

# L'échantillonnage des enquêtes auprès des ménages dans la source fiscale

Sébastien Faivre

Traditionnellement à l'Insee, la base de sondage des enquêtes auprès des ménages est issue du recensement de la population (RP). Toutefois, l'Insee a été amené à envisager une base de sondage alternative compte tenu des trois éléments de contexte suivants :

- la complexité du nouveau recensement en continu, qui permet certes d'avoir des informations plus fraîches, mais qui ajoute de nouvelles contraintes difficiles à contourner dans le cas d'un plan de sondage aréolaire par exemple ;
- l'émergence de nouvelles demandes (enquêtes par internet par exemple, ou enquête « individus ») ;
- la disponibilité et la qualité d'une nouvelle source : les bases fiscales.

L'Insee a ainsi décidé d'explorer les possibilités offertes par les bases fiscales qui sont mises à sa disposition par la Direction Générale des Finances Publiques (DGFIP).

## 1. Les fichiers fiscaux fournis à l'Insee par la DGFIP [1]

Chaque citoyen français est imposé :

- sur ses revenus ;
- sur le logement<sup>1</sup> qu'il occupe au 1<sup>er</sup> janvier de l'année. Cet impôt correspond à la « taxe d'habitation » (TH) ;
- sur le (ou les) logements dont il est propriétaire. Cet impôt correspond à la « taxe foncière ».

Pour ce faire, il doit déclarer les informations concernant ses revenus, son patrimoine et son lieu de vie à la DGFIP. **Une même déclaration peut concerner plusieurs personnes** car elle s'établit **au niveau du foyer fiscal**.

Le terme de « foyer fiscal » désigne l'ensemble des personnes inscrites sur une même déclaration de revenus. **Il peut y avoir plusieurs foyers fiscaux dans un seul ménage** : par exemple, un couple non marié où chacun remplit sa propre déclaration de revenus compte pour deux foyers fiscaux. En outre, **un individu peut appartenir à plusieurs foyers fiscaux**, ces derniers étant « départementaux ». Ainsi, si une personne habite et paye son impôt sur le revenu dans un département D1 et possède une résidence secondaire, pour laquelle il paye la taxe d'habitation, dans un département D2, il appartiendra à deux foyers fiscaux différents. Ce point est d'ailleurs la principale difficulté qu'il faudra surmonter pour pouvoir constituer une base de sondage d'« individus » sans double compte à partir des fichiers fiscaux.

Le foyer fiscal est donc une notion assez proche de celle de foyer (en sociologie), de foyer familial ou de ménage (au sens économique), à quelques particularités près. Ainsi :

- en cas de concubinage, les concubins sont traités comme des célibataires (sauf pour l'impôt de solidarité sur la fortune, où l'imposition est commune), avec, le cas échéant, rattachement des enfants célibataires mineurs à charge de l'un ou de l'autre ;
- les revenus des personnes liées par un PACS (Pacte Civil de Solidarité) sont imposés en commun, depuis la loi de finances promulguée en janvier 2005, sauf si le PACS est rompu dans l'année suivant sa signature ;
- un monastère ou un couvent est considéré comme un foyer fiscal unique.

**A contrario, une même personne peut faire plusieurs déclarations du même type ou aucune selon le cas.**

La DGFIP est en charge de la collecte et du contrôle de ces déclarations, puis par la suite du calcul du montant de l'impôt par foyer et de sa perception.

---

1 qu'il soit propriétaire, locataire ou occupant à titre gratuit.

Pour chaque année N, la DGFIP dispose donc de fichiers informatiques contenant des informations sur les déclarations suivantes :

- le revenu N-1 de l'ensemble des foyers fiscaux français ;
- la composition individuelle des foyers fiscaux et leur adresse au 1<sup>er</sup> janvier de l'année N ;
- la description des logements (y compris l'adresse) achevés et habitables l'année N.

Ce sont ces fichiers qui sont fournis en début d'année N+1 à l'Insee.

À noter cependant qu'un faible nombre d'individus ne sont pas toujours déclarés dans les fichiers fiscaux. Il s'agit principalement des individus qui passent du statut de mineur à celui de majeur (jeunes principalement âgés de 18 ans, mais le phénomène perdure jusqu'à 22 ans). Dans certains cas, ces jeunes ne donnant plus droit à des réductions d'impôts, ils peuvent ne pas être déclarés dans le foyer fiscal de leurs parents. Ceci n'aura en effet aucun impact sur le montant de l'imposition des parents. Et, comme par ailleurs ces jeunes majeurs ne sont pas incités à se déclarer aux impôts s'ils n'ont ni revenu ni lieu d'habitation propre pour lequel ils s'acquittent de la taxe d'habitation, ils sont absents des fichiers ci-dessus.

Par ailleurs, l'information disponible dans ces fichiers pour les enfants mineurs est minimaliste puisque seule l'année de naissance est disponible pour ces individus. De plus les enfants en garde alternée sont présents dans deux déclarations fiscales différentes. S'il est possible de les repérer indépendamment dans les deux foyers, il est impossible d'identifier les deux foyers qui concernent le même enfant.

## **2. Les qualités de la source fiscale comme base de sondage**

L'Insee vise à constituer chaque année une base de logements et une base d'individus exhaustives et sans doubles comptes à partir de la source fiscale.

Un travail méthodologique important a été entrepris pour repérer les résidences principales au sens Insee parmi les logements répertoriés par les services fiscaux, et pour identifier les doubles comptes parmi les individus connus de l'administration fiscale regroupés dans le Fichier d'Imposition des Personnes.

Ces bases peuvent être utilisées de façon naturelle comme bases de sondage pour le tirage d'enquêtes ménages, logements ou individus, dans la mesure où elles présentent, après retraitement statistique, les bonnes propriétés attendues d'une base de sondage :

- exhaustivité
- fraîcheur
- absence de doublons
- présence d'information auxiliaire (composition du ménage, âge des occupants, variables de revenus...).

## **3. L'utilisation de la source fiscale comme base de sondage pour les enquêtes auprès des ménages**

### **3.1. Le tirage de l'échantillon Emploi dans la TH 2006 [2]**

L'échantillonnage de l'enquête emploi est aréolaire : on sélectionne un certain nombre de « petites zones », appelées aires, ou grappes, et on interroge tous les logements des aires sélectionnées, par construction aussi proches que possible les uns des autres. Plus précisément, l'échantillonnage de l'enquête emploi en continu (EEC) est fondé sur un échantillon de 3200 secteurs de 120 logements chacun : chaque secteur est divisé en 6 grappes de 20 logements interrogées successivement, chaque grappe étant enquêtée pendant 6 trimestres consécutifs. L'échantillon actuel a été tiré dans la source fiscale (fichiers de la TH 2006) et est entré en service progressivement à partir de 2009.

### **3.2. Certaines enquêtes ménages spécifiques sont déjà tirées dans la source fiscale**

Le recensement de la population ne permet pas de couvrir l'ensemble des besoins en matière d'échantillonnage, dans la mesure où il permet seulement un échantillonnage de niveau logement (l'information sur l'identité des habitants du logement n'est pas disponible) et où, pour les répondants « papier » au RP on ne dispose de l'adresse que via le scan du cadre adresse de la feuille de logement. Il est de ce fait indispensable de recourir à la source fiscale dans les cas suivants :

- enquêtes ménages en face à face échantillonnant des individus (cas des enquêtes PIAAC<sup>2</sup> 2012 et Formation et Qualification Professionnelle 2015) ;
- enquêtes ménages en face à face échantillonnant des logements en utilisant l'information fiscale dans le plan de sondage pour surreprésenter certaines sous-populations (enquête Patrimoine) ou certains zonages (sur-échantillon Nouveaux Quartiers Prioritaires de l'enquête Cadre de Vie et Sécurité 2016) ;
- enquêtes multimode nécessitant un envoi de courrier massif par publipostage (enquête Vie Quotidienne et Santé, enquêtes expérimentales par internet) ou une recherche de numéros de téléphones via AnnuCapt (enquête Technologies de l'Information et de la Communication et enquête mensuelle de conjoncture Camme).

### 3.3. Un projet de refonte vers un échantillonnage unifié dans la source fiscale à compter de 2020

Le Comité de Direction de l'Insee a acté fin 2015 le lancement du projet Nautile, qui a pour objectif de mettre en place un échantillonnage unifié de l'ensemble des enquêtes ménages de l'Insee au sein de la source fiscale, à l'horizon 2020.

Les gains attendus du projet portent sur notamment sur :

- une amélioration de la précision à taille d'échantillon donnée suite au retour à une base de sondage exhaustive et à l'abandon du caractère rotatif de la base de sondage RP
- une couverture de l'ensemble des besoins en matière d'échantillonnage par une application unique
- une planification facilitée de la collecte terrain.

## Références

[1] Olivier HAAG, « La source fiscale : une base de sondage qui s'impose pour les enquêtes auprès des logements et des individus », Symposium de Statistique Canada 2013.

[2] Vincent LOONIS, « La construction du nouvel échantillon de l'Enquête Emploi en Continu à partir des fichiers de la Taxe d'Habitation », *actes des journées de méthodologie statistique 23, 24 et 25 mars 2009*.



Département des méthodes statistiques  
Version n°1, diffusée le 10 octobre 2017

---

<sup>2</sup> Program for the International Assessment of Adult Competencies, enquête sur les compétences des adultes en littéracie et numératie menée dans le cadre de l'OCDE.