



**Conseil national
de l'information statistique**

Montrouge, le 24 novembre 2022
N° 157 / H030

**AVIS DU CNIS SUR DES DEMANDES D'ACCÈS À DES DONNÉES
AU TITRE DE L'ARTICLE 7BIS DE LA LOI DU 7 JUIN 1951 MODIFIÉE**

La commission « Territoires » a examiné la demande suivante :

Formulée par le département des études économiques de l'Institut national de la statistique et des études économiques (Insee)

Auprès du département des études et statistiques fiscales (DESF) de la direction générale des finances publiques (DGFIP)

Concernant les données fiscales du formulaire n°2044 EB.

La commission émet un **avis favorable** à cette demande d'accès.

**Le président de la commission
Pierre-Philippe Combes**

ANNEXE

Demande d'accès au titre de l'article 7bis de la loi n° 51-7111 du 7 Juin 1951 modifiée à des données concernant le statut du bailleur privé et l'engagement de location détenues par le département des études et statistiques fiscales (DESF).

1. Service demandeur

Direction Générale de l'Insee – Département des Études Économiques (DEE) et Direction de la méthodologie et de la coordination statistique et internationale (DMCSI)

2. Organisme détenteur des données demandées

Le département des études et statistiques fiscales (DESF) de la Direction générale des finances publiques (DGFIP) assure les maîtrises d'ouvrage et d'œuvre du fichier dit RF_2044EB issu du formulaire n° 2044 EB.

3. Nature des données demandées

Les informations demandées correspondent aux différents champs présents sur la déclaration 2044 EB. Le formulaire n°2044 EB doit être rempli par tout propriétaire, personne physique ou société non soumise à l'impôt sur les sociétés, qui souhaite bénéficier d'une déduction au titre de l'amortissement (Robien recentré ou Borloo neuf) ou des réductions d'impôt sur le revenu (Scellier, Scellier Outre-mer, Duflot/Pinel, Denormandie).

Ces informations comprennent : les identifiants fiscaux, les noms et prénoms du déclarant, les noms et prénoms des locataires, ainsi que l'adresse du logement, la date d'acquisition et le prix payé. Toutes ces informations peuvent être plus ou moins bien renseignées par les déclarants (ce sont des champs textuels). La DGFIP a signalé que ces données ne sont pas exhaustives. En effet, ces données concernent uniquement les foyers fiscaux soumis à l'impôt sur le revenu qui ont télé-déclaré leurs revenus fonciers.

Pour information, une demande d'accès à ces données avait été réalisée par le Service des données et études statistiques (SDES), ministère de la Transition écologique et solidaire. La demande avait reçu un avis positif lors de la commission « Environnement et développement durable » du 2 avril 2020.

4. Objectifs généraux des traitements prévus avec les données demandées

L'objectif général est la réalisation d'études sur les impacts des dispositifs d'aide à l'investissement locatif. Les impacts potentiels de tels dispositifs sont multiples : augmentation de la construction de logements, modification des caractéristiques des logements construits, modification des rendements locatifs, inflation des prix du foncier, inflation des prix immobiliers dans le neuf mais potentiellement également dans l'ancien par effet d'entraînement et modification de la structure sociale des quartiers. L'objectif principal serait de quantifier ces effets et notamment d'évaluer l'impact sur la structure sociale des quartiers.

A ce stade, les travaux visent à répondre à trois grandes questions :

- qui sont les ménages qui bénéficient des dispositifs d'aide à l'investissement locatif ?
- quels sont les logements qui font l'objet des dispositifs d'aide à l'investissement locatif ?
- quels sont les effets de ces dispositifs sur le fonctionnement des marchés immobiliers et locatifs locaux ?

5. Nature des travaux statistiques prévus

Sources disponibles à l'Insee

Au moment de la présente demande d'avis, l'Insee dispose de trois sources sur les revenus fonciers : les données de l'impôt sur le revenu (déclarations 2042), et les déclarations annexes de revenus fonciers

(déclarations 2044 et 2044 S, à bien distinguer des déclarations 2044 EB). Ces données ne permettent pas de repérer avec certitude les logements bénéficiant des dispositifs considérés. L'Insee dispose par ailleurs de trois sources sur le logement : le fichier de la taxe d'habitation (fichier PLFC), les données cadastrales (fichiers Majic) et les données sur les transactions immobilières (données DVF).

Limites des sources disponibles à l'Insee

Les données de l'impôt sur le revenu (déclarations 2042) permettent de connaître les revenus fonciers totaux des foyers fiscaux et de savoir quels foyers fiscaux ont bénéficié d'un dispositif de soutien à l'investissement locatif, et l'année de réalisation de l'investissement éligible. En revanche, ces données ne contiennent aucune information sur la localisation du logement concerné, sur sa nature (maison ou appartement, logement neuf ou logement ancien rénové); sur le mode d'acquisition (en nom propre, via une société civile immobilière ou via une SCPI), sur les locataires qui l'occupent et sur le niveau de loyer pratiqué.

Les données sur les revenus fonciers (déclarations 2044 et 2044 S) permettent de connaître les revenus fonciers des foyers fiscaux de façon plus détaillée (immeuble par immeuble). Ces données contiennent l'adresse et les revenus fonciers tirés de chaque immeuble, ainsi que les noms et prénoms des locataires. En revanche, ces données n'indiquent pas quels logements font l'objet d'un dispositif fiscal.

De premiers travaux ont été menés à l'Insee courant 2022, afin de déterminer s'il était possible de repérer dans le cadastre les logements éligibles à un dispositif à partir des seules informations contenues dans les déclarations 2042 (nom des déclarants, existence d'un investissement, et année de l'investissement). Ces tests ont donné des résultats décevants pour de multiples raisons. En particulier, il arrive fréquemment que des foyers fiscaux bénéficiant d'un dispositif aient acquis plusieurs logements la même année, auquel cas on ne peut pas déterminer avec certitude quel est le logement qui bénéficie du dispositif. Il arrive également que les foyers fiscaux déclarent une année d'investissement erronée. Dans l'ensemble, ces tests aboutissent à des données de qualité trop basse pour être exploitables.

Par rapport aux sources disponibles à l'Insee, les données 2044 EB apportent des informations supplémentaires utiles (en particulier l'adresse exacte des logements et la date exacte d'acquisition) qui devraient permettre de repérer beaucoup plus précisément les logements concernés.

Description des traitements envisagés

Le principal traitement consistera à appairer les données 2044 EB avec les données du cadastre (données Majic). Il s'agira de repérer précisément les logements bénéficiant de dispositifs fiscaux (par exemple : tel logement a été acquis en 2017 par tel foyer fiscal dans le cadre de la réduction d'impôt Pinel). Cet appariement mobiliserait essentiellement les noms et prénoms des propriétaires, l'adresse du logement, la date d'acquisition et le nom de la SCI ou de la SPCI le cas échéant. Il n'est en revanche pas certain que les identifiants fiscaux puissent être utilisés, car les données 2044 EB et les données du cadastre ne contiennent *a priori* pas d'identifiant fiscal commun. Les modalités exactes de cet appariement ne pourront toutefois être déterminées qu'après expertise des données 2044 EB.

Une fois cet appariement réalisé, plusieurs tables seront construites. Il est prévu de construire les tables suivantes :

- une table de foyers fiscaux, avec des informations sur les foyers fiscaux (composition, niveau de revenu), les logements acquis par ces foyers, et les dispositifs fiscaux dont ils bénéficient. Cette table sera construite en utilisant l'identifiant de foyer fiscal. Cette table permettrait de répondre à la première question.
- une table de logements, avec des informations sur leurs caractéristiques et leur localisation, la date de construction et les dispositifs dont ils font l'objet. Cette table sera immédiatement dérivée de l'appariement entre le cadastre et les données 2044 EB. Cette table permettrait de répondre à la deuxième question, et partiellement à la troisième question (effet des dispositifs sur la construction de logements neufs).
- une table de transactions immobilières, avec des informations sur les logements vendus, la date de transaction et le montant payé. Cette table sera construite en croisant la table de logements et les données DVF, à l'aide de l'identifiant cadastral de local. Cette table permettrait de répondre partiellement à la troisième question (effet des dispositifs sur les prix immobiliers).

Enfin, ces trois tables pourront être appariées avec d'autres sources disponibles à l'Insee (notamment les

fichiers Fidéli et Filosofi) afin d'étudier les effets des dispositifs sur la composition sociale des quartiers (exemple : le taux d'accession à la propriété est-il plus faible dans les quartiers comprenant un nombre élevé d'investissements Pinel ?). Ces appartements se feront par zone de façon agrégée (commune, IRIS, éventuellement carreau), et ne mobiliseront donc pas de données individuelles.

6. Place dans le dispositif statistique existant sur le même sujet

Les données issues de la déclaration 2044 EB rendront possibles les études d'évaluation des dispositifs d'aide à l'investissement locatif concernés par cette déclaration et sont indispensables pour pouvoir identifier les adresses des biens locatifs concernés.

7. Périodicité de la transmission

Seront fournis les millésimes existants depuis 2006 dès signature de la convention. Le DESF procédera à une transmission annuelle du fichier Revenus_fonciers_2044 relatif à l'année N-1, au courant du mois d'octobre N.

8. Diffusion des résultats

L'Insee sera seul détenteur des données individuelles.

Les données qui seront diffusées par l'Insee le seront donc sous forme agrégée conformément aux dispositions prévues par les textes relatifs au secret professionnel, au secret fiscal ainsi qu'au secret en matière de statistiques, notamment l'article L.103 du LPF et la loi du 7 juin 1951.

Ces règles sont les suivantes :

- la donnée agrégée diffusée doit concerner au moins trois unités statistiques, ce seuil étant porté à onze si l'unité statistique est une personne physique. Il est rappelé qu'en matière de fiscalité professionnelle, une entreprise individuelle correspond à une personne physique ;
- la donnée agrégée ne doit comporter aucune unité statistique représentant à elle seule plus de 85 % du montant agrégé.

Les études d'évaluation des politiques publiques seront diffusées au grand public sur le site de l'Insee et éventuellement dans des revues académiques.

Le service producteur cédant a été informé en amont de la demande.