

**Essai d'évaluation d'un effet du montant de
l'aide sur les prix des logements acquis par des
bénéficiaires du Prêt à Taux Zéro**

**Olivier Dupré
Maryem Saissi**



Mai 2017

Résumé

Cette étude est une contribution au débat sur l'effet inflationniste éventuel des aides publiques au logement sur les prix, et plus particulièrement sur celui du prêt à taux zéro (PTZ). Elle a été rendue possible par la mise à disposition par la DGALN au Cerema d'une base de données issue du SGFGAS¹ qui fournit pour chaque transaction ayant été financée par un PTZ, un NPTZ ou un PTZ+ entre le 1^{er} octobre 1995 et 17 mars 2014 une série de variables relatives au bien objet de la transaction, à l'emprunteur et aux caractéristiques de l'opération.

Après une tentative infructueuse de rapprochement de ces données avec les bases notariales (BIEN et PERVAL), le Cerema a été contraint de comparer les transactions ayant fait l'objet d'un PTZ entre elles. Nous nous sommes donc efforcé de mesurer l'effet toutes choses égales par ailleurs de l'importance de l'aide accordée au ménage (estimée grâce à la variable relative à l'équivalent subvention) sur le montant total de la transaction. Pour ce faire, nous avons mis en œuvre un modèle complet de prix à partir de la méthode des moindres carrés ordinaires.

En introduisant dans le modèle le plus grand nombre possible de variables relatives au logement lui-même et à son environnement, à sa localisation, mais aussi dans certains cas aux caractéristiques des acheteurs, la modélisation permet, à certaines conditions, d'isoler le rôle de l'aide accordée dans la détermination du prix, en éliminant les effets de structure liés aux autres variables. Les modélisations sont réalisées sur des périodes pendant lesquelles le dispositif PTZ est resté relativement inchangé.

Une seule des modélisations que nous avons réalisées (entre 1995 et 2005, toutes transactions et toutes variables confondues) aboutit à un effet inflationniste relativement faible de l'équivalent subvention : le passage de l'équivalent subvention de 5 000 à 10 000 euros conduit selon ce modèle à une croissance de 0,94 % du montant total de l'opération. Toutes les autres modélisations aboutissent à un effet déflationniste. Si on constate une certaine sensibilité des résultats aux choix de paramétrage des modèles, il demeure qu'**aucun modèle testé n'a montré un effet inflationniste fort. Nos résultats, s'ils soulèvent des difficultés d'interprétation et présentent des limites, permettent de conclure à l'absence d'un effet inflationniste net du montant de l'aide sur les prix des logements acquis avec un PTZ.**

Nos travaux ne permettent toutefois pas de conclure sur l'existence d'un effet inflationniste du prêt à taux zéro. En effet, nous n'avons pas pu comparer les transactions ayant fait l'objet d'un prêt à taux zéro à celles financées sans cette aide publique. Une piste à explorer pour aller dans ce sens pourrait être de chercher à identifier, parmi les transactions rassemblées dans la base de données DVF² celles qui ont été financées par un prêt à taux zéro, en rapprochant cette dernière base des données fournies par le SGFGAS.

1 Société de Gestion des Financements et de la Garantie de l'Accession Sociale à la propriété

2 Depuis 2011, les collectivités locales, l'État, et certaines structures publiques peuvent bénéficier des données « Demande de Valeur Foncière » (DVF) de la Direction Générale des Finances Publiques (DGFIP). Ces données gratuites, et actualisées tous les 6 mois, recensent de manière exhaustive toutes les transactions foncières et immobilières onéreuses sur le territoire français à l'exception de l'Alsace, de la Moselle (soumises au livre foncier) et de Mayotte.

Sommaire

Introduction.....	9
Éléments de méthode.....	11
Résultats.....	19
Discussion.....	47
Conclusion et perspectives.....	51
Références bibliographiques.....	53

Introduction

Le prêt taux zéro, des politiques d'aide à l'accession

C'est le décret n° 95-1064 du 29 septembre 1995 qui a créé cette nouvelle aide de l'État à l'accession à la propriété pour l'acquisition d'une résidence principale, sous forme d'avance remboursable sans intérêt. Le décret n° 97-1000 du 30 octobre 1997 a restreint l'accès du PTZ aux « primo-accédants », c'est-à-dire aux personnes n'ayant pas été propriétaires de leur résidence principale au cours des deux dernières années précédant l'offre de prêt³.

Pour bénéficier d'un prêt à taux zéro, les emprunteurs doivent acquérir ou faire construire un logement destiné à être occupé à titre de résidence principale, pendant au moins huit mois dans l'année. Il peut s'agir selon les périodes d'un logement neuf ou ancien (cf. tableau ci-dessous). S'il s'agit d'un logement neuf, l'acquéreur peut financer auprès d'un promoteur son achat ou la construction d'une maison, ainsi que l'acquisition du terrain sur lequel elle est construite.

Les bénéficiaires doivent, sauf en 2011, respecter un plafond de ressources qui varie en fonction de la composition du ménage et de la localisation du logement. Le barème du prêt à taux zéro distingue plusieurs tranches de revenu imposable qui déterminent les caractéristiques du remboursement du prêt. Le montant du prêt est limité à un pourcentage du montant de l'opération hors frais de notaire. En outre, le prêt ne peut dépasser des montants maximaux qui varient, comme les plafonds de ressources, en fonction de la composition du ménage et de la localisation géographique du futur logement. Le montant du prêt est, de plus, limité à un pourcentage du total des autres prêts souscrits pour l'opération⁴. De ce fait, un prêt à taux zéro doit donc être obligatoirement accompagné d'au moins un autre prêt.

Le prêt à taux zéro est remboursé plus ou moins rapidement, en fonction du niveau des revenus fiscaux de référence du ménage. Les durées sont d'autant plus longues que ces revenus sont modestes, les plus basses tranches de revenus bénéficiant en outre d'un différé d'amortissement.

Nom du dispositif	Période	Partition du territoire	Conditions de ressources	Ouverture à l'ancien	Précisions
PTZ	De 1995 à 2005	Province / Île-de-France	oui	non	Ouvert à tous les accédants avant 1997 et ouvert à l'ancien mais sous condition d'un montant minimum de travaux représentant 35 % au moins du montant global de l'opération
NPTZ	De 2005 à 2010	3 zones	oui	oui	Doublement du montant du NPTZ pour les logements neufs à partir de 2009, dans le cadre du plan de relance de l'économie. Cumul possible, entre mi-2009 et fin 2010, du PTZ et du Pass-Foncier pour l'achat d'un logement collectif neuf.
PTZ+	2011	3 zones	non	oui	Le revenu fiscal de référence du ménage influence toutefois les conditions de remboursement du prêt. De plus, le PTZ+ prend en compte la performance énergétique du logement.
PTZ+ dit « recentré »	De 2012 à 2014	3 zones	oui	non	L'ouverture à l'ancien n'est possible que pour les logements anciens du parc social HLM vendus à leurs occupants. À compter du 1 ^{er} janvier 2013, le PTZ+ ne peut pas être accordé à un logement neuf si celui-ci ne respecte pas une condition de performance énergétique

Un dispositif qui a connu principalement 4 périodes entre 1995 et 2014

3 Le décret n° 2000-104 du 8 février 2000 a introduit deux exceptions à la règle de primo-accession : l'acquisition d'un logement adapté pour une occupation par une personne handicapée et le relogement des propriétaires victimes de catastrophes naturelles.

4 Les prêts pris en compte pour l'application de cette quotité doivent avoir une durée supérieure ou égale à deux ans, ce qui exclut donc les prêts relais.

L'effet inflationniste des aides au logement : une question récurrente

La question de l'effet inflationniste éventuel du prêt à taux zéro a déjà été posée à plusieurs reprises (Gobillon et Le Blanc 2005; Delarue, Lecomte, et Vergriete 2013; Beaubrun-Diant et Maury 2015; Ihmaïne, Snoussi et Souche 2017). Elle s'inscrit dans le cadre plus général des controverses sur l'effet inflationniste des aides au logement en général (Fack 2005; Laferrère et Le Blanc 2002; Bouteille 2015; Cavailhès 2015).

Pour apporter des éléments à ce débat, le Cerema et plus particulièrement son pôle « Marchés Locaux de l'Habitat » a réalisé la présente étude visant à identifier un éventuel inflationniste de la politique publique « prêt à taux zéro »⁵.

Cette étude a été rendue possible par la mise à disposition du Cerema, par la DGALN, d'une base de données issue du SGFGAS⁶ qui fournit pour chaque transaction ayant été financée par un PTZ, un NPTZ ou un PTZ+ entre le 1^{er} octobre 1995 et 17 mars 2014 une série de variables relatives au bien objet de la transaction, à l'emprunteur et aux caractéristiques de l'opération.

Le présent document présente la méthode mise en œuvre par le Cerema et les résultats obtenus, qui sont ensuite discutés.

5 On recouvre sous cette expression les dispositifs d'aide à l'accèsion mis en place entre 1995 et aujourd'hui sous les noms de PTZ, NPTZ ou encore PTZ+.

6 Société de Gestion des Financements et de la Garantie de l'Accession Sociale à la propriété

Éléments de méthode

Une première tentative qui échoue

Pour identifier un éventuel effet inflationniste du dispositif « prêt à taux zéro », le Cerema a d'abord tenté de disposer d'une base de données regroupant à la fois des transactions financées par un PTZ, un NPTZ ou un PTZ+ – et identifiables comme telles – et des transactions qui n'auraient pas fait l'objet d'une de ces aides.

Les seules bases de données disponibles au début de l'étude étaient les bases notariales PERVAL (en province) et BIEN (en Île-de-France)⁷. Le Cerema a, dans une première étape, cherché à identifier, à partir de variables communes entre ces fichiers, les transactions du PTZ au sein des bases BIEN et Perval. Les variables communes entre les deux fichiers étaient notamment : la commune du logement, le type d'immeuble (appartement ou maison), le nombre de pièces, l'année de naissance de l'acheteur (de l'emprunteur pour les données SGFGAS), la date de la transaction (ou date de mise en force du prêt pour les données SGFGAS), le montant de la transaction ou encore le montant des prêts. Malheureusement, il n'a pas été possible d'identifier avec un niveau de confiance suffisant les transactions. En effet, nous avons constaté un trop grand nombre de transactions ayant bénéficié d'un PTZ pour lesquelles aucune donnée PERVAL ou BIEN ne pouvait correspondre. Inversement, certaines transactions ayant bénéficié d'un PTZ pouvaient correspondre, selon qu'on s'appuie sur l'une ou l'autre des variables communes aux deux bases, à plusieurs transactions dans BIEN ou PERVAL, sans qu'il soit possible de trancher entre elles avec suffisamment de certitude.

Étudier l'effet du montant de l'aide sur le prix des transactions ayant fait l'objet d'un PTZ

Constatant l'impossibilité de disposer d'une base qui permette de comparer les transactions ayant fait l'objet d'un PTZ des autres, le Cerema a choisi de comparer les transactions ayant fait l'objet d'un PTZ entre elles. En effet, la base de données SGFGAS fournit des informations qui concernent le montant de l'aide reçue : il s'agit en particulier de la variable équivalent subvention (EQUIV_SUBV) à propos de laquelle nous revenons par la suite. Il a donc été envisagé de réaliser une modélisation à partir de la base SGFGAS et d'étudier l'influence de cette variable, toutes choses égales par ailleurs, sur le prix des transactions. Les résultats permettraient de conclure, non pas sur l'effet inflationniste des différentes versions du dispositif « prêt à taux zéro », mais sur la présence ou l'absence d'un effet inflationniste du montant de l'aide sur les prix des logements acquis avec un PTZ. Le Cerema a donc recouru à une modélisation des prix des logements acquis avec l'aide d'un PTZ. À la différence d'autres travaux ayant procédé par double différence, nous tentons de constituer un modèle complet de prix à partir de la méthode des moindres carrés ordinaires.

Qu'est-ce qu'une modélisation des prix ?

L'objet d'un modèle est de construire une fonction qui représente au mieux une variable « à expliquer » en fonction de n variables « explicatives ». La modélisation appliquée aux prix immobiliers consiste à établir une fonction expliquant le rôle de différentes variables relatives au logement dans la constitution du prix. Elle s'appuie sur un ensemble d'observations, ici des transactions.

Pour chaque observation (chaque vente de logement), il faut disposer d'un ensemble de k variables indépendantes (surface par exemple) et d'une variable à expliquer (le prix). Si on représente ces transactions sous la forme d'un nuage de points, chaque point est défini par $k+1$ coordonnées dans un espace à $k+1$ dimensions. La modélisation ou « régression linéaire » consiste alors à déterminer la fonction dont la courbe à k dimensions va passer au plus près des points du nuage.

La régression linéaire vise à réduire au maximum les écarts entre les prix « prédits » par le modèle et les prix observés. Elle est effectuée grâce à la méthode des moindres carrés ordinaire : il s'agit de trouver le modèle qui minimise la somme des carrés des écarts aux valeurs observées, sur l'ensemble des observations.

⁷ Désormais, la base de données DVF (Demandes de Valeurs Foncières) est disponible et une des perspectives de la présente étude (qui sera rappelée en conclusion) serait de tenter une identification dans cette nouvelle base de données des transactions qui ont fait l'objet d'un prêt à taux zéro.

La modélisation permet aussi, à certaines conditions, d'isoler le rôle d'un paramètre (d'une variable explicative) dans la détermination du prix, en éliminant les effets de structure liés aux autres variables.

Première étape : appliquer les filtres préconisés par le SGFGAS

Tout d'abord, la base a été épurée en appliquant certaines recommandations du SGFGAS. Cela a conduit à ne retenir que les observations pour lesquelles :

- le code de statut de l'enregistrement (CSEN) est 'V' (validé)
- le revenu net total pour l'année N (RANN) est strictement supérieur au revenu imposable (RANI)
- le revenu annuel (RANN) est inférieur à 600 000 €
- le nombre de pièces (NPPR) est inférieur ou égal à 10
- la surface habitable (SURH) est comprise entre 14 m² et 350 m²
- le nombre de personnes (NPER) est inférieur ou égal à 10 ;
- le montant de l'opération (VTTO) est compris entre 7 600 € et 450 000 €
- le montant du prêt principal (VTPP) est compris entre 762 € et 350 000 €

D'une base comprenant 2 898 202 observations, nous sommes ainsi passé à une base de 2 412 356 observations. Ces filtres ont donc conduit à la suppression d'un peu moins de 500 000 transactions ce qui représente de l'ordre de 17 % de la base initiale.

Nous avons également corrigé deux variables conformément aux recommandations du SGFGAS :

- lorsque l'emprunteur est marié ou vivant maritalement alors qu'une seule personne est déclarée, nous avons rectifié la variable du nombre de personnes (NPER) à 2.
- Pour les opérations de construction avec ou sans l'achat du terrain (TOPE = 2 ou 3), lorsque le type d'immeuble (TIMM) est collectif, nous l'avons rectifié à individuel.

Deuxième étape : fiabiliser la base

La base SGFGAS que nous avons utilisé contient 2 450 354 observations. 865 d'entre elles sont inutilisables, car la plupart des variables les concernant ne sont pas renseignées. De plus, après de multiples opérations de redressement de la variable relative au code INSEE de la commune du logement objet de la transaction, il demeure 37 133 observations pour lesquelles cette variable n'est pas renseignée. Enfin, nous limitons nos analyses à la France métropolitaine⁸, ce qui nous conduit à exclure 15010 observations. Nous avons donc au total exclu de la base 37 133 + 865 + 15010 = 53 008 observations ce qui représente 2,16 % de l'ensemble. Pour les 2 397 346 observations que nous avons exploitées, toutes les variables utilisées dans la modélisation sont renseignées à 100 %.

Troisième étape : choisir la variable à expliquer

La variable à expliquer est construite à partir de la variable relative au montant total de l'opération TCC (VTTO). Le coût total de l'opération comprend :

- la charge foncière ou la charge immobilière, les honoraires de géomètre et les taxes afférentes ;
- [à compter du 1er juin 2011], les coûts d'aménagement et de viabilisation du terrain (et honoraires afférents) ;
- les honoraires de négociation (les frais d'agence) ;
- le coût des travaux, y compris les honoraires liés à leur réalisation ([à compter du 1er juin 2011] et à leur certification), à l'exception des montants financés par un éco prêt à taux zéro ;
- les frais relatifs à l'assurance responsabilité / assurance dommages (L241-1 et L242-1 du code des assurances) ;
- [jusqu'à 2013] diverses taxes afférentes à la construction (taxe locale d'équipement et autres mentionnées aux articles 1585 A, 1599 octies, 1599 B et 1599-0 B du code général des

⁸ Certaines variables n'étaient d'ailleurs pas disponibles pour les DOM comme le zonage GTC_2011.

impôts ainsi qu'à l'article L.142-2 du code de l'urbanisme) ;

- [à compter de 2014] la taxe d'aménagement mentionnée à l'article L.331-1 du code de l'urbanisme.

Le coût total d'opération comprend, le cas échéant, la valeur d'acquisition du terrain destiné à la construction, acquis depuis moins de deux ans, à la date d'émission de l'offre de prêt.

En revanche, le coût total ne comprend pas :

- les frais d'acte notarié ;
- les droits d'enregistrement pour les terrains à bâtir ou les immeubles anciens.

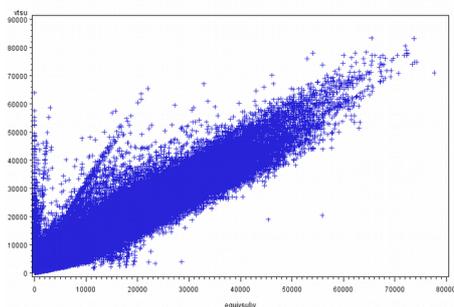
Quatrième étape : choisir la variable explicative permettant de mesurer l'effet du dispositif PTZ sur la capacité d'achat des ménages bénéficiaires

La variable EQUIV_SUBV est retenue pour évaluer l'effet du dispositif PTZ sur le prix des logements. Cette variable est calculée. Elle est supposée égale au supplément d'endettement qu'a permis la mise en place du PTZ, toutes choses égales par ailleurs (c'est-à-dire notamment pour un taux d'effort constant par rapport à la situation où le PTZ n'existerait pas), le coût de la ressource pour l'emprunteur étant supposé être le taux d'intérêt du prêt principal supposé unique et de durée au moins égale à celle du PTZ.

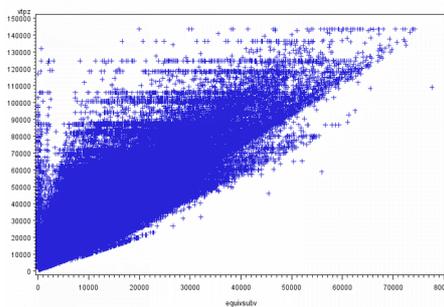
Cette variable a été préférée à la variable relative au montant du prêt à taux zéro (VTPZ) ainsi qu'à la variable relative au montant de la subvention ou du crédit d'impôt (VTSU). En effet, en renseignant le supplément d'endettement rendu possible par le PTZ à taux d'effort constant, la variable EQUIV_SUBV est représentative de l'effet du PTZ sur la capacité d'emprunt supplémentaire, donc à la capacité d'achat supplémentaire en matière de prix. Elle tient compte à la fois du montant du PTZ et de l'effet d'un éventuel différé de remboursement.

Les graphiques ci-dessous montrent la valeur des variables VTSU et VTPZ en fonction de la variable EQUIV_SUBV.

- Le premier graphique montre que les variables VTSU et EQUIV_SUBV sont corrélées pour une large majorité des transactions, sans que l'une soit systématiquement supérieure à l'autre. Par exemple, pour un équivalent subvention de 30 000 euros, le montant de la subvention ou du crédit d'impôt varie essentiellement entre 20 000 et 40 000 euros. Il n'existe pas de transactions pour lesquelles EQUIV_SUBV est élevé alors que VTSU ne l'est pas. Par contre, certaines transactions, relativement peu nombreuses, se caractérisent par une variable VTSU élevée et un équivalent subvention faible.
- Le second graphique montre une certaine corrélation entre les variables VTPZ et EQUIV_SUBV. La variable VTPZ prend presque toujours des valeurs supérieures à la variable EQUIV_SUBV. Par exemple pour un montant de PTZ de 30 000 euros, la variable EQUIV_SUBV varie entre 0 et 25 000 euros.



Variable VTSU en fonction de la variable EQUIV_SUBV



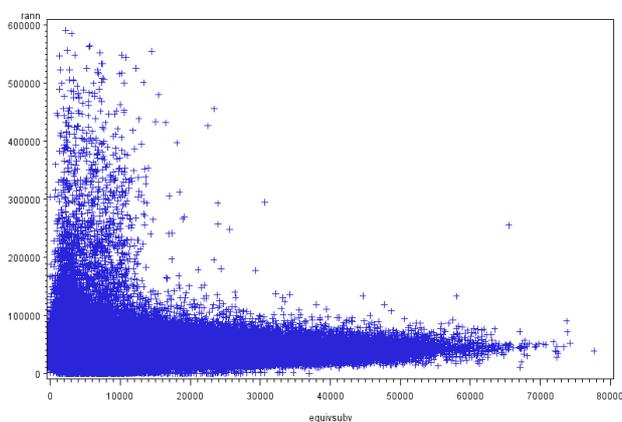
Variable VTPZ en fonction de la variable EQUIV_SUBV

Les graphiques ci-dessous montrent la valeur des variables RANN et VTTO en fonction de la variable EQUIV_SUBV.

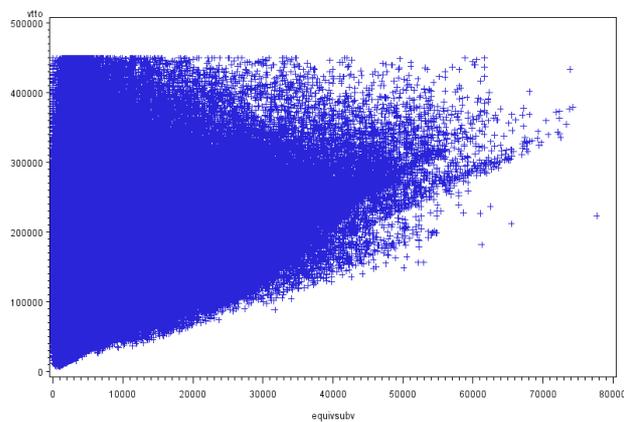
- Le premier graphique montre que les valeurs élevées de RANN (par exemple supérieures à 150 000 euros) correspondent presque toujours à des équivalents subvention relativement faibles. En revanche, un faible revenu ne signifie pas toujours un équivalent subvention élevé.

On observe par exemple des valeurs de EQUIV_SUBV inférieures à 10 000 euros pour un grand nombre de bénéficiaires dont les revenus sont compris entre 0 et 100 000 euros.

- Le second graphique montre que l'EQUIV_SUBV varie fortement pour un même montant d'opération. Pour un prix de 200 000 euros, l'équivalent subvention varie par exemple de 0 à 50 000 euros. S'il y a des valeurs de VTTO élevées associées à un équivalent subvention faible, l'inverse n'est pas observé.



Variable RANN en fonction de la variable EQUIV_SUBV



Variable VTTO en fonction de la variable EQUIV_SUBV

Cinquième étape : choisir les variables explicatives relatives au logement et à son environnement

Dans un premier temps, nous avons retenu uniquement comme variables explicatives des variables relatives au logement qui est l'objet de la transaction et à sa localisation :

- Type d'opération (TOPE). Cette variable comprend les 7 modalités suivantes :
 1. achat neuf – 11,1 % des observations
 2. construction de maison individuelle (achat du terrain compris) – 32,2 % des observations
 3. construction de maison individuelle (hors achat du terrain) – 11,0 % des observations
 4. acquisition / amélioration – 45,6 % des observations
 5. transformation de locaux non destinés à l'habitation en logement
 6. levée d'option en location-accession pour le 1^{er} locataire
 7. acquisition d'un logement avec travaux de remise à neuf TVA

En raison du très faible nombre d'observations pour lesquelles cette variable est égale à 5, 6 ou 7 (0,08 % des transactions), nous avons regroupé ces trois catégories en une seule appelée « autres ».
- Type d'immeuble (TIMM). Cette variable comprend les 2 modalités suivantes :
 1. individuel – 75,7 % des observations
 2. collectif – 24,3 % des observations
- Nombre de pièces principales (NPPR). Cette variable quantitative est comprise entre 1 et 10
- Surface habitable (SURH) en m².
- Performance énergétique du logement (PELG). Cette variable n'est utilisée que pour les transactions pour lesquelles la date d'émission de l'offre de prêt est postérieure au 1^{er} janvier 2011. Nous ne l'utilisons que lorsque nous travaillons sur le segment du neuf (TOPE = 1, 2 ou 3). Elle comprend alors 2 modalités : 0 (non BBC) et 1 (BBC).
- Code INSEE (CINS) de la commune de localisation du logement. Cette variable a fait l'objet de traitements pour améliorer son taux de remplissage et fiabiliser son contenu, par rapprochement avec les variables LCOM (libellé commune) et CPFL (code postal du futur logement). Elle n'est pas utilisée en tant que telle dans les modélisations mais nous sert à construire d'autres variables (cf. quatrième étape).
- Zone Robien (ZOCO). Cette variable n'est pas renseignée pour les offres émises avant le 31 janvier 2005. Nous ne l'utilisons donc pas sur cette période. Entre le 1^{er} février 2005 et le 31

décembre 2010, elle comprend trois modalités (A, B et C). À partir du 1^{er} janvier 2011, elle comprend quatre modalités (A, B1, B2 et C). Il s'agit du zonage ayant cours au moment de l'émission de l'offre de prêt et non d'un zonage donné à une date fixe.

Sixième étape : choisir les variables explicatives relatives aux caractéristiques de l'acheteur

Nous avons également réalisé certaines modélisations en utilisant, en plus des variables relatives au logement et à son environnement, des variables explicatives relatives aux caractéristiques de l'acheteur :

- État matrimonial du ménage emprunteur (EMEM). Cette variable comprend les 5 modalités suivantes :
 1. marié – 38,9 % des observations
 2. veuf – 0,4 % des observations
 3. divorcé ou séparé – 4,9 % des observations
 4. célibataire – 29,6 % des observations
 5. vivant maritalement – 26,2 % des observations
- Nombre de personnes (NPER) du ménage
- Code CSP de l'emprunteur (CCSP). Cette variable comprend les 8 modalités suivantes :
 1. agriculteurs exploitants – 1,1 % des observations
 2. artisans, commerçants et chefs d'entreprise – 2,8 % des observations
 3. cadres et professions intellectuelles supérieures – 12,1 % des observations
 4. professions intermédiaires – 21,8 % des observations
 5. employés – 30,8 % des observations
 6. ouvriers – 28,4 % des observations
 7. retraités – 0,7 % des observations
 8. autres personnes sans activité professionnelle – 2,2 % des observations
- Situation professionnelle de l'emprunteur (SPRE). Cette variable comprend les 6 modalités suivantes :
 1. salarié secteur privé (CDI) – 58,1 % des observations
 2. salarié secteur privé (autres cas) – 14,1 % des observations
 3. salarié secteur public – 16,7 % des observations
 4. non salarié – 9,0 % des observations
 5. chômeur – 0,3 % des observations
 6. inactif – 1,8 % des observations
- Revenu net total du ménage à l'année n (RANN). Il s'agit du montant pris en compte pour le calcul du taux d'effort x 12, primes et compléments de rémunération inclus.

Septième étape : enrichir la liste des variables

À ces variables présentes dans la base, nous en avons ajouté d'autres, calculées :

- Année de l'offre de prêt (ANNEE_DATP). Cette variable est calculée à partir de la variable relative à la date d'émission de l'offre de prêt à taux zéro (DATP).
- Saisonnalité (SAISONNALITE). Cette variable comporte 4 modalités et correspond au trimestre de la date d'émission de l'offre de prêt (DATP) : T1 pour une date comprise entre le 1^{er} janvier et le 31 mars, T2 pour une date comprise entre le 1^{er} avril et le 30 juin, T3 pour une date comprise entre le 1^{er} juillet et le 30 septembre et T4 pour une date comprise entre le 1^{er} octobre et le 31 décembre.
- Âge de l'emprunteur (AGE_EMPR). Cette variable est calculée par différence entre la variable année de naissance de l'emprunteur (ANEM) et la variable année de l'offre de prêt (ANNEE_DATP).
- Surface moyenne par pièce (SURH_N). Cette variable est calculée en rapportant la surface habitable (SURH) au nombre de pièces (NPPR).
- Zonage de tension (GTC_2011). Cette variable, établie à partir du code INSEE (CINS), indique dans quelle catégorie de tension sur les marchés, selon le zonage établi par Guy Taieb pour l'Anah en 2011 à partir des zones d'emploi 2010, se trouve la commune de

localisation du logement. Elle comprend 6 modalités correspondant aux 6 niveaux de tension sur les marchés.

- Zonage en aires urbaines (CODE_GÉOGRAPHIQUE). Cette variable, établie à partir du code INSEE (CINS), indique dans quelle catégorie du zonage en aires urbaines⁹ se trouve la commune de localisation du logement. Elle comprend 10 modalités :
 - 11 : commune donnant son nom à un grand pôle (10 000 emplois ou plus)
 - 111 : commune appartenant à un grand pôle sans lui donner son nom
 - 112 : commune appartenant à la couronne d'un grand pôle
 - 120 : commune multipolarisée des grandes aires urbaines
 - 211 : commune appartenant à un moyen pôle (5 000 à moins de 10 000 emplois)
 - 212 : commune appartenant à la couronne d'un moyen pôle
 - 221 : commune appartenant à un petit pôle (de 1 500 à moins de 5 000 emplois)
 - 222 : commune appartenant à la couronne d'un petit pôle
 - 300 : autre commune multipolarisée
 - 400 : commune isolée hors influence des pôles
- Zonage littoral (ZONAGE_LITTORAL). Cette variable, établie à partir du code INSEE (CINS). Elle comprend 6 modalités :
 - E : communes d'estuaire (au sens de la loi littoral)
 - L4 : communes littorales (au sens de la loi littoral)
 - L3 : communes rétrolittorales situées à moins de 5 km du littoral¹⁰ (approche géomatique)
 - L2 : communes rétrolittorales situées à moins de 10 km du littoral (approche géomatique)
 - L1 : communes rétrolittorales situées à moins de 15 km du littoral (approche géomatique)
 - N : autres communes

Huitième étape : modéliser sur des périodes relativement homogènes du dispositif

Nous avons fait le choix de tester l'effet inflationniste du PTZ (et donc l'effet de la variable EQUIV_SUBV sur le prix) en distinguant les périodes suivantes :

- Période 1 (PTZ). Elle regroupe l'ensemble des transactions pour lesquelles la date d'émission de l'offre de prêt est comprise entre le 1^{er} octobre 1995 et le 31 janvier 2005. Si le dispositif a connu quelques évolutions sur cette période (il ne devient réservé aux primo-accédants¹¹ qu'à partir de novembre 1997 par exemple), ses caractéristiques sont restées grosso modo les mêmes. Le PTZ est réservé en priorité aux logements neufs achetés par des ménages modestes (plafonds de revenus selon la taille de la famille) et primo-accédants. Il est également ouvert à l'ancien mais sous condition d'un montant minimum de travaux représentant 35 % au moins du montant global de l'opération. Les banques recevaient alors une subvention de l'État pour compenser l'absence de perception d'intérêts.
- Période 2 (NPTZ). Elle regroupe l'ensemble des transactions pour lesquelles la date d'émission de l'offre de prêts est comprise entre le 1^{er} février 2005 et le 31 décembre 2010. Le dispositif est alors étendu à tous les logements anciens sans condition de travaux. L'aide est versée sous forme d'un crédit d'impôt aux établissements de crédit. C'est aussi à partir de cette date qu'une partition du territoire en trois zones (A, B et C) est instituée. Selon la zone où est situé le logement objet de la transaction, les conditions d'attribution du prêt à taux zéro varient. Là encore, le dispositif a connu des évolutions sur cette période, notamment le doublement du montant du NPTZ pour les logements neufs à partir de 2009, dans le cadre du plan de relance de l'économie¹², ou encore le cumul possible, entre mi-2009 et fin 2010, du PTZ et du Pass-Foncier¹³ pour l'achat d'un logement collectif neuf.

9 Les données ayant servi à l'établir sont issues du recensement 2008 de l'INSEE. Nous avons légèrement modifié ce zonage en l'enrichissant d'une catégorie : au sein de chaque grand pôle, nous avons séparé des autres communes la commune qui donne son nom au pôle.

10 Par commune rétrolittorales situées à moins de 5 km du littoral, il faut entendre les communes dont tout ou partie des limites sont situées à moins de 5 km de la mer (marée haute) et qui ne sont pas littorales au sens de la loi littoral. Même chose pour les communes rétrolittorales situées à moins de 10 et 15 km de la mer.

11 On entend par primo-accédants les ménages qui n'ont pas été propriétaires de leur logement dans les deux ans précédant l'opération envisagée. Ils peuvent néanmoins être propriétaires d'un logement qui n'est pas leur résidence principale.

12 Cette majoration est réduite à 50 % durant le second semestre 2010.

13 Lors de l'achat d'une résidence principale, mécanisme financier permettant de différer l'achat du terrain pendant toute la

- Période 3 (PTZ+ 2011). Elle regroupe l'ensemble des transactions pour lesquelles l'année d'émission de l'offre est 2011, première année du PTZ+. Le PTZ+ résulte de la fusion du NPTZ, du Pass-Foncier et du crédit d'impôt sur les intérêts d'emprunt. Ce dispositif reste réservé aux primo-accédants pour une acquisition dans le neuf ou l'ancien (sans quotité minimale de travaux), mais l'éligibilité n'est plus soumise à des conditions de ressources. Le revenu fiscal de référence du ménage influence toutefois les conditions de remboursement du PTZ. De plus, le PTZ+ prend désormais en compte la performance énergétique du logement.
- Période 4 (PTZ + à partir de 2012). Elle regroupe l'ensemble des transactions pour lesquelles la date d'émission de l'offre est postérieure au 1^{er} janvier 2012. Le PTZ + est alors dit « recentré », c'est-à-dire qu'il ne finance plus que des logements neufs¹⁴. Des plafonds de ressources sont réintroduits. Le dispositif a connu des évolutions pendant la période considérée. Ainsi, à compter du 1^{er} janvier 2013, le PTZ + ne peut pas être accordé à un logement neuf si celui-ci ne respecte pas une condition de performance énergétique¹⁵.

Neuvième étape : tester un grand nombre de modèles en faisant varier les paramètres

De nombreux essais de combinaisons de variables ont été effectués, générant autant de modèles différents. Ceci a permis de sélectionner les variables ayant le plus grand pouvoir explicatif sur le prix mais aussi de tester diverses transformations de variables (en logarithme par exemple¹⁶). Cela est valable pour les variables explicatives mais aussi pour la variable à expliquer¹⁷. Nous testons ainsi systématiquement l'évolution de la qualité du modèle selon qu'il explique le prix, le logarithme du prix, le prix/m² ou encore le logarithme du prix/m².

Dixième étape : envisager des segmentations

Nous avons également systématiquement testé l'opportunité de modélisations distinctes sur les segments suivants :

- Individuel et collectif (TIMM = 1 ou 2) ;
- Neuf et ancien¹⁸ (TOPE = 4 ou 1/2/3, puis, si la segmentation est jugée pertinente, séparation du neuf entre la promotion, la construction avec terrain et la construction sans terrain).

durée de remboursement du prêt principal tout en sécurisant l'opération en cas de difficultés.

14 Les exceptions concernent les logements anciens du parc social HLM vendus à leurs occupants.

15 Dans le cas général, le respect de la réglementation thermique 2012 ou l'obtention du label « Bâtiment Basse Consommation BBC 2005 ».

16 Une variation marginale d'une variable explicative numérique peut avoir des effets variables sur le prix selon la valeur initiale que prend cette variable explicative. Ainsi, dans le cas de la surface habitable, le passage de 10 à 11 m² a plus d'impact sur le prix que le passage de 100 à 101 m². Pour tenir compte de ce phénomène, on peut choisir d'appliquer un logarithme à la variable explicative concernée. Le logarithme est une fonction croissante dont la dérivée seconde est négative. En transformant dans la fonction, la surface par le logarithme de la surface, le prix va continuer à croître avec l'augmentation de la surface mais cette hausse sera d'autant moins forte que la surface du bien sera importante.

17 L'application d'un logarithme à la variable à expliquer n'a pas du tout le même rôle que dans la note précédente. Il vise à passer d'un modèle additif à un modèle multiplicatif.

18 Ce test n'est réalisé que pour les périodes où l'ancien (TOPE = 4) représente un nombre significatif d'observations. Lorsque le PTZ dans l'ancien est très minoritaire, nous n'avons pas testé de modélisation et nous ne travaillons que sur le neuf (TOPE = 1, 2 ou 3).

Résultats

Les modélisations ont été réalisées à l'aide du logiciel SAS (version 9.2).

Pour chaque période, une première modélisation est réalisée en utilisant uniquement les variables relatives aux caractéristiques intrinsèques du logement, à sa localisation et à la période de la transaction.

Une seconde modélisation est réalisée en ajoutant à cette liste les variables relatives aux caractéristiques de l'emprunteur.

La méthode « STEPWISE » est utilisée. Elle permet de sélectionner les variables explicatives à la fois les moins corrélées entre elles et qui contribuent le plus à l'estimation de la variable expliquée.

Le logiciel SAS s'assure, pour chaque modélisation, que les variables explicatives ne sont pas liées entre elles par une relation linéaire.

Pour la période 4 (PTZ + à partir de 2012)

Nous nous sommes intéressés uniquement aux transactions dans le neuf (TOPE = 1, 2 et 3), étant donné que les autres types d'opérations ne rassemblent qu'environ 2 500 observations. 8 observations sont retirées du modèle car elles présentent des valeurs négatives pour la variable EQUIV_SUBV. La modélisation est donc effectuée sur une base comprenant 112 981 observations.

Les modélisations sont réalisées sans distinguer l'individuel du collectif. En effet, la qualité du modèle (cf. encadré ci-contre) est nettement dégradée pour le collectif en cas de séparation (R^2 ajusté de l'ordre de 0,5). De même, les différents segments du neuf (promotion, construction avec terrain ou sans terrain) ne font pas l'objet de modélisations distinctes car ces dernières n'améliorent pas le modèle.

Le R^2 ajusté, un indicateur de la qualité du modèle

La qualité d'un modèle est d'autant plus grande qu'il permet de minimiser les écarts entre la valeur prédite et la valeur réelle. Pour l'estimer, on calcule d'abord le R^2 , parfois appelé coefficient de détermination. Celui-ci est égal à la part de la variance (c'est-à-dire des écarts de chaque observation à la moyenne) expliquée par le modèle. Il se calcule selon la formule suivante :

$R^2 = \text{SCE} / (\text{SCE} + \text{SCER})$ où SCE = Somme des carrés des écarts de chaque valeur prédite à la moyenne et SCER = Somme des carrés des écarts de chaque valeur observée à la valeur prédite

Le coefficient R^2 est positif et inférieur ou égal à 1 par construction. Il vaut 1 si l'écart résiduel est nul.

Le R^2 est un indicateur de qualité de la modélisation, mais il présente une limite. En effet, il est d'autant plus élevé que le nombre de variables explicatives est important. En effet, si on ajoute une variable au modèle, on augmente les chances de s'approcher de la valeur réelle de l'observation. Pourtant, l'ajout d'une variable n'améliore pas nécessairement la qualité du modèle. En effet, l'ajout d'une variable génère aussi un nouvel écart (une nouvelle erreur) entre valeur prédite et valeur réelle. Il faut donc un indicateur qui permette de juger si l'erreur supplémentaire induite par l'ajout d'une variable n'est pas trop grande. C'est le R^2 ajusté.

La qualité du modèle est donc d'autant plus grande que la valeur du R^2 ajusté est élevée.

1. Hors caractéristiques de l'acheteur

La variable expliquée qui donne le meilleur modèle (au sens du R^2 ajusté) est le logarithme du prix au m².

Les variables explicatives sont : TOPE, TIMM, LOG(NPPR), SURH, PELG, ZOCO, ANNEE_DATP, SAISONNALITE, LOG(SURH_N), GTC_2011, CODE_GÉOGRAPHIQUE, ZONAGE_LITTORAL et LOG(EQUIV_SUBV).

Le tableau ci-dessous présente la liste des variables dans l'ordre croissant de leur contribution cumulée au R^2 ajusté, ainsi que la valeur de cette contribution. Le R^2 ajusté est égal à 0,80.

Libellé de la variable	Code de la variable	Contribution cumulée au R^2 ajusté
Zone Robien	ZOCO	0.5870

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Surface habitable	SURH	0.7408
Zonage de tension	GTC_2011	0.7652
Type d'opération	TOPE	0.7840
Performance énergétique du logement	PELG	0.7930
Zonage en aires urbaines	CODE_GEOG_URB	0.7958
Logarithme de la surface par pièce	LOG_SURH_N	0.7977
Année de l'émission de l'offre de prêt	ANNEE_DATP	0.7993
Logarithme du nombre de pièces	LOG_NPPR	0.7999
Logarithme de l'équivalent subvention	LOG(EQUIV_SUBV)	0.8005
Zonage littoral	ZONAGE_LITTORAL	0.8007
Trimestre de l'émission de l'offre de prêt	SAISONNALITE	0.8007
Type de logement	TIMM	0.8007

Aucune variable n'est rejetée par la méthode, ce qui signifie qu'elles contribuent toutes à une amélioration du R² ajusté. Toutes les variables sont significatives au sens du test de Student¹⁹ (p-value inférieure à 0,05) sauf pour une modalité de la variable SAISONNALITE (p-value = 0,08).

La matrice des corrélations ne présente aucun coefficient supérieur à 0,7 entre deux variables quantitatives. De plus, pour identifier une dépendance excessive de la variable EQUIV_SUBV aux autres variables explicatives du modèle, une modélisation auxiliaire a été réalisée. Elle consiste en une tentative d'explication de LOG(EQUIV_SUBV) par l'ensemble des autres variables explicatives du modèle. Cette modélisation aboutit à un R² ajusté de 0,38²⁰. Ainsi, la variable LOG(EQUIV_SUBV) n'est pas excessivement dépendante des autres variables explicatives du modèle²¹.

Le premier graphique ci-dessous permet de s'assurer que les résidus suivent une loi normale centrée. Le second graphique permet de contrôler l'homoscédasticité du modèle²².

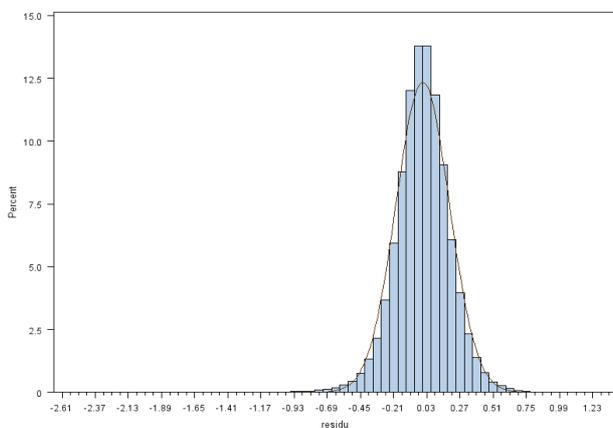
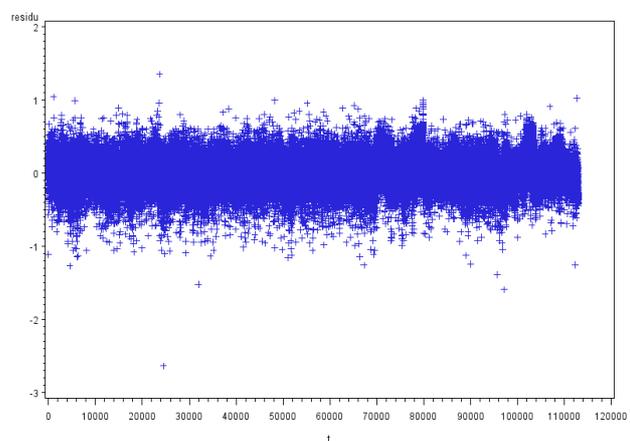


Diagramme des résidus



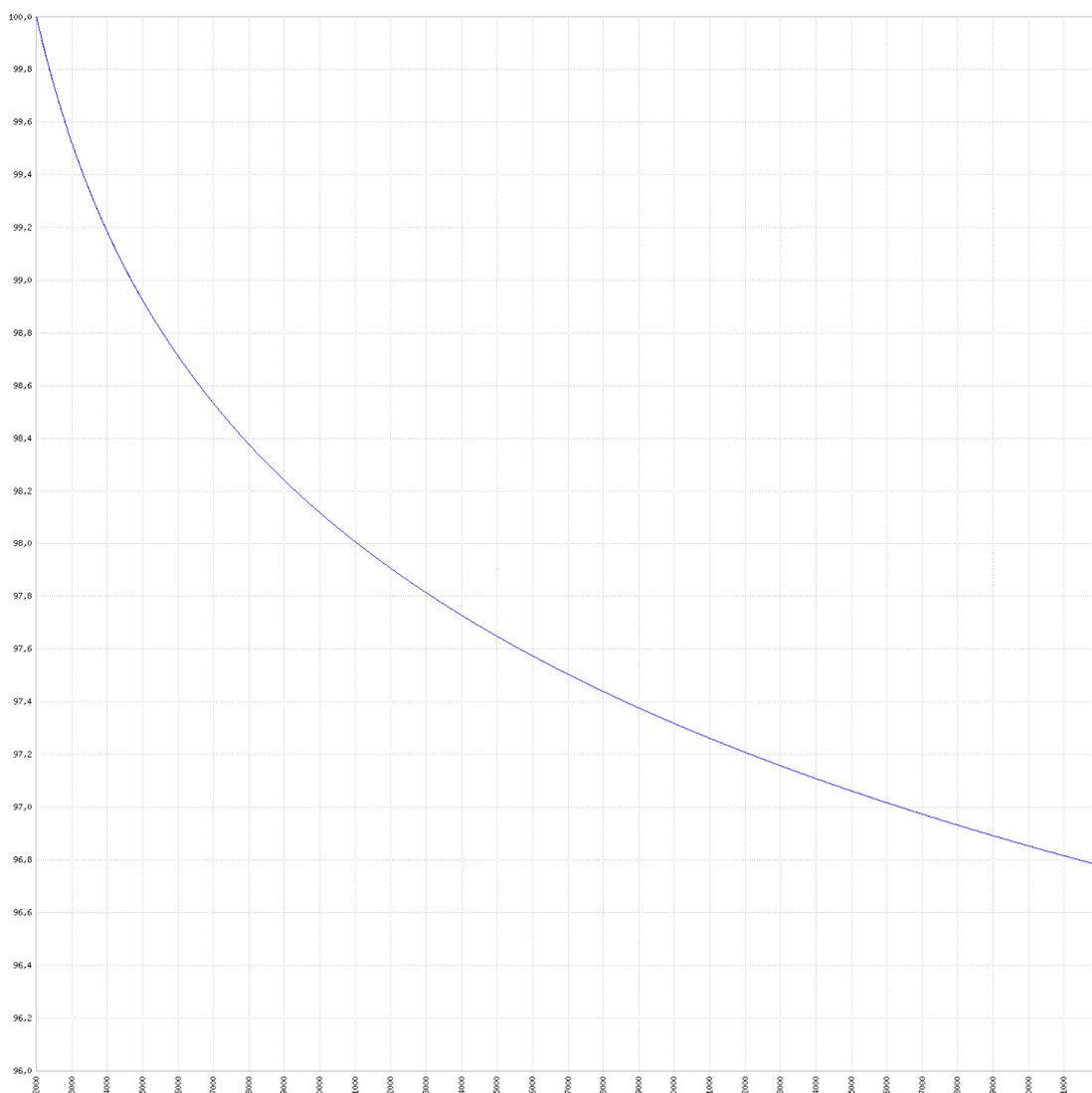
Répartition des résidus

La variable LOG(EQUIV_SUBV) contribue de manière significative au R² ajusté. **Le coefficient associé à cette variable est négatif égal à - 0,011806.** Une variation de 1 euro de

- 19 Ce test permet de savoir si la variable (ou une de ses modalités dans le cas d'une variable qualitative) a ou non une influence sur la variable expliquée. Plus la p-value est élevée, plus le risque que cette variable n'ait aucune influence sur le modèle est important.
- 20 Les deux variables qui contribuent le plus au R² ajusté de cette modélisation auxiliaire sont dans l'ordre ZOCO et PELG qui sont effectivement des paramètres qui entrent en considération dans le calcul du montant de l'aide.
- 21 Pour que les résultats d'une modélisation soient interprétables de manière fiable, il faut s'assurer que les variables explicatives soient autant que possible indépendantes les unes des autres. En effet, si deux variables sont trop liées entre elles, le modèle ne peut pas rendre compte correctement du poids spécifique de celles-ci.
- 22 Le modèle répond à l'exigence d'homoscédasticité lorsque les résidus se répartissent de manière homogène sur le spectre des valeurs prises par la variable à expliquer.

LOG(EQUIV_SUBV) fait diminuer de 0,011806 euros le logarithme du prix au m². La courbe ci-dessous décrit le comportement du prix au m² selon l'évolution de l'équivalent subvention, en base 100 lorsque la variable EQUIV_SUBV est égale à 2000 euros²³. Le tableau ci-dessous présente quelques illustrations de l'importance de l'effet de cette variable sur les prix.

Si	Alors
EQUIV_SUBV passe de 2 000 à 3 000 euros	Le prix au m ² diminue de 0,48 %
EQUIV_SUBV passe de 5 000 à 15 000 euros	Le prix au m ² diminue de 1,29 %
EQUIV_SUBV passe de 20 000 à 21 000 euros	Le prix au m ² diminue de 0,06 %
EQUIV_SUBV passe de 20 000 à 30 000 euros	Le prix au m ² diminue de 0,48 %



Courbe de la fonction $\text{Prix/m}^2 = 109,4 \cdot \text{EQUIV_SUBV}^{-0,011806}$

2. Avec les caractéristiques de l'acheteur

En plus des variables utilisées dans le modèle précédent, les variables suivantes sont ajoutées : EMEM, NPER, CCSP, RANN, AGE_EMPR.

²³ Nous avons représenté la courbe pour un équivalent subvention compris entre 2 000 et 32 000 euros ce qui correspond aux valeurs observées pour cette période (5 % des valeurs de l'équivalent subvention sont inférieures à 2 240 et 5 % sont supérieures à 32 085).

La variable expliquée qui donne le meilleur modèle est le logarithme du prix au m².

L'ajout au modèle de ces variables améliore légèrement le R² ajusté (0,85). Le tableau ci-dessous présente la liste des variables dans l'ordre croissant de leur contribution cumulée au R² ajusté, ainsi que la valeur de cette contribution.

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Zone Robien	ZOCO	0.5870
Surface habitable	SURH	0.7408
Revenu net total pour l'année N	RANN	0.7995
Zonage de tension	GTC_2011	0.8154
Type d'opération	TOPE	0.8284
Performance énergétique du logement	PELG	0.8366
Code CSP de l'emprunteur	CCSP	0.8397
Zonage en aires urbaines	CODE_GEOG_URB	0.8412
Année de l'émission de l'offre de prêt	ANNEE_DATP	0.8416
Logarithme de la surface par pièce	LOG_SURH_N	0.8420
Logarithme du nombre de pièces	LOG_NPPR	0.8476
État matrimonial du ménage emprunteur	EMEM	0.8481
Nombre de personnes du ménage	NPER	0.8483
Zonage littoral	ZONAGE_LITTORAL	0.8484
Logarithme de l'équivalent subvention	LOG(EQUIV_SUBV)	0.8485
Trimestre de l'émission de l'offre de prêt	SAISONNALITE	0.8486
Âge de l'emprunteur	AGE_EMPR	0.8486

La méthode « STEPWISE » écarte dans ce cas la variable relative au type d'immeuble (TIMM), celle-ci n'améliorant pas le R² ajusté. Toutes les variables sont significatives au sens du test de Student (p-value inférieure à 0,05) sauf pour une modalité de la variable SAISONNALITE (p-value = 0,14) et une modalité de la variable ZONAGE_LITTORAL (p-value = 0,06).

De plus, pour identifier une dépendance excessive de la variable EQUIV_SUBV aux autres variables explicatives du modèle, une modélisation auxiliaire a été réalisée. Elle consiste en une tentative d'explication de LOG(EQUIV_SUBV) par l'ensemble des autres variables explicatives du modèle. Cette modélisation aboutit à un R² ajusté de 0,46²⁴. Ainsi, la variable LOG(EQUIV_SUBV) n'est pas excessivement dépendante des autres variables explicatives du modèle.

Le premier graphique ci-dessous permet de s'assurer que les résidus suivent une loi normale centrée. Le second graphique permet de contrôler l'homoscédasticité du modèle.

24 Les quatre variables qui contribuent le plus au R² ajusté de cette modélisation auxiliaire sont dans l'ordre ZOCO, NPER, PELG et RANN, qui sont effectivement des paramètres qui entrent en considération dans le calcul du montant du prêt à taux zéro conféré au ménage et du différé éventuel.

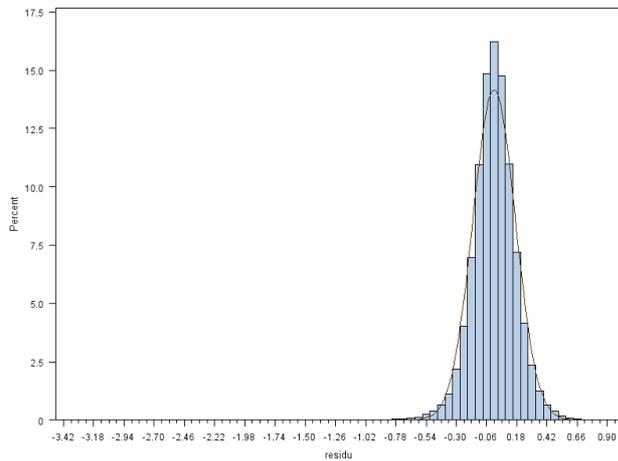
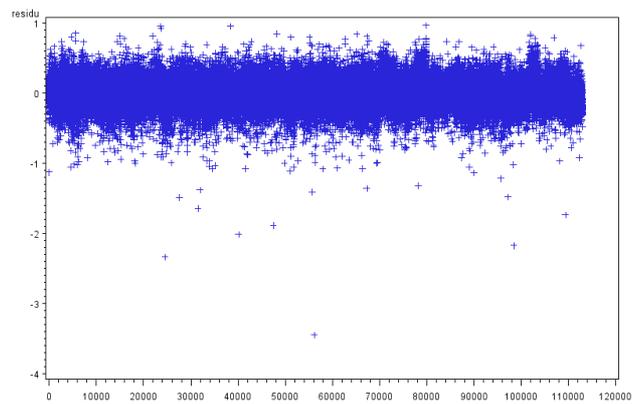


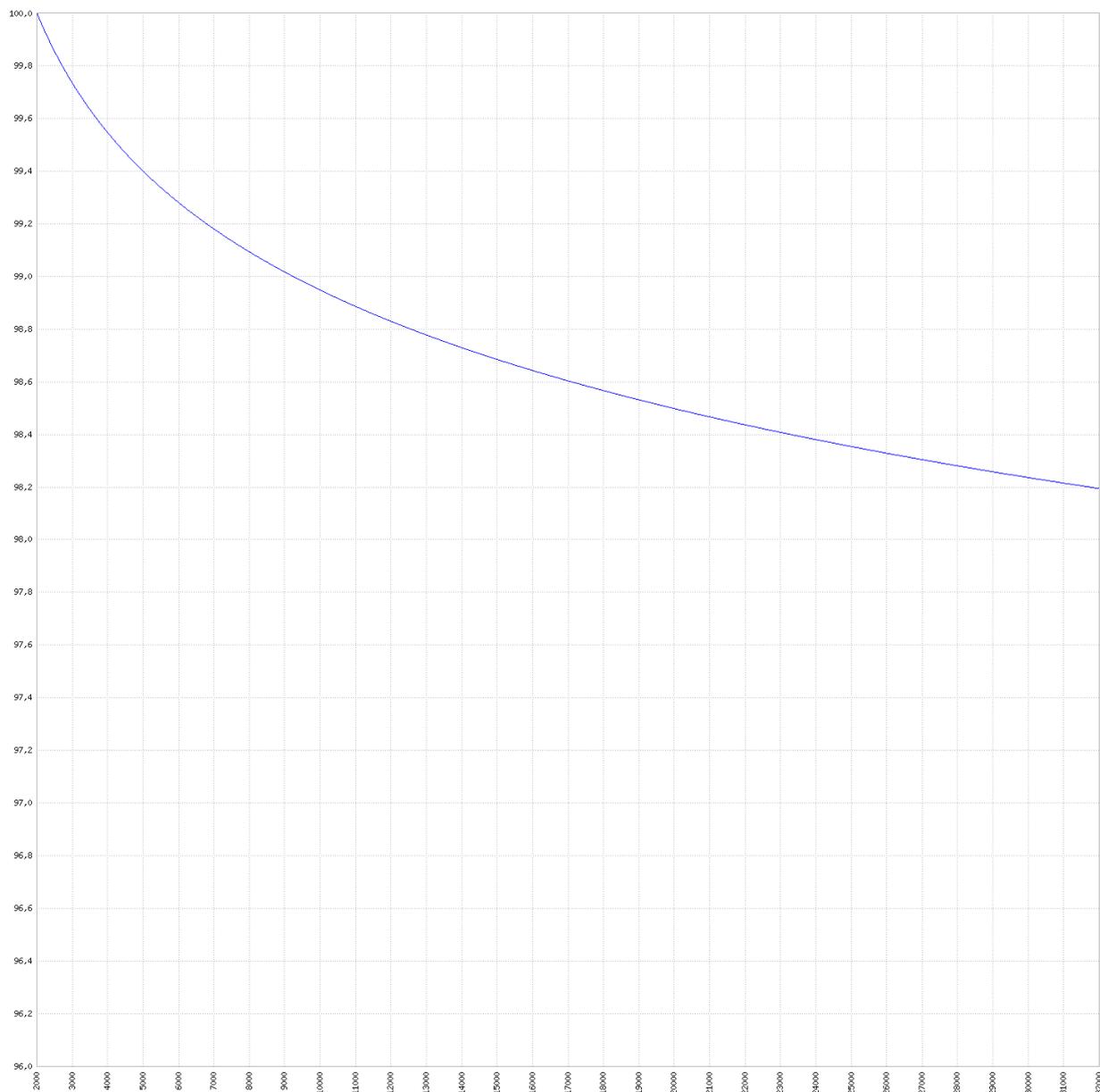
Diagramme des résidus



Répartition des résidus

La variable LOG(EQUIV_SUBV) contribue de manière significative au R^2 ajusté. **Le coefficient associé à cette variable est négatif égal à $-0,006571$.** Une variation de 1 euro de LOG(EQUIV_SUBV) fait diminuer de 0,006571 euros le logarithme du prix au m^2 . La courbe ci-dessous décrit le comportement du prix au m^2 selon l'évolution de l'équivalent subvention, en base 100 lorsque la variable EQUIV_SUBV est égale à 2000. Le tableau ci-dessous présente quelques illustrations de l'importance de l'effet de cette variable sur les prix.

Si	Alors
EQUIV_SUBV passe de 2 000 à 3 000 euros	Le prix au m ² diminue de 0,27 %
EQUIV_SUBV passe de 5 000 à 15 000 euros	Le prix au m ² diminue de 0,72 %
EQUIV_SUBV passe de 20 000 à 21 000 euros	Le prix au m ² diminue de 0,03 %
EQUIV_SUBV passe de 20 000 à 30 000 euros	Le prix au m ² diminue de 0,27 %



Courbe de la fonction $\text{Prix/m}^2 = 105,12 \cdot \text{EQUIV_SUBV}^{-0,006571}$

Pour la période 3 (PTZ + année 2011)

Nous avons privilégié une séparation entre le segment du neuf (TOPE = 1, 2 et 3) et le segment de l'ancien (TOPE = 4). En effet, la qualité du modèle est un peu moins bonne sans cette séparation (R^2 ajusté égal 0,68). De même, les différents segments du neuf (promotion, construction avec terrain ou sans terrain) ne font pas l'objet de modélisations distinctes car ces dernières n'améliorent pas le modèle. Les modélisations sont réalisées sans distinguer l'individuel du collectif.

11 observations sont retirées car la variable TOPE est renseignée à 5, 6 ou 7 (cf. page 6).

1. Le segment du neuf (TOPE = 1, 2 ou 3)

La modélisation est effectuée sur une base comprenant 82 066 observations.

La matrice de corrélations montre que les variables NPPR et SURH sont fortement corrélées. Aussi, seule la variable SURH a été conservée.

La variable expliquée qui donne le meilleur modèle est le logarithme du prix au m².

a. hors caractéristiques de l'acheteur

Les variables explicatives sont : TOPE, TIMM, LOG(SURH), PELG, ZOCO, SAISONNALITE, LOG(SURH_N), GTC_2011, CODE_GÉOGRAPHIQUE, ZONAGE_LITTORAL et EQUIV_SUBV.

Le tableau ci-dessous présente la liste des variables dans l'ordre croissant de leur contribution cumulée au R² ajusté, ainsi que la valeur de cette contribution. Le R² ajusté est égal à 0,74.

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Logarithme de la surface habitable	LOG_SURH	0.4979
Zone Robien	ZOCO	0.6502
Type d'opération	TOPE	0.6852
Zonage de tension	GTC_2011	0.7135
Performance énergétique du logement	PELG	0.7273
Logarithme de la surface par pièce	LOG_SURH_N	0.7349
Équivalent subvention	EQUIV_SUBV	0.7386
Zonage en aires urbaines	CODE_GEOG_URB	0.7411
Zonage littoral	ZONAGE_LITTORAL	0.7414
Type de logement	TIMM	0.7415
Trimestre de l'émission de l'offre de prêt	SAISONNALITE	0.7415

La méthode « STEPWISE » n'écarte aucune variable. Toutes les variables sont significatives au sens du test de Student (p-value inférieure à 0,05) sauf pour deux modalités de la variable SAISONNALITE (p-value = 0,25 et 0,07) et une modalité de la variable ZONAGE_LITTORAL (p-value = 0,20).

De plus, pour identifier une dépendance excessive de la variable EQUIV_SUBV aux autres variables explicatives du modèle, une modélisation auxiliaire a été réalisée. Elle consiste en une tentative d'explication de EQUIV_SUBV par l'ensemble des autres variables explicatives du modèle. Cette modélisation aboutit à un R² ajusté de 0,64²⁵. **La variable EQUIV_SUBV étant expliquée de manière relativement significative par les autres variables explicatives du modèle, il apparaît difficile de tirer d'interpréter de manière fiable le coefficient qui lui est associé.** Les résultats sont fournis dans la suite à titre indicatif.

Le premier graphique ci-dessous permet de s'assurer que les résidus suivent une loi normale centrée. Le second graphique permet de contrôler l'homoscédasticité du modèle.

25 Les deux variables qui contribuent le plus au R² ajusté de cette modélisation auxiliaire sont dans l'ordre ZOCO et PELG qui sont effectivement des paramètres qui entrent en considération dans le calcul du montant de l'aide.

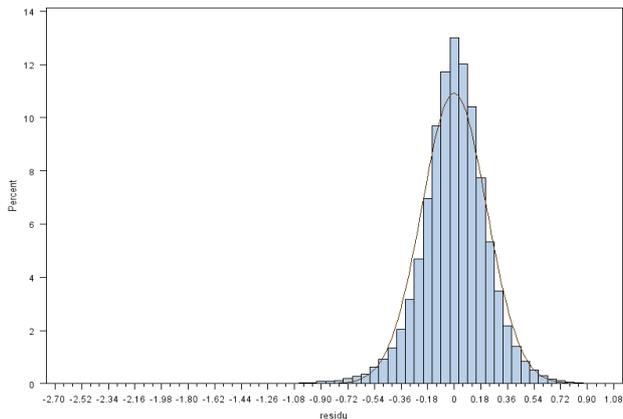
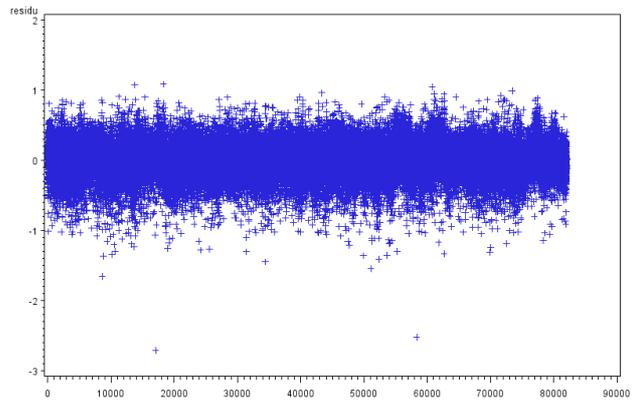


Diagramme des résidus

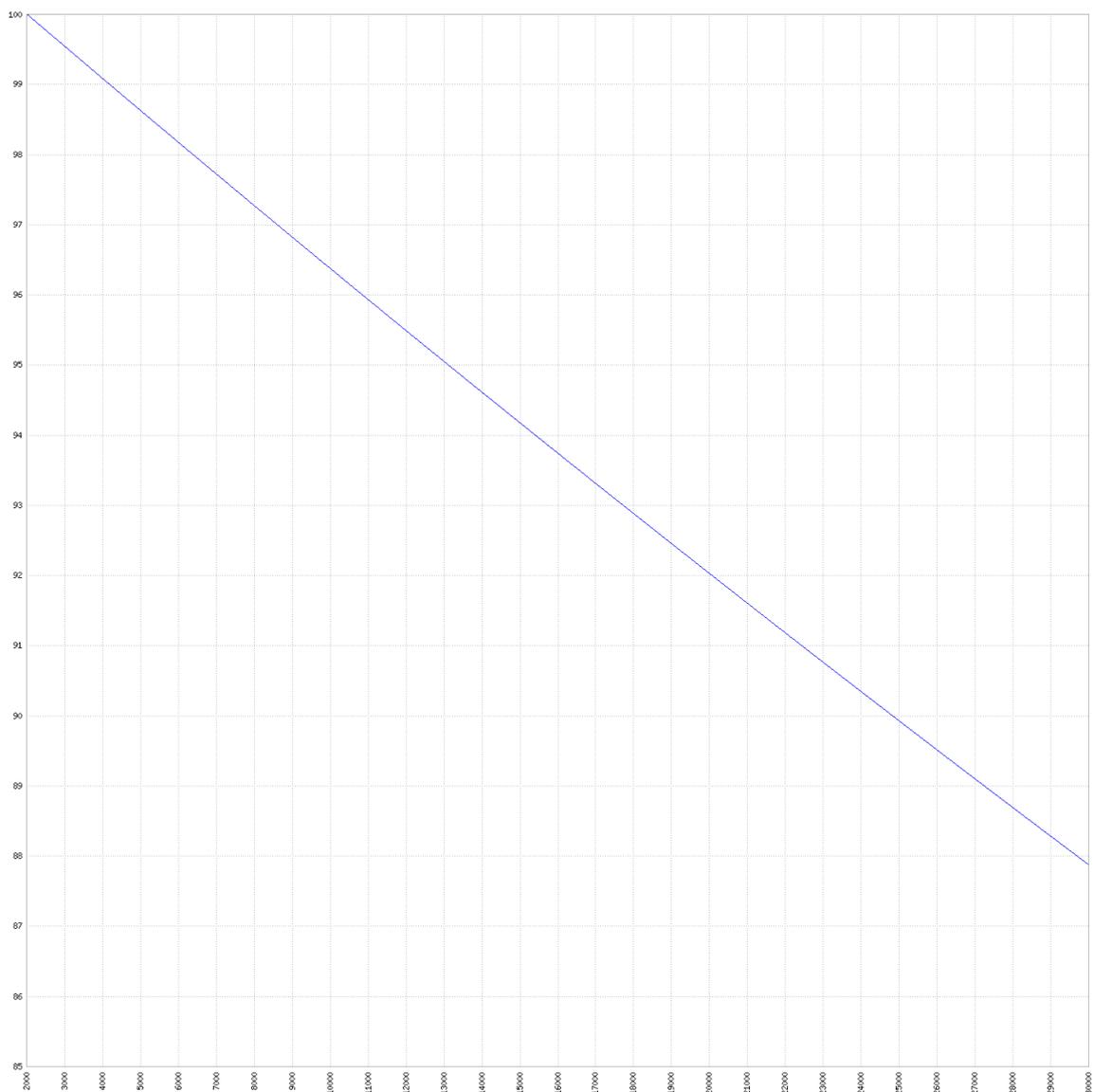


Répartition des résidus

La variable EQUIV_SUBV contribue de manière significative au R^2 ajusté. **Le coefficient associé à cette variable est négatif égal à $-0,00004616$.** Une variation de 1 euro de EQUIV_SUBV fait diminuer de $0,00004616$ euros le logarithme du prix au m^2 . La courbe ci-dessous décrit le comportement du prix au m^2 selon l'évolution de l'équivalent subvention, en base 100 lorsque la variable EQUIV_SUBV est égale à 2000²⁶. Le tableau ci-dessous présente quelques illustrations de l'importance de l'effet de cette variable sur les prix.

Si	Alors
EQUIV_SUBV passe de 2 000 à 3 000 euros	Le prix au m^2 diminue de 0,46 %
EQUIV_SUBV passe de 5 000 à 15 000 euros	Le prix au m^2 diminue de 4,5 %
EQUIV_SUBV passe de 20 000 à 21 000 euros	Le prix au m^2 diminue de 0,46 %
EQUIV_SUBV passe de 20 000 à 30 000 euros	Le prix au m^2 diminue de 4,5 %

²⁶ Nous avons représenté la courbe pour un équivalent subvention compris entre 2 000 et 30 000 euros ce qui correspond aux valeurs observées pour cette période (5 % des valeurs de l'équivalent subvention sont inférieures à 1 848 et 5 % sont supérieures à 29 379).



Courbe de la fonction $\text{Prix/m}^2 = 100,93 \cdot \text{EXP}(-0,000004616 \cdot \text{EQUIV_SUBV})$

b. avec les caractéristiques de l'acheteur

On ajoute au modèle précédent les variables explicatives suivantes : EMEM, RANN, AGE_EMPR, CCSP, NPER.

Le tableau ci-dessous présente la liste des variables dans l'ordre croissant de leur contribution cumulée au R² ajusté, ainsi que la valeur de cette contribution. Le R² ajusté est égal à 0,80.

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Logarithme de la surface habitable	LOG_SURH	0.4979
Revenu net total pour l'année N	RANN	0.6613
Zone Robien	ZOCO	0.7373
Type d'opération	TOPE	0.7642
Zonage de tension	GTC_2011	0.7821
Performance énergétique du logement	PELG	0.7920

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Code CSP de l'emprunteur	CCSP	0.7967
Logarithme de la surface par pièce	LOG_SURH_N	0.8011
État matrimonial du ménage emprunteur	EMEM	0.8028
Zonage en aires urbaines	CODE_GEOG_URB	0.8042
Nombre de personnes du ménage	NPER	0.8044
Zonage littoral	ZONAGE_LITTORAL	0.8045
Équivalent subvention	EQUIV_SUBV	0.8046
Âge de l'emprunteur	AGE_EMPR	0.8046
Type de logement	TIMM	0.8046

La méthode « STEPWISE » écarte dans ce cas la variable relative au trimestre de la transaction (SAISONNALITE), celle-ci n'améliorant pas le R² ajusté. Toutes les variables sont significatives au sens du test de Student (p-value inférieure à 0,05) sauf la variable TIMM (p-value = 0,19) et pour une modalité de la variable ZONAGE_LITTORAL (p-value = 0,13).

De plus, pour identifier une dépendance excessive de la variable EQUIV_SUBV aux autres variables explicatives du modèle, une modélisation auxiliaire a été réalisée. Elle consiste en une tentative d'explication de EQUIV_SUBV par l'ensemble des autres variables explicatives du modèle. Cette modélisation aboutit à un R² ajusté de 0,78²⁷. **La variable EQUIV_SUBV étant expliquée de manière relativement significative par les autres variables explicatives du modèle, il apparaît difficile de tirer d'interpréter de manière fiable le coefficient qui lui est associé.** Les résultats sont fournis dans la suite à titre indicatif.

Le premier graphique ci-dessous permet de s'assurer que les résidus suivent une loi normale centrée. Le second graphique permet de contrôler l'homoscédasticité du modèle.

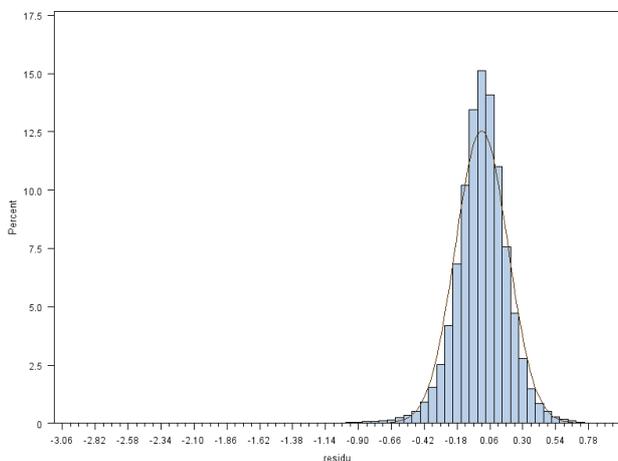
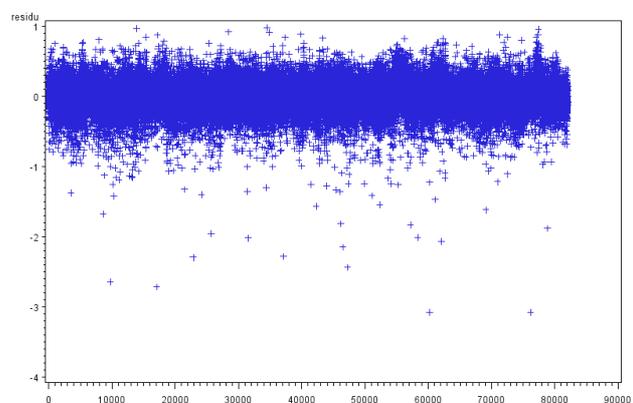


Diagramme des résidus

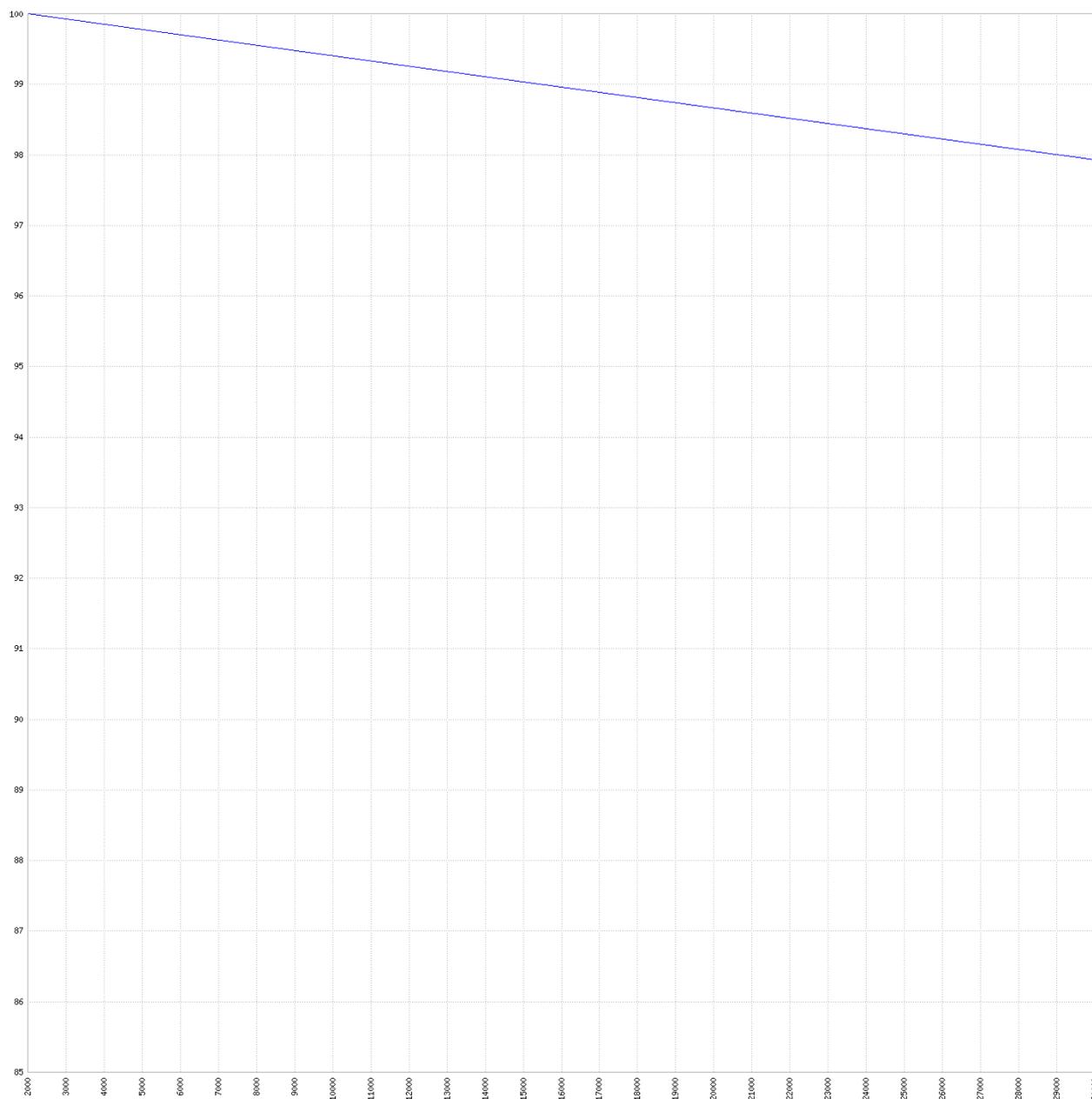


Répartition des résidus

La variable EQUIV_SUBV contribue de manière significative au R² ajusté. **Le coefficient associé à cette variable est négatif égal à - 0,000000747.** Une variation de 1 euro de EQUIV_SUBV fait diminuer de 0,000000747 euros le logarithme du prix au m². La courbe ci-dessous décrit le comportement du prix au m² selon l'évolution de l'équivalent subvention, en base 100 lorsque la variable EQUIV_SUBV est égale à 2000. Le tableau ci-dessous présente quelques illustrations de l'importance de l'effet de cette variable sur les prix.

27 Les quatre variables qui contribuent le plus au R² ajusté de cette modélisation auxiliaire sont dans l'ordre ZOCO, NPER, PELG et RANN, qui sont effectivement des paramètres qui entrent en considération dans le calcul du montant du prêt à taux zéro conféré au ménage et du différé éventuel.

Si	Alors
EQUIV_SUBV passe de 2 000 à 3 000 euros	Le prix au m ² diminue de 0,07 %
EQUIV_SUBV passe de 5 000 à 15 000 euros	Le prix au m ² diminue de 0,74 %
EQUIV_SUBV passe de 20 000 à 21 000 euros	Le prix au m ² diminue de 0,07 %
EQUIV_SUBV passe de 20 000 à 30 000 euros	Le prix au m ² diminue de 0,74 %



Courbe de la fonction $\text{Prix/m}^2 = 100,15 \cdot \text{EXP}(-0,00000747 \cdot \text{EQUIV_SUBV})$

2. Le segment de l'ancien (TOPE = 4)

La modélisation est effectuée sur une base comprenant 225 359 observations.

La matrice de corrélations montre que les variables NPPR et SURH sont fortement corrélées. Dans ce cas, la variable surface habitable SURH a été retirée du modèle.

La variable expliquée qui donne le meilleur modèle est le logarithme du prix au m².

a. hors caractéristiques de l'acheteur

Les variables explicatives sont : TIMM, LOG(NPPR), PELG, ZOCO, SAISONNALITE, LOG(SURH_N),

GTC_2011, CODE_GÉOGRAPHIQUE, ZONAGE_LITTORAL et EQUIV_SUBV.

Le tableau ci-dessous présente la liste des variables dans l'ordre croissant de leur contribution cumulée au R² ajusté, ainsi que la valeur de cette contribution. Le R² ajusté est égal à 0,67.

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Zone Robien	ZOCO	0.5085
Zonage de tension	GTC_2011	0.5609
Logarithme du nombre de pièces	LOG_NPPR	0.5913
Logarithme de la surface par pièce	LOG_SURH_N	0.6374
Performance énergétique du logement	PELG	0.6474
Zonage en aires urbaines	CODE_GEOG_URB	0.6569
Équivalent subvention	EQUIV_SUBV	0.6641
Type de logement	TIMM	0.6705
Trimestre de l'émission de l'offre de prêt	SAISONNALITE	0.6710
Zonage littoral	ZONAGE_LITTORAL	0.6715

La méthode « STEPWISE » n'écarte aucune variable. Toutes les variables sont significatives au sens du test de Student (p-value inférieure à 0,05) sauf pour une modalité de la variable ZONAGE_LITTORAL (p-value = 0,17).

De plus, pour identifier une dépendance excessive de la variable EQUIV_SUBV aux autres variables explicatives du modèle, une modélisation auxiliaire a été réalisée. Elle consiste en une tentative d'explication de EQUIV_SUBV par l'ensemble des autres variables explicatives du modèle. Cette modélisation aboutit à un R² ajusté de 0,40²⁸. La variable EQUIV_SUBV n'est donc pas excessivement dépendante des autres variables explicatives du modèle.

Le premier graphique ci-dessous permet de s'assurer que les résidus suivent une loi normale centrée. Le second graphique permet de contrôler l'homoscédasticité du modèle.

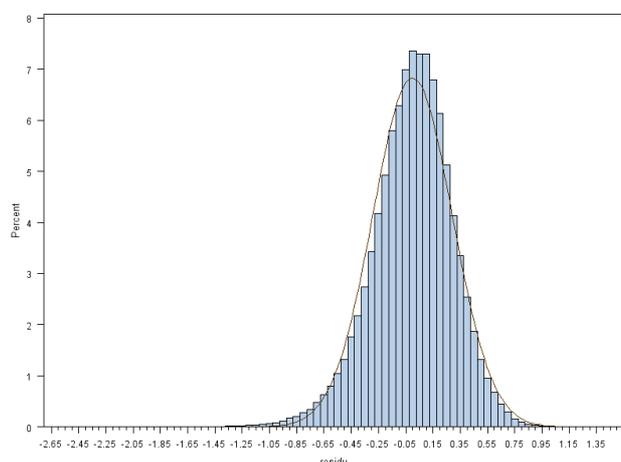
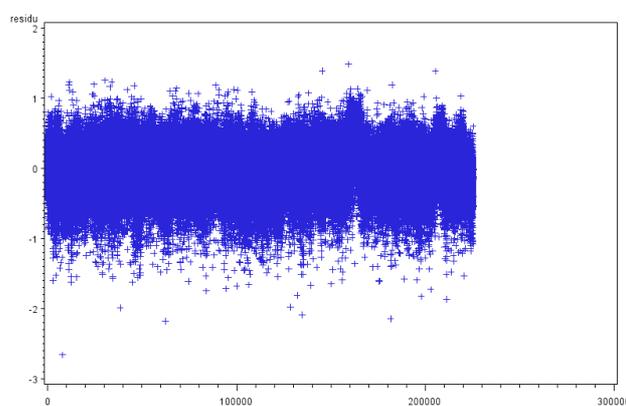


Diagramme des résidus



Répartition des résidus

La variable EQUIV_SUBV contribue de manière significative au R² ajusté. **Le coefficient associé à cette variable est négatif égal à - 0,000014437.** Une variation de 1 euro de EQUIV_SUBV fait diminuer de 0,000014437 euros le logarithme du prix au m². La courbe ci-dessous décrit le comportement du prix au m² selon l'évolution de l'équivalent subvention, en base 100 lorsque la variable EQUIV_SUBV est égale à 1000²⁹. Le tableau ci-dessous présente quelques illustrations de

28 Les deux variables qui contribuent le plus au R² ajusté de cette modélisation auxiliaire sont dans l'ordre PELG et ZOCO qui sont effectivement des paramètres qui entrent en considération dans le calcul du montant de l'aide.

29 Nous avons représenté la courbe pour un équivalent subvention compris entre 1 000 et 14 000 euros ce qui correspond

l'importance de l'effet de cette variable sur les prix.

Si	Alors
EQUIV_SUBV passe de 2 000 à 3 000 euros	Le prix au m ² diminue de 1,43 %
EQUIV_SUBV passe de 5 000 à 10 000 euros	Le prix au m ² diminue de 6,96 %
EQUIV_SUBV passe de 9 000 à 10 000 euros	Le prix au m ² diminue de 1,43 %
EQUIV_SUBV passe de 9 000 à 14 000 euros	Le prix au m ² diminue de 6,96 %



Courbe de la fonction $\text{Prix/m}^2 = 101,45 \cdot \text{EXP}(-0,000014437 \cdot \text{EQUIV_SUBV})$

b. avec les caractéristiques de l'acheteur

On ajoute au modèle précédent les variables explicatives suivantes : EMEM, RANN, AGE_EMPR, CCSP, NPER.

Le tableau ci-dessous présente la liste des variables dans l'ordre croissant de leur contribution cumulée au R² ajusté, ainsi que la valeur de cette contribution. Le R² ajusté est égal à 0,76.

aux valeurs observées pour cette période (5 % des valeurs de l'équivalent subvention sont inférieures à 1 067 et 5 % sont supérieures à 13 607).

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Zone Robien	ZOCO	0.5085
Zonage de tension	GTC_2011	0.5609
Logarithme du nombre de pièces	LOG_NPPR	0.5913
Revenu net total pour l'année N	RANN	0.6415
Logarithme de la surface par pièce	LOG_SURH_N	0.7258
Code CSP de l'emprunteur	CCSP	0.7362
Zonage en aires urbaines	CODE_GEOG_URB	0.7419
Performance énergétique du logement	PELG	0.7476
Type de logement	TIMM	0.7515
État matrimonial du ménage emprunteur	EMEM	0.7545
Âge de l'emprunteur	AGE_EMPR	0.7550
Nombre de personnes du ménage	NPER	0.7554
Zonage littoral	ZONAGE_LITTORAL	0.7558
Équivalent subvention	EQUIV_SUBV	0.7561
Trimestre de l'émission de l'offre de prêt	SAISONNALITE	0.7564

La méthode « STEPWISE » n'écarte aucune variable. Toutes les variables sont significatives au sens du test de Student (p-value inférieure à 0,05).

De plus, pour identifier une dépendance excessive de la variable EQUIV_SUBV aux autres variables explicatives du modèle, une modélisation auxiliaire a été réalisée. Elle consiste en une tentative d'explication de EQUIV_SUBV par l'ensemble des autres variables explicatives du modèle. Cette modélisation aboutit à un R² ajusté de 0,65³⁰. **La variable EQUIV_SUBV étant expliquée de manière relativement significative par les autres variables explicatives du modèle, il apparaît difficile de tirer d'interpréter de manière fiable le coefficient qui lui est associé.** Les résultats sont fournis dans la suite à titre indicatif.

Le premier graphique ci-dessous permet de s'assurer que les résidus suivent une loi normale centrée. Le second graphique permet de contrôler l'homoscédasticité du modèle.

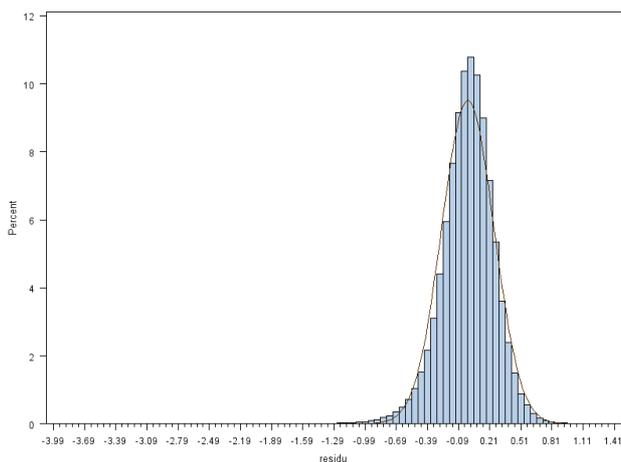
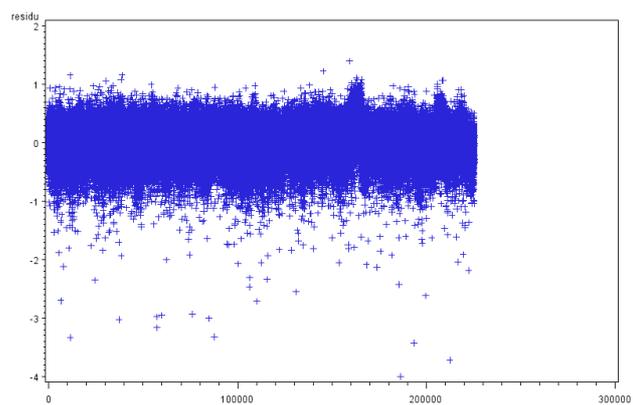


Diagramme des résidus



Répartition des résidus

30 Les quatre variables qui contribuent le plus au R² ajusté de cette modélisation auxiliaire sont dans l'ordre NPER, PELG, ZOCO et RANN, qui sont effectivement des paramètres qui entrent en considération dans le calcul du montant du prêt à taux zéro conféré au ménage et du différé éventuel.

La variable EQUIV_SUBV contribue de manière significative au R² ajusté. **Le coefficient associé à cette variable est négatif égal à - 0,000003781.** Une variation de 1 euro de EQUIV_SUBV fait diminuer de 0,000003781 euros le logarithme du prix au m². La courbe ci-dessous décrit le comportement du prix au m² selon l'évolution de l'équivalent subvention, en base 100 lorsque la variable EQUIV_SUBV est égale à 1000. Le tableau ci-dessous présente quelques illustrations de l'importance de l'effet de cette variable sur les prix.

Si	Alors
EQUIV_SUBV passe de 2 000 à 3 000 euros	Le prix au m ² diminue de 0,38 %
EQUIV_SUBV passe de 5 000 à 10 000 euros	Le prix au m ² diminue de 1,87 %
EQUIV_SUBV passe de 9 000 à 10 000 euros	Le prix au m ² diminue de 0,38 %
EQUIV_SUBV passe de 9 000 à 14 000 euros	Le prix au m ² diminue de 1,87 %



Courbe de la fonction Prix/m² = 100,38.EXP(-0,000003781*EQUIV_SUBV)

Pour la période 2 (NPTZ, de 2005 à 2010)

Nous avons privilégié une séparation entre le segment du neuf (TOPE = 1, 2 et 3) et le segment de l'ancien (TOPE = 4). En effet, la qualité du modèle est un peu moins bonne sans cette séparation. De même, les différents segments du neuf (promotion, construction avec terrain ou sans terrain) ne font pas l'objet de modélisations distinctes car ces dernières n'améliorent pas le modèle. Les modélisations sont réalisées sans distinguer l'individuel du collectif.

1. Le segment du neuf (TOPE = 1, 2 ou 3)

La modélisation est effectuée sur une base comprenant 365 900 observations. La variable expliquée qui donne le meilleur modèle est le logarithme du prix au m².

La matrice de corrélations montre que les variables NPPR et SURH sont fortement corrélées. Aussi, seule la variable NPPR a été conservée.

a. hors caractéristiques de l'acheteur

Les variables explicatives sont : ANNEE_DATP, TOPE, TIMM, LOG(NPPR), ZOCO, SAISONNALITE, LOG(SURH_N), GTC_2011, CODE_GÉOGRAPHIQUE, ZONAGE_LITTORAL et LOG(EQUIV_SUBV). La variable PELG n'est pas introduite dans ce modèle : comme elle n'entre pas en compte dans le dispositif PTZ avant 2011, elle n'est pas renseignée avant cette date.

Le tableau ci-dessous présente la liste des variables dans l'ordre croissant de leur contribution cumulée au R² ajusté, ainsi que la valeur de cette contribution. Le R² ajusté est égal à 0,71.

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Type d'opération	TOPE	0.4222
Zonage de tension	GTC_2011	0.5167
Logarithme de la surface par pièce	LOG_SURH_N	0.5876
Logarithme du nombre de pièces	LOG_NPPR	0.6743
Zone Robien	ZOCO	0.6929
Année de l'émission de l'offre de prêt	ANNEE_DATP	0.7065
Zonage en aires urbaines	CODE_GEOG_URB	0.7101
Logarithme de l'équivalent subvention	LOG_EQUIV_SUBV	0.7125
Type de logement	TIMM	0.7127
Zonage littoral	ZONAGE_LITTORAL	0.7128
Trimestre de l'émission de l'offre de prêt	SAISONNALITE	0.7129

La méthode « STEPWISE » n'écarte aucune variable. Toutes les variables sont significatives au sens du test de Student (p-value inférieure à 0,05) sauf pour une modalité de la variable SAISONNALITE (p-value = 0,43) et une modalité de la variable ZONAGE_LITTORAL (p-value = 0,95).

De plus, pour identifier une dépendance excessive de l'équivalent subvention aux autres variables explicatives du modèle, une modélisation auxiliaire a été réalisée. Elle consiste en une tentative d'explication de LOG_EQUIV_SUBV par l'ensemble des autres variables explicatives du modèle. Cette modélisation aboutit à un R² ajusté de 0,15³¹. La variable LOG_EQUIV_SUBV n'est donc pas excessivement dépendante des autres variables explicatives du modèle.

Le premier graphique ci-dessous permet de s'assurer que les résidus suivent une loi normale centrée. Le second graphique permet de contrôler l'homoscédasticité du modèle.

³¹ La variable qui contribue le plus au R² ajusté de cette modélisation auxiliaire est l'année d'émission de l'offre de prêt.

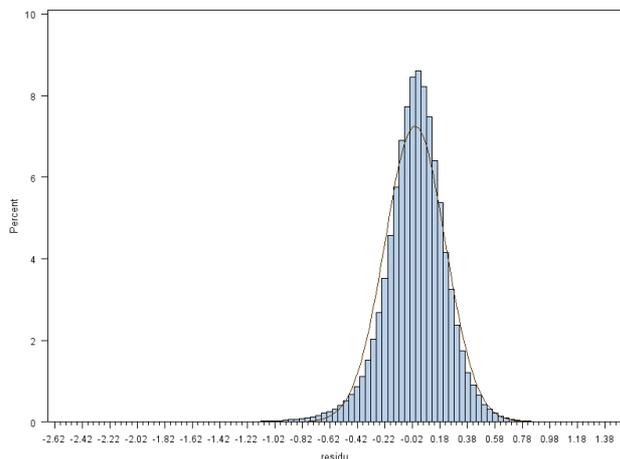
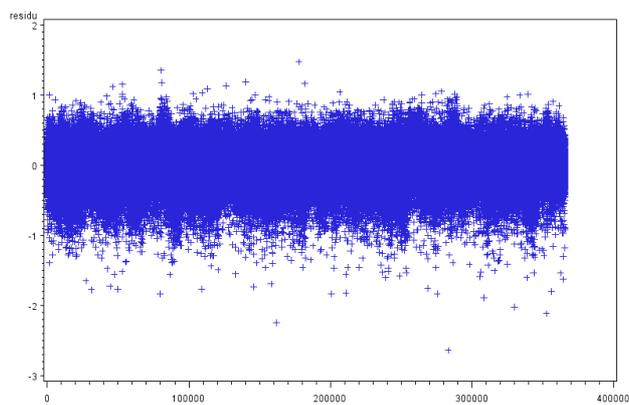


Diagramme des résidus

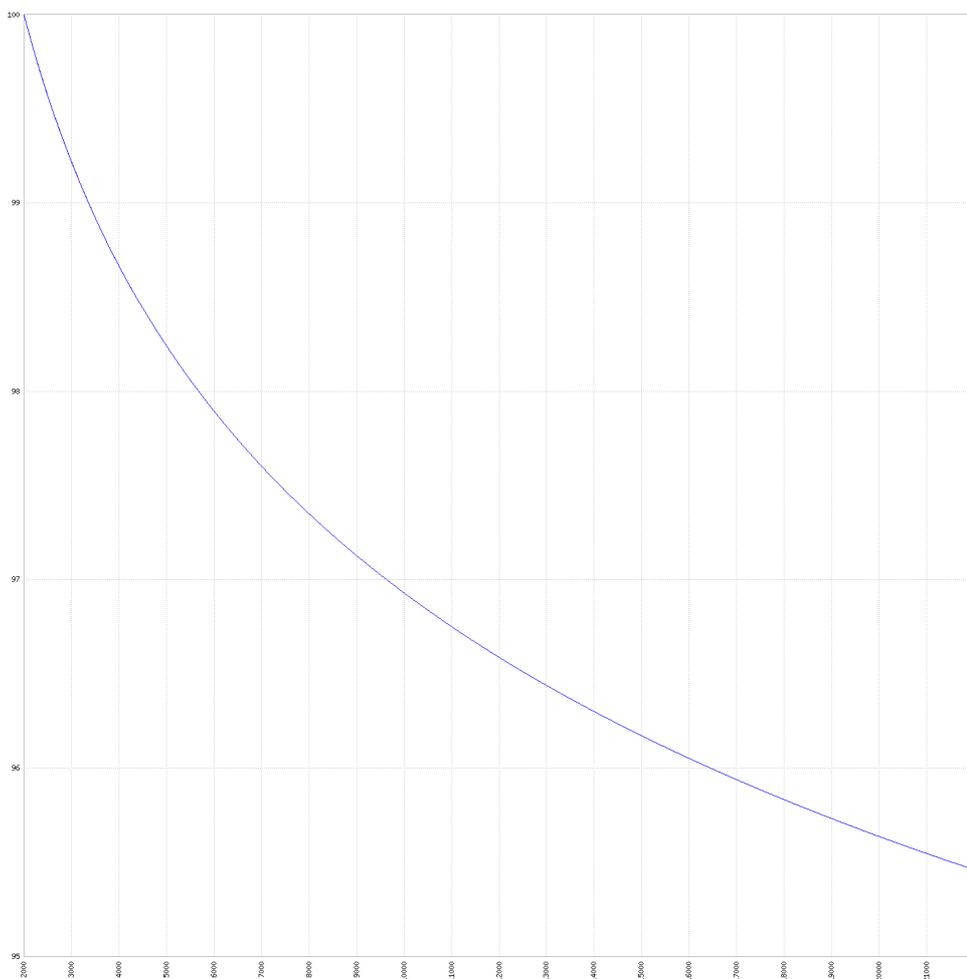


Répartition des résidus

La variable LOG_EQUIV_SUBV contribue de manière significative au R^2 ajusté. **Le coefficient associé à cette variable est négatif égal à $-0,019370$.** Une variation de 1 euro de LOG_EQUIV_SUBV fait diminuer de 0,019370 euros le logarithme du prix au m². La courbe ci-dessous décrit le comportement du prix au m² selon l'évolution de l'équivalent subvention, en base 100 lorsque la variable EQUIV_SUBV est égale à 2000³². Le tableau ci-dessous présente quelques illustrations de l'importance de l'effet de cette variable sur les prix.

Si	Alors
EQUIV_SUBV passe de 2 000 à 3 000 euros	Le prix au m ² diminue de 0,78 %
EQUIV_SUBV passe de 5 000 à 15 000 euros	Le prix au m ² diminue de 2,11 %
EQUIV_SUBV passe de 15 000 à 16 000 euros	Le prix au m ² diminue de 0,12 %
EQUIV_SUBV passe de 15 000 à 20 000 euros	Le prix au m ² diminue de 0,55 %

32 Nous avons représenté la courbe pour un équivalent subvention compris entre 2 000 et 22 000 euros ce qui correspond aux valeurs observées pour cette période (5 % des valeurs de l'équivalent subvention sont inférieures à 2 246 et 5 % sont supérieures à 22 113).



Courbe de la fonction $\text{Prix/m}^2 = 115,86.EQUIV_SUBV^{-0,019370}$

b. avec les caractéristiques de l'acheteur

On ajoute au modèle précédent les variables explicatives suivantes : EMEM, RANN, AGE_EMPR, CCSP, NPER.

Le tableau ci-dessous présente la liste des variables dans l'ordre croissant de leur contribution cumulée au R² ajusté, ainsi que la valeur de cette contribution. Le R² ajusté est égal à 0,76.

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Type d'opération	TOPE	0.4222
Zonage de tension	GTC_2011	0.5167
Logarithme de la surface par pièce	LOG_SURH_N	0.5876
Logarithme du nombre de pièces	LOG_NPPR	0.6743
Revenu net total pour l'année N	RANN	0.7329
Zone Robien	ZOCO	0.7436
Année de l'émission de l'offre de prêt	ANNEE_DATP	0.7530
Code CSP de l'emprunteur	CCSP	0.7580
Zonage en aires urbaines	CODE_GEOG_URB	0.7601
État matrimonial du ménage emprunteur	EMEM	0.7615
Nombre de personnes du ménage	NPER	0.7618
Type de logement	TIMM	0.7620

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Trimestre de l'émission de l'offre de prêt	SAISONNALITE	0.7621
Logarithme de l'équivalent subvention	LOG_EQUIV_SUBV	0.7622
Zonage littoral	ZONAGE_LITTORAL	0.7623
Âge de l'emprunteur	AGE_EMPR	0.7623

La méthode « STEPWISE » n'écarte aucune variable. Toutes les variables sont significatives au sens du test de Student (p-value inférieure à 0,05) sauf pour une modalité de la variable ZONAGE_LITTORAL (p-value = 0,06).

De plus, pour identifier une dépendance excessive de la variable EQUIV_SUBV aux autres variables explicatives du modèle, une modélisation auxiliaire a été réalisée. Elle consiste en une tentative d'explication de LOG_EQUIV_SUBV par l'ensemble des autres variables explicatives du modèle. Cette modélisation aboutit à un R² ajusté de 0,21³³. La variable LOG_EQUIV_SUBV n'est donc pas excessivement dépendante des autres variables explicatives du modèle.

Le premier graphique ci-dessous permet de s'assurer que les résidus suivent une loi normale centrée. Le second graphique permet de contrôler l'homoscédasticité du modèle.

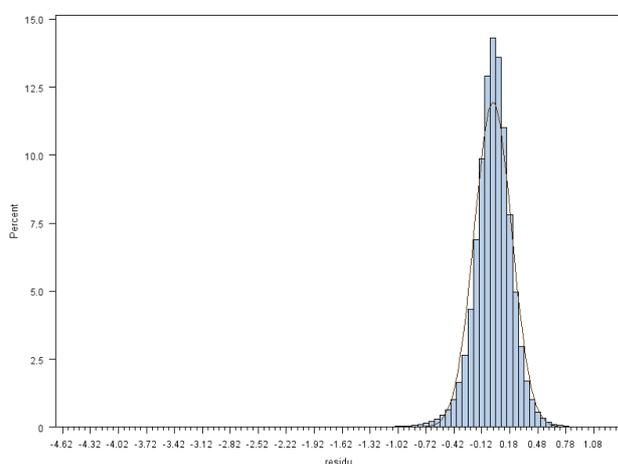
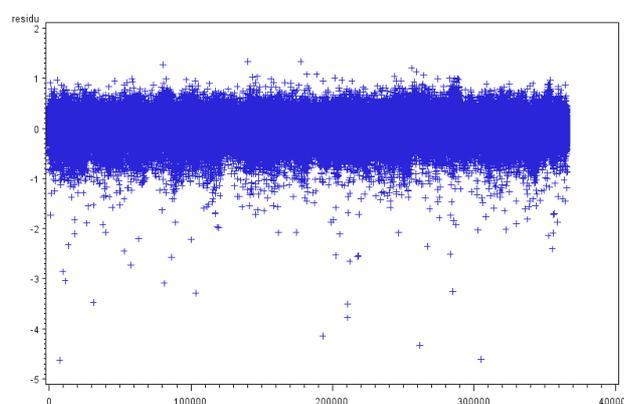


Diagramme des résidus

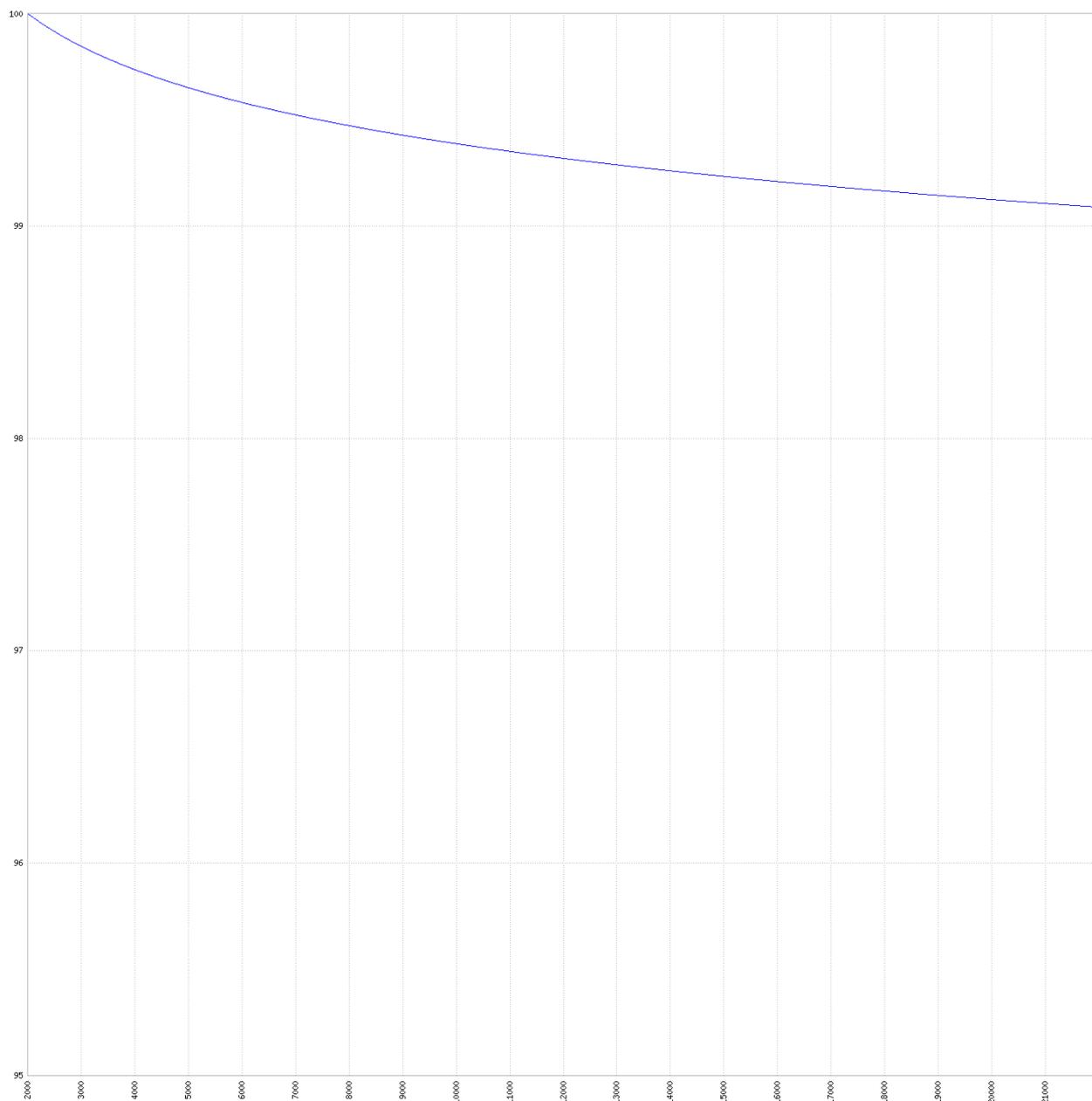


Répartition des résidus

La variable LOG_EQUIV_SUBV contribue de manière significative au R² ajusté. **Le coefficient associé à cette variable est négatif égal à - 0,003821.** Une variation de 1 euro de LOG_EQUIV_SUBV fait diminuer de 0,003821 euros le logarithme du prix au m². La courbe ci-dessous décrit le comportement du prix au m² selon l'évolution de l'équivalent subvention, en base 100 lorsque la variable EQUIV_SUBV est égale à 2000. Le tableau ci-dessous présente quelques illustrations de l'importance de l'effet de cette variable sur les prix.

Si	Alors
EQUIV_SUBV passe de 2 000 à 3 000 euros	Le prix au m ² diminue de 0,15 %
EQUIV_SUBV passe de 5 000 à 15 000 euros	Le prix au m ² diminue de 0,42 %
EQUIV_SUBV passe de 15 000 à 16 000 euros	Le prix au m ² diminue de 0,02 %
EQUIV_SUBV passe de 15 000 à 20 000 euros	Le prix au m ² diminue de 0,11 %

33 Les deux variables qui contribuent le plus au R² ajusté de cette modélisation auxiliaire sont dans l'ordre l'année d'émission de l'offre de prêt et le revenu net total pour l'année n.



Courbe de la fonction $\text{Prix/m}^2 = 102,95 \cdot \text{EQUIV_SUBV}^{-0,003821}$

2. Le segment de l'ancien (TOPE = 4)

La modélisation est effectuée sur une base comprenant 724 759 observations. La variable expliquée qui donne le meilleur modèle est le logarithme du prix au m².

La matrice de corrélations montre que les variables NPPR et SURH sont fortement corrélées. Dans ce cas, la variable surface habitable SURH a été retirée du modèle.

a. hors caractéristiques de l'acheteur

Les variables explicatives sont : ANNEE_DATP, TIMM, LOG(NPPR), ZOCO, SAISONNALITE, LOG(SURH_N), GTC_2011, CODE_GÉOGRAPHIQUE, ZONAGE_LITTORAL et LOG(EQUIV_SUBV). La variable PELG n'est pas introduite dans ce modèle : comme elle n'entre pas en compte dans le dispositif PTZ avant 2011, elle n'est pas renseignée avant cette date.

Le tableau ci-dessous présente la liste des variables dans l'ordre croissant de leur contribution cumulée au R² ajusté, ainsi que la valeur de cette contribution. Le R² ajusté est égal à 0,67.

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Zonage de tension	GTC_2011	0.4631
Logarithme du nombre de pièces	LOG_NPPR	0.5247
Logarithme de la surface par pièce	LOG_SURH_N	0.6241
Zone Robien	ZOCO	0.6498
Année de l'émission de l'offre de prêt	ANNEE_DATP	0.6591
Zonage en aires urbaines	CODE_GEOG_URB	0.6661
Logarithme de l'équivalent subvention	LOG_EQUIV_SUBV	0.6707
Type de logement	TIMM	0.6733
Zonage littoral	ZONAGE_LITTORAL	0.6742
Trimestre de l'émission de l'offre de prêt	SAISONNALITE	0.6744

La méthode « STEPWISE » n'écarte aucune variable. Toutes les variables sont significatives au sens du test de Student (p-value inférieure à 0,05).

De plus, pour identifier une dépendance excessive de l'équivalent subvention aux autres variables explicatives du modèle, une modélisation auxiliaire a été réalisée. Elle consiste en une tentative d'explication de LOG_EQUIV_SUBV par l'ensemble des autres variables explicatives du modèle. Cette modélisation aboutit à un R² ajusté de 0,01. La variable LOG_EQUIV_SUBV n'est que très peu dépendante des autres variables explicatives du modèle.

Le premier graphique ci-dessous permet de s'assurer que les résidus suivent une loi normale centrée. Le second graphique permet de contrôler l'homoscédasticité du modèle.

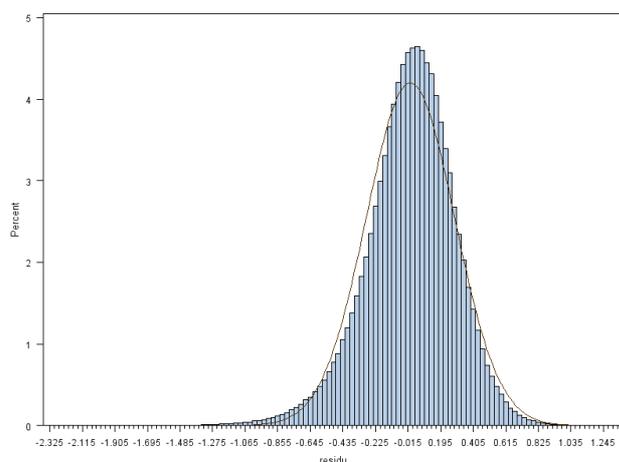
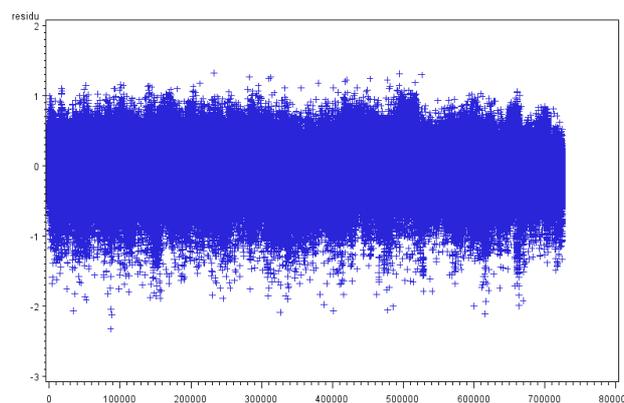


Diagramme des résidus



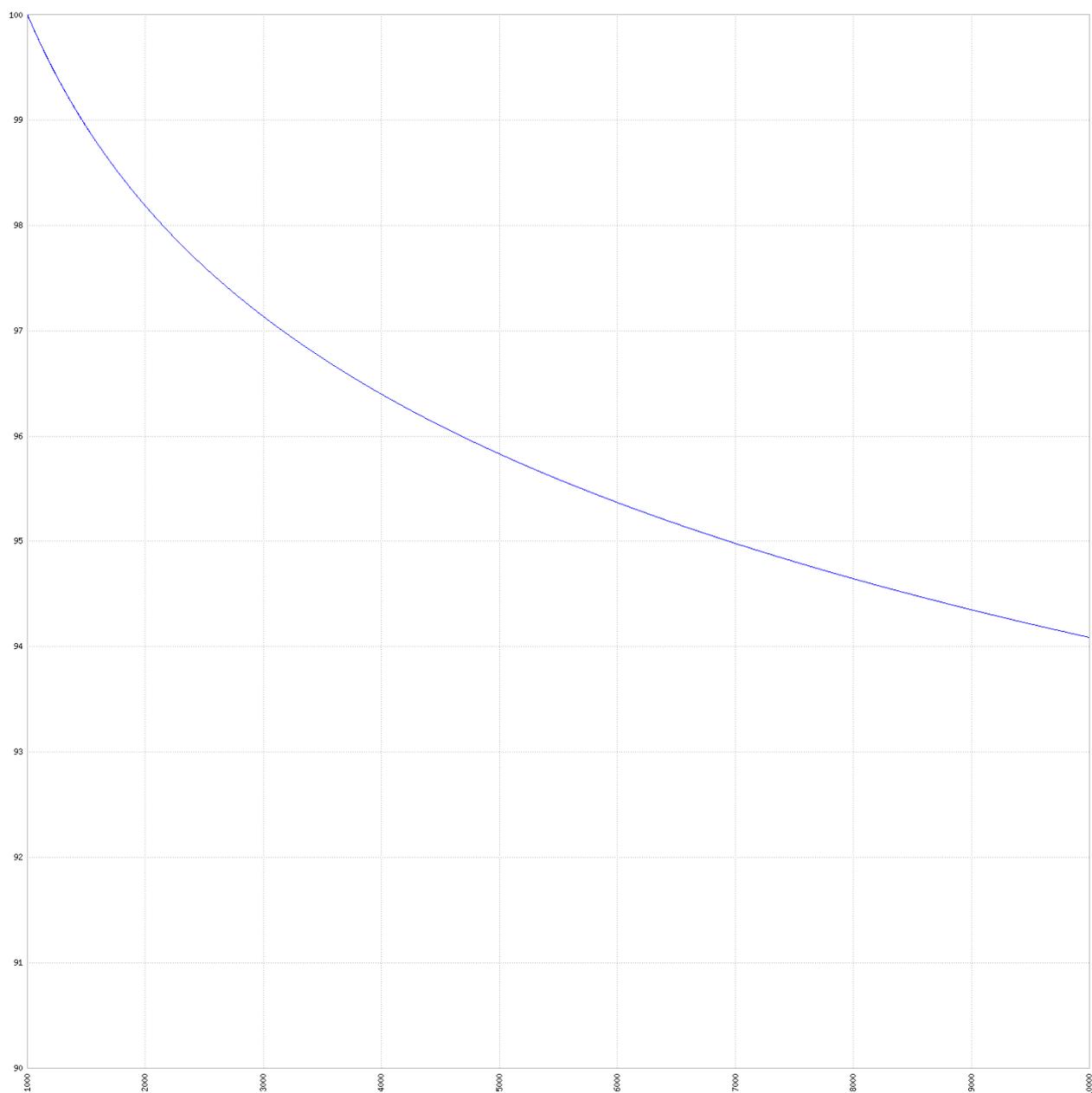
Répartition des résidus

La variable LOG_EQUIV_SUBV contribue de manière significative au R² ajusté. **Le coefficient associé à cette variable est négatif égal à - 0,026462.** Une variation de 1 euro de EQUIV_SUBV fait diminuer de 0,026462 euros le logarithme du prix au m². La courbe ci-dessous décrit le comportement du prix au m² selon l'évolution de l'équivalent subvention, en base 100 lorsque la variable EQUIV_SUBV est égale à 1000³⁴. Le tableau ci-dessous présente quelques illustrations de l'importance de l'effet de cette variable sur les prix.

Si	Alors
----	-------

34 Nous avons représenté la courbe pour un équivalent subvention compris entre 1 000 et 10 000 euros ce qui correspond aux valeurs observées pour cette période (5 % des valeurs de l'équivalent subvention sont inférieures à 1 684 et 5 % sont supérieures à 9 351).

EQUIV_SUBV passe de 1 000 à 2 000 euros	Le prix au m ² diminue de 1,82 %
EQUIV_SUBV passe de 5 000 à 6 000 euros	Le prix au m ² diminue de 0,48 %
EQUIV_SUBV passe de 5 000 à 10 000 euros	Le prix au m ² diminue de 1,82 %
EQUIV_SUBV passe de 9 000 à 10 000 euros	Le prix au m ² diminue de 0,28 %



Courbe de la fonction Prix/m² = 120,06.EQUIV_SUBV^(-0,026462)

b. avec les caractéristiques de l'acheteur

On ajoute au modèle précédent les variables explicatives suivantes : EMEM, RANN, AGE_EMPR, CCSP, NPER.

Le tableau ci-dessous présente la liste des variables dans l'ordre croissant de leur contribution cumulée au R² ajusté, ainsi que la valeur de cette contribution. Le R² ajusté est égal à 0,74.

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Zonage de tension	GTC_2011	0.4631

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Logarithme du nombre de pièces	LOG_NPPR	0.5247
Logarithme de la surface par pièce	LOG_SURH_N	0.6241
Revenu net total pour l'année N	RANN	0.7114
Zone Robien	ZOCO	0.7228
Code CSP de l'emprunteur	CCSP	0.7309
Année de l'émission de l'offre de prêt	ANNEE_DATP	0.7360
Zonage en aires urbaines	CODE_GEOG_URB	0.7401
Type de logement	TIMM	0.7418
État matrimonial du ménage emprunteur	EMEM	0.7429
Zonage littoral	ZONAGE_LITTORAL	0.7437
Nombre de personnes du ménage	NPER	0.7440
Logarithme de l'équivalent subvention	LOG_EQUIV_SUBV	0.7442
Trimestre de l'émission de l'offre de prêt	SAISONNALITE	0.7444
Âge de l'emprunteur	AGE_EMPR	0.7446

La méthode « STEPWISE » n'écarte aucune variable. Toutes les variables sont significatives au sens du test de Student (p-value inférieure à 0,05).

De plus, pour identifier une dépendance excessive de la variable EQUIV_SUBV aux autres variables explicatives du modèle, une modélisation auxiliaire a été réalisée. Elle consiste en une tentative d'explication de LOG_EQUIV_SUBV par l'ensemble des autres variables explicatives du modèle. Cette modélisation aboutit à un R² ajusté de 0,06. La variable LOG_EQUIV_SUBV est donc très peu dépendante des autres variables explicatives du modèle.

Le premier graphique ci-dessous permet de s'assurer que les résidus suivent une loi normale centrée. Le second graphique permet de contrôler l'homoscédasticité du modèle.

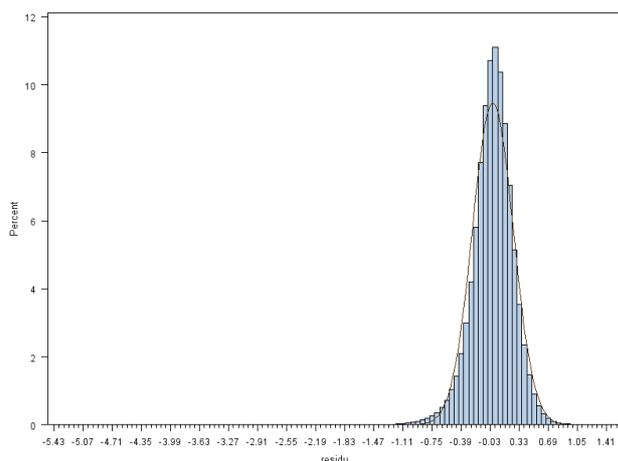
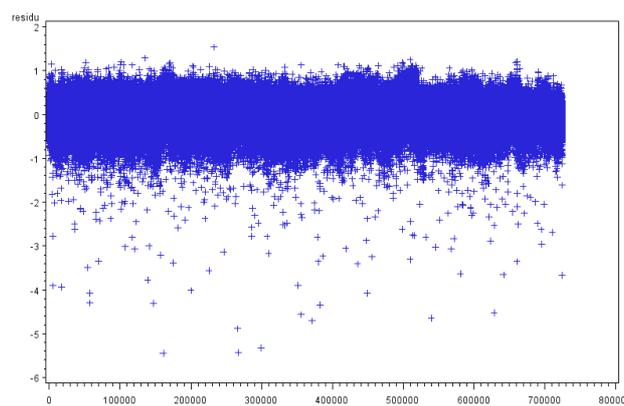


Diagramme des résidus

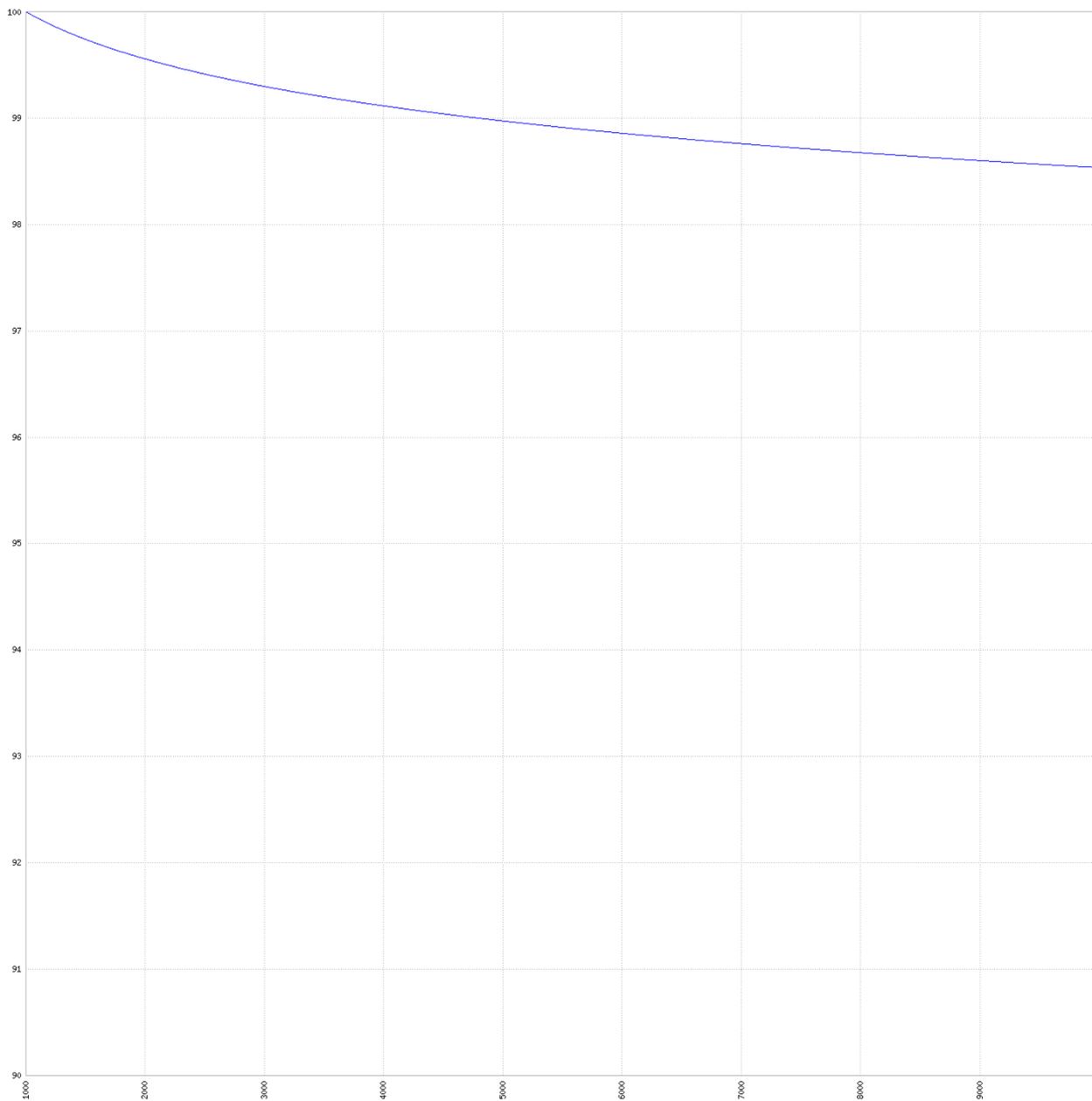


Répartition des résidus

La variable LOG_EQUIV_SUBV contribue de manière significative au R² ajusté. **Le coefficient associé à cette variable est négatif égal à - 0,006413.** Une variation de 1 euro de EQUIV_SUBV fait diminuer de 0,006413 euros le logarithme du prix au m². La courbe ci-dessous décrit le comportement du prix au m² selon l'évolution de l'équivalent subvention, en base 100 lorsque la variable

EQUIV_SUBV est égale à 1 000. Le tableau ci-dessous présente quelques illustrations de l'importance de l'effet de cette variable sur les prix.

Si	Alors
EQUIV_SUBV passe de 1 000 à 2 000 euros	Le prix au m ² diminue de 0,44 %
EQUIV_SUBV passe de 5 000 à 6 000 euros	Le prix au m ² diminue de 0,12 %
EQUIV_SUBV passe de 5 000 à 10 000 euros	Le prix au m ² diminue de 0,44 %
EQUIV_SUBV passe de 9 000 à 10 000 euros	Le prix au m ² diminue de 0,07 %



Courbe de la fonction $\text{Prix/m}^2 = 104,53 \cdot \text{EQUIV_SUBV}^{-0,006413}$

Pour la période 1 (PTZ, de 1995 à 2005)

Nous n'avons procédé à aucune segmentation de la base. En effet, la qualité du modèle est un moins bonne que soit en segmentant selon le type d'immeuble (TIMM) ou le type d'opération (TOPE).

La modélisation est effectuée sur une base comprenant 882 758 observations. La variable expliquée qui donne le meilleur modèle est le logarithme du prix au m².

a. hors caractéristiques de l'acheteur

Les variables explicatives sont : ANNEE_DATP, TOPE, TIMM, SURH, LOG(NPPR), SAISONNALITE, LOG(SURH_N), GTC_2011, CODE_GÉOGRAPHIQUE, ZONAGE_LITTORAL et LOG(EQUIV_SUBV). La variable ZOCO n'est pas introduite dans ce modèle : comme elle n'entre pas en compte dans le dispositif PTZ avant 2005, elle n'est pas renseignée avant cette date.

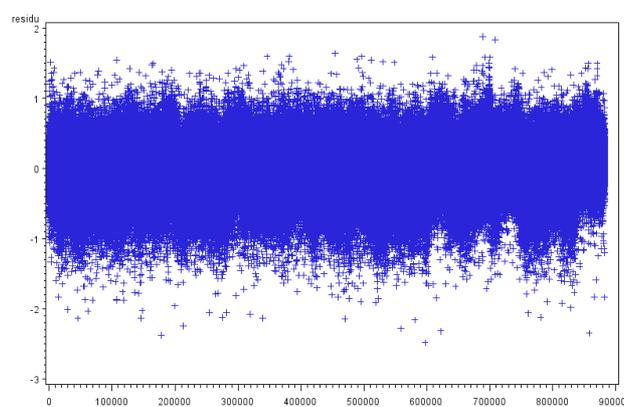
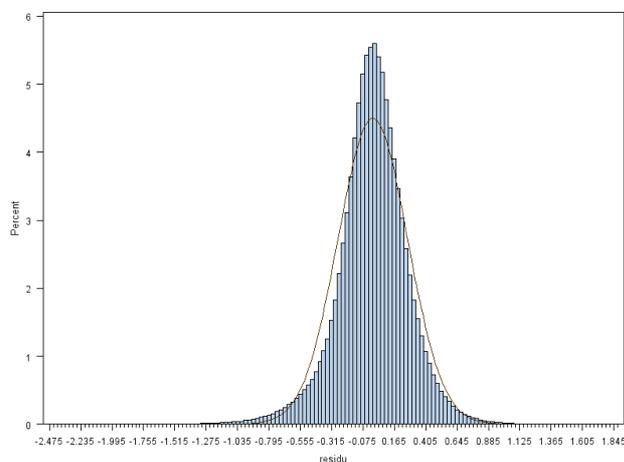
Le tableau ci-dessous présente la liste des variables dans l'ordre croissant de leur contribution cumulée au R² ajusté, ainsi que la valeur de cette contribution. Le R² ajusté est égal à 0,61.

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Type d'opération	TOPE	0.3496
Surface habitable	SURH	0.4955
Année de l'émission de l'offre de prêt	ANNEE_DATP	0.5349
Zonage de tension	GTC_2011	0.5823
Zonage en aires urbaines	CODE_GEOG_URB	0.5922
Logarithme de la surface par pièce	LOG_SURH_N	0.5993
Logarithme du nombre de pièces	LOG_NPPR	0.6051
Zonage littoral	ZONAGE_LITTORAL	0.6062
Type de logement	TIMM	0.6069
Trimestre de l'émission de l'offre de prêt	SAISONNALITE	0.6072
Logarithme de l'équivalent subvention	LOG_EQUIV_SUBV	0.6072

La méthode « STEPWISE » n'écarte aucune variable. Toutes les variables sont significatives au sens du test de Student (p-value inférieure à 0,05).

De plus, pour identifier une dépendance excessive de l'équivalent subvention aux autres variables explicatives du modèle, une modélisation auxiliaire a été réalisée. Elle consiste en une tentative d'explication de LOG_EQUIV_SUBV par l'ensemble des autres variables explicatives du modèle. Cette modélisation aboutit à un R² ajusté de 0,04. La variable LOG_EQUIV_SUBV est donc très peu dépendante des autres variables explicatives du modèle.

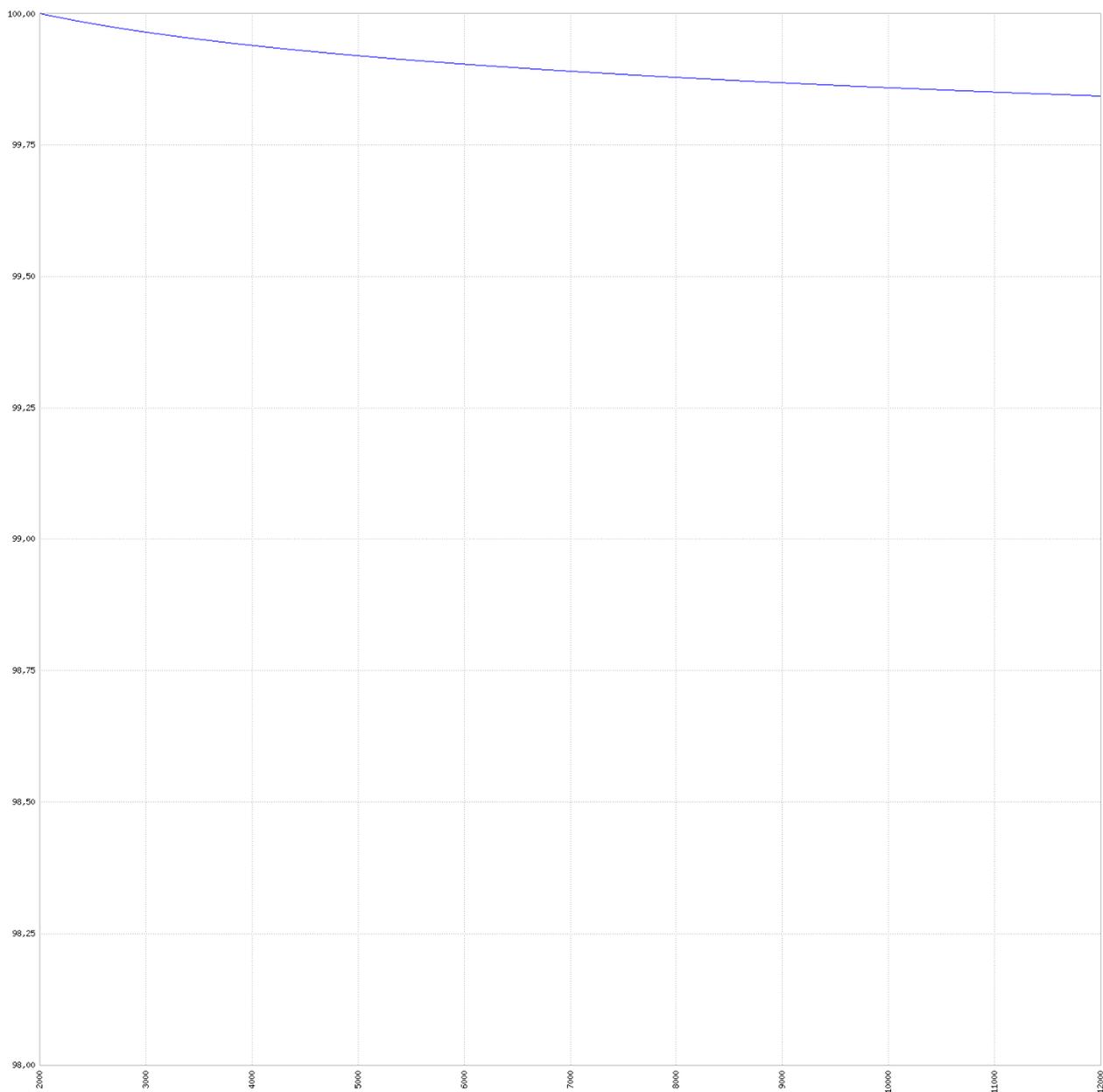
Le premier graphique ci-dessous permet de s'assurer que les résidus suivent une loi normale centrée. Le second graphique permet de contrôler l'homoscédasticité du modèle.



La variable LOG_EQUIV_SUBV contribue de manière significative au R² ajusté. **Le coefficient associé à cette variable est négatif égal à - 0,000878.** Une variation de 1 euro de LOG_EQUIV_SUBV fait diminuer de 0,000878 euros le logarithme du prix au m². La courbe ci-dessous décrit le comportement du prix au m² selon l'évolution de l'équivalent subvention, en base 100 lorsque la variable EQUIV_SUBV est égale à 2000³⁵. Le tableau ci-dessous présente quelques illustrations de l'importance de l'effet de cette variable sur les prix.

Si	Alors
EQUIV_SUBV passe de 2 000 à 3 000 euros	Le prix au m ² diminue de 0,04 %
EQUIV_SUBV passe de 5 000 à 6 000 euros	Le prix au m ² diminue de 0,02 %
EQUIV_SUBV passe de 6 000 à 11 000 euros	Le prix au m ² diminue de 0,05 %
EQUIV_SUBV passe de 10 000 à 11 000 euros	Le prix au m ² diminue de 0,01 %

³⁵ Nous avons représenté la courbe pour un équivalent subvention compris entre 2 000 et 12 000 euros ce qui correspond aux valeurs observées pour cette période (5 % des valeurs de l'équivalent subvention sont inférieures à 2 708 et 5 % sont supérieures à 11 891).



Courbe de la fonction Prix/m² = 100,67.EQUIV_SUBV^(-0,000878)

b. avec les caractéristiques de l'acheteur

On ajoute au modèle précédent les variables explicatives suivantes : EMEM, RANN, AGE_EMPR, CCSP, NPER.

Le tableau ci-dessous présente la liste des variables dans l'ordre croissant de leur contribution cumulée au R² ajusté, ainsi que la valeur de cette contribution. Le R² ajusté est égal à 0,67.

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Type d'opération	TOPE	0.3496
Surface habitable	SURH	0.4955
Revenu net total pour l'année N	RANN	0.5668
Année de l'émission de l'offre de prêt	ANNEE_DATP	0.5989
Zonage de tension	GTC_2011	0.6343
Code CSP de l'emprunteur	CCSP	0.6479

Libellé de la variable	Code de la variable	Contribution cumulée au R ² ajusté
Zonage en aires urbaines	CODE_GEOG_URB	0.6537
Logarithme de la surface par pièce	LOG_SURH_N	0.6579
Logarithme du nombre de pièces	LOG_NPPR	0.6666
État matrimonial du ménage emprunteur	EMEM	0.6699
Logarithme de l'équivalent subvention	LOG_EQUIV_SUBV	0.6708
Zonage littoral	ZONAGE_LITTORAL	0.6714
Nombre de personnes du ménage	NPER	0.6720
Type de logement	TIMM	0.6726
Trimestre de l'émission de l'offre de prêt	SAISONNALITE	0.6728
Âge de l'emprunteur	AGE_EMPR	0.6728

La méthode « STEPWISE » n'écarte aucune variable. Toutes les variables sont significatives au sens du test de Student (p-value inférieure à 0,05) .

De plus, pour identifier une dépendance excessive de la variable EQUIV_SUBV aux autres variables explicatives du modèle, une modélisation auxiliaire a été réalisée. Elle consiste en une tentative d'explication de LOG_EQUIV_SUBV par l'ensemble des autres variables explicatives du modèle. Cette modélisation aboutit à un R² ajusté de 0,07. La variable LOG_EQUIV_SUBV est donc très peu dépendante des autres variables explicatives du modèle.

Le premier graphique ci-dessous permet de s'assurer que les résidus suivent une loi normale centrée. Le second graphique permet de contrôler l'homoscédasticité du modèle, qui consiste en une vérification que les résidus se répartissent de manière homogène sur le spectre des valeurs prises par la variable expliquée.

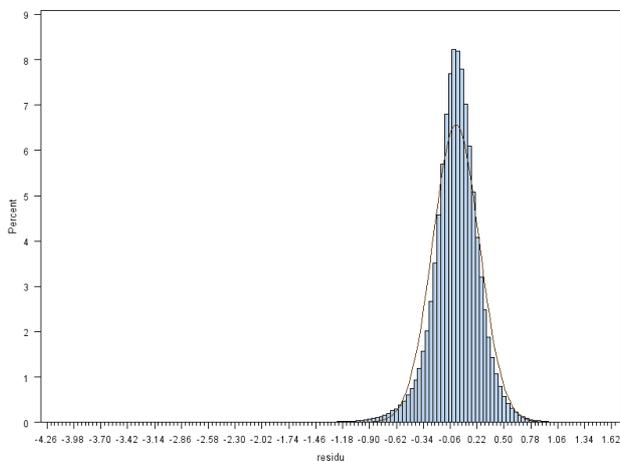
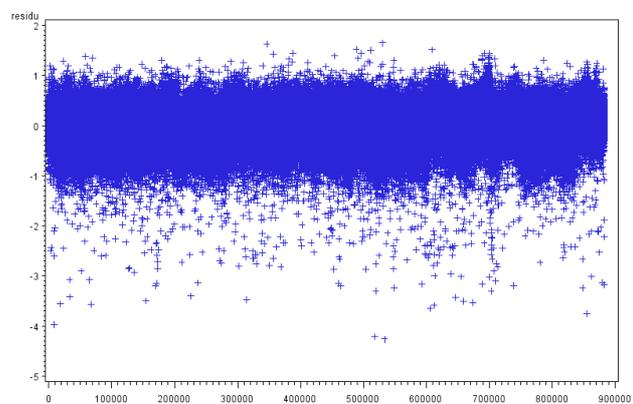


Diagramme des résidus

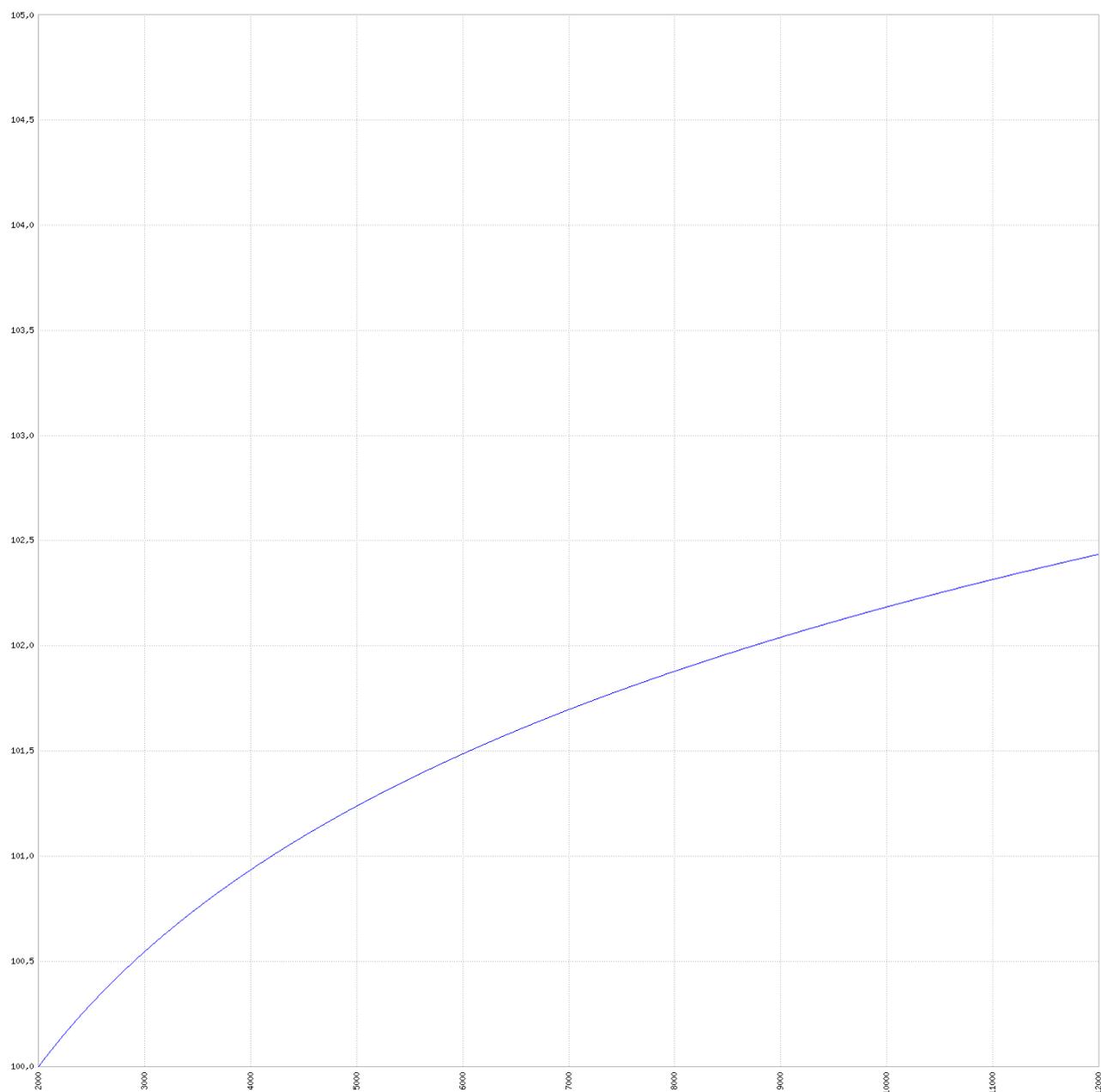


Répartition des résidus

La variable LOG_EQUIV_SUBV contribue de manière significative au R² ajusté. **Le coefficient associé à cette variable est positif égal à + 0,013429.** Une variation de 1 euro de LOG_EQUIV_SUBV fait augmenter de 0,013429 euros le logarithme du prix au m². La courbe ci-dessous décrit le comportement du prix au m² selon l'évolution de l'équivalent subvention, en base 100 lorsque la variable EQUIV_SUBV est égale à 2 000. Le tableau ci-dessous présente quelques illustrations de l'importance de l'effet de cette variable sur les prix.

Si	Alors
EQUIV_SUBV passe de 2 000 à 3 000 euros	Le prix au m ² augmente de 0,55 %
EQUIV_SUBV passe de 5 000 à 6 000 euros	Le prix au m ² augmente de 0,25 %

EQUIV_SUBV passe de 6 000 à 11 000 euros	Le prix au m ² augmente de 0,82 %
EQUIV_SUBV passe de 10 000 à 11 000 euros	Le prix au m ² augmente de 0,13 %



Courbe de la fonction $\text{Prix/m}^2 = 100 \cdot \text{EQUIV_SUBV}^{(+0,013429)}$

Discussion

Rappel des principaux résultats

Une seule des modélisations que nous avons réalisées (entre 1995 et 2005, toutes transactions et toutes variables – y-compris celles relatives aux caractéristiques des acheteurs – confondues) aboutit à un effet inflationniste relativement faible de l'équivalent subvention : le passage de l'équivalent subvention de 5 000 à 10 000 euros conduit selon ce modèle à une croissance de 0,94 % du montant total de l'opération. Toutes les autres modélisations aboutissent à un effet déflationniste. Si on constate une certaine sensibilité des résultats aux choix de paramétrage des modèles, il demeure qu'**aucun modèle testé n'a montré un effet inflationniste fort.**

Il nous semble que les modélisations les plus fiables, parmi celles que nous avons ici présentées,

sont celles incluant les caractéristiques des acheteurs. D'une part, elle améliorent significativement la valeur du R² ajusté et, d'autre part, elles introduisent des variables (revenu, nombre de personnes du ménage, voire performance énergétique) qui ont toutes les chances d'impacter la variable relative à l'équivalent subvention. Ces modèles nous semblent donc devoir être ceux qui minimisent le risque que le coefficient correspondant à l'équivalent subvention soit représentatif d'autres variables non ou mal prises en compte par le modèle.

Remarquons d'ailleurs que l'ajout des caractéristiques des acheteurs au modèle induit, pour toutes les périodes, une diminution de l'importance de l'effet déflationniste observé, et même, pour la période 1, un passage d'un effet déflationniste extrêmement faible à un effet inflationniste (très modéré lui aussi). Cela nous semble pouvoir s'expliquer par le fait qu'en l'absence des variables relatives aux caractéristiques de l'acheteur (et notamment celles qui portent sur les revenus ou la catégorie socio-professionnelle), le coefficient associé à l'équivalent subvention révèle pour partie le fait que celui-ci est d'autant plus élevé que les revenus du ménage bénéficiaires sont faibles. Si les ménages aux revenus les plus faibles de notre base achètent les logements les moins chers (en termes de prix au m²) tout en bénéficiant d'un équivalent subvention plus élevé, l'ajout des caractéristiques des acheteurs au modèle permet de révéler l'effet spécifique des revenus et a tendance logiquement à amener le coefficient associé à l'équivalent subvention vers un effet inflationniste.

Deux autres indicateurs nous semblent également déterminants pour hiérarchiser la robustesse des résultats qui précèdent :

- la valeur du R² ajusté de la modélisation principale : plus celui-ci est élevé, plus le modèle explique une part importante des prix observés :
- la valeur du R² ajusté de la modélisation secondaire : plus celui-ci est faible, plus la dépendance de l'équivalent subvention aux autres variables du modèle est faible, et donc plus le coefficient associé à l'équivalent subvention a de chances de rendre fidèlement compte de l'effet spécifique de cette variable sur le prix.

Le tableau ci-dessous présente les couples de ces valeurs selon les différentes périodes et les différents segments de parc, pour les modélisations incluant les caractéristiques des acheteurs.

	R² ajusté, modélisation principale	R² ajusté, modélisation auxiliaire	Effet d'un passage de l'équivalent subvention de 5 000 à 10 000 euros sur le prix
<i>Période 1</i>	0,67	0,07	Le prix au m² augmente de 0,94 %
<i>Période 2, neuf</i>	0,76	0,21	Le prix au m² diminue de 0,26 %.
<i>Période 2, ancien</i>	0,74	0,06	Le prix au m² diminue de 0,44 %
<i>Période 3, neuf</i>	0,80	0,78	Le prix au m ² diminue de 0,37 %
<i>Période 3, ancien</i>	0,76	0,65	Le prix au m ² diminue de 1,87 %
<i>Période 4</i>	0,85	0,46	Le prix au m ² diminue de 0,45 %.

Ce tableau montre que la modélisation qui apparaît la plus fiable est celle réalisée sur la période 2 dans l'ancien. Le R² ajusté de la modélisation principale est supérieur à 0,7, celui de la modélisation auxiliaire est inférieur à 0,1 et il s'agit de la modélisation pour laquelle l'écart entre ces deux indicateurs est le plus grand. La modélisation réalisée sur la même période mais sur le segment du neuf semble également de bonne qualité selon ces critères. Enfin, la modélisation réalisée sur la période 1 apparaît elle aussi exploitable. Pour les autres, les résultats ne peuvent faire l'objet d'une interprétation de l'effet sur le prix de la variable équivalent subvention en raison d'un R² ajusté de la modélisation auxiliaire que nous jugeons trop élevé (et d'un écart avec le R² ajusté de la modélisation principale jugé trop faible).

Les principales limites de la méthode

D'abord, il convient de rappeler qu'il ne nous a pas été possible de comparer toutes choses égales

par ailleurs les transactions dont le financement a inclus un prêt à taux zéro des autres.

Ensuite, notre modèle n'inclut pas un certain nombre de variables relatives au logement dont on sait par ailleurs le rôle sur les prix. C'est notamment le cas sur le segment du collectif de variables comme l'étage du logement, le nombre d'étage de l'immeuble, la présence d'un ascenseur, d'un balcon ou d'un garage. C'est aussi le cas de la présence d'un jardin et de sa taille pour le segment de l'individuel. Citons enfin l'année de construction qui joue un rôle important sur le prix dans l'ancien.

Enfin, aucune recherche d'autocorrélation spatiale n'a été effectuée en raison notamment de l'absence de donnée relative à l'adresse du logement objet de la transaction. Il en est de même pour l'éventuelle autocorrélation temporelle³⁶.

Interprétation d'un effet inflationniste

Beaubrunt-Diant et Maury (2015) concluent à un effet inflationniste significatif mais transitoire du PTZ [pendant la période dite du doublement du PTZ dans le neuf de janvier 2009 à juin 2010] sur le prix du foncier, de l'ordre de 7,4 %. « L'effet inflationniste moyen est de 2,4% en 2009 sur l'ensemble de l'opération (foncier + bâti) ». « l'effet inflationniste mis en évidence en 2009 absorbe plus de la moitié des gains en pouvoir d'achat consécutifs à la hausse du PTZ ». Les auteurs précisent toutefois que cet effet est transitoire « puisque l'effet inflationniste disparaît à l'horizon de 2 ans ». Ils ont obtenu ces résultats à partir des données issues de l'Enquête sur le Prix des Terrains à Bâti (EPTB) sur le segment de l'individuel pur et en comparant des communes situées en zone B2 (ayant davantage bénéficié du doublement du PTZ en raison d'un relèvement plus important du plafond d'opération) à d'autres de « profil » (taille et prix) similaires mais situées en zone C.

Ihmaïne, Snoussi et Souche (2017) concluent que le reclassement de certaines communes intervenu en octobre 2014 simultanément à une réforme du PTZ n'a pas eu pas d'effet inflationniste significatif dans les communes reclassées. Ils ont travaillé à partir des données SGFGAS sur le segment du neuf et en comparant des communes ayant changé de zone en 2014, à d'autres, limitrophes, n'ayant pas connu de changement de zone.

Les résultats de notre étude sont issus d'un modèle complet de prix et non, comme ces deux études, d'une méthode en doubles différence. Toutefois, nos résultats ne montrent pas davantage que ces deux études.

D'un point de vue théorique, l'effet inflationniste est étroitement lié à l'élasticité de l'offre par rapport à la demande. On considère en effet que l'aide publique « prêt à taux zéro » augmente la demande d'accession, soit en déclenchant un achat auprès de ménages qui sinon seraient restés dans leur logement ou auraient opté pour un logement locatif, soit en augmentant la capacité d'achat des ménages ce qui leur permet d'acheter un logement plus grand, mieux placé ou de meilleure qualité. Or, à long terme, cette demande supplémentaire peut conduire soit à augmenter l'offre (par la construction de logements ou un déplacement de la part des propriétaires occupants dans l'ensemble du parc de logements), soit à augmenter les prix des logements³⁷.

L'élasticité de l'offre par rapport à l'évolution de la demande joue donc un rôle important sur l'effet inflationniste éventuel. Plus l'élasticité de l'offre est faible, plus l'effet inflationniste est a priori probable et d'ampleur importante. La capacité de l'appareil productif à s'adapter aux évolutions de l'aide (et donc de la demande), variable selon la conjoncture et le territoire, constitue donc un facteur

36 L'autocorrélation mesure la corrélation d'une variable avec elle-même, lorsque les observations sont considérées avec un décalage dans le temps (autocorrélation temporelle) ou dans l'espace (autocorrélation spatiale). Ainsi, certaines variables peuvent ne pas être indépendantes d'un individu à l'autre en raison de la proximité géographique de ces individus. Dans le même ordre d'idée, notre modèle ne tient pas compte du fait que les modalités de la variable relative à l'année de la transaction sont ordonnées et que le prix des logements en 2000 est lié au prix des années précédentes.

37 À propos des aides aux locataires, Anne Laferrère et David le Blanc [2002] écrivent par exemple : « La théorie économique prédit que fournir une aide personnelle au locataire stimule la demande de logement locatif : les ménages vont pouvoir mieux se loger. Mais si l'offre de logements locatifs est rigide à court terme (il y a peu de logements vacants), les loyers augmenteront. À plus long terme, le niveau des loyers et la qualité des logements locatifs dépendront de l'élasticité de long terme de l'offre. Une partie de la masse des aides sera donc, au moins à court terme, récupérée par les bailleurs ». Il précisent : « Du côté de l'offre, il convient de distinguer les effets de court terme et de long terme. En réponse à la hausse de la demande de logement, l'offre réagit, plus ou moins rapidement [...]. Mais il se peut, au moins à court terme, que ce soit le même logement que, grâce à l'existence de l'aide et la pression qu'elle induit sur la demande, le propriétaire bailleur puisse louer plus cher, sans création de nouveaux logements ou sans rénovation. À l'extrême limite, dans le cas d'une offre totalement rigide à court terme, ce n'est plus au ménage aidé que bénéficie l'aide, mais au propriétaire bailleur : seule la demande augmente, avec uniquement un effet de hausse des prix ».

a priori important.

Or cette élasticité est différente sur le segment du neuf, où le rythme de la construction peut être impacté rapidement par des évolutions du dispositif PTZ, et sur le segment de l'ancien où l'offre n'évolue que par l'ampleur des mises sur le marché de logements (quels que soient leurs modes et statuts d'occupation antérieurs). Remarquons toutefois que pour la période 2 (2005-2010, caractérisée notamment par l'ouverture du PTZ à l'ancien sans quotité minimale de travaux), nos modèles ne mettent pas davantage en évidence un effet inflationniste sur le segment de l'ancien que sur le segment du neuf. Il est vrai également qu'une partie de la demande peut se reporter d'un segment à l'autre.

L'absence d'effet inflationniste significatif peut a priori s'expliquer par l'absence d'un segment de production directement dédié à la primo-accession sociale. Le vendeur ou le promoteur ne savent pas a priori si l'acheteur bénéficiera ou non d'un prêt à taux zéro pour financer la transaction. Toutefois, « l'offre de logement peut être faiblement élastique pour des raisons techniques de production, mais on peut aussi supposer que l'effet inflationniste des aides s'explique par la structure non concurrentielle des marchés » (Fack 2005). A propos des aides au locataires, « si les propriétaires ont des informations suffisantes sur les ménages locataires, ils peuvent savoir qui sont les bénéficiaires d'allocations et augmenter le loyer des locataires bénéficiaires d'aides plus que celui des ménages non bénéficiaires » (Fack 2005). Cette hypothèse est avancée à propos du prêt à taux zéro par Gobillon et Le Blanc (2005) : « Seuls les ménages relativement peu aisés (au moins relativement aux autres accédants) sont éligibles au PTZ, de sorte qu'il est possible pour les promoteurs de cibler des « produits PTZ » destinés à ce segment de clientèle et de capturer ainsi une partie de la rente constituée par l'aide ».

Ihmaïne, Snoussi et Souche (2017) écrivent toutefois que « le PTZ est protégé de produire des effets inflationnistes par la relative complexité des calculs nécessaires à l'évaluation précise de l'aide publique qui lui est associée. Il est en effet impossible d'évaluer, sans recours à un calcul actuariel, quel est l'avantage que le PTZ confère à l'emprunteur, au contraire par exemple de l'APL locative pour les étudiants pour laquelle il est facile à un bailleur de savoir quel sera le montant de l'aide perçue par son locataire. Il en résulte qu'en PTZ, il est pratiquement impossible pour le vendeur d'un terrain à construire, et a fortiori pour une entreprise de construction ou un artisan, de connaître à l'avance le montant d'aide qui n'est guère connu que de l'emprunteur lui-même et de son banquier, et encore de manière fort peu explicite, puisque c'est seulement en se livrant à des simulations de la mensualité lissée globale, tous prêts confondus, que l'avantage inhérent au PTZ peut être rendu davantage explicite ».

Interprétation d'un effet déflationniste

Un effet déflationniste signifierait que, toutes choses égales par ailleurs, l'augmentation de l'équivalent subvention lié au PTZ a un impact à la baisse sur le prix au m² des logements qui font l'objet de la transaction. Il est difficile d'expliquer ce résultat, sauf à invoquer des défauts du modèle qui conduisent à remettre en cause la réalité de l'acception « toutes choses égales par ailleurs », et par exemple les suivants :

- Ce plus faible prix au m² traduirait la moindre qualité des logements achetés. Il est vrai que notre modèle ne contient aucune information sur la qualité intrinsèque des logements (si ce n'est, dans certains cas, la performance énergétique). Aucune variable directe relative à la qualité des logements (classement cadastral, niveau de confort...) n'est disponible dans la base du SGFGAS. La base ne contient pas non plus de variables indirectes qui seraient susceptibles de donner une information sur la qualité (année de construction par exemple). Sur le segment de l'ancien, ce manque est particulièrement gênant : si les logements neufs sont évidemment habitables en l'état, il n'en est pas de même pour les logements anciens qui, pour certains sont achetés dans un état de vétusté parfois significatif, nécessitant des travaux parfois importants avant l'entrée dans les lieux, voire après celle-ci. Si le montant total de l'opération peut inclure les travaux, ce n'est pas nécessairement le cas, l'acquéreur pouvant choisir de les réaliser lui-même. Aussi, on compare sur ce segment des logements dont les prix peuvent être très différents en raison de leur état général, dont aucune de nos données ne rend compte.
- L'effet négatif de l'équivalent subvention sur le prix au m² pourrait également traduire le fait

que les ménages qui bénéficient d'une subvention élevée sont aussi ceux qui achètent dans les territoires les plus dévalorisés. Dans ce cas, c'est l'insuffisante prise en compte de la géographie des acquisitions objets d'un prêt à taux zéro qui expliquerait l'effet déflationniste observé. « D'une manière qualitative, il est souvent fait mention de « lotissements PTZ » consistant en voisinages entiers de maisons individuelles destinées à des accédants modestes, construits dans des secteurs éloignés des centres (Gobillon et Le Blanc 2005).

- Enfin, l'effet négatif de l'équivalent subvention sur le prix au m² est susceptible de traduire le fait que l'équivalent subvention est d'autant plus élevé que les revenus du ménage sont faibles. Si on suppose que les ménages qui ont les revenus les plus faibles sont aussi ceux qui achètent les biens les moins chers, cela peut expliquer l'effet déflationniste observé. Il aurait alors été souhaitable de travailler sur des ménages homogènes du point de vue des revenus. Cette dernière hypothèse nous semble néanmoins la moins probable dans la mesure où une variable relative au revenu est introduite dans le modèle³⁸.

38 Il est vrai que nous avons fait le choix d'utiliser la variable RANN, qui correspond au revenu net total du ménage à l'année n. Il s'agit du montant pris en compte pour le calcul du taux d'effort x 12, primes et compléments de rémunération inclus. Il est supérieur ou égal au revenu net imposable de l'année de référence (RANI), variable qui fait foi pour vérifier la condition d'éligibilité au PTZ et définir les conditions d'octroi de l'aide.

Conclusion et perspectives

Nos résultats, s'ils soulèvent des difficultés d'interprétation et présentent des limites, permettent de conclure à l'absence d'un effet inflationniste net du montant de l'aide sur les prix des logements acquis avec un PTZ.

Parmi l'ensemble des modélisations que nous avons réalisées, la seule aboutissant à un effet inflationniste de l'équivalent subvention montre que celui-ci demeure très limité. De plus, elle porte sur la période la plus ancienne (1995-2005).

Nos travaux ne permettent toutefois pas de conclure sur l'existence d'un effet inflationniste du prêt à taux zéro. En effet, nous n'avons pas pu comparer les transactions ayant fait l'objet d'un prêt à taux zéro à celles financées sans cette aide publique. Une piste à explorer pour aller dans ce sens pourrait être de chercher à identifier, parmi les transactions rassemblées dans la base de données DVF³⁹ celles qui ont été financées par un prêt à taux zéro, en rapprochant cette dernière base des données fournies par le SGFGAS.

Il serait également souhaitable d'améliorer la qualité des modèles présentés ici, par exemple en enrichissant les variables de localisation, mais aussi en ajoutant des variables ou en testant d'autres transformations de variables (Box Cox par exemple).

De plus, l'effet inflationniste est sans doute très variable selon les caractéristiques des territoires, ce que des études comparatives locales permettraient peut-être de mettre en évidence⁴⁰. De même, l'effet inflationniste est susceptible de varier en fonction de la conjoncture générale des marchés, ce que des études sur des périodes plus restreintes et contrastées permettraient peut-être également d'approcher.

Enfin, rappelons que notre étude n'a pas cherché à corriger l'endogénéité éventuelle du modèle ni les éventuelles auto-corrélations spatiales et temporelles.

39 Depuis 2011, les collectivités locales, l'État, et certaines structures publiques peuvent bénéficier des données « Demande de Valeur Foncière » (DVF) de la Direction Générale des Finances Publiques (DGFIP). Ces données gratuites, et actualisées tous les 6 mois, recensent de manière exhaustive toutes les transactions foncières et immobilières onéreuses sur le territoire français à l'exception de l'Alsace, de la Moselle (soumises au livre foncier) et de Mayotte.

40 Cela supposerait toutefois de réduire le nombre des observations utilisées dans la modélisation, avec des impacts négatifs sur la qualité du modèle.

Références bibliographiques

- Beaubrun-Diant, Kevin E., et Tristan-Pierre Maury. 2015. « Quels sont les effets du prêt à taux zéro sur les prix du foncier? » http://intra.dternp.cerema.i2/IMG/pdf/edhec-position-paper-quels-sont-les-effets-du-pret-a-taux-zero_cle7e15c4.pdf.
- Bouteille, Arnaud. 2015. « Les aides au logement sont-elles inflationnistes? » *La revue foncière*, n° 3: 30-31.
- Cavailhès, Jean. 2015. « Les aides au logement, une bonne aubaine pour le bailleur ». *La revue foncière*, n° 3: 28-29.
- Delarue, François, Alain Lecomte, et Patrick Vergriete. 2013. « Evaluation de la territorialisation des politiques du logement ». http://www.modernisation.gouv.fr/sites/default/files/epp/epp_politique_logement_rapport_analyse.pdf.
- Fack, Gabrielle. 2005. « Pourquoi les ménages pauvres paient-ils des loyers de plus en plus élevés?[L'incidence des aides au logement en France (1973-2002)] ». *Économie et statistique* 381 (1): 17–40.
- Gobillon, Laurent, et David Le Blanc. 2005. « Quelques effets économiques du prêt à taux zéro ». *Economie & Statistique*. <http://search.proquest.com/openview/5f8e60ad61aa116c01acb575e087c77a/1?pq-origsite=gscholar>.
- Sélim Ihmaïne, Sonia Snoussi et Pierre Souche. 2017. « Le prêt à taux zéro dans le neuf a-t-il un effet inflationniste ? | Politique du logement.com ». Consulté le mai 22. <http://politiquedulogement.com/2017/04/le-pret-a-taux-zero-dans-le-neuf-a-t-il-un-effet-inflationniste/>.
- Laferrère, Anne, et David Le Blanc. 2002. « Comment les aides au logement affectent-elles les loyers? » *Economie et statistique* 351 (1): 3–30.