

Quel est l'apport de la taxe d'habitation à l'extrapolation du nombre de résidences principales au *Recensement de la population* ?

Hugues Lermechin *

Depuis la rénovation du recensement au début des années 2000, obligation est faite à l'Insee de publier chaque année les populations légales de toutes les communes. Pour les communes de moins de 10 000 habitants, qui sont recensées exhaustivement tous les cinq ans, la difficulté réside dans l'actualisation des populations durant la période intercensitaire. L'utilisation des données de la taxe d'habitation (TH) répond pour partie à ce défi et constitue une innovation notable dans le processus de calcul des populations. La TH permet d'estimer le nombre de résidences principales, approchant ainsi la dynamique de population. Elle contribue à calculer chaque année la population de 40 % des communes de moins de 10 000 habitants, occasionnant par là-même l'estimation d'environ 20 % de la population française.

La TH apporte-t-elle une réelle plus-value à l'estimation des populations communales et de la population totale concernées par son utilisation ? Si le nombre de résidences principales est surestimé, l'erreur occasionnée est faible. En extrapolant le nombre de résidences principales avec la TH durant cinq années successives, l'erreur est ainsi inférieure à 5 % pour près de trois quart des communes et le nombre total de résidences principales n'est surestimé que de 1 %. En réalité, l'extrapolation n'excédant jamais deux années successives, l'erreur est selon toute vraisemblance moindre les années où la TH participe au calcul des populations. Un effet taille apparaît néanmoins : les plus grandes des petites communes sont plus souvent surestimées.

Rappel :

Les jugements et opinions exprimés par les auteurs n'engagent qu'eux mêmes, et non les institutions auxquelles ils appartiennent, ni a fortiori l'Insee.

Par ailleurs, les populations légales de sept communes sur dix sont mieux estimées avec la TH, qu'elles ne le seraient par prolongement linéaire des derniers recensements. En outre, la TH restitue finement les ruptures de tendance observées dans l'évolution du parc de logements des communes. C'est à la fois un instrument précis et souple car il s'adapte aux différentes dynamiques d'évolution qui peuvent exister.

Codes JEL : J19, C81.

Mots clés : taxe d'habitation, recensement, estimation de population, estimation de résidences principales, démographie, données administratives, enquête.

* Insee.

Remerciements : l'auteur remercie Nathalie Caron, Gwennaëlle Brilhault, Sébastien Hallépée, Anne Reffet-Rochas, Julien Talbot et Lise Lefebvre pour leurs commentaires et leur aide. Les propos tenus dans l'article n'engagent cependant que lui.

Le recensement rénové et le rôle de la taxe d'habitation

Avec le recensement rénové, l'Insee met en œuvre de nouvelles méthodes de calcul des populations

En France, le dernier recensement général exhaustif de la population date de 1999. Il se tenait alors irrégulièrement : tous les six à neuf ans depuis la fin de la deuxième guerre mondiale. Pour des motifs budgétaires, le recensement de 1999 eut lieu deux ans après la date initialement prévue. À l'issue de celui-ci, un processus de rénovation du recensement a été enclenché, dans le but d'améliorer la qualité du service rendu au public tout en contenant le coût d'une telle opération. Ce processus a abouti en 2004 avec les premières *Enquêtes annuelles de recensement (EAR)*, qui se sont substituées au recensement exhaustif. Un important dispositif juridique définit et encadre les nouvelles orientations : l'Institut national de la statistique et des études économiques (Insee) organise et contrôle la collecte des informations, alors que les communes sont désormais tenues de préparer et de réaliser les enquêtes de recensement (Godinot, 2005).

Dans le recensement rénové, on distingue désormais les grandes communes (10 000 habitants ou plus) des petites communes (moins de 10 000 habitants). Les grandes communes sont recensées chaque année, mais seule une partie

des adresses est enquêtée (environ 8 % d'entre elles chaque année). Dans les petites communes, le recensement exhaustif perdure et se tient tous les cinq ans. Il existe cinq groupes équilibrés de petites communes, qu'on appelle « Groupes de rotation » (GR) et qu'on recense à tour de rôle année après année. Par exemple, en 2004, les communes du premier groupe de rotation ont été recensées. En 2005, ce fut au tour du deuxième groupe de rotation et ainsi de suite... À nouveau, en 2009, on a recensé les communes du premier groupe de rotation et, en 2010, celles du deuxième...

L'obligation faite à l'Insee par la loi de calculer chaque année les populations légales de toutes les communes de France¹ constitue une amélioration notable du recensement rénové. Rien de tel n'était proposé jusque-là. Cela a nécessité de mettre en œuvre de nouvelles méthodes de calcul de la population. Dans les grandes communes, chaque *EAR* renouvelle l'information et contribue à actualiser l'estimation de la population au 1^{er} janvier d'une année (cf. encadré 1). L'Insee publie les populations légales de manière décalée : par exemple, au 31 décembre 2013, il livre les populations communales en date du 1^{er} janvier 2011.

1. Exceptions faites des communes de Mayotte, de la Nouvelle-Calédonie, de la Polynésie française et de Wallis-et-Futuna, où des recensements généraux exhaustifs sont organisés tous les cinq ans.

Encadré 1

MÉTHODE DE CALCUL DE LA POPULATION DES MÉNAGES ORDINAIRES DES GRANDES COMMUNES

Dans les grandes communes (10 000 habitants ou plus), une nouvelle méthode de recensement par sondage est mise en œuvre. Le Répertoire d'immeubles localisés (RIL), qui répertorie l'ensemble des adresses de la commune, constitue la base de sondage. L'Insee le met à jour annuellement grâce aux informations données par les communes, différentes sources de données administratives (parmi lesquelles figure la TH) et les constatations faites durant les opérations de collecte.

Chaque année, 8 % des adresses sont tirées au sort et recensées. Pour le calcul de la population des ménages ordinaires (appelée aussi population des résidences principales) d'une commune, l'Insee utilise cinq *EAR* successives, soit au total l'information recueillie

auprès d'environ 40 % des adresses. L'année de calcul de la population est l'année médiane des cinq campagnes utilisées. Par exemple, pour calculer les populations légales de 2010 publiées au 31 décembre 2012, l'Insee a eu besoin des *EAR* 2008 à 2012. La campagne de collecte de 2012 est donc indispensable au calcul des populations légales de 2010 dans le recensement rénové. Ceci explique pourquoi l'Insee publie les populations légales avec un décalage de deux années.

Les cinq *EAR* permettent d'obtenir le nombre moyen de personnes par logement, dont le produit avec le nombre de logements répertoriés dans le RIL conduit à l'estimation de la population des ménages ordinaires (ou population en résidences principales) de la commune.

Dans les petites communes, ce décalage de deux années conduit à mettre en œuvre des méthodes de calcul différentes selon le groupe de rotation. Pour deux groupes de rotation sur cinq, le recensement le plus récent est strictement antérieur à la date de calcul des populations légales. L'Insee utilise alors une source externe d'informations : les données de la taxe d'habitation (TH) produites par la Direction générale des Finances publiques (DGFIP). Ces données permettent d'approcher la dynamique du parc de logements et par extension la dynamique de la

population dans la commune (cf. encadré 2). On dit de ces communes que leurs populations sont estimées par « extrapolation » ou « extrapolation TH ». Pour simplifier le propos, on parlera aussi dans le reste de l'étude de « communes en extrapolation ». Dans cet article, nous nous intéresserons uniquement au champ des communes en extrapolation, sauf mention contraire.

L'utilisation d'une source externe dans le calcul des populations légales a fait l'objet de nombreuses discussions lors de la rénovation

Encadré 2

MÉTHODE DE CALCUL DE LA POPULATION DES MÉNAGES ORDINAIRES DANS LES PETITES COMMUNES

Les petites communes sont réparties dans cinq groupes de rotation. Chaque groupe de rotation comprend un peu plus de 7 000 petites communes.

Les 7 000 communes du groupe de rotation sont recensées la même année, par cycle de cinq ans (cf. tableau a).

Tableau a
Année de recensement des différents groupes de rotation de petites communes

Année de recensement	GR1	GR2	GR3	GR4	GR5
2004	x				
2005		x			
2006			x		
2007				x	
2008					x
2009	x				
2010		x			
2011			x		
2012				x	
2013					x

Lecture : les communes du 1^{er} groupe de rotation ont été recensées en 2004 et 2009.

L'Insee publie depuis 2008 les populations légales issues du recensement rénové. Les premières à avoir été calculées se rapportaient au 1^{er} janvier 2006 et ont été publiées fin 2008. Le décalage entre l'année de diffusion et la date de calcul des populations légales permet de disposer pour deux groupes de rotation sur cinq d'un recensement postérieur à la date de calcul des populations légales. On utilise la méthode dite d'interpolation linéaire pour estimer la population des ménages ordinaires de ces communes. Cette méthode consiste à positionner l'estimation sur la droite qui relie le dernier chiffre de population publié au point de recensement. Elle ne nécessite pas de source d'information externe. Dans le cadre des populations légales au 1^{er} janvier 2006, elle a permis d'estimer la population des communes des 4^e et 5^e groupes de rotation, respectivement recensées en 2007 et 2008.

Pour un troisième groupe, c'est le recensement qui fait foi puisque l'année où il s'est tenu correspond à l'année de calcul des populations légales. Ainsi, les populations légales au 1^{er} janvier 2006 des communes du 3^e groupe de rotation sont issues du recensement.

Pour les deux derniers groupes de rotation, le recensement le plus récent est antérieur à l'année de calcul des populations, ce qui conduit à utiliser une source d'information externe pour mieux approcher la dynamique post-recensement des communes concernées. On a recours aux données de la taxe d'habitation. Cette méthode est appelée « méthode d'extrapolation » ou « méthode d'extrapolation TH ». Dans le cadre des populations légales au 1^{er} janvier 2006, la méthode d'extrapolation a permis d'estimer la population des communes des 1^{er} et 2^e groupes de rotation, respectivement recensées en 2004 et 2005. →

du recensement. Plusieurs pistes, comme les fichiers d'Électricité de France (EDF) et de la Sécurité sociale ont été explorées. Le choix s'est arrêté sur les données de la TH pour plusieurs raisons :

- Les concepts de logement au sens du *Recensement de la population (RP)* et au sens fiscal sont proches ;
- L'Insee manie les données de la TH depuis longtemps ; il existe donc un savoir-faire à l'Institut dans l'exploitation de cette information ;
- Depuis longtemps, la législation de la taxe d'habitation n'a pas évolué dans un sens qui remettrait en cause la manière dont les locaux sont enregistrés par les services fiscaux ; en conséquence, la structure de la base de données est restée stable.

L'utilisation de la TH constitue une innovation mais génère de l'approximation pour le calcul des populations des petites communes en extrapolation

La mise en place d'une base de sondage d'adresses (le RIL) dans les grandes communes a constitué sans nul doute l'élément méthodologique le plus discuté et le plus mis en avant dans la rénovation du recensement. L'utilisation directe d'une source de données externes (en l'occurrence la TH) dans le calcul

des populations légales des petites communes est apparue beaucoup plus anecdotique, alors qu'elle représente pourtant une innovation importante elle aussi. Pour la première fois, des données non produites par l'Insee servent à dénombrer la population française². De plus, la TH contribue chaque année à l'estimation de la population de près de 40 %³ des communes, soit environ 20 %⁴ de la population, puisque les petites communes regroupent sensiblement la moitié de la population française et que deux groupes de rotation sur cinq font l'objet chaque année d'une estimation de leur population reposant sur l'utilisation de la TH. L'enjeu lié à la TH est donc double : il faut approcher correctement la population de chaque petite commune en extrapolation tout en estimant le plus précisément possible l'ensemble de la population habitant dans ces petites communes.

Toutefois, l'utilisation d'une source externe dans le cadre d'un projet, et en particulier d'une source administrative à des fins statistiques, génère nécessairement de l'approximation (Desrosières, 2004).

2. Il faut mentionner l'existence de nombreux travaux antérieurs à la rénovation du recensement et ayant pour objet l'estimation de populations avec des sources de données administratives dont la TH : (Descours, 1991), (Labat et Decaudin, 1996) ou (Bergouignan, 1999)... Toutefois, leurs résultats n'avaient aucune portée légale contrairement à ceux du recensement rénové.
 3. Entre 38,8 % et 38,9 % pour les millésimes de populations légales 2006 à 2010.
 4. Entre 20,1 % et 20,3 % pour les millésimes de populations légales 2006 à 2010.

Encadré 2 (suite)

Le tableau b indique la méthode de calcul de la population des ménages ordinaires employée pour les petites communes de chaque groupe de rotation par millésime de populations légales.

Tableau b
Méthode utilisée pour le calcul de la population des ménages ordinaires par groupe de rotation et millésime de populations légales

Année de populations légales	GR1	GR2	GR3	GR4	GR5
2006	Extrapolation TH	Extrapolation TH	<i>Recensement</i>	Interpolation	Interpolation
2007	Interpolation	Extrapolation TH	Extrapolation TH	<i>Recensement</i>	Interpolation
2008	Interpolation	Interpolation	Extrapolation TH	Extrapolation TH	<i>Recensement</i>
2009	<i>Recensement</i>	Interpolation	Interpolation	Extrapolation TH	Extrapolation TH
2010	Extrapolation TH	<i>Recensement</i>	Interpolation	Interpolation	Extrapolation TH
2011	Extrapolation TH	Extrapolation TH	<i>Recensement</i>	Interpolation	Interpolation

Lecture : la méthode d'interpolation a permis d'estimer la population des ménages ordinaires au 1^{er} janvier 2009 des 2^e et 3^e groupes de rotation.

Malgré la proximité des concepts en matière de logement entre la DGFIP et l'Insee, la TH ne fait pas exception à cette règle. Les données de la taxe d'habitation d'un millésime offrent la photographie de l'état du bâti dans les communes au 1^{er} janvier de l'année considérée. Elles sont collectées par la DGFIP dans le but de répondre à des problématiques d'ordre fiscal et budgétaire. En d'autres termes, elles décrivent la réalité du bâti sous ces angles et ne sont pas dédiées au dénombrement de la population française.

De son côté, l'Insee s'attelle avec le *Recensement de la population* à décrire notamment la réalité de l'occupation des logements afin de dénombrer précisément la population. Si elles peuvent se confondre localement ou ponctuellement, ces deux réalités ne se superposent pas parfaitement.

Les missions de la DGFIP ne sont pas les mêmes que celles de l'Insee et les bases de données reflètent logiquement cette divergence.

Il y a donc lieu de traiter les données de la TH pour les adapter à la problématique du dénombrement de la population

Le traitement central et primordial appliqué par l'Insee aux données TH est un filtre des logements visant à s'approcher au mieux du concept de logement utilisé dans le *Recensement de la population* (cf. annexe 1). Dans la base TH, on trouve en effet l'ensemble du bâti susceptible d'être assujéti à la TH. Or, ce bâti comprend maisons et appartements, mais aussi garages, box, piscines... S'il est facile d'exclure de la liste des logements les locaux dont la nature diffère d'un local d'habitation, il est plus difficile d'affiner le filtre pour les locaux d'habitation eux-mêmes, car certains appartements et maisons définis ainsi au sens fiscal ne sont pas des logements au sens du *RP*. La distinction se fait à l'aide de variables de gestion de la DGFIP qui ne décrivent donc pas la réalité du terrain mais une réalité fiscale (cf. encadré 3).

L'utilisation des variables de gestion dans la définition du logement constitue donc un expédient. Mais de la qualité de cet expédient dépend l'apport de la TH dans le calcul des populations légales. « *Le comptage des pauvres implique au préalable leur codage, c'est-à-dire leur attribution à une classe d'équivalence* » (Desrosières, 2004). Avant de dénombrer, il faut en effet caractériser. Si l'on choisit un déterminant

économique pour définir les pauvres (il aurait pu être sociologique ou culturel par exemple), il faut convenir d'un ou plusieurs critères qui distinguent les pauvres des autres : quel est le seuil de pauvreté ? quels revenus considère-t-on (revenu net, imposable ou disponible ; quelles allocations prend-on en compte...) ? comment intègre-t-on le nombre de personnes composant le foyer dans la fixation du seuil de pauvreté ? Le nombre de pauvres dépend donc des déterminants et des critères retenus pour caractériser la pauvreté. Si on reprend cette idée pour l'usage de la TH, on peut affirmer que le comptage des logements au sens du *RP* implique au préalable leur codage dans la TH. Il s'agit d'identifier parmi l'ensemble des locaux TH ceux qui sont des logements au sens du *RP*. Ici, le déterminant du codage est fiscal et les critères ont trait à des variables de gestion de la TH : la nature du local, le mode de taxation, le type d'occupation... Il faut construire le filtre des logements le plus approprié avec les informations fiscales dont on dispose. Ainsi, qu'il s'agisse de dénombrer des logements ou des pauvres, l'attribution de la classe d'équivalence reste une étape essentielle à tout dénombrement.

C'est précisément la qualité du codage des logements à partir de la source TH appliqué aux petites communes en extrapolation qui est expertisée dans la suite de l'étude. Le traitement central que constitue le filtre des logements est-il pertinent ? Quelle approximation engendret-il à la fois dans l'estimation des populations communales et dans celle de la population nationale ? On cherche à mesurer l'adéquation de la TH à notre processus.

La TH est une composante de l'extrapolation parmi d'autres

Pour juger de la qualité de la TH au processus d'estimation, il importe de situer précisément son apport au calcul des populations. Or, la TH n'est qu'une composante du calcul des populations des petites communes en extrapolation. La méthode d'extrapolation TH du calcul de la population des ménages ordinaires (cf. encadré 4) combine en fait deux éléments de variation :

- Un coefficient d'évolution du nombre de logements ou « coefficient TH », calculé à l'aide des données de la taxe d'habitation ;
- Un coefficient d'évolution du nombre moyen de personnes par logement ou « coefficient de

décohabitation », calculé avec les *EAR* par projection linéaire de l'évolution observée entre les deux derniers recensements.

La TH ne participe pas du tout à l'estimation du coefficient de décohabitation.

De plus, bien que la population des ménages ordinaires constitue l'essentiel de la population municipale, d'autres catégories de populations sont incluses dans la population municipale et sont estimées sans que la TH n'intervienne (cf. annexe 2). C'est le cas de la population des communautés et de la population des habitations mobiles terrestres et fluviales et des

personnes sans-abri (HMSA). En conséquence, la TH n'est pas le seul élément qui participe à l'estimation de la population des petites communes en extrapolation.

La TH sert à estimer le nombre de résidences principales

Mesurer l'apport de la TH dans le calcul des populations des ménages ordinaires nécessite de s'affranchir de l'approximation causée par le coefficient de décohabitation. Pour expliquer simplement la méthode d'extrapolation, le coefficient TH sert à estimer un nombre de

Encadré 3

PRÉSENTATION DES DONNÉES TH ET DU FILTRE DES LOGEMENTS

La base de données TH comporte l'ensemble du bâti susceptible d'être assujéti à la taxe. Il comprend les locaux affectés à l'habitation mais aussi les dépendances immédiates, comme les parkings, les box, les remises, les caves, les piscines... Les locaux sont présents dans la base de données, qu'ils soient soumis ou non à la taxe d'habitation.

L'arrêté du 26 juin 2003 portant création d'un traitement automatisé d'informations individuelles issues des fichiers de la taxe d'habitation encadre l'utilisation de la TH au Recensement de la Population (Godinot, 2005). Son traitement poursuit trois objectifs :

- La mise en œuvre des processus d'estimation de la population des communes (pour 40 % des petites communes de moins de 10 000 habitants chaque année) ;
- La mise à jour du Répertoire d'immeubles localisés (RIL), base de sondage des grandes communes ;
- Le contrôle d'exhaustivité des enquêtes de recensement.

Cet arrêté consigne également la liste des variables TH transmises à usage du recensement de la population, qui donnent des informations relatives aux contribuables et aux locaux (Clanché, 2010). L'Insee utilise une partie de ces variables pour déterminer si le local taxé est un logement au sens du recensement de la population, c'est ce qu'on appelle le « filtre des logements ».

En définitive, un local TH est considéré comme un logement au sens du RP s'il satisfait à trois conditions :

- La condition de nature du local : le local doit être habitable (il s'agit essentiellement de maisons et d'appartements) ;
- La condition de taxation : le local ne doit pas être considéré comme une dépendance au sens fiscal ;
- La condition de local principal : le local doit être le local principal d'un ensemble de locaux occupés par le même foyer fiscal à la même adresse.

Ces conditions sont examinées à l'aune de variables de gestion renseignées par les services fiscaux. Le filtre des logements conduit à considérer qu'environ 25 % des locaux TH ne sont pas des logements au sens du RP. La condition de nature du local contribue à l'exclusion de 90 % d'entre eux, que les deux autres conditions soient satisfaites ou non. Cela signifie que 10 % des locaux exclus par le filtre des logements sont des maisons ou des appartements au sens fiscal : ce peut être des locaux à usage professionnel ou faisant partie d'un ensemble habitable plus grand. Dans le cas de deux appartements réunis en un seul, la base TH conserve deux observations mais le filtre des logements permet théoriquement d'en exclure une. On considère ainsi pour l'ensemble habitable correspondant un seul logement au sens de la TH comme cela est fait au recensement.

Le filtre appliqué reste cependant approximatif : on inclut comme on exclut à tort des locaux parmi les logements. En toute logique, les logements de type communauté (cf. définition en annexe 1) doivent être exclus des logements, car la TH sert à l'extrapolation de la population des ménages ordinaires. Cependant, une parfaite prise en compte de ces logements s'avère impossible. En conséquence, certaines communautés sont présentes dans l'ensemble de logements TH même si la dérive la plus grave réside dans une prise en compte dissymétrique d'une communauté au cours du temps. Cela survient par exemple, quand à une adresse de communauté on comptabilise 80 logements une année et 0 la suivante alors que la communauté n'a pas connu de modification. Cela peut alors avoir une conséquence non négligeable sur l'évolution du nombre de logements TH calculé par l'Insee. Dans l'étude, tous les dysfonctionnements du filtre seront partie intégrante de l'approximation calculée, bien que l'Insee corrige en réalité les erreurs flagrantes, qui induisent une évolution de logements erronée.

résidences principales dans la commune tandis que le coefficient de décohabitation permet d'estimer le nombre moyen de personnes par résidence principale, la combinaison des deux coefficients concourant à l'estimation de population. La TH doit donc être jugée sur sa qualité à estimer le nombre de résidences principales.

Des imprécisions subsistent entre la source TH et la source RP quant au caractère d'occupation du logement. Un contribuable ne peut avoir qu'une seule résidence principale. Si des contrôles sont possibles au niveau fiscal, il n'en est pas de même par le Recensement de la population. Un individu peut donc y déclarer plusieurs résidences principales. En outre, des stratégies d'optimisation fiscale peuvent conduire à des différences déclaratives. Enfin, les logements vacants au sens du RP peuvent être considérés comme des résidences principales ou secondaires au sens fiscal, si par exemple leur propriétaire habite en maison de retraite. Un flou subsiste donc entre les définitions DGFIP et Insee d'une résidence principale. Pour ces raisons, le coefficient TH est calculé sur l'ensemble des logements, même

s'il traduit l'évolution du nombre de résidences principales. Cet ensemble comprend les résidences principales, secondaires et les logements vacants. Ce choix avait également pour but d'empêcher d'éventuelles contestations de communes relatives au caractère d'occupation des logements. Prendre en compte l'ensemble des logements pour le coefficient TH impulse une dynamique plus générale et moins heurtée du parc de logements dans la commune.

Un autre élément d'approximation entre les sources TH et RP concerne la date de recueil de l'information. Les populations légales sont calculées au 1^{er} janvier d'une année. La TH dresse un état fiscal de l'occupation du bâti à la même date. Cependant, le recensement de la population se déroule traditionnellement du 3^e jeudi de janvier à mi-février. L'Insee assimile l'état de l'occupation du logement à la date du recensement à celui du 1^{er} janvier de la même année. Il est impossible d'opérer un recensement en une journée et encore moins au 1^{er} janvier d'une année. Un foyer recensé qui déménage début janvier déclare donc en toute logique des

Encadré 4

FORMULE DE CALCUL DE LA POPULATION DES MÉNAGES ORDINAIRES DES COMMUNES EN EXTRAPOLATION

Lors d'une extrapolation, la population des ménages ordinaires est calculée selon la formule suivante :

$$PopMénage_N^{ExtrapolationTH} = PopMénage_{N-1} \times \underbrace{\frac{NbLog_N^{TH}}{NbLog_{N-1}^{TH}}}_{\text{Coefficient TH}} \times \underbrace{\left(\frac{NbPersLog_{DernièreCollecte}^{EAR}}{NbPersLog_{AvantDernièreCollecte}^{EAR}} \right)^{\frac{1}{5}}}_{\text{Coefficient de décohabitation}}$$

où

- $PopMénage_N^{ExtrapolationTH}$ est la population des ménages ordinaires calculée par la méthode d'extrapolation TH au 1^{er} janvier de l'année N,
- $PopMénage_{N-1}$ est la population des ménages ordinaires calculée au 1^{er} janvier de l'année N-1,
- $NbLog_N^{TH}$ est le nombre de logements filtrés dans la TH au 1^{er} janvier de l'année N,
- $NbLog_{N-1}^{TH}$ est le nombre de logements filtrés dans la TH au 1^{er} janvier de l'année N-1,

- $NbPersLog_{DernièreCollecte}^{EAR}$ est le nombre moyen de personnes par résidence principale calculé lors de la dernière enquête de recensement,
- $NbPersLog_{AvantDernièreCollecte}^{EAR}$ est le nombre moyen de personnes par résidence principale calculé lors de l'avant-dernière enquête de recensement.

L'intervalle intercensitaire étant de cinq ans, on porte à la puissance 1/5^e le rapport des nombres de personnes par logement des deux derniers recensements afin d'obtenir un coefficient de décohabitation annualisé.

informations différentes aux services fiscaux et à l'Insee sur la localisation et les caractéristiques d'occupation de son logement.

En conséquence, l'utilisation de la TH dans le cadre du *RP* est entachée, par construction, d'approximations impossibles à amender. Si la TH sert à estimer le nombre de résidences principales, les différences de concept entre sources sur le caractère d'occupation d'un local justifient que l'on utilise la série de l'ensemble des logements pour le calcul du coefficient TH, d'autant plus que le décalage de date entre la déclaration fiscale et la période de recensement peut générer des doubles comptes ou au contraire des oublis.

L'apport de la TH à l'estimation des résidences principales : méthode et mesures

Comment juger de la valeur de l'apport de la TH ?

L'évaluation du recours à la TH sera faite à travers le nombre de résidences principales estimé. Or, il faut pouvoir confronter cette estimation à un point de référence, en l'occurrence un recensement. Cela revient à comparer, une année de recensement donnée, les résultats du recensement aux résultats d'une extrapolation TH. La valeur de l'estimation faite au moyen de la TH est donc appréhendée sur un intervalle de cinq années, c'est-à-dire le nombre d'années d'extrapolation nécessaires à une telle comparaison.

Pour le 1^{er} groupe de rotation de petites communes recensées en 2004, on confronte en pratique les informations collectées en 2009 à une estimation portant sur l'extrapolation du point de recensement 2004 par l'évolution du nombre de logements observée dans la TH entre 2004 et 2009 (cf. tableau 1). Cela revient à confronter la situation réelle de 2009 à une situation fictive, où l'on imagine que le recensement de 2009 n'a pu avoir lieu.

L'étude porte sur les petites communes des groupes de rotation 1 à 4 de France métropolitaine, soit les communes recensées de 2004 à 2007, puis de 2009 à 2012. Il n'a pas été possible de mener la même étude sur le 5^e groupe de rotation, car on ne disposait pas du millésime 2013 de la base TH lorsque cette étude fut entreprise.

Sont considérées comme « petites » les communes dont la population municipale était inférieure à 10 000 habitants lors du recensement de 1999 et que l'Insee continue de recenser exhaustivement. Cette définition est retenue car le mode de recensement de ces communes a été celui d'une petite commune durant la période analysée, quelle que soit l'évolution de leur population après 1999. Il est strictement nécessaire que ces communes n'aient subi aucune modification géographique durant la période d'analyse. On exclut donc les communes disparues, créées ou modifiées ainsi que les communes absentes des données de la TH. Les communes des DOM n'ont pas été retenues dans l'analyse car, La Réunion mise à part, la TH n'y est pas utilisée pour l'extrapolation de la population.

Tableau 1
Présentation de la méthode pour juger de l'apport de la TH dans le 1^{er} groupe de rotation

Année	Situation réelle (point de référence)		Scénario fictif (estimation)	
	État de recensement	Mode de calcul des populations	État de recensement	Mode de calcul des populations
2004	<i>Recensement</i>	Absence de calcul*	<i>Recensement</i>	<i>Recensement</i>
2005		Absence de calcul*		Extrapolation
2006		Extrapolation		Extrapolation
2007		Interpolation**		Extrapolation
2008		Interpolation**		Extrapolation
2009	<i>Recensement</i>	<i>Recensement</i>	Pas de recensement	Extrapolation

Note : * les premières populations légales issues du recensement rénové concernaient les populations au 1^{er} janvier 2006. L'Insee n'a pas publié de populations légales pour les années 2004 et 2005 ; ** la méthode d'interpolation consiste à positionner l'estimation sur la droite qui relie le dernier chiffre de population publié au point de recensement.

Lecture : dans la réalité, les communes du 1^{er} groupe de rotation ont été recensées en 2004 et 2009 ; leurs populations ont été estimées par extrapolation en 2006, par interpolation en 2007 et 2008 avec le recensement en 2009. Dans un scénario fictif auquel est adossée la méthode d'extrapolation TH, on suppose que le recensement n'a pas eu lieu en 2009 ; entre 2005 et 2009, les populations des communes du 1^{er} groupe de rotation sont donc estimées par extrapolation.

Il sera également procédé à la confrontation de la méthode d'extrapolation TH avec des méthodes d'extrapolation alternatives (cf. annexe 3). Deux méthodes alternatives ont été retenues :

- Une méthode d'extrapolation sans TH, que l'on nommera « PL », qui prolonge linéairement les résultats des deux derniers recensements ;
- Une méthode d'extrapolation TH alternative, que l'on nommera « TH* », découlant d'un filtre simplifié des logements basé uniquement sur la condition de nature du local renseignée dans la TH (cf. encadré 3).

Ces différentes méthodes s'appuient sur un schéma identique à celui utilisé pour juger de la qualité d'estimation de la méthode TH usuelle. On imagine un scénario fictif, où le recensement n'a pas eu lieu et où il est nécessaire d'extrapoler les populations cinq années durant pour rafraîchir l'information.

La méthode de prolongement linéaire « PL » est utilisée pour extrapoler les populations des communes dans lesquelles la méthode TH ne peut être mise en œuvre. C'est le cas des communes de Guadeloupe, Guyane et Martinique ainsi que de quelques communes de France métropolitaine dénuées de base fiscale locale (et donc de TH). Il semble donc opportun de comparer cette méthode à la méthode TH usuelle.

La méthode d'extrapolation alternative « TH* » est séduisante par son efficacité et sa simplicité. En effet, la nature du local est le critère le plus discriminant du filtre appliqué actuellement contribuant à l'exclusion de 90 % des locaux et la méthode est peu sujette à interprétation d'une zone géographique à l'autre et de l'Insee à la DGFIP. Il est en effet difficile d'imaginer qu'un local puisse être appréhendé comme une maison à un endroit et comme une piscine à un autre, ou encore comme une maison au sens du RP et comme une piscine au sens fiscal. Cette méthode ne peut toutefois être considérée comme une alternative crédible, car elle renvoie à un concept du logement par trop éloigné de celui du recensement de la population (cf. annexe 4). De ce fait, elle sera employée marginalement dans l'étude.

Les méthodes « PL » et « TH* » permettront respectivement de vérifier le bien-fondé de l'utilisation d'une source externe de données et d'étalonner la qualité du filtre de logements.

La taille de la commune illustrera les différentes analyses. Le choix d'utiliser l'écart relatif

comme indicateur principal dans l'étude incite à décrire les résultats selon cette dimension. Les communes seront regroupées en quatre classes, selon leur population des résidences principales (cf. annexe 5) : moins de 250 habitants, entre 250 et 499 habitants, entre 500 et 999 habitants, 1 000 habitants ou plus.

On confrontera les différentes méthodes d'estimation dans le cadre des deux enjeux précédemment énoncés : les estimations de population dans les groupes de petites communes extrapolées et l'estimation totale de la population extrapolée. La TH servant à estimer un nombre de résidences principales, on convertit respectivement ces deux enjeux en estimations de résidences principales dans les groupes de petites communes extrapolées et estimation totale des résidences principales.

Nous nous intéresserons d'abord aux estimations par commune, en mesurant l'apport de la TH à celles-ci, puis en comparant les estimations issues des méthodes TH et PL et enfin en examinant l'apport de la TH selon le profil d'évolution des logements de la commune.

Avec la méthode TH, l'erreur d'estimation du nombre de résidences principales est inférieure à 5 % pour près de trois communes sur quatre

La proportion de communes dont l'estimation par la méthode TH usuelle s'écarte peu du point de recensement est forte.

Pour près de quatre communes sur dix, l'écart relatif n'excède pas 2 % en valeur absolue. Pour un peu moins de trois communes sur quatre, il n'excède pas 5 % et pour neuf communes sur dix, il est inférieur à 10 % (cf. tableau 2).

Les différences sont minimes entre groupes de rotation. Les distributions de l'écart relatif entre le nombre de résidences principales recensées et le nombre de résidences principales estimé par la méthode TH sont presque confondues d'un groupe de rotation à l'autre (cf. figure I). Pour chacune des courbes, il y a un pic de distribution proche de 0 (soit l'écart relatif nul).

Une légère tendance à la surestimation de la méthode d'extrapolation TH

La méthode d'extrapolation TH conduit à surestimer le nombre de résidences principales des

quatre groupes de rotation étudiés. La surestimation atteint quatre résidences principales en moyenne, hormis pour le 1^{er} groupe de rotation où elle ne vaut que deux résidences (cf. tableau 3). Pour plus de la moitié des communes, l'écart absolu de résidences principales est compris entre - 5 et 6 unités. C'est un intervalle faible, eu égard au fait que la comparaison

intervient au bout de cinq années d'extrapolation. L'écart médian est proche de 0.

La proportion de communes surestimées avec la méthode TH usuelle augmente avec les groupes de rotation (cf. tableau 4). Pour le 4^e groupe de rotation, 53 % des communes ont ainsi une estimation du nombre de résidences principales

Tableau 2
Qualité de l'estimation de la méthode TH usuelle selon le groupe de rotation
Approche par l'écart relatif entre le nombre de résidences principales recensées
et le nombre de résidences principales estimé

Écart relatif compris entre	En %			
	GR1	GR2	GR3	GR4
[-1 % ; 1 %]	21	20	20	21
[-2 % ; 2 %]	38	37	39	40
[-5 % ; 5 %]	72	71	74	73
[-10 % ; 10 %]	91	91	92	92

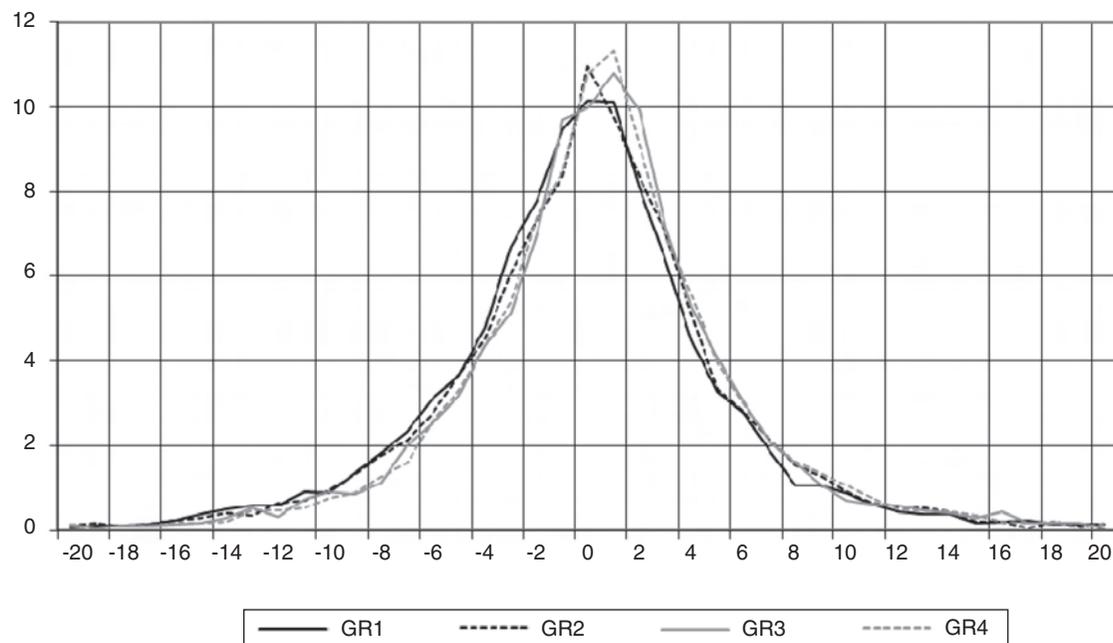
Lecture : pour 20 % des communes du 2^e groupe de rotation, l'écart entre le nombre de résidences principales recensées et l'estimation par la méthode TH usuelle n'excède pas 1 %. 91 % des communes de ce même groupe de rotation ont un nombre de résidences principales estimé par la méthode TH usuelle égal à 10 % près au nombre recensé.

Champ : petites communes de France métropolitaine estimées par extrapolation TH.

Source : EAR, TH.

Figure 1
Distribution de l'approximation de la méthode TH usuelle selon le groupe de rotation
Approche par l'écart relatif entre le nombre de résidences principales recensées
et le nombre de résidences principales estimé

En %



Lecture : la proportion de communes du 1^{er} groupe de rotation dont le nombre de résidences principales estimé avec la méthode TH usuelle surestime de 1 % le nombre de résidences principales recensées vaut 10 %. Par ailleurs, le nombre de résidences principales estimé avec la méthode TH usuelle surestime de 8 % le nombre de résidences principales effectivement recensées pour 2 % des communes du 3^e groupe de rotation.

Champ : petites communes de France métropolitaine estimées par extrapolation TH.

Source : EAR, TH.

par la méthode TH supérieure au recensement. Dans le 1^{er} groupe de rotation, cette proportion n'atteint que 46 % et la part des communes sous-estimées y est supérieure (à 48 %).

La part des communes justement estimées reste en revanche stable dans le temps. Entre 6 et 7 % des communes présentent une estimation du nombre de résidences principales par la méthode TH usuelle (après approximation à l'entier le plus proche) égale au nombre de résidences principales recensées.

Les communes les plus grandes sont plus fréquemment surestimées

La proportion de communes dont le nombre de résidences principales est surestimé avec la méthode TH s'accroît avec la taille des communes (cf. figure II). Plus les communes sont grandes, plus il est fréquent qu'elles soient surestimées. Cette observation se vérifie dans chacun des quatre groupes de rotation.

On constate ainsi un écart avoisinant 20 points entre la proportion de communes de 1 000 habitants ou plus surestimées et celle de moins de 250 habitants, quel que soit le groupe de rotation (cf. tableau 5). Par exemple, dans le 3^e groupe de rotation, 46 % des communes de moins de 250 habitants présentent une estimation strictement supérieure au recensement, contre 67 % pour les communes de 1 000 habitants ou plus.

Il a été impossible de mettre en évidence une explication à ce phénomène. De manière générale, il est difficile de formuler des interprétations de certaines observations faites dans l'étude. Au-delà des multiples facteurs (taille de la commune, géographie, type de logement individuel ou collectif, dimension touristique de la commune...) qui peuvent influencer sur l'adaptabilité de la TH au processus *RP* et qui, pour certains, ont déjà été étudiés et mesurés (SES, 2001), l'étude se base sur la comparaison de deux taux d'évolution. On compare donc non pas un stock de logements à un autre, mais plutôt un flux de logements à un autre. S'extirper

Tableau 3

Approximation de l'estimation de la méthode TH usuelle selon le groupe de rotation Description des écarts relatifs et absolus entre le nombre de résidences principales recensées et le nombre de résidences principales estimé

	En %							
	Écart relatif				Écart absolu			
	Moyenne	1 ^{er} quartile	Médiane	3 ^e quartile	Moyenne	1 ^{er} quartile	Médiane	3 ^e quartile
GR1	- 0,2	- 3,0	- 0,1	2,6	1,9	- 4,6	- 0,2	4,7
GR2	0,2	- 2,8	0,2	2,9	4,3	- 4,0	0,3	5,7
GR3	0,6	- 2,2	0,5	3,1	4,2	- 3,3	0,8	6,0
GR4	0,7	- 2,2	0,5	3,2	3,6	- 3,2	0,9	6,0

Lecture : dans le 4^e groupe de rotation, l'écart relatif moyen atteint 0,7 % ; l'écart absolu moyen y est de 3,6 résidences principales. La médiane de l'écart absolu est à 0,9 résidence principale, ce qui signifie que la moitié des communes ont un écart inférieur à ce niveau et la moitié un écart supérieur.

Champ : petites communes de France métropolitaine estimées par extrapolation TH.

Source : EAR, TH.

Tableau 4

Proportion de communes justement estimées, sous-estimées ou surestimées par la méthode TH usuelle selon le groupe de rotation

	En %		
	Sous-estimation	Estimation juste	Surestimation
GR1	48	6	46
GR2	44	7	49
GR3	42	6	52
GR4	40	7	53

Lecture : dans le 3^e groupe de rotation, 42 % des communes sont sous-estimées en nombre de résidences principales par la méthode TH, 6 % sont parfaitement estimées et 52 % sont surestimées.

Champ : petites communes de France métropolitaine estimées par extrapolation TH.

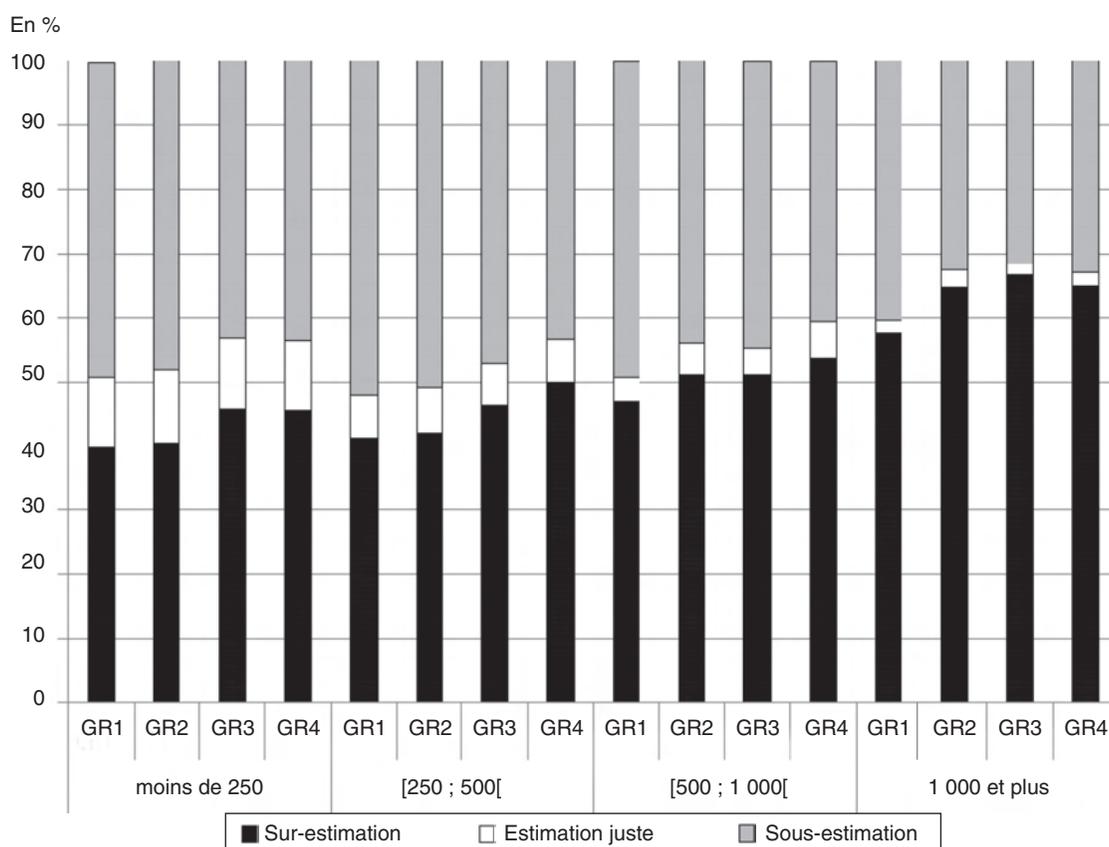
Source : EAR, TH.

du cadre temporel de l'étude et de sa dimension conjoncturelle reste ardu. La surestimation des plus grandes communes est indéniable dans la période d'extrapolation étudiée, mais il n'est pas dit qu'elle caractérisera les prochains millésimes extrapolés. La poursuite de l'étude permettra de

confirmer ou d'infirmer la tendance de la TH à surestimer ce type de communes et peut-être de trouver des clés d'explication au phénomène.

Il n'en reste pas moins que la TH a fait montre de caractéristiques intéressantes pour l'estimation

Figure II
Répartition des communes selon leur niveau d'estimation par la méthode TH usuelle
Déclinaison par taille de commune et groupe de rotation



Lecture : parmi les communes de 500 à 1 000 habitants du 1^{er} groupe de rotation, 47 % ont une estimation du nombre de résidences principales par la méthode TH usuelle strictement supérieure au nombre de résidences principales recensées. Cette proportion atteint respectivement 51 %, 51 % et 54 % pour les communes des 2^e, 3^e et 4^e groupes de rotation.

Champ : petites communes de France métropolitaine estimées par extrapolation TH.

Source : EAR, TH.

Tableau 5
Différence entre la proportion de communes de 1 000 habitants ou plus surestimées avec la méthode TH usuelle et celle des moins de 250 habitants

		En points
		Différence
	GR1	18
	GR2	24
	GR3	21
	GR4	19

Lecture : la proportion de communes de 1 000 habitants ou plus surestimées avec la méthode TH usuelle est supérieure de 21 points à celle des communes de moins de 250 habitants dans le 3^e groupe de rotation.

Champ : petites communes de France métropolitaine estimées par extrapolation TH.

Source : EAR, TH.

des résidences principales dans les périodes étudiées. Dans près de 3 communes sur 4, l'erreur est inférieure à 5 % au bout de cinq années d'extrapolation. En outre, l'erreur est globalement centrée : seules les plus grandes des petites communes (1 000 habitants ou plus) se distinguent par un nombre de résidences principales plus fréquemment surestimé avec la méthode TH.

Comparons maintenant la valeur de la méthode TH à celle de la méthode PL dans les estimations communales. La méthode TH apporte-elle un réel bénéfice ?

Les faiblesses de la méthode de prolongement linéaire mettent en lumière l'apport de la TH dans le calcul des populations

En confrontant les méthodes d'estimation pour les quatre premiers groupes de rotation, on constate que la méthode de prolongement linéaire est de moins bonne qualité que la méthode TH usuelle et même que la méthode TH alternative. Elle est notamment bien moins précise, la variabilité des écarts entre le nombre de résidences principales recensées et le nombre de résidences principales estimé par cette méthode étant importante. En d'autres termes, les méthodes TH estiment mieux le nombre de résidences principales collectées pour plus de communes que la méthode de prolongement linéaire (cf. figure III).

La méthode de prolongement linéaire ne permet d'approcher le nombre de résidences principales recensées à 5 % près que pour une commune sur deux. Au contraire, les méthodes recourant à la TH approchent avec la même erreur le nombre de résidences principales recensées pour près de trois communes sur quatre. Ces résultats sont extrêmement stables d'un groupe de rotation à l'autre (cf. tableau 6).

La méthode TH alternative donne des résultats très proches de la méthode TH usuelle. Quelle que soit l'année et quel que soit le groupe de rotation, la qualité des deux méthodes TH est identique. Ce qui peut apparaître décevant de prime abord doit être vu comme le signe de la robustesse de la source d'informations TH. C'est un élément très rassurant quand on rapporte cela au principal enjeu du recensement : chaque année, la TH sert à estimer la population de plus de 14 000 communes.

L'observation des écarts-types confirme l'étendue de la dispersion des estimations par la

méthode de prolongement linéaire et par ricochet la faible précision de cette méthode (cf. tableau 7). En moyenne, à l'exception du 2^e groupe de rotation, elle est aussi la méthode qui surestime le plus le nombre de résidences principales (entre 3 et 5 résidences principales).

Ces éléments montrent donc l'imprécision de la méthode de prolongement linéaire relativement aux méthodes employant la TH.

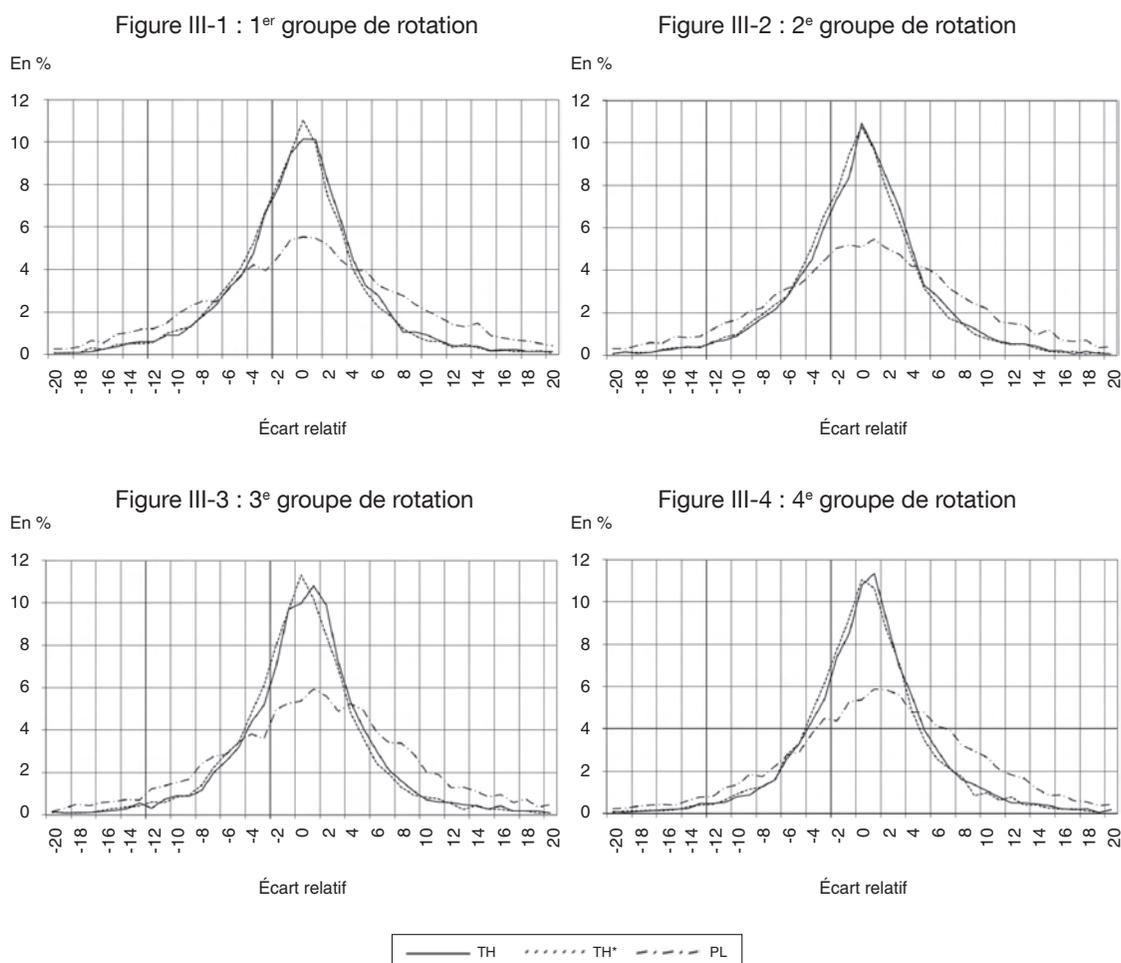
L'utilisation de la TH permet de mieux estimer le nombre de résidences principales que la méthode de prolongement linéaire pour près de sept communes sur dix

Quel que soit le groupe de rotation, la méthode TH usuelle permet d'estimer plus précisément le nombre de résidences principales que la méthode de prolongement linéaire pour plus de 65 % des communes. Dans un peu moins de 10 % des cas, les deux méthodes sont équivalentes. Pour près d'un quart des communes en revanche, la méthode de prolongement linéaire est meilleure que la méthode TH (cf. figure IV).

On considère qu'une méthode est meilleure qu'une autre quand le nombre de résidences principales estimé par la première est plus proche du nombre de résidences principales recensées que le nombre de résidences principales estimé par la seconde méthode. On compare donc les écarts entre le point de recensement et le point estimé. Deux méthodes seront considérées de qualité analogue si les écarts respectifs de leur estimation au nombre de résidences principales recensées sont identiques, même si les estimations sont différentes. Par exemple, si deux méthodes testées fournissent, respectivement, une estimation de 50 et 52 résidences principales, elles seront considérées comme de qualité équivalente si le nombre de résidences principales recensées est 51. Ce parti pris entraîne par construction une occurrence bien plus élevée des equivalences de méthodes dans les groupes de très petites communes. On observe que moins la commune est peuplée, plus il est fréquent que les méthodes TH et de prolongement linéaire aboutissent à la même estimation de nombre de résidences principales (cf. annexe 6).

La taille de la commune n'influe pas sur la qualité de la confrontation des méthodes TH et PL. L'apport de la TH est bénéfique dans l'estimation de 70 % des communes quels que soient leur taille et leur groupe de rotation (cf. tableau 8).

Figure III
Distribution de l'approximation des différentes méthodes d'estimation selon le groupe de rotation
Approche par l'écart relatif entre le nombre de résidences principales recensées
et le nombre de résidences principales estimé



Lecture : la proportion de communes du 1^{er} groupe de rotation dont le nombre de résidences principales estimé avec la méthode TH alternative surestime de 1 % le nombre de résidences principales recensées s'élève à 11 %. Par ailleurs, 2 % des communes du 3^e groupe de rotation possèdent un nombre de résidences principales estimé avec la méthode de prolongement linéaire surestimé de 12 %.

Champ : petites communes de France métropolitaine estimées par extrapolation TH.

Source : EAR, TH.

Tableau 6
Qualité de l'estimation des différentes méthodes utilisées selon le groupe de rotation
Approche par la proportion de communes dont l'estimation de résidences principales
équivalait à 5 % près au recensement

	En %		
	Méthode TH	Méthode TH*	Méthode PL
GR1	72	72	47
GR2	71	72	47
GR3	74	74	49
GR4	73	74	50

Lecture : l'estimation par la méthode TH de 72 % des communes du 1^{er} groupe de rotation ne s'écarte pas de plus de 5 % du nombre de résidences principales recensées.

Champ : petites communes de France métropolitaine estimées par extrapolation TH.

Source : EAR, TH.

Tableau 7

Approximation de l'estimation des différentes méthodes utilisées selon le groupe de rotation
Description de l'écart absolu entre le nombre de résidences principales recensées et le nombre de résidences principales estimé

	GR1		GR2		GR3		GR4	
	Moyenne	Écart-Type	Moyenne	Écart-Type	Moyenne	Écart-Type	Moyenne	Écart-Type
Méthode TH	1,9	26,9	4,3	29,5	4,2	24,9	3,6	24,9
Méthode TH*	0,7	26,1	2,7	28,0	2,5	23,5	2,7	24,5
Méthode PL	3,1	46,9	4,2	47,8	5,1	42,3	5,3	41,7

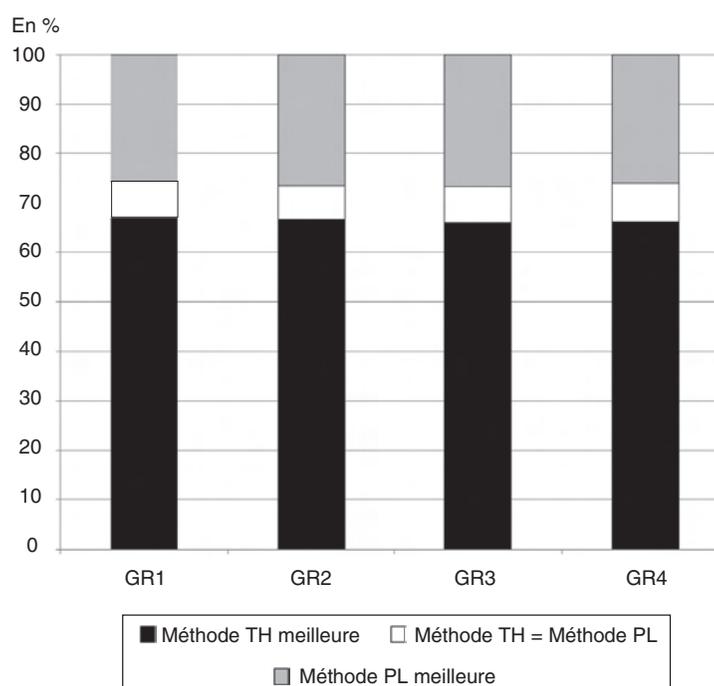
Lecture : la surestimation moyenne par la méthode TH alternative pour les communes du 2^e groupe de rotation atteint 2,7 résidences principales ; l'écart-type de l'écart absolu est de 28,0 résidences principales.

Champ : petites communes de France métropolitaine estimées par extrapolation TH.

Source : EAR, TH.

Figure IV

Confrontation des méthodes TH et PL par groupe de rotation



Lecture : 67 % des communes du 1^{er} groupe de rotation sont mieux extrapolées avec la TH qu'elles ne le seraient avec la méthode de prolongement linéaire. Pour 7 % d'entre elles, les estimations des deux méthodes sont équivalentes. Enfin, la méthode de prolongement linéaire serait meilleure pour 26 % des communes.

Champ : petites communes de France métropolitaine estimées par extrapolation TH.

Source : EAR, TH.

Tableau 8

Part de communes où la méthode TH usuelle est strictement meilleure que la méthode de prolongement linéaire

Déclinaison par groupe de rotation et taille de commune sur l'ensemble des communes où les deux méthodes d'estimation ne sont pas de qualité équivalente

	En %				
	Moins de 250 habitants	Entre 250 et 499 habitants	Entre 500 et 999 habitants	1 000 habitants ou plus	Tous
GR1	72	72	74	71	72
GR2	72	72	71	70	72
GR3	71	73	72	70	71
GR4	70	72	74	72	72

Lecture : pour les communes dont la population des ménages ordinaires est comprise entre 250 et 500 habitants et pour lesquelles les deux méthodes d'estimation ne sont pas de qualité équivalente, l'estimation par la TH est strictement meilleure que l'estimation par prolongement linéaire dans respectivement 72 %, 72 %, 73 % et 72 % des cas selon qu'il s'agisse du 1^{er}, 2^e, 3^e ou 4^e groupe de rotation.

Champ : petites communes de France métropolitaine estimées par extrapolation TH.

Source : EAR, TH.

La confrontation des méthodes fournit un nouvel argument en faveur de l'utilisation de la TH dans le processus de calcul des populations. Relativement à la méthode de prolongement linéaire, son apport est irréfutable. D'une part, la méthode TH fournit des estimations plus précises que la méthode PL : près de trois communes sur quatre sont bien estimées à 5 % près avec la méthode TH au bout de cinq années d'extrapolations, quand environ une sur deux l'est avec la méthode PL. D'autre part, quand il s'agit de fournir la meilleure estimation par commune, la méthode TH l'emporte encore sur la méthode PL dans près de 70 % des cas.

Les éléments présentés jusqu'ici témoignent de la robustesse de la TH et d'une plus-value indéniable dans l'estimation des résidences principales. Son apport peut également être jugé d'un point de vue plus qualitatif : est-il un bon outil pour toutes les communes, y compris pour celles dont l'évolution du nombre de logements est moins linéaire, plus erratique ?

La TH rend compte correctement des ruptures d'évolution des parcs de logements

Si la TH permet de saisir correctement les ruptures de tendance dans la dynamique d'évolution des parcs de logements, elle présentera là un avantage considérable par rapport à la méthode de prolongement linéaire. Par construction, cette dernière ne peut restituer les retournements ou les chocs d'évolution. Lorsque le nombre de logements d'une commune diminue après avoir augmenté, ou lorsqu'il augmente soudainement après une période de stabilité, la méthode de prolongement linéaire délivre une estimation éloignée de la vérité, car elle reproduit la tendance d'évolution observée dans le passé.

En étudiant l'apport de la méthode TH usuelle selon le profil de l'évolution du nombre de résidences principales durant la dernière période intercensitaire, on peut répondre à la question : la TH engendre-t-elle ou non une bonne estimation du nombre de résidences principales dans les communes où la dynamique de logements est instable ?

On convient que l'évolution du parc des logements est stable dans une commune lorsque l'écart entre le nombre de résidences principales estimé par la méthode de prolongement linéaire et le nombre de résidences principales recensées ne dépasse pas 5 %. Cette définition

divise les communes en deux groupes d'effectif comparable : la moitié des communes a une dynamique de logements stable (47 %), et l'autre moitié une évolution heurtée (53 %) (cf. tableau 10). Lorsque la dynamique du parc de logements est stable, la méthode PL estime correctement tandis qu'elle estime mal quand la série de logements subit des à-coups. On utilise le même seuil de 5 % pour juger de la qualité de l'estimation du nombre de résidences principales par la méthode TH avec le nombre de résidences principales recensées : si l'écart relatif par la méthode TH usuelle est inférieur ou égal à 5 %, on considère que la méthode TH estime correctement.

Pour simplifier la présentation des résultats, on se contente ici d'étudier un seul groupe de rotation, en l'occurrence le premier. Les différences sont minimales entre groupes de rotation et l'affichage de l'ensemble des résultats nuit à la lecture. L'analyse comparée des résultats de la méthode TH usuelle et de la méthode PL n'est en tout cas pas modifiée par le choix du groupe de rotation.

En combinant la qualité d'estimation des deux méthodes TH et PL, quatre ensembles de communes se dessinent (cf. tableau 9 et tableau 10) :

- Les communes correctement estimées avec les deux méthodes (ensemble A), soit 39 % des communes ;
- Les communes correctement estimées avec la méthode PL mais pas avec la méthode TH (ensemble B1), soit 8 % des communes ;
- Les communes correctement estimées avec la méthode TH mais pas avec la méthode PL (ensemble B2), soit 32 % des communes ;
- Les communes mal estimées par les deux méthodes (ensemble C), soit 21 % des communes.

L'apport de la TH à l'estimation des populations peut être interprété de la manière suivante : l'ensemble B2 constitue l'ensemble des communes pour lesquelles l'estimation est améliorée grâce à l'utilisation de la TH, tandis que l'ensemble B1 est composé des communes dont l'estimation est dégradée du fait de son utilisation.

Pour l'ensemble des communes du 1^{er} groupe de rotation, la méthode TH améliore l'estimation de 32 % des communes et la dégrade dans 8 % des cas.

Le rapport de l'ensemble A à la somme des ensembles A et B1 constitue la proportion de

communes bien estimées avec la méthode TH quand l'évolution des logements est stable. Le rapport de l'ensemble B2 à la somme des ensembles B2 et C constitue la proportion de communes bien estimées avec la méthode TH quand l'évolution des logements est instable.

Confronter ces rapports permet de mesurer la qualité de la TH selon la stabilité de l'évolution des logements (cf. tableau 11).

La méthode TH estime mieux le nombre de logements des communes dont l'évolution

Tableau 9

Caractérisation des communes selon la qualité de l'estimation des méthodes TH et PL

	Bonne estimation de la méthode PL Évolution stable des logements	Mauvaise estimation de la méthode PL Évolution instable des logements
Bonne estimation de la méthode TH	Ensemble A $ \text{Écart}_{\text{Méthode PL}} \leq 5\% \text{ et } \text{Écart}_{\text{Méthode TH}} \leq 5\%$	Ensemble B2 $ \text{Écart}_{\text{Méthode PL}} > 5\% \text{ et } \text{Écart}_{\text{Méthode TH}} \leq 5\%$
Mauvaise estimation de la méthode TH	Ensemble B1 $ \text{Écart}_{\text{Méthode PL}} \leq 5\% \text{ et } \text{Écart}_{\text{Méthode TH}} > 5\%$	Ensemble C $ \text{Écart}_{\text{Méthode PL}} > 5\% \text{ et } \text{Écart}_{\text{Méthode TH}} > 5\%$

Lecture : les communes bien estimées par les méthodes PL et TH constituent l'ensemble A : l'écart entre l'estimation par chacune des méthodes et le recensement est inférieur à 5 %. Les communes mal estimées avec la méthode PL et bien estimées avec la méthode TH forment l'ensemble B2 : l'écart entre l'estimation par la méthode PL et le recensement est supérieur à 5 %, tandis que l'écart entre l'estimation par la méthode TH et le recensement est inférieur à 5 %.

Tableau 10

Répartition des communes du 1^{er} groupe de rotation selon la qualité de l'estimation des méthodes TH et PL

	Bonne estimation de la méthode PL $ \text{Écart}_{\text{Méthode PL}} \leq 5\%$	Mauvaise estimation de la méthode PL $ \text{Écart}_{\text{Méthode PL}} > 5\%$	Total	En %
Bonne estimation de la méthode TH $ \text{Écart}_{\text{Méthode TH}} \leq 5\%$	Ensemble A : 39	Ensemble B2 : 32	72	
Mauvaise estimation de la méthode TH $ \text{Écart}_{\text{Méthode TH}} > 5\%$	Ensemble B1 : 8	Ensemble C : 21	28	
Total	47	53	100	

Lecture : 39 % des communes du 1^{er} groupe de rotation sont bien estimées par les deux méthodes ; elles constituent l'ensemble A. Champ : petites communes du 1^{er} groupe de rotation de France métropolitaine estimées par extrapolation TH. Source : EAR, TH.

Tableau 11

**Proportion de communes du 1^{er} groupe de rotation estimées correctement par la méthode TH usuelle
Ventilation par taille de commune et qualité d'estimation de la méthode PL**

	Bonne estimation de la méthode PL Évolution stable des logements	Mauvaise estimation de la méthode PL Évolution instable des logements	Ensemble des communes	En %
Moins de 250 habitants	65	45	52	
Entre 250 et 499 habitants	82	62	72	
Entre 500 et 999 habitants	91	76	84	
1 000 habitants ou plus	93	83	89	
Tous	83	61	72	

Lecture : dans le groupe des communes de 1 000 habitants ou plus du 1^{er} groupe de rotation, la proportion de communes correctement estimées avec la méthode TH usuelle (à 5 % près) s'élève à 83 % lorsque l'évolution du nombre de résidences principales est heurtée (mauvaise qualité d'estimation de la méthode PL) et à 93 % lorsqu'elle est stable (bonne qualité d'estimation de la méthode PL). Au total, 89 % des communes de 1 000 habitants ou plus sont correctement estimées avec la méthode TH usuelle. Champ : petites communes du 1^{er} groupe de rotation de France métropolitaine estimées par extrapolation TH. Source : EAR, TH.

du parc de logements est linéaire, quelle que soit la taille de la commune. Toutefois, l'écart s'amenuise avec la taille de la commune. Dans les communes de moins de 250 habitants, la TH estime correctement le nombre de résidences principales recensées dans 65 % des cas quand l'évolution est stable et dans un peu moins d'un cas sur deux quand l'évolution est instable (cf. tableau 11). En revanche, pour les deux ensembles de communes les plus peuplées à l'évolution heurtée, la méthode TH approche correctement le nombre de résidences principales recensées dans 4 cas sur 5, contre un peu plus de 90 % dans les communes à l'évolution linéaire.

En somme, la TH estime de manière plus précise les communes à l'évolution linéaire, mais la dégradation de l'estimation pour les communes à l'évolution heurtée reste acceptable. Ce constat n'a rien de choquant si l'on considère qu'il est plus facile de mesurer des états inertes ou progressant de manière régulière, quel que soit le contexte. Il est même plutôt encourageant compte tenu des écarts observés. Pour les situations heurtées, le recours à la TH dans le cadre du calcul des populations légales est précieux. Les ruptures de tendance sont souvent correctement appréhendées. La TH s'avère ainsi un bon instrument pour les estimations des communes en extrapolation, quel que soit le profil de l'évolution de leur nombre de logements.

On constate également dans le tableau 11 que la part de communes correctement estimées avec la méthode TH augmente avec la taille de la commune. Ainsi, 52 % des communes de moins de 250 habitants sont correctement estimées à 5 % près avec la méthode TH contre 89 % des communes de 1 000 habitants ou plus. Cette observation est à rapprocher de l'usage d'un critère de comparaison d'écart relatif au nombre de logements. Ce choix désavantage les communes de petite taille. Une différence de deux résidences principales pour une commune de 50 résidences principales génère un écart relatif de 10 %, contre 1 % pour une commune de 500 résidences principales. C'est pourquoi la proportion de communes dont l'écart par la méthode TH (et par la méthode PL également) est supérieur à 5 % est bien plus important dans les groupes de petites communes (cf. annexe 7).

Les bénéfices de la TH à l'estimation des petites communes en extrapolation sont donc multiples. D'un point de vue quantitatif, les estimations

sont bien meilleures qu'avec la méthode de prolongement linéaire. D'un point de vue qualitatif, la TH est précieuse pour estimer les populations des communes à l'évolution heurtée.

Il reste à mesurer l'apport de la TH dans le cadre du deuxième enjeu : l'estimation totale du nombre de résidences principales des petites communes en extrapolation.

La TH tend à surestimer le nombre global de résidences principales des communes extrapolées

Quels que soient le groupe de rotation et la plage d'extrapolation utilisée dans le cadre de l'étude (2004-2009 pour le 1^{er} groupe de rotation, 2005-2010 pour le 2^e...), le recours à la TH conduit à surestimer le nombre de résidences principales des communes en extrapolation.

L'estimation par la méthode TH alternative est meilleure que l'estimation par la méthode TH usuelle, et les deux méthodes TH donnent de meilleurs résultats que la méthode de prolongement linéaire (cf. figure V).

L'avantage comparatif de la méthode TH par rapport à la méthode PL reste cependant mesuré dans le cadre de cette problématique. Quelle que soit la méthode utilisée, la surestimation est faible. Au bout de cinq années d'extrapolation, on aboutit à une surestimation d'au plus 1,1 % du nombre total de résidences principales avec la méthode TH usuelle. L'écart lié à la méthode TH alternative n'excède pas 0,7 % quel que soit le groupe de rotation, tandis que celui de la méthode de prolongement linéaire est compris entre 0,8 et 1,4 % (cf. tableau 12).

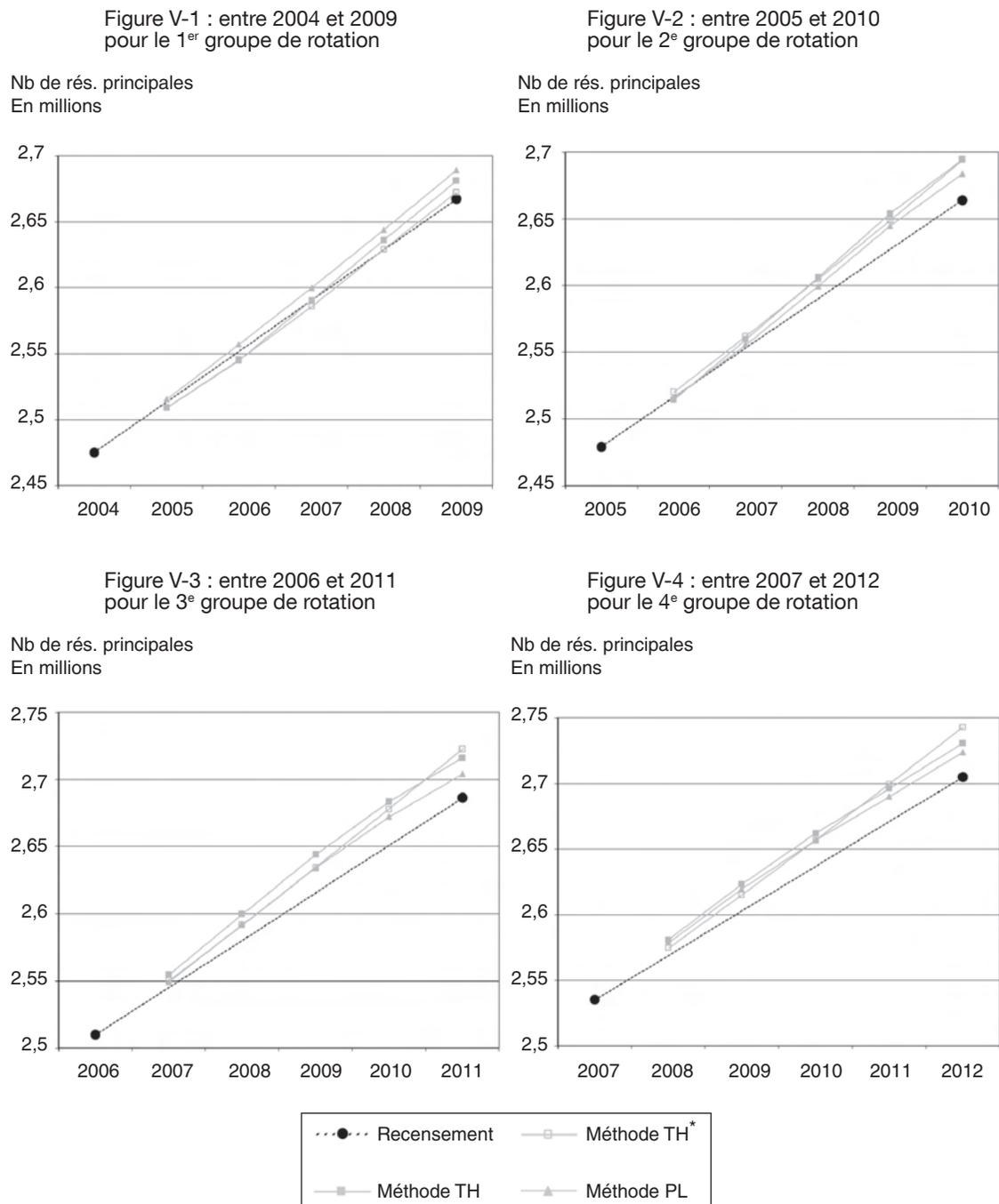
Toutes les méthodes d'extrapolation testées conduisent à une estimation trop élevée du nombre de résidences principales. S'agissant de la méthode de prolongement linéaire, mécaniquement, cela tient au fait que l'évolution des logements est moins dynamique sur la dernière période intercensitaire qu'elle ne l'était lors de l'avant-dernière. La surestimation est donc de nature conjoncturelle. Les travaux entrepris ne permettent pas d'établir la nature de la surestimation de la méthode TH : est-elle conjoncturelle, structurelle ou résulte-t-elle d'un mélange des deux ?

On peut toutefois affirmer que dans les périodes d'extrapolation étudiées de cinq ans,

la TH a conduit à une légère surestimation du nombre de résidences principales. Pour les années d'extrapolation effective, c'est-à-dire une et deux années après le recensement, la surestimation s'avère également extrêmement

plausible et vraisemblablement moindre qu'au bout de cinq années d'extrapolation. Il faudrait imaginer une évolution erratique de l'approximation de l'estimation par la méthode TH durant la période intercensitaire pour que ce ne

Figure V
Évolution du nombre de résidences principales recensées et estimations correspondantes



Lecture : le nombre de résidences principales recensées et ce nombre estimé par les différentes méthodes testées sont compris entre 2 650 000 et 2 700 000 en 2009 pour les communes du 1^{er} groupe de rotation. Dans tous les groupes de rotation, les estimations sont toujours supérieures au point de référence, c'est-à-dire au point de recensement.
Champ : petites communes de France métropolitaine estimées par extrapolation TH.
Source : EAR, TH.

soit pas le cas⁵. Or, la logique voudrait plutôt que l'écart se creuse progressivement au cours du temps.

L'approximation réellement engendrée par la TH pour les communes estimées par extrapolation dans le cadre du calcul des populations légales est moindre que celle calculée dans l'étude

La méthode utilisée dans l'étude jauge la qualité de la TH sur une durée cumulée et extrapolée de cinq ans. Il est impossible d'estimer le niveau de l'erreur lors des années réelles d'extrapolation en ne recourant pas à de nouvelles hypothèses. On peut en revanche avancer raisonnablement que les tendances constatées (notamment la surestimation des résidences principales ou l'apport bénéfique de la TH dans l'estimation de la population) restent vraies quelle que soit la durée de l'extrapolation, et que l'approximation engendrée par la TH est en réalité un minorant de celle calculée dans l'étude. Plusieurs arguments étayaient cette affirmation.

D'une part, si l'on convient que l'extrapolation se révèle d'autant plus hasardeuse que la durée sur laquelle elle porte est longue, le jugement porté dans l'étude exagère le niveau d'erreur imputable à la TH pour les extrapolations à une et deux années, qui sont celles réalisées pour le calcul des populations légales. La deuxième réserve relative à la méthode a trait au fait que l'on utilise quatre points de mesure (deux points TH et deux EAR) alors que dans la réalité de l'extrapolation, on n'en utilise que trois (deux points TH et un EAR). Le 4^e point utilisé (celui de comparaison) n'est, comme les trois autres, pas forcément dénué d'erreurs de mesure. On conviendra sans mal que cela ajoute un surplus artificiel

d'approximation, sans qu'il ne soit possible de l'évaluer formellement.

D'autre part, l'Insee mène chaque année un long et minutieux examen de l'évolution de nombre de logements des petites communes en extrapolation avec le soutien des Directions régionales, du Pôle Recensement et du Pôle Communautés de l'Insee. Des corrections sont engagées lorsque l'ampleur de l'anomalie détectée est significative. Trois formes de contrôle sont engagées :

- Calage du champ de logements TH utilisé pour l'extrapolation,
- Contrôle des artefacts de gestion dans la base TH,
- Contrôle de l'approximation relative à la formule de calcul du coefficient TH.

Premièrement, on cale le champ de logements TH utilisé pour l'extrapolation. On s'assure qu'il est cohérent avec le champ des logements de la dernière EAR pour les communes à l'évolution remarquable ou suspecte. Le risque de prendre en compte de manière fallacieuse une communauté a déjà été évoqué (cf. encadré 3). L'autre danger récurrent provient de l'absence de synchronisation entre TH et RP d'adresses particulièrement contributives à l'évolution.

5. À ce titre, l'exemple récent du recensement allemand est instructif. Le recensement de 2011 s'est tenu plus de vingt ans après les précédents recensements qui dataient d'avant la réunification : en 1987 en RFA et en 1981 en RDA. Dans l'intervalle, l'emploi de registres administratifs a contribué à extrapoler la population. Or, le recensement de 2011 a conduit à réévaluer la population à la baisse de 1,5 million d'habitants. La surestimation a été essentiellement causée par la prise en compte de migrants non radiés des registres des « Länder » où ils n'habitaient plus (Héran, 2013). Il est plausible que la détérioration de l'estimation ait été progressive dans l'exemple allemand avec l'accumulation des non-radiations. Cela signifierait que l'approximation entre la population réelle et la population estimée n'a fait que s'accroître au cours du temps.

Tableau 12
Écart relatif entre le nombre de résidences principales recensées et le nombre de résidences principales estimé par les méthodes TH, TH* et PL

	Année de comparaison	Méthode TH	Méthode TH*	Méthode PL
GR1	2009	0,5	0,2	0,8
GR2	2010	1,1	0,7	1,1
GR3	2011	1,1	0,7	1,4
GR4	2012	1,0	0,7	1,4

Lecture : le nombre de résidences principales estimé par la méthode TH usuelle est supérieur de 0,5 % au nombre de résidences principales recensées du 1^{er} groupe de rotation en 2009.
Champ : petites communes de France métropolitaine estimées par extrapolation TH.
Source : EAR, TH.

Par exemple, si une résidence de 50 logements recensée en 2009 absente de la TH 2009 apparaît dans la TH 2010, il ne faut pas comptabiliser cette adresse dans l'évolution du nombre de logements sous peine d'augmenter artificiellement la population de la commune.

Deuxièmement, on contrôle les artefacts de gestion dans le fichier TH qui déforment le comptage des logements ; cela concerne les variables utilisées pour le filtre des logements. L'ampleur de la déformation nous conduit à juger les modifications apportées dans la base artificielles au sens de la problématique *RP* du calcul des logements. On suspecte alors fortement une mise à jour ciblée et localisée de l'information par les services fiscaux sans lien avec une réelle évolution du parc de logements.

Troisièmement, on contrôle l'approximation de la formule de calcul utilisée pour l'extrapolation, à savoir l'utilisation de l'évolution du nombre de logements total pour extrapoler le nombre de résidences principales. Dans les communes où l'évolution est essentiellement

portée par les logements non principaux, on privilégie la série des résidences principales pour le calcul du coefficient TH.

Cette triple expertise est menée uniquement pour les communes en extrapolation et ne peut être évaluée, bien que l'on prétende assurément améliorer les estimations du nombre de résidences principales grâce aux corrections proposées. Il existe donc une partie du processus de calcul des populations légales non mesuré dans l'étude qui participe à l'amélioration de l'utilisation de la TH. Chaque année, entre 1 et 2 % des estimations des communes en extrapolation sont corrigées à l'issue de l'expertise.

Le processus de contrôle et de correction des séries de logements ainsi que les approximations relatives à la méthode de comparaison sont des gages de la majoration de l'erreur affichée dans l'étude. S'il est impossible de l'évaluer précisément, elle conforte indiscutablement l'apport de la TH au recensement de la population, et plus particulièrement à l'estimation des populations. □

BIBLIOGRAPHIE

Bergouignan C. (1999), « Les sources administratives : un outil pour la démographie locale », Thèse de doctorat soutenue à l'université Montesquieu – Bordeaux IV.

Clanché F. (2010), « L'utilisation de la taxe d'habitation en France dans la nouvelle méthode de recensement », Conférence Eurostat 10-11 mai 2010.

Desrosières A. (2004), « Enquêtes *versus* registres administratifs : réflexions sur la dualité des ressources statistiques », *Courrier des statistiques*, n° 111, Insee, pp. 3-16.

Descours L. (1991), « Estimation de populations locales par la méthode de la taxe d'habitation », *Insee Méthodes*, n° 29-30-31, pp. 355-362.

Direction générale des Finances publiques (2013), « Le rapport d'activité 2012 de la direction générale des finances publiques », *Tableaux statistiques*, pp. 35-54.

Godinot A. (2003), « La rénovation du recensement de la population », *Courrier des statistiques*, n° 105-106, Insee, pp. 5-12.

Godinot A. (2005), « Pour comprendre le recensement de la population », *Insee Méthodes*, n° hors série, mai.

Héran F. (2013), « L'Allemagne et ses minorités ou les surprises du recensement de 2011 », *Note d'actualité*, Ined, 5 juin 2013.

Labat J.-C. et Decaudin G. (1996), « Une méthode synthétique, robuste et efficace pour réaliser des estimations locales de population », *Document de travail Insee*, Méthodologie statistique, n° 9601.

Pan Ké Shon J.-L. (2007), « Le recensement rénové français et l'étude des mobilités », *Population*, n° 1, Ined, pp 123-141.

Pirou D., Poullain N. et Rochelle S. (2013), « La vie en communauté : 1,6 million de personnes en France », *Insee Première*, n° 1434, février.

Service économique et statistique (SES) (2003), « Connaître les logements et leurs occupants à partir des sources Filocom et du Recensement », *Les documents du travail du SES*, Ministère de l'Écologie, du développement durable et de l'énergie).

SIGLES ET DÉFINITIONS**Sigles**

DGFIP : Direction générale des Finances publiques

EAR : Enquête annuelle de recensement

EDF : Électricité de France

GR : Groupe de rotation

HMSA : Habitations mobiles terrestres et fluviales et personnes sans-abri

Insee : Institut national de la statistique et des études économiques

RIL : Répertoire d'immeubles localisés

RP : Recensement de la population

TH : Taxe d'habitation

Méthode TH : Méthode d'extrapolation basée sur le filtre usuel des logements dans la TH

Méthode TH* : Méthode d'extrapolation basée sur un filtre alternatif des logements dans la TH

Méthode PL : Méthode d'extrapolation basée sur le prolongement linéaire des deux derniers recensements

Définitions au sens de l'Insee

Un logement est un local séparé indépendant utilisé pour l'habitation.

Une résidence principale est un logement occupé de façon habituelle et à titre principal par un ménage.

Une résidence secondaire est un logement d'habitation occupé seulement pendant une partie de l'année pour les week-ends, les vacances ou les loisirs.

Une communauté est définie comme un ensemble de locaux d'habitation relevant d'une même autorité gestionnaire et dont les habitants partagent à titre habituel un mode de vie commun.

ANNEXE 2

COMPOSITION ET ESTIMATION DE LA POPULATION DES PETITES COMMUNES

L'ensemble de la population des ménages ordinaires, de la population des communautés et de la population

des habitations mobiles terrestres et fluviales et des personnes sans-abri (HMSA) forme la population municipale.

$$PopMunicipale = PopMénage + PopCommunautés + PopHMSA$$

La population des ménages ordinaires est la catégorie de population la plus représentée, bien que, dans quelques communes, la population des communautés

soit plus importante. Elle représente 98 % de la population municipale (cf. tableau a).

Tableau a

Répartition des différentes catégories de population au 1^{er} janvier 2010

En %

Populations légales 2010	Population des ménages ordinaires	Population des communautés	Population HMSA
Communes en extrapolation	97,9	1,9	0,2

Lecture : la proportion de la population des ménages ordinaires dans l'ensemble de la population municipale est de 97,9 % pour les petites communes en extrapolation des populations légales 2010.

Champ : petites communes de France métropolitaine estimées par extrapolation TH en 2010.

Source : RP.

C'est pourquoi on a tendance, pour les petites communes, à assimiler la méthode de calcul de la population des ménages ordinaires à la méthode de calcul de la population municipale. Or, les différentes catégories de population sont calculées de manière indépendante et la TH ne participe qu'au calcul de la population des ménages ordinaires (cf. tableau b).

des communautés de la commune pour cinq années, aux ouvertures et fermetures de communautés près, dont l'occurrence est enregistrée dans le répertoire des communautés (Pirou *et al.*, 2013).

La population des communautés est calculée à partir du recensement et d'un répertoire que tient à jour l'Insee. Les communautés sont recensées exhaustivement tous les cinq ans. Dans les petites communes, leur recensement a lieu la même année que le recensement des ménages ordinaires. Le recensement fixe la population

Les habitations mobiles terrestres et les personnes sans-abri sont également recensées dans les petites communes la même année que le recensement des ménages ordinaires. Les habitations mobiles fluviales sont recensées une année donnée dans toute la France par cycle de cinq ans. Les effectifs recensés à cette occasion ne sont pas modifiés pour la commune concernée jusqu'au recensement suivant.

Tableau b

Mode de calcul utilisé pour estimer les différentes catégories de population dans les petites communes selon l'année

Année de calcul des populations légales	Mode de calcul par catégorie de population		
	Population des ménages ordinaires	Population des communautés	Population HMSA
Année de Recensement	Recensement	Recensement	Recensement
Une ou deux années avant le Recensement	Interpolation	Recensement + Répertoire des communautés	Recensement
Une ou deux années après le Recensement	Extrapolation TH	Recensement + Répertoire des communautés	Recensement

Lecture : quand l'année de calcul des populations légales se situe une ou deux années après le dernier recensement, la population des ménages ordinaires est estimée par extrapolation, la population des communautés à partir du recensement et du répertoire des communautés et la population HMSA avec le dernier recensement.

En résumé, pour les groupes de rotation en extrapolation, la TH contribue bien au calcul de la population, mais n'est pas le seul élément à intervenir. D'une part, le coefficient de décohabitation influe sur le calcul de la population des

ménages ordinaires ; d'autre part, les populations des communautés et HMSA (le plus souvent de faible effectif dans la population municipale) sont estimées en dehors de toute intervention des données TH.

PRÉSENTATION DES MÉTHODES DE CALCUL MISES EN ŒUVRE

Le nombre de résidences principales estimé par extrapolation dans l'étude, qui constitue l'estimation TH selon la formule usuelle, est le produit du nombre de résidences

principales recensées cinq ans auparavant par l'évolution du nombre de logements comptabilisés dans la TH durant la période intercensitaire. Il est donc calculé comme suit :

$$NbResPr inc_N^{ExtrapolationTH} = NbResPr inc_{N-5}^{EAR} \times \frac{NbLog_N^{TH}}{NbLog_{N-5}^{TH}}$$

où

- $NbResPr inc_N^{ExtrapolationTH}$ est le nombre de résidences principales estimé par extrapolation TH au 1^{er} janvier de l'année N ;
- $NbResPr inc_{N-5}^{EAR}$ est le nombre de résidences principales recensées au 1^{er} janvier de l'année N-5 ;
- $NbLog_X^{TH}$ est le nombre de logements comptabilisés dans la TH au 1^{er} janvier de l'année X.

Le tableau c répertorie les différentes méthodes de calcul mises en œuvre dans l'étude.

Tableau c

Formules de calcul des méthodes mises en œuvre dans l'étude

Méthode	Libellé	Formule
Recensement	RP	$NbResPr inc_N^{EAR}$
Extrapolation TH	TH	$NbResPr inc_N^{TH} = NbResPr inc_{N-5}^{EAR} \times \frac{NbLog_N^{TH}}{NbLog_{N-5}^{TH}}$
Prolongement linéaire	PL	$NbResPr inc_N^{PL} = NbResPr inc_{N-5}^{EAR} \times \left(\frac{NbResPr inc_{N-5}^{EAR}}{NbResPr inc_{1999}^{RP}} \right)^{\frac{5}{N-5-1999}}$
Extrapolation TH alternative	TH*	$NbResPr inc_N^{TH*} = NbResPr inc_{N-5}^{EAR} \times \frac{NbLog_N^{TH*}}{NbLog_{N-5}^{TH*}}$

où

- $NbResPr inc_X^{EAR}$ est le nombre de résidences principales recensées lors du recensement de l'année X ;
- $NbResPr inc_N^{TH}$ est le nombre de résidences principales extrapolé avec la méthode TH usuelle au 1^{er} janvier de l'année N ;
- $NbResPr inc_N^{PL}$ est le nombre de résidences principales extrapolé par prolongement linéaire au 1^{er} janvier de l'année N ;
- $NbResPr inc_{1999}^{RP}$ est le nombre de résidences principales recensées lors du dernier recensement exhaustif de 1999 ;
- $NbResPr inc_N^{TH*}$ est le nombre de résidences principales extrapolé par la méthode alternative TH au 1^{er} janvier de l'année N ;
- $NbLog_X^{TH*}$ est le nombre de logements comptabilisés avec le filtre alternatif TH au 1^{er} janvier de l'année X.

Le point de recensement est le point de référence (« méthode RP »). La méthode d'extrapolation TH est la méthode actuellement mise en œuvre dans le cadre du calcul des populations légales une année et deux années après la dernière enquête annuelle de recensement. La méthode de prolongement linéaire PL est la méthode de calcul utilisée en l'absence de TH. La méthode d'extrapolation alternative TH* est une méthode d'extrapolation simplifiée utilisant la TH.

ANNEXE 4

POURQUOI LA SÉRIE TH ALTERNATIVE N'EST-ELLE PAS UNE ALTERNATIVE CRÉDIBLE ?

Pour que l'utilisation d'une série TH alternative (série TH*) soit légitime en vue du calcul du coefficient TH, il faut que le filtre correspondant qui permet de définir les logements satisfasse à deux conditions :

- 1^{re} condition : le nombre de logements défini par le filtre alternatif TH doit être proche du nombre de logements recensés ; si ce n'est pas le cas, cela signifie en effet que le filtre basé sur les données fiscales renvoie à un concept du logement éloigné de celui du recensement ;
- 2^e condition : à supposer que la 1^{re} condition soit remplie, la qualité de la série alternative pour l'estimation du nombre de résidences principales doit être au moins aussi bonne que la série actuellement utilisée.

La 1^{re} condition n'est pas satisfaite par le filtre des logements alternatif TH*. Il conduit à surestimer le nombre de logements de manière plus prononcée que le filtre usuel. La distribution des écarts relatifs au nombre de logements recensés le montre (cf. figure a) : le pic de distribution de l'écart de la méthode alternative est proche de 5 %, tandis que celui des écarts associés au filtre usuel vaut le plus souvent 1 %. Les distributions étant par ailleurs réparties de manière équitable autour du pic de distribution, on en conclut que le filtre alternatif conduit à une surestimation structurelle forte du nombre de logements.

Figure a
Distribution des écarts relatifs entre le nombre de logements recensés et le nombre de logements comptabilisé par les filtres TH et TH*

Figure a1 : 1^{er} groupe de rotation

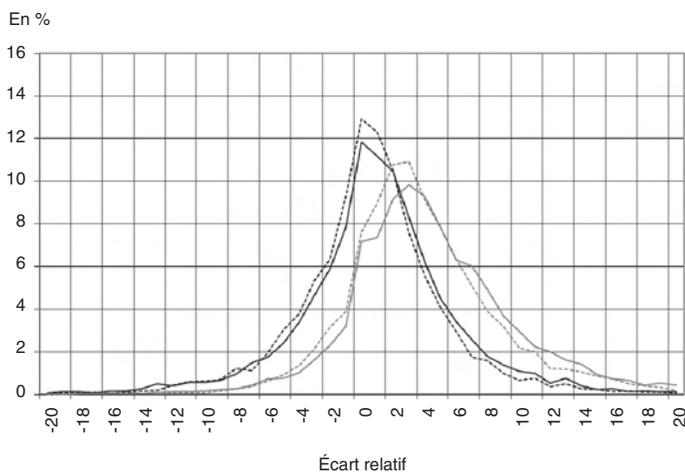


Figure a2 : 2^e groupe de rotation

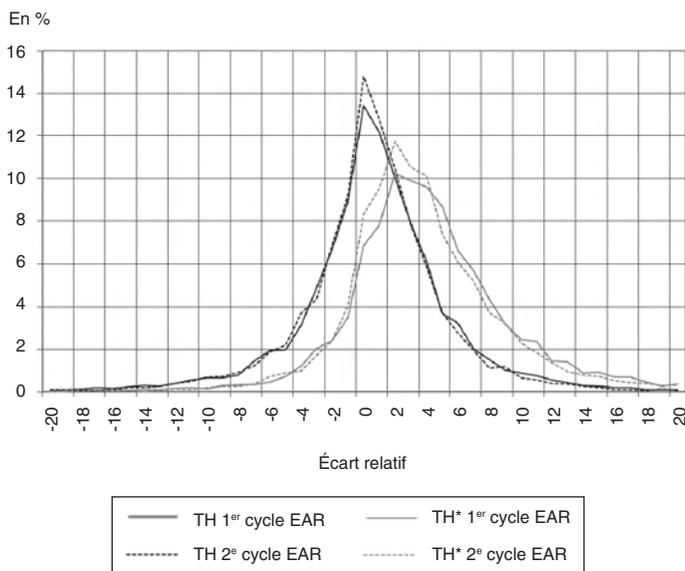


Figure a3 : 3^e groupe de rotation

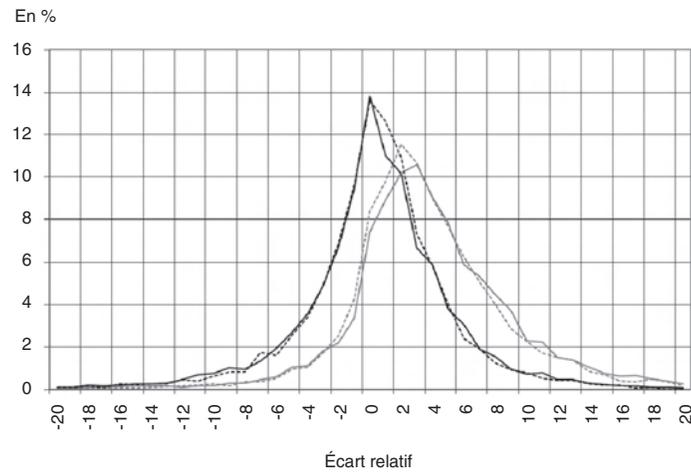
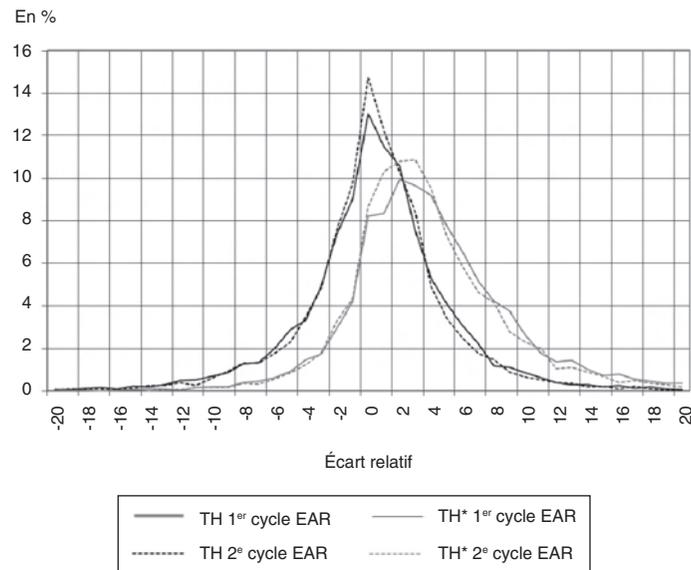


Figure a4 : 4^e groupe de rotation



Lecture : dans le 1^{er} groupe de rotation, la méthode TH usuelle conduit à estimer justement le bon nombre de logements pour 12 % des communes lors de l'EAR de 1^{er} cycle (soit en 2004). La méthode TH alternative sous-estime de 2 % le nombre de logements recensés en 2004 pour 2 % des communes du 1^{er} groupe de rotation.
 Champ : petites communes de France métropolitaine.
 Source : EAR, TH.

L'observation du tableau d confirme l'analyse des distributions des écarts relatifs. En considérant que l'estimation du stock de logements par un des filtres est proche du nombre de logements effectivement recensés si l'écart relatif n'excède pas 5 %, la proportion d'estimations proches est bien plus élevée avec la méthode usuelle qu'avec la méthode alternative.

Avec la méthode usuelle, environ 80 % des communes ont une estimation proche de la réalité. Cette proportion ne dépasse jamais 70 % avec la méthode alternative.

Quelles que soient la taille de la commune et l'année de recensement, la méthode usuelle permet toujours d'approcher le stock de logements dans une proportion plus importante de communes que la méthode alternative.

Par ailleurs, on remarque que la proportion de communes correctement estimées à 5 % augmente avec le temps quel que soit le filtre utilisé. L'amélioration est surtout notable entre le premier cycle de recensement renouvelé (2004 à 2007) et le deuxième (2009 à 2012).

Tableau d
Proportion de communes où l'écart entre le nombre de logements recensés et le nombre de logements comptabilisé avec la TH est inférieur ou égal à 5 %
Déclinaison par méthode, année de recensement et taille de commune

En %

Méthode	Année	GR	< 250 hab	[250 ; 500 [hab	[500 ; 1 000 [hab	> = 1 000 hab	Tous
TH	2004	GR1	64	80	84	86	77
TH	2005	GR2	67	83	87	86	79
TH	2006	GR3	67	81	87	86	79
TH	2007	GR4	68	83	87	86	80
TH	2009	GR1	69	84	87	89	81
TH	2010	GR2	71	84	88	89	82
TH	2011	GR3	71	83	90	89	82
TH	2012	GR4	70	87	90	89	82
TH*	2004	GR1	52	62	64	65	60
TH*	2005	GR2	55	67	66	67	63
TH*	2006	GR3	56	65	70	67	63
TH*	2007	GR4	57	67	70	69	64
TH*	2009	GR1	59	70	71	71	67
TH*	2010	GR2	60	70	73	72	68
TH*	2011	GR3	59	69	73	73	67
TH*	2012	GR4	61	71	75	73	69

Lecture : en 2010, pour les communes du 2^e groupe de rotation, la proportion de communes où l'écart entre le nombre de logements comptabilisés par la méthode TH usuelle et le nombre de logements recensés est inférieur à 5 % atteint 84 % pour les communes de 250 à moins de 500 habitants. Sur le même champ, la proportion de communes présentant un écart inférieur à 5 % entre la méthode TH alternative et le recensement n'est que de 70 %.

Champ : petites communes de France métropolitaine.

Source : EAR, TH.

La méthode TH alternative surestime le nombre de logements pour au moins deux raisons :

- Elle intègre les logements de type « communauté », alors qu'avec le filtre usuel on a le souci de les exclure (même si cela échoue parfois) ;
- Elle intègre des logements qui ne sont pas des logements d'habitation au sens du recensement de la popu-

lation, bien que la nature fiscale du local soit « appartement » ou « maison ».

En conséquence, même si le filtre alternatif répondait à l'exigence de la deuxième condition, il ne peut être considéré comme une alternative crédible au filtre actuellement utilisé dès lors qu'il ne satisfait pas à la première.

FORMAT DE TAILLE DE COMMUNE

La taille de la commune intervient dans l'analyse effectuée. Elle est définie ici par la population comptée dans les résidences principales lors du recensement de premier cycle d'enquête annuelle de recensement (soit 2004, 2005, 2006 et 2007 respectivement pour les 1^{er}, 2^e, 3^e et 4^e groupes de rotation). On regroupe les communes en quatre classes différentes selon le format suivant :

- moins de 250 habitants strictement ;
- entre 250 et 499 habitants ;
- entre 500 et 999 habitants ;
- 1 000 habitants ou plus.

Les classes ont été constituées de manière à être relativement homogènes en termes d'effectifs (cf. tableau e).

Tableau e

Répartition des communes selon leur population des ménages ordinaires lors de la 1^{re} Enquête annuelle de recensement

	En %			
	GR1	GR2	GR3	GR4
Moins de 250 habitants	34	34	34	34
Entre 250 et 499 habitants	22	22	22	23
Entre 500 et 999 habitants	20	20	20	19
1 000 habitants ou plus	24	24	24	24

Lecture : 22 % des communes du 1^{er} groupe de rotation avaient une population des ménages ordinaires comprise entre 250 et 499 habitants lors de la première Enquête annuelle de recensement (c'est-à-dire en 2004 pour ce groupe de rotation).

Champ : petites communes de France métropolitaine.

Source : EAR.

Parce que la population en résidences principales constitue l'essentiel de la population municipale, on ne

précisera pas toujours la nature exacte de la population lorsqu'on évoquera la taille de la population dans l'étude.

ANNEXE 6

**LA PROPORTION DE COMMUNES POUR LESQUELLES LES MÉTHODES TH ET PL
SONT ÉQUIVALENTES DÉCROÎT AVEC LA TAILLE**

Plus la commune est petite, plus il est fréquent que l'estimation du nombre de résidences principales par la méthode TH soit équivalente à l'estimation de la méthode PL (cf. tableau f).

Tableau f

Proportion de communes pour lesquelles les méthodes TH et PL sont équivalentes

En %

	GR1	GR2	GR3	GR4
Moins de 250 habitants	14	12	14	14
Entre 250 et 499 habitants	7	7	6	8
Entre 500 et 999 habitants	4	3	5	4
1 000 habitants ou plus	2	2	2	2
Tous	7	7	7	8

Lecture : la part de communes de 500 à 999 habitants du 2^e groupe de rotation pour lesquelles les estimations par les méthodes TH et PL sont équivalentes est de 3 %. Pour l'ensemble des communes de ce groupe de rotation, la proportion d'équivalence des méthodes grimpe à 7 %

Champ : petites communes de France métropolitaine estimées par extrapolation TH.

Source : EAR, TH.

LA FRAGILITÉ DE L'INDICATEUR D'ÉCART RELATIF POUR LES PLUS PETITES COMMUNES

On a privilégié l'écart relatif pour plusieurs raisons :

- Le coefficient TH est un rapport entre des nombres de logements ; une approche relative semble donc appropriée ;
- Il faut un écart relatif très important dans une petite commune pour que cela ait des conséquences lourdes en termes d'extrapolation de la population. Que cet écart relatif dépasse le seuil de 5 % au bout de cinq ans d'extrapolation ne signifie pas pour autant une dégradation importante de l'estimation, ni même

une dégradation lors des années où la commune était estimée par extrapolation.

Dans l'étude, on utilise donc le plus souvent l'écart relatif comme indicateur principal de l'apport de la TH ou comme indicateur de comparaison entre les méthodes testées. Sa variabilité est plus forte dans les plus petites communes, car l'inclusion ou l'exclusion à tort d'un logement a beaucoup plus de conséquences (relativement au nombre de logements) que dans une commune de plus grande taille (cf. tableau g).

Tableau g

Nombre moyen de logements TH par commune en 2004 par taille de commune dans les communes du 1^{er} groupe de rotation et impact relatif d'un logement sur le stock de logements

	Nombre moyen de logements TH en 2004	Poids relatif d'un logement sur le stock moyen de logements En %
Moins de 250 habitants	85	1,2
Entre 250 et 499 habitants	186	0,5
Entre 500 et 999 habitants	347	0,3
1 000 habitants ou plus	1 268	0,1

Lecture : dans les communes du 1^{er} groupe de rotation dont la population est comprise entre 500 et 1 000 habitants, le nombre moyen de logements s'élève à 347. Un logement représente 0,3 % de ce nombre.

Champ : petites communes de France métropolitaine du 1^{er} GR.

Source : EAR, TH.

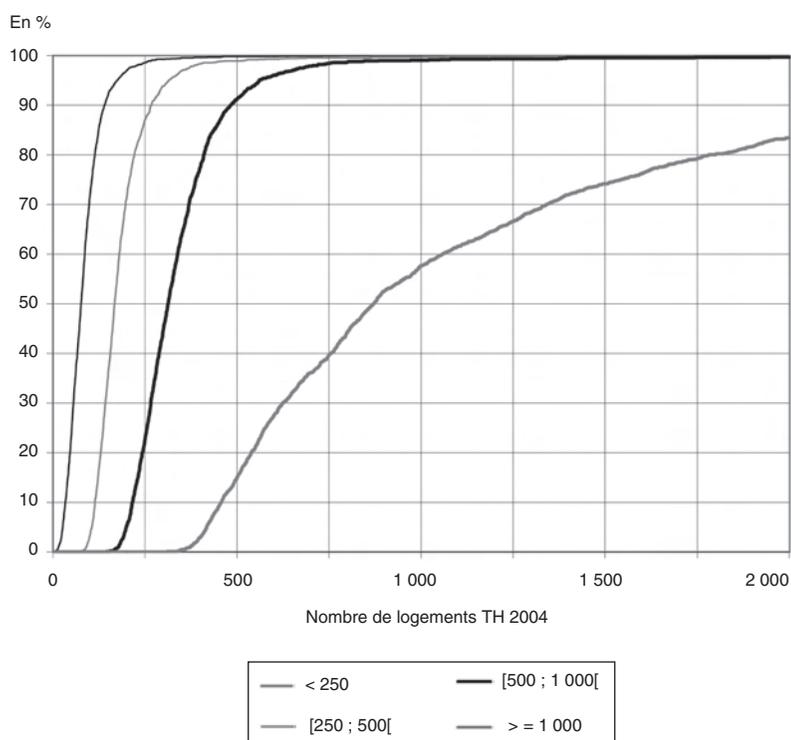
Dans une commune de moins de 250 habitants du 1^{er} groupe de rotation, il y a en moyenne 85 locaux considérés comme des logements dans la TH 2004. L'impact relatif d'un logement exclu ou inclus à tort sur un nombre de 85 logements est de 1,2 %. Pour les communes de 1 000 habitants ou plus, la moyenne du nombre de logements comptabilisés dans la TH est située à 1 268. L'impact relatif d'un logement exclu ou inclus à tort sur le stock de logements moyen est cette fois de 0,1 % seulement.

Logiquement, les communes peu peuplées comptent un nombre de logements restreint dans la TH (cf. figure b) : plus elle est peuplée, plus elle compte de logements.

La dispersion de l'écart relatif est plus importante dans les plus petites des communes. En effet, l'écart-type relatif y est de 7,5 contre moins de 5 pour les communes de 500 habitants ou plus (cf. tableau h).

La situation aurait été inversée si on avait considéré l'écart absolu comme indicateur principal de la qualité d'estimation de la TH. La dispersion de cette variable est plus importante dans les communes de grande taille. L'écart-type n'est que de 10 logements dans les communes de moins de 250 habitants du 4^e groupe de rotation et culmine à 81 logements dans les communes de 1 000 habitants ou plus de ce même groupe.

Figure b
Répartition du nombre de logements comptabilisés dans la TH en 2004 pour les communes du 1^{er} groupe de rotation



Lecture : 40 % des communes du 1^{er} groupe de rotation de 1 000 habitants ou plus ont un nombre de logements TH en 2004 inférieur à 750. Environ 10 % des communes du 1^{er} groupe de rotation de 500 à 1 000 habitants ont un nombre de logements TH en 2004 supérieur à 500.

Champ : petites communes de France métropolitaine du 1^{er} GR.
Source : EAR, TH.

Tableau h
Variabilité des écarts absolu et relatif entre le nombre de logements comptabilisés avec la TH en 2007 et le nombre de logements recensés par taille de commune du 4^e groupe de rotation

	Écart relatif		Écart absolu	
	Moyenne	Écart-type	Moyenne	Écart-type
Moins de 250 habitants	0,2	7,5	- 1	10
Entre 250 et 499 habitants	0,2	5,1	- 0	14
Entre 500 et 999 habitants	0,4	4,5	- 0	28
1 000 habitants ou plus	1,4	4,1	19	81

Lecture : en moyenne, la TH comptabilise 19 logements de plus que le nombre recensé pour les communes de 1 000 habitants ou plus ; l'écart-type est de 81 logements. L'écart relatif moyen est de 1,4 % et l'écart-type relatif de 4,1 %.

Champ : petites communes de France métropolitaine du 4^e GR.
Source : EAR, TH.

