclaré à l'administration fiscale. D'autre part, l'avantage en nature constitué par la participation des employeurs aux contrats collectifs de complémentaire santé est maintenant intégré dans les salaires déclarés à l'administration. Ces ressources sont désormais comptabilisées dans le revenu disponible calculé dans l'ERFS.

En prenant en compte l'impôt effectivement payé en 2013, les majorations de pensions et la participation employeur aux contrats de complémentaire santé, la notion de revenu dans l'ERFS s'est rapprochée de celle de la comptabilité nationale ainsi que de celle de l'enquête européenne *Statistics on income and living conditions (SILC)*.

Autres changements importants intervenus dans les ERFS entre 2014 et 2020 :

→ En 2016, intégration des agrégats individu et prise en compte de la nouvelle enquête logement 2013

Le modèle d'imputation des loyers des locataires et des propriétaires dans ERFS a été modifié afin de prendre en compte la nouvelle enquête logement 2013. De plus, l'enquête Logement 2013 a été enrichie à partir des sources fiscales et sociales. Le recours à des données de revenus issues d'un appariement avec les données administratives plutôt qu'à des données collectées, comme lors des enquêtes précédentes, fournit une variable de revenu plus précise, mais cela induit également une rupture avec les revenus collectés dans les enquêtes logement précédentes (impact sur les aides au logement).

→ En 2017, nouveautés dans l'appariement et l'imputation

Une nouvelle méthode de scoring a été intégrée dans ERFS lors de l'appariement automatique avec le fichier FIP. Les méthodes d'imputation pour la taxe d'habitation, le minimum vieillesse et les minima sociaux ainsi que pour les revenus agricoles ont été modifiés. Deux nouvelles imputations ont également été mises en place suite à l'introduction dans l'enquête emploi de la prime d'activité et de la garantie jeunes.

→ En 2018, évolution de la fiscalité, nouvelle enquête BMS et nouveau fichier annuel exhaustif de la CNAF

La nouvelle enquête sur les Bénéficiaires de Minima Sociaux (BMS) de la DREES de 2018 a été utilisée pour estimer la part d'allocataires ne vivant pas en ménage ordinaire. Un nouveau fichier annuel exhaustif de la Caisse Nationale des Allocations Familiales a servi à déterminer les allocataires de prestations Cnaf et les montants perçus.

Les programmes de la chaîne ERFS se sont adaptés à la mise en place du prélèvement forfaitaire unique (PFU) et à l'évolution de la fiscalité sur les revenus du patrimoine associée. Ils ont également intégré une variable recalculée par la DGFIP, pour le montant de calcul de l'impôt sur les revenus de 2018.

5 Une nouvelle version de l'ERFS 2010 avait également été produite avec l'enquête Patrimoine 2010, afin de pouvoir apprécier les évolutions entre 2010 et 2011 à méthode constante.

→ *En 2019, nouvelle méthode d'enrichissement et imputations de revenus supplémentaires*

Les enrichissements fiscaux et sociaux de l'ERFS s'appuient désormais sur la méthode Rapsodie (RAPprochement des données SOciales, Des Impôts et des Enquêtes), qui rend ces processus plus génériques.

En 2019, des imputations statistiques ont été réalisées pour introduire la prime exceptionnelle de pouvoir d'achat et les heures supplémentaires exonérées, afin de corriger un défaut de couverture des salaires dans la source fiscale.

→ *En 2020, collecte perturbée par les restrictions liées à la crise sanitaire et imputation des dispositifs exceptionnels mis en place durant la crise sanitaire*

Avertissement : du fait de difficultés de production 2020, les données sont fragiles.

Plusieurs dispositifs exceptionnels mis en place durant la crise sanitaire, non mesurés par les sources fiscales et sociales servant à constituer l'ERFS, ont été imputés. Pour la prime exceptionnelle de pouvoir d'achat et les heures supplémentaires exonérées, l'imputation mise en place en 2019 a été reconduite. Des nouvelles imputations ont été introduites pour le Fonds de solidarité des entreprises, les primes versées dans le secteur de la santé, et l'aide exceptionnelle de solidarité versée aux ménages modestes.

Une amélioration des méthodes de correction de la non-réponse et une actualisation des concepts et marges pour les rapprocher de ceux du recensement de la population.

→ *En 2021, l'ERFS a connu une refonte importante pour s'adapter à la nouvelle EEC. L'ERFS a ainsi repris à son compte les évolutions des règles de rattachement (notamment concernant les étudiants) et de définition du concept de personne de référence introduites dans l'EEC.* Cette refonte a également été l'occasion d'intégrer l'amélioration des méthodes de correction de la non-réponse de l'EEC et de conserver les individus de l'EEC non appariés avec les sources administratives dans le champ de l'ERFS (excepté pour les individus de 90 et plus pour lesquels les imputations sont trop fragiles). Enfin, depuis 2021, les marges de population intègrent les évolutions du questionnaire de l'enquête annuelle de recensement qui visent à mieux mesurer les situations de multi-résidence. Les niveaux de vie sont rehaussés et les indicateurs de pauvreté et d'inégalités diminuent (excepté le rapport interdéciles qui demeure stable). Certaines ruptures de séries peuvent être plus importantes par sous-catégories, selon les types de ménage ou le statut d'activité. Une évaluation de ces ruptures de mesure est donc indispensable pour être en mesure de proposer des évolutions historiques cohérentes. Pour plus de précision, se référer à l'[Insee Méthode n°145](#).

Par ailleurs, en 2021, les imputations de la prime exceptionnelle de pouvoir d'achat, du fonds de solidarité et des primes versées dans le secteur de la santé ont été reconduites. L'indemnité inflation a également été imputée pour les personnes éligibles fin 2021 (un tiers des bénéficiaires, les deux tiers restants ayant perçu la prime en 2022). L'imputation des heures supplémentaires a, à l'inverse, été très limitée en 2021, la couverture dans les données fiscales s'étant largement améliorée.

→ *En 2022, l'ERFS a été produite dans la continuité de celle de 2021.* En particulier, les dispositifs exonérés encore existants ont également été imputés : prime de partage de la valeur qui a remplacé la prime exceptionnelle de pouvoir d'achat courant 2022, indemnité inflation et heures supplémentaires non déclarées à l'administration fiscale (très minoritaires). L'aide exceptionnelle de solidarité créée cette année-là pour soutenir le pouvoir d'achat des ménages a également été imputée en 2022.

→ *En 2023, l'ERFS n'a pas connu d'évolution particulière.* Toutefois, aucun dispositif exonéré n'a été imputé du fait de leur qualité dans les sources fiscales (cas des heures supplémentaires et de la prime de partage de la valeur) ou de leur absence (non reconduction en 2023 des dispositifs imputés en 2022 : indemnité inflation ou aide exceptionnelle de solidarité). La taxe d'habitation, du fait de sa complète exonération pour les résidences principales, n'existe désormais plus dans l'ERFS.

Méthodologie et déroulement de l'enquête

1. L'échantillon de départ

L'EEC, socle de l'ERFS, est une enquête par sondage, c'est-à-dire que l'on interroge un échantillon représentatif de l'ensemble de la population. Elle se déroule tout au long de l'année (d'où son appellation « en continu »).

L'échantillon de l'enquête est un échantillon de logements, et non d'individus. Il présente deux caractéristiques principales : il est aréolaire (constitué d'un tirage de groupes de logements proches) et rotatif (les logements entrants dans l'échantillon un trimestre donné sont enquêtés pendant six trimestres consécutifs).

En France métropolitaine, la base de sondage de l'EEC est constituée du fichier démographique des logements et des individus (Fideli). Le tirage de l'échantillon est stratifié par région. Au sein de chaque région, le tirage des secteurs est fait de manière "équilibrée", proportionnellement au nombre de résidences principales, selon des caractéristiques de la base de sondage.

Les logements tirés au 4^{ème} trimestre de l'année N constituent le champ de l'ERFS. L'échantillon de départ de l'ERFS 2023 est constitué, par exemple, de l'EEC du 4^{ème} trimestre de l'année 2023 soit 92 500 individus pour environ 44 000 ménages résidant en France métropolitaine. Sont inclus dans cet échantillon les individus appartenant à un ménage de l'EEC, même s'ils n'ont pas répondu au questionnaire, ainsi que les individus rattachés qui considèrent ce logement comme leur résidence principale. Chaque individu interrogé dans l'EEC représente *in fine* environ 700 personnes compte tenu de la taille de l'échantillon.

2. L'appariement des données avec les fichiers fiscaux

L'appariement a pour objectif de retrouver les déclarations fiscales remplies par les personnes ayant répondu à l'EEC. La direction générale des Finances publiques (DGFiP) livre trois fichiers à l'Insee :

1. Le premier est le fichier d'imposition des personnes ; c'est la liste des contribuables avec leur adresse au 1^{er} janvier N+1.
2. Le deuxième fichier contient l'ensemble des éléments de taxation du foyer fiscal à l'impôt sur le revenu (IR) de l'année N ; il s'agit de l'ensemble des éléments des déclarations fiscales n° 2042 et n° 2042-C (complémentaire) remplies en mai-juin N+1 au titre des revenus de l'année N.
3. Enfin, le dernier fichier contient les données relatives à la taxe d'habitation (TH) au 1er janvier N (et de l'année N+1 si le ménage n'est pas retrouvé dans le fichier de l'année N). Suite à l'exonération totale de taxe d'habitation pour les résidences principales à compter de 2023, ce fichier n'est plus utilisé dans la production de l'ERFS.

Ces données issues des fichiers fiscaux sont « fusionnées » (après appariement statistique) avec celles des individus de l'échantillon de l'EEC collectées au cours du 4^{ème} trimestre de l'année N.

Depuis 2021, l'appariement avec le fichier d'imposition des personnes est effectué grâce à des données identifiantes telles que le nom de famille, le prénom, le sexe, le jour, le mois et l'année de naissance du déclarant, le département de naissance, le numéro de la voie et deux « mots directeurs » extraits du libellé de la voie (les deux plus grands mots du libellé de la voie figurant dans les adresses⁶). Auparavant, le nom de famille n'était pas disponible dans l'EEC et l'appariement avec ces données était dit « statistique ».

Ensuite, pour les individus retrouvés, les données portant sur les revenus sont extraites du fichier contenant l'ensemble des éléments de taxation du foyer fiscal. Mais pour certains individus appariés les déclarations

⁶ Ces mots directeurs sont tirés des deux adresses que les fichiers contiennent : l'adresse fiscale de taxation au 1^{er} janvier N+1 et la dernière adresse connue si celle-ci est différente (cas des déménagements).

ne se retrouvent pas systématiquement. C'est le cas, par exemple, des étudiants redevables d'une taxe d'habitation, mais rattachés à la déclaration de revenus de leurs parents.

Des individus du ménage peuvent ne pas être retrouvés dans les déclarations fiscales. Pour ces ménages, les déclarations manquantes sont reconstituées. C'est le cas lorsqu'il y a eu mariage, séparation ou décès. Dans ces situations, plusieurs déclarations peuvent être rédigées pour une même année.

Les revenus de personnes qui apparaissent dans les déclarations fiscales (conjoints ou personnes à charge) mais qui ne font pas partie du ménage de l'EEC ne sont pas pris en compte. Il s'agit le plus souvent d'enfants à charge vivant hors de la résidence familiale pour lesquels aucune condition de vie sous le toit du déclarant fiscal n'est exigée (enfants de moins de 25 ans poursuivant des études, enfants de moins de 21 ans, enfants infirmes et résidant en instituts spécialisés) et d'enfants en résidence alternée absents au moment de l'enquête.

3. L'appariement avec les fichiers sociaux

Les individus des ménages de l'EEC sont également appariés avec les données des organismes sociaux : caisse nationale des allocations familiales (CNAF), caisse nationale assurance vieillesse (CNAV) et caisse centrale de la mutualité sociale agricole (CCMSA).

Les caisses d'allocations familiales gèrent le versement de la très grande majorité des prestations prises en compte dans l'ERFS. La bonne qualité de l'appariement avec les données de la CNAF est un élément essentiel dans l'ERFS.

Le déroulement de l'appariement avec le fichier CNAF est le suivant. Les données identifiantes des individus issues de l'EEC (nom, prénom, date de naissance et adresse fiscale de taxation, département et commune de résidence) sont utilisées pour retrouver ces individus dans le fichier des bénéficiaires transmis par la CNAF. Depuis 2018, l'appariement est réalisé avec un nouveau fichier de la CNAF recensant l'ensemble des bénéficiaires et des prestations versées au cours de l'année. L'Insee est ainsi en mesure de calculer les prestations versées dès la fin de l'appariement. La méthode d'appariement a également été revue à cette occasion, et s'appuie désormais notamment sur des distances calculées entre les individus. Depuis 2021, le nom utilisé pour retrouver les individus provient directement de l'EEC (il était auparavant récupéré dans les données fiscales).

Certains ménages non retrouvés dans les sources fiscales sont retrouvés dans le fichier de la CNAF. Dans ce cas, certaines informations concernant leurs revenus peuvent être retrouvées dans le fichier de la CNAF. Ces informations sont par la suite intégrées dans les traitements de non-réponse décrits ci-dessous.

La caisse centrale de la mutualité sociale agricole (CCMSA) est l'organisme d'assurance sociale obligatoire des salariés et exploitants du secteur agricole. Elle verse ainsi les prestations familiales, au même titre que la CAF, mais aussi l'ASPA - allocation de solidarité pour les personnes âgées – ou le minimum vieillesse⁷ pour la population agricole. Le rapprochement avec le fichier des allocataires de la MSA est réalisé à partir du fichier des allocataires de l'année N fourni à l'Insee. Les individus qui font l'objet de l'appariement sont les individus de l'EEC âgés de 18 ans ou plus⁸ qui n'ont pas été appariés (en tant que bénéficiaire ou conjoint) avec le fichier CNAF dans le cadre des prestations familiales. Les critères identifiants utilisés sont le nom, le prénom, la date et le lieu de naissance ainsi que l'adresse.

La caisse nationale d'assurance vieillesse (CNAV) gère l'ASPA et le minimum vieillesse des individus relevant de la branche générale. Depuis l'ERFS 2018, l'Insee réalise directement l'appariement avec un fichier des allocataires de l'année N transmis par la CNAV. Les critères identifiants utilisés sont le nom, le prénom, la date et le lieu de naissance ainsi que l'adresse.

⁷ Les données sur le minimum vieillesse ou l'ASPA versées par la MSA ne sont récupérées qu'à partir de l'ERFS portant sur les revenus 2007.

⁸ Âge minimum des allocataires observé dans le fichier de la MSA.

4. Le traitement de la non-réponse et le redressement

Comme pour toutes les enquêtes réalisées auprès des ménages, les résultats bruts de l'ERFS sont traités pour répondre à deux objectifs principaux :

1. traiter la non-réponse pour compenser le biais introduit par les non-réponses totales et partielles ;
2. effectuer un redressement afin de réduire autant que possible les fluctuations d'échantillonnage.

4.1 Traitement de la non-réponse

4.1.1 Non-réponse totale ou partielle ?

Dans l'ERFS, les revenus d'un individu sont tirés des déclarations fiscales qui ont été associées au ménage de l'EEC dans lequel il se trouve. Plusieurs cas peuvent être distingués.

Dès lors que pour un ménage de l'EEC une déclaration fiscale est retrouvée (c'est le cas de plus de 97 % des ménages de l'enquête Emploi au 4^{ème} trimestre 2023), le ménage est considéré comme « **répondant** » à l'ERFS. Ses revenus sont tirés des déclarations fiscales et des données sociales s'il perçoit des prestations.

En revanche, si un dossier fiscal est retrouvé, mais qu'il ne concerne pas toutes les personnes du ménage, alors ce ménage constitue une « **non-réponse partielle** » : seules les données d'une partie des personnes de ce ménage sont disponibles. De même si un dossier social (données CNAF) est retrouvé sans qu'aucun dossier fiscal ne soit retrouvé, le ménage considéré est aussi en non-réponse partielle. En cas de mariage, de séparation ou de décès dans l'année, plusieurs déclarations peuvent être rédigées pour une même année. Si certaines d'entre elles ne sont pas retrouvées, le ménage figure parmi les non-réponses partielles : le revenu de ces individus est de fait incomplet.

Dans le cas de non-réponses partielles, les données des individus ou des dossiers manquants sont imputées afin de ne pas perdre une partie de l'information. Ces imputations concernent les revenus individualisables. Les revenus non individualisables ne sont pas imputés comme ils sont portés au niveau des déclarations fiscales associées au ménage, pour lesquels des individus ont pu être retrouvés.

Lorsque l'ensemble des individus du ménage ne sont pas retrouvés dans les fichiers fiscaux et sociaux (environ 2 % de l'échantillon de l'enquête Emploi en 2023), ces individus sont considérés comme « **non-répondant total** » : en effet, on ne dispose ni de leurs revenus, ni des prestations reçues. Depuis 2021, ces « non-répondants total » sont conservés dans l'ERFS afin de garder un maximum d'informations⁹. Leurs revenus individualisables et non individualisables sont imputés par modèle économétrique grâce aux données recueillies dans l'EEC (en particulier sur la situation d'activité, la PCS, le diplôme, etc.). Font exception les individus de 90 ans ou plus, qui ne sont plus enquêtés au niveau individuel par la nouvelle enquête Emploi. L'absence d'informations sur ces personnes rend l'imputation de leurs revenus trop fragile. Les individus de 90 ans ou plus demeurent donc exclus du champ de l'ERFS lorsqu'ils ne sont retrouvés, ni dans les fichiers fiscaux, ni dans les fichiers sociaux. Le biais introduit par cette exclusion est corrigé par repondération des individus de plus de 90 ans appariés.

4.1.2 Le traitement de la non-réponse partielle : imputation des données manquantes.

Les non-réponses partielles concernent des individus qui sont dans le champ de l'ERFS mais pour lesquels certaines informations manquent alors que d'autres sont disponibles. Il convient de conserver ces individus en opérant des imputations adaptées à chaque type d'information manquante (revenus, cas des mariages, divorces ou décès, etc.). L'objectif de ces imputations est de remplacer la donnée absente par une donnée « plausible ».

⁹ Les individus retrouvés uniquement dans les fichiers sociaux ou uniquement dans les fichiers fiscaux étaient déjà conservés dans le champ de l'enquête.

Revenus manquants pour les individus ayant un dossier CNAF

Pour les individus ayant un dossier CNAF, la méthode d'imputation des revenus manquants diffère selon le type de prestation reçue par les bénéficiaires car les informations disponibles en matière de revenus dans les fichiers de la CNAF varient selon les cas.

Des imputations de revenus sont réalisées pour les individus (allocataires) non appariés avec les fichiers fiscaux mais dont on connaît soit les revenus des 4 trimestres de l'année N (car ils ont perçu le revenu de solidarité active (RSA) ou la prime d'activité sur l'année entière), soit les revenus déclarés de l'année N-1 (ayant servi au calcul d'autres droits pour l'année suivante). En effet, les données fournies par la CNAF contiennent les informations sur les ressources trimestrielles déclarées par les bénéficiaires du RSA ou de la prime d'activité : revenus d'activité, revenus de stages de formation, allocations chômage, etc. Lorsque ces informations sont disponibles pour les 4 trimestres de l'année, les revenus trimestriels sont cumulés afin d'obtenir le revenu annuel. Pour les autres individus du ménage qui ne sont pas appariés avec les fichiers fiscaux, des imputations de revenus sont réalisées selon des modèles économétriques (les individus non retrouvés mais appartenant à un ménage retrouvé dans les fichiers fiscaux servent de base pour ces imputations).

De même, des imputations sont réalisées dès lors que les fichiers CNAF contiennent les revenus déclarés par les apporteurs de ressources (salaires, indemnités chômage, pensions, etc.) pris en compte pour le calcul des droits aux prestations familiales. Ces revenus sont ceux des 4 trimestres de l'année N, si l'allocataire a perçu le RSA ou la prime d'activité sur l'année entière. Sinon ce sont les revenus de l'année N-2 ayant servi au calcul des autres droits pour l'année N. Quand les ressources des bénéficiaires du RSA sont disponibles pour les 4 trimestres de l'année, on les cumule pour obtenir un revenu annuel. Pour les autres allocataires, la correction appliquée à leurs revenus correspond à l'évolution moyenne des revenus constatée entre N-2 et N pour les individus allocataires dont les ressources sont connues en N et N-2 (source DGFiP pour l'année N, source CNAF pour l'année N-2). La répartition des revenus des indépendants (revenus d'agriculteurs, commerciaux, non commerciaux) est établie à partir de la catégorie socioprofessionnelle des individus.

Individus dont la situation familiale a évolué au cours de l'année

Depuis les déclarations de revenus 2011, les individus amenés à remplir plusieurs déclarations sont uniquement ceux dont le conjoint est décédé au cours de l'année. Dans le cas d'un mariage (ou PACS) ou d'un divorce, les individus concernés remplissent désormais soit une seule déclaration commune, soit deux déclarations séparées pour l'année complète.

Les personnes concernées par le décès de leur conjoint remplissent une déclaration fiscale pour chaque période que l'événement délimite. La période de vie commune se traduit par le dépôt d'une seule déclaration pour le couple et la période de vie seul(e) fait l'objet d'une déclaration séparée. À l'issue de l'appariement des individus de l'EEC avec les données fiscales, la récupération des deux déclarations n'est pas toujours assurée. Il faut alors compléter le revenu de ces individus pour estimer leur revenu annuel. Cela couvre les catégories suivantes : salaires, chômage, revenus agricoles, revenus industriels et commerciaux, revenus non commerciaux. La méthode retenue consiste à appliquer à la période manquante le rendement observé (revenus déclarés/nombre de jours) sur la période connue. Cette méthode fait l'hypothèse que ces individus ont une situation au regard de l'activité assez stable sur l'année de revenu étudiée. Pour limiter les risques de surestimation du revenu annuel de ces individus, ces valeurs sont majorées par le revenu maximum observé chez les individus de même catégorie socioprofessionnelle qui ont un revenu annuel complet. Cette méthode n'est appliquée que pour certaines personnes : celles qui disposent d'une période observée de revenus supérieure à 31 jours dans la déclaration retrouvée (la période est suffisamment longue pour que la méthode conduise à une imputation fiable).

Imputation des revenus d'activité ou de remplacement

Pour certaines catégories de personnes, il convient, faute d'information retrouvée, de procéder à des imputations de leurs revenus d'activité (salaires et revenus d'indépendants) ou de remplacement (allocations chômage, retraites). Préalablement aux imputations de ces revenus individuels manquants, revenus bien évidemment liés au statut d'activité des personnes, il faut reconstituer précisément l'activité principale des personnes concernées sur l'année entière (individus de l'enquête emploi du 4^{ème} trimestre de l'année N). Ainsi, pour chaque trimestre et donc sur l'ensemble de l'année une activité principalement exercée est reconstituée en mobilisant notamment les concepts d'activité de l'EEC : activité au sens du BIT

pour la semaine de référence de l'enquête et calendrier rétrospectif (activité déclarée spontanément sur les 11 derniers mois).

La population des personnes à imputer est constituée des individus d'une tranche d'âge donnée et ayant un statut spécifique au regard de l'activité professionnelle, et selon les cas :

- qui ne sont pas retrouvés dans les fichiers fiscaux et ne sont pas non plus allocataires à la CNAF ; ou bien
- qui ont connu le décès de leur conjoint l'année N et pour lesquels la période observée de revenus dans la déclaration fiscale retrouvée est inférieure ou égale à 30 jours.

Plus précisément, pour les salaires, les personnes dont le revenu est à imputer sont celles âgées de 18 à 65 ans¹⁰, ayant pendant au moins un trimestre une activité principale de salariés. Pour le chômage, la tranche d'âge est la même et les personnes ont au moins un trimestre en situation principale d'inactivité ou de chômage, à condition qu'elles aient déclaré percevoir une allocation chômage dans l'enquête Emploi. Concernant les retraites, la population à imputer est constituée des individus de plus de 55 ans, ayant au moins un trimestre d'inactivité.

Pour pouvoir imputer les revenus manquants de ces personnes grâce à un modèle économétrique, la population de référence (ou population des répondants) est celle dont l'information en matière de revenus est utilisée pour imputer les données manquantes. Elle est constituée des individus ayant les mêmes caractéristiques en termes d'âge et de situation professionnelle, individus quant à eux retrouvés dans les fichiers fiscaux dans lesquels ils ont déclaré, selon le cas : des salaires, des montants d'allocations de chômage ou de préretraite, des pensions de retraite ou d'invalidité. Un certain nombre de variables auxiliaires sont prises en compte dans les modèles économétriques : sexe, diplôme, âge, catégorie socioprofessionnelle, etc.

Des méthodes d'imputation similaires sont mises en œuvre pour imputer les revenus des indépendants, en traitant séparément les revenus agricoles, les revenus industriels et commerciaux et les revenus non commerciaux. Au sein des revenus agricoles, les exploitants agricoles qui sont au forfait¹¹ sont traités séparément : leur particularité est qu'au moment de l'appariement avec les fichiers fiscaux, le montant du forfait n'est pas fixé par l'administration fiscale. Les personnes se déclarant « aide familial¹² » ne sont pas intégrées à ces traitements.

Imputation des revenus des ménages d'étudiants

Les ménages d'étudiants de moins de 26 ans (ménages dont la personne de référence est étudiante) font partie de l'échantillon ERFS même s'ils ne sont pas appariés avec les fichiers fiscaux et sociaux. L'hypothèse est que ces étudiants sont rattachés fiscalement à la déclaration de revenus de leurs parents.

On affecte par imputation à ces ménages un revenu observé dans la déclaration fiscale d'un individu comparable, en procédant comme suit :

- Parmi les appariements réussis, on identifie des ménages au sein desquels une personne est présente dans la (ou les) déclaration(s) fiscale(s), comme personne à charge, mais absente du ménage de l'enquête Emploi. On suppose que les personnes ainsi identifiées sont des étudiants qui vivent dans un autre ménage (souvent seuls).
- En prenant soin de considérer des individus comparables en termes de catégorie socioprofessionnelle et d'âge de la personne de référence (informations présentes dans l'enquête Emploi), on attribue alors aux étudiants en échec d'appariement les revenus des personnes identifiées dans l'étape précédente. Leurs revenus sont donc probablement sous-estimés dans la mesure où l'on ne prend pas en compte les transferts qu'ils peuvent recevoir de leurs parents.

10 On ne peut imputer de revenus aux personnes en emploi âgées de moins de 18 ans car elles sont peu nombreuses et leurs revenus atypiques : par défaut, on considère qu'elles n'ont aucun revenu. Pour la même raison, on ne tient pas compte des très jeunes travailleurs dans le modèle d'imputation des revenus.

11 Le régime du forfait est appliqué dès lors que les recettes TTC de l'exploitation agricole ne dépassent pas un seuil fixé par l'administration fiscale.

12 Les aides familiaux participent aux travaux d'une entreprise familiale (exploitation agricole, commerce, artisan, etc.) sans être ni "indépendants ou employeurs" ni "salariés". Il s'agit essentiellement d'épouses et d'enfants des chefs d'entreprise individuelle travaillant pour leur propre compte.

4.1.3 Les corrections de la non-réponse totale

Dans le cas où, pour un ménage de l'enquête Emploi, il n'est retrouvé ni dossier fiscal ni dossier social, il s'agit d'un échec d'appariement qui est donc assimilable à de la non-réponse totale. Les résultats de l'enquête ERFS seraient biaisés si l'on ne corrigeait pas la non-réponse, sauf dans l'hypothèse où les non-répondants ont un comportement identique à celui des répondants. Mais, les échecs d'appariement avec les données fiscales et sociales ne se répartissent a priori pas au hasard parmi les ménages de l'EEC. Aussi, afin de compenser le biais lié à la non-réponse totale, on effectue les traitements suivants.

Jusqu'en 2020, la correction de la non-réponse totale consistait à mettre en œuvre un modèle de régression logistique : il s'agit dans un premier temps d'analyser la non-réponse (hors ménages à personne de référence étudiante de moins de 26 ans au 31 décembre de l'année N) pour mieux la corriger. On postulait un modèle de comportement des individus face à la non-réponse en modélisant la probabilité d'appartenir à l'échantillon apparié (i.e. d'être répondant) plutôt qu'à l'échantillon non-apparié (i.e. d'être non-répondant) en fonction de caractéristiques des individus et de leur ménage (par exemple : l'âge, la localisation géographique, la catégorie socioprofessionnelle, etc.). Corriger de la non-réponse consistait à effectuer une modification des poids initiaux¹³ des ménages pour prendre en compte le fait que tous ne répondent pas. La correction était d'autant meilleure que les poids initiaux étaient transformés de telle sorte que les ménages non-répondants sont représentés par les ménages possédant autant que possible les mêmes caractéristiques : on divisait donc la pondération initiale des ménages répondants par la probabilité estimée d'appartenir à l'échantillon apparié.

Depuis 2021, les « non-répondants total » sont conservés dans le champ de l'ERFS afin de conserver un maximum d'information. Leurs revenus d'activité ou de remplacement sont imputés à l'aide de modèles économétriques décrits dans la partie sur la correction de la non-réponse partielle. Ces imputations de données fiscales sont traitées au niveau individuel si le revenu correspondant est individualisé dans la déclaration fiscale. Ceci concerne les salaires, les indemnités de chômage, les pensions de retraite, ainsi que les revenus d'indépendants. Les revenus non individualisables sont imputés au niveau du ménage : il s'agit des revenus fonciers et des revenus de capitaux mobiliers. L'imputation de ces revenus est réalisée en trois temps à partir des ménages enrichis fiscalement :

- dans un premier temps, la perception est imputée à partir d'un modèle économétrique basé sur l'âge, l'activité et la catégorie socio-professionnelle de la personne de référence du ménage, le type de famille et le statut d'occupation ;
- dans un second temps, les montants perçus sont imputés globalement (revenus fonciers et revenus de capitaux mobiliers) à partir d'un modèle économétrique basé sur l'âge et la catégorie socio-professionnelle de la personne de référence du ménage, le type de famille, le statut d'occupation et le fait de résider dans l'unité urbaine de Paris ;
- enfin, les montants imputés sont répartis entre revenus fonciers et revenus de capitaux mobiliers selon la répartition observée sur les ménages enrichis fiscalement.

Seuls les individus de 90 ans et plus non appariés sont exclus du champ de l'enquête : le biais introduit par cette exclusion est corrigé par repondération des individus de plus de 90 ans appariés.

4.2 Le redressement pour fluctuations d'échantillonnage

Lorsque l'on tire au sort un échantillon pour une enquête, il arrive fréquemment que l'échantillon obtenu ne soit pas exactement représentatif de la population de référence, au regard de certains critères importants pour le sujet de l'enquête. Ces fluctuations d'échantillonnage peuvent être redressées par une modification des poids des individus de l'échantillon. Corriger les fluctuations d'échantillonnage suppose d'avoir des informations issues d'une autre source que l'enquête et de meilleure qualité ou mieux encore la « vraie » valeur. Cette repondération est réalisée à l'aide d'une méthode usuelle de « calage sur marges ». Cette méthode consiste à modifier les poids initiaux¹⁴ des ménages de telle sorte que le total de certaines

13 Ces poids initiaux sont ceux qu'ont les ménages dans l'enquête Emploi.

14 À partir de l'enquête emploi de 2007, cette pondération initiale a été calée sur les nouvelles enquêtes de recensement. Une rétropolation, consistant en une mise à jour de ces pondérations, a également été opérée sur les enquêtes emploi à partir du millésime 2002. Compte tenu de la rupture de série pour l'ERFS en 2005, seules les ERFS 2005 et 2006 ont été modifiées en conséquence (mise à jour des pondérations initiales).

variables, en l'occurrence celles citées ci-dessous, estimé à partir des répondants de l'échantillon, soit égal au « vrai » total (ou « marges ») connu par ailleurs (par les **enquêtes annuelles de recensement**, l'**enquête Logement** et par les **fichiers fiscaux**).

L'enquête ERFS est calée sur des données démographiques (pyramide des âges et nombre de ménages) et sur les masses fiscales déclarées. Concernant la pyramide des âges, les marges sont les effectifs issus de la structure par sexe et âge quinquennal de la population des ménages dits ordinaires de France métropolitaine au 15 novembre de l'année N¹⁵. Cette population est calculée à l'issue de la dernière collecte du recensement de la population. Depuis 2021, les marges utilisées tiennent compte de l'évolution du questionnaire de l'enquête annuelle de recensement visant à mieux mesurer les situations de multi-résidence, qui a eu pour effet de diminuer le nombre d'enfants.

Quant au nombre de ménages au 15 novembre N, il est estimé à partir du nombre connu de résidences principales au 1^{er} janvier N et au 1^{er} janvier N+1¹⁶. On fait l'hypothèse d'une évolution linéaire du nombre de résidences principales dans l'année. On estime donc un nombre de ménages par interpolation linéaire avec la formule suivante pour l'année N :

$$\text{Nb de ménages au 15/11/N} = [\text{Nb au 01/01/N}] + [(10,5 / 12) \times (\text{Nb au 01/01/N+1} - \text{Nb au 01/01/N})]$$

Remarque : les enquêtes ERFS ont été rétropolées, intégrant les mises à jour des données démographiques (ERFS 2005 à 2007) et les corrections de la catégorie socioprofessionnelle (CS)¹⁷ dans l'enquête Emploi (ERFS 2003 à 2007).

Le calage sur les masses de revenus issues des fichiers fournis par la direction générale des Finances publiques (fichiers fiscaux pour la France métropolitaine) est réalisé pour les quatre agrégats : 1. traitements et salaires du ménage, 2. retraites, pensions et rentes, 3. revenus d'indépendants et 4. revenus déclarés du patrimoine¹⁸. Ainsi, les agrégats de calage sont fondés sur des concepts de revenus identiques à ceux de l'ERFS. Les masses fiscales exhaustives utilisées sont corrigées avant le calage afin de correspondre au même champ en termes de ménages : on exclut donc des masses sur lesquelles le calage est effectué les individus ne vivant pas en ménages dits ordinaires. Les parts des ménages ordinaires par composantes de revenus ont été estimées à partir du dispositif d'observation des revenus (FiLoSoFi).

15 Cette date du 15 novembre correspond au milieu du 4^{ème} trimestre (EEC du 4^{ème} trimestre sur lequel est assise l'enquête).

16 Source : Insee et SOeS, estimation annuelle du parc de logements au 1^{er} janvier.

17 En raison d'une erreur de codage de la CS à 2 chiffres (avec impact sur celle à 1 chiffre), les catégories socioprofessionnelles ont été rectifiées dans les enquêtes emploi. Par ailleurs les personnes retraitées âgées de moins de 53 ans sont désormais reclassées en inactives. Ces rectifications ont été intégrées dans les enquêtes ERFS de 2005 à 2007.

18 Dans la précédente série des enquêtes ERF, le calage était réalisé uniquement sur les trois premiers agrégats.

Glossaire

Ménage

Dans les enquêtes sur les revenus fiscaux (et sociaux), il s'agit de l'ensemble des occupants d'une résidence principale, qu'ils aient ou non des liens de parenté. Le ménage peut ne comprendre qu'une seule personne. Ne font pas partie des ménages ordinaires les personnes vivant dans des habitations mobiles (y compris les mariniers et les sans-abri) ou dans des communautés (foyers de travailleurs, maisons de retraite, résidences universitaires, maisons de détention, etc.).

Niveau de vie

Revenu disponible du ménage divisé par le nombre d'unités de consommation (UC). Le niveau de vie est donc le même pour toutes les personnes d'un même ménage.

Pauvreté monétaire

Un individu (ou un ménage) est considéré comme pauvre lorsqu'il vit dans un ménage dont le niveau de vie est inférieur au seuil de pauvreté. L'Insee, comme Eurostat et les autres pays européens, mesure la pauvreté monétaire de manière relative alors que d'autres pays (comme les États-Unis ou l'Australie) ont une approche absolue. Dans l'approche en termes relatifs, le seuil est déterminé par rapport à la distribution des niveaux de vie de l'ensemble de la population. On priviliege en Europe le seuil de 60 % du niveau de vie médian.

Prélèvement libératoire

En principe, quand les revenus des valeurs mobilières sont imposables, ils s'ajoutent aux autres revenus soumis à l'impôt sur le revenu. Toutefois, les produits des placements à revenu fixe (c'est-à-dire les obligations, les titres d'emprunt négociables, les créances, dépôts, cautionnements et comptes courants, les bons du Trésor et assimilés, les bons de caisse, les bons et contrats de capitalisation) peuvent être soumis à un prélèvement libératoire forfaitaire qui est un prélèvement à la source. Lorsqu'il est opéré, le prélèvement libératoire a pour conséquence de libérer de l'impôt sur le revenu les produits auxquels il s'applique.

Le prélèvement libératoire est qualifié de forfaitaire car son taux ne dépend pas du revenu global du contribuable.

A partir de la déclaration de revenus 2013, le prélèvement libératoire forfaitaire est supprimé pour les produits à revenu fixe et les dividendes. La DGFiP parle maintenant de prélèvement forfaitaire obligatoire mais non libératoire car il est ensuite imposé au barème dans l'impôt sur le revenu.

Prélèvement à la source

Le prélèvement (ou retenue) à la source est un mode de recouvrement de l'impôt, consistant à faire prélever son montant par un tiers payeur, le plus souvent l'employeur ou le banquier, au moment du versement au contribuable des revenus sur lesquels porte l'impôt.

Revenu déclaré

Somme des revenus déclarés au fisc avant abattements. Il comprend les revenus d'activité salariée ou indépendante, les indemnités de chômage, les pensions alimentaires, d'invalidité ou de retraite, et les revenus du patrimoine. Ces revenus sont nets de cotisations sociales et de CSG (contribution sociale généralisée) déductible.

La déclaration de revenus est propre au foyer fiscal. Le revenu déclaré d'un ménage correspond donc à l'ensemble des revenus déclarés des foyers fiscaux qui le composent. Dans le cas où un foyer fiscal est éclaté sur plusieurs ménages (par exemple, un enfant à charge vivant hors de la résidence familiale), on ne prend pas en compte les revenus des individus absents du ménage.

Revenu disponible

Il comprend les revenus déclarés au fisc (revenus d'activité, retraites et pensions, indemnités de chômage et certains revenus du patrimoine), les revenus financiers non déclarés et imputés (produits d'assurance-vie, livrets exonérés, PEA, PEP, CEL, PEL), les prestations sociales perçues et la prime pour l'emploi. Tous ces revenus sont nets des principaux impôts directs (impôt sur le revenu, taxe d'habitation, contribution sociale généralisée - CSG -, contribution à la réduction de la dette sociale - CRDS -, prélèvement forfaitaire sur valeurs mobilières et autres prélèvements sociaux sur les revenus du patrimoine). La mesure du revenu disponible présentée ici correspond à un concept microéconomique de revenu mesuré au travers de l'enquête Revenus fiscaux et sociaux qui permet d'étudier les disparités de revenus et la pauvreté monétaire. En parallèle, l'Insee publie des données macroéconomiques relatives au revenu disponible brut dans la Comptabilité nationale, qui font référence pour les évolutions globales. Elles ne sont pas directement comparables aux données microéconomiques, ni en niveau ni en évolution.

Unités de consommation

Les dépenses d'un ménage de plusieurs personnes ne sont pas strictement proportionnelles au nombre de personnes, grâce aux économies d'échelle issues de la mise en commun de certains biens. Aussi, pour comparer les niveaux de vie d'individus vivant dans des ménages de taille ou de composition différente, on utilise une mesure du revenu par équivalent-adulte ou par unité de consommation, à l'aide d'une « échelle d'équivalence ».

L'échelle actuellement la plus utilisée (dite de l'OCDE modifiée) consiste à décompter 1 unité de consommation (UC) pour le premier adulte du ménage, puis 0,5 UC pour les autres personnes de 14 ans ou plus, et enfin 0,3 UC pour les enfants de moins de 14 ans.