MINISTÈRE
DE L'ÉCONOMIE,
DES FINANCES
ET DE LA SOUVERAINETÉ
INDUSTRIELLE ET NUMÉRIQUE
Liberté
Égalité
Fraternité

Documents de Travail

N° 2026/3 • Novembre 2026

Direction générale du Trésor

MIEUX COMPRENDRE L'IMPÔT SUR LE REVENU GRÂCE À UN NOUVEL ÉCHANTILLON DES DÉCLARATIONS D'IMPÔT DES MÉNAGES

Elisa BUSSON
Agathe EUPHERTE
Thomas GONCALVES
Noémie SAHUC
Eva YOUINOU

MIEUX COMPRENDRE L'IMPÔT SUR LE REVENU GRÂCE À UN NOUVEL ÉCHANTILLON DES DÉCLARATIONS D'IMPÔT DES MÉNAGES

Ce document de travail n'engage que ses auteurs. L'objet de sa diffusion est de stimuler le débat et d'appeler commentaires et critiques.

Elisa Busson, Agathe, Eupherte, Thomas Goncalves, Noémie Sahuc, **Eva Youinou** sont en poste à la Direction Générale du Trésor au Ministère de l'Économie, des Finances et de la Souveraineté industrielle et numérique (France)

elisa.busson@dgtresor.gouv.fr

agathe.eupherte@dgtresor.gouv.fr

thomas.goncalves@dgtresor.gouv.fr

noemie.sahuc@dgtresor.gouv.fr

eva.youinou@dgtresor.gouv.fr

Table des matières

Résumé / Abstract 5

I. Présentation du fichier POTE 6

1. Les données issues des déclarations d'impôt sur le revenu 6

2. Contraintes techniques et solutions d'échantillonnage 6

II. Justification de la création d'un nouvel échantillon 6

1. Limites des échantillons existants 6

2. Objectif du nouvel échantillon 6

III. Méthodologie de construction de l'échantillon 7

1. Variables d'intérêt pour l'échantillonnage 7

2. Principe de l'échantillonnage équilibré 8

3. La méthode d'échantillonnage du Cube et ses mécanismes 8

4. Stratification de la population en trois groupes 8

IV. Validation de l'échantillon et statistiques descriptives 10

1. Validation des variables d'intérêt pour l'échantillonnage 10

2. Validation des taux de sondage 10

3. Statistiques descriptives de l'échantillon 12

3.1 Foyers imposables, imposés et non imposés 12

3.2 Concepts de revenus : RFR par part et revenu déclaré par UC 13

3.3 Variables de revenus par catégorie 15

3.4 Répartition de la charge fiscale et dispositifs fiscaux 18

3.5 Caractéristiques démographiques et familiales 20

V. Exploitation du nouvel échantillon 22

1. Utilisation des données de l'échantillon 22

2. Vieillissement et actualisation de l'échantillon 22

3. Simulation et chiffrage des réformes fiscales 23

4. Application pratique 1 : chiffrage d'une sous-indexation du barème de l'impôt sur le revenu 23

5. Prise en compte des effets temporels du prélèvement à la source 24

6. Application pratique 2 : chiffrage d'une réforme du prélèvement forfaitaire unique 25

VI. Conclusion et Perspectives 26

1. Synthèse des améliorations apportées par le nouvel échantillon 26

2. Perspectives 26

Annexes 27

Bibliographie 28

Résumé

Ce document présente la méthodologie de construction de l'échantillon RETRIEVER (Retraitement des déclarations d'impôt sur le revenu en un échantillon représentatif), développé par la Direction Générale du Trésor pour remplacer les échantillons fiscaux précédents et améliorer la précision des simulations de l'impôt sur le revenu français. Le fichier POTE, base exhaustive des déclarations d'impôt de 41,2 millions de foyers fiscaux, pose des contraintes techniques majeures pour les analyses statistiques en raison de sa grande taille, tandis que les échantillons existants souffrent de limites importantes : obsolescence des données (dernier échantillon fiable datant de 2019) et sous-représentation des hauts revenus, particulièrement problématique car ces segments concentrent une part significative de la masse fiscale et des revenus du capital. L'échantillon RETRIEVER est construit en adoptant la méthode d'échantillonnage équilibré *via* l'algorithme du Cube de Deville et Tillé (2004), calibré sur 14 variables fiscales clés (revenus par catégorie, réductions et crédits d'impôt, parts fiscales), garantissant que l'estimateur d'Horvitz-Thompson reproduit au plus proche les totaux de population pour ces variables auxiliaires et élimine les biais structurels des méthodes d'échantillonnage classiques. Face à la très forte hétérogénéité au sein du dernier décile, la population est stratifiée en trois groupes : déciles 1-9 (300 000 observations échantillonnées), 10^{e} décile hors dernier centile (200 000 observations échantillonnées), et dernier centile intégré exhaustivement (412 126 observations), préservant la robustesse de l'échantillonnage équilibré tout en capturant fidèlement les revenus du dernier centile (notamment ceux du capital). Les tests de calibrage démontrent une précision remarquable, avec des écarts relatifs inférieurs à 0,1 % pour la majorité des variables, l'analyse descriptive confirmant la cohérence démographique et fiscale avec une concentration de 73 % de l'impôt sur le revenu dans le 10^{e} décile, dont 32 % pour le seul dernier centile. L'échantillon permet des chiffrages précis de réformes fiscales *via* une approche comparative rigoureuse, comme l'illustre l'exemple du chiffrage du prélèvement forfaitaire unique, tandis que le processus de vieillissement intégré assure la cohérence des chiffrages avec les hypothèses retenues concernant l'évolution prévisionnelle de l'impôt sur le revenu.

Abstract

This document presents the methodology for constructing the RETRIEVER sample (Reprocessing of income tax returns into a representative sample), developed by the French Treasury to replace previous fiscal samples and improve the accuracy of French income tax simulations. The POTE file, an exhaustive database of 41.2 million tax households from income tax returns, presents major technical constraints for statistical analysis, while existing samples suffer from significant limitations: data obsolescence (the latest reliable sample dates from 2019) and under-representation of high-income households, which is particularly problematic given that these segments account for a significant share of total tax revenues and capital income. The RETRIEVER sample is built using the balanced sampling method based on Deville and Tillé's (2004) Cube algorithm, calibrated on 14 key fiscal variables (income by category, tax reductions and credits, tax units), ensuring that the Horvitz-Thompson estimator closely reproduces population totals for these auxiliary variables and eliminates the structural biases of classical sampling methods. Given the extreme heterogeneity within the top decile, the population is stratified into three groups: deciles 1 to 9 (300,000 sampled observations), the 10th decile excluding the top 1% (200,000 sampled observations), and full integration of the top centile (412,126 observations), preserving the robustness of balanced sampling while accurately capturing the high incomes of the top 1%—notably capital income. Calibration tests show remarkable precision, with relative deviations below 0.1% for most variables. Descriptive analysis confirms demographic and fiscal consistency, with 73% of income tax concentrated in the 10th decile, including 32% for the top 1% alone. The sample enables accurate fiscal reform costing through rigorous comparative analysis, as illustrated by the flat tax (prélèvement forfaitaire unique) example, while the integrated aging process ensures that the cost estimates are consistent with the assumptions made regarding the projected evolution of income tax.

I. Présentation du fichier POTE

1. Les données issues des déclarations d'impôt sur le revenu

Le **POTE** (Permanent des Occurrences de Traitement des Émissions) constitue la population de référence pour les travaux de micro-simulation de l'impôt sur le revenu. Il regroupe l'ensemble des déclarations d'impôt sur le revenu des foyers fiscaux en France, issues des formulaires 2042 et de leurs annexes. Cette base de données exhaustive et détaillée, produite et mise à disposition par la DGFiP via le Centre d'accès sécurisé aux données (CASD), offre une couverture complète des foyers fiscaux, permettant ainsi d'étudier avec précision leurs caractéristiques économiques et fiscales.

Le POTE inclut des informations essentielles, telles que :

- Les revenus déclarés (salaires, pensions, revenus du capital, revenus des indépendants, *etc.*).
- Les réductions et crédits d'impôt.
- La composition des foyers (nombre de parts fiscales, personnes à charge).

2. Contraintes techniques et solutions d'échantillonnage

En raison de sa taille importante, l'utilisation du POTE présente des contraintes techniques majeures. Avec plus de 40 millions de foyers fiscaux et une grande granularité des données, son traitement direct dans des logiciels statistiques tels que **R** ou **SAS** est particulièrement complexe et chronophage. La manipulation de cette population dans sa totalité peut entraîner des limitations en termes de mémoire et de temps de calcul, rendant difficile la réalisation d'analyses et de chiffrages dans des délais contraints.

Pour surmonter ces défis, des échantillons représentatifs, comme **RETRIEVER** et **FELIN**, sont élaborés. **FELIN**, produit par la DGFiP, est un échantillon d'environ **500 000 déclarations de revenus**, conçu pour être représentatif de l'ensemble des foyers fiscaux selon le statut imposable, le montant de revenu imposable et la nature dominante des revenus. Chaque foyer est associé à un **coefficient de pondération** permettant une extrapolation à la population totale. Ces échantillons permettent de réduire considérablement la taille des données tout en garantissant une représentativité optimale, offrant ainsi une alternative performante pour mener des études approfondies et des simulations fiscales.

II. Justification de la création d'un nouvel échantillon

1. Limites des échantillons existants

La création d'un nouvel échantillon vise à remédier aux limites constatées dans les échantillons existants et à répondre aux exigences croissantes en termes de précision et de représentativité des simulations fiscales. En effet, le dernier échantillon FELIN accessible via le CASD est celui de 2020. Toutefois, l'année 2020 a été marquée par la crise du Covid-19, événement qui a pu affecter le contenu des déclarations fiscales et la pertinence de les utiliser pour simuler des réformes en dehors du contexte particulier de cette année. C'est pourquoi, **jusqu'à la mise en place de RETRIEVER**, l'échantillon FELIN de 2019 **servait de référence** pour les travaux de simulation fiscale de la Direction Générale du Trésor. Par ailleurs, les échantillons actuels ne capturent pas de manière suffisamment précise les foyers fiscaux situés dans les hauts déciles. Cette insuffisance est problématique, car ces segments concentrent une part significative de la masse fiscale, notamment en ce qui concerne les revenus du capital et les revenus exceptionnels, et leur sous-représentation limite la qualité des analyses et la portée des simulations fiscales, qui portent parfois spécifiquement sur ces segments.

2. Objectif du nouvel échantillon

Le nouvel échantillon a pour objectif de mieux couvrir l'ensemble des foyers fiscaux, avec une attention particulière portée au décile le plus élevé de revenu. Cette amélioration vise à renforcer la qualité des simulations fiscales, en tenant compte de la diversité des situations fiscales, y compris celles des foyers aux revenus les plus élevés. Une meilleure représentativité permet d'affiner les chiffrages, notamment

en matière d'impôt sur le revenu, et d'améliorer la précision des estimations issues des modèles de micro-simulation.

Par ailleurs, la méthodologie retenue pour la construction de l'échantillon permet une **actualisation plus régulière des données utilisées pour la simulation fiscale**. Le recours direct aux données issues du fichier POTE rend possible une **construction annuelle de l'échantillon**, dès la mise à disposition des données fiscales, sans dépendre exclusivement de techniques de vieillissement sur plusieurs années. Cela contribue à limiter les incertitudes liées à l'évolution des revenus dans le temps et à améliorer la réactivité du dispositif pour les exercices budgétaires.

En complément des outils existants, ce nouvel échantillon offre ainsi un cadre actualisé et flexible pour accompagner les travaux de prévision et d'évaluation des politiques fiscales.

III. Méthodologie de construction de l'échantillon

Pour construire l'échantillon **RETRIEVER** (Retraitement des déclarations d'impôt sur le revenu en un échantillon représentatif), nous avons adopté la méthode de **l'échantillonnage équilibré**, qui repose sur l'utilisation de variables auxiliaires corrélées aux variables d'intérêt. Cette méthode vise à sélectionner un sous-ensemble d'observations représentatif de la population totale, en respectant au mieux les totaux connus pour certaines variables clés.

Deux objectifs principaux sont poursuivis :

- **Assurer la représentativité** : garantir que la distribution des variables auxiliaires dans l'échantillon reflète fidèlement celle de l'ensemble des foyers fiscaux, y compris dans les hauts déciles de revenu ;
- **Optimiser la représentativité statistique de l'échantillon** : grâce à l'application de l'**estimateur d'Horvitz-Thompson**, qui permet d'estimer correctement les agrégats de population à partir de données échantillonnées.

Dans cette approche, **chaque unité de la population a une probabilité d'inclusion différente**, qui traduit son importance relative au regard de l'impôt. Par exemple, les foyers à très hauts revenus, bien que peu fréquents, sont essentiels à représenter correctement pour ne pas biaiser les résultats. Pour assurer la représentativité de l'échantillon malgré ces différences de probabilité d'inclusion, **chaque observation se voit attribuer un poids égal à l'inverse de sa probabilité d'inclusion** : un foyer ayant 1 chance sur 100 d'être tiré représentera 100 foyers dans la population, tandis qu'un foyer avec 1 chance sur 5 en représentera 5. Cette pondération permet de reconstituer fidèlement les totaux de la population, **comme si l'on disposait des données pour l'ensemble des foyers** (selon le principe de l'estimateur d'Horvitz-Thompson).

1. Variables d'intérêt pour l'échantillonnage

Les variables utilisées pour calibrer la sélection de l'échantillon sont choisies en fonction de leur rôle déterminant dans le calcul de l'impôt sur le revenu. Elles permettent de s'assurer que les dimensions fiscales et sociales essentielles sont bien représentées dans l'échantillon final. Elles incluent :

- Impôts sur les revenus soumis au barème (avant réductions d'impôt) (€).
- Réductions d'impôt (€).
- Crédits d'impôt (€).
- Nombre de parts fiscales.
- Montant de salaire déclaré (€).
- Montant des bénéfices agricoles déclarés y compris déficits (€).
- Montant des bénéfices industriels et commerciaux déclarés y compris déficits (€).

- Montant des bénéfices non commerciaux déclarés y compris déficits (€).
- Montant des préretraites et allocations chômage déclarés (€).
- Montant des pensions déclarées (€).
- Montant des revenus fonciers déclarés (€).
- Montant des revenus de capitaux mobiliers (RCM) déclarés (€).
- Montant total des plus-values non salariées et autres (€).
- Montant des revenus variables taxés hors système du quotient (RVTO) (€).

2. Principe de l'échantillonnage équilibré

L'échantillonnage équilibré consiste à tirer un échantillon tel que, pour chaque variable auxiliaire x, l'estimateur d'Horvitz-Thompson du total

$$\hat{t}_x = \sum_{i \in Retriever} \frac{x_i}{\pi_i}$$

soit le plus proche possible du total réel de la population

$$\hat{t}_x = \sum_{i \in POTE} x_i$$

Ici, $\pi_{i,}$ désigne la probabilité d'inclusion de l'unité *i* dans l'échantillon. Cette propriété garantit que les agrégats calculés à partir de l'échantillon reflètent fidèlement ceux de la population, tout en limitant la variance due au tirage. L'approche est particulièrement performante dans les contextes où la population est hétérogène, car elle impose des contraintes fortes sur les totaux. Là où un simple plan aléatoire ou un échantillonnage stratifié peut échouer à reproduire les extrêmes de la distribution, l'échantillonnage équilibré assure une meilleure représentativité des agrégats structurants.

3. La méthode d'échantillonnage du Cube et ses mécanismes

Dans notre démarche, nous avons retenu **la méthode dite du *Cube***, développée par Deville et Tillé (2004), qui permet d'obtenir des échantillons équilibrés sur un nombre quelconque de variables auxiliaires, pour des probabilités d'inclusion $\pi = (\pi_{1,} \pi_2, \dots, \pi_N)$. L'algorithme est **itératif** et se décompose en deux types d'opérations, décrites de manière imagée comme un *vol* et un *atterrissage*.

- La **phase de vol** correspond au déplacement progressif dans l'espace des probabilités, à la recherche d'un équilibre optimal. À chaque itération, les probabilités d'inclusion de certaines unités sont ajustées, de façon à réduire peu à peu le nombre d'unités encore « indécises » (celles dont la probabilité reste comprise strictement entre 0 et 1).
- Lorsque l'équilibre parfait ne peut être atteint, intervient la **phase d'atterrissage**. Celle-ci exploite la position atteinte au terme du vol : l'espace des solutions a été réduit et il devient nécessaire de relâcher partiellement certaines contraintes pour finaliser la sélection. Le choix est déterminé automatiquement par un critère d'optimalité intégré à l'algorithme, qui vise à minimiser l'écart global entre l'échantillon obtenu et l'équilibre recherché.

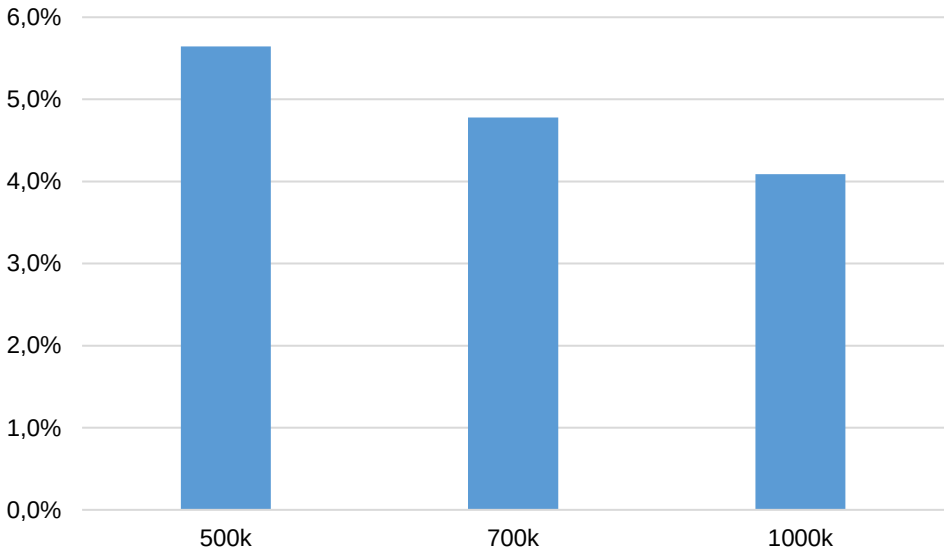
Ces dénominations imagées reflètent bien le comportement de l'algorithme : le vol prépare le terrain en guidant la trajectoire vers des zones favorables, tandis que l'atterrissage assure la faisabilité du tirage en arbitrant entre contraintes incompatibles. Ce mécanisme garantit in fine un compromis optimal entre respect des agrégats connus et faisabilité statistique.

4. Stratification de la population en trois groupes

En pratique, il est particulièrement difficile de réaliser un échantillonnage équilibré sur l'ensemble de la population, en raison de l'**hétérogénéité extrême du dernier centile**, notamment sur les revenus du

capital. Les tests empiriques menés montrent que l'augmentation de la taille brute de l'échantillon ne permet pas de corriger cette difficulté : la moyenne absolue des écarts relatifs (MAER) passe de 5,6 % pour un échantillon de 500 000 foyers à 4,81 % pour 700 000, puis 4,1 % pour 1 000 000, sans parvenir à réduire suffisamment l'écart avec la population d'intérêt. Autrement dit, **accroître la taille de l'échantillon ne règle pas le problème structurel posé par la queue de distribution**.

Graphique 1 : Moyenne absolue des écarts relatifs sur les 14 variables de calibrage selon la taille de l'échantillon hors stratification



Source : Calculs DG Trésor, échantillonnage du POTE23.

- Pour remédier à cette difficulté, la population a été **stratifiée en trois groupes distincts** :
  Groupe 1 : Les foyers appartenant aux déciles 1 à 9, c'est-à-dire l'ensemble des foyers situés dans les neuf premiers déciles de revenu.
- **Groupe 2 :** Les foyers du 10e décile, à l'exclusion du dernier centile. Ce groupe comprend des foyers avec des revenus élevés mais dont la variabilité reste moins extrême.
- **Groupe 3 :** Le dernier centile, caractérisé par une dispersion très importante des revenus, en particulier pour le revenu du capital.

Le dimensionnement des trois strates a été défini selon des logiques différenciées, combinant considérations statistiques et contraintes opérationnelles. Le **Groupe 1** (déciles 1 à 9) et le **Groupe 2** (10e décile hors dernier centile) ont été fixés respectivement à 300 000 et 200 000 observations, sur la base de plusieurs tests cherchant un équilibre entre représentativité de la population et temps de calcul du modèle de micro-simulation. Pour le Groupe 1, ce dimensionnement correspond à un taux de sondage d'environ 0,86 % (chaque observation représentant en moyenne 123,36 foyers), tandis que le Groupe 2 présente un taux de sondage de 5,39 % (chaque observation représentant en moyenne 18,55 foyers). Pour ces deux groupes, un échantillonnage équilibré est réalisé à l'aide de la méthode d'échantillonnage du Cube décrite ci-dessus, afin d'assurer une bonne estimation des totaux fiscaux et des caractéristiques économiques. À l'inverse, le Groupe 3, correspondant au dernier centile, particulièrement hétérogène et dont la bonne représentation est cruciale pour le chiffrage de certains types de mesures fiscales (*e.g.* sur la fiscalité du capital), est intégré de manière exhaustive dans l'échantillon, soit 412 126 observations, ce qui revient à un taux de sondage de 100 %. Ce choix vise à préserver la robustesse de l'échantillonnage pour l'ensemble de la population tout en évitant que la forte variabilité du haut de la distribution ne biaise les résultats globaux. Enfin, la stratification retenue permet de tenir compte du poids différencié de chaque strate dans le total de l'impôt sur le revenu : le **Groupe 1** représente 31 % de l'impôt, le **Groupe 2** en concentre 41 %, et le **Groupe 3** (dernier centile) en représente 27 %.

IV. Validation de l'échantillon et statistiques descriptives

1. Validation des variables d'intérêt pour l'échantillonnage

L'échantillon RETRIEVER reproduit l'ensemble de la population fiscale française avec plus de 41,2 millions de foyers fiscaux en 2023. Sur l'ensemble des foyers, 21,3 millions sont imposables (52 %), 18,5 millions sont effectivement imposés (45 %), et 65 000 foyers sont assujettis à la contribution exceptionnelle sur les hauts revenus (CEHR), soit 0,2 % du total.

Tableau 1 : Validation du calibrage de l'échantillon RETRIEVER : comparaison avec la population POTE 2023

Variables	POTE 2023	Retriever23	Écart relatif au POTE (%)	Écart au POTE
Nombre de parts fiscales (en milliers)	67 612	67 612	0,00%	0,0
Impôts sur les revenus soumis au barème (avant réductions et d'impôt) (M€)	92 213,1	92 199,9	−0,01%	−13,2
Réductions d'impôt (M€)	8 727,9	8 727,2	−0,01%	−0,7
Crédits d'impôt (M€)	14 664,0	14 655,6	−0,06%	−8,4
Montant de salaire déclaré (M€).	868 485,5	868 480,8	0,00%	−4,6
Montant des bénéfices agricoles déclarés y compris déficits (M€)	12 121,8	12 231,1	0,90%	109,3
Montant des bénéfices non commerciaux déclarés y compris déficits (M€)	42 497,5	42 493,1	−0,01%	−4,4
Montant des bénéfices industriels et commerciaux déclarés y compris déficits (M€)	20 928,7	20 936,7	0,04%	8,1
Montant des préretraites et allocations chômage déclarés (M€)	35 100,4	35 101,4	0,00%	1,0
Montant des pensions déclarées (retraites, invalidité et alimentaires) (M€)	374 695,9	374 699,7	0,00%	3,8
Montant des revenus fonciers déclarés (M€).	40 052,2	40 056,6	0,01%	4,4
Montant des revenus de capitaux mobiliers (RCM) déclarés (M€)	53 148,4	53 121,2	−0,05%	−27,2
Montant total des plus-values non salariées et autres (M€).	61 237,6	61 305,2	0,11%	67,6
Montant des rentes viagères à titre onéreux (RVTO) (€).	935,0	934,6	−0,05%	−0,4

Source : Calculs DG Trésor, échantillon de revenus RETRIEVER23 et POTE 2023.

Les résultats présentés dans le Tableau 1 illustrent l'efficacité de la méthode d'échantillonnage équilibré. Les écarts entre l'échantillon et la population POTE sont particulièrement faibles pour l'ensemble des 14 variables de calibrage, avec une majorité d'écarts relatifs inférieurs à 0,1 %. Cette précision confirme que l'estimateur d'Horvitz-Thompson reproduit fidèlement les totaux de population, validant ainsi la qualité de l'équilibrage obtenu par le choix des strates et la méthode du Cube. Les écarts les plus significatifs concernent les bénéfices agricoles (0,90 %) et les plus-values non salariées et autres (0,11 %), différences qui restent néanmoins acceptables.

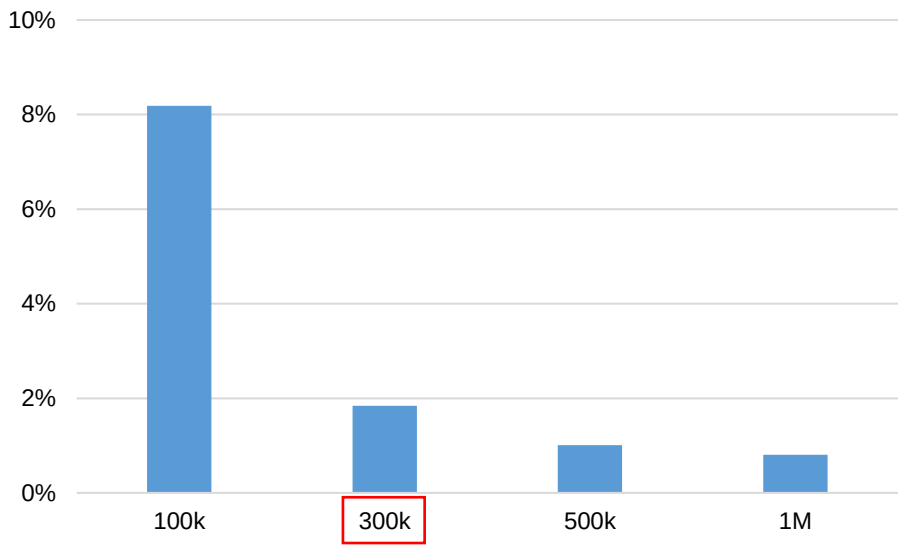
Au regard de ces résultats et de la précision sur le haut de la distribution liée aux spécificités de l'échantillonnage, l'utilisation de l'échantillon RETRIEVER s'avère préférable pour les usages de microsimulation de la DG Trésor, à celle du FELIN, d'autant plus que le dernier échantillon FELIN disponible est constitué à partir des revenus 2020, contre 2024 pour RETRIEVER.

2. Validation des taux de sondage

Afin de valider empiriquement les choix de dimensionnement réalisés, des tests ont été menés pour mesurer la précision statistique obtenue selon différentes tailles d'échantillon. Les résultats sont présentés ci-dessous à travers la moyenne absolue des écarts relatifs (MAER).

- **Groupe 1 : Déciles 1 à 9.**
  Le Graphique 2 présente la moyenne absolue des écarts relatifs (MAER) pour différentes tailles d'échantillon allant de 100 000 à 2 000 000. On constate que les gains de précision deviennent négligeables au-delà de 300 000 unités. Le dimensionnement retenu (300 000 observations) offre donc un compromis satisfaisant entre représentativité et coût de calcul.

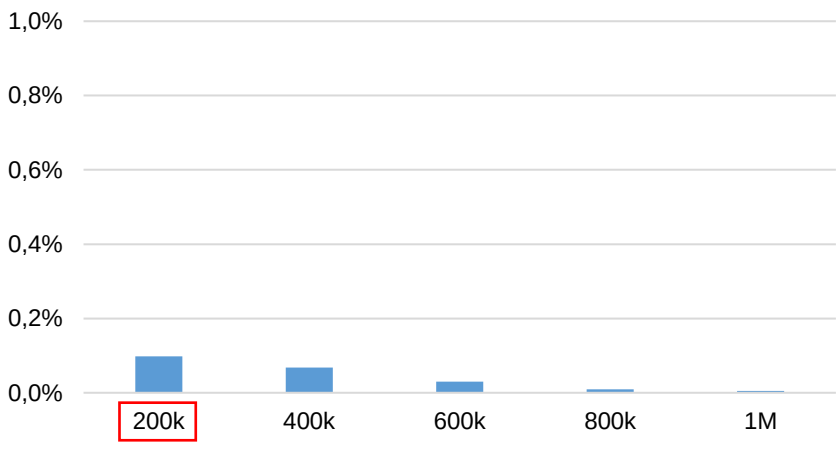
Graphique 2 : Moyenne absolue des écarts relatifs selon la taille de l'échantillon (Décile 1 à 9, Groupe 1)



Source : Calculs DG Trésor, échantillonnage du POTE23.

- **Groupe 2 : Dixième décile hors dernier centile** : le Graphique 3 illustre la variation de la moyenne absolue des écarts relatifs sur les variables de calibrage pour différentes tailles d'échantillon allant de 100 000 à 2 000 000. On constate qu'au-delà de 200 000 observations, les gains de précision deviennent marginaux. Ce résultat, qui peut sembler contre-intuitif, s'explique par le recours à un échantillonnage équilibré : dès lors qu'un effectif suffisant permet à la méthode du Cube de respecter simultanément les contraintes de calage sur les variables auxiliaires, l'ajout d'unités supplémentaires n'apporte plus d'amélioration sensible dans la perspective d'un équilibrage correct de l'échantillon. Cela justifie le choix d'un dimensionnement fixé à 200 000 pour ce groupe, garantissant à la fois la robustesse statistique et une taille raisonnable pour des calculs.

Graphique 3 : Moyenne absolue des écarts relatifs selon la taille de l'échantillon (Dernier décile, Groupe 2)

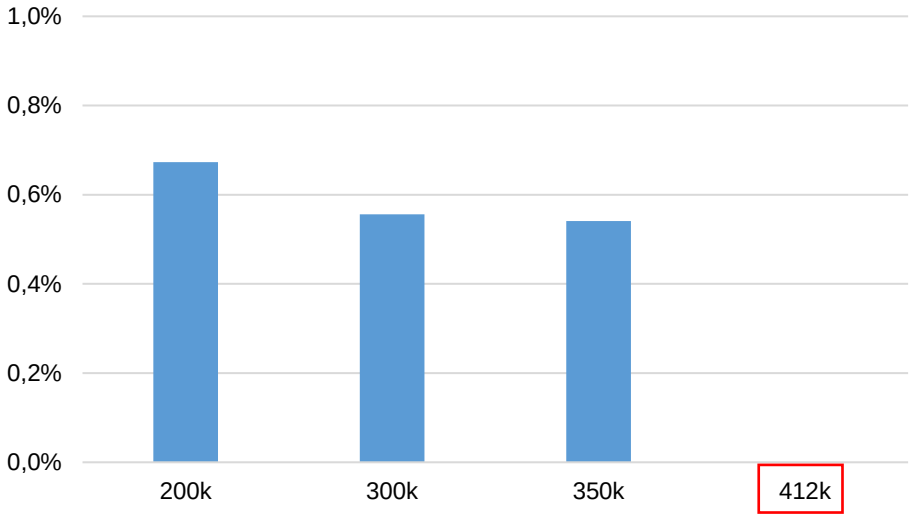


Source : Calculs DG Trésor, échantillonnage du POTE23.

- **Groupe 3 : Dernier centile.**

  Le Graphique 4 montre que les options d'échantillonnage non exhaustives dégradent fortement la précision : même en augmentant l'échantillon de 200 000 à 350 000 individus (sur un total de 412 126), les erreurs sur les totaux demeurent importantes. Cela justifie l'intégration exhaustive du dernier centile dans RETRIEVER.

Graphique 4 : Moyenne absolue des écarts relatifs sur les 14 variables de calibrage selon la taille de l'échantillon (Dernier centile, Groupe 3)



Source : Calculs DG Trésor, échantillonnage du POTE23.

3. Statistiques descriptives de l'échantillon

Pour les analyses statistiques descriptives qui suivent, nous limitons notre étude aux revenus des résidents français car l'imposition des revenus français des non-résidents diffère substantiellement de celle des résidents. En effet, leur retenue à la source comprend quatre tranches (0 %, 12 %, 20 % et 30 %), les deux premières tranches étant libératoires de l'impôt sur le revenu (IR). Les non-résidents représentent environ 0,3 million de foyers fiscaux, contre 41,0 millions de résidents français, et ne contribuent qu'à hauteur de 0,9 Md€ d'IR, soit à peine plus de 1 % du produit total de l'impôt (77,4 Md€ pour les résidents), selon l'échantillon RETRIEVER23.

3.1 Foyers imposables, imposés et non imposés

Il convient de distinguer deux concepts fiscaux. Un foyer est imposable à l'impôt sur le revenu si son IR après réductions d'impôt est strictement positif. En revanche, un foyer imposé est un foyer imposable qui acquitte effectivement l'impôt, après application de l'ensemble des mécanismes correcteurs (décote, crédit d'impôt, *etc*.).

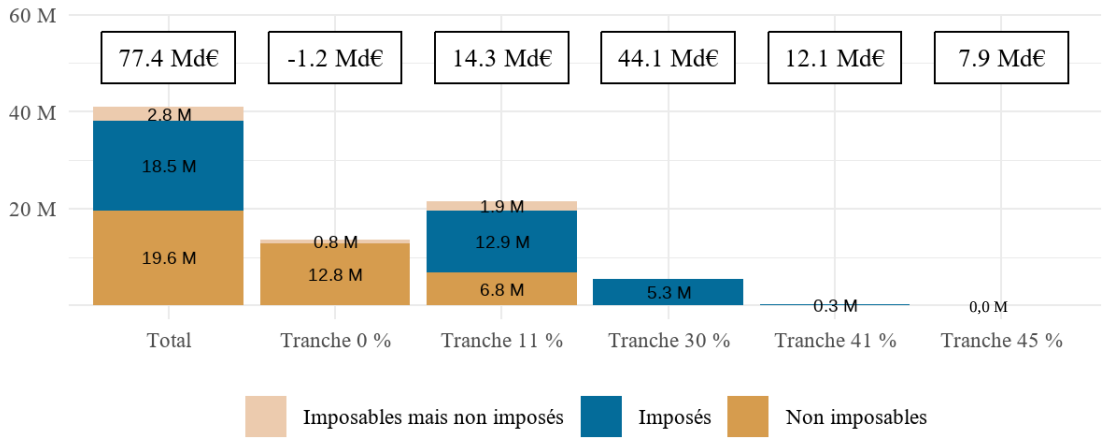
L'analyse révèle un partage en deux parts quasi égales de la population fiscale : 52 % des foyers sont imposables contre 48 % de non imposables. Parmi les foyers imposables, la grande majorité (18,5 millions sur 21,3 millions) sont effectivement imposés, la différence s'expliquant principalement par l'effet des crédits d'impôt qui peuvent annuler l'impôt dû. Une infime minorité de foyers très aisés (65 000 soit 0,2 % du total) sont assujettis à la contribution exceptionnelle sur les hauts revenus.

Tableau 2 : Foyers imposables et imposés à l'IR 2024 sur revenus 2023 (millions)

Statut	Nombre (en millions)	Pourcentage du total
Imposables	21,3	52 %
dont imposés	*18,5*	*45 %*
dont imposés à la CEHR	*0,065*	*0,2 %*
Non imposables	19,7	48 %
Ensemble des foyers	41	100 %

Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus RETRIEVER23.

Graphique 5 : Nombre de foyers non imposables, imposables et imposés à l'IR 2024 sur revenus 2023 par tranches d'imposition au barème (millions)



Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus RETRIEVER23.

Note de lecture : Parmi les foyers dans la tranche d'imposition au barème à 11 %, 6,8 M sont non imposables, 1,9 M sont imposables mais non imposés et 12,9 M sont imposés.

Le **Graphique 1** met en évidence la coexistence de foyers non imposables dont le revenu imposable relève de certaines tranches d'imposition au barème, notamment dans la tranche à 11 %. Cette situation résulte de l'application de la décote et des réductions d'impôt qui peuvent ramener à zéro l'impôt de foyers dont les revenus imposables les situent théoriquement dans des tranches d'imposition positives. À l'inverse, certains foyers sont imposables bien que leurs revenus imposables au barème soient nuls (tranche à 0 %), en raison de l'imposition à taux proportionnel de certains revenus spécifiques tels que les capitaux mobiliers, les plus-values ou les retraites en capital. Ainsi, est classé comme imposable tout foyer supportant une charge à l'impôt sur le revenu positive, y compris sur revenus non soumis au barème progressif

3.2 Concepts de revenus : RFR par part et revenu déclaré par UC

L'analyse des revenus dans l'échantillon RETRIEVER repose sur deux concepts distincts qui répondent à des logiques différentes et donnent lieu à des distributions contrastées.

Le revenu fiscal de référence et les parts fiscales constituent des concepts fiscaux définis dans le code général des impôts. Le RFR par part sert principalement au calcul de l'éligibilité à certaines prestations sociales et avantages fiscaux. Il correspond au montant des revenus et plus-values retenus pour le calcul de l'IR (après application de l'abattement pour frais professionnels sur les salaires et de son équivalent pour les retraites et du quotient pour les revenus exceptionnels ou différés), majoré de certains revenus spécifiques.

Ces revenus complémentaires incluent certains revenus exonérés d'IR (heures supplémentaires, salaires des salariés détachés à l'étranger), certains revenus soumis à prélèvement libératoire (revenus des capitaux mobiliers), certains abattements déductibles (abattement de 40 % sur les dividendes) et certaines charges déductibles du revenu (cotisations et primes d'épargne-retraite déduites du revenu global).

Le nombre de parts du quotient familial suit une logique familiale précise, avec une part pour un célibataire, deux parts pour un couple, et des demi-parts supplémentaires pour les enfants selon le barème défini réglementairement (**Tableau 3**).

Tableau 3 : Nombre de parts du foyer fiscal

Nombre de parts de quotient familial		
Nombre d'enfants	**Célibataire, divorcé ou veuf**	**Couple marié ou pacsé**
0	1	2
1	1,5	2,5
2	2	3
3	3	4
par enfant supplémentaire[2]	1	1

Source : Direction générale des Finances publiques.

Le revenu déclaré constitue un concept économique, calculé comme la somme des revenus catégoriels diminuée des déficits et moins-values. Ce concept diffère du revenu disponible des ménages qui intègre les transferts (prélèvements obligatoires et prestations sociales). Le champ du revenu déclaré diverge sur plusieurs points de celui du RFR.

Pour permettre la comparaison entre foyers de tailles ou compositions différentes, le revenu déclaré est divisé par le nombre **d'unités de consommation (UC),** calculé selon l'échelle d'équivalence de l'OCDE : 1 UC pour le premier adulte, 0,5 UC pour les autres personnes de 14 ans ou plus, 0,3 UC pour les enfants de moins de 14 ans.

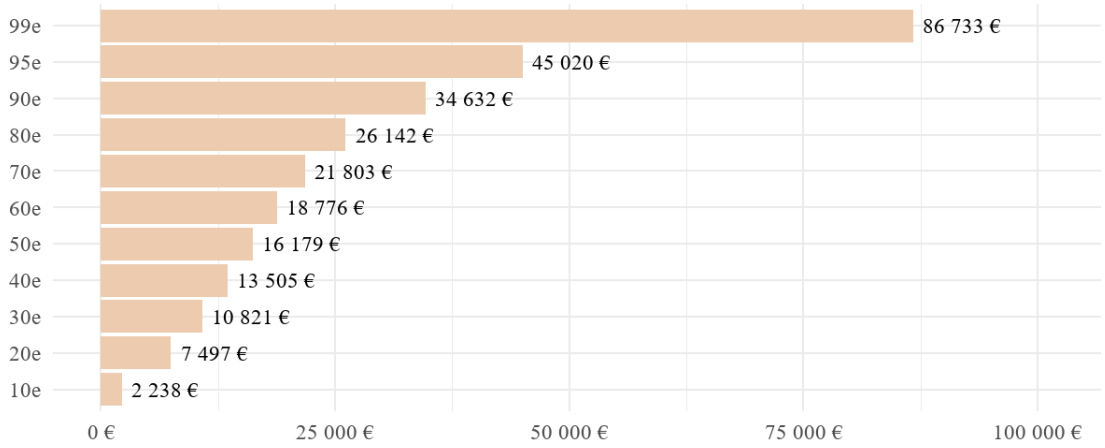
Ces deux approches génèrent des distributions de revenus distinctes car elles répondent à des objectifs différents. Le **RFR par part** privilégie la logique fiscale et administrative, ce qui le rend particulièrement adapté à l'analyse des dispositifs fiscaux et de leur ciblage, à l'étude de l'éligibilité aux prestations sociales conditionnées par le RFR, ainsi qu'aux simulations d'impact des réformes fiscales. Cette approche reflète directement les mécanismes institutionnels du système fiscal français et permet une correspondance immédiate avec les barèmes réglementaires.

Le **revenu déclaré par UC** adopte une perspective économique qui se révèle plus adaptée à l'analyse des inégalités de revenus et des niveaux de vie, aux comparaisons internationales utilisant les échelles d'équivalence standard, et à l'étude de la redistribution et de la progressivité du système fiscal. Cette approche permet une meilleure comparabilité entre foyers de compositions différentes en neutralisant les effets de taille du ménage selon une logique économique plutôt que fiscale.

Le choix entre ces deux concepts dépend donc de l'objectif analytique poursuivi : privilégier le RFR pour les analyses centrées sur les mécanismes fiscaux et sociaux existants, ou retenir le revenu par UC pour les analyses économiques et redistributives.

Le **Graphique 4** révèle que 10 % des foyers fiscaux présentent en 2023 un RFR par part supérieur à 34 632 €. Cette distribution reflète la logique fiscale du RFR, avec des seuils qui correspondent aux barèmes d'éligibilité aux dispositifs sociaux et fiscaux.

Graphique 6 : Déciles de revenu fiscal de référence 2023 par part

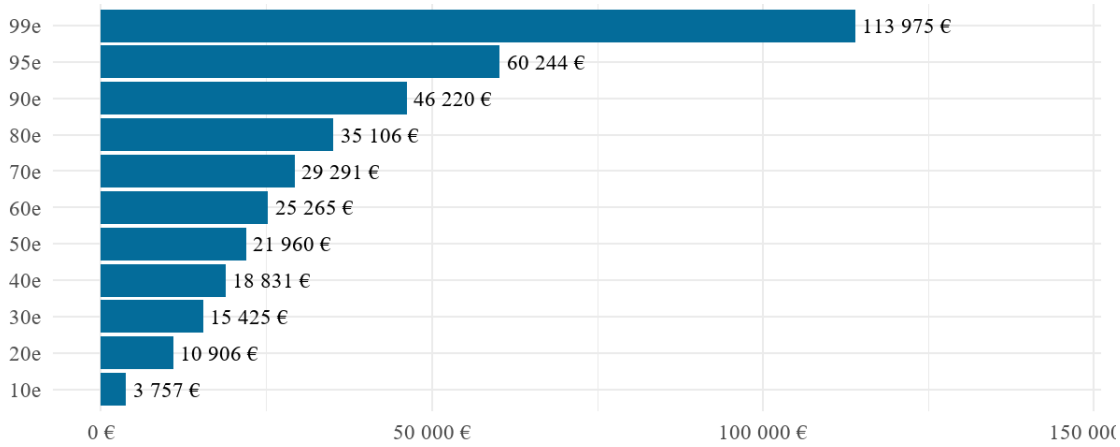


Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus RETRIEVER23.

Note de lecture : 90 % des foyers fiscaux ont en 2023 un RFR par part supérieur à 2 238 €, au titre de l'IR 2024 sur revenus 2023.

Le **Graphique 5** montre que 10 % des foyers affichent un revenu déclaré par unité de consommation supérieur à 46 220 €. Cette valeur plus élevée s'explique par les différences de calcul entre les deux concepts : l'échelle d'équivalence en UC tend à moins réduire le revenu que la division par le nombre de parts fiscales, particulièrement pour les foyers sans enfant.

Graphique 7 : Déciles de revenu déclaré 2023 par unité de consommation



Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus RETRIEVER23.

Note de lecture : 90 % des foyers fiscaux ont en 2023 un revenu déclaré par unité de consommation supérieur à 3 757 €.

3.3 Variables de revenus par catégorie

Les revenus déclarés ont été regroupés en cinq catégories :

- Les traitements et salaires ;
- Les revenus du capital, qui regroupent les intérêts et dividendes, les plus-values mobilières, les revenus fonciers, les rentes viagères et les plus-values des indépendants, qu'ils soient imposés au barème ou au PFU ;
- Les pensions de retraites ;
- Les revenus du chômage/préretraite ;
- Les revenus des indépendants qui regroupent les bénéfices agricoles (BA), les bénéfices industriels et commerciaux (BIC) et les bénéfices non commerciaux (BNC), déclarés selon le régime réel ou le régime de la micro entreprise.

Le **Tableau 4** met en regard les principaux postes de revenus issus de l'échantillon Retriever23 avec ceux publiés dans le POTE 2023. Les résultats sont bien entendu similaires à ceux du **Tableau 1** : les écarts sont faibles, systématiquement inférieurs à 0,2 % pour chaque catégorie. Sur les salaires (poste principal), l'écart est quasi nul, à moins de 5 M€ sur un total de plus de 868 Mds, soit 0,00 %. Il en va de même pour les revenus du chômage, des retraites, et du capital, avec des écarts inférieurs à 0,03 %. Le poste le plus écarté — les revenus d'indépendants — n'affiche qu'une différence de 0,15 %, soit 113 M€ sur plus de 75 Mds. L'écart total entre l'échantillon Retriever23 et les comptes POTE 2023 s'établit ainsi à 157,5 M€ sur plus de 1 509 Mds, soit un écart relatif global de 0,01 %. Cette quasi-coïncidence valide le bon calibrage de l'échantillon par rapport aux données exhaustives et confirme sa robustesse pour des simulations fiscales.

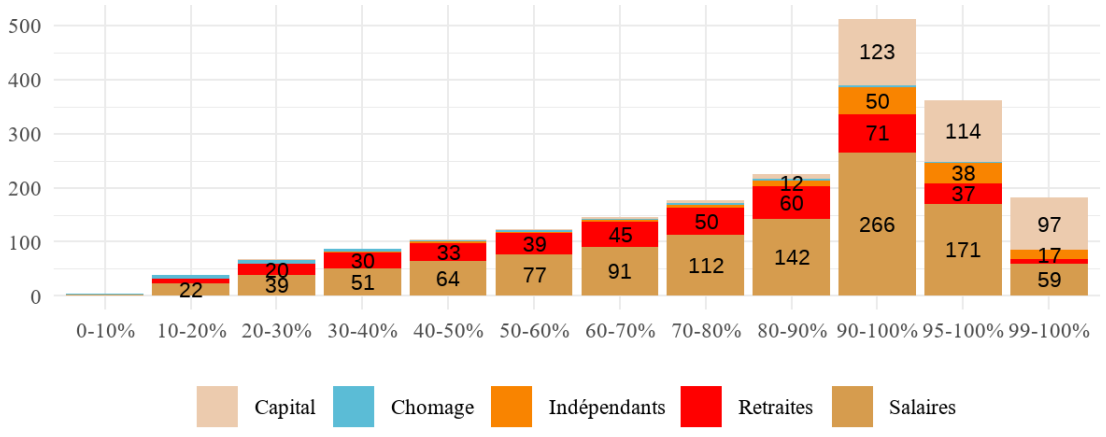
Tableau 4 : Comparaison des revenus déclarés entre l'échantillon Retriever23 et population POTE 2023 (en M€)

Revenu	Pote 2023	Retriever23	Écart relatif au POTE (%)	Écart au POTE
Salaires	868 485,5	868 480,8	0,00 %	−4,70
Chômage	35 100,4	35 101,4	0,00 %	1,00
Pensions	374 695,9	374 699,7	0,00 %	3,80
Indépendants	75 548,0	75 661,0	0,15 %	113,00
Capital	155 373,2	155 417,6	0,03 %	44,40
Total	1 509 203,0	1 509 360,5	0,01 %	157,50

Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus RETRIEVER23 et POTE 2023

Le **Graphique 6** et **le Tableau 5** mettent en évidence la structure des revenus selon les déciles de revenu déclaré par unité de consommation (UC). Dans les neuf premiers déciles, les revenus des ménages sont dominés par les salaires, qui constituent la principale source de revenu tant en masse qu'en moyenne individuelle. Pour ces déciles, les revenus salariaux progressent régulièrement, de 2,1 k€ en moyenne dans le 1er décile à 49,3 k€ dans le 9e, traduisant une montée progressive avec le niveau de vie. En masse agrégée (**Graphique 6**), ces salaires représentent plus de 598 Md€ pour les neuf premiers déciles, en raison de leur rôle central dans la distribution des revenus pour l'essentiel de la **population.**

Graphique 8 : Revenus déclarés par dixième de revenu déclaré par unité de consommation (Md€)



Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus RETRIEVER23.

Note de lecture : Pour les 4,1 M de foyers fiscaux du dernier décile de revenus, les salaires représentent 266 Md€, les revenus de pension 74 Md€, les revenus des indépendants 50 Md€ et les revenus du capital 123 Md€.

Tableau 5 : Revenus moyens déclarés par décile de revenu déclaré par unité de consommation

Revenu annuel moyen (en k€)	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	P95	P99
Nombre de déclarants (en M)	*4,1*	*4,1*	*4,1*	*4,1*	*4,1*	*4,1*	*4,1*	*4,1*	*4,1*	*4,1*	*2,0*	*0,4*
Salaires	2,1	9,0	15,1	20,0	24,0	28,2	33,0	39,5	49,3	87,2	109,5	185,0
Nombre de déclarants	*0,9*	*2,5*	*2,6*	*2,5*	*2,7*	*2,7*	*2,8*	*2,8*	*2,9*	*3,0*	*1,6*	*0,3*
Pensions	2,2	7,5	12,7	17,1	20,5	24,0	27,9	31,9	38,4	51,6	58,1	70,8
Nombre de déclarants	*0,4*	*1,5*	*1,7*	*1,8*	*1,7*	*1,7*	*1,7*	*1,6*	*1,6*	*1,4*	*0,7*	*0,1*
Chômage	2,3	5,4	6,1	5,6	5,8	6,2	7,4	8,3	9,6	14,7	17,3	19,6
Nombre de déclarants	*0,2*	*1,1*	*1,0*	*0,8*	*0,6*	*0,4*	*0,4*	*0,3*	*0,3*	*0,3*	*0,1*	*0,0*
Indépendants	2,0	7,2	10,5	12,0	13,7	14,8	16,3	19,2	24,0	54,9	70,1	125,6
Nombre de déclarants	*0,1*	*0,2*	*0,2*	*0,1*	*0,2*	*0,2*	*0,2*	*0,3*	*0,5*	*0,9*	*0,5*	*0,1*
Capital	0,2	0,6	0,7	0,8	0,9	1,1	1,4	1,9	3,0	37,8	67,0	259,4
Nombre de déclarants	*0,6*	*1,0*	*1,3*	*1,6*	*1,8*	*2,0*	*2,2*	*2,5*	*2,8*	*3,2*	*1,7*	*0,4*

Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus RETRIEVER23.

Note de lecture : Les 2,7 M de foyers déclarant des salaires au sein du 5ème décile de revenu déclaré par UC déclarent un salaire annuel moyen de 23 968€. Le salaire moyen est calculé sur les seuls foyers déclarant des salaires.

Les pensions (retraites, invalidité et alimentaires) suivent un profil similaire mais avec une amplitude moindre. En moyenne, elles vont de 2,2 k€ pour le 1er décile à 51,6 k€ pour le 10e. Leur poids relatif diminue au sein des revenus supérieurs.

Les allocations chômage, quant à elles, se distinguent par leur relative stabilité à travers les déciles : les montants moyens varient peu, entre 5 et 9 k€, avec un léger pic dans le dernier décile, lié à des cas spécifiques (hauts revenus antérieurs, *etc*.). Le caractère universel mais plafonné de ces prestations explique leur faible différenciation.

En revanche, les revenus des indépendants (BIC, BNC, BA) et surtout les revenus du capital ont une distribution qui se distingue fortement des variables précédentes. Les revenus des indépendants sont peu présents dans les bas déciles, mais progressent lentement jusqu'à 54,9 k€ dans le 10e décile, et même 125,6 k€ dans le dernier centile. Cela témoigne d'une grande hétérogénéité dans les statuts indépendants, certains ayant des revenus modestes, d'autres des revenus très élevés.

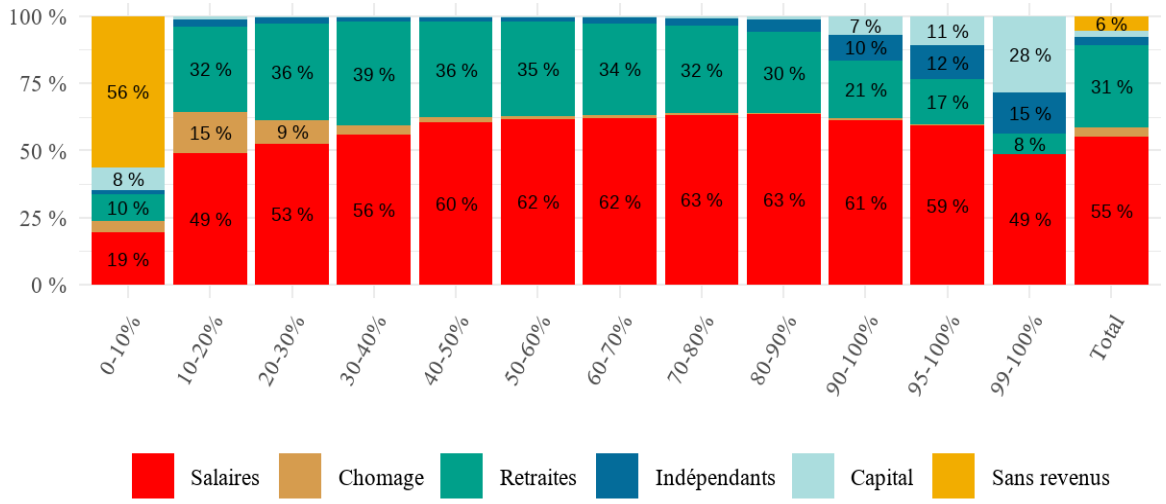
Ce sont les revenus du capital qui connaissent la concentration la plus forte. Quasi inexistants dans les premiers déciles (moins de 1 k€ en moyenne), ils commencent à croître dans le 9e décile (3,0 k€), puis deviennent massifs dans le 10e décile (37,8 k€ en moyenne), et surtout dans les tranches supérieures de ce dernier : 67,0 k€ pour les 5 % les plus aisés, 259,4 k€ pour le dernier centile. En masse, les 4,1 millions de foyers du dernier décile déclarent 123 Md€ de revenus du capital, contre moins de 10 Md€ pour l'ensemble des neuf premiers déciles. La part des revenus du capital devient majoritaire pour le dernier centile et ce fait stylisé est cohérent avec des estimations du CPO (2024), selon lesquelles la part du capital dans les revenus dépasse 86 % pour les 0,01 % les plus riches. Le rapport du CPO rappelle que les 5 % des ménages les plus riches détiennent un tiers du patrimoine, et que ce patrimoine alimente un niveau de revenu du capital important en comparaison avec le reste de la distribution.

Cette dynamique justifie pleinement la méthodologie retenue dans l'échantillon RETRIEVER23, qui prévoit une stratification fine dans le haut de la distribution afin de capturer cette concentration. Sans cette approche, la masse et la distribution des revenus du capital serait moins bien estimée, biaisant potentiellement toute analyse des réformes touchant la fiscalité du capital.

Le **Graphique 7** vient confirmer ces éléments en s'intéressant à la source principale de revenu des foyers, définie comme celle dont le montant est le plus élevé parmi les salaires, pensions/chômage, revenus des indépendants et revenus du capital. Dans l'ensemble de la population, les salaires dominent très largement, représentant la source principale pour plus de la moitié des foyers dans les déciles 2 à 9. Les revenus de remplacement sont importants dans les déciles les plus modestes, notamment dans les deux premiers, où ils concernent environ 45 à 48 % des foyers. Mais cette hiérarchie s'inverse à mesure que l'on s'élève dans la distribution. Dans le 10e décile, on observe une diversification accrue : 60 % des foyers ont toujours les salaires comme première source, mais une part

non négligeable est dominée par les revenus d'indépendants (11 %) ou du capital (8 %). Cette part du capital augmente encore dans les tranches supérieures : 11 % pour les 5 % les plus aisés, et 28 % dans le dernier centile.

Graphique 9 : Source principale de revenus déclarés par décile de revenu déclaré par unité de consommation (% des foyers du décile)



Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus RETRIEVER23.

Note de lecture : La source principale de revenus est définie comme la source de revenus ordinaires déclarés dont le montant est le plus important parmi les salaires, les revenus de remplacement, les revenus du capital et les revenus des indépendants. Parmi les foyers du 2e décile de revenu déclaré par UC, 49 % des foyers ont pour source principale de revenu des salaires, 1 % ont pour source principale de revenu du capital, 3 % ont pour source de revenu principale des revenus des indépendants, 32 % ont pour source principale de revenu des retraites et 15 des revenus de chômage.

3.4 Répartition de la charge fiscale et dispositifs fiscaux

L'analyse de l'impôt sur le revenu par déciles illustre la forte progressivité du système fiscal et valide la capacité de l'échantillon à reproduire cette caractéristique essentielle (*cf.* **Graphiques 8 et 9**). L'impôt moyen par décile passe de 0,0 k€ pour les quatre premiers déciles à 16,4 k€ pour le dixième décile, atteignant 25,9 k€ aux cinq derniers centiles et 70,5 k€ au dernier centile. La répartition de la charge fiscale confirme cette concentration : le dixième décile concentre 73 % de l'impôt sur le revenu total, dont 58 % pour les 5 % les plus aisés et 32 % pour le seul dernier centile. Cela confirme l'utilité de l'attention particulière portée aux plus hauts déciles dans la construction de l'échantillon RETRIEVER.

Le taux moyen d'imposition progresse tout au long de la distribution des revenus déclarés par unité de consommation, de manière cohérente avec la progressivité du barème, atteignant 13 % pour le dixième décile. Cette progression témoigne de l'effet redistributif à l'œuvre dans le système d'imposition. Il convient de souligner que l'ensemble des analyses menées sur l'échantillon Retriever s'appuient sur les revenus déclarés, qui diffèrent du concept de revenu économique employée par l'IPP[1], ce dernier incluant notamment les bénéfices non distribués des sociétés contrôlées par les foyers fiscaux. En rapportant l'ensemble des impôts payés par les ménages à ce revenu économique, l'IPP montre que l'imposition des revenus devient régressive au sein des 0,1 % des ménages les plus aisés. L'utilisation du revenu économique constitue ainsi un instrument complémentaire de mesure de la progressivité dans le haut de la distribution, et n'entre pas en contradiction avec le constat de progressivité d'ensemble du système d'imposition rapporté aux revenus déclarés.

[1] IPP (2023), « Quels impôt les milliardaires paient-ils ? »

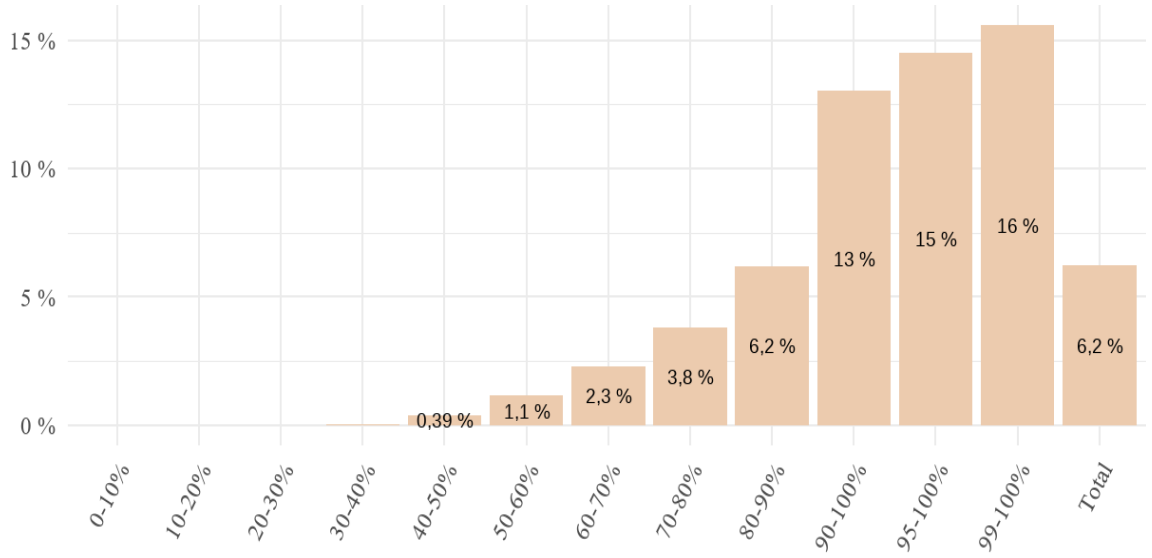
Tableau 6 : Impôt sur le revenu moyen par décile de revenu déclaré par unité de consommation

	0-10 %	10-20 %	20-30 %	30-40 %	40-50 %	50-60 %	60-70 %	70-80 %	80-90 %	90-100 %	95-100 %	99-100 %	Total
Impôt moyen du décile (k€)	0,0	0,0	0,0	0,0	0,1	0,3	0,8	1,6	3,3	16,4	25,9	70,5	2,2
Poids du décile dans l'IR	0 %	0 %	0 %	0 %	0 %	1 %	3 %	7 %	15 %	73 %	58 %	32 %	100 %

Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus RETRIEVER23.

Note de lecture : Les foyers du 10e décile de revenu déclaré par unité de consommation ont un IR après RI moyen de 16,4 k€ (moyenne arithmétique simple sur les ménages du décile) et concentrent 73 % de l'IR total.

Graphique 10 : Taux moyen d'imposition par décile de revenu déclaré par unité de consommation (pondéré)



Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus RETRIEVER23.

Note de lecture : Le taux moyen d'imposition est calculé à l'échelle du décile (IR après RI total du décile rapporté au revenu déclaré total du décile). Le taux moyen d'imposition du 10e décile est de 13 %. Le taux moyen pour un décile est ici calculé en rapportant l'IR payé par le décile aux revenus du décile, ce qui revient à calculer la moyenne des taux moyens d'imposition des ménages du décile en pondérant chaque ménage par son poids dans les revenus du décile, et non à calculer la moyenne arithmétique simple des taux d'imposition moyens individuels du décile. Pour cette raison, on parle de taux moyen « pondéré » du décile.

L'analyse des crédits d'impôt révèle une distribution différente de celle de l'impôt brut, avec une concentration moins marquée mais néanmoins significative dans les hauts déciles. Le crédit d'impôt moyen croît de 0,0 k€ dans le premier décile à 2,0 k€ dans le dixième décile, atteignant 3,3 k€ aux cinq derniers centiles et 11,0 k€ au dernier centile. Le dixième décile concentre 57 % des crédits d'impôt totaux, illustrant l'effet des dispositifs d'investissement et d'épargne principalement utilisés par les contribuables aisés.

Tableau 7 : Montant de crédits d'impôt par décile de revenu déclaré par unité de consommation

	0-10 %	10-20 %	20-30 %	30-40 %	40-50 %	50-60 %	60-70 %	70-80 %	80-90 %	90-100 %	95-100 %	99-100 %	Total
CI moyen du décile (k€)	0,0	0,1	0,1	0,1	0,1	0,2	0,2	0,3	0,4	2,0	3,3	11,0	0,3
Poids du décile dans les CI	1 %	2 %	3 %	4 %	4 %	5v%	6 %	8 %	11 %	57 %	48 %	32 %	100 %

Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus RETRIEVER23.

Note de lecture : Les foyers du 9e décile de revenu déclaré par unité de consommation bénéficient en moyenne d'un montant de CI de 395 € et concentrent 11 % des crédits d'impôt au titre des revenus 2023.

Cette concentration des crédits d'impôt, bien que marquée, reste inférieure à celle de l'impôt sur le revenu (57 % contre 73 % pour le D10). Cette différence s'explique par la nature diversifiée des dispositifs : l'impôt sur le revenu présente une progressivité qui augmente avec les revenus, tandis que les crédits d'impôt regroupent des mesures incitatives de portée variable. Certains dispositifs concernent principalement les classes moyennes et supérieures (emploi à domicile, frais de garde, dons), tandis que d'autres s'adressent davantage aux contribuables aisés (investissements défiscalisants). Par ailleurs, de nombreux crédits d'impôt sont soumis à des plafonnements qui limitent les montants perçus par les très hauts revenus, contrairement à l'impôt sur le revenu qui croît sans plafond supérieur avec les revenus déclarés.

Le gain associé au quotient familial présente une distribution plus équilibrée entre les déciles, avec un gain moyen croissant de 0,1 k€ dans le premier décile à 3,0 k€ dans le dixième décile. La distribution du gain associé au quotient familial révèle une progression plus lissée entre les déciles, avec un pic dans les déciles moyens-supérieurs (D8-D10) qui correspond aux familles avec enfants disposant de revenus intermédiaires à élevés. Le plafonnement du gain explique la relative stabilisation au niveau du dernier centile (3,3 k€ contre 3,0 k€ pour l'ensemble du D10).

Tableau 8 : Gain lié au quotient familial par décile de revenu déclaré par UC

	0-10 %	10-20 %	20-30 %	30-40 %	40-50 %	50-60 %	60-70 %	70-80 %	80-90 %	90-100 %	95-100 %	99-100 %	Total
Gain total du décile (en Md€)	0,0	0,1	0,7	1,1	1,4	1,5	1,7	2,3	3,3	4,1	2,1	0,5	16,3
Gain moyen (en k€)	0,1	0,3	0,7	0,9	1,0	1,1	1,3	1,7	2,4	3,0	3,1	3,3	1,5

Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de RETRIEVER23.

Note de lecture : Pour les foyers fiscaux bénéficiant du quotient familial, le montant moyen du gain lié au quotient familial est de 3,0 k€ pour le dernier décile. Le montant global du gain lié au quotient familial pour l'ensemble du dernier décile est de 4,1 Md€.

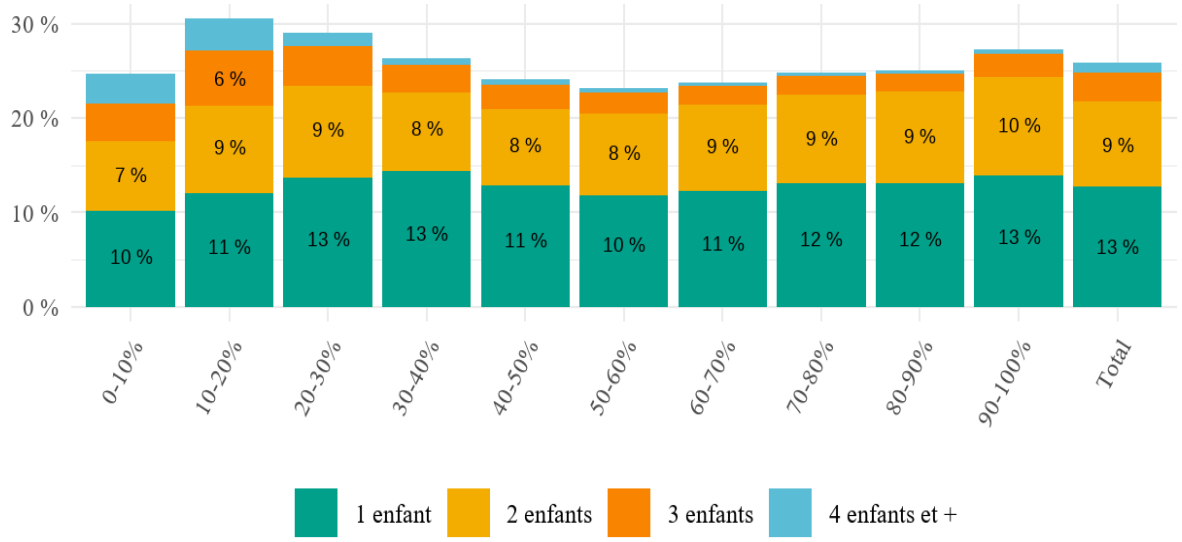
3.5 Caractéristiques démographiques et familiales

Le **Graphique 10** montre une répartition relativement stable en ce qui concerne le nombre d'enfants par foyer : dans tous les déciles, la part de foyers ayant un ou deux enfants reste relativement constante, autour de 10 à 13 % pour un enfant et 8 à 10 % pour deux enfants. Les foyers avec trois enfants sont légèrement plus fréquents dans les premiers déciles. Ces résultats s'accordent avec les données de fécondité publiées par l'Insee, qui montrent une certaine stabilité des tailles de fratrie selon le revenu, bien que les femmes les plus modestes aient un nombre d'enfants moyen légèrement plus élevé que les plus aisées (2,27 enfants contre 1,98 pour le dernier décile selon les données fiscales de 2018).

Au-delà des écarts de fécondité, cette relative stabilité dans le haut de la distribution peut aussi s'expliquer par des comportements de **rattachement fiscal des enfants majeurs**. Les foyers aisés y recourent plus fréquemment, à la fois parce que les enfants poursuivent davantage d'études supérieures, retardant leur autonomie fiscale, et parce que ce rattachement est fiscalement avantageux. En effet, il permet soit une majoration du quotient familial (pour un enfant majeur célibataire), soit un abattement forfaitaire d'environ 6 794 € par personne rattachée (pour un enfant majeur marié ou pacsé).

Ces pratiques contribuent à maintenir une proportion relativement stable de foyers déclarant un ou deux enfants dans les derniers déciles, alors même que la fécondité y est en moyenne plus faible que dans les déciles inférieurs.

Graphique 11 : Nombre d'enfants rattachés au foyer fiscal par décile de revenu déclaré par unité de consommation (% de foyers du décile)



Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus 2023 vieillis pour 2024.

Note de lecture : Afin d'éviter le double compte d'enfants en garde partagée déclarés dans deux foyers fiscaux, un enfant en garde partagée compte pour 0,5 enfant dans chaque foyer fiscal le déclarant. Parmi les 4,1 M de foyers du premier décile de revenu déclaré par UC, 25 % ont au moins un enfant : 10 % ont un enfant, 7 % ont deux enfants, 4 % ont 3 enfants, 3 % ont plus de 3 enfants.

En outre, l'âge moyen du déclarant principal augmente significativement avec le revenu, passant de 44,2 ans dans le premier décile à 55,1 ans dans le dernier. Cette progression traduit un effet clair de cycle de vie : les revenus progressent avec l'âge en activité. L'Insee observe des profils similaires, où les 50–64 ans sont surreprésentés parmi les hauts déciles de niveau de vie, tandis que les plus jeunes adultes sont davantage présents dans les déciles inférieurs.

Cette lecture doit toutefois être nuancée par le rôle des retraités. Bien qu'ils représentent une part importante de la population, leur présence est moins marquée dans les tout derniers déciles, car leurs ressources, majoritairement constituées de pensions (en 2022, près de 80 % du revenu disponible des ménages dont au moins l'un des membres est retraité selon la DREES[2]), plafonnent à des niveaux généralement inférieurs à ceux des actifs en fin de carrière. L'Insee[3] montre d'ailleurs que les retraités sont répartis de façon relativement homogène entre les 2e et 9e déciles de niveau de vie.

Ainsi, le fait que l'âge moyen culmine à 55 ans dans le dernier décile s'explique surtout par la prédominance des actifs en fin de carrière dans le haut de la distribution, tandis que les retraités se concentrent davantage dans les déciles intermédiaires à supérieurs, sans dominer les extrêmes.

[2] DREES (2025), « Les retraités et les retraites ».
[3] Insee (2024), « Niveau de vie et pauvreté selon l'activité ».

Tableau 9 : Âge moyen du déclarant principal par décile de revenu déclaré par unité de consommation

	0-10 %	10-20 %	20-30 %	30-40 %	40-50 %	50-60 %	60-70 %	70-80 %	80-90 %	90-100 %
Âge du déclarant	44,2	47,6	50,9	53,0	52,2	52,8	53,4	53,6	54,3	55,1

Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus RETRIEVER23.

Note de lecture : Dans les foyers du premier décile de revenu déclaré par unité de consommation, le déclarant principal a en moyenne 44,2 ans.

Ces statistiques descriptives confirment globalement la capacité de l'échantillon RETRIEVER à reproduire les caractéristiques complexes de la population fiscale française, notamment dans ses dimensions les plus critiques pour les analyses et simulations fiscales.

V. Exploitation du nouvel échantillon

La présente section détaille la méthode employée pour utiliser l'échantillon RETRIEVER dans le cadre des simulations fiscales et du chiffrage des réformes de l'impôt sur le revenu.

1. Utilisation des données de l'échantillon

L'échantillon RETRIEVER, construit selon la méthodologie d'échantillonnage équilibré décrite précédemment, constitue la base de données de référence pour les travaux de micro-simulation fiscale de la direction générale du Trésor. Il permet également d'estimer des élasticités de l'impôt aux différentes assiettes (salaires, pensions, *etc*.) utilisées pour la prévision d'impôt sur le revenu des années N+2 et N+3, notamment dans le cadre de l'élaboration des projets de loi de finances. Cette utilisation prospective nécessite une étape préalable de « vieillissement » de l'échantillon pour actualiser les données aux années cibles.

2. Vieillissement et actualisation de l'échantillon

Pour simuler l'impôt sur le revenu des années postérieures à celle des revenus de l'échantillon, la première étape consiste à actualiser les paramètres de la législation fiscale. Cette mise à jour s'effectue selon deux modalités complémentaires. D'une part, concernant les paramètres indexés sur l'inflation, tels que les tranches du barème de l'impôt ou les seuils de réductions et crédits d'impôt, sont appliqués les taux d'inflation inscrits dans la loi et / ou prévisionnels jusqu'à l'année cible. D'autre part, les paramètres explicitement redéfinis par les lois de finances successives sont actualisés selon leurs nouvelles valeurs votées.

Le processus de vieillissement des revenus constitue l'étape centrale de cette actualisation. Il débute par le calcul des revenus bruts de l'année de référence, reconstitués en appliquant les paramètres de la législation sociale correspondante. Cette reconstitution est nécessaire car les prévisions macroéconomiques portent sur des revenus bruts, alors que l'échantillon contient principalement des revenus nets déclarés.

L'application des évolutions macroéconomiques s'effectue ensuite en appliquant les taux de croissance prévisionnels par catégorie de revenus, d'après les premières données disponibles sur l'exécution de l'année N+1 et les hypothèses macroéconomiques retenues dans le cadre des exercices de prévision pour les années postérieures. Cette évolution s'applique uniformément au sein de chaque catégorie de revenus, sous l'hypothèse simplificatrice que l'ensemble des revenus d'une catégorie donnée évoluent comme le revenu moyen de cette catégorie. Par exemple, les salaires déclarés par chaque ménage de l'échantillon sont « vieillis » selon l'évolution du salaire moyen brut, les pensions selon l'évolution moyenne des pensions, et ainsi de suite. Enfin, pour tenir compte de l'évolution démographique, la pondération de l'échantillon est simultanément ajustée par un coefficient d'accroissement démographique issu des projections de population de l'Insee. Ce coefficient reflète l'évolution globale de la population résidente, mais ne distingue pas explicitement l'emploi et le nombre de retraités.

3. Simulation et chiffrage des réformes fiscales

Une fois les revenus vieillis et les paramètres actualisés, la simulation de l'impôt sur le revenu s'effectue en recalculant l'ensemble des composantes fiscales selon la législation de l'année cible. Cette simulation nécessite de recalculer un certain nombre de variables intermédiaires, comme la contribution sociale généralisée ou le revenu fiscal de référence, qui entrent dans le calcul de diverses composantes de l'impôt. Les différents crédits et réductions d'impôt sont également calculés individuellement selon leurs modalités spécifiques, sans vieillissement, permettant d'obtenir un impôt net après application de l'ensemble des dispositifs fiscaux. L'échantillonnage équilibré de RETRIEVER garantit que ces simulations respectent les agrégats fiscaux de référence, assurant ainsi la cohérence des résultats avec les données macroéconomiques connues.

Le chiffrage d'une réforme fiscale repose sur une approche comparative rigoureuse consistant à simuler l'impôt sur le revenu dans deux configurations distinctes. Le scénario de référence applique la législation en vigueur sans modification, tandis que le scénario réformé applique la même législation modifiée selon les dispositions du projet de réforme étudié. La comparaison systématique des résultats obtenus dans ces deux configurations permet de quantifier précisément l'impact budgétaire global de la réforme (hors effet des éventuelles réactions de comportement des contribuables), d'analyser la répartition des gains et pertes et leur montant moyen par catégories de contribuables (par exemple par décile de revenu) ou de dénombrer les foyers fiscaux gagnants et perdants.

L'échantillon RETRIEVER permet de traiter différents types de réformes avec des niveaux de complexité variables. Les réformes purement paramétriques, consistant en la modification de paramètres existants tels que les taux marginaux, les abattements ou les plafonds, sont généralement les plus simples à simuler, notamment lorsque les variables concernées sont bien couvertes dans l'échantillon. Les réformes d'assiette telles que la fiscalisation des heures supplémentaires, en revanche, sont plus complexes à modéliser, car elles impliquent parfois des données qui ne sont pas directement disponibles et nécessitent alors des hypothèses complémentaires ou des appariements avec d'autres sources. La simulation de la création de nouveaux dispositifs fiscaux (crédits ou réductions d'impôt) peut également s'avérer délicate, en particulier en l'absence d'informations sur les dépenses éligibles, ce qui conduit à recourir à des modélisations *ad hoc* fondées sur des données externes ou des enquêtes spécifiques.

Les réformes structurelles, qui modifient en profondeur les modalités de calcul de l'impôt, constituent le niveau de complexité le plus élevé et impliquent souvent des adaptations méthodologiques importantes. L'échantillon RETRIEVER, du fait de sa taille et de la structure de son échantillonnage, permet dans une certaine mesure de mieux appréhender ces réformes. En particulier, la couverture étendue des hauts déciles et l'inclusion exhaustive du dernier centile peuvent contribuer à une meilleure estimation des effets sur les revenus les plus élevés, notamment pour les réformes portant sur les revenus du capital ou les revenus exceptionnels.

4. Application pratique 1 : chiffrage d'une sous-indexation du barème de l'impôt sur le revenu

Le barème progressif de l'impôt sur le revenu est généralement revalorisé chaque année à hauteur de l'inflation afin d'éviter une augmentation mécanique de la charge fiscale liée à la progression des revenus nominaux. La réforme étudiée consiste de manière illustrative à appliquer une sous-indexation d'un point : au lieu d'une revalorisation de 1,8 %, le barème n'est relevé que de 0,8 % en 2025.

Cette mesure accroît mécaniquement l'impôt payé par l'ensemble des foyers imposables et augmente le nombre de foyers imposables. Le chiffrage repose sur une méthode comparative classique : l'impôt est simulé pour chaque foyer dans deux scénarios – avec et sans sous-indexation – puis les écarts sont agrégés en tenant compte des poids de sondage.

L'application de cette méthode conduit à estimer qu'une sous-indexation d'un point aurait généré un supplément de recettes fiscales d'environ **1,4 Md€** d'IR 2025 assis sur les revenus 2024. Ce montant serait essentiellement réparti sur la moitié haute de la distribution des revenus (la moitié basse étant peu imposable) et concentrée de manière prépondérante sur les ménages les plus aisés, comme l'IR lui-même. Comme pour les autres réformes, cette estimation est réalisée en statique, sans prise en compte d'éventuels effets de comportement ni d'adaptation des revenus.

Tableau 10 : Gain fiscal associé à une sous indexation du barème de 1 point, par rapport à une situation contrefactuelle sans sous indexation

	0-10 %	10-20 %	20-30 %	30-40 %	40-50 %	50-60 %	60-70 %	70-80 %	80-90 %	90-100 %	Total
Gain fiscal total dans le décile (Md€)	0,0	0,0	0,0	0,0	0,1	0,1	0,2	0,2	0,3	0,5	1,4
Perdants dans le décile (en %)	0,1	0,1	0,1	6,3	45,7	71,2	90,9	96,9	98,0	97,3	50,6

Source : Modèle de microsimulation de l'impôt sur le revenu de la DG Trésor, échantillon de revenus 2023 vieillis pour 2024.

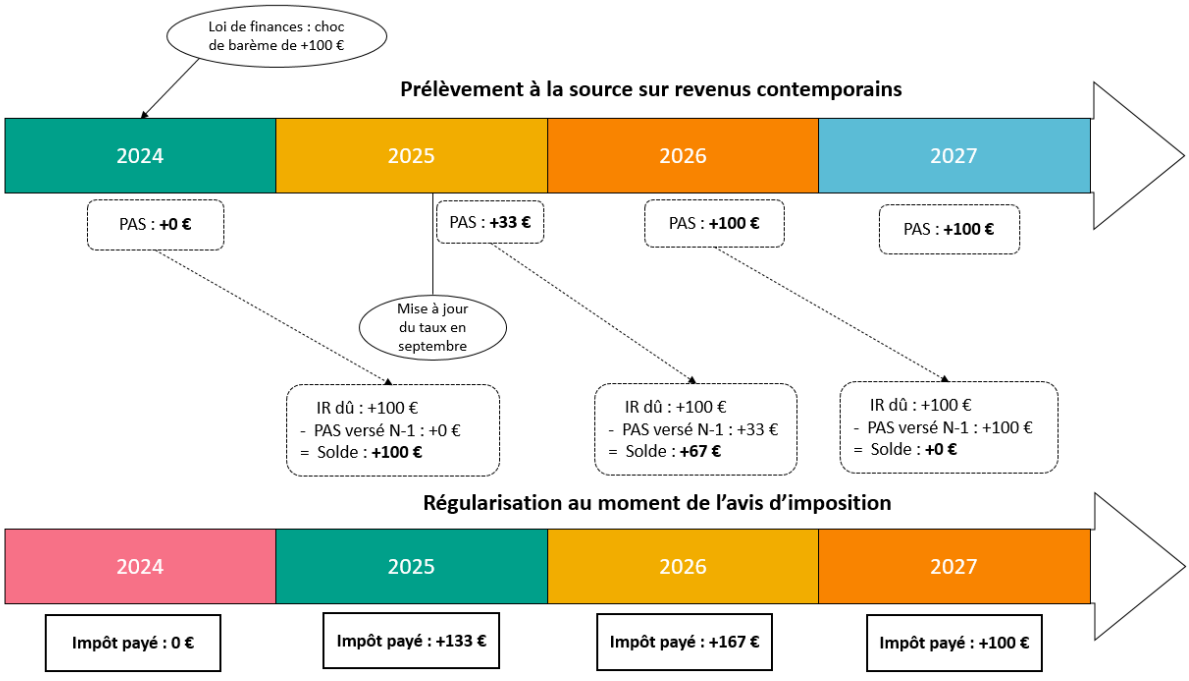
Note de lecture : Une sous-indexation d'un point du barème en 2025 se traduit par un gain budgétaire supporté de 0,1 Md€ pour le 5e décile, contre 0,5 M€ pour le dernier décile. Dans ce dernier décile, 97,3 % des foyers imposables sont affectés.

5. Prise en compte des effets temporels du prélèvement à la source

Les variations calculées sur l'IR dû par les ménages au titre de chaque année, comme dans la section précédente, ne tiennent pas compte du prélèvement à la source (PAS), qui induit une montée en charge sur trois ans des impacts des évolutions du barème sur les recettes d'IR net budgétaire.

En effet, le calendrier de mise à jour des taux de prélèvement à la source retarde l'absorption d'un choc d'imposition affectant le taux d'imposition d'un ménage en raison, par exemple, d'un changement de situation. Un tel choc peut également être dû à un changement de législation non anticipé dans le prélèvement à la source. C'est le cas de la revalorisation du barème de l'impôt sur le revenu sur l'inflation chaque année. Prenons l'exemple d'une revalorisation du barème sur l'inflation qui résulte en une augmentation de 100 € d'impôt pour un contribuable en 2025. L'impact du prélèvement à la source dans la diffusion de ce choc est illustré avec le **Graphique 12**.

Graphique 12 : Choc de barème et prélèvement à la source



Source : DG Trésor.

Le **Tableau 11** présente la montée en charge d'une mesure d'évolution du barème de l'IR N+1 au titre des revenus N.

Tableau 11 : Chronique de l'impact total pour un choc de 100 sur le barème

Année de revenus	N	N+1	N+2	N+3
Impact total sur l'IR net budgétaire		**133**	**167**	**100**
PAS		33	100	100
Solde		100	67	0
dont IR barème		*100*	*100*	*100*
dont Remboursement de PAS			*-33*	*-100*

Note de lecture : Un choc de 100 sur l'IR barème N+1 au titre des revenus N aura un impact de 33 sur le prélèvement à la source N+1 : la mise à jour des taux de PAS a lieu en septembre et donc l'impact sur le PAS ne sera que d'un tiers du montant total. En N+2, le prélèvement à la source prend bien tout l'effet de l'impact, mais le solde n'est pas encore pleinement affecté car le PAS de l'année précédente n'était pas pleinement affecté. En N+3, toutes les chroniques ont fini de monter en charge.

Ainsi, la réforme simulée dans la section précédente aurait rapporté respectivement 1,8 Md€, 2,3 Md€ et 1,4 Md€ en 2025, 2026, 2027. En pratique, certaines réformes conduisent à la fois à des effets sur l'assiette et des effets sur les taux moyens d'imposition, pouvant conduire à des dynamiques plus complexes.

6. Application pratique 2 : chiffrage d'une réforme du prélèvement forfaitaire unique

Un second exemple d'utilisation de l'échantillon concerne une modification du prélèvement forfaitaire unique (PFU) applicable aux revenus de l'épargne et du capital. Le PFU, institué par la loi de finances pour 2018, constitue un impôt proportionnel au taux global de 30 % (12,8 % au titre de l'impôt sur le revenu et 17,2 % au titre des prélèvements sociaux) qui s'applique aux revenus de l'épargne et du capital hors immobilier. La réforme étudiée consiste en une augmentation d'un point du taux d'impôt sur le revenu applicable dans le cadre du PFU, portant ce taux de 12,8 % à 13,8 %.

Les revenus de capitaux mobiliers étant concentrés sur les ménages aux plus hauts revenus, les estimations sont sensibles à la qualité de la représentation des foyers les plus aisés. Les échantillons précédents sous-représentaient ces foyers, introduisant des biais potentiels dans l'évaluation des réformes ciblant les revenus du patrimoine. L'échantillon RETRIEVER permet de mieux appréhender ces effets.

Le chiffrage s'appuie sur une méthode comparative classique : l'impôt est simulé pour chaque foyer de l'échantillon dans deux scénarios – avec et sans réforme – en simulant également le choix optimal d'imposition entre PFU et barème progressif. L'impact budgétaire est ensuite obtenu en agrégeant les variations individuelles pondérées selon les poids de sondage.

L'application de cette méthode conduit à estimer qu'une hausse d'un point du taux du PFU aurait généré un supplément de recettes fiscales d'environ 0,6 Md€ en année pleine[4] en 2025, concentré sur le dernier décile. Ce montant reflète à la fois l'effet direct de la hausse pour les foyers qui conservent le PFU, et l'effet indirect du basculement de certains contribuables vers le barème. Il convient toutefois de souligner que cette estimation est réalisée *en statique*, c'est-à-dire sans prise en compte des effets de comportement ni de réallocation de l'épargne qu'une telle réforme pourrait induire.

[4] C'est-à-dire, hors prise en compte des effets du prélèvement à la source.

VI. Conclusion et Perspectives

1. Synthèse des améliorations apportées par le nouvel échantillon

La création de l'échantillon RETRIEVER répond à un besoin d'actualisation et d'amélioration des outils de micro-simulation fiscale disponibles pour l'analyse de l'impôt sur le revenu. Ce nouvel échantillon apporte des améliorations substantielles par rapport aux dispositifs précédents.

L'innovation méthodologique principale réside dans l'adoption de la méthode d'échantillonnage équilibré *via* la méthode du Cube, qui garantit une représentativité optimale sur l'ensemble des variables fiscales d'intérêt. Cette approche, calibrée sur quatorze variables clés incluant les différentes catégories de revenus, les réductions et crédits d'impôt, assure une fidélité remarquable entre l'échantillon et la population totale des déclarants. L'estimateur d'Horvitz-Thompson appliqué à ces variables auxiliaires produit des estimations très proches des totaux, limitant les biais structurels qui pouvaient affecter les échantillons précédents.

L'amélioration la plus significative concerne la représentation des hauts déciles de revenus, et notamment la décision d'intégrer exhaustivement le dernier centile dans l'échantillon. Cette stratification permet de capturer fidèlement les revenus du capital et les revenus exceptionnels, caractérisés par une forte variabilité et une concentration importante dans cette strate de population. Cette représentation exhaustive des hauts revenus se traduit directement par une précision accrue des chiffrages de réformes fiscales, particulièrement pour les mesures touchant aux revenus du capital.

L'échantillon RETRIEVER offre également une réactivité renforcée face aux évolutions économiques et législatives. Sa construction méthodologique permet une actualisation plus fréquente des données utilisées pour les prévisions fiscales, réduisant ainsi la dépendance aux techniques de vieillissement et limitant l'accumulation d'incertitudes liées à l'extrapolation temporelle des données.

2. Perspectives

L'échantillon RETRIEVER, bien qu'offrant des améliorations substantielles, ouvre des perspectives d'évolutions futures qui pourraient renforcer sa performance, notamment concernant l'optimisation des deux premières strates de l'échantillon (groupes 1 et 2), afin d'obtenir une représentativité encore plus fidèle des déciles 1 à 9 et du 10^{e} décile hors dernier centile. Cette recherche d'une proximité maximale avec les données observées pourrait s'appuyer sur deux approches complémentaires.

D'une part, l'exclusion des valeurs aberrantes (*outliers*) présentes dans ces strates permettrait de réduire l'impact des observations atypiques susceptibles d'introduire des biais dans l'estimation ou d'affecter la stabilité des résultats de simulation. Cette approche nécessiterait le développement de critères statistiques robustes pour identifier ces valeurs extrêmes et les corriger ou les exclure tout en préservant la richesse informative de l'échantillon.

D'autre part, l'optimisation du modèle d'échantillonnage pourrait permettre d'augmenter la taille des deux premières strates, renforçant ainsi leur puissance statistique et la précision des estimations qui en découlent. Cette expansion contrôlée de l'échantillon sur les déciles inférieurs et moyens pourrait s'accompagner d'un affinement des techniques d'équilibrage, permettant de maintenir la calibration sur les variables auxiliaires tout en améliorant la représentativité de segments spécifiques de la population fiscale. Une telle approche contribuerait à améliorer la robustesse des chiffrages de réformes affectant principalement les revenus moyens.

Annexes

Variables de calibrage : correspondance avec les cases fiscales du POTE

Variables	Cases mobilisées dans le POTE
Nombre de parts fiscales	nbp
Impôts sur les revenus soumis au barème (avant réductions et d'impôt) (€).	mnimxg
Réductions d'impôt (€)	mcirrd
Crédits d'impôt (€)	rnimjw
Montant de salaire déclaré (€).	z1aj + z1bj + z1cj + z1dj + z1aa + z1ba + z1ca + z1da + z1gh + z1hh + z1ih + z1jh + z1gb + z1hb + z1ib + z1jb + z1gf + z1hf + z1if + z1jf + z1gg + z1hg + z1ig + z1jg + z1af + z1ag + z1bf + z1bg + z1cf + z1cg + z1df + z1dg + z1tt + z1ut + z1tz + z1uz + z1wz + z1vz + z1nx + z1ox + z1ny + z1oy + z1aq + z1bq + z1dy + z1ey + z1sm + z1dn + z1ac + z1bc + z1cc + z1dc + z1ah + z1bh + z1ch + z1dh + z1pm + z1qm + z1tp + z1up + z3vj + z3vk
Montant des bénéfices agricoles déclarés y compris déficits (€)	(z5xa + z5ya + z5za) + (z5xb + z5yb + z5zb) + (z5hd + z5id + z5jd) + (z5hw + z5iw + z5jw - (z5xo + z5yo + z5zo)) + (z5hb + z5ib + z5jb) + (z5hc + z5ic + z5jc) + (z5ak + z5bk + z5ck) + (z5xt + z5xu) + (z5ea + z5ei + z5eu) - (z5hf + z5if + z5jf) + (z5hm + z5im + z5jm) - (z5qf + z5qg + z5qn + z5qo + z5qp + z5qq)
Montant des bénéfices non commerciaux déclarés y compris déficits (€)	(z5hp + z5ip + z5jp + z5hq + z5iq + z5jq + z5hv + z5iv + z5jv - (z5kz + z5lz + z5mz)) + z5qb + z5rb + z5sb + z5qc + z5rc + z5sc + z5xj + z5yj + z5zj - (z5qe + z5re + z5se) + z5ql + z5rl + z5sl + z5qm + z5rm + (z5th + z5uh + z5vh + z5ku + z5lu + z5mu + z5ky + z5ly + z5my - (z5ju + z5ld + z5md)) + z5hk + z5jk + z5lk + z5jg + z5rf + z5sf + z5xs + z5ys + z5zs - (z5jj + z5rg + z5sg) + z5sv + z5sw + z5sx - (z5ht + z5it + z5jt + z5kt + z5lt + z5mt)
Montant des bénéfices industriels et commerciaux déclarés y compris déficits (€)	(z5kn + z5ln + z5mn) + (z5ko + z5lo + z5mo + z5kp + z5lp + z5mp) + (z5kx + z5lx + z5mx) - (z5kj + z5lj + z5mj) + (z5kb + z5lb + z5mb + z5kc + z5lc + z5mc + z5df + z5ef + z5ff - (z5kf + z5lf + z5mf)) + (z5nd + z5od + z5pd + z5ng + z5og + z5pg + z5qs + z5rs + z5ss + z5nw + z5ow + z5pw + z5nj + z5oj + z5pj + z5qt + z5rt + z5st) + (z5na + z5oa + z5pa + z5ey + z5fy + z5gy + z5nm + z5om + z5pm) - (z5ny + z5oy + z5py + z5we + z5xe + z5ye) - (z5ga + z5gb + z5gc + z5gd + z5ge + z5gf + z5gg + z5gh + z5gi + z5gj) + (z5nn + z5on + z5pn + z5no + z5oo + z5po + z5np + z5op + z5pp + z5nx + z5ox + z5px - (z5iu + z5rz + z5sz)) + (z5nb + z5ob + z5pb + z5nc + z5oc + z5pc + z5ur + z5vr + z5wr) - (z5nf + z5qf + z5pf) - (z5rn + z5ro + z5rp + z5rq + z5rr + z5rw)
Montant des préretraites et allocations chômage déclarés (€)	z1ap + z1bp + z1cp + z1dp
Montant des pensions déclarées (€)	z1as + z1bs + z1cs + z1ds + z1at + z1bt + z1ct + z1dt + z1ai + z1bi + z1ci + z1di + z1az + z1bz + z1cz + z1dz + z1ao + z1bo + z1co + z1do + z1al + z1bl + z1cl + z1dl + z1am + z1bm + z1cm + z1dm
Montant des revenus fonciers déclarés (€).	z4be + z4ba + z4bd
Montant des revenus de capitaux mobiliers (RCM) déclarés (€)	z2dc + z2fu + z2go + z2ts + z2tr + z2tt + z2tq + z2tz + z2dm + z2dh + z2ch + z2ee + z2vv + z2ww + z2xx + z2yy + z2zz + z2vm + z2vn + z2vo + z2vp - z2aa - z2al - z2am - z2an - z2aq - z2ar
Montant total des plus-values non salariées et autres (€).	z5hx + z5ix + z5jx + z5he + z5ie + z5je + z5ha + z5ia + z5ja + z5kq + z5lq + z5mq + z5ke + z5le + z5me + z5ui + z5vi + z5wi + z5tf + z5uf + z5vf + z5nq + z5oq + z5pq + z5ne + z5oe + z5pe + z5hr + z5ir + z5jr + z5qd + z5rd + z5sd + z5kv + z5lv + z5mv + z5so + z5nt + z5ot + z5qa + z5ra + z5sa + z5qj + z5rj + z5sj + z5tc + z5uc + z5vc+z3vg - z3vh + z3ua + z3vt + z3pi + z3vq - z3vr + z3vc + z3se + z3we + z3wh + z3wi + z3wj + z3wn + z3xn + z3sz + z3an - z3bn + z3wm + z3xm + z3xa + z3wd + z3xd + z3ya + z3vd + z3vi + z3vf + z3vn + z3sj + z3tj + z3sk + z3vz + z3vw
Montant des rentes viagères à titre onéreux (RVTO) (€).	z1aw + z1bw + z1cw + z1dw + z1ar + z1br + z1cr + z1dr

Source : DG Trésor.

Bibliographie

CPO (2018), « Les prélèvements obligatoires sur le capital des ménages », Conseil des prélèvements obligatoires ,*Rapport*.

CPO (2024), « Conforter l'égalité des citoyens devant l'imposition des revenus », Conseil des prélèvements obligatoires, *Rapport*.

Insee (2024), « Revenu des ménages en comptabilité nationale ». *Insee Références.*

Insee (2024), « Foyers fiscaux imposés et revenus déclarés en 2022 Niveau de vie et pauvreté des adultes selon l'âge », *Insee Références*.

Insee (2025), « Niveau de vie et pauvreté selon l'activité », *Insee Références.*

Drees (2025), « Les retraités et les retraites ».

Insee (2022), « Fécondité selon le niveau de vie : une nouvelle estimation », *Insee Analyses*, n° 72.

IPP (2023), « Quels impôts les milliardaires paient-ils ? »

Deville J.-C. et Tillé Y. (2004), "Efficient balanced sampling : the cube method", *Biometrika*, vol. 91, no 4, p. 893–912.