



The Dynamics of Property Ownership with Spatial Markets and Policies

Rémi LEI

► To cite this version:

Rémi LEI. The Dynamics of Property Ownership with Spatial Markets and Policies. Economics and Finance. Université Bourgogne - Franche-Comté, 2023. English. NNT : 2023UBFCG003 . tel-04291427

HAL Id: tel-04291427

<https://theses.hal.science/tel-04291427>

Submitted on 17 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**THESE DE DOCTORAT DE L'ETABLISSEMENT UNIVERSITE BOURGOGNE FRANCHE-COMTE
PREPAREE AU CENTRE D'ÉCONOMIE ET DE SOCIOLOGIE APPLIQUÉES À L'AGRICULTURE ET AUX ESPACES
RURAUX**

École doctorale n°593

École Doctorale, Droit, Gestion, Sciences Économiques et Politiques

Doctorat de Sciences Économiques

Par

M. LEI Rémi

Dynamique de la propriété immobilière :
entre marchés et politiques spatialisés

Thèse présentée et soutenue à Dijon, le 26/05/2023

Composition du Jury :

M. Driant Jean-Claude
Mme Goffette-Nagot Florence
M. Gaigné Carl
Mme Dienesch Elisa
Mme Le Gallo Julie
M. Ay Jean-Sauveur
M. Dupré Olivier

Professeur Émérite, Université Paris Est Créteil
Directrice de Recherche, GATE CNRS
Directeur de Recherche, INRAE
Maître de Conférences, AMSE
Professeure, L'Institut Agro Dijon
Chargé de Recherche, INRAE
Ingénieur, CEREMA

Président
Rapporteuse
Rapporteur
Examinatrice
Directrice de thèse
Codirecteur de thèse
Co-encadrant de thèse

Remerciements

Bien que la rédaction d'une thèse puisse sembler être un travail solitaire, de nombreuses personnes ont permis à ce travail d'exister. Leur soutien et leur contribution ont été inestimables tout au long de ce parcours. Par ces mots, je souhaite les remercier.

Tout d'abord, je souhaite remercier profondément Julie et Jean-Sauveur. Ils ont été pour moi pendant ces dernières années les premiers soutiens et critiques de ce travail. Je retiendrai d'eux leur présence de tous les instants, leurs nombreux conseils, leur intransigeance, mais surtout leur bienveillance. Je mesure la chance qui est la mienne d'avoir appris à leurs côtés, m'efforçant de fournir un travail sincère afin d'honorer leur confiance. Je remercie aussi Olivier d'avoir porté ce projet dans ces prémisses, puis pour ses remarques précieuses. Je le remercie aussi pour son attention particulière pendant la crise sanitaire, et plus généralement pour sa prévenance. Je n'aurais pu imaginer meilleur encadrement pour réaliser ce travail, qu'ils en soient tous les trois remerciés.

Je remercie aussi Florence Goffette-Nagot, Élisabeth Dienesch, Jean-Claude Driant et Carl Gagné d'avoir accepté de faire partie de mon jury. Leur présence m'honore, et je ne doute pas que leurs retours riches permettront d'améliorer ce travail. Je remercie aussi Eva Simon, Benoît Schmutz et Sophie Legras pour leurs retours sur mon travail.

Je souhaite ensuite remercier le CEREMA, et plus particulièrement le département DATHA, pour son accueil. Je souhaite avoir une attention particulière pour mes collègues du troisième étage qui ont vu évoluer ce projet, tout en m'assurant de leur soutien. Parmi eux, une pensée pour Martin pour sa présence et nos discussions, Sabrina pour son apport précieux pour les relectures et conseils en anglais, Frédéric pour nos repas heureux pendant la crise sanitaire. J'ai également une pensée pour Laure, Guilhem, et Antoine pour leurs retours riches et précis à propos de la constitution de la base de données. Je remercie enfin mes collègues du CESAER, bien que mes présences aient été réduites, pour les moments échangés lors de mes différents passages. Par ailleurs, j'ai une pensée émue pour Jean Cavailhès, dont les remarques ont permis la création de ce projet.

Outre les personnes que cette thèse m'a permis de rencontrer, ce travail n'aurait pu exister sans le soutien continu de mes amis. Grâce à eux, j'ai eu la chance de partager

les joies et d'oublier les moments plus délicats. Je pense tout particulièrement à Maxime, Mary, Constance, Rémy, Grégoire, Alice, Annaïck, Lisa, Baptiste, Anoushka, Mehdi, Justine, Viktor, Quentin et Camille. Je remercie aussi Stan, Ludo et Perrine pour les moments simples, notamment pendant la crise sanitaire, qui ont rendu ces instants plus légers. Je me sens chanceux de partager leurs vies.

Je souhaite aussi remercier plus particulièrement trois personnes qui ont une empreinte personnelle sur ce travail. Gauthier, pour sa présence lors de ces années pour avoir été un soutien constant avec simplicité et compréhension. Je me souviendrai longtemps de nos discussions si futiles, mais tellement nécessaires. Jérôme, pour son soutien sans faille depuis le lycée, de demeurer si proche et pour les moments spéciaux partagés. Cécile, d'avoir été présente au début de ce travail et d'avoir suivi ce projet. Je la remercie aussi pour sa présence quotidienne sur la fin de cette aventure, de sa compréhension, et de son réconfort dans les périodes les plus difficiles. Partager cette réussite avec eux rend cet accomplissement plus intense.

Je souhaite ensuite remercier ma famille. Tout d'abord, mes parents pour m'avoir transmis le goût de l'effort et permis de m'épanouir dans les meilleures conditions. Je ne pourrais les remercier assez, et ce travail est aussi leur réussite. Je pense aussi particulièrement à Magali, Robin, Nabile, Séverine, Audrey, Lucas, Christel, Joaquín, Hugo, Léo, Didier, Isabelle, Étienne, Lucas, Valérie, Dominique, Lola, Enzo et Sarah. Je mesure l'importance de leur présence. Je souhaite aussi remercier particulièrement Franck pour sa présence continue, son regard bienveillant et franc ainsi que son affection. Je pense aussi à mes grands-parents dont les moments partagés n'ont été que trop courts, mais dont les souvenirs ne s'altèrent pas avec le temps.

Je souhaite ensuite rendre hommage à Gillette, Christiane et Yves, dont je n'accepte toujours pas l'absence. Je retiendrai d'eux qu'il n'y a pas d'apprentissage sans humilité, de réussite sans travail et de bonheur sans partage. Bien qu'il me soit douloureux de ne pas pouvoir partager ces moments avec eux, je garde près de moi leurs sourires heureux et leur amour profond. Je leur dédie naturellement ce travail.

Je souhaite terminer ces lignes en remerciant mon frère, mon premier soutien. Je le remercie pour l'exemple qu'il a été pour moi, à la fois dans la vie et dans le travail, pour les moments de joie partagés dans notre enfance, ou plus simplement notre complicité. Je souhaite surtout mettre en avant sa capacité à rendre les choses simples et à se focaliser sur l'essentiel. Sa simple présence rend les choses intenses à mes yeux, et c'est selon moi sa plus grande qualité. Partager sa vie est une de mes plus grandes chances.

à Gillette, Christiane et Yves,

à mon frère,

Contents

Résumé	17
Summary	19
Introduction Générale	21
1 Spatial Sorting and Housing Wealth Persistence: Insights from the Universe of French Homeowners	35
1.1 Introduction	36
1.2 Data Processing	39
1.2.1 Raw Administrative Sources	39
1.2.2 Housing Value Imputation	41
1.2.3 Housing Transfers	43
1.2.4 First-time Ownership	44
1.3 Empirical Results	45
1.3.1 Housing Wealth Persistence	45
1.3.2 HWP for homeowners with a constant housing portfolio	47
1.3.3 HWP for homeowners that change their main residences	49
1.3.4 HWP for homeowners that receive a housing inheritance	52
1.3.5 HWP for homeowners with rental assets in their portfolios	54
1.4 Conclusion	56
Appendices	63
1.A Data	63
1.B Additional Material	81
2 Cost-Efficiency of Supporting Homeownership: Welfare-Based Measure from France	89
2.1 Introduction	90
2.2 Context and Data	93
2.2.1 The French IFL Policy	93

2.2.2	The ABC Zoning	94
2.2.3	Data	95
2.3	Welfare-based measures of cost-efficiency	97
2.3.1	Welfare Framework	97
2.3.2	Defining Cost-efficiency	98
2.3.3	Dose-response Functions	100
2.4	Empirical Methods	101
2.4.1	Identifying Assumptions	101
2.4.2	Specification of the Propensity Scores	102
2.4.3	Specification of the Outcomes	103
2.5	Results	104
2.5.1	First-stage Models from ABC Zoning	104
2.5.2	Second-stage Models for the Outcomes	107
2.5.3	Treatment Effects and Dose-response Functions	108
2.5.4	Measures of Cost-effectiveness	109
2.5.5	Placebo Analysis for Unconfoundedness	112
2.6	Conclusion	113
	Appendices	119
2.A	Additional Descriptive Statistics	119
2.B	Credit Cost	123
2.C	First-Step: GPS Specification	124
2.D	Policy Effects	128
2.E	Placebo Analysis: Difference in Treatment	139
3	Price Capitalization of Demand-side Policies with Segmented Housing Markets	141
3.1	Introduction	142
3.2	Framework	145
3.3	Data	148
3.4	Identification Strategy	150
3.4.1	Policy Design	150
3.4.2	Strategy	154
3.4.3	Models Specification	155
3.4.4	Empirical Specification	156
3.5	Results	157
3.5.1	Results Under Homogeneity Assumption	157
3.5.2	Local Housing Market Segmentation	158
3.5.3	Falsification Test	161

3.5.4 Robustness Checks	162
3.6 Conclusion	163
Appendices	168
3.A Additional Material	168
 Conclusion Générale	 183
 References	 188
 List of Figures	 200
 List of Tables	 202

List of Acronyms

ATE Average Treatment Effect

ATT Average Treatment effect on Treated

CBD Central Business District

CEREMA Centre d'Expertise sur les Risques, l'Environnement, la Mobilité et l'Aménagement

DGFiP Direction Générale des Finances Publiques

DiD Difference-in-Differences

DVF Demande de Valeurs Foncières

DV3F Demande de Valeurs Foncières et Fichier Fonciers

GAM Generalised Additive Models

GPS Generalised Propensity Score

GWR Geographically Weighted Regression

HWP Housing Wealth Persistence

IFL Interest Free Loans

LATE Local Average Treatment Effect

INSEE Institut National des Statistiques et des Études Économiques

LIHTC Low-Income Housing Tax Credit

MID Mortgage Interest Deduction

OLS Ordinary Least Squares

PIB Produit Intérieur Brut

SDES Service des Données et Études Statistiques

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

TWFE Two-Way Fixed Effect

WLS Weighted Least Squares

Résumé

Dynamique de la propriété immobilière : entre marchés et politiques spatialisés

Cette thèse contient trois analyses économiques concernant l'évolution récente de la propriété immobilière en France. Les variations des prix des logements et des politiques publiques dédiées sont plus particulièrement abordées dans leur dimension spatiale.

Le premier chapitre s'appuie sur la constitution et l'analyse de bases de données originales, qui permettent de localiser la valeur de l'ensemble des logements en France et de les relier à leurs propriétaires privés. La relation entre les niveaux initiaux de richesse immobilière et les taux de croissance sur la période 2011-2019 met en évidence une forte persistance des inégalités entre propriétaires. La précision spatiale des données permet ensuite de préciser quelques mécanismes sous-jacents. Plus que les variations locales des prix immobiliers, il apparaît que les changements de résidences principales et les transmissions par héritages sont des déterminants clés de l'accumulation de richesse immobilière.

Le second chapitre propose d'évaluer l'efficacité de la politique française de soutien à la transition vers la propriété occupante, le Prêt à Taux Zéro. Il distingue les effets de la politique à la marge extensive (sur le nombre de propriétaires additionnels) des effets à la marge intensive (sur la valeur des logements, subventionnés ou non). La stratégie d'identification repose sur les variations spatiales et temporelles des modalités d'application de la politique, en utilisant un score de propension généralisé et des estimateurs doublement robustes. Nous ne pouvons pas rejeter l'absence d'effet à la marge extensive, alors que nous obtenons des effets positifs significatifs à la marge intensive. Pour des valeurs admissibles d'externalités (positives et négatives) produites aux deux marges, il apparaît que le bien-être social diminue avec l'augmentation des dépenses dédiées à la politique.

Le troisième chapitre traite de la segmentation des marchés immobiliers par l'effet de capitalisation causé par les politiques d'investissement locatif. Nous exploitons une singularité de ces dispositifs pour isoler un choc de demande restreint à un segment spécifique. Utilisant une approche en différence de différences, nous isolons deux chocs exogènes affectant uniquement la demande pour l'investissement locatif. Les résultats indiquent que le statut d'occupation est un élément structurant de la segmentation locale des marchés puisque le segment défini par la propriété occupante n'est pas affecté par les chocs de demande pour l'investissement locatif. Ainsi, les politiques du logement en soutien à la

propriété occupante et l'investissement locatif n'interagissent pas à court terme. En outre, la capitalisation de la subvention est prononcée sur les segments locatifs, affectant aussi le segment de l'ancien par un effet de report, bien que non subventionné.

Summary

The Dynamics of Property Ownership with Spatial Markets and Policies

The thesis report contains three economic analyses dealing with the recent evolution of housing ownership in France. Housing market dynamics and related housing policies are addressed more specifically through their spatial dimension.

The first chapter takes advantage of the construction and analysis of an original dataset, which allows to precisely locate the value of the entire housing stock and its attribution to private owners. The relationship between the initial level and the gross rate of housing wealth for the 2011–2019 period reveals a strong persistence of inequality within homeowners. The spatial accuracy of data allows to specifically discuss some potential underlying drivers of this persistence. More than the variation due to housing market dynamics, moving to another main residence and inheritance of housing appear to be the main determinants of persistence.

The second chapter aims to assess the cost-efficiency of a French policy supporting the transition to homeownership: the Interest-Free Loan policy. We disentangle the policy effects according to whether they concern the extensive margin (number of additional homeowners) or the intensive margin (price capitalisation, including for unsubsidised units). The identification strategy leverages the spatio-temporal variations of the subsidy parameters, using a Generalised Propensity Score approach and doubly robust estimators. We cannot reject that the treatment variation has no effect on the number of first-time owners, while we recover a positive and significant effect at the intensive margin. Thus, for credible values of externalities at both margins, it appears that social welfare decreases with policy cost.

The third chapter deals with the segmentation of the housing market and its implications for price capitalisation resulting from demand-side policies. We exploit the specificity of the rental investment scheme to isolate a demand shock limited to particular segments. Using a difference-in-differences approach, we isolate two successive demand shocks that specifically affect demand for rental investment. We show that tenure status is a structural determinant of local housing market segmentation. Indeed, the owner-occupied segments remain unaffected by successive demand shocks to rental investment. In this case, housing policies that jointly support owner-occupied and rental investment do not interact, at least in the short term. Nevertheless, price capitalisation is pronounced

in the rental segments and also affects the existing segments resulting from demand shifts, albeit unsubsidised.

Introduction Générale

Les marchés immobiliers ont connu un essor prononcé au début du XXI^e siècle. La valeur des logements a en effet divergé des autres biens de consommation courante, augmentant significativement la part des dépenses liées au logement dans le budget des ménages. Alors qu'il représentait 20.0% de la consommation des ménages en 1986, le logement est devenu la principale dépense atteignant 26.5% en 2019 [Sources: Portrait Social, INSEE]. Cette hausse importante est de nature à affecter la distribution des richesses entre les ménages. En effet, outre son occupation, le logement représente le principal patrimoine des ménages français puisque 61% d'entre eux possèdent au moins un logement en 2021 [Sources: Enquête Patrimoine, INSEE]. Considérant le logement comme un actif financier de confiance, une majorité des Français estime qu'il est désirable de détenir une propriété immobilière et particulièrement sa résidence principale. La hausse importante des valeurs immobilières, bien que bénéficiant aux propriétaires existants, renforce la difficulté d'accès à un patrimoine immobilier pour les locataires.

Outre les aspirations individuelles, la propriété est soutenue par les pouvoirs publics. Ainsi, le futur président Nicolas Sarkozy appela en 2007 lors de la campagne présidentielle à une "France des propriétaires" reprenant à son compte une formule de Valéry Giscard d'Estaing lors de la campagne victorieuse de 1974. La propriété immobilière s'est diffusée progressivement dans la société française, le taux de propriétaire occupant progressant de 5.7 points entre 1980 et 2021 [Sources: Chiffres Clés du Logement, Édition 2022, SDES]. De plus, le développement de la propriété immobilière s'est opéré de pair avec la libéralisation des crédits immobiliers, permettant aux ménages de s'endetter à moindre coût pour accumuler un patrimoine. Les taux d'intérêt immobiliers n'ont cessé de décroître depuis 1990 pour atteindre un plafond historiquement bas fin 2021 (10% en 1990 contre 1% fin 2021 en moyenne [Sources: Banque de France]). Cette libéralisation dans l'accès aux prêts immobiliers n'est toutefois pas sans risques, puisque la principale crise économique du XXI^e siècle (la crise des *subprimes*) trouve son origine dans la financiarisation de l'immobilier. Les répercussions ont affecté l'ensemble de l'économie, illustrant le caractère central de l'actif immobilier dans son fonctionnement.

De par son caractère immobile et localisé, le bien immobilier se distingue des autres actifs financiers. Les équilibres de marchés présentent donc une hétérogénéité spatiale prononcée au regard de leur valeur. À titre d'exemple, le prix d'un mètre carré en 2019 dans l'aire métropolitaine de Paris avoisine 7,250 euros en moyenne, alors qu'il atteint 2,600 euros à Lille ou 2,000 euros à Dijon [Sources: Calcul de l'auteur issu des données DVF]. La décision d'acquérir un bien immobilier dépend donc, outre leurs capacités financières, de l'aire métropolitaine dans laquelle les individus choisissent de résider, souvent déterminée par le lieu d'emploi. En retour, la localisation de la résidence principale détermine les coûts liés aux transports quotidiens, l'accès aux services publics ou encore la fiscalité locale. En conséquence, les choix de localisations de leur résidence principale affectent durablement les ménages par un choix de consommation, de dépenses associées (souvent variables) et d'investissement. Les conséquences économiques résultant de chocs tels que l'augmentation des coûts de l'énergie, de la mobilité ou encore des taxes sont de nature à affecter les ménages de manière différenciée en fonction de leur localisation. De fait, outre les potentiels vecteurs définis au niveau macroéconomique, l'évolution des marchés immobiliers dépend de l'évolution de l'environnement local, aboutissant à une seconde hétérogénéité spatiale en termes de dynamique.

Du fait des récentes crises sanitaire et géopolitique, l'inflation soutenue affectant les dépenses associées au choix de logement comme l'augmentation des coûts de transports et de l'énergie renforce la dimension centrale du logement comme actif financier et objet de consommation. Cette hausse des coûts associés questionne l'attrait pour la propriété immobilière dans un contexte économique incertain, incluant notamment une récente augmentation des charges d'emprunts¹. En conséquence, les politiques publiques ont un rôle déterminant pour atténuer les inégalités entre les ménages en réduisant la part de la dépense en logement dans leur budget. Le soutien à la transition vers la propriété occupante participe aussi à accroître l'offre disponible en logements locatifs par les mobilités résidentielles (Driant and Madec, 2018). Néanmoins, puisque les aides au logement représentent 38.2 milliards d'euros en 2021² [Sources: Rapport du Compte du Logement 2021] soit 1.5% du PIB, l'efficacité de ces investissements publics doit être questionnée. Ce besoin d'évaluation est appuyé par le Conseil d'État, qui, dans son rapport annuel de 2020, appelle à une plus large exploitation des travaux d'évaluation académiques pour éclairer le débat public en fournissant de nouveaux éléments de connaissance. Outre la complexité des dispositifs et de leur articulation (Driant, 2015), le principal frein à l'évaluation des politiques du logement réside dans l'accès à des données détaillées et fines. Du fait de la spécificité des marchés immobiliers, résultant du caractère immobile du logement, il est

¹1.77% au quatrième trimestre 2022 contre 1.09% en 2021. Sources: Banque de France

²Cette somme comprend l'ensemble des aides individuelles, incluant par exemple les APL ou le chèque énergie.

nécessaire d'accéder à des données géolocalisées selon la Cour des Comptes en 2022.

Contexte Scientifique et Problématique de la Recherche

Le logement se distingue des autres actifs financiers puisqu'il remplit simultanément un besoin de consommation et une demande d'investissement (Henderson and Ioannides, 1983; Henderson and Ioannides, 1986). L'actif immobilier apparaît donc particulièrement attrayant pour la classe moyenne (Garbinti et al., 2021), leur assurant la consommation d'un bien fondamental sur le long terme tout en accumulant un patrimoine par l'accès aux crédits immobiliers. En outre, bien que le logement présente un risque de perte en capital comme tout actif financier, la préférence pour la sécurisation des coûts de logement au cours du cycle de vie prédomine sur ces risques (Sinai and Souleles, 2005). Le logement, par l'inflation des valeurs immobilières, devient donc un vecteur d'accumulation de richesse, mais aussi et en conséquence, un vecteur d'inégalités économiques entre les ménages.

Une hétérogénéité spatiale multi-échelle. La dimension spatiale est prépondérante dans la compréhension des marchés immobiliers et de la dynamique d'accumulation. La valeur d'un bien immobilier dépend intrinsèquement de sa localisation à plusieurs échelles spatiales. En effet, l'aire métropolitaine, le quartier ou encore la rue sont autant de niveaux susceptibles d'affecter la valeur d'un bien. Du fait de contraintes géographiques plus fortes (Saiz, 2010), de contraintes foncières locales (Turner et al., 2014), ou d'une spécialisation spatiale des emplois et des activités (Behrens and Robert-Nicoud, 2015), l'élasticité de la construction ou la demande pour la propriété immobilière diffèrent entre les zones d'emploi. Il en résulte donc une première hétérogénéité dans les prix de marché, entre zones d'emploi. En outre, les localisations internes à la zone d'emploi ne confèrent pas des aménités équivalentes. Par exemple, la qualité de l'école de quartier (Black, 1999), l'accessibilité aux réseaux de transports urbains (Gibbons and Machin, 2005), la qualité de l'air (Anselin and Le Gallo, 2006), ou les opportunités économiques accessibles (Bilal and Rossi-Hansberg, 2021) sont autant d'aménités qui se différencient spatialement, et sont capitalisées dans la valeur du bien (Bayer et al., 2016). Il résulte donc un tri spatial en fonction des capacités financières des ménages concernant le choix du quartier de résidence (Kuminoff et al., 2013). Ce tri spatial est aussi constitutif d'inégalités de richesses par la valeur du bien acquis.

Un déficit d'information sur la propriété immobilière. La propriété immobilière demeure peu documentée au regard de son aspect central pour la richesse des ménages,

mais aussi de la société. La majorité des travaux académiques existants concernant les inégalités de richesse se focalisent sur les très hauts patrimoines (Zucman, 2019) ou encore les transferts générationnels (Boserup et al., 2016), et considèrent l’immobilier comme un actif parmi d’autres, en dépit de son aspect localisé et des conséquences dérivées. En outre, alors que la primo-accession est étudiée avec attention dans son rôle sur l’accumulation de richesse, peu de travaux documentent la concentration des actifs immobiliers parmi les ménages. Cette sous-représentation de la propriété immobilière dans l’explication des inégalités est en partie liée au manque de données accessibles. En effet, les principales sources d’informations statistiques reposent sur des données d’enquêtes menées sur des échantillons statistiques réduits³. Par ailleurs, les plans d’échantillonnage ne comportent pas nécessairement de stratification spatiale autre que la région, limitant cet axe d’analyse. Lorsque l’échantillonnage comporte une stratification spatiale, la maille est imposée à un niveau souvent agrégé, ne permettant pas une descente d’échelle nécessaire pour les besoins de l’évaluation. De récents développements issus de sources administratives ouvrent néanmoins des possibilités d’exploitation, notamment au regard de la localisation fine des observations et des individus. Les travaux pionniers exploitant des sources administratives exhaustives pour documenter les inégalités de patrimoines sont réalisés sur données scandinaves (voir par exemple Bach et al., 2020; Fagereng et al., 2020). Ces données sont les plus fines et les plus complètes actuellement au niveau européen pour documenter les patrimoines des individus et ainsi détailler les inégalités de richesse. Les économistes peuvent déterminer la maille de restitution adéquate à leur question de recherche, sans se soucier de la représentativité statistique du fait de l’exhaustivité. Néanmoins, à notre connaissance, peu de travaux traitent de la localisation de la résidence principale comme dimension des inégalités de richesse.

La nécessité de considérer le caractère spatialisé. Bien que le contexte macroéconomique puisse affecter indistinctement l’évolution des valeurs immobilières, des chocs spécifiques à certaines localisations peuvent modifier les équilibres de marchés localement. Ainsi, les variations de l’environnement local comme la création d’une ligne de transport ou la diminution de la pollution sonore sont de nature à affecter la valeur des biens par une modification des aménités locales. La prise en compte de la dimension spatiale à plusieurs échelles est donc nécessaire pour comprendre les dynamiques d’accumulation et les choix individuels qui en découlent. En outre, de nombreux facteurs définis localement déterminent au moins en partie l’efficacité d’une politique publique. L’élasticité de la construction, l’adaptation des bénéficiaires à l’introduction d’une politique ou encore les politiques locales menées conjointement sont autant de variables susceptibles d’affecter

³Par exemple, l’enquête nationale logement en cours de réalisation se focalise sur 27,000 logements, alors que l’enquête patrimoine concerne 12,035 ménages pour le millésime 2014–2015.

l'efficacité de l'intervention publique. L'absence de prise en compte de ces différences induit une évaluation partielle de la politique. En retour, la majorité des politiques du logement en France sont territorialisées, que ce soit à propos de l'éligibilité ou du montant de la subvention. L'évaluation des politiques du logement doit tenir compte de cet aspect fondamental du marché du logement, tant la territorialisation des politiques du logement est mise en débat (Madec, 2018).

Des politiques du logement peu évaluées. Bien qu'élément central de la répartition des richesses, la distribution spatiale de la propriété immobilière demeure peu documentée à ce jour et l'efficacité de l'intervention publique pour atténuer les inégalités a fait l'objet de peu d'évaluations. Dans ce contexte, nous proposons dans ce travail d'analyser sous l'angle économique ces deux thématiques. Premièrement, il s'agit d'établir la relation entre localisation et évolution du patrimoine. En effet, bien que la littérature associée à l'évolution des prix immobiliers soit abondante, il existe peu de travaux relatifs aux relations entre tri spatial des propriétaires et accumulation de patrimoine. Au regard de l'importance de la résidence principale comme vecteur d'accumulation de richesse et d'opportunités économiques, nous contribuons à lier les deux éléments. Deuxièmement, nous proposons d'établir des relations causales entre les politiques du logement et les variables d'intérêt pour leur évaluation. De par la nature du marché sur lesquelles les politiques publiques interviennent, de potentiels effets contradictoires peuvent apparaître, notamment une capitalisation de la subvention comme démontrée par Fack (2006) pour les aides locatives. Ainsi, en dépit de coûts importants engagés, la politique du logement peut produire des effets contre-productifs malgré une popularité importante, renforçant le phénomène qu'elle vise pourtant à atténuer.

Données et Méthodes

Nos trois chapitres répondent aux questions structurantes posées par l'exploitation d'une nouvelle base de données. Notre travail propose une contribution originale permettant d'observer la distribution du patrimoine immobilier au niveau individuel ainsi que son évolution. Nous tirons avantage du caractère centralisé de l'État français, en comparaison d'autres pays tels que l'Allemagne, l'Espagne ou les États-Unis. Le système fiscal français est en effet unifié avec une division spatiale fine, permettant de disposer de sources de données exhaustives, géolocalisées et normalisées à propos de la propriété et des transactions immobilières. Ces sources de données permettent d'étudier finement la propriété, à la fois dans ses dimensions individuelle et spatiale, bien qu'elles soient restreintes aux actifs immobiliers contrairement aux données scandinaves. Nous détaillons

les deux sources administratives principales, ainsi que les sources additionnelles exploitées dans notre travail.

Fichiers Fonciers. Les Fichiers Fonciers dérivent directement de la taxe foncière et fournissent des informations à propos des propriétés bâties et non bâties, ainsi que de leurs propriétaires. Cette source de données mise à disposition par services fiscaux (DG-FiP⁴) bénéficie d'un retraitement spécifique réalisé par le CEREMA⁵. La base de données est annuelle couvrant à ce jour la période 2011–2021⁶. L'identification des propriétés est exhaustive et l'information géolocalisée à la parcelle cadastrale. En outre, les caractéristiques structurelles du logement, telles que la surface, l'année de construction, le type, ou encore la présence d'équipements spécifiques (cave, piscine, terrasse, etc.) sont disponibles. Puisque ces éléments servent au calcul de la taxe foncière, nous considérons que les services fiscaux renseignent précisément ces données. L'exploitation des Fichiers Fonciers peut éclairer de nombreux phénomènes. Par exemple, la taxe foncière peut ainsi être reconstituée au niveau local, voire individuel. De plus, la consommation foncière est calculée annuellement, servant de base aux discussions sur la politique de lutte contre l'artificialisation.

Demande de Valeurs Foncières. La seconde source de données exploitée complète les Fichiers Fonciers (information de stock) en listant l'ensemble des transactions foncières et immobilières depuis 2010. Ces données permettent ainsi de qualifier les changements de propriétaires, mais aussi d'observer l'évolution des marchés locaux par l'analyse des prix de vente. Bien que la base de données brutes soit en *open-access*⁷, un retraitement réalisé par le CEREMA permet d'enrichir les informations disponibles à propos des logements, des acquéreurs et des vendeurs. La base est augmentée par les Fichiers Fonciers (passage de DVF à DV3F) en exploitant l'identifiant fiscal du logement.

Autres données. Au-delà des données fiscales exploitées dans l'ensemble des trois chapitres, nous mobilisons les données relatives aux bénéficiaires du Prêt à Taux Zéro fournies par le SGFGAS⁸. Cette base est à notre connaissance une des seules bases recensant l'ensemble des bénéficiaires d'une politique du logement depuis son introduction. Ainsi, cela nous permet de caractériser les biens acquis par les bénéficiaires de la politique, ainsi qu'un dénombrement exhaustif du nombre d'opérations par commune et par

⁴Direction Générale des Finances Publiques

⁵Centre d'Étude sur les Risques, l'Environnement, la Mobilité et l'Aménagement.

⁶Une version existe pour l'année 2009, mais l'année 2010 n'étant pas disponible, nous nous focalisons sur les années consécutives.

⁷La base de données est disponible ici.

⁸Société de Gestion des Financements et de la Garantie de l'Accession Sociale à la propriété.

année. Enfin, notre travail exploite le zonage des zones d'emploi construit par l'INSEE. Nous considérons que les zones d'emploi constituent une maille spatiale pertinente pour qualifier les marchés immobiliers puisqu'elles sont définies à la fois à partir du lieu de résidence et du lieu d'emploi des ménages.

En introduisant un identifiant invariant unique combiné au caractère annuel des Fichiers Fonciers, nous construisons des données de panel individuelles et exhaustives, à la fois relatives aux propriétaires (listant ainsi l'ensemble de leurs propriétés) et aux logements (listant ainsi l'ensemble des propriétaires par année). La géolocalisation des informations permet de mener une analyse spatiale fine, non soumise aux problèmes d'échantillonnage. Il est ainsi possible de mener des analyses locales, de choisir la maille de restitution adéquate, ou de sélectionner des observations individuelles sur la base de leur localisation, dépassant les problèmes de représentativité statistique. Enfin, le caractère longitudinal couplé à l'exhaustivité des données fiscales permet de construire des bases de données additionnelles telles que l'identification des primo-accédants ou les héritages de patrimoines immobiliers. Ces deux extensions alimentent ainsi nos travaux en documentant empiriquement des champs jusqu'ici non couverts par les données existantes.

Alors que l'hétérogénéité spatiale fournit en soi des éléments nécessaires à la compréhension du fonctionnement des marchés, il convient souvent de la contrôler, notamment lorsqu'elle reste inobservée dans le cadre d'une approche causale. En particulier, bien que la méthode hédonique représente l'approche commune pour prendre en compte les hétérogénéités observables du logement (Rosen, 1974), la spécification de la maille spatiale demeure une question empirique centrale. Il s'agit d'établir un compromis entre l'échelle à laquelle la localisation interne à la maille choisie n'affecte pas la variable d'intérêt, et le nombre d'observations permettant d'assurer une puissance statistique suffisante aux analyses empiriques. Puisque de multiples échelles spatiales sont susceptibles d'avoir une influence sur le prix, le compromis est difficile à trouver de manière exogène. L'utilisation de modèles non paramétrique représente une alternative crédible pour le contrôle de l'hétérogénéité spatiale sans définir *a priori* la maille d'intérêt (McMillen, 2010). Cette approche apparaît d'autant plus pertinente sur des données individuelles, permettant de contrôler finement l'hétérogénéité en exploitant la continuité de la dimension spatiale.

L'approche semi-paramétrique spatialisée. Parmi les méthodes qui ont émergé de pair avec le développement des capacités de calcul informatique, la structure proposée par les *Generalised Additive Models* (Wood, 2017) est attractive pour modéliser les relations spatiales. En introduisant des fonctions de lissage bivariées dont les paramètres de lissage sont définis de manière endogène afin de minimiser l'erreur, cette approche tire avantage

du caractère continu de l'espace. La contribution spatiale est en effet déterminée à partir des données, ne nécessitant pas la définition d'une unité spatiale d'intérêt souvent dérivée des frontières administratives. En outre, leur utilisation relâche l'hypothèse d'homogénéité interne à la maille spatiale imposée par l'inclusion d'effets fixes dans les modèles de régressions. En conséquence, nous introduisons une fonction de lissage spatiale en interaction avec la dimension temporelle dans notre modèle d'imputation de valeur immobilière pour contrôler de l'hétérogénéité spatiale non observée concernant le prix de marché (Ahlfeldt et al., 2023).

Les méthodes d'inférence causale. L'évaluation causale nécessite, outre le contrôle de l'hétérogénéité non observée, une approche spécifique afin de prendre en considération le caractère endogène de l'assignation du traitement. En effet, dans un objectif d'optimisation de la dépense publique, les politiques publiques étudiées dans la thèse ciblent certaines populations ou zones d'intérêt, rendant l'assignation du traitement endogène. Suivant la révolution en crédibilité connue par l'économétrie appliquée (Angrist and Pischke, 2010), les expériences naturelles basées sur les variations exogènes des politiques publiques se sont fortement développées (Wooldridge, 2016), y compris pour l'économie urbaine et régionale (Baum-Snow and Ferreira, 2015). De fait, notre travail exploite des méthodes d'inférences causales usuelles, telles que la sélection sur variables observables ou la différence de différences pour contrôler l'assignation endogène du traitement. En conséquence, nous pouvons établir des relations causales entre les politiques du logement et les variables économiques d'intérêts (*e.g.* nombre de propriétaires, prix de marché), condition nécessaire à leurs évaluations.

Les méthodes d'inférence causale et l'approche semi-paramétrique spatialisée sont complémentaires. Nous introduisons les fonctions de lissages spatiales dans la spécification du score de propension, diminuant la probabilité d'une mauvaise définition du fait du contrôle de potentielles variables non observables spatialisées (Gilbert et al., 2023). En effet, des variables non-observées avec une structure spatiale (*e.g.* l'accessibilité aux réseaux de transports, la qualité de l'air) peuvent confondre le traitement. Ainsi, la crédibilité de l'hypothèse d'*unconfoundedness*, nécessaire pour assurer la validité interne de la méthode de sélection sur observable, est accrue. Les méthodes semi-paramétriques permettant de contrôler l'hétérogénéité spatiale non-observée par l'utilisation de variables spatiales continues contribuent à accroître la crédibilité des travaux empiriques établissant des relations causales.

Plan

Cette thèse d'économie se compose de trois chapitres indépendants, traitant des inégalités de richesses, des politiques publiques ainsi que des marchés immobiliers locaux. La base de données originale que nous construisons dérivées des sources administratives alimente les trois chapitres. Nous détaillons les choix méthodologiques nécessaires à son élaboration dans le premier chapitre.

Chapitre 1. Le premier chapitre exploite la base de données pour dériver quatre résultats empiriques justifiant la nécessité de traiter simultanément le tri spatial et les inégalités de richesse. Fort du constat que la richesse immobilière persiste au niveau des zones d'emploi sur la période 2011-2019, nous documentons quatre mécanismes sous-jacents potentiels. Premièrement, l'évolution des patrimoines résultant des marchés immobiliers bénéficie aux individus avec un niveau de richesse initial important, dont la résidence principale est située proche des centres des agglomérations les plus attractives. Deuxièmement, le choix de la localisation de la résidence principale est fortement corrélé au niveau de richesse initiale des acquéreurs, mécanisme clé dans la persistance de la richesse. Troisièmement, la somme reçue de dons ou héritages augmente avec le niveau de richesse des bénéficiaires, et le statut économique de la zone d'emploi. Quatrièmement, les investissements locatifs de longue distance, bien que minoritaires, atténuent la persistance de la richesse immobilière par une redistribution partielle. En conséquence, bien que les dynamiques de marchés participent à cette persistance, ce sont surtout les choix patrimoniaux, incluant le choix de la résidence principale, et les dons et héritages qui semblent entretenir les inégalités.

Chapitre 2. Le second chapitre propose une évaluation de l'efficacité de la politique française de soutien à la transition vers la propriété occupante, le Prêt à Taux Zéro. Cette politique réduit les charges d'emprunt à la charge des acquéreurs, octroyant aux établissements bancaires une diminution d'impôt en compensation. Notre travail distingue les effets de la politique à la marge extensive (*i.e.* le nombre de propriétaires additionnels) des effets à la marge intensive (*i.e.* valeur des biens, subventionnés ou non). En lien avec la littérature sur les statistiques suffisantes (Chetty, 2009), nous menons une estimation de l'efficacité-coût pour des variations faibles des modalités de distribution de subvention. La définition de l'efficacité repose sur une fonction de bien-être social, alors que la variation de coût de la politique est distinguée selon qu'elle est causée par une décision politique ou par l'évolution du contexte économique. Notre méthode d'identification repose sur les variations spatio-temporelles de traitement, utilisant un score de propension généralisé et des estimateurs doublement robustes. Enfin, notre travail exploite le jeu de données

exhaustif pour les primo-accédants issus du premier chapitre, ainsi que le fichier exhaustif des bénéficiaires de la politique. Bien que nous ne pouvons pas rejeter que les variations des modalités d'application aient une influence sur le nombre de propriétaires, nous évaluons précisément que les externalités à la marge intensive excèdent celles produites à la marge extensive. En conséquence, pour des valeurs crédibles des externalités aux deux marges, une augmentation du coût de la politique décidée par la puissance publique affecte négativement le bien-être social. Cependant, pour une définition alternative de la fonction de bien-être valorisant positivement les externalités à la marge intensive, l'efficacité apparaît être positive.

Chapitre 3. Enfin, notre troisième chapitre traite de la segmentation des marchés locaux et des conséquences en termes de capitalisation de subvention publique. Alors que la littérature existante souligne que les politiques d'investissement locatif stimulant la demande pour accroître l'offre causent un effet inflationniste, nous discutons d'un effet de capitalisation différencié en fonction de la segmentation des marchés locaux selon le statut d'occupation. Dans ce but, nous exploitons une singularité des dispositifs d'investissement locatif (*e.g.* Pinel, Duflot) pour mettre en œuvre une méthode de différence de différences permettant d'isoler un choc exogène sur le segment subventionné. Nous identifions deux changements majeurs dans la politique étudiée entre 2013 et 2016, générant successivement un choc positif et négatif restreint à la demande en investissement dans le neuf. En supposant l'offre comme inélastique à court terme, nous vérifions l'hypothèse de segmentation parfaite en fonction du statut d'occupation par l'effet de capitalisation. Nos résultats confirment que la politique d'investissement locatif génère des effets inflationnistes à court terme. Toutefois, ces effets sont restreints aux segments locatifs résultant d'une segmentation locale prononcée en fonction du statut d'occupation. En conséquence, les politiques publiques supportant différents statuts d'occupation n'interagissent pas à court terme. L'effet de capitalisation sur le segment subventionné est toutefois prononcé, bien que ne dépassant pas le montant de l'aide moyenne.

Les principaux apports de notre travail reposent sur trois points. Premièrement, nous apportons une nouvelle base de données exhaustive à propos des propriétaires immobiliers en France ainsi que de leur patrimoine. L'exhaustivité et le caractère longitudinal ouvrent de nouvelles possibilités quant aux sujets couverts et aux méthodes accessibles pour l'inférence causale. Deuxièmement, notre travail se distingue dans la manière de traiter l'hétérogénéité spatiale. En effet, alors que la majorité des travaux existants contrôlent l'hétérogénéité par l'ajout d'effets fixes spatiaux dérivés de frontières administratives, nous préférons exploiter une fonction de lissage bivariée pour tirer parti de la localisation fine des observations. En outre, les questionnements autour de la maille pertinente pour les

effets fixes sont éludés, tout en gardant une parcimonie de la spécification de la fonction de lissage, et donc une précision suffisante. Troisièmement, notre travail contribue à la connaissance relative à la propriété immobilière, les marchés locaux ainsi que les effets des politiques du logement. Alors que la propriété immobilière entretient les inégalités individuelles et spatiales, certes par l'évolution des marchés, mais surtout par les décisions patrimoniales, les politiques du logement peinent à atténuer ce creusement des inégalités. En effet, l'efficacité de la politique de soutien à l'accession est au mieux faible, avec un effet de capitalisation de la subvention dans les prix de vente, tout comme l'investissement locatif. Ces résultats questionnent donc globalement le recours aux aides subventionnant directement la demande.

References

- Ahlfeldt, Gabriel M., Stephan Heblich, and Tobias Seidel (2023), “Micro-Geographic Property Price and Rent Indices”. *Regional Science and Urban Economics* 98.103836.
- Angrist, Joshua D. and Jörn-Steffen Pischke (2010), “The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics”. *Journal of Economic Perspectives* 24.2, pp. 3–30.
- Anselin, Luc and Julie Le Gallo (2006), “Interpolation of Air Quality Measures in Hedonic House Price Models: Spatial Aspects”. *Spatial Economic Analysis* 1.1, pp. 31–52.
- Bach, Laurent, Laurent E. Calvet, and Paolo Sodini (2020), “Rich Pickings? Risk, Return, and Skill in Household Wealth”. *American Economic Review* 110.9, pp. 2703–2747.
- Baum-Snow, Nathaniel and Fernando Ferreira (2015), “Causal Inference in Urban and Regional Economics”. *Handbook of Regional and Urban Economics, Vol 5a*. Boston, MA: Elsevier.
- Bayer, Patrick, Robert McMillan, Alvin Murphy, and Christopher Timmins (2016), “A Dynamic Model of Demand for Houses and Neighborhoods”. *Econometrica* 84.3, pp. 893–942.
- Behrens, Kristian and Frédéric Robert-Nicoud (2015), “Agglomeration Theory with Heterogeneous Agents”. *Handbook of Regional and Urban Economics, Vol 5a*. Boston, MA: Elsevier.
- Bilal, Adrien and Esteban Rossi-Hansberg (2021), “Location as an Asset”. *Econometrica* 89.5, pp. 2459–2495.
- Black, Sandra (1999), “Do Better Schools Matter? Parental Valuation of Elementary Education”. *Quarterly Journal of Economics* 114.2, pp. 577–599.
- Boserup, Simon H., Wojciech Kopczuk, and Claus T. Kreiner (2016), “The Role of Bequests in Shaping Wealth Inequality: Evidence from Danish Wealth Records”. *American Economic Review* 106.5, pp. 656–661.
- Chetty, Raj (2009), “Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods”. *Annual Review of Economics* 1.1, pp. 451–488.
- Driant, Jean-Claude (2015), *Les Politiques Du Logement En France*. 2e édition. Les Études 5414-15. France: La Documentation Française.
- Driant, Jean-Claude and Pierre Madec (2018), “Mobilité Résidentielle : Tensionmètre Des Crises Du Logement?” *Les Crises Du Logement*. 1re édition. La Vie Des Idées. Paris: Puf : La vie des idées.fr.
- Fack, Gabrielle (2006), “Are Housing Benefit an Effective Way to Redistribute Income? Evidence from a Natural Experiment in France”. *Labour Economics* 13.6, pp. 747–771.
- Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri (2020), “Heterogeneity and Persistence in Returns to Wealth”. *Econometrica* 88.1, pp. 115–170.

- Garbinti, Bertrand, Jonathan Goupille-Lebret, and Thomas Piketty (2021), “Accounting for Wealth-Inequality Dynamics: Methods, Estimates, and Simulations for France”. *Journal of the European Economic Association* 19.1, pp. 620–663.
- Gibbons, Stephen and Stephen Machin (2005), “Valuing Rail Access Using Transport Innovations”. *Journal of Urban Economics* 57.1, pp. 148–169.
- Gilbert, Brian, Abhirup Datta, Joan A. Casey, and Elizabeth L. Ogburn (2023), “A Causal Inference Framework for Spatial Confounding”. *arXiv:2112.14946*.
- Henderson, J. Vernon and Yannis M. Ioannides (1983), “A Model of Housing Tenure Choice”. *American Economic Review* 73.1, pp. 98–113.
- Henderson, J. Vernon and Yannis M. Ioannides (1986), “Tenure Choice and the Demand for Housing”. *Economica* 53.210, pp. 231–246.
- Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins (2013), “The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets”. *Journal of Economic Literature* 51.4, pp. 1007–1062.
- Madec, Pierre (2018), “La territorialisation de la politique du logement : un enjeu d’égalité”. *Hérodote* 3.170, pp. 175–194.
- McMillen, Daniel P. (2010), “Issues in Spatial Data Analysis”. *Journal of Regional Science* 50.1, pp. 119–141.
- Rosen, Sherwin (1974), “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition”. *Journal of Political Economy* 82.1, pp. 34–55.
- Saiz, Albert (2010), “The Geographic Determinants of Housing Supply”. *Quarterly Journal of Economics* 125.3, pp. 1253–1296.
- Sinai, Todd and Nicholas S. Souleles (2005), “Owner-Occupied Housing as a Hedge Against Rent Risk”. *Quarterly Journal of Economics* 120.2, pp. 763–789.
- Turner, Matthew, Andrew Haughwout, and Wilbert Van der Klaauw (2014), “Land Use Regulation and Welfare”. *Econometrica* 82.4, pp. 1341–1403.
- Wood, Simon N. (2017), *Generalized Additive Models: An Introduction with R*. Second. Chapman and Hall/CRC.
- Wooldridge, Jeffrey M. (2016), *Introductory Econometrics: A Modern Approach*. Sixth edition. Boston, MA: Cengage Learning.
- Zucman, Gabriel (2019), “Global Wealth Inequality”. *Annual Review of Economics* 11.1, pp. 109–138.

Chapter 1

Spatial Sorting and Housing Wealth Persistence: Insights from the Universe of French Homeowners

Abstract

This article studies the relationship between initial levels and subsequent changes in gross housing wealth. Using administrative microdata on housing ownership and transactions in France over the period 2011-2019, we impute a market value for all housing portfolios held by private homeowners. We document strong housing wealth persistence (HWP) for this population, particularly with respect to the location of the main residences. We then examine HWP for three sub-populations identified from the panel structure of our data. HWP appears to be mainly driven by homeowners who change their main residence or receive a housing inheritance, while it is less pronounced for homeowners with a constant housing portfolio. The spatial sorting operated by the location choices of main residences (both between and within commuting zones) appears to be a key determinant of HWP, rather than capital gains from price variations of given housing portfolios.

JEL classification: R31 ; R12 ; D31 ; C21

Keywords: Private ownership ; individual data ; housing portfolios ; capital gains ; housing transfers ; spatial inequality.

1.1 Introduction

Since it was recognized as the highest source of monetary inequality in rich countries, wealth and its distribution has received large attention since the papers by (Piketty and Zucman, 2014; Benhabib et al., 2017). Because housing is the most widespread asset in individual portfolios (Jordà et al., 2019), housing wealth persistence (HWP) has been given a particular focus in subsequent studies (Saez and Zucman, 2016; Alvaredo et al., 2018; Blanco et al., 2021; Garbinti et al., 2021).

From a financial perspective, HWP may result from increasing returns to scale of housing investments (Bach et al., 2020) and, consequently, higher returns for better endowed homeowners (Fagereng et al., 2020). In addition, access to homeownership is a significant determinant for HWP (Pfeffer and Waitkus, 2021), stemming both from monetary (Gabriel and Painter, 2020) and non-monetary transfers such as inheritances (Boserup et al., 2016). These differences of initial endowments would persist and condition inequality over the life-cycle (Huggett et al., 2011).

However, housing is also an illiquid durable consumption good (Grossman and Laroque, 1987) used by most individuals when making their location choices (Ortalo-Magné and Prat, 2016). Indeed, the location of the main residence determines the access to local amenities (Roback, 1982) and more generally, life perspectives (Bilal and Rossi-Hansberg, 2021) that in turn are capitalized in housing prices (Glaeser et al., 2005). As a spatially-fixed asset, cross-sectional variations of housing prices then induce homeowners to sort themselves across neighbourhoods (Kuminoff et al., 2013).

In this article, we provide first insights on HWP in France, by studying the relationship between initial levels and subsequent changes of gross housing wealth. For the 2011–2019 period, we leverage a new data-set about the population of private homeowners in order to trace back both the initial levels and the variations of gross housing wealth. Wealth includes rented accommodations, second homes, and inherited dwellings in addition to main residences. By estimating partial correlations between cross-sectional and longitudinal wealth variations, we document the interactions between the financial and the consumption sides of housing wealth that, taken together, are first-order drivers of HWP.

In analysing this relationship, we particularly consider the spatial dimension of HWP. Indeed, it is well known that housing derives most of its value from its location at different spatial scales (Kiel and Zabel, 2008). Considering jointly that rent-to-price ratio has decreased in land-constrained areas (Hilber and Mense, 2021) and that housing values diverge from income (Albouy et al., 2016), the location of main residences might be increasingly driven by wealth over income. Because of the larger moving cost associated to ownership (Van Ommeren and Van Leuvensteijn, 2005), the spatial sorting of households

in their location choices could amplify HWP through spatially differentiated dynamics of housing prices. For instance, as wealthier individuals are more likely to invest in local public goods (Hilber and Mayer, 2009), the attractive features of their neighbourhood are strengthened through price capitalisation (see, e.g., the numerous works about the capitalization of local amenities that followed Black, 1999; Chay and Greenstone, 2005; Banzhaf and Farooque, 2013; Diao et al., 2017). Finally, spatial sorting across neighbourhoods would produce externalities that strengthen the housing price dynamics of the most attractive locations (Guerrieri et al., 2013).

Despite the potential high effect of spatial sorting on HWP, few empirical papers simultaneously consider both processes. Using both spatial equilibrium and asset pricing models, Ortalo-Magné and Prat (2016) aim to provide “a first step” in this direction and derive some important propositions. Firstly, the decreasing marginal utility of consumption, a determinant mechanism in spatial sorting models, is not sufficient to explain location decisions when investment in housing is considered. Indeed, as tenure choices derive partly from the balance between risks associated with expected rent variations and potential capital gains (Sinai and Souleles, 2005), portfolio considerations from the whole individual wealth also affect location choices. Secondly, spatial sorting affects overall asset pricing (including bonds and stocks) through differentiated investment capabilities related to housing wealth. Thirdly, transaction and moving costs do not affect the preference for local investment.

Nonetheless, the lack of individual data about housing wealth limits the empirical approach despite recent development (Eggum and Larsen, 2021). The administrative data used in this paper come from French fiscal sources to construct a panel of the universe of French homeowners with both the locations of their main residences and their whole housing portfolio. Firstly, using property tax records (*Fichiers Fonciers*), we identify unique private homeowners and their detailed housing portfolio. Note that, independently of our work, André and Meslin (2021) also use similar data sources to provide the cross-sectional distribution of housing wealth between private owners. Our cross-sectional results for the year 2017 are consistent with theirs. Secondly, we appraise housing wealth to owners using housing value imputation model making the best of exhaustive geocoded transaction datasets (DV3F). Our method captures spatial heterogeneity both in cross-section and longitudinal way using bivariate smoothing splines in interaction with time (Wood, 2017). Thirdly, we improve current transaction datasets derived from tax sources by identifying changes in ownership due to inheritance and first-time ownership.

The definition of the spatial dimension to provide empirical insights on the HWP is based on the commuting zone. Commuting zones are commonly assumed to be a

consistent spatial unit for local housing markets as decision for residence places and job opportunities are closely related (Zabel, 2012). To draw partial correlation of interest, we choose to rank commuting zones according to a continuous variable over the introduction of fixed effects. Our choice is motivated by the tractability it provides, especially for interaction specifications. We therefore use the built-up area in 2000 as a proxy for the economic status of commuting areas. The underlying assumption is that built-up area is correlated with the level of amenities and productivity (Saiz, 2010, Proposition 2). In addition, we treat heterogeneity within commuting zones using distance from the Centre Business District (CBD), following the usual monocentric approach (Duranton and Puga, 2015). Remark that the simplicity of the spatial dimensions enables to interact both, without loss of generality. Empirically, we define exogenously commuting zones and their CBD using commuting zones supplied by the INSEE.¹

Our contribution concerns both the data we bring, and the empirical results we derive from it. The new exhaustive dataset on homeowners and their entire housing portfolios for the period 2011-2019 in France is a major contribution, and is likely to feed further empirical work. From their use, we provide first insights based on partial correlations about the HWP through the spatial and individual dimensions. Our four empirical findings concern four different populations that suggest potential drivers of HWP. First, wealth accumulation that stems from the housing market dynamics is correlated with the initial level of wealth at the individual level and, more importantly, with the location of the main residence. Second, HWP is likely to result from housing choices for the main residence. Both transaction price and spatial sorting are correlated with the initial level of housing wealth. Third, housing transfers mainly benefit the wealthiest homeowners, as the share received increases with the initial level of wealth. Moreover, the value of the transfer increases with the economic status of the commuting zone in which the recipient lives. Fourth, the ownership of long-distance rental investments mitigate the HWP through the spatial dimension.

Our paper is structured as follows. We detail the administrative data sources we gather to build exhaustive panel data about owners over the 2011–2019 period (Section 1.2). It includes methodological choices associated to the main extension we make. Then, we provide our empirical analysis in Section 1.3. Conclusions and possible extensions are presented in Section 1.4.

¹French National Institute of Economic Studies and Statistics

1.2 Data Processing

We first present the administrative data sources we use, and how we construct panel data about owners and housing properties over the 2011–2019 period. In addition, we detail our housing market imputation (Section 1.2.2), which is required to assess both initial levels and trends for the HWP, the identification of transfers (Section 1.3.4) and the identification of first-time owners (Section 1.2.4). Our results in the cross-sectional dimension are close to those obtained by André and Meslin (2021).

1.2.1 Raw Administrative Sources

Our database exploits three raw data sources i) a housing stock dataset with matched owners ii) a housing transaction dataset iii) commuting zone perimeter. Except from the commuting zone perimeter, all data sources are derived from fiscal sources.

Housing Stock Database This database, entitled *Fichiers Fonciers* provided by the CEREMA,² contains information about housing stock using to property tax collection. For each January 01st, we observe the housing stock in France and detailed observations about current owners. Data sources are exhaustive about both owners and housing properties. Private owners are defined according to their civil state, with the date of birth, gender and current address, while legal persons are classified according to the structure type (public, private, social landlords) and identified with national ID. Furthermore, each housing is identified by a unique national ID and its structural features, including all relevant characteristics for tax assessment such as surface, housing type, building years or presence of particular facilities (e.g. swimming pool, cellar, parking lot). We expect these characteristics to be precisely reported as they condition the property tax value. In addition, we observe the housing location using the centroid of the parcel where the housing belongs. As parcel sizes are tiny,³ observations are precisely located. Finally, a property right table is available to assign properties to owners each year. Each property right is described according to its type, following the French law definition (full owners, usufruct, bare owner). We provide detailed statistics about the housing transaction dataset in Section 1.A.1 including a description of the housing stock through detailed housing characteristics and spatial distribution.

²Centre d'Expertise sur les Risques, l'Environnement, la Mobilité et l'Aménagement.

³We estimate that the average size for a parcel in France about 0.57 hectares. It represents the area covered by a 75-metre square.

Housing Transaction Dataset We also access administrative files registering all housing transactions in France since 2010 (DV3F).⁴ Housing is defined with similar variables as the housing stock database (as they both derive from fiscal sources), including structural characteristics, housing location and unique identifier. Furthermore, additional information with respect to transactions, such as dates, nature (sale, exchange, expropriation, etc.) and the purchase price are available. Nonetheless, we do not observe the financial conditions related to the operation such as equity and mortgage reimbursement conditions. Finally, the data source provides information on the type of seller and buyer according to whether they are private individuals or legal entities. Transaction datasets contain 9,158,323 transactions about 10,774,349 housings over the 2010–2020 period. We report additional descriptive statistics are available in Section 1.A.2 about transaction price, housing characteristics and spatial distribution.

Commuting Zones We adopt commuting zone as stable unit for local housing market. We exploit the zoning supplied by the INSEE for 2010. Commuting zones are defined based on working and residence place for most inhabitants, leading to stable and consistent geographical units for both housing and labour markets. We adopt as the centre of the commuting zone, the chief town of the municipality with the highest density within the area. Instead of introducing commuting zone fixed effects, we classify the commuting zone using a continuous variable. This allows us to interact both dimensions of spatial heterogeneity (within and between commuting zones) with tractable and transparent results that do not result from the fixed effects approach. While the fixed effects approach is more precise, it requires a univariate analysis that does not allow for interactions. Nevertheless, we expect heterogeneity within commuting zones to vary with their characteristics. Our continuous classification for commuting zone proxy for economic status with the built-up area in 2000, in line with Saiz (2010). We assume that top commuting zones in terms of economic status have high built-up areas. Empirically, we consider as built-up parcels, land with at least one construction built prior to 2000 using the housing stock database. We report additional statistics about the spatial distribution of the built-up area in 2000 in Section 1.A.3.

Making the best of administrative data sources, we construct panel data about French homeowners over the 2011–2019 period. We introduce a new homeowner ID that is based on civil status rather than address or department. This allows the ID to be considered invariant over time, which is an appealing property for building panel data. Our ID construction takes account for potential misspecification in civil state using names trunca-

⁴Only the region Alsace-Moselle is not available due to historical reasons. These areas have an alternative system, named as *Livre Fonciers*, inherited from a German law in 1896. Recall that Alsace-Moselle belonged to Germany between 1870 and 1918.

ture.⁵ Consequently, we overcome the main shortcoming of the current ID, which is based on French departmental boundaries.⁶ Then, the current ID is irrelevant to study housing wealth both in cross-section (e.g. homeowners can have assets in distinct departments) and longitudinal (e.g. individuals can change their residence place) dimensions.

Empirically, we retain as civil state variables birth name (being more stable over time), first name and days of birth. The improvement is sizeable: while we estimate the number of unique private owners to 41.6M over the 2011–2019 period based on the current ID, this estimation decreases to 30.7M using the unified ID.

1.2.2 Housing Value Imputation

Beyond the construction of panel data about private owners and their properties, we improve fiscal data through the imputation of housing market value to appraise owners' housing wealth. Although the number of properties is a good proxy for wealth, the valuation of housing allows the derivation of the gross housing wealth for owners and its evolution required for the HWP. Our imputation method to estimate unitary housing price for each observation account specifically for spatially heterogeneous trend of the housing market over the 2011–2019 period. These approaches are mainly used for tax property purposes as it aims to retrieve transparent and fair values for tax purpose that overcome main issues related to self-reported values (Tur-Sinai et al., 2020). In addition, researches have focus on appraisal models to understand the determinants of housing or rent prices especially for the spatial dimension (Ahlfeldt et al., 2023).

The housing value imputation procedure must account for various sources of heterogeneity. Apart from housing characteristics, the location is determinants for multiple spatial scale (Lee and Myers, 2003; Kiel and Zabel, 2008). While controlling for housing characteristics such as surface or housing type is most of the time straightforward, spatial heterogeneity is more difficult to model. Our imputation model introduces spatial coordinates using bivariate smoothing splines, in line with the Generalised Additive Models framework (Wood, 2017), to capture spatial heterogeneity. It takes advantage of the continuity of spatial coordinates available in administrative sources, with data-driven definition. Indeed, the effective degree of freedom for the spatial contribution is endogenously defined to fit the best to the data. The intuition is simple, as models with higher degree of freedom are more likely to capture the local singularity, but decrease the precision.

Our imputation model has many advantages for an implementation at the country level

⁵We provide an error assessment using type I and type II distinction according to civil state truncating to find the option that fits the best.

⁶Departments are administrative boundaries for France introduced in 1789 following the *French Revolution*. It corresponds to NUTS-3 level in Nomenclature of Territorial Units for Statistics.

compared to common approaches to mass appraisal, such as a spatial fixed effects approach estimated by Ordinary Least Squares (OLS) or a Geographically Weighted Regression (GWR Brunsdon et al., 1996; Fotheringham et al., 2002). Firstly, it relaxes the need to find the most appropriate spatial unit for the fixed effects approach based on the trade-off between statistical power and granularity. Secondly, it relaxes the assumption that price is homogeneous within the spatial unit conditional on housing characteristics for fixed effects as well (McMillen, 2010). Thirdly, although recent developments have addressed computational cost issues (Li et al., 2019; Murakami et al., 2020; Li and Fotheringham, 2020), GWR remains difficult to implement on a national scale. Although it does not allow the marginal contribution of covariates to vary across space, we expect the spatial smoothing approach to be the best compromise between accuracy and tractability. Making an accuracy comparison on smaller datasets, we consider that the GAM approach is more accurate than the OLS, and performs at least as well as the GWR (for more details about the accuracy comparison procedure, see Section 1.A.6).

In addition to spatial locations, we introduce variables to control for heterogeneity in housing characteristics. Nonetheless, our variable set is exogenously selected for computational reasons and consistency at a national level. Each housing transaction is defined based on its housing surface, housing type, dependence surface, and building age. We do not introduce additional variables such as the average surface per rooms to reduce computational costs. We introduce the spatial coordinates taking advantage of geocoded nature of the observations, using bivariate smoothing thin plates, in interaction with time dummies. Our objective is to capture both cross-sectional and dynamic heterogeneity. The introduction of additive smoothing splines aims to capture for different scale levels that determine housing prices Kiel and Zabel (2008). The mass appraisal model is reported in Equation 1.1.

$$y_{it} = \alpha + h(z_i, t) + \sum_{j=1}^J f_j(x_{it}) + \mathbf{X}\beta + \varepsilon_{it} \quad (1.1)$$

where y_{it} is the outcome for observation i at time t (unitary price); z_i are the spatial coordinates for the location of housing i ; h represents the spatial smoothing function; x_{it} is the $i - th$ selected variable in our mass appraisal; f_j represents transformations functions using additive splines; and ε_{it} is the idiosyncratic error term. Empirically, we define f_j and h as additive variable transformation. Variance of errors is minimised by penalised least squares, while smoothing parameters are estimated using restricted maximum likelihood approach. Using the estimation of α , f_j , h and β (respectively noted $\hat{\alpha}$, \hat{f}_j , \hat{h} and $\hat{\beta}$), we estimate housing market price for housing asset i at time t

$$\hat{y}_{it} = \hat{\alpha} + \hat{h}(z_i, t) + \sum_{j=1}^J \hat{f}_j(x_{it}) + \mathbf{X}\hat{\beta} \quad (1.2)$$

As our dataset is large (more than 7M transactions, 36M of housing), we split the procedure by commuting zone to mitigate computational issues. Our imputation model is then composed of 293 independent models. It also has practical advantages, as it handles geographical discontinuities such as the Mediterranean Sea for a pooled model as Corsican commuting zones are spatially disjoint from others. We also introduce observations located in a 5-km area from the commuting zone area to prevent from border effects. According to Equation 1.2, we estimate the appraised value of the entire housing stock each January 01st during the 2011–2019 period. We report from Section 1.A.6 to Section 1.A.10 additional results for the imputation procedure, including partial correlation for independent variables, joint significance for additive smoothing splines, and imputed housing values.

Our housing market price imputation confirms sizeable heterogeneity both in the cross-sectional and the longitudinal dimensions. Although the unitary housing price based on location is left-skewed,⁷ unitary housing price ranges from 0.5k to 10k euros per square metres. It reinforces the spatial dimension of housing fundamental values. Moreover, the housing price dynamics also exhibit spatial heterogeneity, as annual gross rate varies from -2.5% to 5.0% .

Our price imputation model confirms the ability of local housing markets to generate heterogeneous capital gains (and loss). In addition, we retrieve sizeable cross-sectional differences both within and between metropolitan areas that support the ability of housing assets to drive inequalities.

1.2.3 Housing Transfers

We exploit the joint exhaustive feature of the transaction dataset and the panel data to identify housing transfers indirectly. We merge the transaction dataset to housing experiencing a change in their owner's composition making the best of common housing ID.⁸ We condition the merging process using the temporal dimension as owners' change must occur in the same year or the year after than the transaction to prevent for potential delay in fiscal source update. We consider our merging process as valid as we match more than 99.15% of housing transactions restricted to private owners, which are those likely to be

⁷For detailed results, see distribution on Section 1.A.11

⁸From the panel data with housing as the observation unit, we identify sequentially, housing which experience change in the owner's composition.

concerned by housing transfers (for detailed results about methodological choices for the transfers identification, see Section 1.A.13). We expect to remain unmatched observations to arise from reverse mortgage operations and transactions for usufruct rights.

Our main contribution to the exhaustive identification of housing transfers with detailed characteristics has two shortcomings resulting from methodological choices. Firstly, as we observe owners' changes sequentially every year, we omit multiple changes within a year. Then, if one transfer occurs the same year as a housing transaction for a particular observation,⁹ our method fails to identify the change resulting from housing transfers. Secondly, we only identify housing transfers *between* private owners leaving rental investment companies out of scope. As a consequence, it brings two recommendations for the interpretation of the results. Firstly, our measure of the amount of housing transfers is likely a downward estimation. Secondly, our resulting transfer dataset is not suitable to study housing portfolio choices consecutive to receiving a transfer, as quick resale are not identified. Despite these two shortcomings, the transfer dataset provides detailed information about housing being transferred or characteristics of both legatees and recipients.

1.2.4 First-time Ownership

Finally, we make the best of the longitudinal feature of the data derived from fiscal sources to identify first-time owners according to the housing policy definition.¹⁰ Empirically, we observe annually whether individuals are owners of their main residence. We consider as first-time owners, new homeowners who have not owned their residence for two years. The housing policy definition has practical advantages in that it does not require tenure status to be considered over the whole life cycle. Consequently, the reduced time span for the panel data is not bounding for the identification.¹¹

The lack of common definition for first-time ownership makes the comparison difficult. The housing policy definition is likely to provide an upward estimation of first-time owners in comparison with the statistical definition provided by the INSEE. In addition, despite the growing interest for first-time ownership in academic works according to their implications on a life-time perspective, our identification is the first to exhaustively identify this population as far as we know. Making the best of the fiscal data, we observe their housing choices, including surface or location choices.

⁹Assuming that housing change of owner from O_1 to O_2 , and consecutively O_2 to O_3 within the same year, we only observe change O_1 to O_3 from the housing stock database.

¹⁰The policy definition is less restrictive in regards with the statistical one. It requires that households have not owned their main residence for a period of two years.

¹¹The only limitation is that we precisely estimate the number of first-time owners from 2013, as we require to observe tenure status in 2011 and 2012.

1.3 Empirical Results

Our panel data set for private homeowners and housing in France covers a 9-year period. We identify 34.35M unique owners with at least one property over the 2011–2019 period. We estimate that the housing stock is composed of 38.17M unique housing, including both new housing and demolished one.

The average gross housing wealth within owners equals in 2019 (respectively 2011) 170.1k euros (respectively 158.4k) per owner, each owner having 1.58 properties (respectively 1.58) on average. Average owner is 55.9 years old in 2019, while it is 54.5 years old in 2011. We estimate that the number of individuals with at least one property in 2019 (respectively 2011) equals 29.17M (respectively 27.67M). In addition, the average housing assets in 2019 is appraised to 176.7k euros (respectively 165.6k). Finally, we estimate that transfer represents 36.2% of owners change over the 2011–2018 period. The mean age recipient is 47.3 years old, while average legatees is 86.6 years old.

We illustrate the HWP (recall Housing Wealth Persistence) at a commuting zone scale. Then, we seek potential drivers, including the housing market dynamics (Section 1.3.2), the location decisions for main residence (Section 1.3.3), the housing transfers (Section 1.3.4), and the property of rental assets (Section 1.3.5).

1.3.1 Housing Wealth Persistence

We first document the HWP for the commuting zone area perimeter. We simply compute the average level of housing wealth for residents in commuting zone in 2011 (Panel A, Figure 1.1) and the average annual gross rate for wealth (Panel B, Figure 1.1), without targeting specific population of homeowners. We then regress the gross rate over the 2011–2019 period on the initial level of wealth to provide evidence that the cross-section differences in housing wealth between commuting zones persist over time.

Although it is not deterministic ($R^2 = 0.26$), we estimate a positive relationship between the gross rate and the initial level of wealth for commuting zones (Figure 1.1, Panel C). Some commuting areas are an exception, such as the Luxembourg border and the Bordeaux area, which have the highest gross rate. Conversely, the Mediterranean coast, despite its wealthy inhabitants, offers a lower accumulation rate for the 2011–2019 period. Nevertheless, the wealth evolution of housing wealth is more pronounced for the commuting zones with the highest average wealth in 2011.

The HWP can be driven by several factors. On the one hand, the dynamics of the housing market affect the accumulation of housing wealth through differentiated capital gains. Our imputation procedure supports the potential for heterogeneous wealth evo-

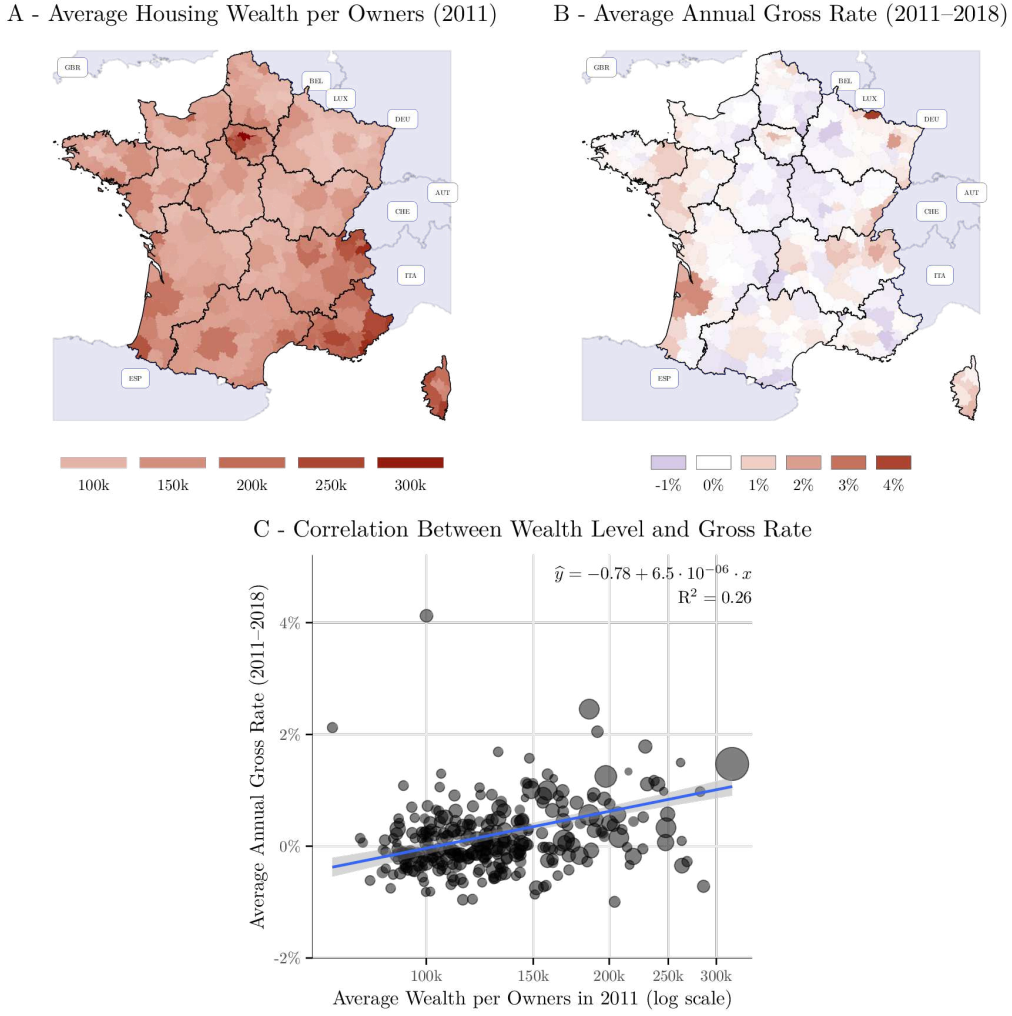


Figure 1.1: Housing Wealth Persistence between Commuting Zones of homeowners' main residences

Notes: We report the average initial level of housing wealth for homeowners (Panel A) and the average annual gross rate over 2011–2019 (Panel B) according to their place of residence. Our results are reported by commuting zone. We also report the linear relationship between the two variables (Panel C). It represents the simple WLS estimation between the average annual gross rate (y-axis) and the initial level of wealth (x-axis). We weight the commuting zone according to the number of homeowners in 2011 (size of the dot).

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

lution arising from housing market dynamics. However, the potential spatial differences between the location of assets and the owners' main residence may redistribute capital gains across space. The wealth evolution for homeowners is thus the average of asset evolution based on their locations. On the other hand, changes in the housing portfolio resulting from the moving to another main residence, inheriting a house or acquiring a rental property have a direct influence on housing wealth evolution. We then document the relationships between housing market dynamics, portfolio choices and HWP.

1.3.2 HWP for homeowners with a constant housing portfolio

While the local dynamics of housing markets are well described in the housing literature, we first document how it contributes to the housing wealth persistence at an individual level. Yet, we expect the housing wealth appreciation to differ at least slightly from local housing markets dynamics. Indeed, we identify two drivers with opposite effects. On the one hand, the prevalence of owner-occupation in individual portfolios suggests that wealth accumulation follows the same pattern as housing market dynamics. On the other hand, as the majority of the French housing stock is held by multiple owners,¹² the locations of owners and assets are likely to differ. It is then mainly the prevalence of local assets that determines the similarities between housing market dynamics and local housing wealth development.

To isolate the correlation between housing market dynamics and HWP, we select owners who do not experience portfolios' change between two consecutive years. By doing so, any variation of individual portfolios arises from changes in housing asset imputation, and therefore market evolution. Every year, we estimate that more than 93.7% of homeowners have no changes in their housing portfolios. We annually select nearly 25M of individuals between 2011 and 2018.

Our outcome of interest is the housing wealth variation (noted $w_{it+1} - w_{it}$) between two years. The logarithm transformation normalises the wealth accumulation according to the sum being invested.¹³ We introduce as explanatory variable the initial level of wealth and spatial location of residence place to discuss the HWP. The location of residence place is defined according to the economic status of the commuting zone (heterogeneity between) and the distance from the CBD (heterogeneity within). In addition, we control by individuals' age. It yields annually to the estimation of

$$\log(w_{it+1} - w_{it}) = \alpha_t + h_t(d_{it}, \ell_{it}) + f_{1t}(w_{it}) + f_{2t}(a_{it}) + \varepsilon_{it} \quad (1.3)$$

with w_{it} gross housing wealth for individual i at time t ; α_t intercept, h_t bivariate smoothing function to account for the distance to the CBD (d_{it}) and built-up area for commuting zone (ℓ_{it}) based on individuals' location; a_{it} individual age; while ε_{it} represents idiosyncratic error term. We do not adopt pooled models to avoid computational issues and estimate Equation 1.3 separately for each year from 2012 to 2018.

Relationship of interest are respectively f_{1t} and h_t as it relates respectively to the ini-

¹²We estimate that almost 60% of the housing stock is held by individuals with at least two dwellings. André and Meslin (2021) estimate this share to be two thirds in 2017.

¹³It mainly overcomes the main drawbacks of level variation. For instance, a 10k increase is not as meaningful for a 500k housing wealth than a 50k one.

tial level of wealth and spatial locations. We consider that both functions are relevant to the HWP. Firstly, any positive relationship between accumulation rate and initial level of wealth support the HWP individually. Second, spatial heterogeneity can either amplify or mitigate differences in housing wealth. Empirically, both functions are specified using additive spline transformations. The variance of errors is minimised by penalised least squares, while smoothing parameters are estimated using restricted maximum likelihood criteria. For clarity reasons, we report the spatial contribution for three distinct commuting zones based on their economic status. We range built-up area from 30% (low commuting zone) to 70% (top commuting zone) with 50% as intermediate value.

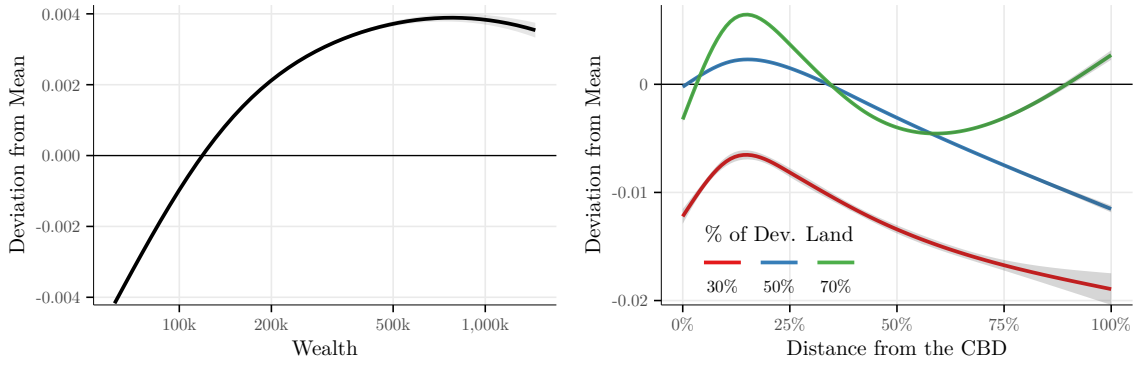


Figure 1.2: Partial Correlation Between Housing Wealth Variation, Initial Level of Wealth and Location of Main Residence in 2018

Notes: We report the relationship between the initial level of wealth and the variation in housing wealth at the individual level (left panel). We also report the relationship between housing wealth variation and the location of the main residence (right panel). We select three commuting zones based on their economic status to account for potential heterogeneity. These results come from the estimation of Equation 1.3. Our observation unit is distinct owners who do not experience a portfolio change in 2018. We report confidence intervals at the 95% confidence level. For annual results over the period 2012–2018, see Section 1.B.1. We use the `mgcv` R package to implement additive transformations of variable with endogenous definition of degree of freedom.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

We find that portfolio appreciation stemming from housing markets increases with initial level of wealth regardless the studied year (Figure 1.2, left panel).¹⁴ Thus, housing appears to behave similarly to other assets, by providing larger gains to the most expensive properties. Yet, these greater returns are not the result of higher risk. In a negative economic context (e.g. the 2012-2014 period of falling housing values), the potential capital loss is smoother for the wealthiest, while in a positive economic context (e.g. the 2017-2019 period) the capitalisation is more pronounced. As a result, housing wealth is persistent, as the wealthiest benefit from the highest growth and hold the most secure assets.

In addition, conditional on individual wealth, we find that the gross rate is heterogeneous according to the place of residence. Firstly, the wealth accumulation depends on

¹⁴We do not report annual results for clarity reasons. See in Section 1.B.1

the characteristics of the commuting zone. The deviation from the mean increases with the economic status of the economic zone. This is in line with (Eggum and Larsen, 2021) findings that individuals who have held assets in the top commuting zones (in her case, Oslo) are able to accumulate more wealth over the life cycle. Secondly, the gross rate appears to decrease with distance from the CBD. As a result, owners living in central areas with easier access to desirable amenities benefit from a higher accumulation rate than those living on the urban fringe.

The dynamics of housing markets are likely to reinforce the spatial inequalities for wealth. Indeed, while the housing wealth appreciation is increasing with individual wealth, we document spatial heterogeneity according to the residence place. Individuals living in top commuting zones and central areas are less likely to experiment capital losses in non-favourable economic context and more likely to receive higher capital gains in positive one. However, although housing markets generate heterogeneous capital gains, it mainly depends on the residence place locations, and thus housing portfolio choices.

1.3.3 HWP for homeowners that change their main residences

In line with the importance of the residence place for homeowners, we document spatial sorting for owner-occupiers according to previous housing endowments. As housing prices diverge from income, the financial capabilities are more likely to depend on wealth rather than income. Focusing on purchase achieved for main residence over the 2012-2017 period,¹⁵ we empirically assess the relationship between housing characteristics, including location, and previous housing endowment for owner-occupiers purchases. We motivate our choice to restrict to owner-occupying as it directly relates to the spatial sorting, unlike rental investment.

Our outcome of interest to discuss spatial sorting is the distance from the CBD following the traditional monocentric model. In addition, we use the transaction price to provide partial correlations between transaction price and initial level of housing wealth (in this case, the year prior to the purchase). We enable the wealth relationship to vary with commuting zone characteristics. We select the total housing wealth of buyers rather than the average value as it better reflects financial capabilities, although this specification does not affect the results. Moreover, as the distance from the CBD might also capture heterogeneity in housing characteristics, we control for housing size. Our model specification contains time fixed effects to capture for the general trend and average purchaser age. It yields to the estimation of

¹⁵Given that tenure status is provided with a delay of one year, we cannot identify main residences for the 2019 year.

$$y_{ijt} = \alpha_t + h_t(w_{it}, \ell_{jt}) + \mathbf{X}\beta + g(q_{jt}) + \varepsilon_{ijt} \quad (1.4)$$

with y_{ijt} outcome for purchasers groups i for housing j at time t ; α_t time fixed effects; w_{it} sum of initial level of wealth for purchasers group i ; ℓ_{jt} built-up area in commuting zone where housing j is located; q_{jt} housing size; \mathbf{X} structural characteristics j including housing type and construction period; while ε_{ijt} represents idiosyncratic error term. Our sample is composed of 3,156,974 observations, including housing purchased by first-time owners. Their wealth is set to zero.

To discuss HWP for homeowners who change their main residence, we focus on the relationship between housing choice (dependent variable) and initial level of wealth (covariate). Indeed, it relates to both the value of the housing stock (for the price outcome) and spatial sorting (for the distance outcome). Any positive relationships between initial level of housing wealth and housing transaction would support the HWP, while the location choices reflect spatial sorting. Given the cross-sectional heterogeneity of local housing markets, we allow the partial correlation with wealth to vary with the economic status of commuting zone. However, we cannot infer a wealth effect on housing decisions for main residence as initial level of housing wealth is endogenous and may reflect unobservable variables such as heterogeneous preferences. Empirically, we define h_t as additive bivariate transformation, with endogenous shrinkage procedures to set effective degree of freedom. We report the relevant contributions for the 2017 year for clarity purposes, but full results are detailed in Section 1.B.3 despite similar patterns over time.

The choice of the main residence of homeowners supports the HWP. The transaction price increases continuously with the initial level of wealth, regardless of the type of commuting zone considered (Figure 1.3, left panel). The difference is particularly pronounced for first-time owners, who buy the cheapest housing within the commuting zone. We expect that the lower level of deposits resulting from the lack of previous housing wealth is at least partly responsible for these differences. Moreover, the cross-sectional differences between commuting zones are substantial. For example, we estimate that owners with an initial wealth of 250k purchase housing that is 2.7 times more expensive than owners in the medium commuting zone with a similar initial wealth. This suggests that either debt or capital gains being reinvested are likely to increase with the economic status of the commuting zones, given similar financial capabilities. It reinforces that housing market is local, with no spatial sorting *between* commuting zones according to financial capabilities.

The heterogeneity observed for transaction price stems partly from the location choices (Figure 1.3, right panel). We observe that on average, distance from the CBD decreases with financial capabilities, except for the bottom commuting zones. Moreover, first-time

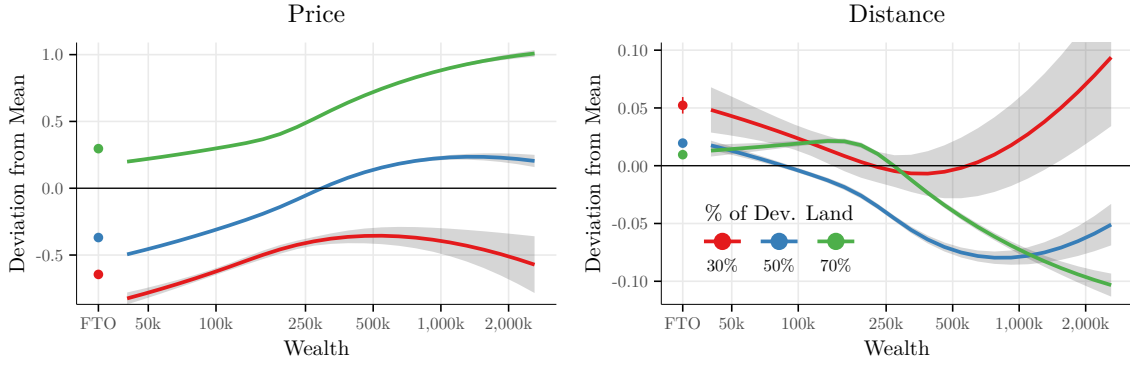


Figure 1.3: Partial Correlation of Initial Level of Wealth with Housing Decisions within Commuting Zones for Owner-Occupiers for 2017

Notes: We report the partial correlation for the interaction between the economic status of the commuting zone and the initial wealth level of the buyer from Equation 1.4. The left panel corresponds to the transaction price, while the right panel corresponds to the distance from the CBD. Our results are limited to the year 2017, but full results are available in Section 1.B.3. We introduce variation in commuting zone characteristics as we estimate the wealth correlation for commuting with 30%, 50% and 70% of the built-up area in 2000. We also introduce first-time homeowners (dot, left side of the plot). We use the *mgcv* R package to implement additive transformations of variable with endogenous definition of degree of freedom.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

owners locate themselves at higher distance from the average housing choices, although we find no significant temporal trends that indicate a reinforcement of the fringe locations for this population. Then, the difference in transaction price is at least partly due to spatial sorting according to initial wealth.

Despite a general relationship indicating that distance from the CBD decreases with initial level of wealth, there remains spatial heterogeneity according to the economic status of the commuting zone. In fact, for similar initial level of wealth, individuals in the top commuting zones are located further away than those in the middle zones, despite the more expensive purchase. The difference in location decisions is likely to reflect important heterogeneity in the fundamental value of housing between commuting zones. Finally, low commuting zones exhibit U-shaped relationship. Then, we expect the spatial distribution of amenities to diverge from those in the top commuting zones ones (Brueckner et al., 1999) according to their nature (Lee and Lin, 2018). In addition, the assumption that jobs are concentrated in the centre of the commuting zone may be less credible.

The spatial sorting appears to contribute to the HWP as the wealthiest individuals purchase the most expensive asset within commuting zones, at closer distance from the CBD. Beyond the direct effect of the utility derived from the consumption of local amenities, it increases the likelihood to benefit from higher capital gains according to the housing market dynamics. Indeed, the central areas seem to offer better returns in the long term, resulting in the most secure asset in a negative context and more favourable in a positive one. Consequently, the characteristics of main residence for housing consumption,

especially the location decision, is likely to drive future wealth accumulation.

1.3.4 HWP for homeowners that receive a housing inheritance

In addition to the relationship between homeowners' decisions for consumption and HWP, non-monetary transfers are commonly assumed to drive wealth inequality on the long-term through intergenerational persistence (De Nardi, 2004; Garbinti et al., 2021). Then, we look for potential persistence in housing wealth resulting from housing transfers. As housing transfers represent nearly one third of owners' change for housing over the 2012–2018 period,¹⁶ their ability to either contribute or mitigate HWP is significant.

The average share being received by individuals is estimated to be close to 100k (detailed results in Section 1.B.5). However, the 2012 year is singular with the highest average housing transfers (nearly 120k per recipients) over the studied period. We lay out this singularity by potential adaptation in behaviour due to policy reform introduced in August 2012. The reform raised inheritance tax which might have caused anticipation for legatees to donate prior the introduction of the reform. Agents behaved similarly for tax reforms applied to gasoline price (Coglianese et al., 2017), although economic goods differ significantly. Our results estimate that restricted to housing wealth, average recipient does not have to pay any inheritance tax to assume direct ascending line between recipients and legatees, except for the 2012 year.¹⁷

We therefore restrict our sample to housing transfers recipients over the 2012–2018 period. For each recipient, we observe the received share derived from our housing value imputation and their residence place. In addition, we observe whether recipients were previously housing owners, and if so, their initial level of housing wealth (i.e. prior to the housing transfers). Additional individual characteristics such as age or gender are introduced as control variables. Consequently, we regress the received share by individual characteristics, including previous housing wealth and individual location:

$$\log(s_{it}) = \alpha_t + h_t(d_i, \ell_i) + f_{1t}(w_{it}) + f_{2t}(a_{it}) + \varepsilon_{it} \quad (1.5)$$

with s_{it} share received by transfer recipients i at time t ; α_t time fixed effects; d_i distance from the CBD of individual i ; ℓ_i built-up area of commuting zone where individual i is living; w_{it} individual housing wealth prior to the transfers; a_{it} individuals age prior to the transfers; finally ε_{it} corresponds to idiosyncratic error term. As previously, unknown functions h_t , f_{1t} and f_{2t} are empirically specified using additive splines, with smoothing

¹⁶*Sources:* Authors' Calculation based on enhanced property tax files.

¹⁷Indeed, currently, legatees' fees concerns value above 100k.

parameters being defined endogenously using restricted maximum likelihood criteria. Our sample is composed of 1,852,126 observations about transfer recipients.

The contribution of housing transfers to the HWP derives from two facts. Firstly, it stems from the individual aspect following the wealth relationship. Indeed, if the wealthiest owners benefit from larger transfers, the housing wealth would persist over time. Secondly, it results from the spatial heterogeneity. Conditionally on recipients' wealth, any positive relationship between housing transfers and economic status of commuting zone would support the HWP spatially. Consequently, the relevant parameters are f_{1t} for the individual effect and h_t for the spatial one. As we allow both heterogeneity to vary over time, we report results for the 2018 year despite no sizeable differences over the 2011–2019 period, only for clarity reasons.¹⁸

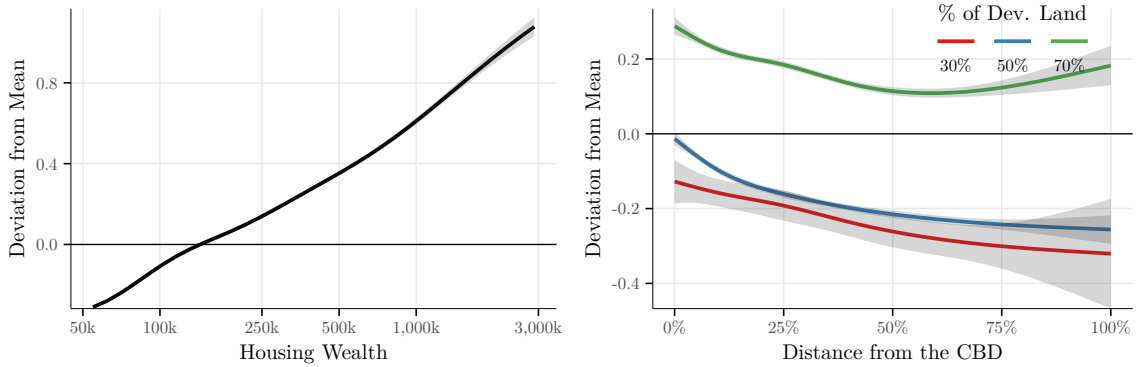


Figure 1.4: Partial Correlation of Initial Level of Wealth and Individual Locations with Transfer Share in 2017

Notes: Our results are derived from estimating Equation 1.5 restricted to housing transfer recipients. The left panel reports the partial correlation according to the initial level of wealth. We use a log transformation for the initial level of wealth. The right panel reports the partial correlation for homeowners' place of residence, using heterogeneity both between and within commuting zones. Our observation unit is recipients of housing transfers between 2012 and 2018. We report confidence intervals at the 95% level. We report all our studied period ranging from 2012 to 2017 for clarity reasons in Section 1.B.6. We use the `mgcv` R package to implement additive transformations of variable with endogenous definition of degree of freedom.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

However, the housing share being transferred increases with the wealth of recipients (see in Figure 1.4, left panel), supporting the intergenerational persistence over time (De Nardi, 2004; Garbinti et al., 2021), and thus HWP. Although we report partial correlations for 2018 year, the overall patterns remain unaffected over time (see in Section 1.B.6) despite slight differences among the wealthiest individuals.¹⁹

In addition to the observed heterogeneity according to individual characteristics, hous-

¹⁸We report detailed results per year over the 2012–2018 period in Section 1.B.6.

¹⁹Indeed, while we estimate that the share received by recipients with 1M of housing wealth was 71.6% higher than the mean in 2012, it reaches 78.6% in 2018. The increase is, however, not driven by the singularity of the 2012–year. For instance, the difference in 2014 is estimated to 71.8%.

ing transfers spatially concentrate housing wealth both within and between commuting zone (see in Figure 1.4, right panel). Firstly, regardless of the location within the commuting zone, the share received in the top commuting zones is at least 10% higher than the average share. Meanwhile, received shares are lower than the national average for beneficiaries living in the middle and low commuting zones. Secondly, the share received decreases with distance from the CBD, regardless of the economic status of the commuting zone. Recipients living in central areas benefit from larger housing transfers than those living on the urban periphery. The gradient is homogeneous and remains similar over the period 2012-2018.²⁰

Both observations support the HWP. While the consequence of increasing the share received with the initial level of housing wealth on the HWP is straightforward, we suggest a potential mechanism for spatial heterogeneity. If housing transfers are capitalised directly (e.g. through sales), the cash flow may be reinvested in the residence place either directly (e.g. through renovation) which capitalise into price, or indirectly by changing the location of the main residence. Recall that the choice of location for main residence is highly correlated with initial level of wealth (Section 1.3.3), and affects future wealth accumulation (Section 1.3.2). Finally, if housing transfers are not capitalised, they provide additional insurance for housing wealth, which may influence household behaviour with increased risk aversion.

1.3.5 HWP for homeowners with rental assets in their portfolios

Finally, we conclude the section on empirical results by focusing on homeowners with rental properties in their portfolio. These homeowners hold most of the housing stock in France and part of the location of their housing wealth is different from their place of residence. To understand how the HWP potentially results from the interaction between both locations, we consider two sub-populations. The first, defined as local investors, concerns individuals whose entire housing portfolio is located within the same commuting zone as their place of residence. Conversely, the latter refers to distant investors who have at least one rental property located in a different commuting zone from their place of residence.

The first sub-population is dominant as almost 80% of rental properties are within the same commuting zone as their owner-occupied home. Given that owner-occupation is widespread among private owners (more than 90% of homeowners are owner-occupiers), it provides a basis for HWP spatially as housing wealth depends on the dynamics of local housing markets. However, rental investment property may have mixed effects on HWP.

²⁰Full results are available in Section 1.B.6

On the one hand, if homeowners with the highest initial level of wealth hold rental assets in the most dynamic areas,²¹ rental facilities would contribute to the HWP. On the other hand, if homeowners with the lowest initial level of housing wealth benefit from a higher gross rate due to rental properties located in different commuting zones, rental properties are likely to spatially redistribute housing wealth and thus mitigate the HWP.

We therefore calculate the average annual gross rate for each sub-population at the commuting zone level and compare it to the local housing market dynamics. We then regress the average annual gross rate for wealth evolution on the gross rate of housing located within the commuting area for both sub-populations.

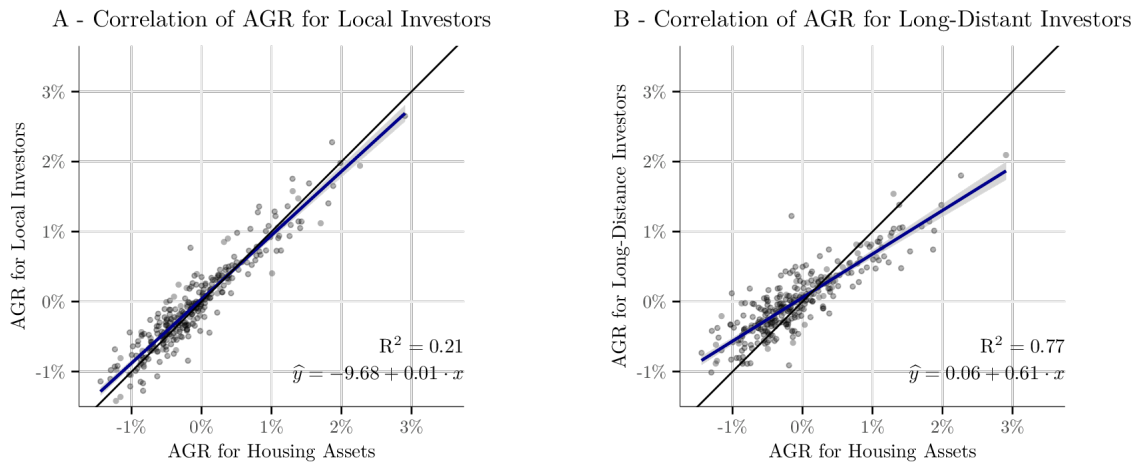


Figure 1.5: Correlation of Annual Gross Rates Between Wealth and Housing Market Dynamics for Local and Distant Investors

Notes: The left (respectively right) panel is the bivariate distribution of commuting zones according to average asset evolution within the commuting zone (x-axis) and local (respectively long-distance) investors (y-axis). We report the linear regression and the first bisector in addition to the coefficient of determination. Our population of interest is composed of homeowners with at least one rental property in their housing portfolio. AGR is the abbreviation for Average Growth Rate, expressed in percent. We exclude Alsace-Moselle homeowners as we cannot calculate their AGR due to data limitations.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

The long-distance investment weakens the relation between housing market dynamics and wealth evolution over the 2011–2019 period as expected. Nonetheless, the relationship we provide mitigate the HWP, as the coefficient of correlation is lower than 1 (Figure 1.5, right panel). In fact, long-distance investors living in commuting zone with low or negative asset evolution benefit from larger evolution. Hence, the long-distance investments smooth the wealth evolution over space, lowering expected results through an increase of the safety. In comparison, the relationship for local investors is strong, despite some deviations resulting from the within location of their rental assets (Figure 1.5, left panel). Again,

²¹Recall that, despite the correlation, the relationship between initial level of wealth and average annual gross rate is not deterministic, see in Section 1.3.1.

the internal location for rental assets mitigate the HWP spatially as we observe a similar pattern than long-distance investment, despite lower intensity.

It therefore results that rental assets are likely to mitigate the HWP spatially, both within and between commuting zones, despite the fact that the intensity is more pronounced for the first-one. Consequently, the housing market by itself is not a driver for wealth inequality, as we illustrate that it mainly results from decisions investments including location for rental assets.

1.4 Conclusion

The housing wealth is a good proxy for overall wealth (Garbinti et al., 2021) resulting from its widespread aspect in individual portfolio (Jordà et al., 2019). In addition, as it defines the main residence for homeowners, it conditions the access to economic opportunities over the life-cycle (Bilal and Rossi-Hansberg, 2021). Nevertheless, following Piketty and Zucman (2014) work, few works have specifically focused on housing assets and individuals' locations. We provide empirical insight supporting the need to consider jointly spatial sorting and wealth inequality to understand the housing wealth persistence (HWP).

Using exhaustive individual data for 2011-2019 on French homeowners and their properties, we document four insights based on different populations. All of our insights seek to understand the persistence of housing wealth, both spatially and individually. Firstly, variations that stem from the housing market mainly depends on the initial level of wealth and residence place. Secondly, the spatial sorting for owner-occupier depends on the initial level of wealth. The wealthiest individuals locate at the closest distance from the CBD. In addition, with similar financial capabilities, the distance from the CBD increases with the economic status of commuting zone despite larger housing investments. Thirdly, housing transfers contribute to the HWP, both individually and spatially. The share received increases with previous housing wealth, while residents of the top commuting zone benefit from a larger amount. Fourthly, the long-distance rental investment tends to mitigate the housing wealth persistence at a commuting zone scale. It allows homeowners with long-distance investments to have a higher annual gross rate than housing market dynamics in low commuting areas.

The implications of HWP are numerous, and relates to different topics. Firstly, the correlation between distance from the CBD and previous housing wealth for owner-occupier yields to the observation that financially constrained households are likely to be more sensitive to the evolution of the commuting cost. In light of the evolution of the economic

context (rising of gasoline price, land constraints), the attractiveness of fringe locations may decrease with potential consequences on spatial sorting and housing market dynamics. Secondly, the role of transfers, although being commonly discussed in inequality literature, is likely to determine further portfolio choices, including location of the main residence. For instance, the access to unaffordable locations could be made conditional on the number of housing transfers received. Beyond the direct impact on the tenure status, transfers are likely to affect the location decision for recipients, with sizeable implications on the long-term.

Considering the contribution we make in terms of individual data to study housing wealth and spatial sorting, we consider that numerous contributions should follow and suggest potential leads. Firstly, despite the evidence on the persistence of housing wealth, we provide indications that rental assets are likely to have redistributive effects. Given that rental assets provide additional income flows to owners, it would be interesting to further investigate the potential redistributive effects of rental assets through the spatial dimension. Secondly, we suggest drawing causal relationships between housing transfers and further portfolio choices. Indeed, while most of the existing literature focuses on the impact on tenure decisions, we expect housing transfers to affect housing decisions within beneficiaries. This is an outstanding challenge due to the endogeneity of the share received. Thirdly, although beyond the scope of this paper, we suggest further discussion of the relationship between wealth accumulation and land restrictions. While the existing literature focuses on the price capitalisation that stems from restrictions, we suggest to detail the consequences on capital gains and housing portfolio decisions.

References

- Ahlfeldt, Gabriel M., Stephan Heblich, and Tobias Seidel (2023), “Micro-Geographic Property Price and Rent Indices”. *Regional Science and Urban Economics* 98.103836.
- Albouy, David, Gabriel Ehrlich, and Yingyi Liu (2016), “Housing Demand, Cost-of-Living Inequality, and the Affordability Crisis”. *NBER Working Paper* 22816, p. 62.
- Alvaredo, Facundo, Anthony B. Atkinson, and Salvatore Morelli (2018), “Top Wealth Shares in the UK over More than a Century”. *Journal of Public Economics* 162, pp. 26–47.
- André, Mathias and Olivier Meslin (2021), “Et Pour Quelques Appartements de plus : Étude de La Propriété Immobilière Des Ménages et Du Profil Redistributif de La Taxe Foncière”. *INSEE Working Paper*, p. 93.
- Bach, Laurent, Laurent E. Calvet, and Paolo Sodini (2020), “Rich Pickings? Risk, Return, and Skill in Household Wealth”. *American Economic Review* 110.9, pp. 2703–2747.
- Banzhaf, H. Spencer and Omar Farooque (2013), “Interjurisdictional Housing Prices and Spatial Amenities: Which Measures of Housing Prices Reflect Local Public Goods?” *Regional Science and Urban Economics* 43.4, pp. 635–648.
- Benhabib, Jess, Alberto Bisin, and Mi Luo (2017), “Earnings Inequality and Other Determinants of Wealth Inequality”. *American Economic Review* 107.5, pp. 593–597.
- Bilal, Adrien and Esteban Rossi-Hansberg (2021), “Location as an Asset”. *Econometrica* 89.5, pp. 2459–2495.
- Black, Sandra (1999), “Do Better Schools Matter? Parental Valuation of Elementary Education”. *Quarterly Journal of Economics* 114.2, pp. 577–599.
- Blanco, Miguel Artola, Luis Bauluz, and Clara Martínez-Toledano (2021), “Wealth in Spain 1900–2017: A Country of Two Lands”. *The Economic Journal* 131.633, pp. 129–155.
- Boserup, Simon H., Wojciech Kopczuk, and Claus T. Kreiner (2016), “The Role of Bequests in Shaping Wealth Inequality: Evidence from Danish Wealth Records”. *American Economic Review* 106.5, pp. 656–661.
- Brueckner, Jan K., Jacques-François Thisse, and Yves Zenou (1999), “Why Is Central Paris Rich and Downtown Detroit Poor?” *European Economic Review* 43.1, pp. 91–107.
- Brunsdon, Chris, A. Stewart Fotheringham, and Martin E. Charlton (1996), “Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity”. *Geographical Analysis* 28.4, pp. 281–298.
- Chay, Kenneth Y. and Michael Greenstone (2005), “Does Air Quality Matter? Evidence from the Housing Market”. *Journal of Political Economy* 113.2, pp. 376–424.

- Coglianese, John, Lucas W. Davis, Lutz Kilian, and James H. Stock (2017), “Anticipation, Tax Avoidance, and the Price Elasticity of Gasoline Demand”. *Journal of Applied Econometrics* 32.1, pp. 1–15.
- De Nardi, Mariacristina (2004), “Wealth Inequality and Intergenerational Links”. *The Review of Economic Studies* 71.3, pp. 743–768.
- Diao, Mi, Delon Leonard, and Tien Foo Sing (2017), “Spatial-Difference-in-Differences Models for Impact of New Mass Rapid Transit Line on Private Housing Values”. *Regional Science and Urban Economics* 67, pp. 64–77.
- Duranton, Gilles and Diego Puga (2015), “Urban Land Use”. *Handbook of Regional and Urban Economics*. Vol. 5A. Boston, MA: Elsevier, p. 698.
- Eggum, Terje and Erling Røed Larsen (2021), “Is the Housing Market an Inequality Generator?” *Housing Lab Working Paper* 2, p. 43.
- Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri (2020), “Heterogeneity and Persistence in Returns to Wealth”. *Econometrica* 88.1, pp. 115–170.
- Fotheringham, A. Stewart, Chris Brunsdon, and Martin Charlton (2002), *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester, England ; Hoboken, NJ, USA: Wiley.
- Gabriel, Stuart and Gary Painter (2020), “Why Affordability Matters”. *Regional Science and Urban Economics* 80.103378.
- Garbinti, Bertrand, Jonathan Goupille-Lebret, and Thomas Piketty (2021), “Accounting for Wealth-Inequality Dynamics: Methods, Estimates, and Simulations for France”. *Journal of the European Economic Association* 19.1, pp. 620–663.
- Glaeser, Edward L, Joseph Gyourko, and Raven E Saks (2005), “Why Have Housing Prices Gone Up?” *American Economic Review* 95.2, pp. 329–333.
- Grossman, Sanford T and Guy Laroque (1987), “Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumer Goods”. *Econometrica : journal of the Econometric Society* 58, pp. 25–21.
- Guerrieri, Veronica, Daniel Hartley, and Erik Hurst (2013), “Endogenous Gentrification and Housing Price Dynamics”. *Journal of Public Economics* 100, pp. 45–60.
- Hilber, Christian A.L. and Christopher Mayer (2009), “Why Do Households without Children Support Local Public Schools? Linking House Price Capitalization to School Spending”. *Journal of Urban Economics* 65.1, pp. 74–90.
- Hilber, Christian A.L. and Andreas Mense (2021), “Why Have House Prices Risen so Much More than Rents in Superstar Cities?” *Geography and Environment Discussion Paper Series* 30, p. 51.
- Huggett, Mark, Gustavo Ventura, and Amir Yaron (2011), “Sources of Lifetime Inequality”. *American Economic Review* 101.7, pp. 2923–2954.

- Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M Taylor (2019), “The Rate of Return on Everything, 1870–2015”. *The Quarterly Journal of Economics* 134.3, pp. 1225–1298.
- Kiel, Katherine A. and Jeffrey E. Zabel (2008), “Location, Location, Location: The 3L Approach to House Price Determination”. *Journal of Housing Economics* 17.2, pp. 175–190.
- Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins (2013), “The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets”. *Journal of Economic Literature* 51.4, pp. 1007–1062.
- Lee, Sanghoon and Jeffrey Lin (2018), “Natural Amenities, Neighbourhood Dynamics, and Persistence in the Spatial Distribution of Income”. *The Review of Economic Studies* 85.1, pp. 663–694.
- Lee, Seong Woo and Dowell Myers (2003), “Local Housing-Market Effects on Tenure Choice”. *Journal of Housing and the Built Environment* 18.2, pp. 129–157.
- Li, Ziqi and A. Stewart Fotheringham (2020), “Computational Improvements to Multi-Scale Geographically Weighted Regression”. *International Journal of Geographical Information Science* 34.7, pp. 1378–1397.
- Li, Ziqi, A. Stewart Fotheringham, Wenwen Li, and Taylor Oshan (2019), “Fast Geographically Weighted Regression (FastGWR): A Scalable Algorithm to Investigate Spatial Process Heterogeneity in Millions of Observations”. *International Journal of Geographical Information Science* 33.1, pp. 155–175.
- McMillen, Daniel P. (2010), “Issues in Spatial Data Analysis”. *Journal of Regional Science* 50.1, pp. 119–141.
- Murakami, Daisuke, Narumasa Tsutsumida, Takahiro Yoshida, Tomoki Nakaya, and Binbin Lu (2020), “Scalable GWR: A Linear-Time Algorithm for Large-Scale Geographically Weighted Regression with Polynomial Kernels”. *Annals of the American Association of Geographers*, pp. 1–22.
- Ortalo-Magné, François and Andrea Prat (2016), “Spatial Asset Pricing: A First Step”. *Economica* 83.329, pp. 130–171.
- Pfeffer, Fabian T and Nora Waitkus (2021), “The Wealth Inequality of Nations”. *American Sociological Review* 86.4, pp. 567–602.
- Piketty, Thomas and Gabriel Zucman (2014), “Capital Is Back: Wealth-income Ratios in Rich Countries 1700–2010”. *Quarterly Journal of Economics* 129.3, pp. 1255–1310.
- Roback, Jennifer (1982), “Wages, Rents, and the Quality of Life”. *Journal of political Economy* 90.6, pp. 1257–1278.

- Saez, Emmanuel and Gabriel Zucman (2016), “Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data”. *Quarterly Journal of Economics* 131.2, pp. 519–578.
- Saiz, Albert (2010), “The Geographic Determinants of Housing Supply”. *Quarterly Journal of Economics* 125.3, pp. 1253–1296.
- Sinai, Todd and Nicholas S. Souleles (2005), “Owner-Occupied Housing as a Hedge Against Rent Risk”. *Quarterly Journal of Economics* 120.2, pp. 763–789.
- Tur-Sinai, Aviad, Larisa Fleishman, and Dmitri Romanov (2020), “The Accuracy of Self-Reported Dwelling Valuation”. *Journal of Housing Economics* 48.101660.
- Van Ommeren, Jos and Michiel Van Leuvensteijn (2005), “New Evidence of the Effect of Transaction Costs on Residential Mobility”. *Journal of Regional Science* 45.4, pp. 681–702.
- Wood, Simon N. (2017), *Generalized Additive Models: An Introduction with R*. Second. Chapman and Hall/CRC.
- Zabel, Jeffrey E. (2012), “Migration, Housing Market, and Labor Market Responses to Employment Shocks”. *Journal of Urban Economics* 72.2-3, pp. 267–284.

Appendix

1.A Data

1.A.1 Descriptive Statistics about Housing Stock Derived From Fiscal Sources

Table 1.A.1: Descriptive Statistics for Housing Stock

Variable	N	Mean	Median	Std Dev	Quantile 1	Quantile 3	Min	Max
House								
Surface	19,960,874	102.3	95.0	45.0	76.0	122.0	11.0	4,122.0
Dep. Surface	19,960,874	14.7	0.0	35.7	0.0	18.0	0.0	3,503.0
Building Year	19,960,874	1941	1971	77	1900	1992	1300	2019
Flat								
Surface	18,206,318	57.8	57.0	28.6	38.0	73.0	11.0	6,646.0
Dep. Surface	18,206,318	7.5	3.0	17.1	0.0	8.0	0.0	6,610.0
Building Year	18,206,318	1923	1970	152	1914	1993	1300	2019

Notes: The top panel reports main descriptive statistics for individual housings (named as house) for housing characteristics (surface, dependence surface, building year). Bottom panel provide similar features for collective housing (named as flat).

Sources: *Fichiers Fonciers*.

1.A.2 Descriptive Statistics for Housing Transactions Derived From Fiscal Sources

Table 1.A.2: Descriptive Statistics for the DV3F database

Variable	N	Mean	Median	Std Dev	Quantile 1	Quantile 3	Min	Max
House								
Surface	4,042,638	100.9	94.0	41.5	75.0	120.0	11.0	1,440.0
Dep. Surface	4,042,638	61.2	47.0	57.3	21.0	85.0	0.0	8,749.0
Building Year	4,042,638	1938	1966	74	1900	1988	1300	2019
Price	4,042,638	200,699	167,000	159,252	109,000	249,000	10,001	15,750,000
Price per m ²	4,042,638	2,009.1	1,769.3	1,193.3	1,216.2	2,523.6	17.3	19,949.7
Flat								
Surface	3,407,612	57.9	56.0	27.1	39.0	72.0	11.0	1,500.0
Dep. Surface	3,407,612	7.6	4.0	14.3	0.0	9.0	0.0	3,025.0
Building Year	3,407,612	1927	1973	158	1930	2002	1300	2019
Price	3,407,612	195,529	152,000	184,297	98,500	228,150	10,001	17,865,830
Price per m ²	3,407,612	3,493.2	2,954.5	2,259.0	1,985.3	4,210.5	39.5	19,837.8

Notes: The top panel reports main descriptive statistics for individual housings (named as house), for housing characteristics (surface, dependence surface, building year), and transaction feature including the purchase price and unitary price. Bottom panel provides similar features for collective housing (named as flat).

Sources: DV3F.

1.A.3 Share of Built-up Area in 2000 Within Each Commuting Area

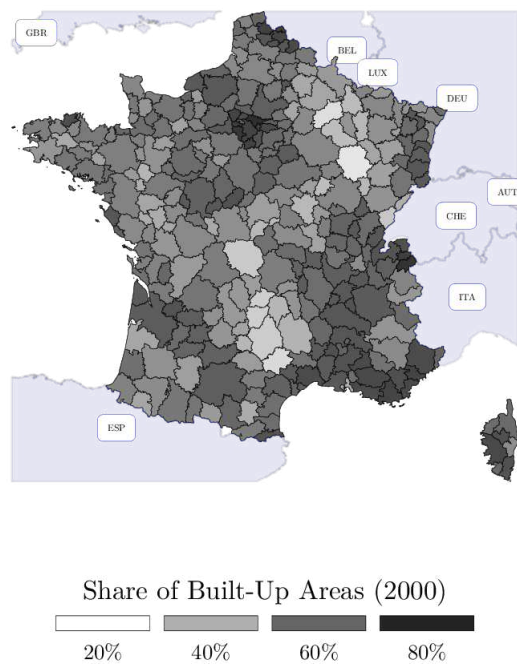


Figure 1.A.1: Share of Developed Land According to Commuting Area

Notes: We report share of developed land in 2000 for each commuting area defined by the INSEE in 2010. We consider as developed parcels, parcel with at least one built premises prior to 2000 using the *Fichiers Fonciers*. We remove from the parcel considered as natural areas.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*. Commuting areas are derived from the INSEE.

1.A.4 ID Construction

We create a national ID to overcome this issue by using the civil state available in the *Fichiers Fonciers*. Our new ID relies on the available civil stat. However, as data include some errors (not filled data, inconsistent date of birth, typographic errors, etc.), we aim to select the best approach to minimize errors (both type I and type II errors, see definition in the following).

Potential IDs First, we review the potential IDs we consider. Our potential IDs aim to reduce the likelihood of typographic errors stemming from misspecification in the fiscal data. Hence, we test 3, 5 and all characters derived from the first part to characterise names. In addition, we introduce variation according to the gender. Finally, we remove particle for names to avoid potential misspecification.

Table 1.A.3: Potential National ID Based on Civil State and Name Truncating

Names	Gender	Place of Birth	Results
Part 1	Yes	No	M-PIERREDUPONT-DDMMYYYY
Part 1	Yes	Yes	M-PIERREDUPONT-DDMMYYYY-75PAR
3 char.	Yes	No	M-PIEDUP-DDMMYYYY
3 char.	Yes	Yes	M-PIEDUP-DDMMYYYY-75PAR
5 char.	Yes	No	M-PIERRDUPON-DDMMYYYY
5 char.	Yes	Yes	M-PIERRDUPON-DDMMYYYY-75PAR
Part 1	No	No	PIERREDUPONT-DDMMYYYY
Part 1	No	Yes	PIERREDUPONT-DDMMYYYY-75PAR
3 char.	No	No	PIEDUP-DDMMYYYY
3 char.	No	Yes	PIEDUP-DDMMYYYY-75PAR
5 char.	No	No	PIERRDUPON-DDMMYYYY
5 char.	No	Yes	PIERRDUPON-DDMMYYYY-75PAR

Notes: Example for a man who its last name is Dupont and its first name Pierre born in Paris (75).

Error Assessment Type I error corresponds to distinct individuals registered under the same national ID whereas type II error corresponds to same individual registered under different national ID. Empirically, we adopt the following definitions for type I and type II errors. A **type I error** is identified when the national ID is equal but there are differences in either last name, first name or date of birth (when they are available). Conversely, a **type II error** is defined by different national ID with similar civil state. We define similarity by enabling slight differences in the character strings that composed civil state. Specifically, the Levenshtein distance²² must be lower or equal to one for both

²²Levenshtein distance between two strings of characters corresponds to the number of characters to either remove, add or replace to obtain equivalent character strings.

first name and last name. We also add the condition of the date of birth equality when the date of birth is available.

Results We provide results from type I and type II errors according to the definition we adopt. For each ID, we estimate the overall error resulting from both type I and type II.

Table 1.A.4: Overall Results for Potential Errors Assessments according to Potential ID

Names	Gender	Place of Birth	All				Without Incomplete Data	
			False	Omitted	Errors	Error Rate	Errors	Error Rate
Part 1	Yes	No	241,932	92,076	334,008	0.80%	92,126	0.22%
Part 1	Yes	Yes	203,298	926,014	1,129,312	2.71%	926,024	2.23%
3 char.	Yes	No	2,901,967	53,472	2,955,439	7.10%	2,711,382	6.52%
3 char.	Yes	Yes	266,576	920,024	1,186,600	2.85%	982,991	2.36%
5 char.	Yes	No	379,899	98,181	478,080	1.15%	291,158	0.70%
5 char.	Yes	Yes	177,742	928,664	1,106,406	2.66%	952,451	2.29%
Part 1	No	No	249,514	81,533	331,047	0.80%	81,587	0.20%
Part 1	No	Yes	209,907	922,285	1,132,192	2.72%	922,295	2.22%
3 char.	No	No	3,640,464	32,321	3,672,785	8.83%	3,420,467	8.22%
3 char.	No	Yes	285,328	914,532	1,199,860	2.88%	989,466	2.38%
5 char.	No	No	449,585	78,601	528,186	1.27%	334,623	0.80%
5 char.	No	Yes	189,188	923,657	1,112,845	2.67%	953,311	2.29%

Notes: The ID composed by the first part of names, the gender without the place of birth leads to 241,932 departmental accounts affected by false matching and 92,076 by omitted matchings. The overall error rate is then 0.80%. However, the error rate is only 0.22% for well-coded data.

The ID which minimizes the number of errors is the ID with the first character string without both place of birth and gender. The number of departmental accounts potentially affected by errors is 0.80% (if we consider all accounts). If we remove omitted matchings for observation with incomplete data, the error rate falls to 0.20%. This can be justified as we have not enough selective criteria to say whether or not they are true errors.

Consequently, we adopt the current ID for observation with misspecified civil state. It leads to an increase of type II errors. We rather prefer type II errors than type I to avoid the over-estimation of housing wealth, and consequently wealth inequalities. To close our choice, we select the ID composed of the first part of first and last names, in addition to the date of birth. For civil state with misspecification, we adopt the current ID. Finally, we anonymise the ID using integer through random assignment. As variables selected in the civil state are invariant, we can apply the national ID from 2011 to 2020.

1.A.5 Comparison With Existing Database

The administrative data about housing stock we use are similar than those exploited by André and Meslin (2021) to analyse the distribution of housing properties in 2017 within private owners. We both select ownership rights that correspond to wealth definition (rather than income flow), i.e. full and bare ownership. They identify unique household based on address (as they have a cross-sectional approach) and assign detained housing properties.

However, as we extend the study period from 2011 to 2019 (while they solely focus on 2017), place of residence is no longer a determinant in identifying stable households due to moving.. Hence, we adopt unique private owners for stability reasons over time (households might experience separation for instance), being consistent with previous empirical works exploiting panel data (Bach et al., 2020; Fagereng et al., 2020). In addition, we do not impute housing properties detained through rental investment companies due to data limitations. On the contrary, André and Meslin (2021) have partly assigned housing properties detained via rental investment companies to individual owners. However, they have imposed stricter assumptions to deal with missing variables about shareholders. Consequently, our population of interest differs in regards with the housing properties under consideration.

Despite differences in methodological choices to identify unique owners that are the consequences of the differences in pursued objectives, our results with respect to multi-ownership and gender gaps are close to theirs for 2017. We estimate that nearly 60% of housing stock detained by private persons belongs to multiowners; it reaches two third according to André and Meslin (2021). Moreover, the gender gap is increasing with the number of properties detained by private persons. We report detailed results on Section 1.A.5. Hence, although our methodological choices differ from theirs to extend studied period and provide panel data about owners and housing properties from 2011 to 2019, our results are close. It thus reinforces the credibility of both approaches.

Table 1.A.5 reports the main methodological differences between both dataset. Although we focus on methodological differences, we also adopt similar choices, especially for property rights and housing definition.

Using the database provided by André and Meslin (2021) as reference, we compare the results obtained with our database to their main ones for 2017. Our estimation leads to differences in magnitudes of measured effects, but our results about multiownership and gender gaps are consistent with those provided by André and Meslin (2021).

Table 1.A.5: Summary of Methodological Choices

Topic	André and Meslin (2021)	Current Paper
Observation Unit	Household	Individual
Period	2017	2011–2019
Rental Investment Companies	Assign to natural owners	Not assigned
Owners Location	French Resident	All
Spatial Perimeter	Metropolitan France	Metropolitan France
Housing Properties	Ind. and Coll.	Ind. and Coll.
Property Rights	Full owners and Bare-owners	Full owners and Bare-owners

Notes: We report main differences we identify between our methodological choices to overcome current ID issue of administrative data sources.

1.A.6 Comparison of Methods for Housing Value Imputation

Procedure We choose to adopt the Leave-One Out Cross Validation (LOOCV) procedure to compare methods. Our procedure is as follows: We sample 10% of the transaction datasets and we remove it from the transaction data. Then, we run the calibration step on the remaining transaction and we predict the value of the transaction based on the model. We itemize this process in order to have multiple predicted value. As this procedure is computationally intensive, we work on 50 iterations.

Perimeter We restrict our perimeter of interest to the *Côte d’Or* department. It is composed of metropolitan areas such as Dijon, middle-size cities such as Beaune, and rural areas. Hence, the department provides sufficiently diversified local contexts to assess the accuracy of housing value imputation to local characteristics. For the temporal dimension, we take advantage of the entire period for the dataset, namely 2010 to 2020. We select 39,799 transactions.

Model Specification We perform three distinct imputations through three different specifications. Firstly, we perform the estimation through OLS, with different spatial units for fixed effects. We also test as benchmark the OLS estimation without spatial fixed effects to account for spatial heterogeneities. The GWR is performed through the `scgwr` package (Murakami et al., 2020) which adapt seminal GWR for large datasets. Finally, we perform the estimation through GAM using the `mgcv` package (Wood, 2017) that enables to introduce spline transformations for continuous variable. The effective degree of freedom for the spline transformation is endogenously shrinked using iterated procedure. We model the spatial heterogeneity using bi-variate smoothing function using spatial coordinates of housing transactions.

Selected Variables To perform this comparison, we select housing characteristics to control for their contribution to the price. The following variables are introduced either in linear form or spline transformations in regards with the estimation procedure: housing surface; outbuilding surface (including cellar, garage, etc.); building age; housing type; balcony presence; time fixed effects using quarter aggregation.

Results We report results according to the average relative error resulting from the 50 iterations procedure we perform. It represents 198,995 individual estimations. Firstly, accounting for spatial heterogeneities appears crucial to obtain accurate estimation of housing value. Indeed, the benchmark (OLS with no spatial fixed effects) have poor accuracy. In addition, the ScaGWR and GAM procedures provide more accurate estimation

than OLS with spatial fixed effects. Both procedure are close, either for mean or median relative errors. The dispersion is however more important for the ScaGWR.

Table 1.A.6: Distribution of Relative Errors According to the Estimation Procedure

	Method	Mean	Median	Std Dev	Quantile 1	Quantile 3	Decile 1	Decile 9
	Benchmark	25.44	20.46	20.59	9.54	35.95	3.76	54.71
	OLS FE Municipality	19.29	15.51	16.17	7.50	26.17	3.02	39.72
	OLS FE IRIS	18.05	14.18	15.76	6.66	24.64	2.62	37.69
	ScaGWR	16.59	12.32	15.63	5.64	22.33	2.20	35.81
	GAM	16.57	12.76	14.83	6.08	22.41	2.42	35.03

Notes: We report main descriptive statistics resulting from our comparison procedure. Our outcome of interest is the relative error for a prediction. We use similar dataset for each estimation procedure under consideration. The benchmark corresponds to an OLS estimation without spatial fixed effects. The ScaGWR is estimated using the `scgwr` R package, while we use the `mgcv` R package to implement the GAM procedure. Our results stem from 198,995 individuals estimation from 50 iterations. Our perimeter of interest is the *Côte d'Or* department, with transactions from 2010 to 2020.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

Spatial Heterogeneity We also assess the accuracy heterogeneity of each method according to the local characteristics. We exploit the urban areas providing by the INSEE. We qualify municipality within each urban areas according to the distance from the centre. In addition, rural areas are clustered.

Table 1.A.7: Distribution of Relative Errors According to the Estimation Procedure and Housing Locations

	Method	Center	1st Ring	2nd Ring	3rd Ring	Rural
	Benchmark	26.64	23.16	24.31	27.33	30.40
	OLS FE Municipality	19.08	17.67	19.93	21.99	22.02
	OLS FE IRIS	17.39	15.86	19.84	21.88	22.02
	ScaGWR	14.88	13.76	20.00	22.67	24.47
	GAM	15.45	14.87	18.66	20.56	21.31

Notes: We report the average relative error according to the municipality characteristics. Our outcome of interest is the relative error for a prediction. We use similar dataset for each estimation procedure under consideration. The benchmark corresponds to an OLS estimation without spatial fixed effects. The ScaGWR is estimated using the `scgwr` R package, while we use the `mgcv` R package to implement the GAM procedure. Our results stem from 198,995 individuals estimation from 50 iterations. Our perimeter of interest is the *Côte d'Or* department, with transactions from 2010 to 2020.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

Although the GAM and the GWR are close when comparing the entire *Côte d'Or* department, they both recover spatial heterogeneities in their accuracy. Indeed, the accuracy decreases with distance from the CBD. Nonetheless, the GAM approach provides more consistent results according to transaction location, in line with the lower dispersion observed on general results. The GWR performs poorly for rural areas and third ring of urban areas, as OLS with spatial fixed effects provide better results. Finally, compu-

tational times are sizeable for the GWR, while the GAM procedure is ten times faster. Hence, we exploit the GAM procedure to implement the imputation model.

1.A.7 Relationship for 1D Variables with Unitary Housing Price

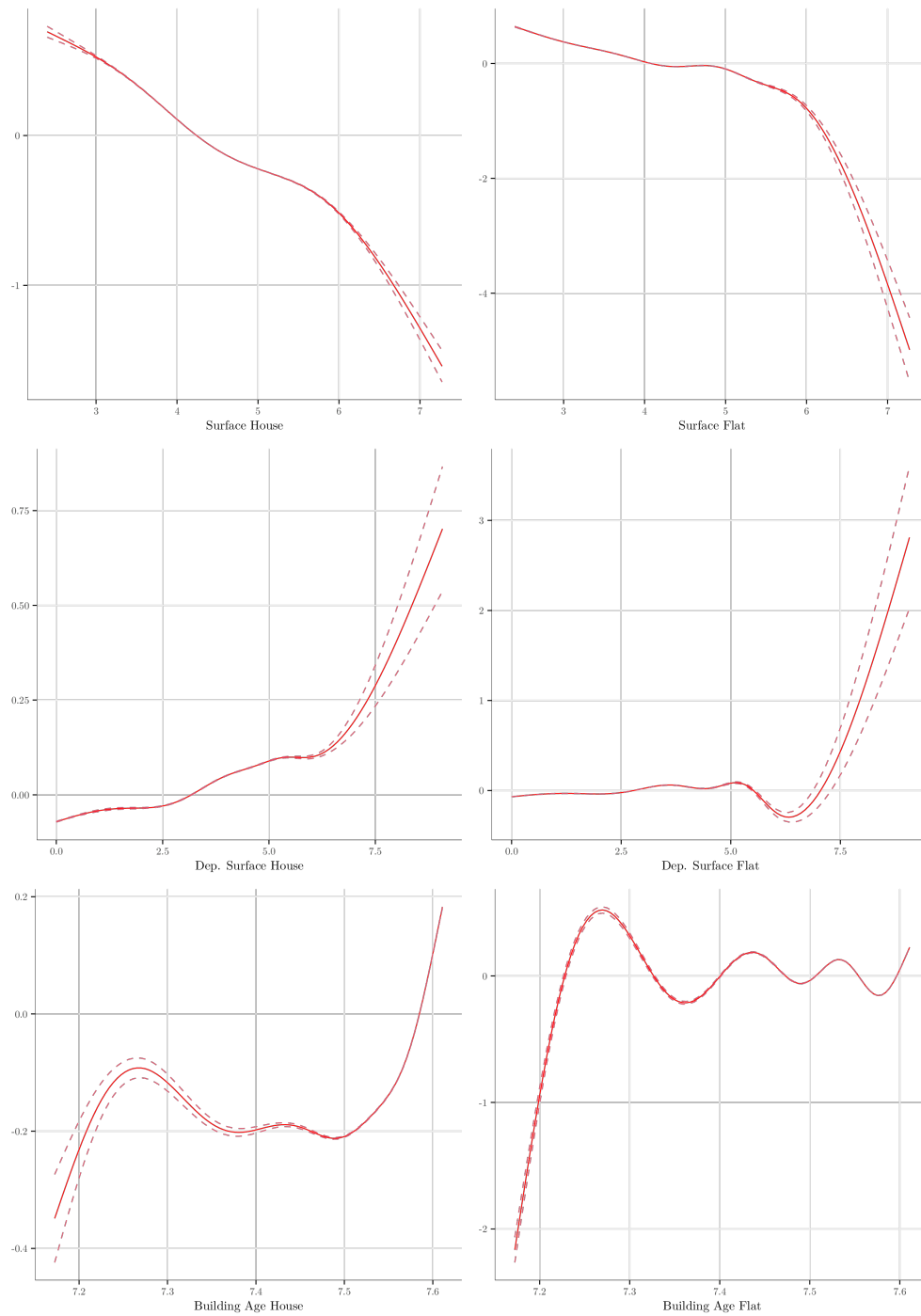


Figure 1.A.2: Relationship for 1D Variables with Unitary Housing Price

Notes: We report the marginal contribution for continuous variables introduced in the housing value imputation model. It stems from the estimation of pooled national model with spatial coordinates using bivariate spline transformation (non reported). The dependent variable is the unitary housing price. We distinguish marginal contribution according to whether it concerns collective or individual housing. The estimation is performed using the `mgcv` R package.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

1.A.8 Joint Significance for Country Models to Explain Price and Unitary Housing Price

Table 1.A.8: Joint Significance of Covariates in Pooled Models for Auto Valuation Method

Variable	Unit. Price			Price		
	(1)	(2)	(3)	(4)	(5)	(6)
Spatial Coordinates		12,149*** [786.8]	8,716*** [733.3]		2,825*** [768.3]	8,663*** [744.0]
Ind. Housing						
Surface	28,446*** [7.0]		50,344*** [6.9]	145,757*** [7.0]		258,894*** [7.0]
Age	53,358*** [7.0]		54,191*** [7.0]	53,363*** [7.0]		54,998*** [7.0]
Dep. Surface	1,069*** [6.9]		7,858*** [6.9]	1,069*** [6.9]		7,668*** [6.9]
Coll. Housing						
Surface	51,909*** [7.0]		45,529*** [7.0]	157,603*** [7.0]		386,225*** [7.0]
Age	40,300*** [7.0]		53,235*** [7.0]	40,289*** [7.0]		53,282*** [7.0]
Dep. Surface	10,510*** [7.0]		2,907*** [6.9]	10,510*** [7.0]		2,951*** [6.9]
N	7,176,945	7,176,945	7,176,945	7,176,945	7,176,945	7,176,945
McFadden R ²	31.44	57.47	65.19	30.90	23.91	64.80
AIC	10,349,171	6,924,063	5,486,871	10,349,089	11,041,679	5,508,224

Notes: We report joint significance estimated using χ^2 for pooled appraisal models for unitary housing price (columns 1 to 3) and overall price (columns 4 to 6). Our pooled model is estimated using endogenous spline transformations for covariates, with GAM. The effective degree of freedom for variable transformations is reported in brackets. Our spline transformation differs between individual and collective housings. The bottom panel reports metrics about regression, including the number of transactions (N), the McFadden R² and Aikike Criterion Information.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

*** p < 0.01, ** p < 0.05 * p < 0.1

1.A.9 Joint Significance for Commuting Areas Models to Explain Unitary Housing Price

Table 1.A.9: Results from Joint Significance of Smoothing Splines Transformation

Covariates	Mean	Q1	Q3	Min	Max
Ind. Housing					
Surface	577.8 [4.7] 292/292	215.0 [4.2]	679.8 [5.3]	13.6 [1.3]	7,163.1 [6.8]
Dep. Surface	164.1 [4.3] 292/292	43.2 [3.7]	197.1 [5.0]	8.2 [1.4]	1,644.4 [6.7]
Building Age	332.9 [6.3] 292/292	121.5 [6.2]	394.9 [6.7]	8.0 [2.1]	3,256.9 [6.9]
Coll. Housing					
Surface	301.3 [3.5] 287/292	26.5 [2.1]	219.8 [4.8]	1.7 [1.0]	5,054.2 [6.9]
Dep. Surface	54.8 [3.0] 227/292	3.4 [1.7]	46.0 [4.1]	0.0 [1.0]	955.0 [6.2]
Building Age	582.7 [5.2] 264/292	11.5 [4.1]	322.0 [6.7]	0.0 [1.0]	15,422.0 [7.0]
Spatial Coordinates	33.6 [199.2] 292/292	1.2 [118.8]	18.4 [268.7]	0.1 [17.8]	2,162.8 [731.3]
R ²	0.42	0.32	0.52	0.18	0.79
N	32,938	7,923	31,298	1,007	910,207
AIC	-51	1,020	4,523	-158,996	21,596

Notes: We report the average joint significance, the average effective degree of freedom and the number of joint significance for our independent mass appraisal models. Unit of observations is the commuting zone. We also report in the bottom panel average statistics about regression (average). Finally, we provide in columns the distribution of statistics for each covariate introduced in imputation model.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

1.A.10 Descriptive Statistics about Imputed Housing Value

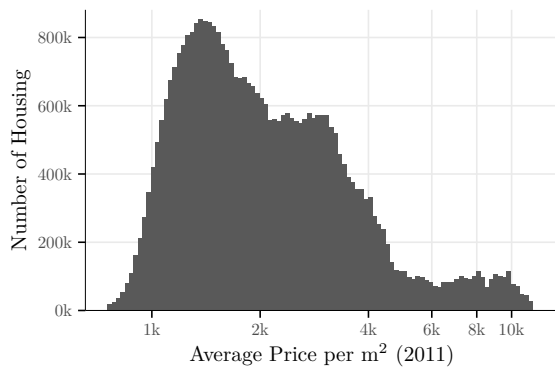
Table 1.A.10: Descriptive Statistics for Housing Value Imputation

Year	N	Mean	Median	Std Dev	Quantile 1	Quantile 3	Min	Max
2011	33,105,876	166,583	138,480	112,829	94,784	203,126	751	10,567,560
2012	33,458,690	168,337	138,750	116,966	94,612	204,768	1,705	10,388,280
2013	33,810,212	167,049	137,390	116,928	93,568	203,040	1,705	10,213,980
2014	34,130,178	164,550	135,548	115,081	92,308	200,016	1,716	42,460,440
2015	34,423,797	163,714	134,910	114,549	91,770	199,064	1,705	37,254,888
2016	34,784,340	165,309	135,892	116,801	92,106	201,110	1,716	34,046,808
2017	35,069,421	168,671	138,208	120,309	93,360	205,248	1,738	31,275,807
2018	35,355,911	172,675	141,120	123,925	94,966	210,315	1,635	30,342,829
2019	35,650,295	177,126	144,432	127,717	96,768	216,087	715	31,036,918

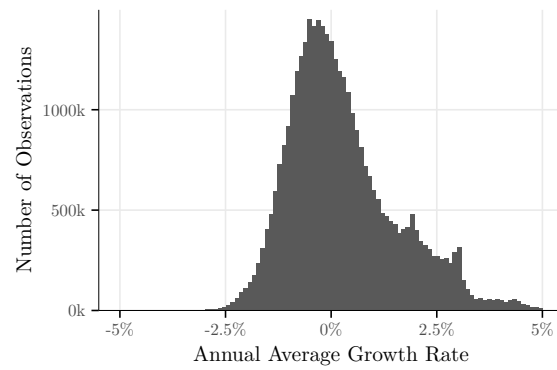
Notes: We report the distribution of imputed housing value through our procedure. It exploits the `mgcv` R package. Imputed values concern the metropolitan France, except from the Alsace-Moselle.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

1.A.11 Distribution of Imputed Housing Value (Cross-sectional and Evolution)



(a) Housing Prices per m² in 2019



(b) Annual Average Growth Rate

Notes: We report the distribution of appraised unitary housing price in 2019 (left panel) and average annual growth rate between 2011 and 2019 (right panel). Our observation unit arises from 200×200 square grid provided by INSEE to provide demographic files.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

1.A.12 Imputed Housing Value According to Locations (Cross-sectional and Evolution)

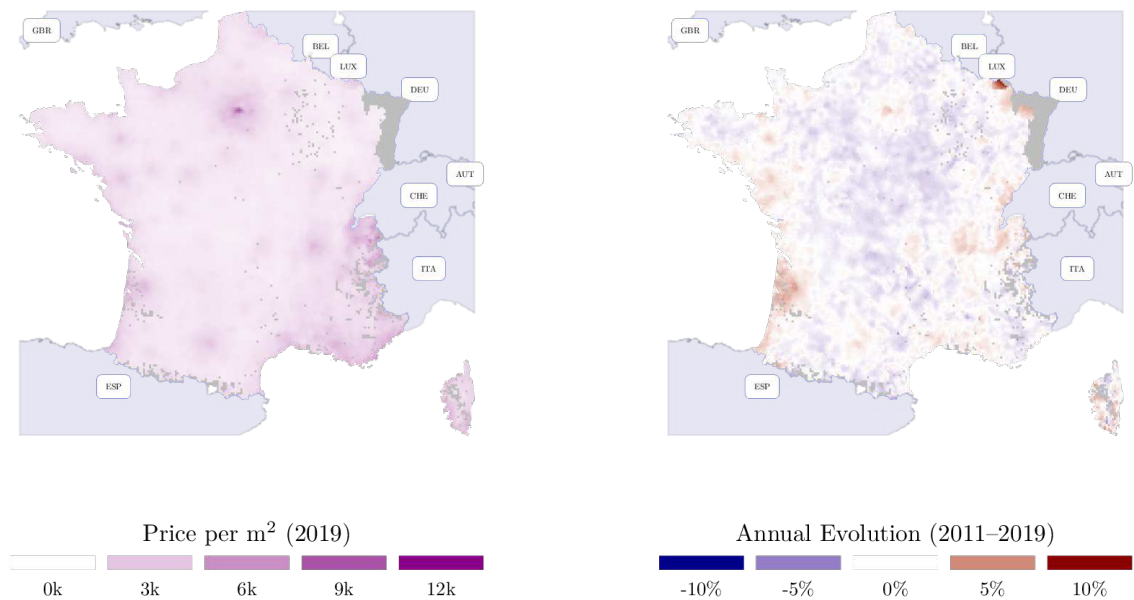


Figure 1.A.4: Descriptive Statistics about Appraised Values According to Location

Notes: We report on the left panel average unitary housing price using 200×200 meters square grid. The right panel represents the average annual evolution for unitary housing price between 2011 and 2019.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

1.A.13 Transfers

To identify transfers, we exploit the exhaustive character of the DV3F registering all transactions for pecuniary purposes, combined with the panel data about housing we construct previously. We indeed identify all housing being concerned by change of owners every year. As both dataset provides unique ID about housing (from fiscal sources), we can merge both dataset and then filter on the change year. That way, we can distinguish change of owners occurring with transfers from purchase. Furthermore, it enables us to distinguish partial changes of owners (*e.g.* consecutive to a divorce widowhood) from full ones.

Our method, to be valid, requires the missing rate, defined as the share of transaction for pecuniary interest associated to no change of owners, to be as low as possible. We then run the merging process as defined previously and estimate the missing rate. Remark that we do not restrict our datasets (for instance remove outliers). However our identification might underestimate the number of transfers as we cannot observe within year change based on the panel data. Hence, if a dwelling is transferred and consecutively sold within the same year, our method qualifies this change as purchase, as we match the change of owners with a transaction for pecuniary purposes. Then, it limits the possibility to exploit these data to study the decision to sell after receiving a transfer.

We report on Table 1.A.11 the missing rate based on the simple merging process between owners change (derived from panel data we construct) and the transaction dataset (derived from the DV3F database). We chose to characterise according to the previous type of owners and the following one, using the moral – natural person distinction. Indeed, as we mainly focus on natural persons, we are more interested to lower the missing rate for natural – natural transactions as they are more likely to be of interest for transfers.

Table 1.A.11: Missing Rate According to Owner Type and Temporal Assumption

	Legal	Natural		Legal	Natural
Legal	26.40	14.61	Legal	9.10	2.03
Natural	15.17	11.96	Natural	3.75	0.85
(a) Restricted Temporal Assumption			(b) Loosen Temporal Assumption		

Notes: We report the missing rate, defined as the ratio between matched transactions in DV3F and overall transactions in DV3F. Our studied period is 2011–2019. The column represents a type of sellers while the row specification represents purchasers. The left panel provide results with greater restriction for the temporal criterion, while the left panel relaxes the assumption (with one additional year).

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

1.A.14 Age Profile for Housing Changes

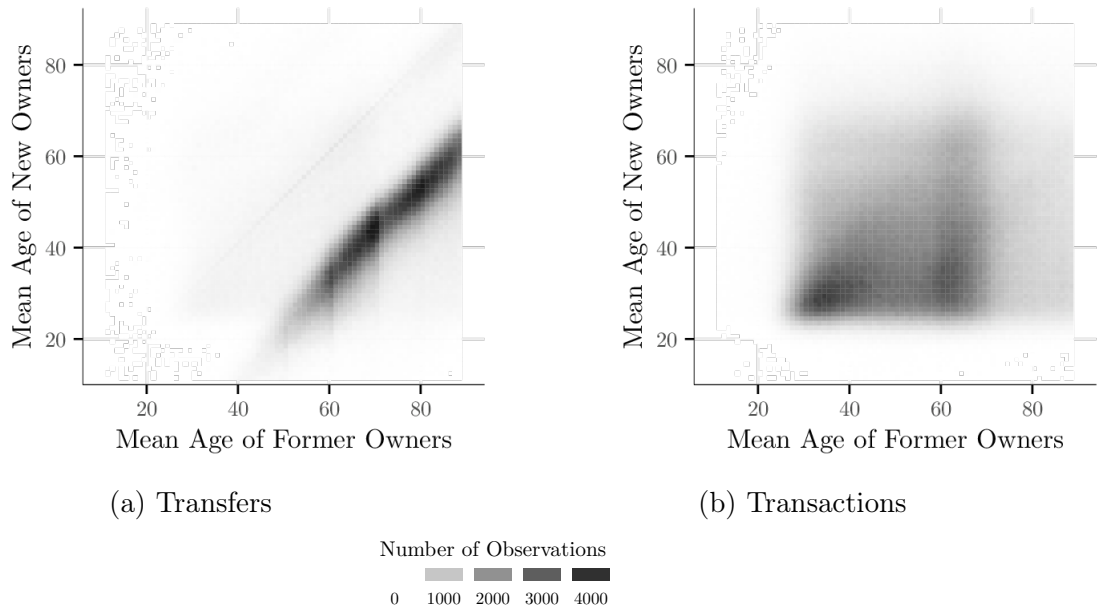


Figure 1.A.5: Age of New and Former Owners According to Change Type

Notes: We report the number of observations according to the mean age of both former (x-axis) and new owners (y-axis). We distinguish between transfer transaction (Figure a) and purchase (Figure b). We pool all year from 2012 to 2018. We remove observations with missing observation about the mean age of either former or new owners. *Sources:* Authors' calculation based on DV3F and *Fichiers Fonciers*.

1.B Additional Material

1.B.1 Partial Correlation Between Housing Wealth Variation, Initial Level of Wealth and Location of Main Residence

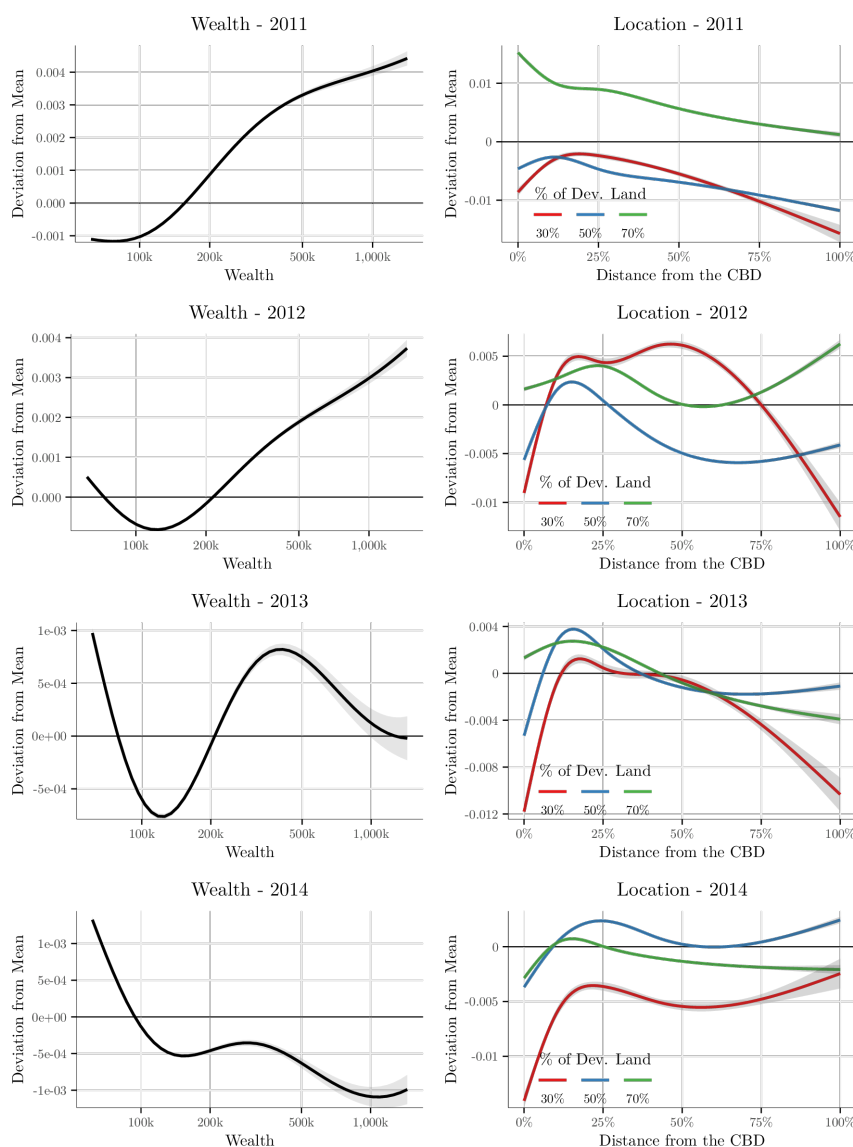


Figure 1.B.1: Partial Correlation Between Housing Wealth Variation, Initial Level of Wealth and Location of Main Residence (2011-2014)

Notes: We report partial correlation with wealth evolution according to the previous of housing wealth (left panel) and locations (right panel). These results stems from the estimation of Equation 1.3. Our observation unit is distinct owners that does not experience portfolio changes in 2018. We report confidence interval to a 95% confidence interval. We use the `mgcv` R package.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

1.B.2 Partial Correlation Between Housing Wealth Variation, Initial Level of Wealth and Location of Main Residence

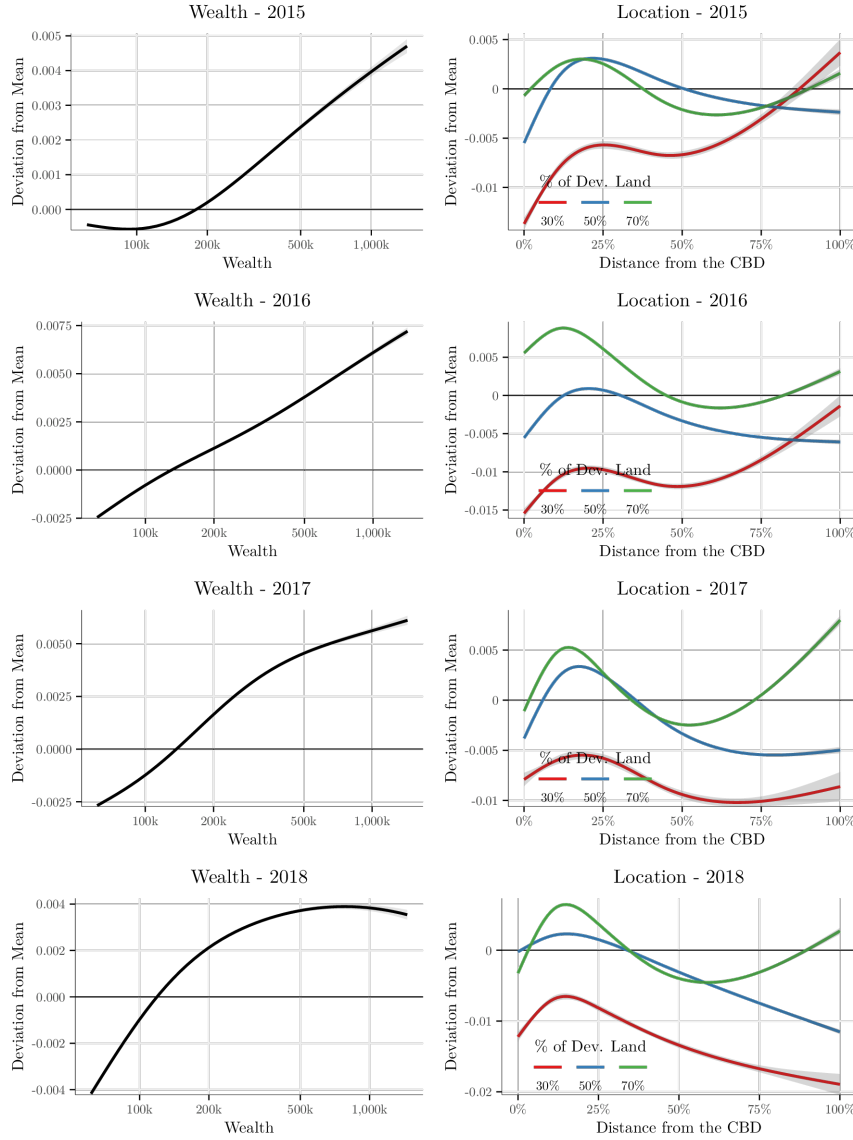


Figure 1.B.2: Partial Correlation Between Housing Wealth Variation, Initial Level of Wealth and Location of Main Residence (2015-2018)

Notes: We report partial correlation with wealth evolution according to the previous of housing wealth (left panel) and locations (right panel). These results stems from the estimation of Equation 1.3. Our observation unit is distinct owners that does not experience portfolio changes in 2018. We report confidence interval to a 95% confidence interval. We use the `mgcv` R package.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

1.B.3 Partial Correlation of Initial Level of Wealth with Housing Decisions within Commuting Zones for Owner-Occupiers

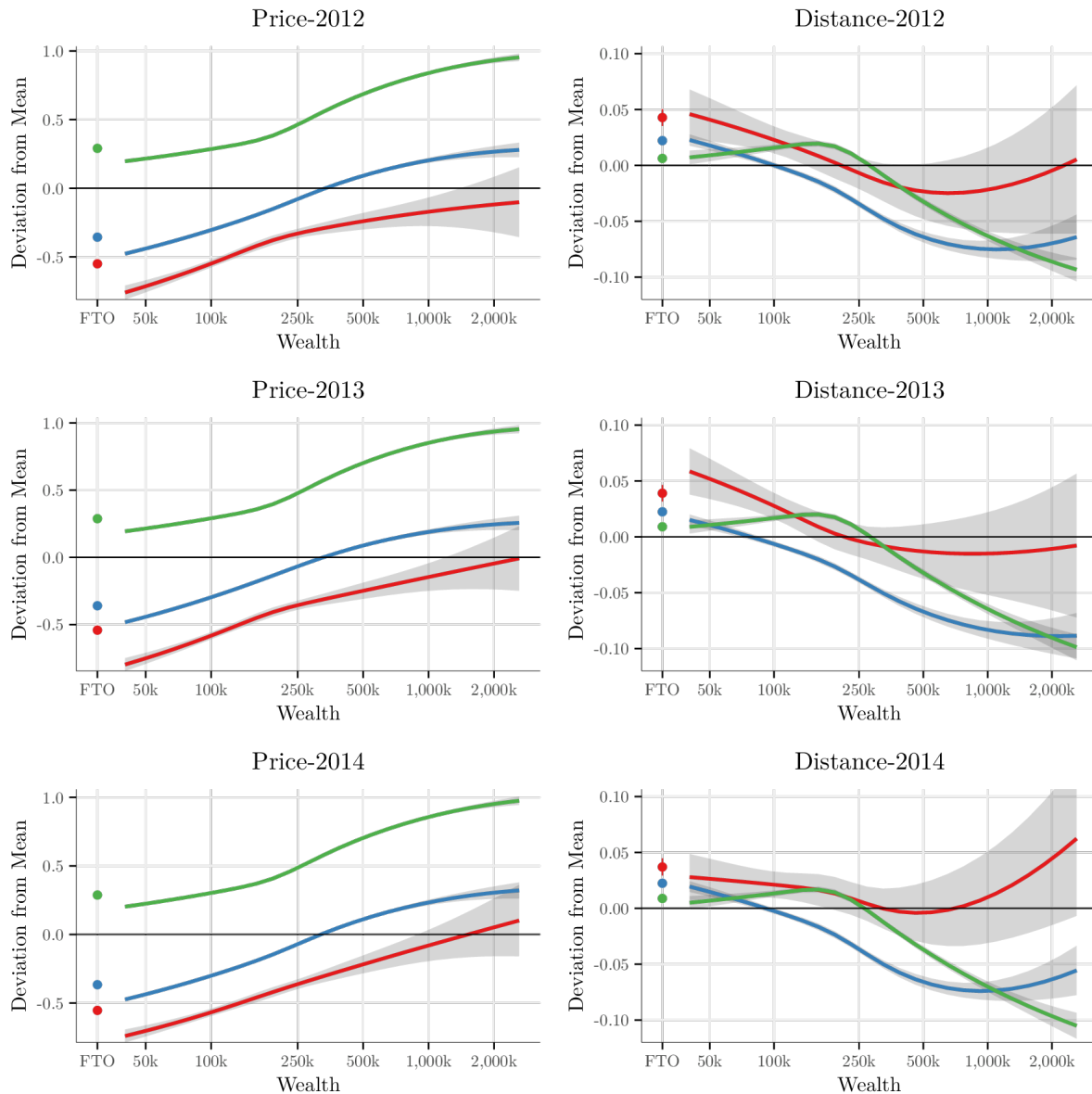


Figure 1.B.3: Correlation between Previous Wealth, Location Choices and Transaction Price for Owner-Occupiers (2012-2014)

Notes: We report the partial correlation for the interaction between the economic status of the commuting zone and the initial wealth level of the buyer from Equation 1.4. The left panel corresponds to the transaction price, while the right panel corresponds to the distance from the CBD. We introduce variation in commuting zone characteristics as we estimate the wealth correlation for commuting with 30%, 50% and 70% of built-up area in 2000. We also introduce first-time homeowners (dot, left side of plot). We use the *mgcv* R package to implement additive transformations of variable with endogenous definition of degree of freedom.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

1.B.4 Partial Correlation of Initial Level of Wealth with Housing Decisions within Commuting Zones for Owner-Occupiers (Bis)

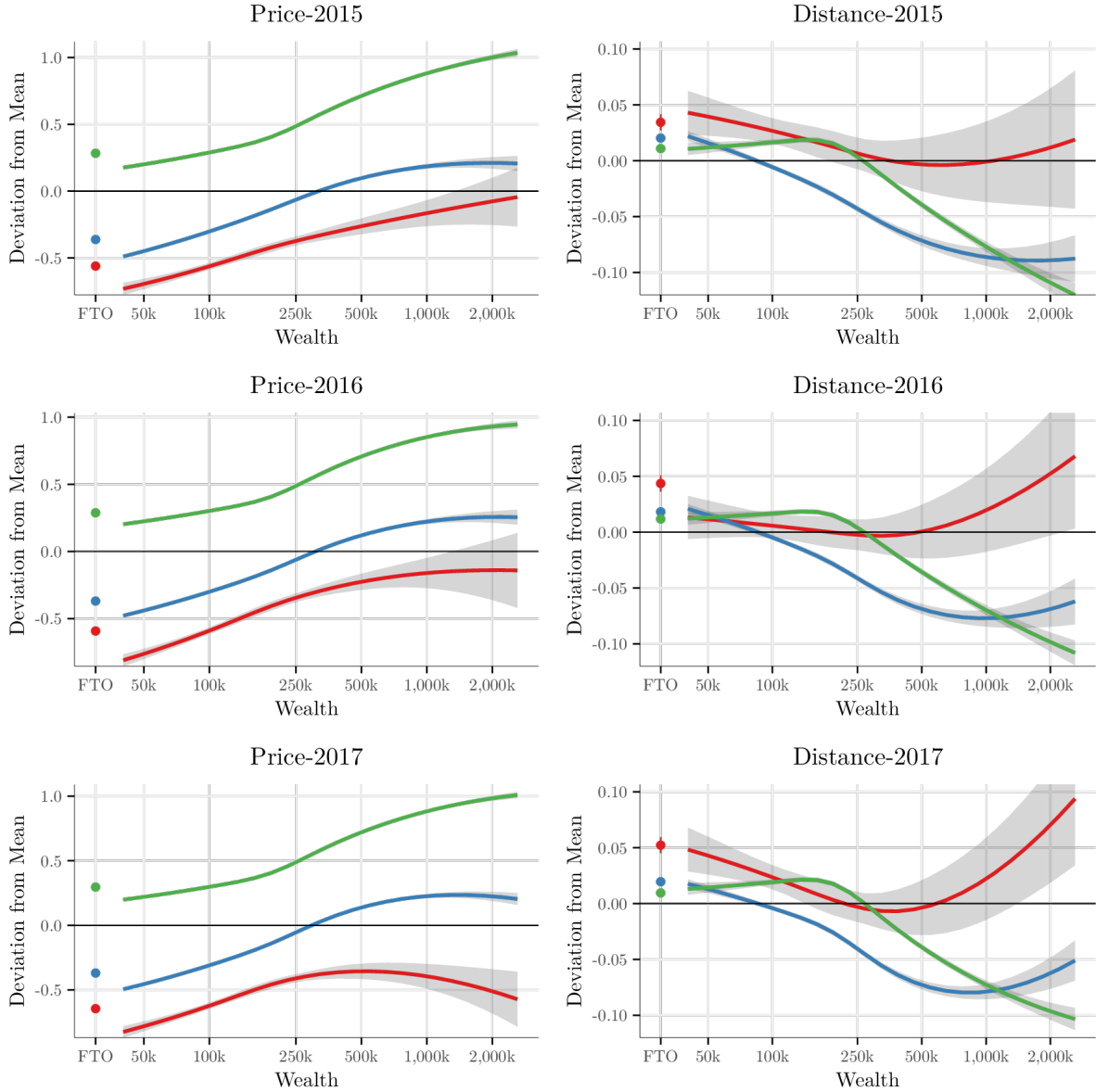


Figure 1.B.4: Correlation between Previous Wealth, Location Choices and Transaction Price for Owner-Occupiers (2015-2017)

Notes: We report the partial correlation for the interaction between the economic status of the commuting zone and the initial wealth level of the buyer from Equation 1.4. The left panel corresponds to the transaction price, while the right panel corresponds to the distance from the CBD. We introduce variation in commuting zone characteristics as we estimate the wealth correlation for commuting with 30%, 50% and 70% of built-up area in 2000. We also introduce first-time homeowners (dot, left side of plot). We use the *mgcv* R package to implement additive transformations of variable with endogenous definition of degree of freedom.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

1.B.5 Mean Housing Transfers per Year

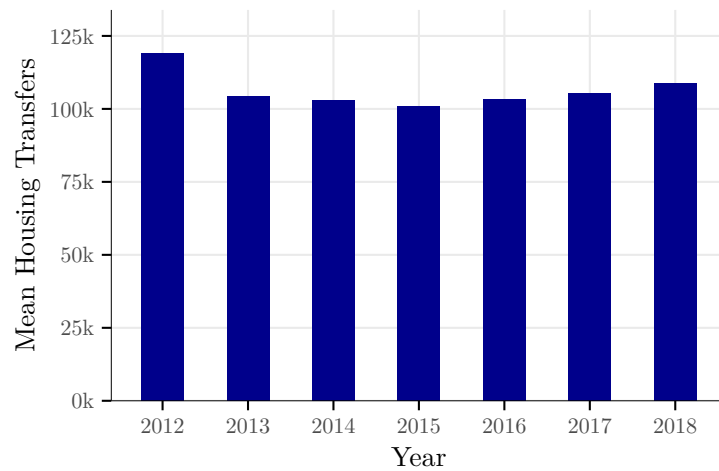


Figure 1.B.5: Mean Housing Transfers per Year per Individuals

Notes: We report the estimation of the average housing transfers per year. The housing price is estimated using our housing market value imputation. If the same housing is split between recipients, we assign equal share of housing value to recipients. Transfers are restricted to housing one.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

1.B.6 Partial Correlation of Initial Level of Wealth and Residence Place with Housing Transfers

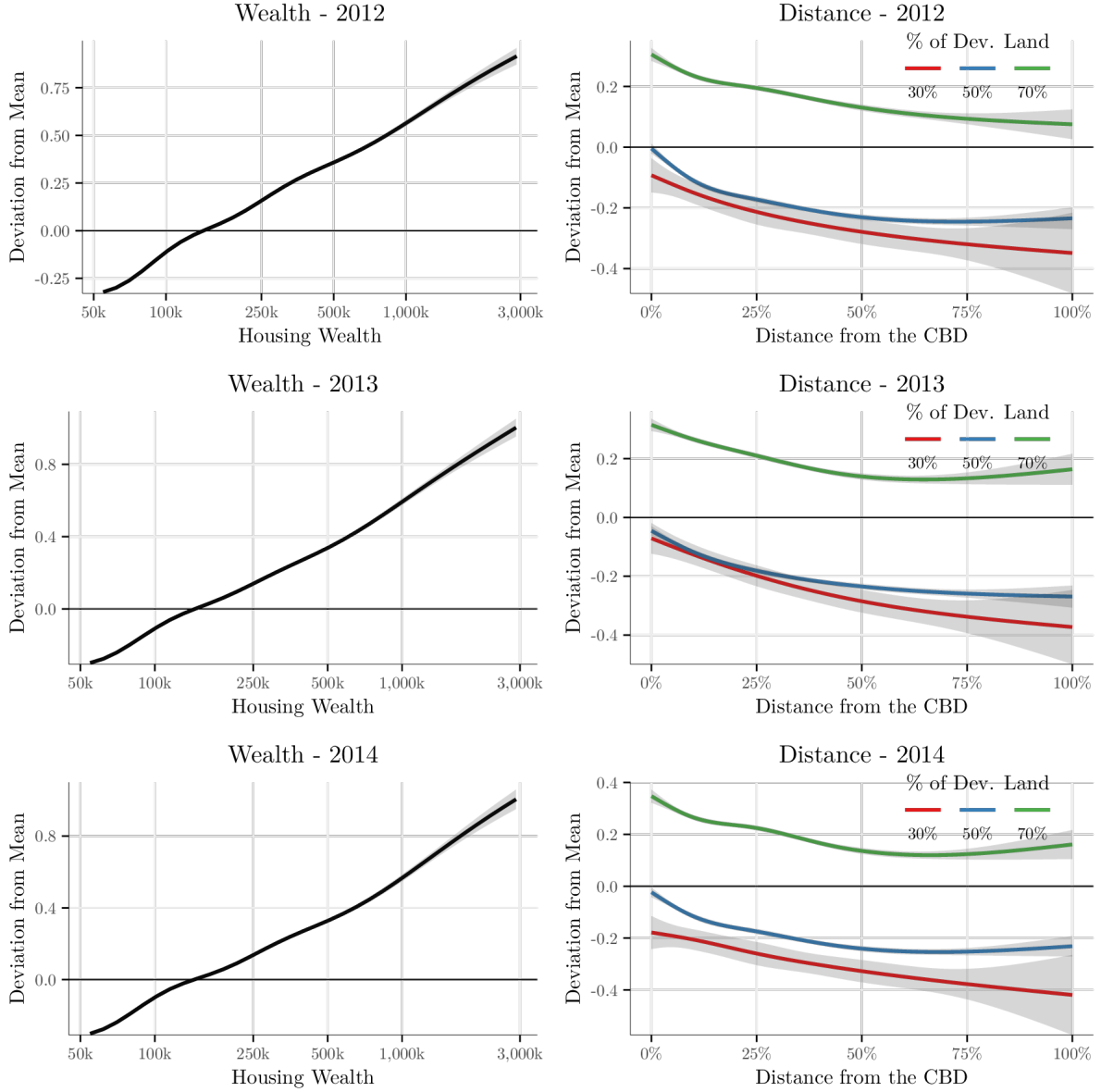


Figure 1.B.6: Partial Correlation of Initial Level of Wealth and Residence Place with Housing Transfers (2012–2014)

Notes: Our results are derived from estimation of Equation 1.5 restricted to housing transfers recipients. The left panel reports marginal contribution according to previous current wealth. The right panel reports spatial heterogeneity based on recipients location. We report pooled results over year for the spatial contribution. We report confidence interval at a 90% level. We use the `mgcv` R package.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

1.B.7 Partial Correlation of Initial Level of Wealth and Residence Place with Housing Transfers (Bis)

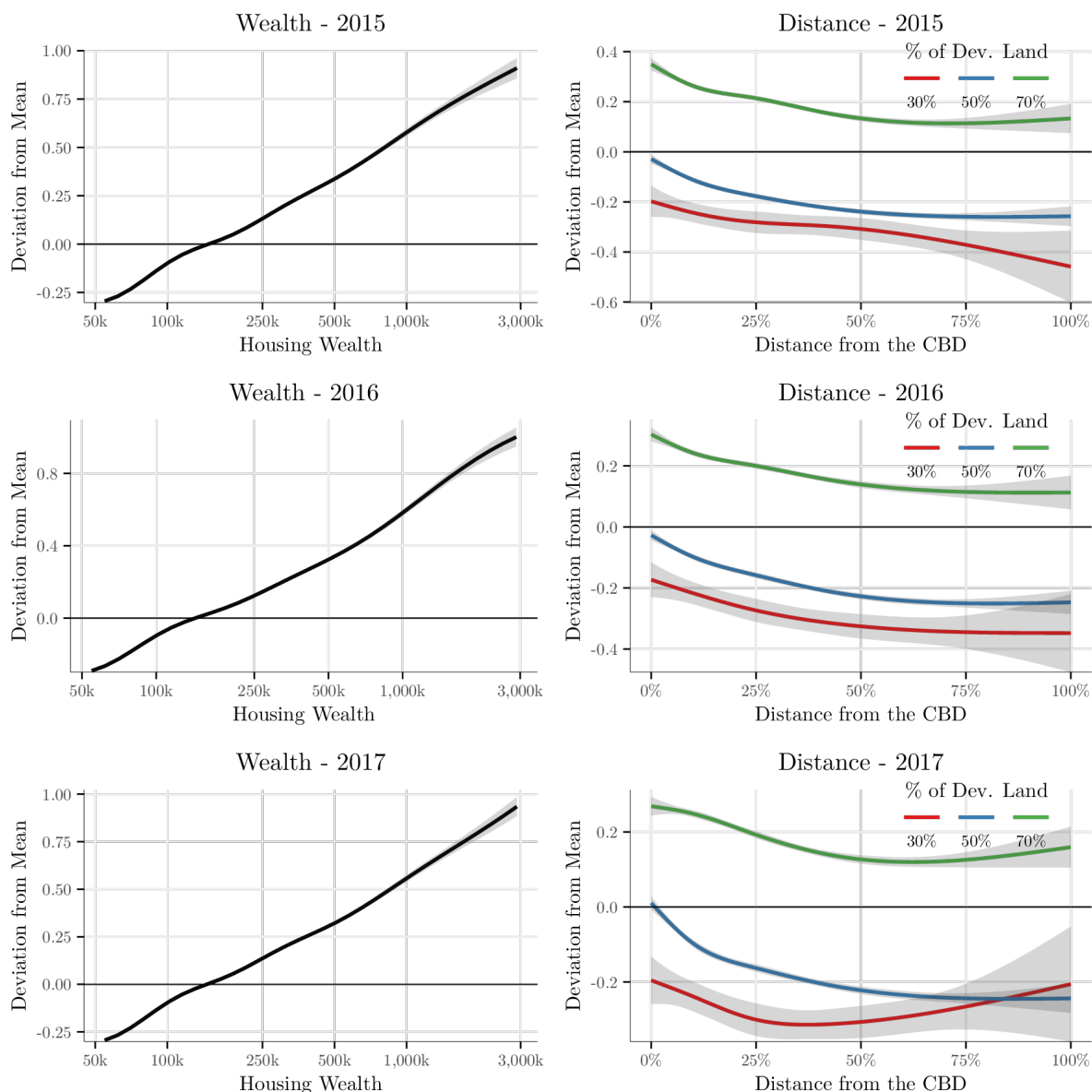


Figure 1.B.7: Partial Correlation of Initial Level of Wealth and Residence Place with Housing Transfers (2015–2017)

Notes: Our results are derived from estimation of Equation 1.5 restricted to housing transfers recipients. The left panel reports marginal contribution according to previous current wealth. The right panel reports spatial heterogeneity based on recipients location. We report pooled results over year for the spatial contribution. We report confidence interval at a 90% level. We use the `mgcv` R package.

Sources: Authors' calculation based on DV3F and *Fichiers Fonciers*.

Chapter 2

Cost-Efficiency of Supporting Homeownership: Welfare-Based Measure from France

Abstract

This article develops a welfare analysis of subsidized loans for the access to homeownership, both in terms of number of additional homeowners (extensive margin) and change in the value of housings (intensive margin). We leverage pre-treatment variables and the precise location of homeowners to ensure a selection-on-observables identification strategy. We estimate doubly robust causal effects both on the policy recipients and the whole population of homeowners. Spatial and temporal variations of implementation details and credit market conditions allow us to relate these policy-relevant treatment effects to the budget costs of the French interest-free loan policy. We cannot reject that the policy has no causal effect on the extensive margin, whereas it significantly impacts the intensive margin. For different values of externalities (e.g. better housing maintenance, price capitalisation or reduced mobility) for both margins, we find that increasing policy intensity through more public spending negatively affects its cost-efficiency. This is true independently from credit market conditions that do not directly impact policy efficiency.

JEL classification: H81, R21, R38

Keywords: Housing policy ; social welfare ; unconfoundedness ; generalised additive model ; spatial smoothing.

2.1 Introduction

In most developed countries, supporting homeownership is one of the largest housing policies (Andrews and Caldera Sánchez, 2011). These public interventions are motivated by the positive externalities from homeownership, expected to be larger than the negative ones (Coulson and Li, 2013). It was found that homeownership increases children school performance (Green and White, 1997; Harkness and Newman, 2003), housing maintenance (Harding et al., 2000), and political involvement (Hall and Yoder, 2022) despite a risk of negative equity (Cunningham and Reed, 2013) and lower mobility (Green and Hendershott, 2001). Furthermore, this public support is a response to the affordability crisis (Fisher and Gervais, 2011; Carozzi, 2020), which could mitigate the growing importance of parental assistance (Boehm and Schlottmann, 1999; Lee et al., 2020) and provide a more equal access to housing across the population (Green and Vandell, 1999).

Despite the positive externalities expected from the extensive margin (i.e. number of additional homeowners), such policies may have unintended consequences by distorting individual behaviours at the intensive margin (though the value of housings sold). Previous research found that subsidized households purchase more expensive (Davis et al., 2020) and larger houses (Hanson, 2012), and that their location choices change (Daminger and Dascher, 2022). It was also shown that the policy-induced increase of demand leads to higher housing prices (Martin and Hanson, 2016; Kunovac and Zilic, 2022) and to lower supply of affordable housings (Sommer and Sullivan, 2018; Gabriel and Painter, 2020). The balance between the two margins is not clear-cut in the empirical literature, even if the unintended effects from the intensive margin seem to dominate the expected ones. Causal effects at the extensive margin are either small or insignificant (Glaeser and Shapiro, 2003; Hanson, 2012; Hilber and Turner, 2014; Bäckman and Lutz, 2020; Kunovac and Zilic, 2022). An exception is Hembre (2018), who found a sizeable effect for a US policy that comes at a large cost for the administration in a specific context (post-2008 crisis). The efficiency of these policies then crucially depends on the balance between the two margins and their budget costs for the governments, the elements that are under study in this article.

Whereas homeownership support is based on mortgage contracts in almost all countries, the lack of clear evidence is mainly due to policy designs that vary across countries. We can distinguish two main approaches in existing policies. In the first one, the subsidy is channelled through income tax as households deduct their payments of interests from their taxable income. The main US housing policy (the Mortgage Interest Deduction Glaeser and Shapiro, 2003; Hilber and Turner, 2014) follows this pattern, such as policies of other countries like Belgium (Hoebeeck and Inghelbrecht, 2017) and Denmark (Gruber

et al., 2021). In these cases, the treatment intensity depends on household income through differentiated marginal tax cuts. In the second approach, subsidized loans cover a share of the overall mortgage amount independently from the income of recipients. Treatment intensity is capped by an eligible share of the overall mortgage value and a specific population is targeted by income thresholds. Countries such as England (Carozzi et al., 2020), Croatia (Kunovac and Zilic, 2022), Germany (Daminger and Dascher, 2022), and the US (in a complementary policy to support homeownership after the 2008 crisis, Hembre, 2018) adopt this design. The French IFL (interest-free loan) policy follows this second approach with, given household eligibility, a public support that depends exclusively on housing loan characteristics (overall value, loan maturity, and interest rate).

Existing causal evidences largely depend on the context of the evaluation and make the comparison of different policy options and the derivation of general recommendations difficult (Angrist and Krueger, 2001). The analyses rely on difference-in-differences or spatial discontinuity designs from particular policy shifts or differences between neighbouring regions. They provide context-dependent local treatment effects that cannot be readily extrapolated to other policy changes or other regions (Hahn et al., 2001; de Chaisemartin and D’Haultfoeuille, 2020). Conversely, we provide a general framework to study the welfare effects of supporting homeownership, balancing budget costs with social benefits. Our identification strategy overcomes common LATE issues arising from quasi-experimental methods and allows to derive more robust policy recommendations. We consider that policy efficiency depends on the trade-off between budget costs and the extent of externalities from the extensive and intensive margins. Externalities at the extensive margin are assumed to be positive considering the homeownership benefit (e.g. better housing maintenance, wealth accumulation) over pitfalls (e.g. lower mobility), while the intensive margin ones are considered as negative in line with the affordability issues (Gabriel and Painter, 2020). We make the best of a selection-on-observables restriction to account for heterogeneous treatment effects from varying policy parameters and different economic contexts. We leverage differences in the IFL policy across space and time to fit dose-response functions between cost-efficiency according to the parameters that define treatment intensity named as primitive variables. We consider primitives controlled by policy makers (loan ceiling value and coverage share) and primitives from the credit market (loan maturity and interest rate). As such, we provide insights for other forms of homeownership support as well as for other policies involving public spending (Hendren and Sprung-Keyser, 2020).

As the criteria of the IFL policy depend on the tenseness of local housing markets by design and responds to the economic context, we control the implied endogenous treatment assignment by a two-step procedure. In the first step, we estimate GPS (Generalised

Propensity Scores, Imbens, 2000) by gathering a large set of pre-treatment variables accounting for demand-side (*e.g.*, population density, income, socio-economic status) and supply-side (*e.g.*, housing construction, previous prices, previous neighbouring prices) determinants of tenseness. To control for other spatial confounders, we introduce smooth functions of spatial coordinates of homeowners' locations in a semi-parametric GAM (Generalised Additive Model, Wood, 2017). In the second step, we regress the outcomes of the extensive (new homeowners) and intensive (housing values) margins on the same pre-treatment variables and spatial coordinates with an inverse probability weighting scheme derived from the GPS. This provides doubly robust estimates for a large panel of policy-relevant treatment effects, which are consistent if either the first or the second step is well specified (Robins and Rotnitzky, 1995; Słoczyński and Wooldridge, 2018). Importantly, we gather external data about the universes of first-time homeowner and housing transactions (subsidized or not), in order to distinguish the opportunistic effects on subsidized homeowners from the social externalities impacting all homeowners. A placebo analysis comparing subsidized homeowners with the same treatment level while in housing markets with different tenseness cannot allow to reject the selection-on-observables restriction.

We contribute to the literature by discussing the welfare implications of small variations of policy costs. Not only do we find consistent results with existing ones, we also accurately estimate that the policy produces most externalities at the intensive margin. This is noticeable knowing that the extensive margin is generally targeted by such policies, whereas the intensive margin is an unintended consequence. Firstly, we find the number of subsidized homeowners increases with the ceiling values set by the policy, with the interest rate in effect, and so with the cost of the policy. Contrariwise, we do not find any causal effect of these primitives on the total number of first-time homeowners (subsidized or not). In the counterfactual framework under consideration, this indicates that the additional subsidized homeownership would still have existed in the absence of additional spending. Secondly, the intensive margin is impacted by policy spending both for subsidized housing transactions and all transactions. Opportunistic behaviours from subsidized homeowners (taking advantage of the policy to get more expensive housings) appear to be more significant than inflationary effects from housing market feedback. Thirdly, we provide evidence that cost-efficiency would be higher for lower coverage shares of subsidized loans and so for lower budget costs for the government. This result applies for a wide range of social values given to homeownership-related externalities and is robust to credit market conditions.

The remaining of the paper is structured as follows. Section 2.2 presents the institutional context of the French IFL policy and the datasets we gather. In Section 2.3, we introduce the welfare-based measures of cost-effectiveness, taking into account the exter-

nalities from the extensive and the intensive margins. Our identification restriction to handle endogenous treatment intensity and the doubly robust estimation procedure are presented in Section 2.4. The following Section 2.5 provides the empirical results from the two-step approach, including the placebo analysis, and the last Section 2.6 concludes.

2.2 Context and Data

2.2.1 The French IFL Policy

The IFL policy was introduced in 1995 in France to encourage first-time homeownership. Recipients benefit from a subsidized loan with no interest to pay for a given share of a maximum ceiling value of their overall housing credit. The cost for public finance equals the sum of interests not at charge for recipients, as the government consents a tax reduction to private lenders (banks) supplying such contracts. As shown in Figure 2.A.1 of Online Appendix (OA), the number of recipients since the beginning of the policy is about 3.2 millions, for an overall budget cost of about 26.1 billions euros (about 8,000 euros by recipient). The policy excludes a small share of high-income households (about 10% of richer tenants according to Sotura, 2020) and targets specifically the newly built housings (although existing housings are eligible conditionally on renovation in certain periods). Any IFL contract is associated with a classical loan with interests to pay, so recipients must comply with the usual conditions to access to the credit market.¹

The mechanism pertaining to the IFL remains stable over time and relies mainly on four primitive parameters. For each period and each group of municipalities (to which we will return below), policy makers fix s , the maximum coverage that the IFL can represent in the total loan amount \tilde{V} (ranging from 10% to 40%) and \bar{V} a maximum ceiling value on the loan (ranging from 100 to 150 thousand euros for one-person household). These are the two primitives under control for the government to implement the IFL policy.² The budget cost for the policy also depends on two other primitives from the credit market, which are not chosen by the government and are considered as exogenous to the IFL policy: the interest rate r and the loan maturity m . As shown in Section 2.B of OA, the budget cost c of a IFL contract equals the subsidy equivalent for the recipient, which is:

¹For the French market, a usual condition to access to a classical loan is that reimbursement payments cannot overcome 33% of income. Consequently, this also restricts the access to the IFL policy.

²We do not consider the income cap as a primitive because this parameter conditions eligibility, with a marginal impact according to their restriction (Sotura, 2020). Moreover, we expect credit constraints to be less restrictive for the wealthiest households in their tenure decisions. This is also the case for allowed deferred reimbursements that are not of main policy concern.

$$c = \left[\frac{m \times r}{1 - (1 + r)^{-m}} - 1 \right] \times s \times \min(\tilde{V}, \bar{V}). \quad (2.1)$$

Section 2.B of OA also shows that the budget cost of a IFL contract weakly increases with the four primitives, which indicates that increasing one of the four primitives is equivalent to increasing the cost of the policy. This defines the treatment intensity of the policy, equivalently as an increase of one of the four considered primitives or an increase of the budget cost for the government. We restrict our studied period on the last three IFL waves of the 2015–2019 period as eligibility conditions remains unaffected (Figure 2.A.1 of OA) and the classification of municipalities with given primitives does not change.

2.2.2 The ABC Zoning

Both ceiling values and eligibility shares depend on the location of IFL contracts, from an exhaustive and mutually exclusive classification of French municipalities based on the tenseness of the housing markets.³ This ABC zoning introduces four ordered degrees of tenseness, from C the lowest level, to B₂ and B₁ the intermediate levels, and A the highest level. This official zoning was updated four times since its introduction in 2003, the latest update of October, 2014 is stable for the 2015–2019 period under study.

Most French municipalities are rural and belong to zone C, the lowest level of the zoning (Panel A of Table 2.2.1). Table 2.2.1 also shows that the distribution of the ABC classification follows the intuition, as municipalities with higher population density and higher housing price per living area (unitary prices hereafter) are higher in the hierarchy. Despite the correlations between the ABC hierarchy and reported pre-treatment variables, it is well recognised that the classification assigns quite similar municipalities to different levels. The French administration in charge of monitoring public expense noted in 2012 the lack of transparency of the zoning (Cour des Comptes, 2012). It concluded that the zoning does not depend exclusively on objective characteristics, suggesting potential subjectivity in the assignment. Most existing quasi-experimental approaches dealing with the endogeneity of IFL treatment assignment rely on the arbitrariness of this zoning (Labonne and Welter-Nicol, 2015; Beaubrun-Diant and Maury, 2021; Chareyron et al., 2021).

Beyond variations introduced by the design of the IFL policy, credit market conditions (interest rate and loan maturity) also vary over the considered period (Panel C of Table 2.2.1). These variations are not clearly correlated with the ABC zoning, as they mainly

³According to official documents, tenseness is defined “from the imbalance between the housing supply and the housing demand” (French Ministry of Ecological Transition).

depend on the exogenous economic context. From the variations of the four primitives, the budget cost of the policy experienced sizeable changes across multiple cross-sectional and temporal dimensions while it is a national policy (Figure 2.A.1 of OA). Hence, it results that the causal evaluation of the IFL policy has to be considered with a varying treatment intensity framework rather than a more classical counterfactual experiment of policy removal.

2.2.3 Data

We aggregate three exhaustive individual data sources at the municipal level ($N = 34,970$) and match them with demographic data. We filter outliers, missing observations or data inconsistencies to obtain a final sample of 26,819 municipalities. We report descriptive statistics in Section 2.A for the relevant variables used in the empirical analysis.

IFL files. The first individual database (SGFGAS) concern all recipients subsidized by the IFL policy. Each recipient is located at the municipality level of its new home, with variables informing the loan contract (total value of the main and subsidized loans, total subsidies, interest rates, and maturity). These data also contain households' characteristics such as annual income, number of members, matrimonial status and previous location when tenants. Finally, these data include characteristics of the housing concerned by the loan, such as the construction date, surface, purchase price, and purchase date. We use them to construct aggregated values for each municipality, by computing for each year the number of IFL contracts and averaging loan, housing and household characteristics.

Tax files. To circumvent the problem of having only subsidized new homeowners from the IFL files, we use exhaustive tax files about French homeownership (*Fichiers Fonciers*) to determine the total number of new homeowners (subsidized or not). Using the temporal dimension of these administrative data, we identify first-time homeowners as defined by the IFL policy, *i.e.*, homeowners that were not homeowners at least two years ago. We obtain for each municipality the number of such new homeowners, which were eligible to the IFL, by counting the number of homeowners that were absent from the tax file in the previous two years. While the tax files and IFL files are independent data, the total numbers of first-time homeowners estimated from tax files are always higher than the numbers of contracts from IFL files (except for two municipalities that are removed from our sample). Finally, we recover the total number of newly built housing over the 2010–2013 period based on the construction year reported in the tax files.

Table 2.2.1: Main variables and primitives for municipalities along the ABC zoning

Variable	Period	Country	ABC Zoning Areas			
			A	B ₁	B ₂	C
A - Pre-treatment variables						
Number of Municipalities (thousand of units)	2013	34.970	0.724	1.535	3.828	28.883
		100%	2.07%	4.39%	10.95%	82.59%
Housing Price (thousand euros)	2010–2013	153.1 (68.0)	284.0 (124.8)	234.9 (74.8)	188.6 (59.6)	139.0 (54.9)
Unit. Housing Price (euros by squared meter)	2010–2013	1,608.7 (691.8)	3,558.7 (1,054.5)	2,597.0 (557.5)	2,003.5 (569.3)	1,430.9 (502.6)
Unit. Price of Neighbors (euro per squared meter)	2010–2013	1,561.0 (724.0)	3,654.1 (1,099.1)	2,618.6 (587.2)	1,975.2 (582.8)	1,371.6 (517.1)
Population Density (inhabitants by hectare)	2013	1.9 (8.1)	26.4 (38.2)	6.8 (9.4)	3.1 (4.6)	0.7 (1.0)
Median Household Income (thousand euros by year)	2013	20.0 (3.4)	25.3 (6.0)	24.0 (4.8)	22.1 (3.7)	19.2 (2.6)
Number of New Housings (number of units)	2010–2013	41.9 (226.7)	364.5 (918.7)	226.8 (648.3)	72.4 (151.7)	15.9 (28.7)
B - IFL Policy Values						
Maximum Ceiling Value (thousand of euros)	2015	-	150.0	135.0	110.0	100.0
	2016–2017	-	150.0	135.0	110.0	100.0
	2018–2019	-	150.0	135.0	110.0	100.0
Maximum Coverage Share (percent)	2015	-	26.0	26.0	21.0	18.0
	2016–2017	-	40.0	40.0	40.0	40.0
	2018–2019	-	40.0	40.0	20.0	20.0
Maximum IFL Amount (thousand of euros)	2015	-	39.0	35.1	23.1	18.0
	2016–2017	-	60.0	54.0	44.0	40.0
	2018–2019	-	60.0	54.0	22.0	20.0
Average Subsidy (thousand euros)	2015	5.21	10.31	9.41	5.81	4.02
	2016–2017	9.63	13.96	12.35	10.45	8.99
	2018–2019	5.03	12.44	10.89	4.64	3.98
C - Mortgage Market Conditions						
Mortgage Maturity (percent)	2015	228	244	254	238	221
	2016–2017	260	268	269	265	258
	2018–2019	258	267	267	262	255
Annual Interest Rate (number of months)	2015	2.51	2.46	2.45	2.52	2.52
	2016–2017	1.87	1.77	1.77	1.89	1.88
	2018–2019	1.62	1.53	1.52	1.65	1.63

Notes: The French municipalities are classified according to the ABC zoning in columns. Panel A reports the average and standard deviation of pre-treatment variables used to control the endogenous treatment assignment. The first three variables of panel B correspond to the IFL parameters for each period (constant between municipalities) with a Maximum IFL Amount that equals the maximum ceiling value times the Maximum Covering Share. The Average Subsidy is computed from IFL data and Equation 2.1. The last panel C reports the average of loan maturities and interest rates, also extracted from IFL files.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

Transaction data. We use a third exhaustive individual dataset about housing transactions (DV3F) to compute, at the municipal level, the average unitary price of housings over the pre-treatment period 2010–2013. In order to mitigate border effects in the delineation of housing markets and tenseness, we also compute the averages of the unitary prices over the same pre-treatment period for neighbouring municipalities. We remove the municipalities of the Corsican island due to geographic constraints related to the spatial smoothing we use in the econometric methodology (360 observations) and municipalities of the *Alsace-Moselle* region (1,605 observations) as transaction data are missing for this region for administrative and historical reasons. These data also allow building variables about the post-treatment outcomes for the overall housing markets, including the average housing prices, the average surface, and the average unitary price for the three periods.

Demographic data. For each municipality considered, we obtain from the *French National Institute of Statistics and Economic Studies* (INSEE) pre-treatment 2013 data on population density, median income, and socio-professional categories. This implies removing 6,186 small rural municipalities with low population and low population density, because of missing values on median income due to statistical secrecy.

2.3 Welfare-based measures of cost-efficiency

2.3.1 Welfare Framework

We develop a welfare-based framework for IFL policy evaluation, in order to compare various policy options in different economic contexts. In line with the varying treatment intensities of the French policy, we derive measures of cost-effectiveness from small changes in policy primitives and budget costs. Our approach is inspired from the sufficient statistic literature and the marginal value of public funds (Finkelstein and Hendren, 2020) although we adapt the general framework to fit our case.

Consider a social planner concerned by the welfare of both IFL recipients and the rest of the population. Its objective function is the sum of recipient utilities u_i (including IFL cuts in interests) plus the aggregated externalities of both margins, minus the budget costs for public finances. The externalities come from the extensive margin (noted N for the number of homeowners) and the intensive margin (noted V for the value of housings, as a proxy for housing characteristics and the prices of these characteristics). We note respectively ϕ and ψ the marginal social values of these externalities, such that:

$$W = \sum_i u_i + \phi N + \psi V - \sum_i c_i. \quad (2.2)$$

This welfare function distinguishes targeted policy recipients i that perceive IFL subsidies from the aggregated externalities N and V concerning all the population (recipients included). Each aggregated externalities are characterised according to their marginal valuation (respectively ϕ and ψ). A general agreement considers the whole externalities generated by homeowners to be positive (Coulson and Li, 2013), while affordability issues are negative (Gabriel and Painter, 2020). Then, we expect externalities from the extensive margin to be positive and externalities from the intensive margin to be negative such that $\phi > 0$ and $\psi < 0$. Yet, given that the balance between positive and negative externalities at both margins is not clear-cut,⁴ we introduce variations for credible values to assess the cost-efficiency in various economic context. Finally, we also consider value of externalities opposite to the expected value to obtain a comprehensive overview of policy efficiency, and estimate efficiency according to targeted populations by policy makers. Indeed, while the price capitalisation might be negative for social welfare of the whole society, it likely turns positive for current homeowners.

In addition to the two welfare-based externalities measured by the number of homeowners and the average housing values, we study intermediate outcomes to investigate in more details the causal effects of the IFL policy. We consider the average housing surface S and the average unitary housing price P in order to decompose the aggregated intensive margin effect with $V = S \times P$. The variations of housing values could come from variations in surfaces or variations in unitary prices. To distinguish the external effects for the recipients to those for the whole population, we use the average value \tilde{V} , surface \tilde{S} , and unitary price \tilde{P} of housings chosen by recipients. These latter outcomes represent opportunistic behaviours produced by the policy, which are different from housing market feedback that applies to the whole population. We also consider the number of IFL recipients \tilde{N} that does not correspond to the full extensive margin, as increasing \tilde{N} without changing N cannot be considered as favouring homeownership. Without significant difference in N , all additional subsidized new homeowners \tilde{N} are opportunistic and would still have existed without the policy. These last outcomes are intermediate variables that could be policy objectives, but do not represent full welfare and are not used to measure cost-efficiency.

2.3.2 Defining Cost-efficiency

Each term of the welfare function (2.2) implicitly depends on the four policy primitives of Equation 2.1. Hence, the following Equation 2.3 presents the welfare variations from small variations in a given primitive $a \in \{s, \bar{V}, m, r\}$:

⁴For instance, supporting homeownership, despite positive externalities such as better housing maintenance or political involvement, raise mobility cost in comparison to renters.

$$\frac{\partial W}{\partial a} = \sum_i \frac{\partial u_i}{\partial a} + \phi \frac{\partial N}{\partial a} + \psi \frac{\partial V}{\partial a} - \sum_i \frac{\partial c_i}{\partial a}. \quad (2.3)$$

For small variations of the primitive a , it is recognised that behavioural responses from recipients have no first-order response on their utilities (Kleven, 2021). Assuming optimal consumption from recipients, envelop theorem ensures that changes in the primitives are zero-sum monetary transfers between the government and recipients (independently from the produced externalities). The willingness-to-pay for a IFL transfer equals the budget cost for the government, as for any lump-sum payment.⁵ Thus, the utility for the recipients and the budget costs cancel in the derivative of the welfare function and the welfare variations rely exclusively on the externalities from both margins. We define the cost-efficiency measure of increasing a given primitive, noted e_a , by the marginal variation of the value of externalities for one monetary unit of additional budget cost, such that:

$$e_a \equiv \frac{\phi \cdot \partial N / \partial a + \psi \cdot \partial V / \partial a}{\sum_i \partial c_i / \partial a}. \quad (2.4)$$

As the total cost of the policy $\tilde{C} \equiv \sum_i \frac{\partial c_i}{\partial a}$ increases with the four primitives (Equation 2.1), this efficiency measure has the same sign as the derivative of the welfare function W (under the conditions of the envelop theorem). If $e_a > 0$, the welfare increases in a whereas it decreases in a if $e_a < 0$ (welfare does not change if $e_a = 0$). Comparing these efficiency measures e_a with -1 and 1 is also relevant for policy makers, as marginal benefits of increasing a are higher than marginal costs if $e_a > 1$. This indicates that one more monetary unit invested in the policy produces more than one euro of additional externalities. If $e_a < -1$, one unit allocated to support homeownership produces a welfare loss of more than one monetary unit. This cost normalisation allows us to interpret welfare changes in monetary units consequently to one additional unit of budget cost, while both variations transit through primitive parameters a . As such, the efficiency of different current policy options could be compared, as other public policies or other economic contexts. The tractability of these efficiency measures is also convenient to study alternative outcomes (replacing V by \tilde{V} for example) in regards to different policy objectives.

⁵This result is borrowed from the literature about the marginal value of public funds (MVPF) that does not account for the deadweight cost of raising tax (Finkelstein and Hendren, 2020). The main difference is that we do not include the net government cost by not considering the increase of housing or land taxes from additional homeowners. The efficiency measure cannot be directly compared to the MVPF.

2.3.3 Dose-response Functions

We recover the cost-efficiency measures for the IFL policy from the counterfactual framework (Rubin, 1974; Finkelstein and Hendren, 2020), through dose-response functions relating policy-relevant treatment effects to the four primitives of interest. Variations of the IFL policy across the four ABC zones and the three periods define a multi-valued treatment taking $G = 12$ levels. Let g denote a level of treatment and T_g a dummy variable indicating whether the municipality receives this level. This means:

$$Y = \sum_{g=1}^G T_g Y_g, \quad (2.5)$$

where Y is the observed outcome, equals to its potential value Y_g only if a municipality receives treatment g . The main outcomes of interest are $Y = N$ for the extensive margin, $Y = V$ for the intensive margin, and $Y = \tilde{C}$ for budget costs, while we are also interested in the other intermediate outcomes described above. Each bilateral combination of different treatment levels g and g' corresponds to a variation of at least one policy primitive. We exploit this structure of the IFL policy to map policy-relevant treatment effects to primitive variations. As the efficiency measures are based on small variations of primitives, we retain a set linear dose-response functions for each outcome Y with:

$$\mathbb{E}(Y_g - Y_{g'}) = \beta_0^Y + \sum_a \beta_a^Y (a_g - a_{g'}) + \xi. \quad (2.6)$$

with mean-independent errors ξ . The Ordinary Least Squares (OLS) coefficients β_a^Y provide a summary of the effects of primitives a on the heterogeneity of treatment effects $\mathbb{E}(Y_g - Y_{g'})$ and allow to recover the derivatives of the efficiency measures from Equation 2.4. For outcomes concerning the whole population, $Y \in \{N, V, P, S\}$, the average treatment effects (ATEs) are clearly policy relevant as they appear in the left-hand side of Equation 2.6. In effect, ATEs represent the change of Y caused by the policy g relatively to g' for the whole population and β_a^Y summaries how these changes can be attributed to the differences between a_g and $a_{g'}$. For outcomes concerning only recipients, $Y \in \{\tilde{N}, \tilde{V}, \tilde{P}, \tilde{S}, \tilde{C}\}$, the policy-relevant treatment effects are about the recipients (ATTs). The left-hand side of the dose-response function (Equation (2.6)) are then $\mathbb{E}(Y_g - Y_{g'} \mid T = g)$. As we study bilateral combinations *within* three distinct periods, this gives $4 \times (4 - 1) \times 3 = 36$ policy-relevant treatment effects. Therefore, each set of dose-response functions is estimated on 36 observations for each of the nine outcomes. This allows to recover efficiency measures and decompose them between the extensive and intensive externalities (with $e_a^N \equiv \beta_a^N / \beta_a^{\tilde{C}}$ and $e_a^V \equiv \beta_a^V / \beta_a^{\tilde{C}}$):

$$e_a = (\phi \cdot \beta_a^N + \psi \cdot \beta_a^V) / \beta_a^{\tilde{C}} = \phi \cdot e_a^N + \psi \cdot e_a^V \quad (2.7)$$

2.4 Empirical Methods

2.4.1 Identifying Assumptions

Facing the endogeneity of the ABC zoning due to the official criteria for defining housing markets tenseness, we maintain two assumptions to recover causal treatment effects. The first is that, conditionally on pre-treatment variables, treatments are weakly unconfounded.

Assumption 1 *Weak Unconfoundedness.*

$$\forall (g, \mathbf{X}), Y_g \perp T \mid \mathbf{X}$$

According to this assumption, the set of pre-treatment variables \mathbf{X} ensures a conditional randomisation of the IFL policy between municipalities. This selection-on-observables restriction considers that all the structural differences between municipalities are controlled by pre-treatment variables, and that the differences between the conditional outcomes can only be attributed to policy changes. As g describes both spatial and time variations, we use this assumption both between areas of the ABC zoning and between policy periods.

The well-known property of dimension reduction of well-specified propensity scores (Hahn, 1998) allows to parsimoniously model the conditional expectation of the outcomes, as long as we have $Y_g \perp T \mid p_g(\mathbf{X})$ with $p_g(\mathbf{X}) \equiv \mathbb{P}(T = g \mid \mathbf{X})$ from Assumption 1. This is the definition of the Generalised Propensity Score (GPS, Imbens, 2000) as the probability to receive a level of treatment knowing the pre-treatment variables. As Crump et al. (2009) show, the propensity to receive a treatment should not be too close to zero or one to ensure precise and robust estimates. This leads to the following overlap assumption, particularly important in the case of multi-valued treatments as in the IFL policy:

Assumption 2 *Overlap*

$$\forall (g, \mathbf{X}), p_g(\mathbf{X}) > 0$$

Under the two previous assumptions, Słoczyński and Wooldridge (2018, Lemma 3.2) demonstrate that counterfactual treatment effects can be identified from usual data. The average outcome $Y_{g'}$ for a counterfactual treatment level g' , respectively for the whole

population and for municipalities that actually receive the treatment level g , are respectively:

$$\mathbb{E}(Y_{g'}) = \mathbb{E} \left[\frac{T_{g'}}{p_{g'}(\mathbf{X})} Y \right] \quad \text{and} \quad \mathbb{E}(Y_{g'} | T = g) = \frac{1}{\mathbb{P}(T = g)} \cdot \mathbb{E} \left[\frac{p_g(\mathbf{X})}{p_{g'}(\mathbf{X})} T_{g'} Y \right]. \quad (2.8)$$

These statistics concern respectively the full population of homeowners impacted by the externalities at both margins and the recipients targeted by the policy support. They are the building blocks of the policy-relevant treatment effects under consideration, as the ATE of g instead of g' on the outcome Y is $\mathbb{E}(Y_g) - \mathbb{E}(Y_{g'})$ and the related ATT is $\mathbb{E}(Y_g | T = g) - \mathbb{E}(Y_{g'} | T = g)$. These counterfactual statistics are used to build policy-relevant treatment effects as they are related to different populations.

2.4.2 Specification of the Propensity Scores

In accordance with the concept of housing market tenseness used to set the ABC zoning, we define an unobserved latent variable η_i^* crossing thresholds to determine the classification of municipalities. The propensity for a municipality i to be high in the hierarchy depends on the J pre-treatment variables x_{ji} used to proxy the political decision, a bivariate smooth function of the geographical coordinates of its centroid \mathbf{z}_i (longitude and latitude, Gilbert et al., 2023), and a random term ε_i representing the arbitrary part of the zoning. This latter term is assumed to be logistically distributed to produce an ordered logit model. The latent variable describing housing market tightness η_i^* is then:

$$\eta_i^* = \alpha + \sum_{j=1}^J f_j(x_{ji}) + h(\mathbf{z}_i) + \varepsilon_i. \quad (2.9)$$

The J univariate functions f_j are specified as additive spline transformations of pre-treatment variables, in accordance with the generalised additive model framework (GAM, Wood, 2017). The spline coefficients are shrunk endogenously by penalised iterated weighted least squares while the smoothing parameters are estimated using a separate criterion from the restricted maximum likelihood (Wood et al., 2016). The same estimation procedure is used simultaneously for the bivariate smooth function h of coordinates, the main difference is the *a priori* specification of the spline that is bivariate thin plate.

By noting Λ the cumulative function of the logistic distribution and $\mu_0 < \mu_1 < \dots < \mu_5$ the unknown ordered thresholds related to the four ABC zones, the GPS for the IFL policy are (with $\eta_i \equiv \eta_i^* - \varepsilon_i$ the deterministic parts of the latent variable):

$$p_g(\eta_i) = \Lambda(\mu_g - \eta_i) - \Lambda(\mu_{g-1} - \eta_i). \quad (2.10)$$

Because municipalities designed as A are more tense than others (B_1 , B_2 , C) and because η_i is a measure of tenseness, values of the latent variable lie between the thresholds μ_4 and μ_5 . As the ABC zoning does not change in the 2015–2019 period under study, the probability of being in a given zone is constant over time. Then, an appealing property of the ordered structure of the ABC zoning is that, if the GPS is well specified, conditioning on the deterministic part of the latent variable η_i is sufficient to reach weak unconfoundedness (instead of the full set of pre-treatment variables \mathbf{X}). To prevent from GPS misspecification, we favour a doubly robust estimation relying on a specification of the outcomes. In this case, the estimation is consistent if at least one specification is well-specified (Robins and Rotnitzky, 1995; Słoczyński and Wooldridge, 2018).

2.4.3 Specification of the Outcomes

The outcomes are specified using the same semi-parametric GAM framework. The main difference is that each outcome Y is modelled separately for each subsample defined from the treatments g received by the municipalities. The smooth functions f_j and h are now indexed by the outcome y and the treatment g such that:

$$y_{gi} = \alpha_g^y + \sum_{j=1}^J f_{gj}^y(x_{ji}) + h_g^y(\mathbf{z}_i) + \varepsilon_{gi}^y. \quad (2.11)$$

The same pre-treatment variables and geographical coordinates are used, with different smoothing parameters shrunk during the estimation procedure. As we have nine outcomes, four treatment levels and three periods, Equation 2.9 corresponds to 108 GAM estimations in order to estimate the full set of functions f_{gj}^y and h_g^y for a given GPS. From the quasi-loglikelihood arguments of Słoczyński and Wooldridge (2018), the double robustness property requires that outcome regressions are weighted according to GPS ratios as in Equation 2.8. To recover the average counterfactual outcome for the treatment g' for the municipalities actually receiving g , each municipality is weighted by $\hat{p}_g(\eta_i)/\hat{p}_{g'}(\eta_i)$ predicted from the first stage. As generally advised in the literature, we use normalised weights by dividing them by their sum within each treatment subsample.

We close this section with the formulas we use to assess the efficiency of the IFL policy. The counterfactual building blocks of Equation 2.8 are recovered from the regression of the outcome Y on the sub-sample of municipalities with the treatment g' using respectively $1/p_{g'}(\eta_i)$ and $p_g(\eta_i)/p_{g'}(\eta_i)$ as weights. Under assumptions 1 and 2, noting $\mu_g \equiv \mathbb{P}(T =$

g) the share of municipalities that receive the treatment g , averaging the fitted values provides a consistent estimation as:

$$\mathbb{E}(Y_{g'}) = N^{-1} \times \sum_{\ell=1}^N \hat{y}_{\ell}(g') \quad \text{and} \quad \mathbb{E}(Y_{g'} | T = g) = \mu_g^{-1} \times \sum_{\ell=1}^N T_{g\ell} \times \hat{y}_{\ell}(g') \quad (2.12)$$

where $\hat{y}_{\ell}(g') \equiv \hat{\alpha}_{g'}^Y + \sum_{j=1}^J \hat{f}_{g'j}^Y(x_{j\ell}) + \hat{h}_{g'}^Y(z_{\ell})$ comes from the estimation of the outcome Y for the subset of municipalities that indeed receive treatment g' . It is simply the predicted outcome values for the whole population of municipalities with $\ell = 1, \dots, N$.

2.5 Results

We first present the estimation results for the first-stage models, followed by second-stage models, the estimation of policy-relevant treatment effects and dose-response functions, allowing the measures of cost-effectiveness of the French IFL policy. We close our results section with a placebo analysis.

2.5.1 First-stage Models from ABC Zoning

We estimate a semi-parametric ordered logit GAM on the ABC classification of municipalities, given a set of pre-treatment variables on housing supply and demand. To reach our identification restriction of unconfoundedness, we include a maximum of variables that can be used by the French administration to construct this classification, including the pre-treatment unitary house prices of the municipality and the neighbouring ones. We include the geographical coordinates of the centroids of each municipality through bivariate smoothing splines to control for spatial confounders (Gilbert et al., 2023). As allowed by the GAM framework, all variables enter semi-parametrically with a degree of smoothing that is endogenously shrunk by the penalised estimation procedure. Table 2.5.1 provides the joint significance of the spline transformations of each variable according to different specifications and maximum degree of spatial smoothing.

The results present both high pseudo- R^2 and share for good predictions in the bottom panel in rows. The specification with the lowest spatial smoothing (that allows fine-scale spatial heterogeneity to be taken into account, reported in the last column) yields 90.1% of correct predictions of the ABC classification for municipalities. This indicates the relevance of the ordered framework for modelling the ABC classification and increases the likelihood of having a well-specified GPS. Although models with higher maximum degrees of freedom allowed for spatial coordinates yield better predictions (91.2% for

Table 2.5.1: Covariates' joint significance from first-stage ordered GAMs

Max. degrees of freedom	<i>Outcome: Ordered ABC Zoning</i>					
	No Spatial Smoothing		With Spatial Smoothing			
	df = 0	df = 0	df = 50	df = 50	df = 100	df = 200
Population Density	1,991.3*** [6.1]	1,723.1*** [5.7]	2,003.3*** [5.8]	1,656.4*** [5.8]	1,688.5*** [6.0]	1,479.1*** [6.0]
New Housing Unit	468.7*** [6.0]	99.0*** [5.3]	295.2*** [5.3]	126.1*** [5.0]	127.4*** [4.9]	141.2*** [4.8]
Median Annual Income	1,647.6*** [6.7]	353.5*** [6.6]	654.7*** [6.7]	208.4*** [6.2]	200.5*** [6.1]	182.7*** [6.0]
Professional Occupations	984.1*** [37.0]	819.4*** [28.4]	312.9*** [30.6]	317.3*** [32.6]	273.8*** [25.1]	267.8*** [26.8]
Unitary Housing Price		214.8*** [6.6]		70.3*** [5.5]	67.7*** [5.2]	51.3*** [1.0]
Neighboring Unitary Price		110.4*** [1.1]		37.1*** [3.1]	26.6*** [1.0]	23.1*** [4.2]
Spatial Coordinates			4,018.4*** [47.9]	2,211.0*** [47.3]	2,575.9*** [90.2]	3,048.8*** [165.2]
Number of Observations	26,818	26,818	26,818	26,818	26,818	26,818
McFadden R2	52.60	61.31	67.08	69.17	70.98	73.81
Percent of Good Predictions	85.88	87.31	88.94	89.29	89.70	90.13
Akaike Information Criterion	18,625.4	16,406.0	14,602.8	14,152.3	13,736.9	13,156.0

Notes: The top panel reports χ^2 statistics of joint significance for each covariate of the first-stage GPS. Professional Occupations are coded as population shares of eight categories according to the one-digit French *Catégories Socio-Professionnelles*. The effective degrees of freedom reported in brackets indicate the smoothing intensity, low values correspond to more smoothing. The unit of observation is the French municipality, columns reports different specifications with different covariates and different maximum spatial smoothing. Estimations come from the `gam` function of the `mgcv` R package (Wood et al., 2016). *Sources:* Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

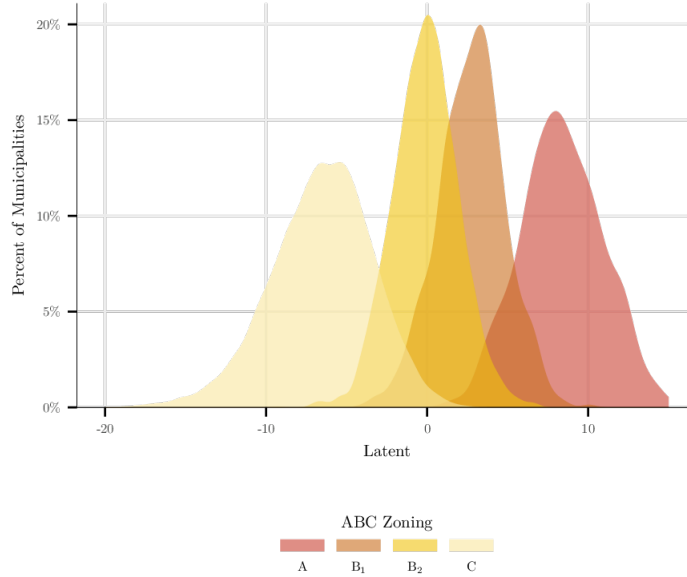
df=400), their computational cost and associated risk of overfitting leads us to not report them in the paper. Nevertheless, our main results regarding the effects of IFL policies are robust to the specification of the maximum degrees of freedom for spatial coordinates, albeit our preferred specification provide the most accurate estimation. In particular, the introduction of spatial coordinates affects the joint significance associated with the unitary housing price being consistent with the local characteristics of the housing market. The joint significance is the highest of the pre-treatment variables, confirming our expectation about the presence of unobservable spatial variables. In our preferred specification (df=200), the contribution of the unitary housing price is linear and increasing, consistent with the ABC perimeter definition.

Although prediction errors are limited, most of the remaining errors (77.2%) concern municipalities that are commonly used as the basis for natural experiments for housing policy assessment (Chapelle et al., 2018; Bono and Trannoy, 2019). The underlying assumption is that variations in treatment derived from the ABC classification are as

good as random for municipalities close to the area boundary. Furthermore, 29.2% of the prediction errors concern municipalities that experience a change in treatment level under the 2014 reform. Our prediction is consistent with the previous classification for 72.7% of these errors. Therefore, our errors mainly concern observations that are considered to be quasi-randomly assigned in the ABC classification or that have experienced a recent change in classification, which strengthens the credibility of the GPS estimate.

Since overlap is crucial to recover consistent effects and likely to be reduced for high-dimension model variables (D’Amour et al., 2021), we compare the distribution of the latent variable underlying the classification process (Figure 2.5.1). Latent distributions follow the ordered structure of the ABC classification, as consecutive treatment levels have greater common support than non-consecutive ones. Although the overlap is reduced, there is still common support for the extreme levels. This is probably due to the spatial proximity of some A- and C-tier municipalities. However, although the treatment assignment is based on the characteristics of the municipality, it still contains some arbitrariness, which we exploit.

Figure 2.5.1: Overlap between predictions of tenseness between the different ABC zones



Notes: The distributions of the latent tenseness variable (x-axis) are predicted from the first stage GPS with a maximum degree of freedom sets to 200 (6th column of Table 2.5.1). As a latent variable, $\hat{\eta}^*$ is unit-less and is displayed between municipalities according to the ABC classification. We report the distribution *within* each classification level (rather than the distribution of the entire population) for clarity reasons.

Sources: Authors’ Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

2.5.2 Second-stage Models for the Outcomes

Beyond the GPS specification, we assess the relevance of our control variables in the outcome specification using pooled models for our doubly robust estimator. We do not weight observations according to their probability derived from the GPS as this is only meaningful in the binary approach (Słoczyński and Wooldridge, 2018). We report in Section 2.D.1 the joint significance of each covariate in addition to the effective degree of freedom. We also report in Section 2.D.2 the marginal contribution of spatial coordinates in each outcome specification.⁶

Table 2.5.2: Covariates' joint significance from second-stage pooled GAMs

	<i>Outcome variables from...</i>								
	Tax	Transaction Data			IFL Files				
	N	V	S	P	\tilde{N}	\tilde{V}	\tilde{S}	\tilde{P}	\tilde{C}
Population Density	517.1*** [8.7]	9.6*** [6.6]	102.7*** [8.6]	50.8*** [7.8]	51.9*** [8.7]	17.6*** [6.8]	88.0*** [7.7]	58.9*** [7.9]	32.2*** [4.2]
Number of New Housing	2,380*** [8.2]	38.7*** [5.6]	93.3*** [4.6]	120.1*** [5.6]	1,128*** [7.1]	8.8*** [4.3]	11.5*** [4.3]	10.6*** [4.1]	3.0** [4.1]
Median Income	53.4*** [8.4]	68.4*** [7.0]	198.2*** [6.8]	3.3** [3.0]	16.9*** [6.0]	105.3*** [6.5]	27.3*** [5.7]	22.3*** [8.5]	1.5** [1.8]
Professional Occupations	771.3*** [50.8]	55.8** [34.1]	21.3** [49.2]	48.6** [38.7]	357*** [41.4]	6.6** [38.3]	6.6** [32.9]	12.1** [47.0]	34.0** [20.4]
Lagged Unitary Price	9.7*** [8.2]	21.7*** [4.5]	8.3*** [8.2]	31.9*** [3.5]	9.7*** [9.0]	4.7*** [3.1]	2.8*** [6.4]	3.8*** [8.6]	4.5*** [7.6]
Lag. Neighbor. Unit. Price	6.7*** [8.3]	96.7*** [7.5]	11.7*** [7.2]	83.7*** [8.1]	9.1*** [8.8]	10.1*** [8.7]	18.3*** [7.3]	32.8*** [7.6]	7.1*** [3.6]
Spatial Coordinates	33.7*** [189]	16.6*** [182]	37.6*** [188]	8.3*** [179]	22.0*** [186]	21.8*** [181]	8.5*** [168]	20.5*** [193]	7.0*** [112]
Number of observations	54,993	54,993	54,993	54,993	54,993	54,993	54,993	54,993	54,991
McFadden R2	77.72	56.12	36.53	56.99	54.52	45.27	18.63	55.81	9.16

Notes: For the nine outcomes of interest (in columns), the table reports the F statistics for the joint significance of each covariates (in rows). N accounts for the number of new homeowners, V for housing value, S for surface, and P for unitary housing price. The variables with a \sim are the same variables computed for IFL recipients, \tilde{C} is the IFL cost. We report pooled GAMs on all treatment levels for the sake of clarity, different GAMs are estimated for each treatment level in the policy-relevant treatment effects reported in the text. Professional Occupations are coded as population shares of eight categories according to the one-digit French *Catégories Socio-Professionnelles*. The effective degrees of freedom reported in brackets indicate the smoothing intensity, low values correspond to more smoothing. The unit of observation is the French municipality and the maximum degree of freedom we allow for the spatial coordinates is 200.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

Pre-treatment variables introduced as regression adjustment in the pooled models explain more than 74% of the observed variance in the number of first-time owners. The development of housing supply as measured by the number of new housing is highly significant in explaining both the number of first-time owners and the number of beneficiaries.

⁶For clarity reason, we do not report marginal contribution of other covariates.

In addition, the local housing market price and median income are significant in explaining the number of transitions to homeownership, highlighting the importance of affordability for transitions. It supports our approach to the introduction of pretreatment variables in the regression adjustment to control for potential heterogeneity.

2.5.3 Treatment Effects and Dose-response Functions

From the second step, we estimate combinations of bilateral effects (g, g') , following our strategy described in Section 2.3.3. We report these bilateral estimations and standard errors, using a bootstrap approach with 500 iterations, in Section 2.D. They constitute our building blocks to compute the causal effect of each primitive for both extensive and intensive margin outcomes.

We estimate each β_a^Y from Equation 2.6 by regressing the bilateral combinations of treatments level on differences in primitive values between treatment level g and g' . Since primitive variations are small, our dose-response functional form is linear. In addition, our weighting scheme to estimate the relevant dose-response parameters depends on the nature of the estimands. We weight observations according to the number of municipalities which currently received treatment level g for ATT estimands. We do not introduce weights for the ATE as it concerns the entire population, unlike ATT. We report our results from linear dose-response specifications in Table 2.5.3 estimated by WLS (for ATT) and OLS (for ATE). Standard errors are estimated using the 500-iterations bootstrap procedure.

Despite significant effect on policy cost (column 9, Table 2.5.3), primitives have no sizeable effect on the number of homeowners. Meanwhile, raising the ceiling value or interest rates increases the number of beneficiaries. Considering jointly the effect on first-time owners and beneficiaries, the results indicate that increasing IFL subsidies trigger opportunistic behaviour rather than favour homeownership. However, because our identification method enables to recover causal effect at a municipality level, we cannot reject that the policy indeed affects tenure decision at an individual level.

Besides a weak effect at the extensive margin, an increase in the subsidy causes significant effect at the intensive margin, except for the loan maturity. Then, a 1 point increase in the coverage share results in a 2.0% increase in the housing market price. Considering the inelastic feature of the housing supply on the short-term, these results are consistent with expected results for boosting demand. Although the distortion of housing choices resulting from the variation in the subsidy has no first-order effect on welfare (recall that we assume optimal consumption for recipients), the additional cost of raising primitive sources has a positive effect on the housing characteristics purchased, capitalised through price. Finally, the number of beneficiaries covaries with the parameters defined by pol-

Table 2.5.3: OLS coefficients for policy primitives from dose-response functions

	<i>Outcome variables from...</i>								
	Tax	Transaction Data				IFL Files			
	N	V	S	P	\tilde{N}	\tilde{V}	\tilde{S}	\tilde{P}	\tilde{C}
Covering Share	-0.005 (0.008)	0.020*** (0.003)	0.007*** (0.001)	-0.006 (0.004)	0.002 (0.001)	-0.002*** (0.000)	-0.006*** (0.000)	0.004*** (0.000)	0.030*** (0.000)
Ceiling Value	0.000 (0.004)	0.004** (0.002)	-0.005*** (0.001)	0.007*** (0.002)	0.021*** (0.001)	-0.006*** (0.000)	-0.010*** (0.000)	0.006*** (0.000)	0.003*** (0.000)
Interest Rate	-0.004 (0.118)	0.087** (0.041)	-0.068*** (0.021)	-0.059 (0.045)	0.374*** (0.036)	-0.193*** (0.006)	-0.236*** (0.005)	0.077*** (0.006)	0.139*** (0.008)
Loan Maturity	0.002 (0.003)	-0.001 (0.001)	0.001** (0.001)	-0.002 (0.001)	-0.006*** (0.001)	0.004*** (0.000)	0.004*** (0.000)	-0.001*** (0.000)	0.022*** (0.000)
Constant	-0.000 (0.000)	0.000* (0.000)	0.000 (0.000)	0.000 (0.000)	-0.146*** (0.010)	0.015*** (0.002)	0.020*** (0.002)	0.003 (0.002)	0.017*** (0.003)
Number of obs.	36	36	36	36	36	36	36	36	36
R2	0.021	0.656	0.503	0.704	0.118	0.379	0.404	0.760	0.977
Adj. R2	-0.106	0.611	0.439	0.666	0.004	0.299	0.327	0.729	0.974

Notes: For the nine outcomes Y in columns, the table reports the β_a^Y coefficients associated to each primitive in rows. They are estimated from dose-response functions of Equation 2.6. N accounts for the number of new homeowners, V for housing value, S for surface, and P for unitary housing price. The variables with a \sim are the same variables computed for IFL recipients, \tilde{C} is the IFL cost. The unit of observation is the bilateral combination of four ABC zones for the three periods of interest, the full set of policy-relevant treatment effects is reported in the Table 2.D.3 of OA. Standards errors in parentheses are estimated by 500-iterations bootstraps accounting for the uncertainty of treatment effects. ATEs for tax and transaction data are weighted according to the inverse of their bootstrapped standard errors, ATT for IFL variables are additionally weighted according to the number of municipalities receiving the considered treatment levels.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

icy makers. Increasing the coverage share or the ceiling value makes the subsidy more attractive to potential homeowners. However, as it does not affect the total number of homeowners, we conclude that it mainly causes a shift from unsubsidised to subsidised units.

2.5.4 Measures of Cost-effectiveness

From the linear dose-response parameters (Table 2.5.3), we assess policy efficiency following Equation 2.7. For each primitive, we compute the ratio between the $\widehat{\beta}_a^Y$ related to both positive and negative externalities, and the margin effect on policy cost based on the variation of both cost per beneficiary $\widehat{\beta}_a^C$ and number $\widehat{\beta}_a^{\tilde{N}}$. This approach allows us to compare the welfare effect of an additional euro invested in each primitive. Recall that we assume that recipients optimally perform their housing choices, leading to no first order response from recipients' behavioural changes. We compute standard errors using the 500-iteration bootstrap procedure.

As the definition of social welfare depends on the marginal value of externalities at

both margins (namely ϕ and ψ) being unobservable, we set their value exogenously. Recall that we introduce variation in the marginal value for externalities given that the balance between desirable and non-desirable feature of externalities is not clear-cut. Nonetheless, our preferred specification sets the marginal value of the externalities at the extensive margin to 15k, while the value for intensive margin is $-5k$. Our choice is motivated by the fact that externalities for one homeowner is estimated to 1.3k per year (Coulson and Li, 2013). On the opposite, our decision to value intensive margin negatively results from the negative implications of affordability (Gabriel and Painter, 2020).

In our preferred specification, increasing the cost of policy through primitive parameters defined by policy markers has a negative impact on social welfare. We estimate that investing one additional euro through the coverage share (respectively the ceiling value) leads to a loss of 2.74 (respectively 0.51) euros for society. It results mainly from the fact that the externalities at the intensive margin, which are evaluated negatively in our preferred specification are sizeable with weak effect at the extensive margin. Differences in intensity between the two primitives controlled by policy makers could result from the subpopulation affected. Indeed, an increase in the coverage share would benefit the whole population of beneficiaries, whereas an increase in the ceiling would specifically affect the most expensive operations. Therefore, the leverage effect on demand is likely to be higher for the coverage share, leading to greater price capitalisation. Moreover, we cannot reject that the policy efficiency does not depend on credit market behaviour. Neither an increase in the interest rate nor in the loan maturity has a significant impact on social welfare.

While we consider cost-effectiveness from a welfare perspective, policy makers may choose to target specific populations, such as homeowners, that deviate from social welfare considerations. Assuming that policy makers target the welfare of this population, the expected marginal value of intensive externalities is likely to be positive (we set $\psi = 5.0k$ holding ϕ constant). On the one hand, the price capitalisation caused by the subsidy would benefit to current homeowners by raising their potential for capital gains. On the other hand, current homeowners are at least insensitive or even favourable to additional homeowners, in line with the impact on local decisions (Ortalo-Magné and Prat, 2014). As a result, policy efficiency becomes positive, with more benefits relative to public expenditure. Investing one additional euro by raising coverage share induces a 2.67 euros of benefit, while the ceiling value is a less efficient channel for policy makers (0.51 euros of benefit). As for the whole society, the efficiency is mainly driven by externalities at the intensive margin.

Table 2.5.4: Measures of cost-effectiveness for different marginal values of externalities

	$\phi = -10.0k$				$\phi = 15.0k$				$\phi = 30.0k$			
	-10.0k	-5.0k	0.0k	5.0k	-10.0k	-5.0k	0.0k	5.0k	-10.0k	-5.0k	0.0k	5.0k
Covering Share	-5.393*** (0.87)	-2.685*** (0.44)	0.022 (0.04)	2.730*** (0.44)	-5.449*** (0.87)	-2.741*** (0.44)	-0.034 (0.06)	2.674*** (0.44)	-5.482*** (0.88)	-2.775*** (0.45)	-0.067 (0.12)	2.641*** (0.45)
Ceiling Value	-1.012*** (0.42)	-0.506*** (0.21)	-0.000 (0.02)	0.505*** (0.21)	-1.011*** (0.42)	-0.505*** (0.21)	0.001 (0.03)	0.506*** (0.21)	-1.010*** (0.43)	-0.504*** (0.22)	0.001 (0.06)	0.507*** (0.22)
Interest Rate	-15.774 (11.29)	-7.882 (5.67)	0.011 (0.54)	7.904 (5.67)	-15.803 (11.31)	-7.910 (5.70)	-0.017 (0.81)	7.876 (5.70)	-15.820 (11.40)	-7.927 (5.87)	-0.034 (1.63)	7.859 (5.87)
Loan Maturity	0.159 (0.32)	0.075 (0.16)	-0.008 (0.01)	-0.092 (0.16)	0.180 (0.32)	0.096 (0.16)	0.013 (0.02)	-0.071 (0.16)	0.193 (0.33)	0.109 (0.17)	0.025 (0.04)	-0.058 (0.17)

Notes: Exploiting coefficients derived from the dose-response function (Table 2.5.3), we calculate efficiency measure defined in Equation (2.7) for the four sources of primitives. It represents the economic returns for the society according to one euro invested. Negative value indicates that increasing the primitive source is counterproductive, while efficiency measure above 1 indicates higher benefit than investment. As our efficiency measure depends on the marginal valuation of externalities at the extensive (ϕ) and the intensive (ψ) margins, we simulate different scenarios. Although extensive margin externalities are likely to be positive unlike intensive margin ones, we also assess scenarios with opposite sign according to our expectations. We report between parentheses the standard errors using 500-iterations bootstrap procedure.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

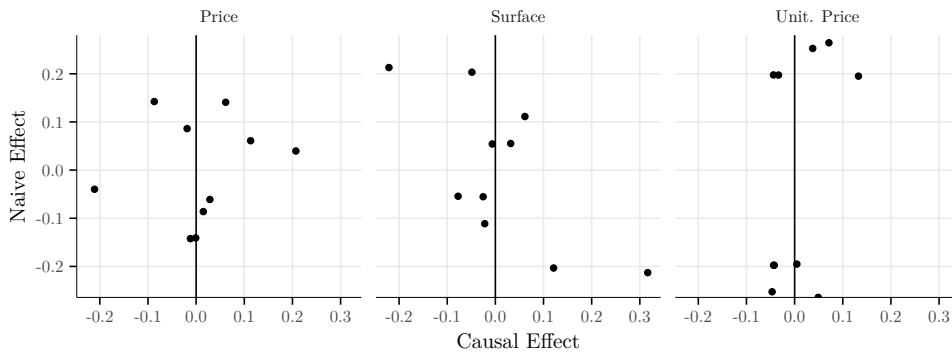
*** p < 0.01, ** p < 0.05 * p < 0.1

2.5.5 Placebo Analysis for Unconfoundedness

Our placebo analysis relies on a specific feature of the IFL design. As shown in Equation 2.1, the IFL amount is characterised by the two policy parameters, the coverage share and the ceiling value being spatially heterogeneous in line with the ABC perimeter. We take advantage of the fact that differences in IFL amounts concern transactions with a price above the cap, when the IFL share does not vary between two levels (we report additional explanations on Section 2.E). Hence, observations with purchase price under the lower ceiling value for both levels of treatment in period with similar coverage share, benefit from the same IFL amount.

Our procedure is as follows. For IFL aggregated outcomes restricted to observations with no difference of IFL amount, we first estimate cross-sectional differences for treatment level using naive regressions corresponding to the difference in average, without correcting for ABC perimeter under the assumption that the treatment assignment is random. Then, we estimate treatment effect using the estimation procedure we adopt to deal with endogeneity of treatment. We present bilateral effects according to whether it is the naive estimator or the doubly robust one. As we can select observations not submitted to difference in treatment for IFL outcomes, we restrict our placebo analysis at the intensive margin related to IFL beneficiaries housing choices. We report in Figure 2.5.2 bivariate graphs for policy-relevant treatment effects to compare magnitude between the naive and the doubly robust estimators.

Figure 2.5.2: Naive and policy-relevant treatment effects used in placebo analysis



Notes: We report the 36 bilateral combinations of the IFL effects on outcomes restricted to beneficiaries for observations with no difference in treatment intensity. In Y-axis, we report the naive effect, *i.e.* without weighting scheme according to treatment intensity and regression adjustment. In the X-axis, we report the doubly robust estimator using the GPS specification and the regression adjustment. Our choice to restrict placebo analysis to the IFL outcomes is driven by the possibility to select precisely observations with no difference in treatment (see Section 2.E).

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

The placebo analysis confirms the validity of our two-step procedure. Indeed, while naive estimated effects are sizeable and significant, our policy relevant treatment effects estimated on the population with similar treatment intensity are not significant for unitary housing price outcomes. It includes policy relevant treatment with restricted overlap, namely the A-C comparison. However, some significant differences still remain for surface and overall purchase price.

2.6 Conclusion

The French IFL policy aims to induce additional homeowners through interest cuts. We leverage spatial variation of treatment using selection-on-observables restriction to assess the cost-efficiency of the policy derived from the sufficient statistics literature. Note this framework could also be useful to compare homeownership support to other public spendings in a broader approach. Our GPS specification and regression adjustment involve, among other variables, spatial coordinates to prevent for omitted variables. Assuming optimal consumption for recipients, we derive welfare implication for small variation of the primitive sources that defines the IFL subsidy by assessing externalities at the extensive (derived from additional homeowners) and the intensive margins (price capitalisation).

We cannot reject the possibility that increasing policy expenditure on the IFL has no effect on the number of homeowners, especially at the individual level, according to our identification strategy. However, we precisely estimate that externalities at the intensive margin exceed potential ones at the extensive margin. It turns out that the cost-effectiveness of the IFL depends mainly on the marginal value of externalities at the intensive margin, which is directly related to inflationary effects. Given the growing concerns about affordability, we expect these externalities to be detrimental to society as a whole, leading to negative efficiency measures.

However, if policy makers' objectives deviate from welfare considerations, increasing the subsidy may be efficient. The cost-efficiency of the policy, beyond the potential debate about the valuation of externalities, depends on the population that policy makers are trying to benefit. For example, if the targeted population consists of current homeowners, increasing the coverage share will generate greater benefits than public expenditure. Then, given that major reforms of the IFL are transitory and coincide with runs for the presidency (namely the 2010 and 2016 reforms), we cannot reject that the election goal (rather than the entire welfare) has been foremost sought by policy makers. Alternative considerations such as generating additional property tax revenue on the long term are unlikely, as newly built housing are exempted for at least a two-year period and tax

property is now perceived by municipalities.

Our paper leaves open questions for further research about externalities induced by supporting homeownership. As housing market capitalisation is related to the housing supply and land availability, externalities produced by interest cuts are likely to depend on local characteristics. Assessing the nature of such heterogeneity is relevant to discuss more precisely welfare implications and spatial equity. Finally, as supporting homeownership affects beneficiaries housing choices, it raises concerns about the impact of interest cuts on land consumption. Since French administration aims to reduce land consumption through higher constraints (the net-zero-artificialisation implementation), policy contribution is of concern.

References

- Andrews, Dan and Aida Caldera Sánchez (2011), “The Evolution of Homeownership Rates in Selected OECD Countries: Demographic and Public Policy Influences”. *OECD Journal: Economic Studies* 2011.1, pp. 1–37.
- Angrist, Joshua D and Alan B Krueger (2001), “Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments”. *Journal of Economic Perspectives* 15.4, pp. 69–85.
- Bäckman, Claes and Chandler Lutz (2020), “The Impact of Interest-Only Loans on Affordability”. *Regional Science and Urban Economics* 80.103376.
- Beaubrun-Diant, Kevin and Tristan-Pierre Maury (2021), “Implications of Homeownership Policies on Land Prices: The Case of a French Experiment”. *Economics Bulletin* 41.3, pp. 1256–1265.
- Boehm, Thomas P. and Alan M. Schlottmann (1999), “Does Home Ownership by Parents Have an Economic Impact on Their Children?” *Journal of Housing Economics* 8, pp. 217–232.
- Bono, Pierre-Henri and Alain Trannoy (2019), “The Impact of the ‘Scellier’ Income Tax Relief on Building Land Prices in France”. *Economie et Statistique / Economics and Statistics* 507d, pp. 91–114.
- Carozzi, Felipe (2020), “Credit Constraints and the Composition of Housing Sales. Farewell to First-Time Buyers?” *Journal of the European Economic Association* 18.3, pp. 1196–1237.
- Carozzi, Felipe, Christian A. L. Hilber, and Xiaolun Yu (2020), “On the Economic Impacts of Mortgage Credit Expansion Policies: Evidence from Help to Buy”. *CEPR Discussion Paper* DP14620.
- Chapelle, Guillaume, Benjamin Vignolles, and Clara Wolf (2018), “Impact of a Housing Tax Credit on Local Housing Markets: Evidence from France”. *Annals of Economics and Statistics* 130, pp. 101–131.
- Chareyron, Sylvain, Tidiane Ly, and Yohann Trouvé-Sargison (2021), “Ownership Incentives and Housing Affordability: Evidence from France”. *SSRN Electronic Journal*.
- Coulson, N. Edward and Herman Li (2013), “Measuring the External Benefits of Homeownership”. *Journal of Urban Economics* 77, pp. 57–67.
- Cour des Comptes (2012), *Rapport Public Annuel 2012, Tome I*. Tech. rep. Cour des Comptes, p. 1033.
- Crump, Richard K., V. Joseph Hotz, Guido W. Imbens, and Oscar A. Mitnik (2009), “Dealing with Limited Overlap in Estimation of Average Treatment Effects”. *Biometrika* 96.1, pp. 187–199.

- Cunningham, Chris and Robert R. Reed (2013), “Negative Equity and Wages”. *Regional Science and Urban Economics* 43.6, pp. 841–849.
- D’Amour, Alexander, Peng Ding, Avi Feller, Lihua Lei, and Jasjeet Sekhon (2021), “Overlap in Observational Studies with High-Dimensional Covariates”. *Journal of Econometrics* 221.2, pp. 644–654.
- Daminger, Alexander and Kristof Dascher (2022), “Homeowner Subsidy Repeal and Housing Recentralization”. *Land Economics*, 081821–0095R.
- Davis, Morris A., Stephen D. Oliner, Tobias J. Peter, and Edward J. Pinto (2020), “The Impact of Federal Housing Policy on Housing Demand and Homeownership: Evidence from a Quasi-Experiment”. *Journal of Housing Economics* 48.101670.
- de Chaisemartin, Clément and Xavier D’Haultfœuille (2020), “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects”. *American Economic Review* 110.9, pp. 2964–2996.
- Finkelstein, Amy and Nathaniel Hendren (2020), “Welfare Analysis Meets Causal Inference”. *Journal of Economic Perspectives* 34.4, pp. 146–167.
- Fisher, Jonas D. M. and Martin Gervais (2011), “Why Has Home Ownership Fallen Among the Young?” *International Economic Review* 52.3, pp. 883–912.
- Gabriel, Stuart and Gary Painter (2020), “Why Affordability Matters”. *Regional Science and Urban Economics* 80.103378.
- Gilbert, Brian, Abhirup Datta, Joan A. Casey, and Elizabeth L. Ogburn (2023), “A Causal Inference Framework for Spatial Confounding”. *arXiv:2112.14946*.
- Glaeser, Edward L. and Jesse M. Shapiro (2003), “The Benefits of the Home Mortgage Interest Deduction”. NBER Chapters. National Bureau of Economic Research, Inc, pp. 37–82.
- Green, Richard K. and Patric H. Hendershott (2001), “Home-Ownership and Unemployment in the US”. *Urban Studies* 38.9, pp. 1509–1520.
- Green, Richard K. and Kerry D. Vandell (1999), “Giving Households Credit: How Changes in the U.S. Tax Code Could Promote Homeownership”. *Regional Science and Urban Economics* 29.4, pp. 419–444.
- Green, Richard K. and Michelle J. White (1997), “Measuring the Benefits of Homeowning: Effects on Children”. *Journal of Urban Economics* 41.3, pp. 441–461.
- Gruber, Jonathan, Amalie Jensen, and Henrik Kleven (2021), “Do People Respond to the Mortgage Interest Deduction? Quasi-experimental Evidence from Denmark”. *American Economic Journal: Economic Policy* 13.2, pp. 273–303.
- Hahn, Jinyong (1998), “On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects”. *Econometrica* 66.2, pp. 315–331.

- Hahn, Jinyong, Petra Todd, and Wilbert Klaauw (2001), “Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design”. *Econometrica* 69.1, pp. 201–209.
- Hall, Andrew B. and Jesse Yoder (2022), “Does Homeownership Influence Political Behavior? Evidence from Administrative Data”. *The Journal of Politics* 84.1, pp. 351–366.
- Hanson, Andrew (2012), “Size of Home, Homeownership, and the Mortgage Interest Deduction”. *Journal of Housing Economics* 21.3, pp. 195–210.
- Harding, John, Thomas J. Miceli, and C.F. Sirmans (2000), “Do Owners Take Better Care of Their Housing Than Renters?” *Real Estate Economics* 28.4, pp. 663–681.
- Harkness, Joseph M and Sandra J Newman (2003), “Effects of Homeownership on Children: The Role of Neighborhood Characteristics and Family Income”. *Economic Policy Review* 9.2, pp. 87–107.
- Hembre, Erik (2018), “An Examination of the First-Time Homebuyer Tax Credit”. *Regional Science and Urban Economics* 73, pp. 196–216.
- Hendren, Nathaniel and Ben Sprung-Keyser (2020), “A Unified Welfare Analysis of Government Policies”. *Quarterly Journal of Economics* 135.3, pp. 1209–1318.
- Hilber, Christian A. L. and Tracy M. Turner (2014), “The Mortgage Interest Deduction and Its Impact on Homeownership Decisions”. *Review of Economics and Statistics* 96.4, pp. 618–637.
- Hoebbeck, Annelies and Koen Inghelbrecht (2017), “The Impact of the Mortgage Interest and Capital Deduction Scheme on the Belgian Mortgage Market”. *Working Paper Research: Bank of Belgium* 327.
- Imbens, Guido W. (2000), “The Role of the Propensity Score in Estimating Dose-Response Functions”. *Biometrika* 87.3, pp. 706–710.
- Kleven, Henrik J. (2021), “Sufficient Statistics Revisited”. *Annual Review of Economics* 13.1, pp. 515–538.
- Kunovac, Davor and Ivan Zilic (2022), “The Effect of Housing Loan Subsidies on Affordability: Evidence from Croatia”. *Journal of Housing Economics* 55.101808.
- Labonne, Claire and Cécile Welter-Nicol (2015), “Cheap Credit, Unaffordable Houses?” *Working Paper*.
- Lee, Hyojung, Dowell Myers, Gary Painter, Johanna Thunell, and Julie Zissimopoulos (2020), “The Role of Parental Financial Assistance in the Transition to Homeownership by Young Adults”. *Journal of Housing Economics* 47, p. 101597.
- Martin, Hal and Andrew Hanson (2016), “Metropolitan Area Home Prices and the Mortgage Interest Deduction: Estimates and Simulations from Policy Change”. *Regional Science and Urban Economics* 59, pp. 12–23.

- Ortalo-Magné, François and Andrea Prat (2014), “On the Political Economy of Urban Growth: Homeownership versus Affordability”. *American Economic Journal: Microeconomics* 6.1, pp. 154–181.
- Robins, James M. and Andrea Rotnitzky (1995), “Semiparametric Efficiency in Multivariate Regression Models with Missing Data”. *Journal of the American Statistical Association* 90.429, pp. 122–129.
- Rubin, Donald B. (1974), “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.” *Journal of Educational Psychology* 66.5, pp. 688–701.
- Słoczyński, Tymon and Jeffrey M. Wooldridge (2018), “A General Double Robustness Result for Estimating Average Treatment Effects”. *Econometric Theory* 34.1, pp. 112–133.
- Sommer, Kamila and Paul Sullivan (2018), “Implications of US Tax Policy for House Prices, Rents, and Homeownership”. *American Economic Review* 108.2, pp. 241–274.
- Sotura, Aurélie (2020), “Housing Policies, Spatial Inequalities and Housing Markets”. PhD thesis. Paris: EHESS.
- Wood, Simon N. (2017), *Generalized Additive Models: An Introduction with R*. Second. Chapman and Hall/CRC.
- Wood, Simon N., Natalya Pya, and Benjamin Säfken (2016), “Smoothing Parameter and Model Selection for General Smooth Models”. *Journal of the American Statistical Association* 111.516, pp. 1548–1563.

Online Appendix (not for publication)

2.A Additional Descriptive Statistics

2.A.1 IFL Summary

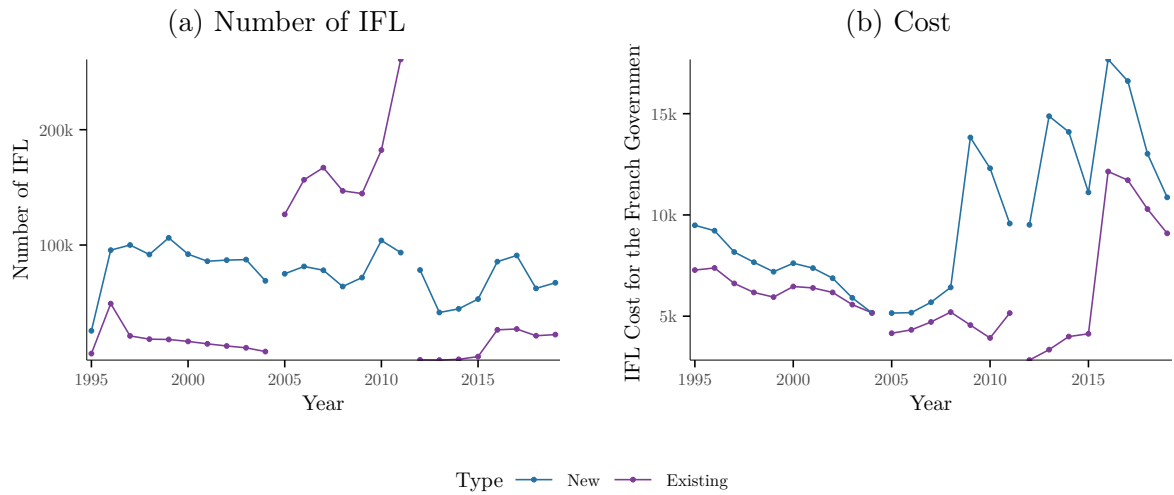


Figure 2.A.1: Number of IFL Policy and Average Subsidy

Notes: We report for each year the number of households who benefit from the IFL (Figure a) and the cost of the policy per household (Figure b). We distinguish both variables according to whether it concerns existing or newly built housing.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

2.A.2 Municipalities Pre-Treatment Characteristics

Table 2.A.1: Pre-Treatment Variables About Municipality Exploit for GPS Specification

	N	Mean	Std Dev	Median	Q1	Q3	Min	Max
Density	26,819	1.926	8.058	0.497	0.244	1.142	0.007	259.982
CS1	26,819	2.91	3.86	1.59	0.18	4.00	0.00	55.00
CS2	26,819	4.34	2.82	3.93	2.57	5.65	0.00	31.25
CS3	26,819	5.81	4.47	4.88	2.79	7.84	0.00	38.46
CS4	26,819	13.50	5.60	13.33	9.62	17.13	0.00	45.00
CS5	26,819	15.87	4.83	15.87	12.98	18.64	0.00	60.14
CS6	26,819	15.43	6.23	15.00	11.11	19.21	0.00	55.00
CS7	26,819	29.92	9.14	29.07	23.65	35.38	0.00	87.50
Price	26,819	153,108	68,002	141,975	110,323	181,032	20,518	2,261,166
Price per m^2 (2010–2013)	26,819	1,608.7	691.8	1,471.0	1,181.4	1,855.4	159.3	19,306.5
Neigh Price per m^2	26,819	1,561.0	724.0	1,420.2	1,123.6	1,807.3	0.0	35,686.1
New Housing (2010–2013)	26,819	42	227	9	4	25	1	15,748
Median Income (2013)	26,819	19,954	3,399	19,432	17,774	21,546	8,774	47,316
Longitude (WGS 84)	26,819	653,319	187,946	653,382	511,822	802,857	124,073	1,072,432
Latitude (WGS 84)	26,819	6,651,138	243,230	6,677,060	6,448,774	6,858,734	6,139,677	7,108,696

Notes: The average density of the municipality being exploited to estimate the GPS is 193 inhabitants per kilometre square. Our sample is composed of 26,819 observations. CS1 corresponds to share of socio-professional categories within the municipality. 1 corresponds to farmers, 2 to artisans and merchants, 3 to managers, 4 to intermediate professions, 5 to employees, 6 to labour works, 7 to retired.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

2.A.3 Descriptive Statistics about the Outcomes

Table 2.A.2: Descriptive Statistics about Outcomes Exploit for Econometric Model

	N	Mean	Std Dev	Median	Q1	Q3	Min	Max
A								
FTO	1,874	258.7	614.6	118.0	48.0	48.0	1.0	12,760.0
Price (Transaction)	1,874	443,572	426,877	320,763	261,133	261,133	107,760	5,999,507
Surface (Transaction)	1,874	80	17	78	68	68	39	177
Unit. Price (Transaction)	1,874	7,279	11,407	4,416	3,470	3,470	1,197	238,899
Beneficiaries	1,874	46.8	102.6	17.0	5.0	5.0	1.0	2,182.0
Price (IFL)	1,874	281,426	67,118	273,749	232,633	232,633	107,000	660,000
Surface (IFL)	1,874	81	25	75	62	62	22	271
Unit. Price (IFL)	1,874	3,662	1,096	3,407	3,003	3,003	641	10,897
Cost	1,874	19,068	5,295	18,568	15,694	15,694	2,264	40,806
B ₁								
FTO	3,295	122.4	348.7	48.0	24.0	24.0	1.0	8,268.0
Price (Transaction)	3,295	309,657	263,593	254,787	208,695	208,695	111,964	7,386,864
Surface (Transaction)	3,295	90	17	89	79	79	39	175
Unit. Price (Transaction)	3,295	4,174	5,056	3,095	2,536	2,536	1,442	119,469
Beneficiaries	3,295	18.7	38.4	8.0	3.0	3.0	1.0	848.0
Price (IFL)	3,295	238,583	53,601	232,025	202,959	202,959	68,441	620,610
Surface (IFL)	3,295	95	23	95	80	80	30	280
Unit. Price (IFL)	3,295	2,595	613	2,499	2,197	2,197	615	10,753
Cost	3,295	17,445	4,734	17,189	14,408	14,408	2,422	41,496
B ₂								
FTO	6,200	55.2	129.4	25.0	12.0	12.0	1.0	2,572.0
Price (Transaction)	6,200	213,863	201,699	186,010	154,125	154,125	20,000	6,297,033
Surface (Transaction)	6,200	96	16	95	86	86	38	191
Unit. Price (Transaction)	6,200	2,630	3,528	2,117	1,713	1,713	345	125,116
Beneficiaries	6,200	8.2	11.5	4.0	2.0	2.0	1.0	127.0
Price (IFL)	6,200	207,007	41,762	200,911	179,365	179,365	60,691	479,954
Surface (IFL)	6,200	103	21	101	93	93	1	500
Unit. Price (IFL)	6,200	2,117	3,142	1,978	1,766	1,766	372	172,197
Cost	6,200	11,906	5,642	10,576	7,370	7,370	995	47,635
C								
FTO	29,463	17.5	23.1	11.0	6.0	6.0	1.0	424.0
Price (Transaction)	29,463	161,697	137,386	144,000	114,000	114,000	12,000	6,258,743
Surface (Transaction)	29,463	100	18	99	90	90	20	400
Unit. Price (Transaction)	29,463	1,849	2,728	1,534	1,229	1,229	138	169,770
Beneficiaries	29,463	3.7	5.7	2.0	1.0	1.0	1.0	169.0
Price (IFL)	29,463	181,554	37,411	178,500	157,928	157,928	40,000	492,888
Surface (IFL)	29,463	108	22	104	95	95	1	700
Unit. Price (IFL)	29,463	1,769	2,769	1,706	1,504	1,504	165	243,577
Cost	29,463	9,851	5,117	8,489	5,729	5,729	322	41,061

2.A.4 ABC Perimeter

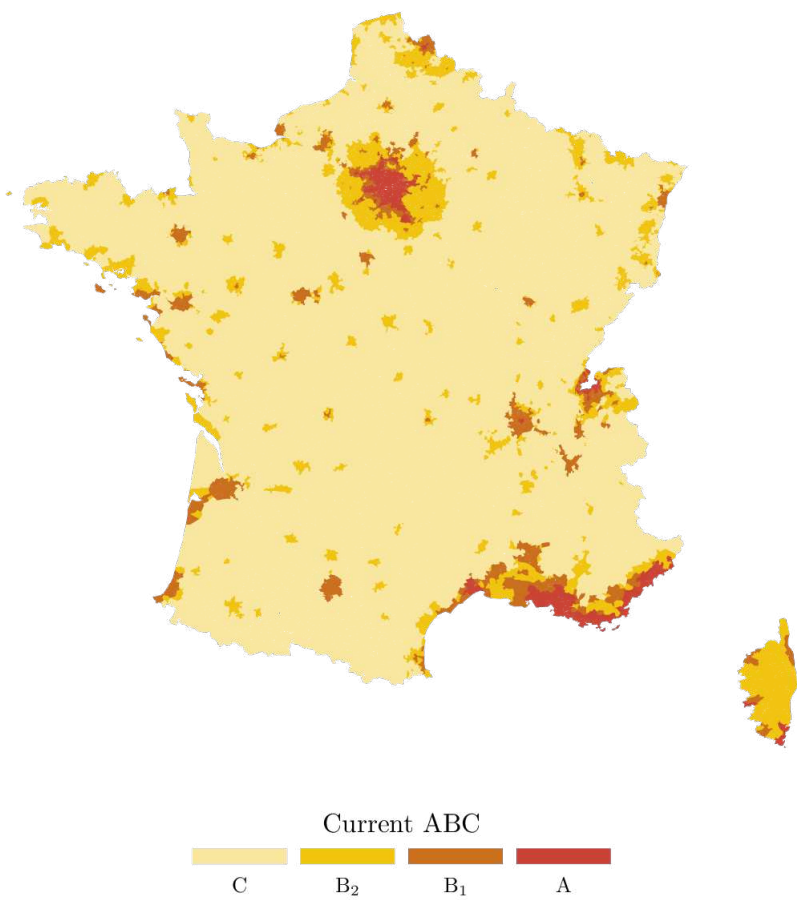


Figure 2.A.2: Current ABC Zoning

2.B Credit Cost

The monetary benefit of IFL for the subsidized first-time owner is equal to the cost for the government, and without extensive and intensive margins, the IFL policy is just a transfer. Consider a loan of total value V_b for a loan duration d at a yearly interest rate of r . For each due date, t , the new homeowner reimburses a fixed payment m . The remaining capital to reimburse at the end of the year is:

$$X_t = X_{t-1} - m + rX_{t-1} = (1 + r) X_{t-1} - m \quad (2.13)$$

Then after calculation, we obtain, using the condition $X_0 = V_b$

$$X_t = (1 + r)^t \left[V_b - \frac{m}{r} \right] + \frac{m}{r} \quad (2.14)$$

Thus, we estimate the monthly payment using $X_D = 0$, corresponding to the loan maturity. Hence, we obtain:

$$m = \frac{rX_0}{1 - (1 + r)^{-D}} \quad (2.15)$$

yielding an overall cost for the household to

$$C = \sum_{k=1}^D m - V_b = \left[\frac{D r}{1 - (1 + r)^{-D}} - 1 \right] V_b \quad (2.16)$$

2.C First-Step: GPS Specification

2.C.1 Estimated Spline Functions for the GPS specification

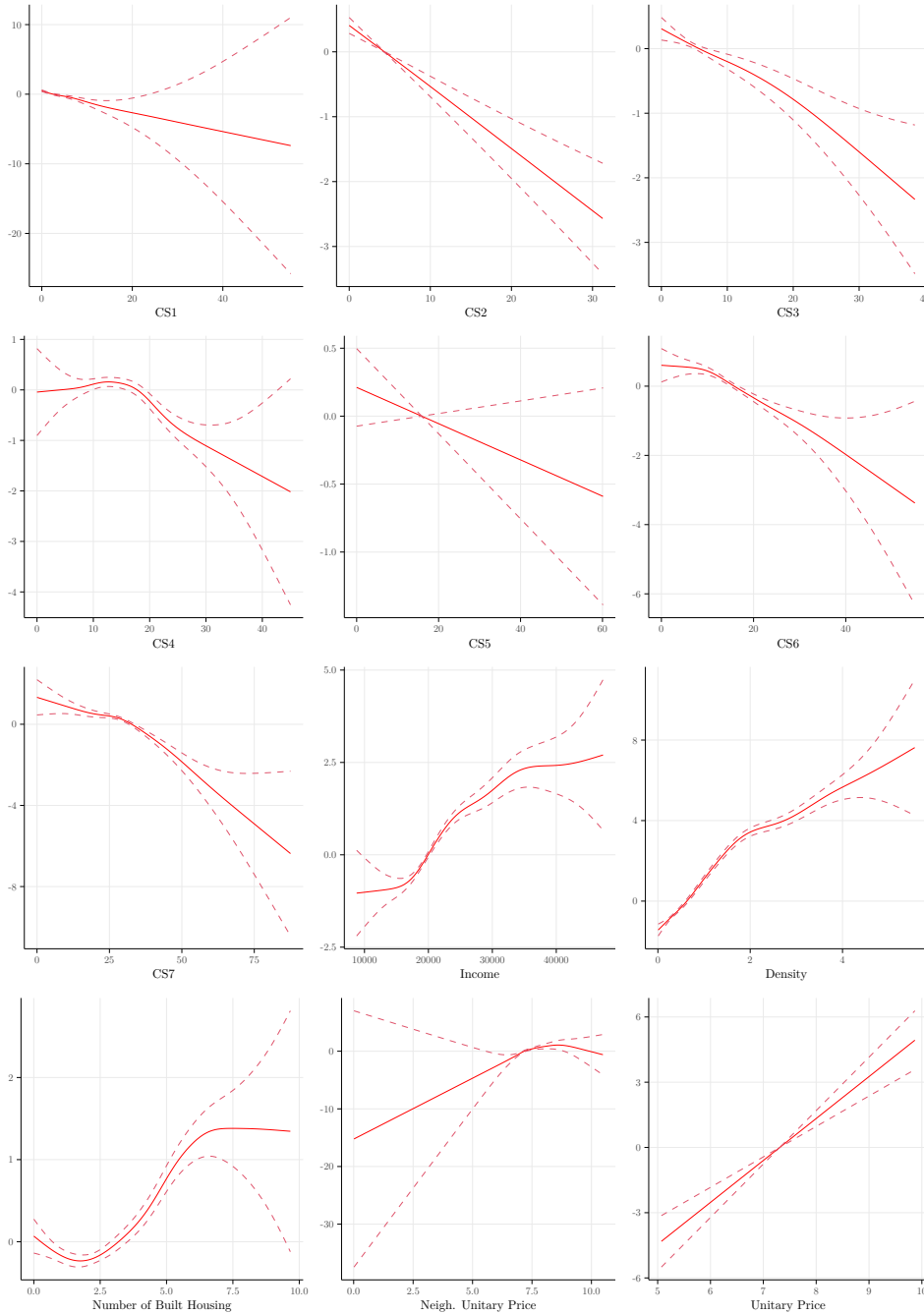


Figure 2.C.1: Contribution for 1D-variable in the GPS Estimation

Notes: For each continuous covariate, we report the functional form in the GPS estimation following the endogenous shrinkage procedure to set the effective degree of freedom. In addition, we report the confidence interval for a 95% level. We exploit the `gam` function from the `mgcv` package.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

2.C.2 Spatial Smoothing Splines for GPS Estimation

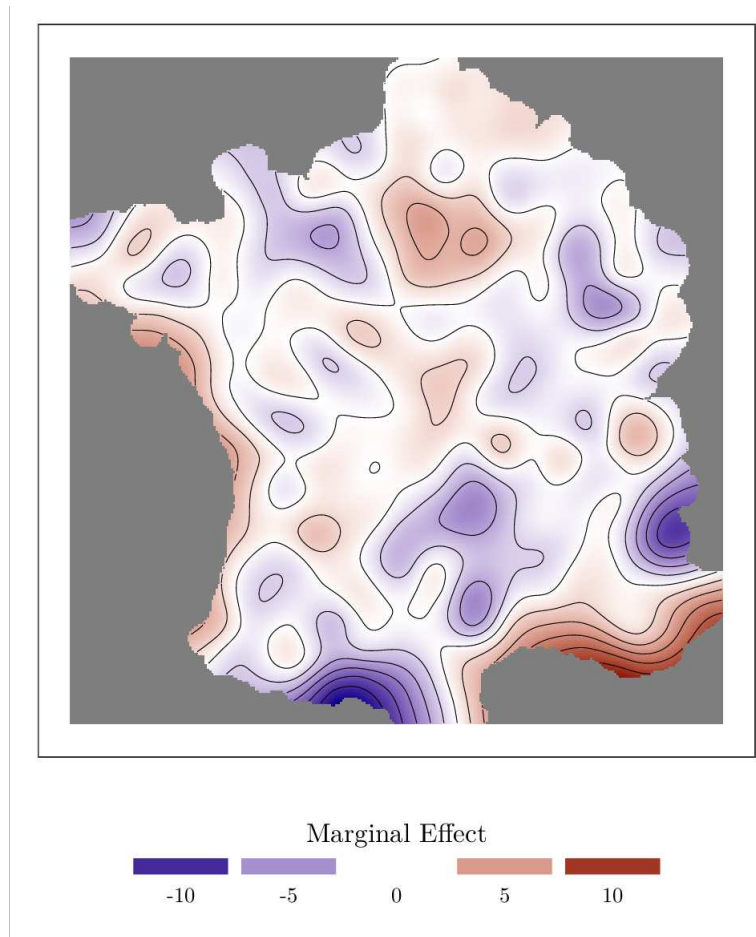


Figure 2.C.2: Spatial Smoothing Function for the GPS Estimation Based On Municipality Coordinates

Notes: We report the spatial smoothing function for the GPS estimation, using bi-variate additive splines. Spline parameters are endogenously shrunk using restricted maximum likelihood approach. The maximum degree of freedom is set to 200. Red (respectively blue) values indicate that the outcome is locally higher than the average. We exploit the `gam` function from the `mgcv` package.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

2.C.3 Predicted Zoning from GPS Estimation

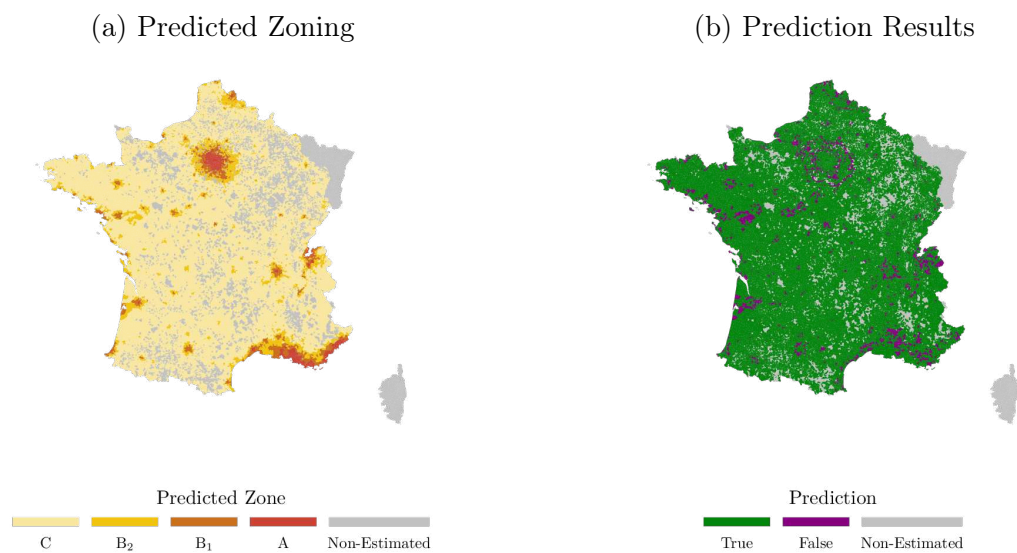


Figure 2.C.3: Estimated Municipalities Classification

Notes: Figure a reports the ABC classification resulting from the estimation. We compare the ABC classification and provide the map of error on Figure b. Municipalities with no values correspond to observations with at least one missing variable.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

2.C.4 Descriptive Statistics about Overlap

Table 2.C.1: Overlap

	Share				Number of Obs.			
	A	B ₁	B ₂	C	A	B ₁	B ₂	C
A	95.0%	26.6%	3.9%	0.0%	679	190	28	0
B ₁	30.3%	95.0%	73.6%	4.7%	398	1,247	966	62
B ₂	3.0%	75.5%	95.0%	34.6%	89	2,266	2,850	1,038
C	0.0%	4.2%	20.9%	95.0%	0	909	4,565	20,701

Notes: We report for each pair of treatment level the overlap measured by the share of observations in treatment level g belonging to the 95% range of the latent distribution of the treatment level k . For instance, using the second row of the table, 30.3% of observations classified as B_1 belong to the 95% range of distribution restricted to A observations, according to the latent variable.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

2.D Policy Effects

2.D.1 Joint Significance for Outcomes Prediction (Pooled Models)

Table 2.D.1: Joint Significance for Control Variables (Second Step)

	Outcome									
	Transaction					IFL				
	FTO	Price	Surface	Unit. Price	Number	Price	Surface	Unit. Price	Cost	
Density	30.0 [3.8] 24/24	3.3 [2.8] 13/24	10.8 [3.2] 19/24	4.6 [3.1] 16/24	8.6 [3.4] 21/24	4.3 [3.6] 15/24	8.6 [3.4] 18/24	4.7 [3.0] 15/24	6.1 [2.8] 15/24	
New Housing	147.5 [4.6] 24/24	5.6 [3.1] 14/24	8.4 [3.7] 19/24	10.9 [3.0] 20/24	47.5 [4.2] 24/24	3.8 [3.4] 14/24	5.1 [3.5] 17/24	7.7 [3.3] 17/24	3.7 [2.7] 12/24	
Median Income	20.2 [3.8] 24/24	8.3 [3.6] 21/24	44.3 [3.9] 24/24	4.6 [3.6] 14/24	4.7 [3.5] 16/24	7.8 [3.7] 17/24	6.2 [3.7] 19/24	6.0 [3.6] 20/24	4.3 [3.2] 15/24	
Price per m ²	3.4 [3.1] 10/24	3.0 [2.6] 8/24	3.6 [3.1] 13/24	3.1 [2.8] 10/24	4.3 [3.2] 16/24	8.8 [3.1] 13/24	3.4 [2.7] 14/24	3.7 [3.3] 15/24	3.5 [3.1] 12/24	
Neigh. Price per m ²	4.3 [3.4] 16/24	4.5 [2.9] 8/24	6.3 [3.5] 21/24	4.4 [3.0] 10/24	3.6 [3.4] 13/24	4.2 [3.2] 15/24	5.6 [3.2] 20/24	6.2 [3.1] 18/24	2.5 [2.9] 7/24	
Spatial Coordinates	6.8 [133.0] 24/24	6.4 [122.3] 24/24	8.0 [136.0] 24/24	5.7 [119.7] 24/24	5.8 [131.7] 24/24	5.2 [124.2] 24/24	4.5 [120.8] 24/24	5.7 [125.6] 24/24	4.4 [96.9] 23/24	
Mean R ²	0.82	0.50	0.74	0.47	0.65	0.58	0.47	0.58	0.38	
Mean N	2,336	2,336	2,336	2,336	2,336	2,336	2,336	2,336	2,336	
Mean AIC	6,390	5,750	-608	6,267	8,086	1,007	2,544	2,007	5,125	

Notes: For the nine outcomes Y in columns, we report the average effective degree of freedom, the average χ^2 test and the number of joint significance for control variable in the second step estimation. In addition, we report average regression statistics (bottom rows). These statistics are derived from the estimation of bilateral combinations effects required to obtain our dose-response functions. The unit of observations is municipality. We exploit the `gam` function from the `mgcv` package.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

2.D.2 Spatial Smoothing Splines for Outcomes Prediction (Pooled Models)

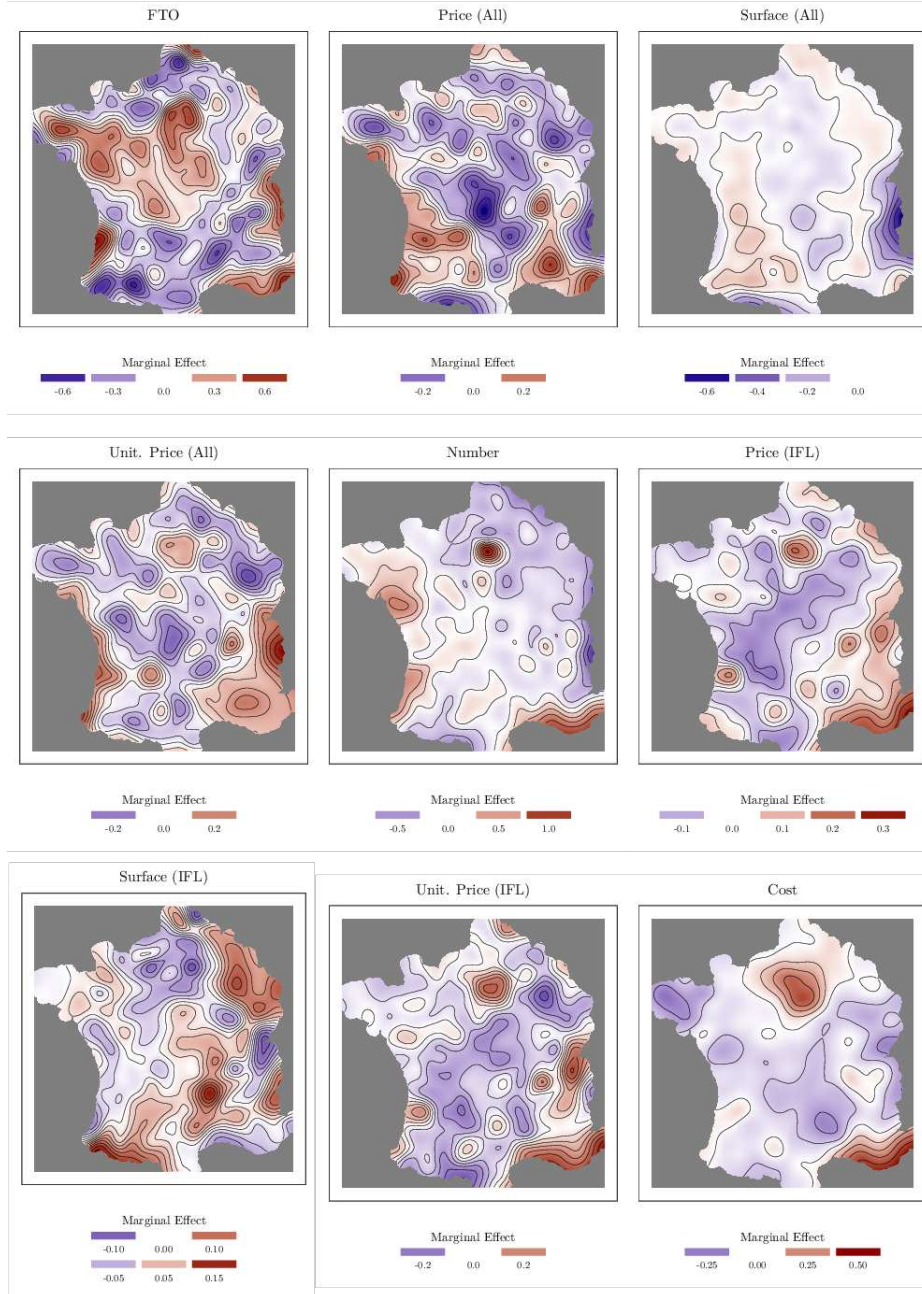


Figure 2.D.1: Marginal Contribution for Spatial Coordinates (Second Step)

Notes: For the nine outcomes Y in columns, we report the spatial smoothing functions for pooled regressions. Our outcome respectively comes from fiscal data or recipients' files. The effective degree of freedom for each function is endogenously shrunk. Red (respectively blue) values indicate that the outcome is locally higher than the average. We exploit the `gam` function from the `mgcv` package.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

Table 2.D.2: Policy Relevant Treatment Effects (ATE)

Area	2015				Treatment (ATE) 2016-2017				2018-2019			
	A	B ₁	B ₂	C	A	B ₁	B ₂	C	A	B ₁	B ₂	C
					<i>N_g</i> : Number of FTO							
A	-	0.242 (0.980)	0.451 (0.965)	0.469 (0.965)	-	-1.180 (1.261)	-0.359 (1.245)	-0.508 (1.245)	-	1.168 (0.882)	1.026 (0.872)	0.979 (0.870)
B ₁	-0.242 (0.980)	-	0.209 (0.175)	0.227 (0.176)	1.180 (1.261)	-	0.821*** (0.184)	0.672*** (0.182)	-1.168 (0.882)	-	-0.142 (0.163)	-0.189 (0.159)
B ₂	-0.451 (0.965)	-0.209 (0.175)	-	0.018 (0.027)	0.359 (1.245)	-0.821*** (0.184)	-	-0.149*** (0.029)	-1.026 (0.872)	0.142 (0.163)	-	-0.048* (0.025)
C	-0.469 (0.965)	-0.227 (0.176)	-0.018 (0.027)	-	0.508 (1.245)	-0.672*** (0.182)	0.149*** (0.029)	-	-0.979 (0.870)	0.189 (0.159)	0.048* (0.025)	-
<i>V_g</i> : Average Housing Price (Overall Transaction)												
A	-	-0.431** (0.190)	-0.129 (0.150)	-0.199 (0.150)	-	0.446 (0.321)	0.488 (0.319)	0.456 (0.319)	-	0.294 (0.286)	-0.202 (0.280)	-0.248 (0.280)
B ₁	0.431** (0.190)	-	0.301*** (0.112)	0.232** (0.111)	-0.446 (0.321)	-	0.043 (0.044)	0.010 (0.043)	-0.294 (0.286)	-	-0.496*** (0.068)	-0.542*** (0.067)
B ₂	0.129 (0.150)	-0.301*** (0.112)	-	-0.070*** (0.011)	0.321 (0.319)	-0.043 (0.044)	-	-0.033*** (0.011)	0.202 (0.280)	0.496*** (0.068)	-	-0.046*** (0.008)
C	0.199 (0.150)	-0.232** (0.111)	0.070*** (0.011)	-	-0.456 (0.319)	-0.010 (0.043)	0.033*** (0.011)	-	0.248 (0.280)	0.542*** (0.067)	0.046*** (0.008)	-
<i>Q_g</i> : Average Housing Size (Overall Transaction)												
A	-	-0.081 (0.161)	0.132 (0.156)	0.113 (0.156)	-	0.177 (0.160)	0.147 (0.157)	0.168 (0.157)	-	0.270* (0.157)	0.128 (0.157)	0.158 (0.157)
B ₁	0.081 (0.161)	-	0.213*** (0.031)	0.194*** (0.031)	-0.177 (0.160)	-	-0.030 (0.035)	-0.010 (0.035)	-0.270* (0.157)	-	-0.142*** (0.024)	-0.112*** (0.023)
B ₂	-0.132 (0.156)	-0.213*** (0.031)	-	-0.019*** (0.006)	0.160 (0.157)	0.030 (0.035)	-	0.021*** (0.005)	0.157 (0.157)	-	-	0.030*** (0.004)
C	-0.113 (0.156)	-0.194*** (0.031)	0.019*** (0.006)	-	-0.168 (0.157)	0.010 (0.035)	-0.021*** (0.005)	-	-0.158 (0.157)	0.112*** (0.023)	-0.030*** (0.004)	-
<i>P_g</i> : Average Housing Price per m ² (Overall Transaction)												
A	-	-0.116 (0.198)	0.092 (0.181)	0.061 (0.181)	-	0.209 (0.363)	-0.021 (0.361)	-0.103 (0.359)	-	0.259 (0.230)	-0.146 (0.202)	-0.227 (0.202)
B ₁	0.116 (0.198)	-	0.209** (0.088)	0.177** (0.087)	-0.209 (0.363)	-	-0.229*** (0.051)	-0.312*** (0.050)	-0.259 (0.230)	-	-0.404*** (0.116)	-0.485*** (0.116)
B ₂	-0.092 (0.181)	-0.209** (0.088)	-	-0.031*** (0.011)	0.021 (0.361)	0.229*** (0.051)	-	-0.082*** (0.011)	0.146 (0.202)	0.404*** (0.116)	-	-0.081*** (0.010)
C	-0.061 (0.181)	-0.177** (0.087)	0.031*** (0.011)	-	0.103 (0.359)	0.312*** (0.050)	0.082*** (0.011)	-	0.227 (0.202)	0.485*** (0.116)	0.081*** (0.010)	-

Notes: We report the bilateral combinations effect for ATE type estimator. The four panels correspond to the four outcome concerned by ATE estimation and derived from fiscal data. Then, we have three main columns that represent the stable period for the IFL scheme, with four subcolumns related to the ABC classification. In rows, we have again the levels contained in the ABC classification. Hence, the bilateral combinations are reported for each intersection, and must be understand as “if (rows) have received (cols), difference in outcome would be (results)”. We also report in parentheses the standard error obtain through a 500-iteration bootstrap.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data. *** p < 0.01, ** p < 0.05 * p < 0.1

Table 2.D.3: Policy Relevant Treatment Effects (ATT)

Area	2015				Treatment (ATT) 2016-2017				2018-2019			
	A	B ₁	B ₂	C	A	B ₁	B ₂	C	A	B ₁	B ₂	C
					\tilde{N}_g : Number of IFL							
A	-	-0.999*** (0.041)	-0.807*** (0.150)	-1.200*** (0.049)	-	-0.257*** (0.049)	-0.620*** (0.115)	-1.985*** (0.105)	-	-0.200*** (0.053)	-0.585*** (0.105)	-1.711*** (0.109)
B ₁	3.446*** (0.047)	-	-0.196*** (0.040)	-0.313*** (0.013)	-0.338*** (0.080)	-	-0.207*** (0.043)	-0.361*** (0.018)	-2.286*** (0.041)	-	-0.214*** (0.039)	-0.517*** (0.016)
B ₂	3.815*** (0.325)	-0.144*** (0.043)	-	-0.193*** (0.018)	0.018 (0.286)	-0.186*** (0.055)	-	-0.193*** (0.016)	-1.534*** (0.132)	-0.200*** (0.045)	-	-0.235*** (0.019)
C	5.386*** (1.064)	0.275 (0.226)	-0.008 (0.032)	-	0.538 (0.974)	0.131 (0.283)	0.044 (0.038)	-	-3.022*** (0.456)	-1.093*** (0.210)	0.003 (0.036)	-
\tilde{V}_g : Average Housing Price (Subsidized Housing)												
A	-	-0.028*** (0.008)	-0.191*** (0.023)	-0.076*** (0.013)	-	-0.015* (0.008)	-0.076*** (0.015)	-0.104*** (0.015)	-	-0.101*** (0.008)	-0.140*** (0.019)	-0.054*** (0.010)
B ₁	-0.801*** (0.011)	-	-0.043*** (0.009)	-0.036*** (0.002)	-0.105*** (0.008)	-	-0.044*** (0.006)	-0.043*** (0.003)	-0.279*** (0.008)	-	-0.043*** (0.006)	-0.037*** (0.002)
B ₂	-1.003*** (0.082)	0.096*** (0.008)	-	-0.028*** (0.004)	-0.008 (0.048)	0.037*** (0.010)	-	-0.026*** (0.002)	0.236*** (0.026)	0.071*** (0.009)	-	-0.031*** (0.003)
C	-1.509*** (0.272)	0.160*** (0.038)	0.081*** (0.010)	-	-0.214 (0.147)	0.063 (0.069)	0.075*** (0.004)	-	0.391*** (0.061)	0.162*** (0.033)	-0.003 (0.004)	-
\tilde{Q}_g : Average Housing Size (Subsidized Housing)												
A	-	-0.109*** (0.008)	0.166*** (0.031)	0.701*** (0.016)	-	-0.052*** (0.007)	0.031* (0.018)	0.158*** (0.014)	-	-0.135*** (0.010)	-0.019 (0.021)	0.348*** (0.006)
B ₁	-1.027*** (0.014)	-	0.022*** (0.009)	0.142*** (0.002)	-0.385*** (0.009)	-	0.028*** (0.006)	0.040*** (0.002)	-0.589*** (0.014)	-	0.017*** (0.007)	0.138*** (0.001)
B ₂	-1.064*** (0.170)	0.026*** (0.009)	-	0.016*** (0.004)	-0.847*** (0.095)	0.081*** (0.010)	-	0.003 (0.002)	0.102*** (0.027)	0.027*** (0.009)	-	-0.019*** (0.002)
C	-1.997*** (0.494)	0.216*** (0.038)	0.054*** (0.008)	-	-1.831*** (0.380)	0.409*** (0.059)	0.042*** (0.006)	-	0.553*** (0.125)	0.224*** (0.026)	-0.005 (0.006)	-
\tilde{P}_g : Average Housing Price per m ² (Subsidized Housing)												
A	-	0.100*** (0.009)	-0.355*** (0.024)	-0.481*** (0.014)	-	0.037*** (0.007)	-0.101*** (0.015)	-0.261*** (0.013)	-	0.042*** (0.007)	-0.134*** (0.021)	-0.341*** (0.010)
B ₁	0.231*** (0.013)	-	-0.071*** (0.008)	-0.137*** (0.002)	0.242*** (0.007)	-	-0.075*** (0.006)	-0.083*** (0.002)	0.273*** (0.009)	-	-0.063*** (0.007)	-0.171*** (0.001)
B ₂	0.115 (0.178)	0.037*** (0.009)	-	-0.044*** (0.004)	0.486*** (0.045)	-0.043*** (0.013)	-	-0.028*** (0.002)	0.282*** (0.026)	0.047*** (0.008)	-	-0.012*** (0.004)
C	0.484 (0.500)	-0.046 (0.049)	0.019*** (0.010)	-	2.370*** (0.151)	-0.317*** (0.006)	0.046*** (0.011)	-	-0.008 (0.084)	0.061*** (0.030)	0.008 (0.008)	-
\tilde{C}_g : Average Cost per IFL												
A	-	-0.125*** (0.010)	-0.516*** (0.064)	-1.514*** (0.020)	-	-0.035*** (0.010)	-0.180*** (0.021)	-0.155*** (0.022)	-	-0.090*** (0.008)	-1.115*** (0.028)	-0.071*** (0.014)
B ₁	-0.096*** (0.023)	-	-0.527*** (0.017)	-1.148*** (0.003)	0.316*** (0.013)	-	-0.204*** (0.009)	-0.237*** (0.004)	-0.196*** (0.015)	-	-0.856*** (0.009)	-0.844*** (0.002)
B ₂	0.243*** (0.080)	0.509*** (0.015)	-	-0.490*** (0.008)	0.148*** (0.041)	0.115*** (0.009)	-	-0.111*** (0.003)	1.208*** (0.027)	0.828*** (0.019)	-	-0.138*** (0.006)
C	0.531*** (0.255)	0.905*** (0.101)	0.555*** (0.016)	-	0.148 (0.128)	0.419*** (0.056)	0.161*** (0.011)	-	-0.756*** (0.133)	0.816*** (0.075)	0.185*** (0.010)	-

Notes: We report the bilateral combinations effect for ATT type estimator. The five panels correspond to the five outcomes concerned by ATT estimation and derived from recipients' files. Then, we have three main columns that represent the stable period for the IFL scheme, with four subcolumns related to the ABC classification. In rows, we have again the levels contained in the ABC classification. Hence, the bilateral combinations are reported for each intersection, and must be understand as "if (rows) have received (cols), difference in outcome would be (results)". We also report in parentheses the standard error obtain through a 500-iteration bootstrap.

2.D.3 Partial Plots (1)

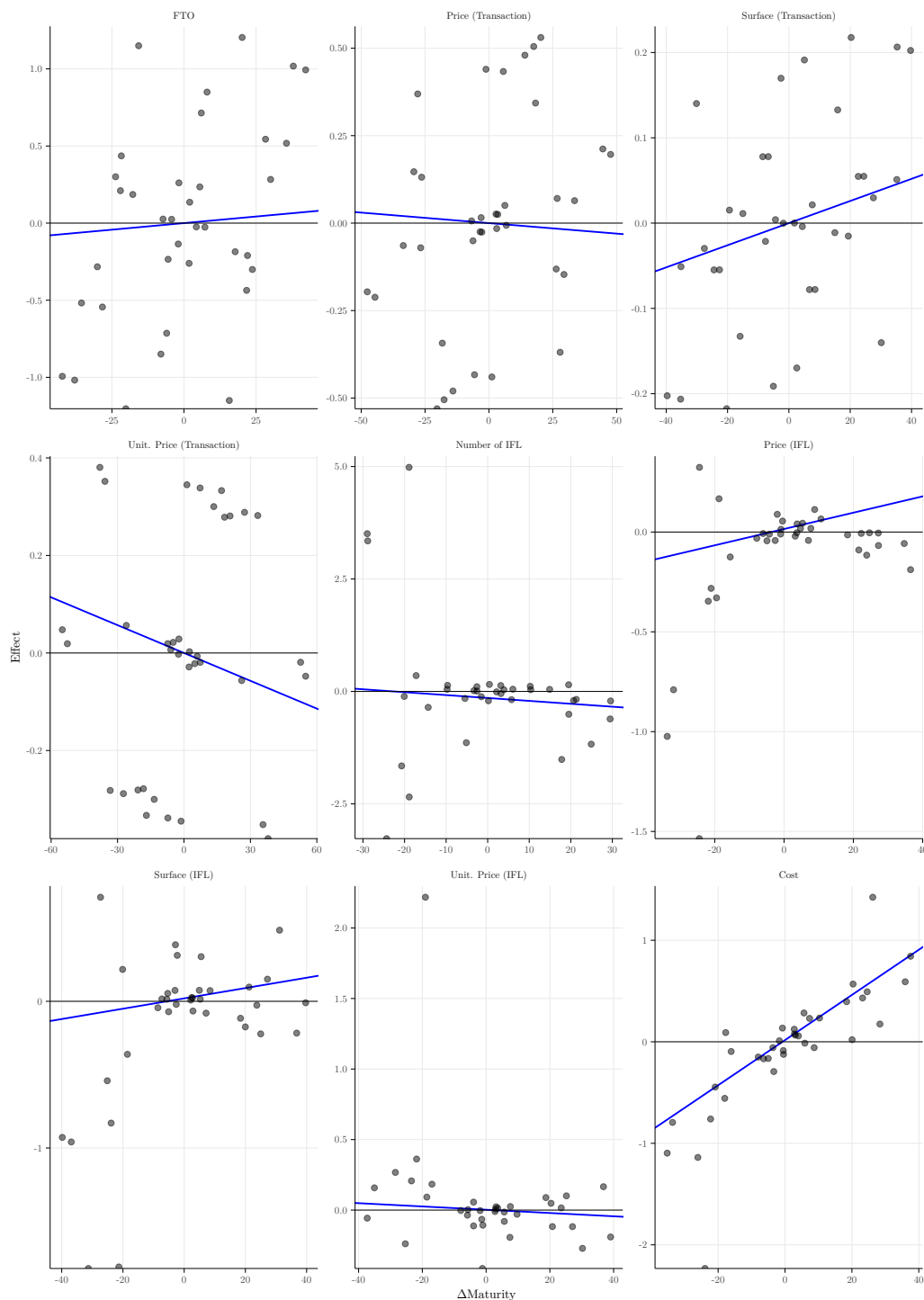


Figure 2.D.2: Dose-Response Plots

Notes: We report the partial plot for dose-response function. The nine partial plot corresponds to our nine selected outcomes, while the x-axis represents a variation in primitive data sources. The observation unit is the bilateral combinations of treatment (either ATE or ATT). The regression is performed using OLS.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

2.D.4 Partial Plots (2)

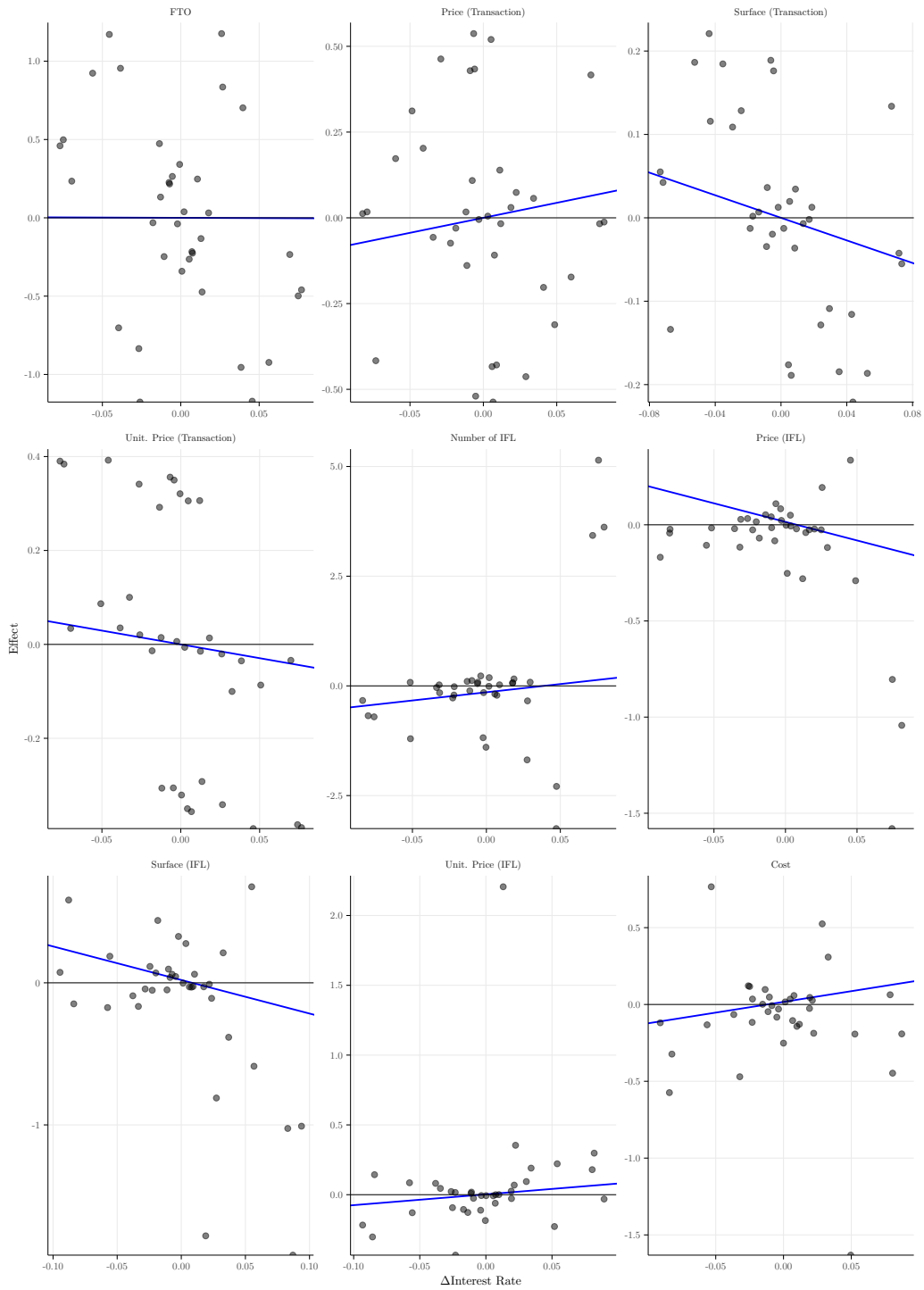


Figure 2.D.3: Dose-Response Plots

Notes: We report the partial plot for dose-response function. The nine partial plot corresponds to our nine selected outcomes, while the x-axis represents a variation in primitive data sources. The observation unit is the bilateral combinations of treatment (either ATE or ATT). The regression is performed using OLS.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

2.D.5 Partial Plots (3)

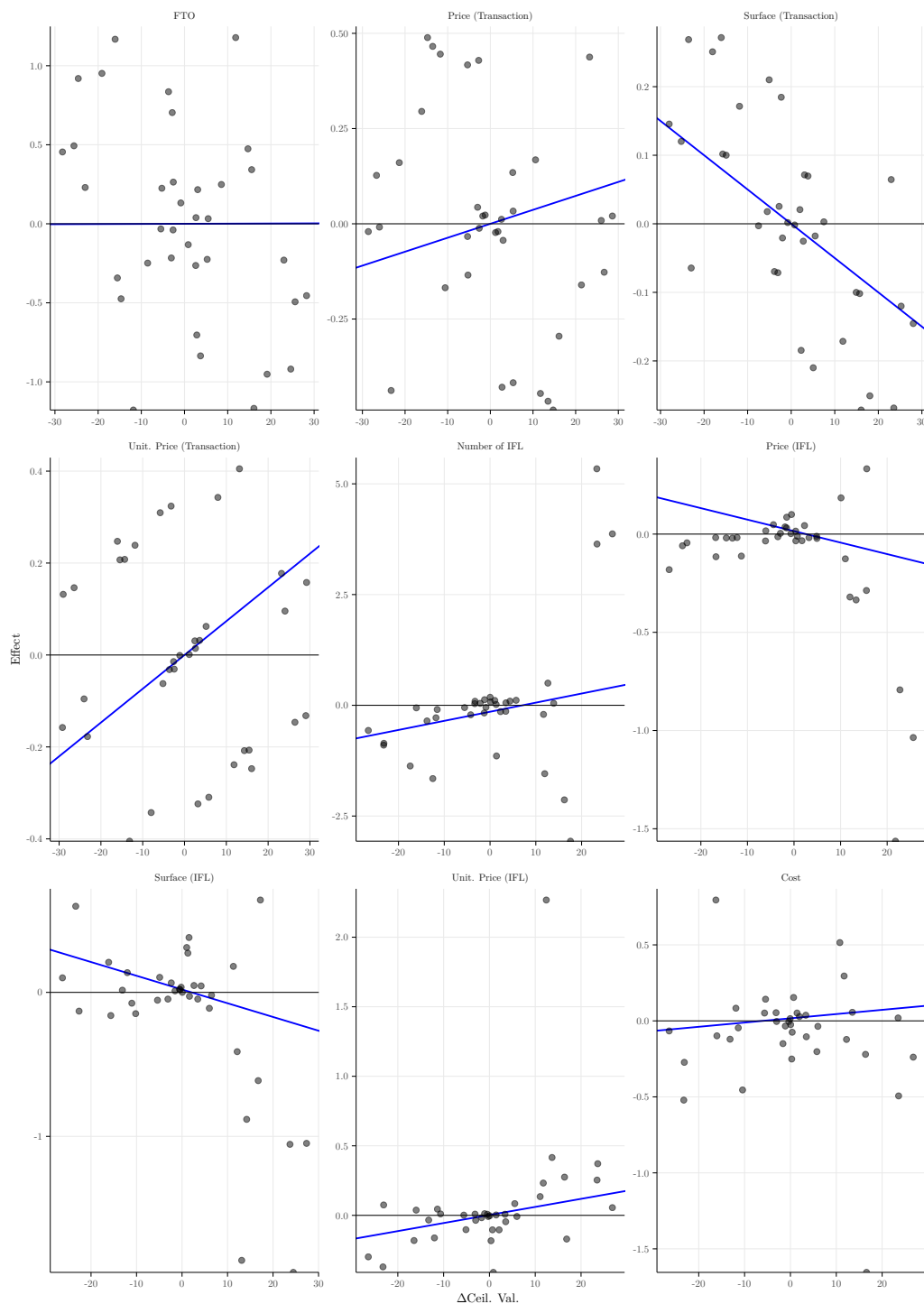


Figure 2.D.4: Dose-Response Plots

Notes: We report the partial plot for dose-response function. The nine partial plot corresponds to our nine selected outcomes, while the x-axis represents a variation in primitive data sources. The observation unit is the bilateral combinations of treatment (either ATE or ATT). The regression is performed using OLS.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

2.D.6 Partial Plots (4)

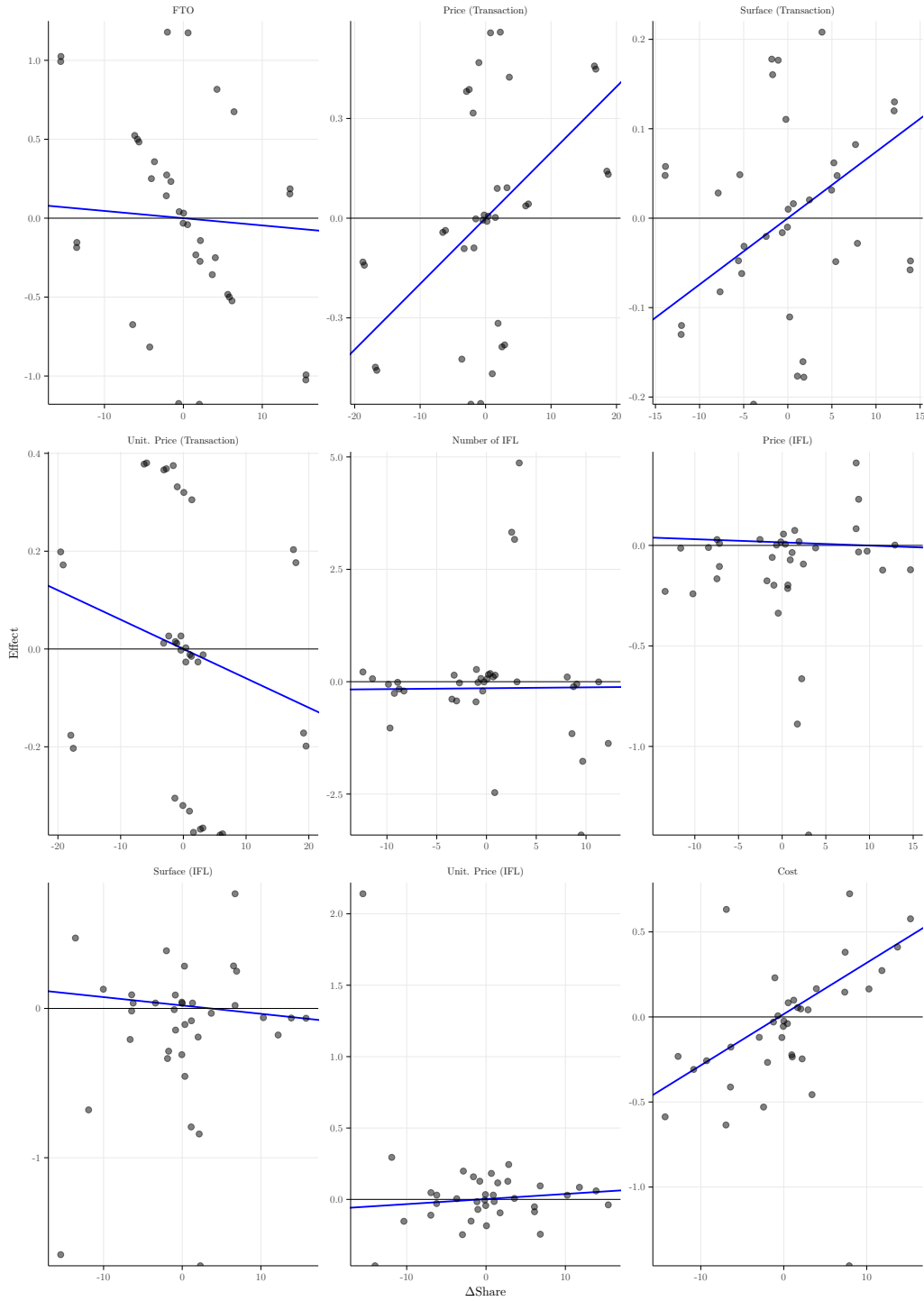


Figure 2.D.5: Dose-Response Plots

Notes: We report the partial plot for dose-response function. The nine partial plot corresponds to our nine selected outcomes, while the x-axis represents a variation in primitive data sources. The observation unit is the bilateral combinations of treatment (either ATE or ATT). The regression is performed using OLS.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

2.D.7 Measures of cost-effectiveness for different marginal values of externalities (Alternative GPS Specification with $df = 50$ for Spatial Smoothing)

Table 2.D.4: Measures of cost-effectiveness for different marginal values of externalities (Alternative GPS Specification with $df = 50$ for Spatial Smoothing)

Primitive	$\phi = -10.0k$				$\phi = 15.0k$				$\phi = 30.0k$			
	-10.0k	-5.0k	0.0k	5.0k	-10.0k	-5.0k	0.0k	5.0k	-10.0k	-5.0k	0.0k	5.0k
Covering Share	-3.299 (4.54)	-1.661 (2.29)	-0.023 (0.31)	1.615 (2.29)	-3.241 (4.56)	-1.603 (2.31)	0.035 (0.47)	1.672 (2.31)	-3.206 (4.63)	-1.569 (2.45)	0.069 (0.93)	1.707 (2.45)
Ceil. Value	0.900 (5.47)	0.692 (2.76)	0.484 (0.46)	0.275 (2.76)	-0.309 (5.49)	-0.517 (2.81)	-0.725 (0.69)	-0.934 (2.81)	-1.034 (5.62)	-1.243 (3.06)	-1.451 (1.38)	-1.659 (3.06)
Int. Rate	-2.671 (338.82)	-0.796 (169.99)	1.080 (16.25)	2.955 (169.99)	-5.370 (339.31)	-3.495 (170.96)	-1.619 (24.38)	0.256 (170.96)	-6.989 (341.93)	-5.114 (176.10)	-3.239 (48.75)	-1.363 (176.10)
Maturity	-0.785 (3.66)	-0.489 (1.84)	-0.193 (0.26)	0.103 (1.84)	-0.302 (3.67)	-0.006 (1.86)	0.290 (0.39)	0.586 (1.86)	-0.013 (3.73)	0.283 (1.99)	0.579 (0.79)	0.875 (1.99)

Notes: We report cost-efficiency with 50 as maximum degree of freedom for the GPS specification and the specification of the outcome. Exploiting coefficients derived from the dose-response function (Table 2.5.3), we calculate efficiency measure defined in Equation (2.7) for the four sources of primitives. It represents the economic returns for the society according to one euro invested. Negative value indicates that increasing the primitive source is counterproductive, while efficiency measure above 1 indicates higher benefit than investment. As our efficiency measure depends on the marginal valuation of externalities at the extensive (ϕ) and the intensive (ψ) margins, we simulate different scenarios. Although extensive margin externalities are likely to be positive unlike intensive margin ones, we also assess scenarios with opposite sign according to our expectations. We report between parentheses the standard errors using 500-iterations bootstrap procedure.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

2.D.8 Measures of cost-effectiveness for different marginal values of externalities (Alternative GPS Specification with No Spatial Smoothing)

Table 2.D.5: Measures of cost-effectiveness for different marginal values of externalities (Alternative GPS Specification with No Spatial Smoothing)

Primitive	$\phi = -10.0k$				$\phi = 15.0k$				$\phi = 30.0k$			
	-10.0k	-5.0k	0.0k	5.0k	-10.0k	-5.0k	0.0k	5.0k	-10.0k	-5.0k	0.0k	5.0k
Covering Share	0.379 (3.80)	0.161 (1.91)	-0.057 (0.24)	-0.275 (1.91)	0.522 (3.81)	0.304 (1.93)	0.086 (0.36)	-0.133 (1.93)	0.608 (3.86)	0.390 (2.03)	0.171 (0.71)	-0.047 (2.03)
Ceil. Value	2.452 (1.73)	1.209 (0.87)	-0.035 (0.10)	-1.279 (0.87)	2.540 (1.73)	1.297 (0.88)	0.053 (0.15)	-1.191 (0.88)	2.593 (1.75)	1.349 (0.91)	0.106 (0.30)	-1.138 (0.91)
Int. Rate	6.141 (49.49)	3.096 (24.85)	0.051 (2.69)	-2.995 (24.85)	6.015 (49.58)	2.969 (25.04)	-0.076 (4.03)	-3.121 (25.04)	5.938 (50.07)	2.893 (25.99)	-0.152 (8.07)	-3.197 (25.99)
Maturity	-1.822 (1.18)	-0.879 (0.59)	0.065 (0.07)	1.008* (0.59)	-1.984 (1.18)	-1.040* (0.60)	-0.097 (0.11)	0.846 (0.60)	-2.081* (1.20)	-1.137* (0.63)	-0.194 (0.21)	0.749 (0.63)

Notes: We report cost-efficiency with no spatial smoothing in GPS specification and specification of the outcomes. Exploiting coefficients derived from the dose-response function (Table 2.5.3), we calculate efficiency measure defined in Equation (2.7) for the four sources of primitives. It represents the economic returns for the society according to one euro invested. Negative value indicates that increasing the primitive source is counterproductive, while efficiency measure above 1 indicates higher benefit than investment. As our efficiency measure depends on the marginal valuation of externalities at the extensive (ϕ) and the intensive (ψ) margins, we simulate different scenarios. Although extensive margin externalities are likely to be positive unlike intensive margin ones, we also assess scenarios with opposite sign according to our expectations. We report between parentheses the standard errors using 500-iterations bootstrap procedure.

Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.
 *** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

2.E Placebo Analysis: Difference in Treatment

The policy-maker can decrease the price of home ownership through two channels: the ceiling value and the share of the loan among the purchase. These two channels produce different effects on the price of homeownership, as the ceiling value may introduce difference for the higher purchase while the share of IFL produce effects on all operations. Indeed, for the first situation, the difference in homeownership cost for two levels of

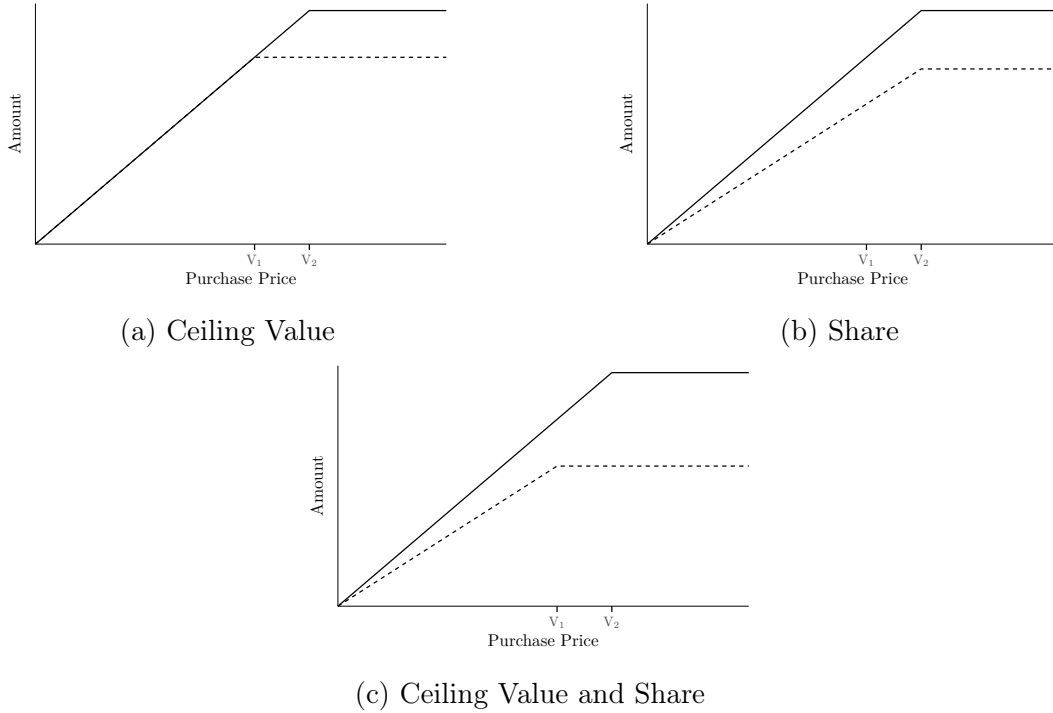


Figure 2.E.1: Difference on Homeownership Cost Induced by the IFL

Notes: Figure a corresponds to difference in treatment based only on the ceiling value. Then, difference in treatment only arises for the more expensive operations. Figure b corresponds to difference in treatment based only on the share of the IFL. Then, difference is homogeneous for all operations. The Figure c corresponds to difference in treatment for both ceiling value and share of the IFL. Then, the difference of treatment is homogeneous for the less expensive operations and increase for the most expensive ones.

treatment being different only about the ceiling value arises for operations above the lowest ceiling value and remains stable for purchase above the higher value. Hence, differences in ceiling value only affects the cost of homeownership for the more expensive operations (Figure 2.E.1a). Conversely, two levels of treatment being different about the share of the loan with no interest decrease the cost of homeownership for all operations, in a proportional manner (Figure 2.E.1b). Finally, if the level of treatment combines both differences in ceiling value and share, both effects add up to, and difference in the cost of homeownership concerns all operations, with a more pronounced difference for the more expensive housing (Figure 2.E.1c).

Chapter 3

Price Capitalization of Demand-side Policies with Segmented Housing Markets

Abstract

I consider the segmentation of the housing market locally by tenure status and type of construction assumed by policy-makers. The segmentation assumption is inherently associated to the nature of the price capitalisation (i.e. whether it concerns the whole local housing markets or specific segments) that derive from demand-side policies in inelastic supply context. Moreover, I seek potential underlying drivers for the price capitalisation based on investors' residence place. I use policy reforms that solely affect demand for rental purchase in new housing to implement difference-in-difference strategy. The results support the assumption that local housing markets are segmented according to tenure status. Owner-occupier segments remain unaffected by the demand shock for rental investment. The price capitalisation is considerable within the subsidised segment (-13.5% for the policy repeal), accounting for more than half of the subsidy. I provide evidence that distant investors contribute to the price capitalisation. Finally, the rental investment scheme excludes potential buyers from the subsidised segment who switch to the unsubsidised segments (existing housing), which also indirectly leads to price capitalisation ($+6.4\%$ for positive demand shock).

JEL classification: R31 ; R38 ; C21

Keywords: Difference-in-differences ; rental investment policies ; housing market segments.

3.1 Introduction

In the early 2000s, house values began to diverge from incomes. Although the dynamic is spatially uneven, economic consequences of the housing boom are numerous. Among them, the reduction of the supply of affordable housing impact both individual and collective outcomes. For instance, it forces individuals to smooth their consumption (Etheridge, 2019; Waxman et al., 2020) with greater constraints on lower incomes (Dustmann et al., 2022), raises income inequality (Albouy et al., 2016) or decreases economic activity (Glaeser and Gyourko, 2018; Szumilo, 2019).

Naturally, questions have emerged about the determinants of house price dynamics that reduce the supply of affordable housing. Given that the housing markets are local (Hwang and Quigley, 2006), in line with its immobile nature, most attention has been paid to supply determinants. The availability of land (Anenberg and Kung, 2020; Molloy, 2020), which is affected by local regulations (Turner et al., 2014; Gyourko and Molloy, 2015; Hilber and Vermeulen, 2016) or geographical constraints (Saiz, 2010), is considered to be a key determinant to affordable housing. Beyond land considerations, housing market dynamics appear to depend on the nature of the housing demand. Hence, the spatial sorting for skills (Behrens and Robert-Nicoud, 2015) is likely to affect the housing market dynamics through differentiated financial capabilities or heterogeneous preferences. In addition, the balance between local and broad purchasers is a key factor for the dynamics of the housing market (and thus affordable housing) as it defines the excess demand in relation to the supply of housing (Piazzesi et al., 2020).

Most countries implement housing policies to mitigate the negative effects of the lack of affordable housing. They are mainly structured according to tenure status, i.e. whether policies favour owner occupiers or investors to increase rental supply. For instance, the US simultaneously implement the Mortgage Interest Deduction (see for instance Glaeser and Shapiro, 2003) to support homeownership and the Low Income Housing Tax Credit (see for a discussion on the LIHTC Baum-Snow and Marion, 2009) to increase rental supply. Numerous countries adopt rental investment schemes (Wijburg, 2022) in addition to the support for homeownership, and France is no exception. However, housing policies that aim to mitigate negative effects from housing market dynamics subsidise demand rather than using the supply mechanism. While support for home ownership is weakly efficient (see for example Gruber et al., 2021), the support to rental investment has a small effect on rental supply (Malpezzi and Vandell, 2002) due to crowding out effects, either to unsubsidised units (Eriksen and Rosenthal, 2010) or to nearby untreated locations (Baum-Snow and Marion, 2009). Moreover, both policies reinforce the housing market dynamics through the capitalisation of subsidies into price. However, given the joint

implementation of both policies, policy-makers assume that local housing markets are perfectly segmented by tenure status and that there are no negative interactions between policy implementation.

In this paper, we assess whether the local segmentation by tenure assumed by policy-makers holds. We derive the local segmentation from the nature of price capitalisation resulting from demand-side policies, i.e. whether it is restricted to the subsidised segment (i.e. heterogeneous effect) or concerns the whole housing market (i.e. homogeneous effect). Moreover, we look for potential mechanisms that drive price capitalisation in the context of low supply elasticity based on purchasers' residence place. In line with the results of Piazzesi et al. (2020), we extend the local-broad distinction to rental assets. Indeed, since rental investment is frictionless in terms of mobility costs compared to owner occupation (Ha et al., 2021), long-distance investments are eased. Empirically, we adopt similar definition as Piazzesi et al. (2020) by considering long-distance investors as purchasers that does not reside within the local housing market.

To identify local housing market segmentation by tenure, we take a different approach from identifying spatial segmentation (Goodman and Thibodeau, 1998; Goodman and Thibodeau, 2007; Bourassa et al., 2021; Coën et al., 2022). The common definition for segmentation is mainly based on three criteria (Rothenberg, 1991)¹ to identify homogeneous markets according to substitutability in the matching process (Wheaton, 1990). The empirical implementation is a two-step procedure: hedonic model estimation and spatial clustering (for a review see Bhattacharjee et al., 2016). However, while the spatial clustering approach combined with the hedonic method reveals spatial segmentation within local housing markets, it does not identify tenure segmentation specifically. Tenure status is expected to be spatially heterogeneous, but cannot be recovered using the spatial approach due to the presence of confounding variables such as income (Duranton and Puga, 2015), skills (Diamond and Gaubert, 2022) or age of buildings (Brueckner and Rosenthal, 2009). In addition, while the spatial consistency for clustering is likely to hold for main residence due to the utility derived from consumption of local amenities (Kuminoff et al., 2013), the assumption of spatial delineation of segments may be less credible for rental investment purchases.

Our approach rather exploits a demand shock limited to a specific segment in an inelastic supply context. In order to isolate the demand shock specific to one segment, we use several reforms of the French rental investment policies. These successive policies (the Pinel-Duflot) boost demand for rental units in new housing, leaving demand in other segments unaffected. Eligible areas are known to have changed several times over the

¹Housing market segments are defined based on similar hedonic prices, similar hedonic characteristics and similar housing characteristics.

2013-2018 period, and are independent of other housing policies for specific municipalities (designated as B_2 on a five-level classification).² In fact, while the scheme was available to all B_2 municipalities in early 2013, the policy was heterogeneously repealed within this subpopulation in July 2013. The scheme was then gradually reintroduced over the period 2015–2018, to finally be repealed entirely for B_2 municipalities in March 2019. We estimate the price adjustment following the demand shock in each segment using a difference-in-differences-type estimator to derive local segmentation according to tenure. We make the best use of individual data on transactions over the period 2010-2020, supplemented by tax property files to track post-purchase ownership status.

Based on our identification strategy, we show that price capitalisation is limited to rental purchases. Yet, the demand shocks for rental investment do not affect owner-occupier segments. Our results therefore confirm that the local segmentation assumed by policy-makers according to tenure status is likely to hold empirically. In addition, the relaxation of the assumption that local housing markets are not segmented according to tenure inflates price effect in the subsidised segment (-13.5% for policy repeal, nonsignificant without local segmentation). We estimate that the attraction of distant investors is at least partly responsible for the price capitalisation in the subsidised market. Moreover, we provide evidence that the introduction of the policy causes a proportion of potential investors to switch from new to existing housing. We expect these results to come from the higher substitutability between new and existing housing, leading to a 6.4% increase in the non-subsidised segment for rental purposes. Nevertheless, we cannot conclude that the out-pricing process is driven by excess demand from distant investors. An additional distinction, such as income, could drive the out-pricing process, although income is unobservable according to the data we use.

We contribute in various ways to the existing literature. First, we demonstrate that local housing markets are segmented according to the tenure status, as assumed by policy-makers. Hence, the joint implementation of housing policies that support homeownership and rental investment do not interact in the short term through price capitalisation. Second, we identify housing market segmentation from demand shocks limited to one segment, rather than a spatial approach combined with the hedonic method. It allows to reveal underlying mechanisms for spatial segmentation. Third, the price capitalisation for demand-side policies in subsidised segments is substantial. Our results suggest that price capitalisation accounts for almost two thirds of the theoretical subsidy. Fourth, rental investment schemes favour long-distance investments that contribute to the price capitalisation.

²An additional subclassification is introduced for municipality classification common to most housing policies that define the eligibility to the rental investment scheme specifically within the B_2 tier.

The paper proceeds as follows. Firstly, we discuss expected price adjustment resulting from demand shock in one specific segment (Section 3.2). Then, we document the data sources we use, and how we enhance them to provide relevant information about tenure status (Section 3.3). We present our identification strategy and detail the housing policy under consideration to generate demand shocks (Section 3.4). Then, we present our results (Section 3.5) and conclude (Section 3.6).

3.2 Framework

To motivate our identification strategy, we first illustrate that housing market segmentation can be inferred from price adjustment, consecutive to demand shock restricted to one segment. Indeed, under perfectly inelastic supply, any shift in demand would transit into price capitalisation. Yet, price capitalisation can affect only the subsidised segment only (perfect segmentation) or the entire market (homogeneous market), depending on the market segmentation.

We consider a housing market with two types of housing being heterogeneous according to their building type (new housing and existing). Then, housing being purchased can be placed either on the rental market or achieve owner consumption. The market is therefore composed of four distinct *a priori* defined segments, according to building type and purchase purpose. In addition, housing supply is perfectly inelastic with a finite number of available housing. At equilibrium, we impose market-clearing and define the number of units in each segment (named as π_ℓ^k , with $\sum_{k,\ell} \pi_\ell^k = 1$). Each segment is defined by an independent demand curve, while supply is inelastic for all segments. Conditions at the equilibrium provide market prices within each segment (Figure 3.2.1).

We now introduce a financial incentive that causes a demand shock limited to a specific segment. We assume that this is the segment defined by new rental housing to fit the rental investment scheme we are considering (more details in the following sections). Firstly, we assume that housing markets are perfectly segmented, i.e. segments cannot provide asset substitutes. In this case, price capitalisation occurs in the subsidised segment due to the assumption of inelastic supply, while the non-subsidised segments remain unaffected (point A', Figure 3.2.1). The financial incentive is then capitalised into the housing price only in the subsidised segment.

We now relax the assumption of perfect segmentation. Firstly, we assume that the segmentation by tenure status is imperfect, i.e. that the housing asset may indistinctly serve rental or owner-occupation purposes. Hence, *within* new housing, the balance between housing intended to be placed on rental markets and for owner-occupancy can be

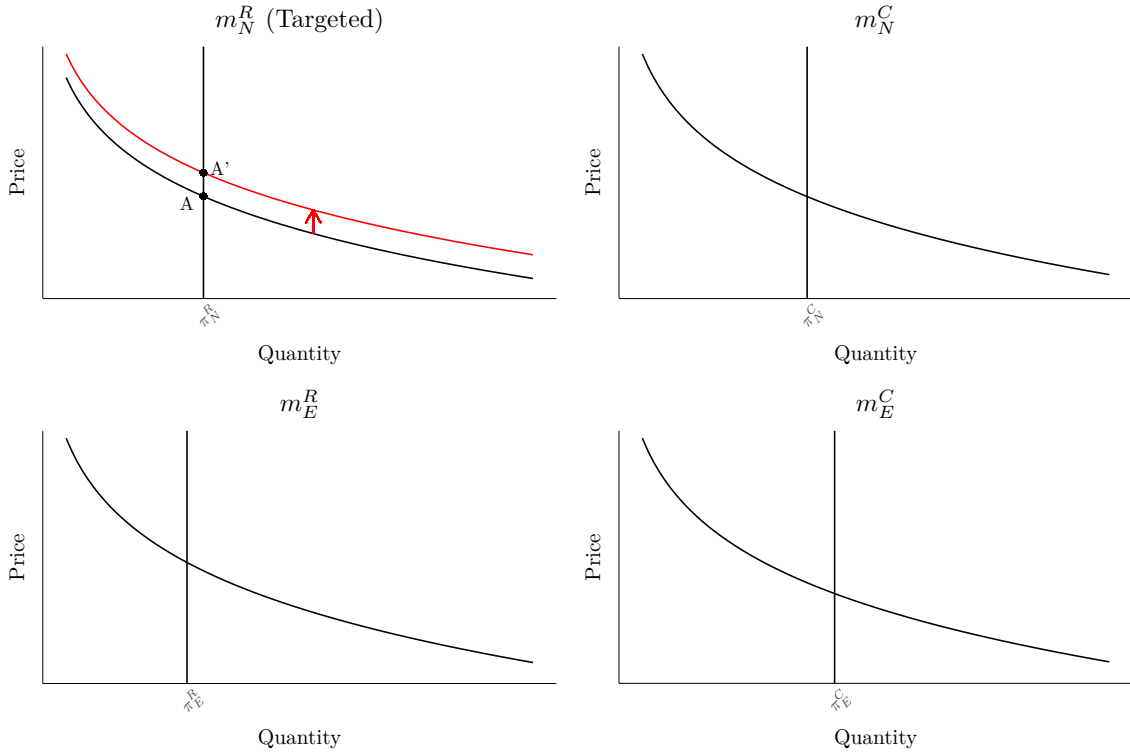


Figure 3.2.1: Price Adjustment According to Market Segmentation

Notes: The top-left panel is the segment subsidised by policy-makers through the implementation of financial incentives that cause demand shift. The demand shock is fully capitalised into price resulting from the inelastic supply conditions. Considering the perfect segmentation, other markets segments remains unaffected by the demand shock.

affected by the implementation of financial incentive. The shift in demand is then partly captured by an increase in the number of housing units purchased for rental, in order to smooth the demand shock (noted $\delta\pi$). Due to inelastic supply conditions, available units for owner-occupier purposes is reduced. It results that the unitary housing price increases in this specific segment although unsubsidised (point B, Figure 3.2.2).

Secondly, we assume that the segmentation by type of construction is imperfect. We then consider that the type of construction is substitutable within each tenure segment. However, the introduction of a financial incentive has ambiguous effects on the unsubsidised segment, depending mainly on whether it reallocates existing demand or creates additional one. We first consider a closed market with reallocation of existing demand. Then, the introduction of a financial incentive drives demand from the non-subsidised segment to the subsidised segment for a similar purchase purpose, in line with the imperfect segmentation by type of construction. It implies that demand in non-subsidised segments, which are imperfectly segmented with the subsidised one, falls and is reflected in falling prices (point C, Figure 3.2.2).

If, on the other hand, the demand comes from incremental demand (e.g. long-distance

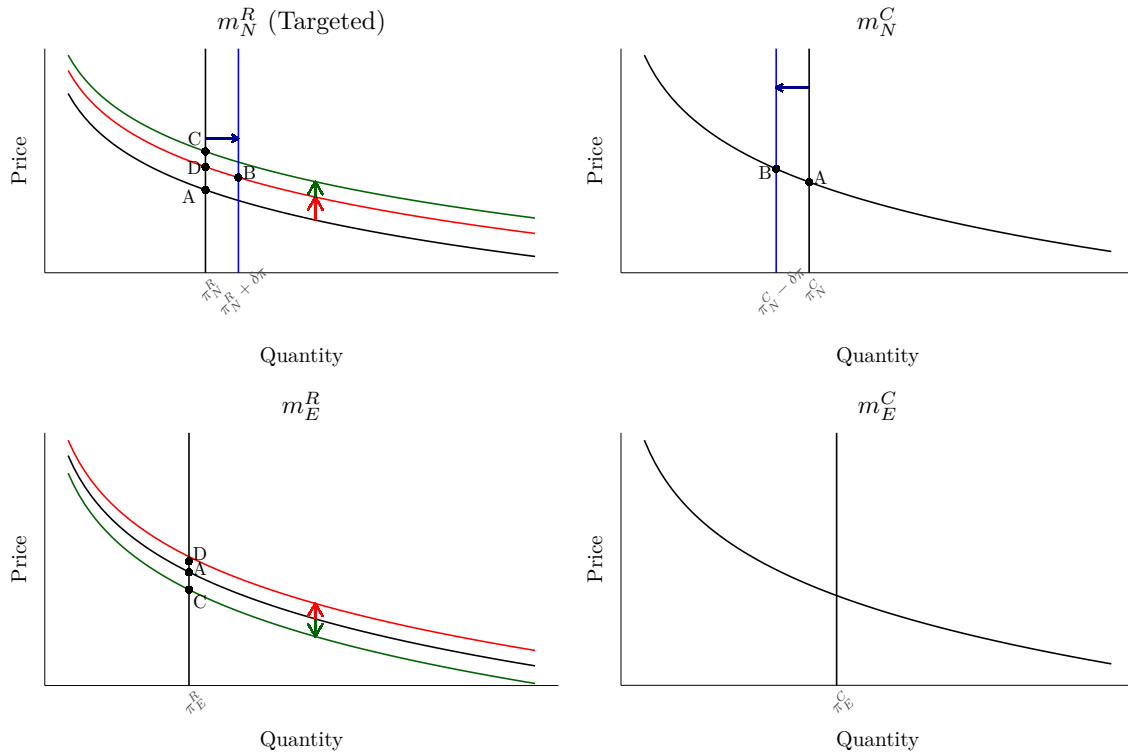


Figure 3.2.2: Price Adjustment According to Market Segmentation

Notes: The top-left panel is the segment subsidised by policy-makers through the implementation of financial incentive, that cause demand shift. The top-right panel correspond to segment with similar housing type and different purchase purpose. We review the case where housing type can indistinctly fulfil rental or owner-occupancy purposes (point B). The bottom-left panel concerns segment with similar purchase purpose and different housing types. We implement the two possibilities dealing with whether financial incentive originates from local investors (point C) or long-distance ones (point D). The bottom-right panel with characteristics being opposite from the subsidised segment.

investment) in line with an open market, the price adjustment in unsubsidised segments is likely to be the opposite, following an outpricing process. In fact, current investors may face increased competition following the demand shift and be priced out of the subsidised segment. Consequently, in order to satisfy their investment needs, they might shift from subsidised segments to non-subsidised segments with similar purchase purposes, according to the substitutable feature. It leads to an increase in demand in non-subsidised segments and thus to price capitalisation (point D, Figure 3.2.2).

Price adjustment following a demand shock is therefore relevant to the discussion of integration between segments within local housing markets. In addition, the nature of the demand shock, i.e. whether financial incentives attract additional investors or reallocate existing investors within the local housing market, determines the price adjustment in non-subsidised segments. However, our strategy requires us to isolate a demand shock limited to a specific segment in order to discuss the segmentation of the housing market and to extend the price capitalisation conclusion.

3.3 Data

To observe price adjustment on *a priori* defined segment, we exploit transaction datasets with detailed information for tenure status and building type. Our main dataset is derived from fiscal sources and covers housing transactions in France since 2010.³ For each observation, housing characteristics are provided, including surface area, type, year of construction and location using spatial coordinates. In addition, transaction details such as purchase price or closing date are available. We restrict the transaction dataset to operations related to housing and remove specific operations such as exchange or adjudication.

To assign each transaction to an *a priori* defined segment based on building type and purchase purpose, we enrich this dataset by exploiting tax property records with detailed information about the housing stock over the 2011–2021 period. It includes tenure status every 01st January. In addition, we add specific filters to remove singular transactions, such as partial purchases following divorce, which result from the panel data approach.⁴ Making the best use of tax records, we assign each housing transaction to potential segments defined *a priori*, according to the type of building (whether the housing is new or existing) and the purpose of the purchase (rented or owner-occupied). Empirically, we proxy the purpose of purchase with the tenure status after purchase. For vacant units, we use the tenure status after one year, which is particularly relevant for new dwellings. We keep transactions related to either rental or owner-occupation purposes.

In addition to detailed information on tenure and characteristics of the property, tax records provide information on the owners. We use the location of owners to determine whether they are local or long-distance purchasers. Empirically, we exploit commuting areas provided by INSEE⁵ to compare housing location with purchasers' ones. We define local purchasers as individuals who live in the same commuting area *prior* to the purchase. Commuting areas are stable geographical units based on both residence and workplace, and assumed to be a good proxy for local housing market. Long-distance investors are defined in opposition, including individuals living abroad. Our empirical definition has a drawback due to data limitations. Since first-time owners are not registered prior to their purchase, we cannot observe their previous place of residence exactly. We therefore choose to adopt their first place of residence as registered in the property tax records. This increases the share of local buyers for owner occupation. However, the effect is assumed

³Our data originates from DV3F. It does not include Alsace-Moselle departments for historical reasons.

⁴For more details, see Chapter 1.

⁵French National Institute of Economic Studies and Statistics. Commuting areas are defined as stable areas for both housing and labour markets. There are 304 commuting areas for the 2010-specification.

to be marginal for rent-to-buy as it does not affect the place of residence. We assess the robustness of this choice, which does not affect our main conclusions. Our resulting dataset contains 5,537,962 observations over the period 2010-2020.

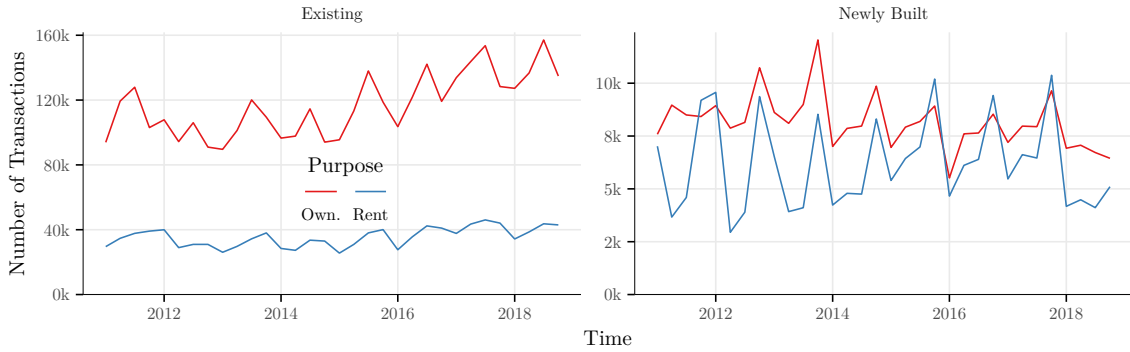


Figure 3.3.1: Number of Transactions According to Defined Segmentation

Notes: We report overall number of transactions according to the construction type (left panel for existing dwellings, right panel for new ones) and tenure purposes (red for owner-occupying, blue for rent). We aggregate number of transactions per quarter.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

Most housing transactions are for owner-occupation purposes, accounting for almost three quarters of transactions in existing segments (left panel, Figure 3.3.1), while new housing transactions more evenly spread by tenure purpose (right panel, Figure 3.3.1). However, despite the fact that most transactions are for owner-occupy purposes, rental purchases are not marginal.

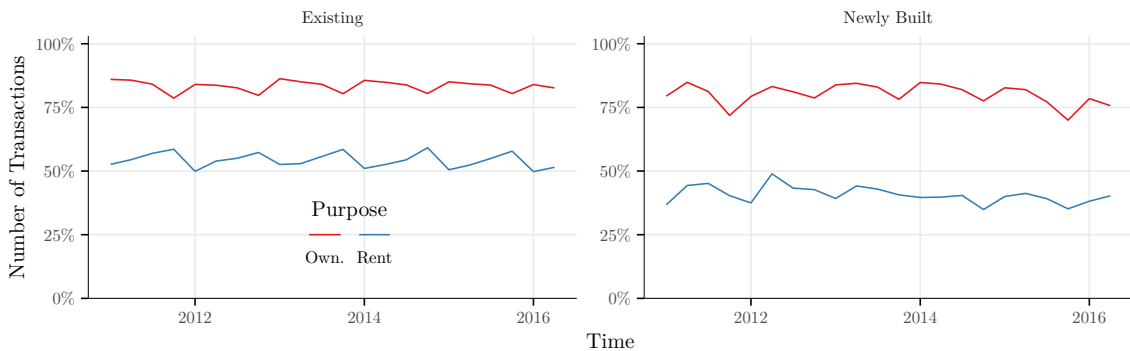


Figure 3.3.2: Share of Narrow Purchasers According to Housing Market Segmentation

Notes: We report share of local purchasers according to the construction type (left panel for existing dwellings, right panel for new ones) and tenure purposes (red for owner-occupying, blue for rent). We define local purchasers according to whether housing is located within the same commuting area as purchasers. We aggregate number of transactions per quarter.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

Share of local purchasers mainly depends on the tenure status (Figure 3.3.2) we consider. Most owner-occupier purchase corresponds to local households (more than 80%),

which is consistent with the local nature of the housing market. Furthermore, for owner-occupiers, there are no significant differences by type of building. Conversely, the share of local investors is lower for rental purposes and varies by building type. On average, almost 50% of existing dwellings purchased for renting are owned by local households, while this falls to 40% for new housing units. Differences according to the purpose of the property are likely to reflect a lack of consumption value from the buyer's perspective for rental purposes. As a result, location choices for rental assets may be less sensitive to local amenities. The elasticity of demand to financial incentives to buy-to-let from distant investors is positive and probably higher than for owner-occupiers.

3.4 Identification Strategy

Our identification of housing market segmentation requires us to isolate a demand shock that is limited to a specific segment (Section 3.2). To this end, we exploit the French housing policy's design of lending incentives to favour rental investment in new housing. Taking into account their place-based nature, we exploit differences in the areas eligible for the rental investment scheme to identify an exogenous increase in demand for investment purposes. We first review the French policy design we exploit being close to the LIHTC in the US (for a presentation, see Baum-Snow and Marion, 2009), and then detail our identification strategy based on the singularity restricted to rental investment policies.

3.4.1 Policy Design

In order to increase the supply of affordable rental housing, French policy-makers have over time introduced various demand-side policies in addition to social housing. A seminal policy design was introduced in 2003 (known as Robien), and financial incentives have been available continuously under various subsequent forms. The last stable scheme was introduced in January 2013 and is called Duflot-Pinel⁶ and is scheduled to be abolished in December 2024. In the rest of the paper, we refer to Pinel to define the Duflot-Pinel scheme.

The Pinel scheme introduces an incentive for households to purchase new housing for rental purposes by reducing their tax income. The financial incentive does not apply to

⁶In practice, Duflot and Pinel schemes are defined as different policies, but Pinel scheme is only a more flexible version of the Duflot scheme. The Duflot scheme was introduced in January 2013, while the Pinel scheme was introduced in July 2013.

existing housing.⁷ It allows households to deduct a proportion of their purchase price from their taxable income over a period of years (three options are available, 6 years, 9 years or 12 years). In return, households commit to renting their purchase for a similar period, with capped rents and income conditions for tenants. The tax reduction increases with the duration of the commitment, ranging from 12% of the purchase price for 6 years to 22% for 12 years. For example, a household that invests 100k in a house under the Pinel scheme over a period of 9 years will benefit from an annual tax reduction of 2k, giving a total subsidy of 18k over a period of 9 years. Note that the financial incentive is therefore more attractive to households with higher taxable incomes, as those below the 2k limit do not take full advantage of the subsidy.

Pinel scheme introduces spatial variation to target specific areas where policy-makers have identified a lack of affordable rental housing. Using the ABC perimeter,⁸ operations located in A and B₁ areas are eligible for the tax reduction, while those in C areas are not. For B₂ municipalities, an important change occurred in July 2013, as operations in these areas are no longer eligible for the tax reduction, except for certain municipalities. The policy design therefore introduces spatial discontinuities, which are commonly used to assess policy effects (Chapelle et al., 2018; Bono and Trannoy, 2019), in line with the spatial continuity assumption (Keele and Titiunik, 2015). The ABC perimeter also determines maximum rents and income limits for tenants.

However, spatial discontinuities in treatment are not suitable to identify an exogenous demand shock limited to a specific segment. Indeed, since the ABC perimeter simultaneously determines the treatment intensity for other housing policies (e.g. interest-free loan policy), the demand shock could affect several *a priori* defined segments.

To tackle this issue, we exploit the singularity in the definition of eligibility, which is limited to a specific tier in the ABC specification. In fact, B₂ municipalities (ineligible for the Pinel scheme since July 2013) can apply for an exception to remain eligible for the Pinel scheme. Municipalities must meet at least five out of ten criteria,⁹ such as demographic dynamics, housing price, rent level or construction activity. These criteria are aimed at municipalities that, according to policy-makers, are facing affordability problems within the B₂ level, introducing an additional distinction to the ABC perimeter. However, this sub-classification within the ABC is specific to the rental investment scheme, holding other policy treatments constant. Finally, the rental investment scheme has been unavailable to B₂ municipalities since January 2018, regardless of whether the municipality has benefited

⁷Only existing unhealthy dwellings are eligible to the scheme in exchange of massive renovation works, but we assume that it is marginal.

⁸The ABC perimeter is common to most housing policies and defines the tension of the housing market locally.

⁹We report the full list on Table 3.A.1 based on French *décret*.

from an exception. In total, 1,251 municipalities benefited from a derogation representing 29.7% of B_2 municipalities.

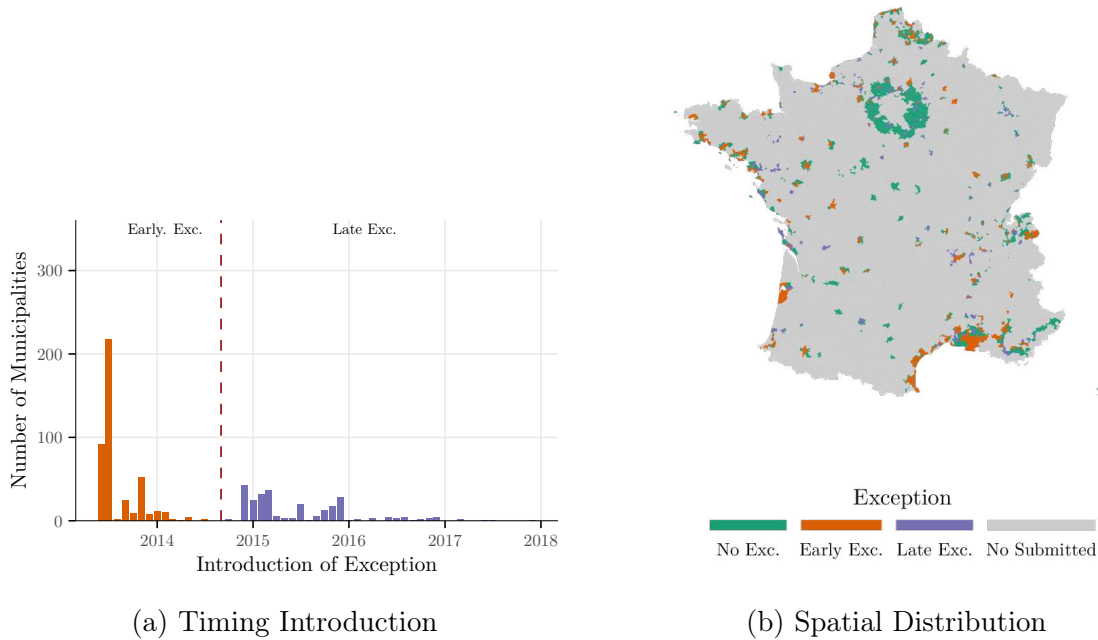
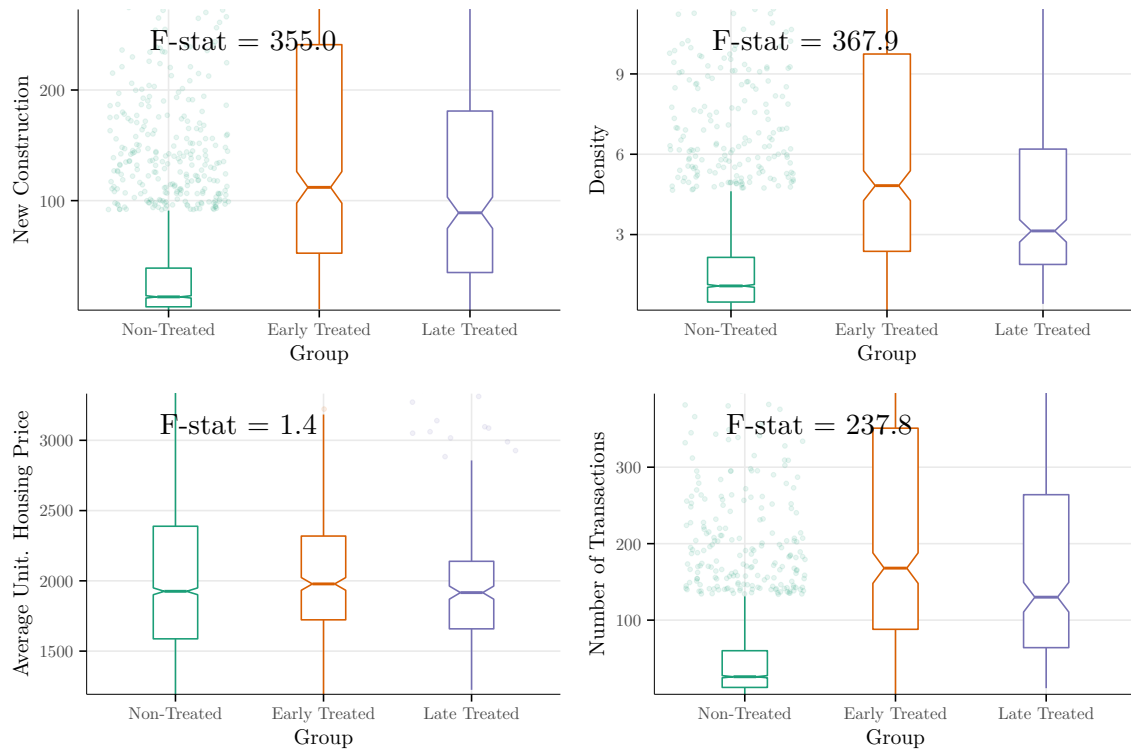


Figure 3.4.1: Temporal and Spatial Distribution of Exception Acceptance
Notes: We report the number of accepted exceptions per month in France (left panel). Exceptions make municipalities eligible to the Pinel scheme. Exceptions are restricted to B_2 municipalities. The right panel represents their location according to timing introduction. We also report non- B_2 for which exceptions are irrelevant.

As municipalities must apply to the administration for an exemption, exemptions are not granted at a similar time (Figure 3.4.1a). Yet, most exceptions occurred in July-August 2013 (45.6% of accepted exceptions, 571 municipalities). We assume that these municipalities remain eligible for the rental investment scheme without interruption, as all B_2 municipalities were eligible for the Pinel scheme before July 2013. After the 2014 reform of the ABC classification, further exceptions are less frequent. These later exceptions mainly concerned municipalities affected by the ABC reform and reclassified as B_2 (whereas they were previously classified as C or B_1). Nevertheless, the exemption process generates multiple changes in the eligibility status for the rental investment scheme within the B_2 tier. Firstly, most municipalities experienced a policy repeal in July 2013, as they do not benefit from an exception. Secondly, the staggered adoption of the derogation over the period 2014-2018 leads to the reintroduction of the policy. Finally, municipalities that benefited from the exception should have experienced a policy repeal in early 2018, while it was finally introduced in March 2019.

The process is endogenous to the characteristics of the municipalities, as observations receiving an exception are significantly different from other B_2 observations (Figure 3.4.2). This is consistent with the requirements to remain eligible. Although we cannot accurately

Figure 3.4.2: Descriptive Statistics About B₂ Municipalities

Notes: We report the distribution of the average per municipalities for the number of new constructions (2010–2012), population density (2013), average unitary housing price (2010–2012) and number of transactions (2010–2012). We classify each group according to whether municipalities benefit from an exception. In addition, we introduce a subcategory to distinguish early exceptions from late ones. Finally, we report F -value from ANOVA tests to test the equality of values between groups.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

collect individual criteria for each municipality due to data unavailability, we compare municipalities based on a set of pre-treatment variables (before July 2013) related to the housing market and demographics. Our data are derived from fiscal sources or from the population in January 2013.

On average, municipalities benefiting from an exception tend to have higher construction activity, higher density and a higher number of transactions (Figure 3.4.2). However, our groups do not differ according to the unitary housing price (F -test = 1.4). Thus, we cannot be sure that the exemption process is systematic. In addition, the timing of the exception acceptance differs according to the characteristics of the municipalities. Early exemptions concern municipalities with a higher density, number of buildings and transactions. Differences may be explained by the fact that late exemptions are granted to municipalities classified as C prior to the 2014 reform, which is the lowest level according to the ABC classification.

3.4.2 Strategy

Our identification strategy exploits these exceptions for B_2 municipalities to identify exogenous demand shock restricted to investment purpose. To deal with endogeneity from the exception process, we adopt a difference-in-differences (DiD) type design. Despite a large number of main policy changes and the staggered introduction of exemptions, there are few appropriate policy changes to induce a demand shock limited to specific segments. For example, the 2014 reform that shifts the municipalities of the tiers (and thus affects the eligibility for the rental investment scheme) also affects third housing policies that are likely to affect demand for other segments (e.g. French policy to support home ownership), leading to compound effects. Furthermore, the repeal of the Pinel scheme for all B_2 municipalities in 2018 was effectively introduced in March 2019. Early announcements are likely to affect both supply and demand, violating the necessary assumptions for segmentation identification (supply inelasticity).

Given the staggered introduction of the exemptions, we adopt two different strategies, exploiting two policy reforms. Firstly, we exploit the repeal of the rental investment policy in July 2013 for most B_2 municipalities, while municipalities benefiting from an exception remain eligible. The policy repeal is assumed to cause a negative demand shock. Secondly, we use the wave of exemptions that occurred in early 2015, which affected a sufficient number of municipalities. We expect our demand shock to be positive and limited to rental purposes, as it does not correspond to other policy changes in terms of introduction timing. Therefore, there is unlikely to be a simultaneous demand shock in other specific segments and any price adjustment would result from a lack of segmentation.

Despite the fact that our sample of municipalities is constant over time (current B_2 municipalities), we do not adopt pooled models for two reasons. Firstly, staggered DiD models require stricter assumptions such as the effect to be homogeneous over time and groups (de Chaisemartin and D'Haultfoeulle, 2018), likely to fail in our setting. For example, economic conditions have changed over time, particularly the cost of borrowing through mortgage interest rates. Pooled models also require always treated groups to be dropped to avoid negative weighting issues arising from multiple decomposition (Goodman-Bacon, 2021). Despite growing literature dealing with staggered DiD strategy (see for a review de Chaisemartin and D'Haultfoeulle, 2022), needed assumptions do not fit our identification strategy. restricting the period by conducting independent natural experiments meets the necessary assumptions for segmentation identification (Section 3.2).

3.4.3 Models Specification

The model specification is similar for the two natural experiments we consider in subsection 3.4.2. To make the best use of the presence of two potential groups (namely the always treated and never treated groups), we adopt a two-way fixed effects estimator. The model specification is

$$Y_{ijt}^{k\ell} = \alpha_t^{k\ell} + \beta_j^{k\ell} + \tau^{k\ell} W_{jt} + \varepsilon_{ijt}^{k\ell} \quad (3.1)$$

with $Y_{ijt}^{k\ell}$ is the dependent variable for transaction i belonging to group j at time t restricted to segment (k, ℓ) ; $\alpha_t^{k\ell}$ are time fixed effects; $\beta_j^{k\ell}$ are group fixed effects; W_{jt} indicates whether group j is treated at time t ; $\varepsilon_{ijt}^{k\ell}$ corresponds to idiosyncratic error term.

The parameter of interest to discuss price adjustment is $\tau^{k\ell}$ as it represents how dependent variable Y adjusts consecutive to the demand shock. As we assume that the rental investment scheme induces a demand shock limited to new rental housing, the price adjustment in other specific markets following the implementation of the financial incentive is captured by $\tau^{k\ell}$. The existence of segmentation can be concluded and discussed from the significance of the price adjustment in the unsubsidised segment.

To discuss whether financial incentives reallocate local demand or attract external demand (see section 3.2), we also assess the effect of financial incentives on the balance between local and distant investors. We then explain the likelihood that the buyer will be a local rather than a distant investor. We use the same design as in Equation 3.1

$$\log \left(\frac{p_{ijt}^{k\ell}}{1 - p_{ijt}^{k\ell}} \right) = \lambda_j^{k\ell} + \theta_t^{k\ell} + \gamma^{k\ell} W_{jt} + \xi_{it}^{k\ell} \quad (3.2)$$

where $p_{ijt}^{k\ell}$ is the probability for observation i to be purchased by local individuals at time t belonging to group j on segment (k, ℓ) ; $\lambda_j^{k\ell}$ corresponds to group fixed effects; $\theta_t^{k\ell}$ represents time fixed effects and W_{jt} is a dummy indicating whether group j is treated at time t ; $\xi_{it}^{k\ell}$ is the idiosyncratic term error. We estimate Equation 3.2 for both positive and negative demand shocks using similar identification strategies than previously.

Similarly to the linear specification, the parameter of interest is $\gamma^{k\ell}$ as it represents the shift in probability to have local investors rather than long-distance ones. Recall that the interest in understanding the balance between local and long-distance investors is relevant for understanding the potential price adjustment in integrated segments.

3.4.4 Empirical Specification

Our outcome of interest to discuss the price adjustment is the unitary housing price derived from the transaction dataset presented on Section 3.3. For the logistic regression, our outcome of interest is whether the purchaser is a local or a long-distance investor. Recall that we define empirically long-distance investors as purchasers living outside of commuting area prior to the purchase.

Although the TWFE estimator handles non-observable heterogeneity, we introduce commuting area dummies to account for sizeable cross-sectional differences in price between local housing markets. Our conclusions are nonetheless robust to the introduction of additional control variables to account for housing heterogeneity (see Section 3.5.4). Finally, we cluster standard errors to the commuting area to account for potential correlation in outcomes within clusters.

Table 3.4.1: Composition of Groups According to the Nature of the Demand Shock

Group	Positive Shock				Negative Shock			
	Existing		New Construction		Existing		New Construction	
	Own	Rent	Own	Rent	Own	Rent	Own	Rent
	Number of Transactions							
Never Treated	7,688	2,720	590	247	53,346	12,404	2,412	684
Shifter	24,712	5,701	1,099	457	5,251	1,361	265	118
Always Treated	42,132	16,534	3,253	2,906	39,916	13,898	2,276	2,370
Group	Number of Municipalities							
	Own	Rent	Own	Rent	Own	Rent	Own	Rent
	Own	Rent	Own	Rent	Own	Rent	Own	Rent
	Transaction per Municipality (Average)							
Never Treated	158	158	158	158	2,348	2,348	2,348	2,348
Shifter	1,901	1,901	1,901	1,901	72	72	72	72
Always Treated	457	457	457	457	312	312	312	312
Never Treated	48.7	17.2	3.7	1.6	22.7	5.3	1.0	0.3
Shifter	13.0	3.0	0.6	0.2	72.9	18.9	3.7	1.6
Always Treated	92.2	36.2	7.1	6.4	127.9	44.5	7.3	7.6

Notes: We report for each natural experiments we implement (either positive or negative demand shock), the composition of each group exploited in the two-way fixed effects specification. The top panel corresponds to the number of transactions, middle one to the number of municipalities with at least one transaction, while the bottom panel compute the average number of transactions per municipalities. For the positive shock, the time range from demand shock is set to 9 months (from February 2015), while it is 6 months for the negative shock (July 2013).

Sources: Data derived from DV3F and property tax files. Authors' calculation.

For each natural experiment, we take both the never treated and the always-treated groups as control groups. Indeed, for the positive demand shock, the always treated group corresponds to B₂ municipalities that benefit from an early exception and thus remain eligible for the Pinel scheme. Conversely, the never treated group corresponds to B₂ municipalities that never benefit from an exception. The shifter group is composed of municipalities that receive the exception. For the negative demand shock, the always

treated group is composed of municipalities that benefit from the exception in July and August 2013. The never treated group is composed of C municipalities that shift to the B_2 level after the 2014 reform. Then, the set of municipalities does not benefit from the Pinel scheme, while being comparable to the B_2 municipalities. Finally, the shifter group concerns B_2 that does not benefit from exception in mid-2013. We report the composition of each group in Table 3.4.1.

Our preferred specification for natural experiments to isolate negative shock is a time range of six months from the policy repeal, namely July 2013. For positive shocks, we adopt a different time range because our shifter group consists of fewer municipalities, resulting in fewer observations (Table 3.4.1). Consequently, we choose to extend the time range to nine months in order to collect enough observations and to obtain consistent estimates, especially for the smaller segments. However, our main results are robust to the specification of the time span (see Section 3.5.4).

3.5 Results

We organise our results section as follows. Firstly, we provide results assuming that the housing market as homogeneous according to tenure and construction type. Secondly, we relax the homogeneous assumption by running independent regressions on each housing market segment defined *a priori*. We provide pre-trend comparisons to discuss the credibility of the parallel trends assumption and additional robustness checks to conclude this section.

3.5.1 Results Under Homogeneity Assumption

Firstly, we estimate the price adjustment by assuming that housing markets are locally homogeneous, i.e. that there is no segmentation by tenure status and type of construction. Then, assuming perfect homogeneity of the housing market, we estimate the average price adjustment resulting from the demand shocks. We estimate Equation 3.1 and report results in Table 3.5.1 for both natural experiments. In addition, we estimate Equation 3.2 following the logit specification, to discuss the impact on the balance between local and long-distance investors.

We obtain mixed results with respect to the price adjustment in the overall housing market resulting from the demand shock. Indeed, we find that the introduction of the policy in specific B_2 municipalities increases the overall unit housing price by 2.9% (column 1, Table 3.5.1), while the policy repeal has no significant effect on housing price (columns 3 and 4, Table 3.5.1). However, despite the differences between positive and negative

Table 3.5.1: Results for Price Adjustment (OLS Estimation) and Probability to be Purchased by Local Investors (Logit) Under Homogeneous Assumption

Covariate	Dependent Variable: Unit. Price (<i>log</i>)				Dependent Variable: Prob. of Local Invest.			
	Positive Shock		Negative Shock		Positive Shock		Negative Shock	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Group (Shift)	-0.003 (0.031)	-0.019 (0.033)	0.018 (0.061)	0.030 (0.063)	-0.603*** (0.204)	-0.383** (0.159)	0.202 (0.269)	0.067 (0.193)
Group (Always)	-0.055** (0.026)	-0.067*** (0.025)	0.016 (0.064)	0.026 (0.063)	-0.201* (0.113)	-0.259*** (0.082)	0.134 (0.263)	0.007 (0.180)
Treatment Effect	0.029** (0.011)	0.030*** (0.011)	-0.005 (0.006)	-0.005 (0.006)	0.140** (0.055)	0.136** (0.058)	-0.083** (0.039)	-0.092** (0.038)
Monthly FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control	No	Yes	No	Yes	No	Yes	No	Yes
R ²	0.248	0.256	0.234	0.246	0.079	0.123	0.073	0.125
Adj. R ²	0.247	0.255	0.233	0.245	-	-	-	-
Log-Lik	-	-	-	-	-65,036	-61,937	-53,869	-50,829
N	134,301	134,301	108,039	108,039	134,301	134,301	108,039	108,039

Notes: We report regression results from the OLS estimation of Equation 3.1 and logistic regression of Equation 3.2 for positive and negative shocks. The housing market is assumed to be locally homogeneous. Observation unit is transaction. For positive shock, we restrict transaction from June 2014 to Nov. 2015. For negative shock, we restrict transaction from Jan. 2013 to Dec. 2013. Reference group is never treated. Time-fixed effects are specified to month. For the lack specification without control variables, we only control for commuting area heterogeneity. Standard errors are reported between parentheses and clustered to the commuting area, using the `fixest` R package. For control variables, we introduce distance to the Central Business District (exogenously defined), housing surface and housing type. Continuous variables

Sources: Data derived from DV3F and property tax files. Authors' calculation.

*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

shocks, our results confirm that the introduction of a rental investment programme is likely to increase the unitary price of housing locally, in line with Chapelle et al. (2018) on French policy (they found a 1% decrease). The introduction of control variables has little impact on the estimation intensity and standard errors.

Moreover, our results demonstrate that the financial incentive in rental markets favours short-distance purchasers. Indeed, the effect on the probability of being bought by local rather than long-distance investors follows a similar movement to that of the demand shock (columns 5 to 8, Table 3.5.1). Then, under homogeneous conditions, the financial incentive has mixed effects on price, while it appears to favour local demand.

3.5.2 Local Housing Market Segmentation

We now relax the homogeneous assumption and observe price movement consecutive to a demand shock as described in Section 3.2. Hence, we estimate Equation 3.1 and Equation 3.2 using respectively OLS and logistic regression and report results on Table 3.5.2. The left panel corresponds to positive demand shock (columns 1 to 4), while the right panel corresponds to negative demand shock (columns 5 to 8). We introduce monthly dummies for time fixed effects. Finally, we report robust standard errors clustered at the commuting area level.

Table 3.5.2: Results from TWFE Regressions According to Demand Shock Nature and Housing Market Segments

Covariates	<i>Dependent Variable: Unit. Housing Price (log)</i>							
	Positive Shock				Negative Shock			
	Existing		New Construction		Existing		New Construction	
	Own.	Rent	Own.	Rent	Own.	Rent	Own.	Rent
Group (Shift)	0.002 (0.022)	-0.034 (0.047)	0.164** (0.064)	0.072 (0.081)	0.021 (0.055)	0.023 (0.065)	-0.056 (0.076)	-0.022 (0.064)
Group (Always)	-0.045* (0.023)	-0.121*** (0.032)	0.128*** (0.049)	0.154* (0.093)	-0.005 (0.057)	0.010 (0.067)	0.121 (0.076)	0.081 (0.067)
Treatment Effect	0.012 (0.011)	0.064*** (0.020)	0.006 (0.036)	0.117 (0.075)	-0.008 (0.006)	-0.002 (0.013)	-0.013 (0.030)	-0.135*** (0.037)
Monthly FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R ²	0.239	0.263	0.256	0.374	0.232	0.240	0.268	0.348
N	98,403	27,651	4,948	3,171	73,794	24,787	4,901	3,597

Notes: We report results from the estimation through OLS of Equation 3.1 to explain unitary housing price. We relax homogeneous assumptions, and run independent regression according to nature of demand shock (columns 1 to 4 corresponds to positive demand shock, columns 5 to 8 corresponds to negative shock) and the housing market segment. Reference group is never treated. Standard errors are reported between parentheses, and clustered to commuting area using the `fixest` R package. Observation unit is transaction. For positive shock, we restrict transaction from June 2014 to Nov. 2015. For negative shock, we restrict transaction from Jan. 2013 to Dec. 2013. We do not report time fixed effects and commuting areas fixed effects for clarity reasons, but results are available on Figure 3.A.2 and Figure 3.A.4.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The relaxation of the homogeneity assumption for local housing markets yields to heterogeneous price capitalisation resulting from demand-side policies. Indeed, the policy repeal causes a 13.5% decrease in unitary housing price in the subsidised segment. Conversely, policy introduction does not lead to a significant price increase, although the estimated effect is close to that of policy repeal (+11.7%, p -value = 0.119). The estimated effects are stronger than those obtained under the homogeneous assumption. Moreover, the policy repeal has a significant impact on the balance between local and distant investors (Table 3.5.3, last column). Indeed, in contrast to the pooled models, the removal of the financial incentive appears to significantly increase the share of local investors in the subsidised segments. The policy therefore tends to stimulate demand from long-distant investors.

Apart from direct price capitalisation in the subsidised segment, the demand shock does not affect the owner-occupied segments. Indeed, we do not recover any significant price adjustment following demand shocks, either for existing or new housing (Table 3.5.2, columns 1, 3, 5 and 7). We interpret these results, in line with our predictions in Section 3.2, as a perfect segmentation by tenure status. This corresponds to a lack of substitutability of housing assets between owner occupiers and rental investment characteristics. While we cannot discuss the underlying mechanism to explain the segmentation, heterogeneous preferences resulting from greater indifference to local amenities or housing characteristics

for investment purposes could be a cause.

Conversely, the housing market is not perfectly segmented by type of construction. In fact, we recover a price adjustment following the introduction of financial incentives for the purchase of existing housing. Recall that this segment is excluded from the rental investment scheme. Existing housing is therefore a substitute for new housing for rental purchase. The price adjustment is positive (+6.4% Table 3.5.2, column 2) and indicates a shift from the subsidised to unsubsidised segment.

Table 3.5.3: Results from Logistic Regression to Explain Probability to Be Purchased by Local Investors

Covariates	<i>Dependent Variable: Unit. Housing Price (log)</i>							
	Positive Shock				Negative Shock			
	Existing		New Construction		Existing		New Construction	
	Own.	Rent	Own.	Rent	Own.	Rent	Own.	Rent
Group (Shift)	-0.675*** (0.225)	-0.196 (0.158)	-0.718* (0.417)	0.294 (0.483)	0.036 (0.461)	0.075 (0.148)	0.418 (0.406)	0.058 (0.480)
Group (Always)	0.127 (0.137)	-0.209 (0.129)	0.044 (0.453)	-0.552 (0.588)	0.379 (0.458)	-0.073 (0.147)	0.108 (0.453)	0.322 (0.513)
Treatment Effect	0.225*** (0.073)	-0.041 (0.103)	0.029 (0.414)	0.115 (0.605)	-0.086 (0.054)	-0.002 (0.056)	-0.345 (0.254)	1.036*** (0.277)
Monthly FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log-Lik	-36,134	-18,048	-2,207	-1,898	-26,671	-16,375	-2,050	-2,219
N	98,403	27,651	4,948	3,171	73,794	24,787	4,901	3,597

Notes: We report results from the estimation of Equation 3.2 to explain the probability to be purchased by local investors. We relax the homogeneous assumption, and run independent regression according to nature of demand shock (columns 1 to 4 corresponds to positive demand shock, columns 5 to 8 corresponds to negative shock) and the housing market segment. Reference group is never treated. Standard errors are reported between parentheses, and clustered to commuting area using the `fixest` R package. Observation unit is transaction. For positive shock, we restrict transaction from June 2014 to Nov. 2015. For negative shock, we restrict transaction from Jan. 2013 to Dec. 2013. We do not report time fixed effects and commuting areas fixed effects for clarity reasons, but results are available on Figure 3.A.6 and Figure 3.A.8.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

However, the shift is not driven by additional long-distance investment. In fact, while we find a significant effect of policy repeal on the balance between local and long-distance purchasers, we do not find a significant effect for policy introduction in the segment of interest.¹⁰ Therefore, we cannot conclude that the outpricing process is driven by additional demand for long-distance investors. An orthogonal unobservable variable could be responsible for the shift.

Considering the policy design, the subsidy is particularly attractive for those with the highest incomes. Indeed, in order to benefit from the full subsidy, expected taxable income must be higher than the tax cut for at least 6 years. Therefore, consecutive

¹⁰We recover a positive effect for owner-occupied purchases in existing housing, but the parallel trend assumption is then likely to fail (see subsection 3.5.4)

to the policy implementation, the highest-income households (including the local ones) might shift their investment from financial markets to housing. However, as we do not observe either purchasers' income or financial market activity, we cannot empirically test our suggestion.

Finally, demand-side policies that support rental investment purchases to increase the supply of affordable housing have inflationary effects on housing markets. However, we specify that these effects are mainly restricted by policy to subsidised segments (namely rental investment in new housing), although we also recover an outpricing process. However, we do not find a significant effect for owner-occupiers, implying that tenure status appears to be structural for housing market segmentation. A lack of substitutability due to heterogeneous preferences between owner occupiers and investors, resulting from differentiated utility derived from purchase, could explain the segmentation. The underlying presence of tenure segmentation suggests that there are no short-term pitfalls in the joint implementation of housing policies that support both tenure statuses.

3.5.3 Falsification Test

The TWFE estimator we adopt requires the parallel trend assumption to hold to provide consistent and unbiased estimator. Although the assumption is by definition non-verifiable (Roth, 2022), we perform a pre-trend comparison to assess its credibility. Despite the fact that it does not ensure that the shifter group would have evolved in a similar pattern than control groups in absence of treatment, it enables to identify potential issues. Then, we falsified the date for the introduction of the treatment for the shifter group. The chosen period was set from three to eight months prior to the current breakdown point.

We therefore estimate Equation 3.1 and Equation 3.2 with similar estimation procedure than main results except that the introduction date is falsified. Hence, if pre-trends are similar, our parameter of interest should not be significant. Recall that it is a necessary but not a sufficient condition for parallel trends assumption to hold. We report the average effect and the number of significant estimation at a 95% confidence level (Table 3.5.4), using standard errors clustered at the commuting area.

The pre-trend comparison is consistent with the credibility of the parallel trends assumption. Indeed, except for one segment in the logistic regression design, all estimated effects are not significant. We can therefore assume that the parallel trends assumption is likely to hold.

The violation of the parallel trends assumption is likely to lead to unexpected results on the shift in the probability of local investors purchasing for owner occupation. Therefore,

Table 3.5.4: Results From TWFE With Falsified Introduction Timing

	Positive Shock				Negative Shock			
	Existing		New Construction		Existing		New Construction	
	Own	Rent	Own	Rent	Own	Rent	Own	Rent
	<i>Dependent Variable: Unit Housing Price</i>							
Mean Estimate	-0.010	-0.014	0.030	0.013	0.004	-0.007	-0.019	-0.017
Significance	0/6	0/6	0/6	0/6	0/6	0/6	0/6	1/6
	<i>Dependent Variable: Probability to be Local Investor</i>							
Mean Estimate	-0.002	0.022	-0.193	0.151	0.041	-0.046	0.556	-0.020
Significance	1/6	0/6	0/6	0/6	1/6	0/6	3/6	0/6

Notes: The top panel represents results for the OLS estimation of Equation 3.1 with unitary housing price as dependent variable. The bottom panel provide results about the logistic regression to explain the probability to be purchased by local investors rather than long-distance ones. We distinguish results according to the demand shock nature, and housing market segments. However, the policy breakdown introduction is falsified (6-months earlier). We report group dummies (the reference category is the never treated group) in addition to the parameter of interest (treatment effect). Local investors are empirically defined by purchasers originates from the same commuting area as the housing. Standard errors are clustered to commuting area using the `fixest` R package and reported between parentheses.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

*** p < 0.01, ** p < 0.05 * p < 0.1

we cannot rule out that the introduction of financial incentives has no effect on the balance between local and distant owner occupiers, which is close to our expectations.

3.5.4 Robustness Checks

To conclude our results section, we assess the robustness of our conclusions to the time range adopted from the exploitation and to the introduction of control variables to account for potential heterogeneity in characteristics. We introduce variation in the time span assumed from the introduction, from three to twelve months after the shock under consideration. In addition, we introduce control variables in the model specification to determine whether our results are driven by heterogeneous housing characteristics. The set of control variables includes exogenously defined distance from the CBD, housing area and housing type. We test two functional forms for continuous variables, namely linear and spline transformations, using an endogenous definition of the effective degree of freedom (Wood, 2017).

Considering the number of models under consideration, we report for each segment and nature of the shock, the number of times the parameter of interest is significant at a 90% confidence interval level. Standard errors are clustered at the commuting area. We also report the average estimation derived from the ten estimations and introduce a distinction according to the specification of control variables.

Firstly, our main results on price adjustment are robust to the time range specification. In fact, the price adjustment due to the positive demand shock for existing rented housing

Table 3.5.5: Robustness Checks From TWFE Results

Type	Positive Shock				Negative Shock			
	Existing		New Construction		Existing		New Construction	
	Own.	Rent	Own.	Rent	Own.	Rent	Own.	Rent
Spline Transformations	0.014	0.069	0.009	0.078	-0.006	-0.010	-0.011	-0.094
	0/10	10/10	0/10	4/10	1/10	2/10	1/10	10/10
Linear	0.016	0.067	0.011	0.077	-0.007	-0.010	-0.009	-0.102
	1/10	10/10	1/10	3/10	1/10	1/10	0/10	9/10
Benchmark	0.013	0.066	0.007	0.092	-0.007	-0.009	-0.008	-0.126
	0/10	10/10	0/10	3/10	1/10	1/10	0/10	9/10

Notes: We introduce variation in the period range from the policy reform (either repeal and introduction) from 3 months to 12 months. We report mean estimate from 10 individual estimation, and report the number of significant estimation for a 90% confidence interval.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

is significant (10/10, average effect to 0.066), as well as for the negative demand shock in the subsidised segment (9/10, average effect to -0.126). The price adjustment due to the positive demand shock in the subsidised segment remains insignificant, with only two-time ranges yielding significant results at the 90% confidence interval. These estimates mainly concern the largest time ranges, which may confirm sample size problems or lags in price capitalisation (see details on Section 3.A.12). Secondly, the introduction of control variables does not affect our conclusions, although it does affect the intensity of the price adjustment. Then, our results and derived conclusions are robust to the specification of the time span and to the heterogeneity of housing characteristics.

3.6 Conclusion

While the affordability issue appears to mainly relate to supply mechanisms, most housing policies focus on the demand side. As a result, it yields price capitalisation in the short term, which is contrary to the pursued objectives. However, given that housing policies commonly assumed tenure segmentation for local housing markets, we discuss whether it holds. From the nature of price capitalisation (homogeneous or heterogeneous), we infer local housing market segmentation by tenure status and type of construction. We exploit exogenous demand shocks restricted to a specific segment resulting from the French rental investment scheme.

Regardless of the nature of the demand shock, segments defined by owner-occupier purposes remain unaffected, due to strong local segmentation by tenure status. It therefore results from pronounced segmentation according to tenure within local housing markets. Moreover, our results confirm the negative short-term impact on housing prices following the implementation of demand-side policies. The price effect is in fact limited to specific housing segments defined by rental purpose with larger effects (-13.5% for policy repeal,

insignificant under homogeneous assumption). We estimate that the price capitalisation is at least partly caused by long-distance demand. Yet, it raises doubts about the efficiency of the policy, as the price capitalisation represents nearly two thirds of the average subsidy (18% of the purchase for 9-year commitment). In addition, we highlight that financial incentives lead to an outpricing process, as the demand for rental purposes in existing housing, albeit unsubsidised, increases following the demand shock. We consider that it results from the higher substitutability by type of construction. However, the outpricing process is not driven by additional long-distance investment, and income sorting may be responsible in terms of policy design.

The presence of a pronounced segmentation by tenure status is highly relevant from a policy perspective. Indeed, we cannot reject that the joint support of both objectives (homeownership and rental objectives) by different policies has negative effects in the short run. It also suggests that there are differences in characteristics according to tenure status, which may explain the lack of substitutability. Although this is the result of purchaser preferences, it raises the question of whether the units placed on the rental market meet the needs of tenants.

Whereas we find no evidence of adverse effects between housing policies in the short run, our identification strategy does not allow us to discuss the long-term effects. For example, increased competition for land in a context favourable to land restrictions could lead to price capitalisation for new housing, regardless of tenure status. Moreover, our results are only valid for a specific set of municipalities classified as intermediate in terms of housing market tension. Therefore, it remains to be discussed whether the segmentation of the housing market depends on local characteristics.

References

- Albouy, David, Gabriel Ehrlich, and Yingyi Liu (2016), “Housing Demand, Cost-of-Living Inequality, and the Affordability Crisis”. *NBER Working Paper* 22816, p. 62.
- Anenberg, Elliot and Edward Kung (2020), “Can More Housing Supply Solve the Affordability Crisis? Evidence from a Neighborhood Choice Model”. *Regional Science and Urban Economics* 80.103363.
- Baum-Snow, Nathaniel and Justin Marion (2009), “The Effects of Low Income Housing Tax Credit Developments on Neighborhoods”. *Journal of Public Economics* 93.5-6, pp. 654–666.
- Behrens, Kristian and Frédéric Robert-Nicoud (2015), “Agglomeration Theory with Heterogeneous Agents”. *Handbook of Regional and Urban Economics, Vol 5a*. Boston, MA: Elsevier.
- Bhattacharjee, Arnab, Eduardo Castro, Taps Maiti, and João Marques (2016), “Endogenous Spatial Regression and Delineation of Submarkets: A New Framework with Application to Housing Markets”. *Journal of Applied Econometrics* 31.1, pp. 32–57.
- Bono, Pierre-Henri and Alain Trannoy (2019), “The Impact of the ‘Scellier’ Income Tax Relief on Building Land Prices in France”. *Economie et Statistique / Economics and Statistics* 507d, pp. 91–114.
- Bourassa, Steven C., Martijn Dröes, and Martin Hoesli (2021), “Hedonic Models and Market Segmentation”. *SSRN Electronic Journal* 21-62.
- Brueckner, Jan K and Stuart S Rosenthal (2009), “Gentrification and Neighborhood Housing Cycles: Will America’s Future Downtowns Be Rich?” *Review of Economics and Statistics* 91.4, pp. 725–743.
- Chapelle, Guillaume, Benjamin Vignolles, and Clara Wolf (2018), “Impact of a Housing Tax Credit on Local Housing Markets: Evidence from France”. *Annals of Economics and Statistics* 130, pp. 101–131.
- Coën, Alain, Alexis Pourcelot, and Richard Malle (2022), “Macroeconomic Shocks and Ripple Effects in the Greater Paris Metropolis”. *Journal of Housing Economics* 56.101823.
- de Chaisemartin, Clément and Xavier D’Haultfœuille (2018), “Fuzzy Differences-in-Differences”. *The Review of Economic Studies* 85.2, pp. 999–1028.
- de Chaisemartin, Clément and Xavier D’Haultfœuille (2022), “Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey”. *The Econometrics Journal*.
- Diamond, Rebecca and Cecile Gaubert (2022), “Spatial Sorting and Inequality”. *Annual Review of Economics* 14.1, pp. 795–819.
- Duranton, Gilles and Diego Puga (2015), “Urban Land Use”. *Handbook of Regional and Urban Economics*. Vol. 5A. Boston, MA: Elsevier, p. 698.

- Dustmann, Christian, Bernd Fitzenberger, and Markus Zimmermann (2022), “Housing Expenditure and Income Inequality”. *The Economic Journal* 132.645, pp. 1709–1736.
- Eriksen, Michael D. and Stuart S. Rosenthal (2010), “Crowd out Effects of Place-Based Subsidized Rental Housing: New Evidence from the LIHTC Program”. *Journal of Public Economics* 94.11-12, pp. 953–966.
- Etheridge, Ben (2019), “House Prices and Consumption Inequality”. *International Economic Review* 60.4, pp. 1781–1822.
- Glaeser, Edward L. and Joseph Gyourko (2018), “The Economic Implications of Housing Supply”. *Journal of Economic Perspectives* 32.1, pp. 3–30.
- Glaeser, Edward L. and Jesse M. Shapiro (2003), “The Benefits of the Home Mortgage Interest Deduction”. NBER Chapters. National Bureau of Economic Research, Inc, pp. 37–82.
- Goodman, Allen C. and Thomas G. Thibodeau (1998), “Housing Market Segmentation”. *Journal of Housing Economics* 7.2, pp. 121–143.
- Goodman, Allen C. and Thomas G. Thibodeau (2007), “The Spatial Proximity of Metropolitan Area Housing Submarkets”. *Real Estate Economics* 35.2, pp. 209–232.
- Goodman-Bacon, Andrew (2021), “Difference-in-Differences with Variation in Treatment Timing”. *Journal of Econometrics* 225.2, pp. 254–277.
- Gruber, Jonathan, Amalie Jensen, and Henrik Kleven (2021), “Do People Respond to the Mortgage Interest Deduction? Quasi-experimental Evidence from Denmark”. *American Economic Journal: Economic Policy* 13.2, pp. 273–303.
- Gyourko, Joseph and Raven Molloy (2015), “Regulation and Housing Supply”. *Handbook of Regional and Urban Economics, Vol 5b*. Boston, MA: Elsevier.
- Ha, Sejeong, Christian A.L. Hilber, and Olivier Schöni (2021), “Do Long-Distance Moves Discourage Homeownership? Evidence from England”. *Journal of Housing Economics* 53.101766.
- Hilber, Christian A. L. and Wouter Vermeulen (2016), “The Impact of Supply Constraints on House Prices in England”. *The Economic Journal* 126.591, pp. 358–405.
- Hwang, Min and John M. Quigley (2006), “Economic Fundamentals in Local Housing Markets: Evidence from U.S. Metropolitan Regions”. *Journal of Regional Science* 46.3, pp. 425–453.
- Keele, Luke and Rocío Titiunik (2015), “Geographic Boundaries as Regression Discontinuities”. *Political Analysis* 23.1, pp. 127–155.
- Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins (2013), “The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets”. *Journal of Economic Literature* 51.4, pp. 1007–1062.

- Malpezzi, Stephen and Kerry Vandell (2002), “Does the Low-Income Housing Tax Credit Increase the Supply of Housing?” *Journal of Housing Economics* 11.4, pp. 360–380.
- Molloy, Raven (2020), “The Effect of Housing Supply Regulation on Housing Affordability: A Review”. *Regional Science and Urban Economics* 80.103350.
- Piazzesi, Monika, Martin Schneider, and Johannes Stroebe (2020), “Segmented Housing Search”. *American Economic Review* 110.3, pp. 720–759.
- Roth, Jonathan (2022), “Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends”. *American Economic Review: Insights* 4.3, pp. 305–322.
- Rothenberg, Jerome, ed. (1991), *The Maze of Urban Housing Markets: Theory, Evidence, and Policy*. Chicago: University of Chicago Press.
- Saiz, Albert (2010), “The Geographic Determinants of Housing Supply”. *Quarterly Journal of Economics* 125.3, pp. 1253–1296.
- Szumilo, Nikodem (2019), “The Spatial Consequences of the Housing Affordability Crisis in England”. *Environment and Planning A: Economy and Space* 51.6, pp. 1264–1286.
- Turner, Matthew, Andrew Haughwout, and Wilbert Van der Klauw (2014), “Land Use Regulation and Welfare”. *Econometrica* 82.4, pp. 1341–1403.
- Waxman, Andrew, Yuanning Liang, Shanjun Li, Panle Jia Barwick, and Meng Zhao (2020), “Tightening Belts to Buy a Home: Consumption Responses to Rising Housing Prices in Urban China”. *Journal of Urban Economics* 115.103190.
- Wheaton, William C. (1990), “Vacancy, Search, and Prices in a Housing Market Matching Model”. *Journal of Political Economy* 98.6, pp. 1270–1292.
- Wijburg, Gertjan (2022), “Tax-Incentivized Housing Production and the Affordability Crisis: International Lessons from the Low-Income Housing Tax Credit Program in the United States”. *Housing Studies* 0.0, pp. 1–26.
- Wood, Simon N. (2017), *Generalized Additive Models: An Introduction with R*. Second. Chapman and Hall/CRC.

3.A Additional Material

3.A.1 Criterion for Municipalities to Obtain Exception

Table 3.A.1: Requirement for Municipalities to Obtain Exception

Topic	Data Sources	Require Criterion
Share of household with rent higher than 39% of income of	CAF*	Higher than 15%
Mobility rate in social housing	Housing minister survey*	Lower than 15%
Average unitary housing price for existing apartments	Perval Data ⁺	Higher than 1.75k
Number of transactions	Perval Data ⁺	Higher than 10.25 per 1k inhabitants
Average unitary housing price for newly built apartments	Commercialisation survey ⁺	Higher than 2,075 euros
Average unitary rent	Clameur Data *	Higher than 7.3 euros per months
Population evolution	INSEE	Higher than 0.4%
Vacancy rate	INSEE	Lower than 9%
Average annual new construction over a three year period	SITADEL	Higher than 200 per year
Average of all criterion	Multi-data	Must be higher than 25/100

Notes: We report criterion municipalities must comply to obtain exception for rental investment scheme.

3.A.2 Descriptive Statistics for Transactions Selected for Negative Demand Shock (6-month range)

Table 3.A.2: Descriptive Statistics for Transactions Selected for Repeal Natural Experiments

Group	Variable	Metrics	Existing		Newly Built	
			Rent	Own.	Rent	Own.
Always Treated	N		2,720	7,688	247	590
	Narrow Purch.	Mean	0.61	0.86	0.54	0.78
	Unit. Price	Mean	1,632	1,788	2,523	2,326
		Median	1,604	1,750	2,576	2,318
		Std. Dev.	(534)	(569)	(683)	(676)
	Price	Mean	116,691	154,452	157,619	169,335
		Median	100,000	141,000	154,000	170,000
		Std. Dev.	(71,360)	(79,716)	(59,653)	(70,420)
Shifters	N		5,701	24,712	457	1,099
	Narrow Purch.	Mean	0.64	0.86	0.52	0.83
	Unit. Price	Mean	1,535	1,765	2,386	2,184
		Median	1,483	1,699	2,340	2,148
		Std. Dev.	(600)	(646)	(748)	(727)
	Price	Mean	126,752	165,979	151,141	182,646
		Median	112,000	155,000	148,669	175,000
		Std. Dev.	(80,818)	(82,289)	(55,932)	(75,359)
Never Treated	N		16,534	42,132	2,906	3,253
	Narrow Purch.	Mean	0.57	0.84	0.40	0.78
	Unit. Price	Mean	1,766	1,925	2,960	2,690
		Median	1,705	1,850	2,957	2,667
		Std. Dev.	(652)	(700)	(618)	(722)
	Price	Mean	110,681	158,306	156,425	192,063
		Median	90,000	140,000	149,000	174,700
		Std. Dev.	(78,752)	(90,386)	(57,723)	(81,859)

Notes: We report for each group (never treated, shifters, always treated) descriptive statistics about transactions. We select transactions six months from the policy repeal, i.e. from January 2013 to December 2013. We detail the number of transactions, the share of local purchasers, the average unitary housing price, and the average price. We also report for relevant metrics standard deviation between parentheses. We also introduce median. All these statistics are provided according to our segment definition (columns).

Sources: Data derived from DV3F and property tax files. Authors' calculation.

3.A.3 Descriptive Statistics for Transactions Selected for Positive Demand Shock (9-month range)

Table 3.A.3: Descriptive Statistics for Transactions Selected for Introduction Natural Experiments

Group	Variable	Metrics	Existing		Newly Built	
			Rent	Own.	Rent	Own.
Always Treated	N		12,404	53,346	684	2,412
	Narrow Purch.	Mean	0.62	0.85	0.53	0.78
	Unit. Price	Mean	1,582	1,800	2,524	2,405
		Median	1,519	1,731	2,496	2,271
		Std. Dev.	(644)	(680)	(879)	(878)
	Price	Mean	130,353	167,979	168,999	193,890
		Median	115,000	155,000	158,450	179,000
		Std. Dev.	(84,099)	(85,999)	(80,182)	(89,873)
Shifters	N		1,361	5,251	118	265
	Narrow Purch.	Mean	0.59	0.74	0.65	0.74
	Unit. Price	Mean	1,540	1,814	2,504	2,498
		Median	1,481	1,749	2,499	2,476
		Std. Dev.	(563)	(661)	(575)	(675)
	Price	Mean	115,272	150,000	169,032	194,365
		Median	100,000	136,000	169,162	181,350
		Std. Dev.	(73,315)	(84,556)	(49,873)	(74,448)
Never Treated	N		13,898	39,916	2,370	2,276
	Narrow Purch.	Mean	0.58	0.84	0.41	0.76
	Unit. Price	Mean	1,541	1,738	2,926	2,608
		Median	1,500	1,683	2,943	2,605
		Std. Dev.	(569)	(627)	(558)	(764)
	Price	Mean	105,430	151,740	160,337	195,267
		Median	85,000	135,000	154,831	176,773
		Std. Dev.	(73,861)	(85,998)	(51,122)	(88,175)

Notes: We report for each group (never treated, shifters, always treated) descriptive statistics about transactions. We select transactions six months from the policy introduction, i.e. from June 2014 to November 2013. We detail the number of transactions, the share of local purchasers, the average unitary housing price, and the average price. We also report for relevant metrics standard deviation between parentheses. We also introduce median. All these statistics are provided according to our segment definition (columns).

Sources: Data derived from DV3F and property tax files. Authors' calculation.

3.A.4 Commuting Area Fixed Effects for Positive Demand Shock to Explain Unitary Housing Price

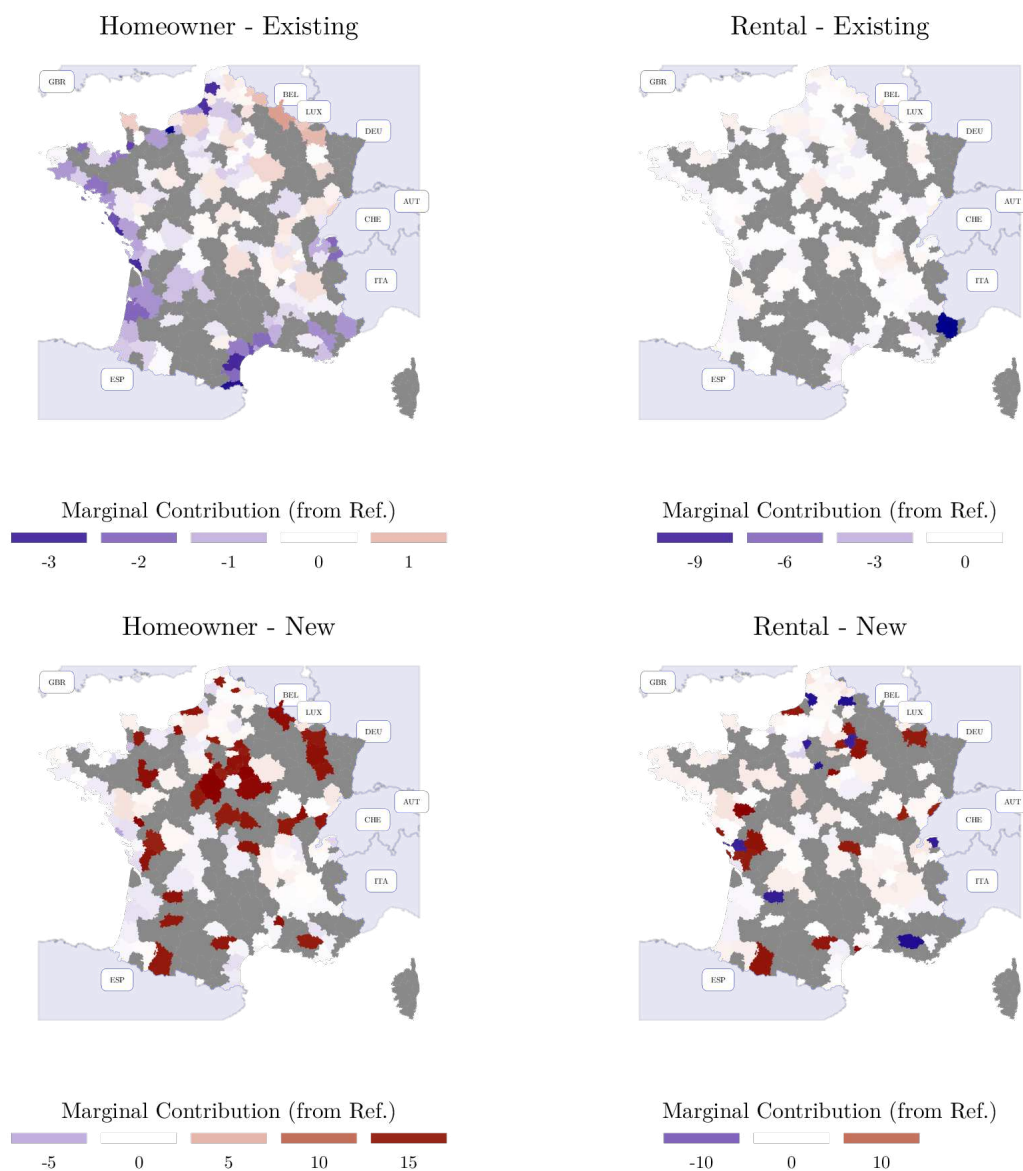


Figure 3.A.1: Commuting Area Fixed Effects for Positive Demand Shock to Explain Unitary Housing Price

Notes: We report marginal contributions for commuting area fixed effects resulting from the estimation of Equation 3.1 for the positive demand shock experiment. Each map corresponds to the four estimations we make based on the *a priori* defined market segment. The model is estimated with OLS and standard errors are clustered to the commuting area.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

3.A.5 Time Fixed Effects for Positive Demand Shock to Explain Unitary Housing Price

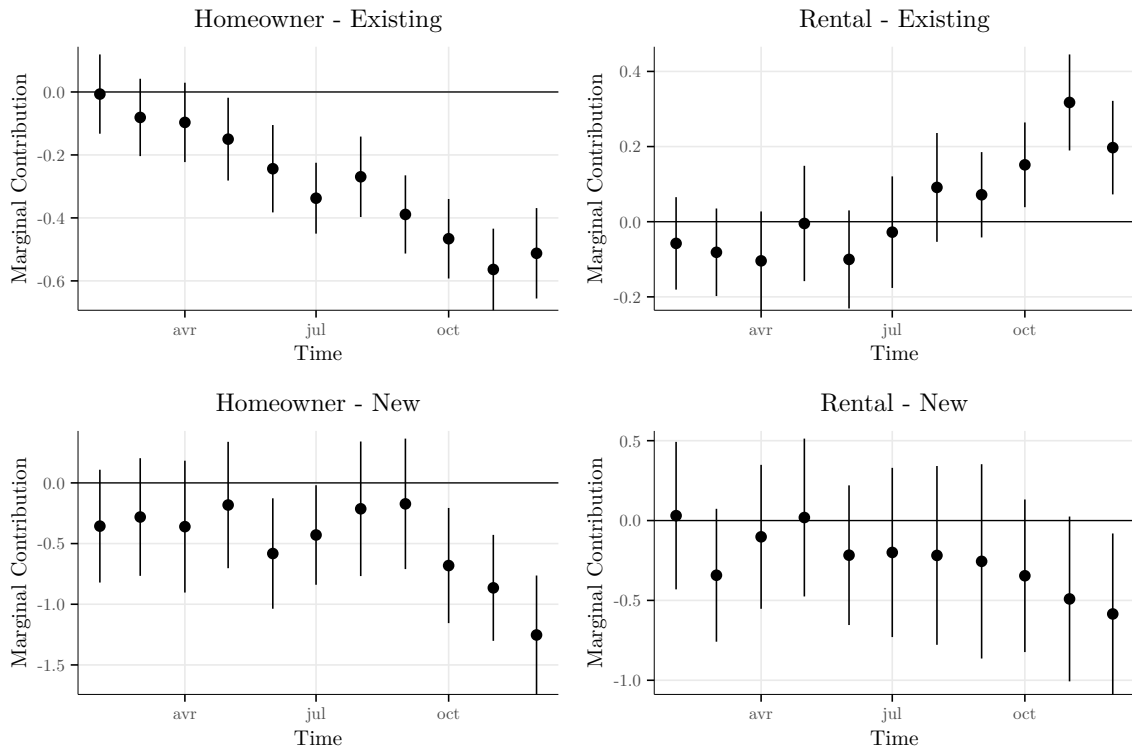


Figure 3.A.2: Time Fixed Effects for Positive Demand Shock to Explain Unitary Housing Price

Notes: We report marginal contributions for time fixed effects resulting from the estimation of Equation 3.1 for the positive demand shock experiment. Each plot corresponds to the four estimations we make based on the *a priori* defined market segment. The model is estimated with OLS and standard errors are clustered to the commuting area. We report confidence intervals to 95% level. The model specification contains commuting areas fixed effects.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

3.A.6 Commuting Area Fixed Effects for Negative Demand Shock to Explain Unitary Housing Price

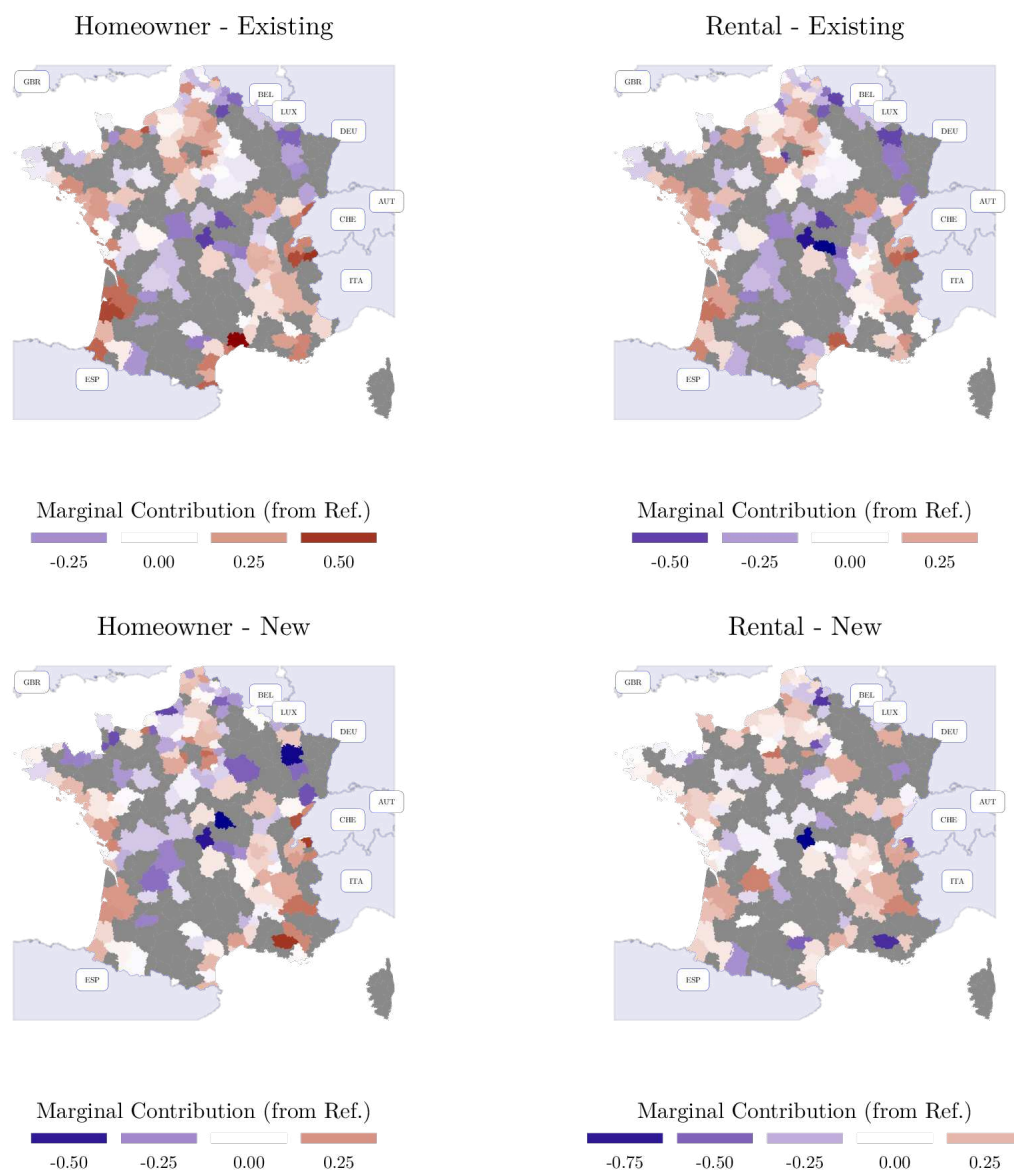


Figure 3.A.3: Commuting Area Fixed Effects for Negative Demand Shock to Explain Unitary Housing Price

Notes: We report marginal contributions for commuting area fixed effects resulting from the estimation of Equation 3.1 for the negative demand shock experiment. Each map corresponds to the four estimations we make based on the *a priori* defined market segment. The model is estimated with OLS and standard errors are clustered to the commuting area.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

3.A.7 Time Fixed Effects for Positive Demand Shock to Explain Unitary Housing Price

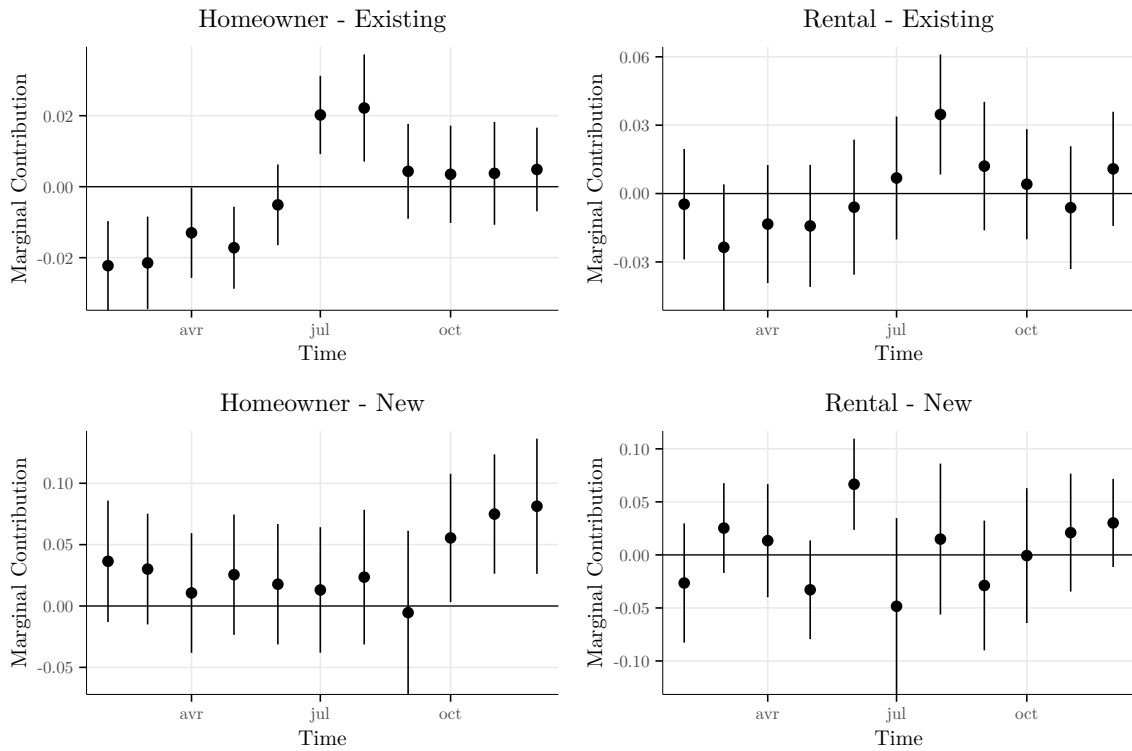


Figure 3.A.4: Time Fixed Effects for Negative Demand Shock to Explain Unitary Housing Price

Notes: We report marginal contributions for time fixed effects resulting from the estimation of Equation 3.1 for the negative demand shock experiment. Each plot corresponds to the four estimations we make based on the *a priori* defined market segment. The model is estimated with OLS and standard errors are clustered to the commuting area. We report confidence intervals to 95% level. The model specification contains commuting areas fixed effects.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

3.A.8 Commuting Area Fixed Effects for Positive Demand Shock to Explain Probability to be Purchased by Local Investors using Logistic Regression

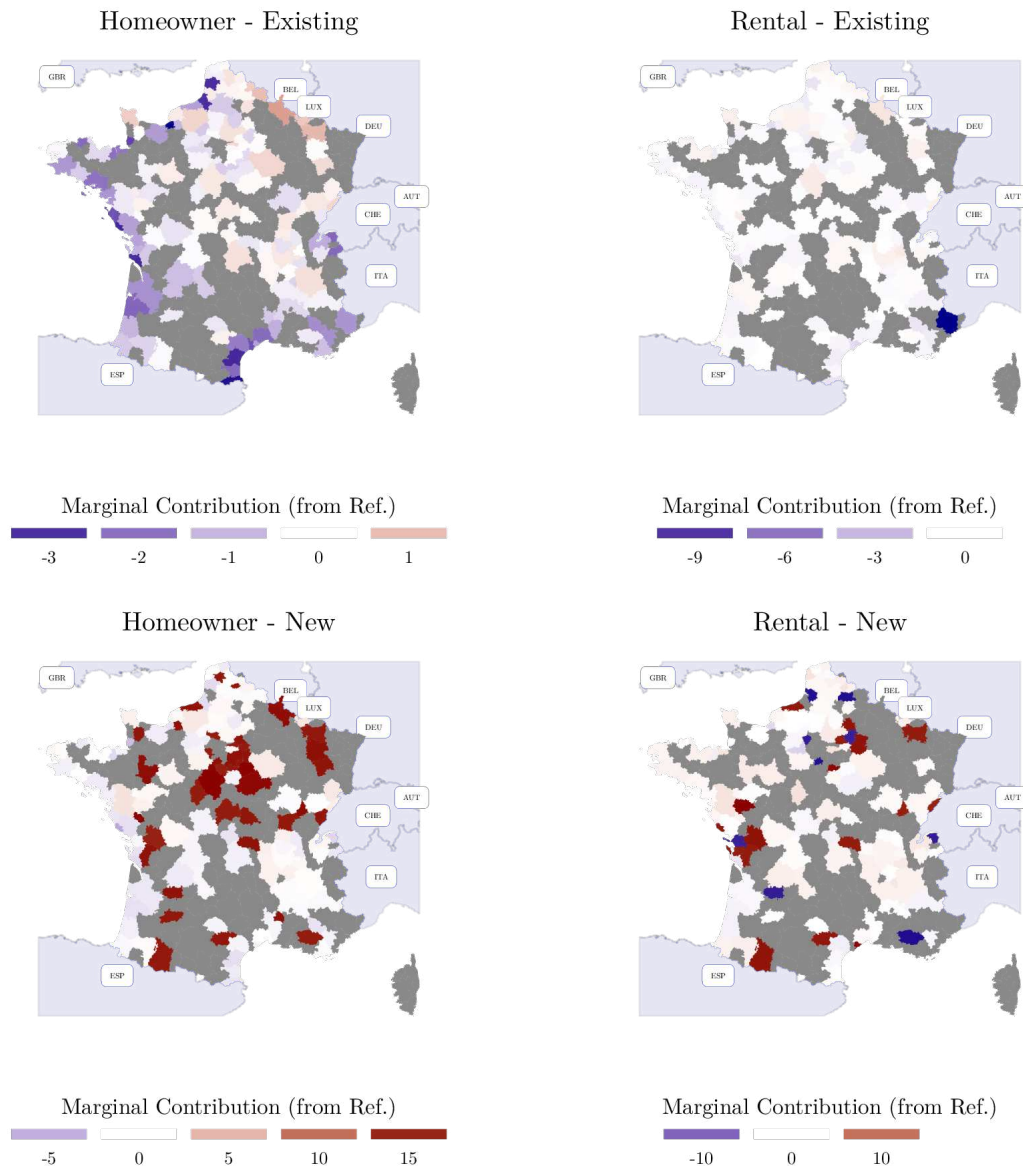


Figure 3.A.5: Commuting Area Fixed Effects for Positive Demand Shock to Explain Probability to be Purchased by Local Investors

Notes: We report marginal contributions for commuting area fixed effects resulting from the estimation of Equation 3.2 to explain probability to be purchased by local investors rather than long-distance one for the positive demand shock experiment. Each map corresponds to the four estimations we make based on the *a priori* defined market segment. The model is estimated with OLS and standard errors are clustered to the commuting area.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

3.A.9 Time Fixed Effects for Positive Demand Shock to Explain Probability to be Purchased by Local Investors using Logistic Regression

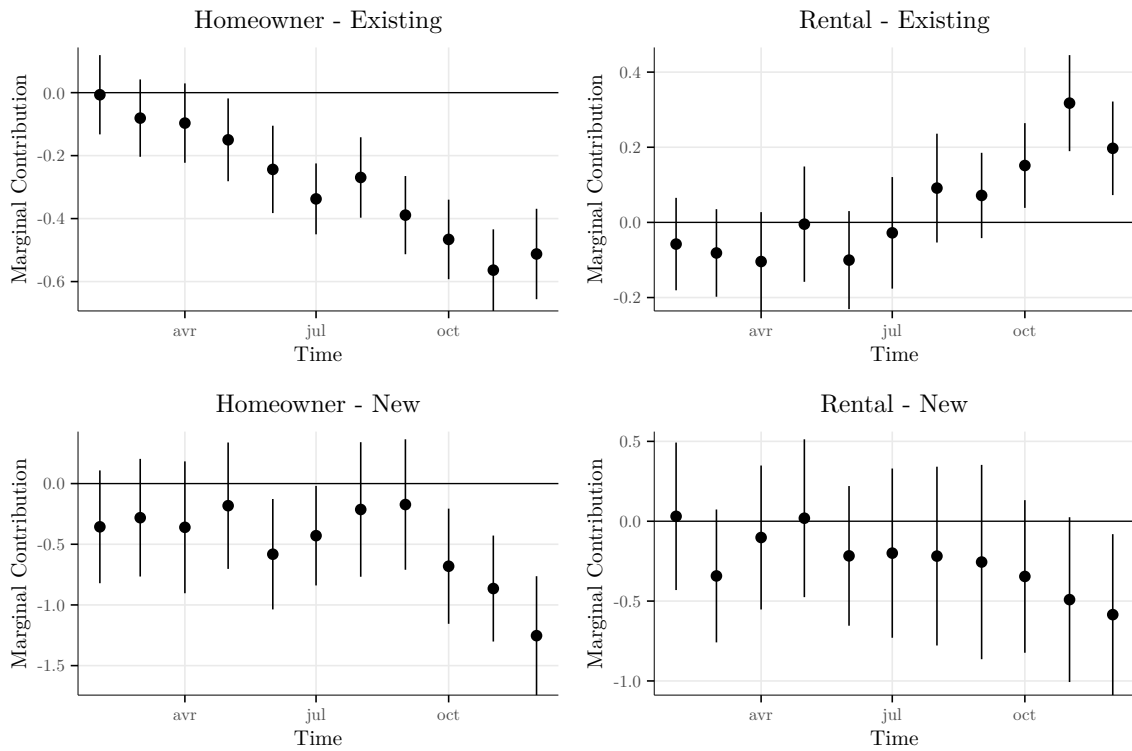


Figure 3.A.6: Time Fixed Effects for Positive Demand Shock to Explain Probability to be Purchased by Local Investors

Notes: We report marginal contributions for time fixed effects resulting from the estimation of Equation 3.2 to explain probability to be purchased by local investors rather than long-distance one for the positive demand shock experiment. Each plot corresponds to the four estimations we make based on the *a priori* defined market segment. The model is estimated with OLS and standard errors are clustered to the commuting area. We report confidence intervals to 95% level. The model specification contains commuting areas fixed effects.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

3.A.10 Commuting Area Fixed Effects for Negative Demand Shock to Explain Probability to be Purchased by Local Investors using Logistic Regression

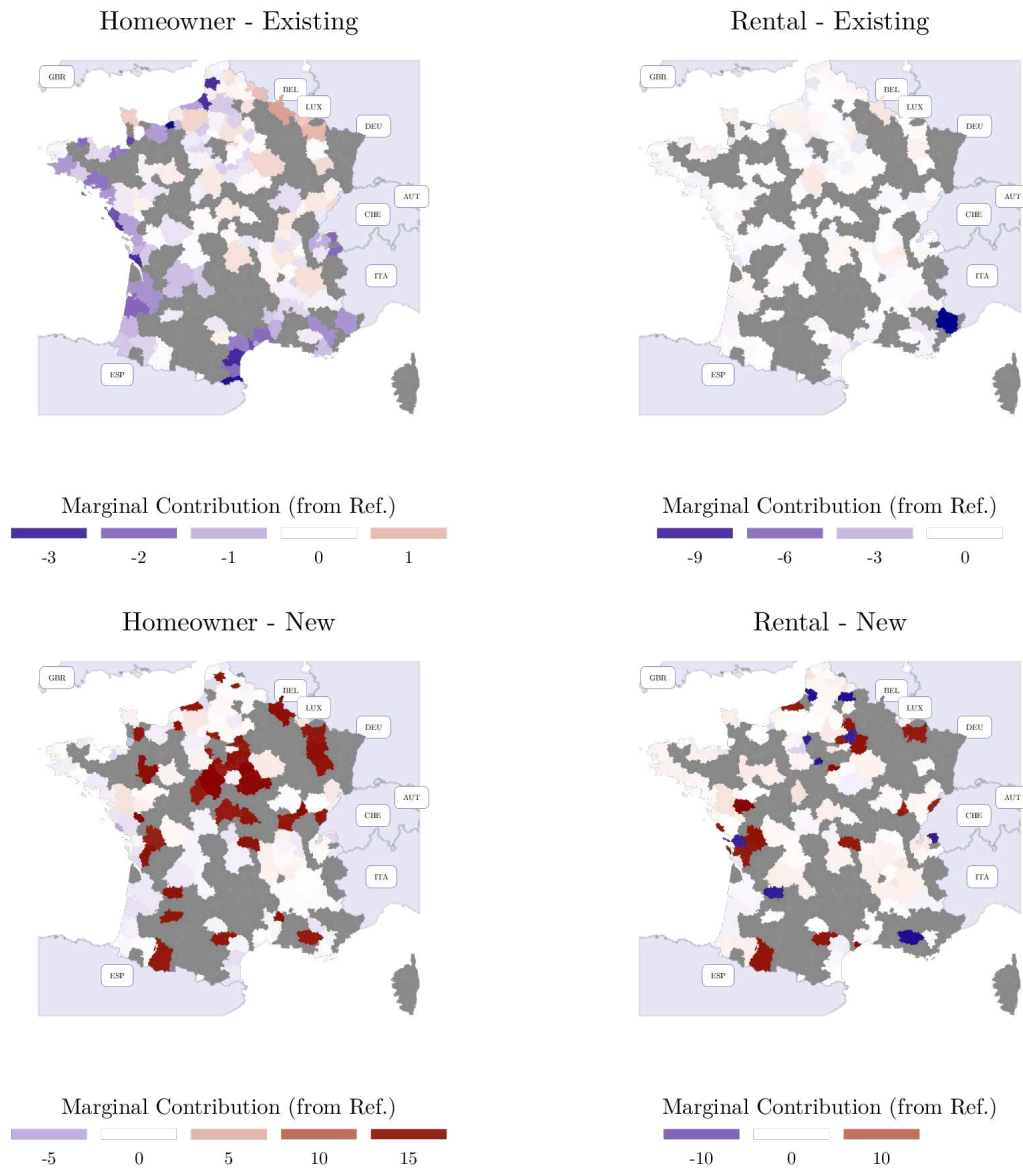


Figure 3.A.7: Commuting Area Fixed Effects for Negative Demand Shock to Explain Probability to be Purchased by Local Investors

Notes: We report marginal contributions for commuting area fixed effects resulting from the estimation of Equation 3.2 to explain probability to be purchased by local investors rather than long-distance one for the negative demand shock experiment. Each map corresponds to the four estimations we make based on the *a priori* defined market segment. The model is estimated with OLS and standard errors are clustered to the commuting area.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

3.A.11 Time Fixed Effects for Negative Demand Shock to Explain Probability to be Purchased by Local Investors using Logistic Regression

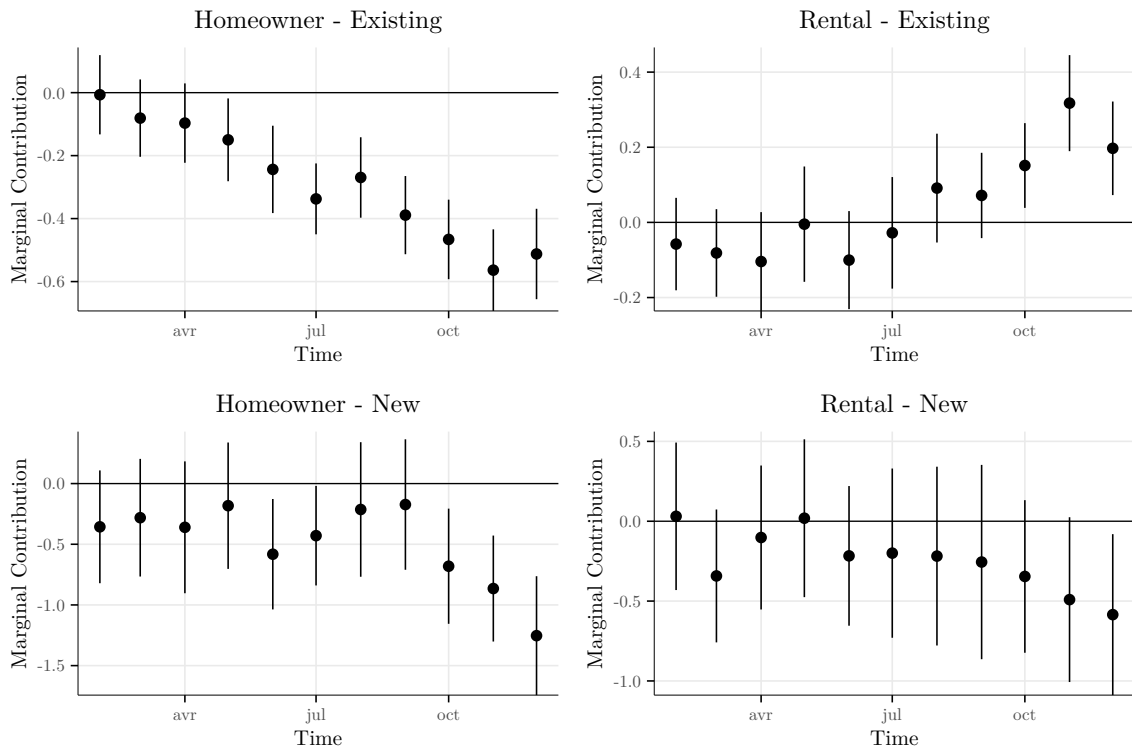


Figure 3.A.8: Time Fixed Effects for Negative Demand Shock to Explain Probability to be Purchased by Local Investors

Notes: We report marginal contributions for time fixed effects resulting from the estimation of Equation 3.2 to explain probability to be purchased by local investors rather than long-distance one for the negative demand shock experiment. Each plot corresponds to the four estimations we make based on the *a priori* defined market segment. The model is estimated with OLS and standard errors are clustered to the commuting area. We report confidence intervals to 95% level. The model specification contains commuting areas fixed effects.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

3.A.12 Detailed Results According to Time Range, Demand Shock Nature, and Market Segments for the OLS Estimation to Explain Unitary Housing Price

Table 3.A.4: Treatment Effects According to Time Range From Policy Introduction

Period Range	<i>Dependent Variable: Unit. Price</i>							
	Positive Shock				Negative Shock			
	Own	Rent	Own	Rent	Own	Rent	Own	Rent
3 months	0.007 (0.020)	0.073* (0.044)	-0.048 (0.055)	0.108 (0.070)	-0.007 (0.009)	-0.037** (0.018)	-0.017 (0.046)	-0.088 (0.078)
4 months	0.017 (0.017)	0.068* (0.035)	-0.016 (0.046)	0.060 (0.065)	-0.005 (0.008)	-0.024 (0.016)	0.005 (0.037)	-0.128** (0.055)
5 months	0.015 (0.016)	0.077*** (0.027)	-0.007 (0.040)	0.054 (0.063)	-0.007 (0.007)	-0.014 (0.015)	-0.009 (0.034)	-0.111** (0.046)
6 months	0.017 (0.013)	0.086*** (0.026)	0.038 (0.046)	0.076 (0.059)	-0.008 (0.006)	-0.002 (0.013)	-0.013 (0.030)	-0.135*** (0.037)
7 months	0.016 (0.011)	0.079*** (0.024)	0.044 (0.038)	0.075 (0.057)	-0.007 (0.006)	-0.001 (0.012)	-0.000 (0.030)	-0.120*** (0.037)
8 months	0.011 (0.011)	0.075*** (0.023)	0.021 (0.035)	0.069 (0.060)	-0.008 (0.005)	0.003 (0.011)	-0.000 (0.028)	-0.131*** (0.037)
9 months	0.012 (0.011)	0.064*** (0.020)	0.006 (0.036)	0.117 (0.075)	-0.007 (0.005)	-0.001 (0.010)	-0.005 (0.028)	-0.126*** (0.039)
10 months	0.012 (0.009)	0.056*** (0.017)	0.029 (0.032)	0.143** (0.065)	-0.008* (0.005)	-0.004 (0.010)	-0.003 (0.027)	-0.151*** (0.037)
11 months	0.013 (0.008)	0.047*** (0.017)	0.014 (0.031)	0.119** (0.058)	-0.006 (0.004)	-0.006 (0.009)	-0.014 (0.027)	-0.139*** (0.039)
12 months	0.012 (0.007)	0.033* (0.019)	-0.010 (0.026)	0.099* (0.050)	-0.005 (0.004)	-0.002 (0.009)	-0.020 (0.025)	-0.135*** (0.037)

Notes: We report estimated treatment effects from Equation 3.1 to explain unitary housing price according to nature of demand shock, housing market segments and period ranges from the policy reform. We do not report group fixed effects, time fixed effects and commuting area fixed effects for clarity reasons. We report standard errors clustered to commuting areas between parentheses.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

3.A.13 Detailed Results According to Time Range, Demand Shock Nature, and Market Segments for the Logistic Regression to Explain Probability to be Purchased by Local Investors

Table 3.A.5: Treatment Effects for Logistic Regression According to Time Range From Policy Introduction

Period Range	<i>Dependent Variable: Unit. Price</i>							
	Positive Shock				Negative Shock			
	Rent	Own.	Rent	Own	Rent	Own.	Rent	Own
3 months	0.068 (0.043)	0.001 (0.022)	0.147** (0.066)	-0.204* (0.111)	-0.152* (0.089)	0.005 (0.060)	1.114** (0.482)	-0.244 (0.378)
4 months	0.060* (0.033)	0.009 (0.019)	0.076 (0.060)	-0.138 (0.089)	-0.051 (0.071)	0.003 (0.057)	1.286*** (0.315)	-0.269 (0.305)
5 months	0.073*** (0.025)	0.011 (0.016)	0.041 (0.068)	-0.077 (0.081)	-0.051 (0.071)	0.003 (0.057)	1.286*** (0.315)	-0.269 (0.305)
6 months	0.076*** (0.025)	0.013 (0.014)	0.046 (0.061)	-0.038 (0.075)	-0.030 (0.054)	-0.067 (0.052)	0.928*** (0.274)	-0.223 (0.266)
7 months	0.073*** (0.023)	0.013 (0.012)	0.046 (0.058)	-0.028 (0.065)	-0.042 (0.052)	-0.052 (0.046)	0.803*** (0.253)	-0.269 (0.254)
8 months	0.063*** (0.023)	0.010 (0.011)	0.037 (0.056)	-0.054 (0.057)	-0.042 (0.052)	-0.052 (0.046)	0.803*** (0.253)	-0.269 (0.254)
9 months	0.053*** (0.020)	0.011 (0.012)	0.081 (0.064)	-0.060 (0.050)	-0.051 (0.048)	-0.002 (0.043)	0.623*** (0.209)	0.087 (0.199)
10 months	0.050*** (0.018)	0.012 (0.010)	0.093* (0.050)	-0.035 (0.050)	-0.046 (0.043)	0.001 (0.041)	0.579*** (0.175)	0.132 (0.192)
11 months	0.038** (0.018)	0.012 (0.009)	0.071 (0.047)	-0.047 (0.047)	-0.046 (0.043)	0.001 (0.041)	0.579*** (0.175)	0.132 (0.192)
12 months	0.027 (0.020)	0.011 (0.008)	0.059 (0.039)	-0.055 (0.046)	-0.014 (0.041)	0.015 (0.040)	0.621*** (0.171)	0.182 (0.181)

Notes: We report estimated treatment effects from Equation 3.2 to explain probability to be purchased by local investors rather than long-distance ones, according to nature of demand shock, housing market segments and period ranges from the policy reform. We do not report group fixed effects, time fixed effects and commuting area fixed effects for clarity reasons. We report standard errors clustered to commuting areas between parentheses.

Sources: Data derived from DV3F and property tax files. Authors' calculation.

*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

Conclusion Générale

Construisant une nouvelle base de données dérivée de sources administratives, nous documentons dans ce travail les liens entre propriété immobilière, politiques publiques et marchés immobiliers, avec une attention particulière pour leur dimension spatiale. De nos travaux, nous justifions la nécessité de considérer conjointement la localisation des propriétaires et l'accumulation de richesses, tout en démontrant une efficacité relative des politiques publiques en soutien à la propriété immobilière.

Chapitre 1 Le premier chapitre documente empiriquement les liens entre tri spatial des propriétaires et persistance de richesse immobilière. En exploitant une base de données construite en appariant différentes sources administratives, représentant une contribution originale de la thèse, nous proposons quatre résultats empiriques illustrant les potentiels mécanismes sous-jacents à la persistance de la richesse immobilière. Premièrement, l'accumulation relative de richesse résultant des marchés immobiliers s'accroît avec la richesse initiale et dépend du lieu de résidence. Deuxièmement, le choix de localisation est corrélé à la richesse initiale, entraînant un tri spatial en fonction de la distance au centre. Troisièmement, la valeur reçue lors d'un transfert non monétaire s'accroît avec la richesse initiale du bénéficiaire, et le statut d'économique de la zone d'emploi de résidence. Quatrièmement, les logements locatifs atténuent la persistance, opérant une redistribution spatiale résultant d'une différence entre lieu de résidence et localisation de l'actif. Ces éléments empiriques illustrent que la persistance de la richesse résulte principalement des choix patrimoniaux plutôt que de l'évolution des marchés immobiliers.

Chapitre 2 Le deuxième chapitre propose une évaluation de l'efficacité du Prêt à Taux Zéro, avec une distinction des effets aux marges extensive et intensive. Sur la base d'une mesure d'efficacité dérivée d'une fonction de bien-être social, nous estimons l'effet général de la politique pour des variations faibles de subvention. Une augmentation de la subvention semble peu efficace au regard des objectifs fixés. En effet, bien que nous ne pouvons rejeter qu'une augmentation de la politique ait une influence sur le nombre de propriétaires, nous estimons que les externalités à la marge intensive (*e.g.* capitalisation de la

subvention) sont plus importantes que celles à la marge extensive (*e.g.* meilleur entretien d'un logement, engagement social accru). Toutefois, l'efficacité de la politique dépend principalement de la valorisation des externalités à la marge intensive, directement liée à la population d'intérêt.

Chapitre 3 Le troisième chapitre traite spécifiquement des effets de capitalisation résultant des politiques subventionnant la demande. Nous montrons que, du fait d'une segmentation locale des marchés prononcée concernant le statut d'occupation, l'effet de capitalisation introduit dans un contexte d'offre inélastique est restreint au segment subventionné. Ainsi, la politique d'investissement locatif n'affecte pas l'accession à la propriété. En outre, l'effet de capitalisation semble principalement causé par des investissements de longue distance. L'introduction de cette politique affecte aussi le segment locatif ancien, bien que non subventionné. Nous estimons que cette capitalisation résulte d'un report d'une partie de la demande pour l'investissement locatif du segment subventionné vers le segment non subventionné du fait d'une concurrence accrue.

Le patrimoine immobilier participe donc au renforcement des inégalités de richesse, à la fois au niveau individuel, mais aussi dans sa dimension spatiale. De fait, les inégalités de richesses doivent être analysées au regard de la localisation des individus. Les conséquences d'un renforcement des inégalités de richesse dans leur dimension spatiale sont nombreuses. Alors que la richesse immobilière est fortement corrélée à la richesse nette des ménages, la relation est probablement étendue à la richesse nette et localisation de la résidence principale. En retour, les revenus résultants d'investissement sont susceptibles de bénéficier aux ménages résidant dans les localisations les plus attractives, renforçant ces dernières.

En contrepartie, les politiques du logement focalisées sur les décisions individuelles des ménages peinent à atteindre les objectifs fixés. Le soutien à la propriété occupante est en effet inefficace, voire contre-productif, puisque l'aide est capitalisée, soit dans les choix individuels sans modification de leur statut d'occupation, soit dans les dynamiques locales de marché. Il ne permet donc pas un développement de la propriété immobilière, notamment pour les ménages les plus contraints financièrement. En outre, le soutien à l'investissement locatif est en partie capitalisé dans les valeurs sur le segment subventionné. La capitalisation est significative, proche de l'aide moyenne apportée. La pertinence des aides focalisées sur la demande dans un contexte de faible élasticité de la construction est donc mise au débat. Des politiques alternatives, bien qu'indirectes, comme le renforcement de l'offre en logement social, représentent un meilleur moyen d'action pour faciliter l'accès à un patrimoine immobilier (Goffette-Nagot and Sidibé, 2016).

Une alternative aux politiques nationales pourrait être une territorialisation accrue

des politiques du logement. En effet, alors que de nombreux paramètres structurants sont définis au niveau local (y compris une partie de la régulation foncière par les documents d'urbanisme), la question de l'adaptation de l'échelle pour la mise en place des politiques du logement se pose. En effet, un alignement entre politiques du logement et décisions locales d'urbanisme pourrait permettre de limiter les effets négatifs. Néanmoins, cela ajouterait un nouveau morcellement des politiques du logement avec un nouveau niveau de complexité, perdant ainsi son caractère universel. En outre, l'initiative des politiques du logement serait reportée au niveau local, pouvant ainsi renforcer les comportements malthusiens des résidents actuels afin de conserver, voire accroître leur valeur immobilière (Ortalo-Magné and Prat, 2014). Enfin, le financement des politiques du logement dans un contexte de territorialisation pourrait introduire de nouvelles différences entre les agglomérations, voire aboutir à une compétition. Bien que nous n'apportons pas de réponses concrètes à propos de ces questions, la spatialisation des politiques du logement demeure un enjeu de compréhension majeur.

Au regard des résultats obtenus, la pertinence des politiques affectant la demande plutôt que l'offre en logement se pose. Les mécanismes de soutien tels que les subventions aux particuliers sont en réalité captés en partie par les promoteurs ou propriétaires existants. Il résulte donc qu'en l'état, du fait de l'inélasticité de l'offre en logement, la majorité des aides publiques associées au logement peinent à produire les effets escomptés. Considérant que les mécanismes liés à l'offre en logement semblent difficiles à mobiliser dans un contexte d'intensification de la restriction de la consommation foncière (la politique ZAN en est un exemple), les mécanismes alternatifs permettant de poursuivre des objectifs similaires à ceux actuellement recherchés par les politiques du logement sont peu nombreux. Alors que l'encadrement des prix des loyers dans certaines métropoles apparaît difficilement transposable aux marchés immobiliers, nous suggérons de nouvelles pistes de réflexion.

La première question concerne le rôle des transferts dans les décisions patrimoniales. Outre l'effet sur la primo-accession par l'octroi d'un apport financier, il s'agit de questionner plus globalement son rôle dans les choix d'investissements (immobiliers ou non) au cours du cycle de vie. À ce stade, nous mettons en lumière l'importance des choix patrimoniaux dans l'accumulation de richesse, sans établir de relation causale ni considérer les effets de long terme sur la nature du patrimoine. Il nous apparaît donc crucial de questionner le rôle précis des transferts intergénérationnels sur les choix d'investissement, y compris immobiliers. Ces travaux seront particulièrement pertinents dans un débat plus général sur la fiscalité des transmissions de patrimoine et les adaptations individuelles (Goupille-Lebret and Infante, 2018; Jakobsen et al., 2020). Le principal défi pour une évaluation causale dans ce cadre (*e.g.* l'effet d'un euro supplémentaire reçu sur la valeur

d'achat, toutes choses égales par ailleurs) concerne l'endogénéité de la variable explicative puisque la richesse des parents est fortement corrélée avec la richesse individuelle, le niveau d'éducation ou encore les revenus.

La deuxième question interroge le rôle de la fiscalité associée à la propriété immobilière. Bien que ce sujet soit complexe et que nous avançons avec prudence, la fiscalité représente un levier de redistribution pouvant faciliter l'accès à un patrimoine immobilier, bien que les propriétaires existants y soient opposés (Brunner et al., 2015). En effet, un accroissement de la fiscalité sur la propriété immobilière pourrait accroître l'offre par une mise en vente plus importante du fait d'une baisse de rendement. Néanmoins, cette hausse est susceptible d'affecter simultanément les propriétaires les plus contraints, voire freiner la transition vers la propriété du fait de coûts additionnels, puisque l'exemption de taxe sur la propriété permet d'augmenter le nombre de propriétaires occupants (Banzhaf et al., 2021). Une taxe potentiellement différenciée en fonction du statut d'occupation peut représenter une solution de compromis. Toutefois, il est difficile *a priori* de préjuger des comportements individuels aux variations de taxe sur la propriété.

Enfin, la dernière question concerne la prise en compte des coûts potentiels associés à la propriété immobilière dans les décisions individuelles. Ces coûts supplémentaires *a priori* omis dans le processus de décision affectent durablement le budget des ménages au cours du cycle de vie et sont peu considérés en ce qui concerne les politiques publiques. Par exemple, les coûts liés à la mobilité dans un contexte d'augmentation du prix des carburants affectent le budget des ménages en fonction de leur localisation (Saber et al., 2017). En conséquence, les préférences individuelles peuvent évoluer au regard du contexte économique, favorisant les localisations avec un faible coût de mobilité du fait d'un accès au réseau de transport public. Ainsi, il s'agit d'établir si ces critères sont pertinents dans le choix de localisation pour un actif immobilier, et dans quelle mesure l'évolution des coûts de mobilité peut affecter les choix de localisation et donc le tri spatial.

References

- Ahlfeldt, Gabriel M., Stephan Heblich, and Tobias Seidel (2023), “Micro-Geographic Property Price and Rent Indices”. *Regional Science and Urban Economics* 98.103836.
- Albouy, David, Gabriel Ehrlich, and Yingyi Liu (2016), “Housing Demand, Cost-of-Living Inequality, and the Affordability Crisis”. *NBER Working Paper* 22816, p. 62.
- Alvaredo, Facundo, Anthony B. Atkinson, and Salvatore Morelli (2018), “Top Wealth Shares in the UK over More than a Century”. *Journal of Public Economics* 162, pp. 26–47.
- André, Mathias and Olivier Meslin (2021), “Et Pour Quelques Appartements de plus : Étude de La Propriété Immobilière Des Ménages et Du Profil Redistributif de La Taxe Foncière”. *INSEE Working Paper*, p. 93.
- Andrews, Dan and Aida Caldera Sánchez (2011), “The Evolution of Homeownership Rates in Selected OECD Countries: Demographic and Public Policy Influences”. *OECD Journal: Economic Studies* 2011.1, pp. 1–37.
- Anenberg, Elliot and Edward Kung (2020), “Can More Housing Supply Solve the Affordability Crisis? Evidence from a Neighborhood Choice Model”. *Regional Science and Urban Economics* 80.103363.
- Angrist, Joshua D and Alan B Krueger (2001), “Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments”. *Journal of Economic Perspectives* 15.4, pp. 69–85.
- Angrist, Joshua D. and Jörn-Steffen Pischke (2010), “The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics”. *Journal of Economic Perspectives* 24.2, pp. 3–30.
- Anselin, Luc and Julie Le Gallo (2006), “Interpolation of Air Quality Measures in Hedonic House Price Models: Spatial Aspects”. *Spatial Economic Analysis* 1.1, pp. 31–52.
- Bach, Laurent, Laurent E. Calvet, and Paolo Sodini (2020), “Rich Pickings? Risk, Return, and Skill in Household Wealth”. *American Economic Review* 110.9, pp. 2703–2747.
- Bäckman, Claes and Chandler Lutz (2020), “The Impact of Interest-Only Loans on Affordability”. *Regional Science and Urban Economics* 80.103376.

- Banzhaf, H. Spencer and Omar Farooque (2013), “Interjurisdictional Housing Prices and Spatial Amenities: Which Measures of Housing Prices Reflect Local Public Goods?” *Regional Science and Urban Economics* 43.4, pp. 635–648.
- Banzhaf, H. Spencer, Ryan Mickey, and Carlianne Patrick (2021), “Age-Based Property Tax Exemptions”. *Journal of Urban Economics* 121, p. 103303.
- Baum-Snow, Nathaniel and Fernando Ferreira (2015), “Causal Inference in Urban and Regional Economics”. *Handbook of Regional and Urban Economics, Vol 5a*. Boston, MA: Elsevier.
- Baum-Snow, Nathaniel and Justin Marion (2009), “The Effects of Low Income Housing Tax Credit Developments on Neighborhoods”. *Journal of Public Economics* 93.5-6, pp. 654–666.
- Bayer, Patrick, Robert McMillan, Alvin Murphy, and Christopher Timmins (2016), “A Dynamic Model of Demand for Houses and Neighborhoods”. *Econometrica* 84.3, pp. 893–942.
- Beaubrun-Diant, Kevin and Tristan-Pierre Maury (2021), “Implications of Homeownership Policies on Land Prices: The Case of a French Experiment”. *Economics Bulletin* 41.3, pp. 1256–1265.
- Behrens, Kristian and Frédéric Robert-Nicoud (2015), “Agglomeration Theory with Heterogeneous Agents”. *Handbook of Regional and Urban Economics, Vol 5a*. Boston, MA: Elsevier.
- Benhabib, Jess, Alberto Bisin, and Mi Luo (2017), “Earnings Inequality and Other Determinants of Wealth Inequality”. *American Economic Review* 107.5, pp. 593–597.
- Bhattacharjee, Arnab, Eduardo Castro, Taps Maiti, and João Marques (2016), “Endogenous Spatial Regression and Delineation of Submarkets: A New Framework with Application to Housing Markets”. *Journal of Applied Econometrics* 31.1, pp. 32–57.
- Bilal, Adrien and Esteban Rossi-Hansberg (2021), “Location as an Asset”. *Econometrica* 89.5, pp. 2459–2495.
- Black, Sandra (1999), “Do Better Schools Matter? Parental Valuation of Elementary Education”. *Quarterly Journal of Economics* 114.2, pp. 577–599.
- Blanco, Miguel Artola, Luis Bauluz, and Clara Martínez-Toledano (2021), “Wealth in Spain 1900–2017: A Country of Two Lands”. *The Economic Journal* 131.633, pp. 129–155.
- Boehm, Thomas P. and Alan M. Schlottmann (1999), “Does Home Ownership by Parents Have an Economic Impact on Their Children?” *Journal of Housing Economics* 8, pp. 217–232.

- Bono, Pierre-Henri and Alain Trannoy (2019), “The Impact of the ‘Scellier’ Income Tax Relief on Building Land Prices in France”. *Economie et Statistique / Economics and Statistics* 507d, pp. 91–114.
- Boserup, Simon H., Wojciech Kopczuk, and Claus T. Kreiner (2016), “The Role of Bequests in Shaping Wealth Inequality: Evidence from Danish Wealth Records”. *American Economic Review* 106.5, pp. 656–661.
- Bourassa, Steven C., Martijn Dröes, and Martin Hoesli (2021), “Hedonic Models and Market Segmentation”. *SSRN Electronic Journal* 21-62.
- Brueckner, Jan K and Stuart S Rosenthal (2009), “Gentrification and Neighborhood Housing Cycles: Will America’s Future Downtowns Be Rich?” *Review of Economics and Statistics* 91.4, pp. 725–743.
- Brueckner, Jan K., Jacques-François Thisse, and Yves Zenou (1999), “Why Is Central Paris Rich and Downtown Detroit Poor?” *European Economic Review* 43.1, pp. 91–107.
- Brunner, Eric J., Stephen L. Ross, and Becky K. Simonsen (2015), “Homeowners, Renters and the Political Economy of Property Taxation”. *Regional Science and Urban Economics* 53, pp. 38–49.
- Brunsdon, Chris, A. Stewart Fotheringham, and Martin E. Charlton (1996), “Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity”. *Geographical Analysis* 28.4, pp. 281–298.
- Carozzi, Felipe (2020), “Credit Constraints and the Composition of Housing Sales. Farewell to First-Time Buyers?” *Journal of the European Economic Association* 18.3, pp. 1196–1237.
- Carozzi, Felipe, Christian A. L. Hilber, and Xiaolun Yu (2020), “On the Economic Impacts of Mortgage Credit Expansion Policies: Evidence from Help to Buy”. *CEPR Discussion Paper* DP14620.
- Chapelle, Guillaume, Benjamin Vignolles, and Clara Wolf (2018), “Impact of a Housing Tax Credit on Local Housing Markets: Evidence from France”. *Annals of Economics and Statistics* 130, pp. 101–131.
- Chareyron, Sylvain, Tidiane Ly, and Yohann Trouvé-Sargison (2021), “Ownership Incentives and Housing Affordability: Evidence from France”. *SSRN Electronic Journal*.
- Chay, Kenneth Y. and Michael Greenstone (2005), “Does Air Quality Matter? Evidence from the Housing Market”. *Journal of Political Economy* 113.2, pp. 376–424.
- Chetty, Raj (2009), “Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods”. *Annual Review of Economics* 1.1, pp. 451–488.
- Coën, Alain, Alexis Pourcelot, and Richard Malle (2022), “Macroeconomic Shocks and Ripple Effects in the Greater Paris Metropolis”. *Journal of Housing Economics* 56.101823.

- Coglianese, John, Lucas W. Davis, Lutz Kilian, and James H. Stock (2017), “Anticipation, Tax Avoidance, and the Price Elasticity of Gasoline Demand”. *Journal of Applied Econometrics* 32.1, pp. 1–15.
- Coulson, N. Edward and Herman Li (2013), “Measuring the External Benefits of Homeownership”. *Journal of Urban Economics* 77, pp. 57–67.
- Cour des Comptes (2012), *Rapport Public Annuel 2012, Tome I*. Tech. rep. Cour des Comptes, p. 1033.
- Crump, Richard K., V. Joseph Hotz, Guido W. Imbens, and Oscar A. Mitnik (2009), “Dealing with Limited Overlap in Estimation of Average Treatment Effects”. *Biometrika* 96.1, pp. 187–199.
- Cunningham, Chris and Robert R. Reed (2013), “Negative Equity and Wages”. *Regional Science and Urban Economics* 43.6, pp. 841–849.
- D’Amour, Alexander, Peng Ding, Avi Feller, Lihua Lei, and Jasjeet Sekhon (2021), “Overlap in Observational Studies with High-Dimensional Covariates”. *Journal of Econometrics* 221.2, pp. 644–654.
- Daminger, Alexander and Kristof Dascher (2022), “Homeowner Subsidy Repeal and Housing Recentralization”. *Land Economics*, 081821–0095R.
- Davis, Morris A., Stephen D. Oliner, Tobias J. Peter, and Edward J. Pinto (2020), “The Impact of Federal Housing Policy on Housing Demand and Homeownership: Evidence from a Quasi-Experiment”. *Journal of Housing Economics* 48.101670.
- de Chaisemartin, Clément and Xavier D’Haultfœuille (2018), “Fuzzy Differences-in-Differences”. *The Review of Economic Studies* 85.2, pp. 999–1028.
- de Chaisemartin, Clément and Xavier D’Haultfœuille (2020), “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects”. *American Economic Review* 110.9, pp. 2964–2996.
- de Chaisemartin, Clément and Xavier D’Haultfœuille (2022), “Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey”. *The Econometrics Journal*.
- De Nardi, Mariacristina (2004), “Wealth Inequality and Intergenerational Links”. *The Review of Economic Studies* 71.3, pp. 743–768.
- Diamond, Rebecca and Cecile Gaubert (2022), “Spatial Sorting and Inequality”. *Annual Review of Economics* 14.1, pp. 795–819.
- Diao, Mi, Delon Leonard, and Tien Foo Sing (2017), “Spatial-Difference-in-Differences Models for Impact of New Mass Rapid Transit Line on Private Housing Values”. *Regional Science and Urban Economics* 67, pp. 64–77.
- Driant, Jean-Claude (2015), *Les Politiques Du Logement En France*. 2e édition. Les Études 5414-15. France: La Documentation Française.

- Driant, Jean-Claude and Pierre Madec (2018), “Mobilité Résidentielle : Tensionmètre Des Crises Du Logement?” *Les Crises Du Logement*. 1re édition. La Vie Des Idées. Paris: Puf : La vie des idées.fr.
- Duranton, Gilles and Diego Puga (2015), “Urban Land Use”. *Handbook of Regional and Urban Economics*. Vol. 5A. Boston, MA: Elsevier, p. 698.
- Dustmann, Christian, Bernd Fitzenberger, and Markus Zimmermann (2022), “Housing Expenditure and Income Inequality”. *The Economic Journal* 132.645, pp. 1709–1736.
- Eggum, Terje and Erling Røed Larsen (2021), “Is the Housing Market an Inequality Generator?” *Housing Lab Working Paper* 2, p. 43.
- Eriksen, Michael D. and Stuart S. Rosenthal (2010), “Crowd out Effects of Place-Based Subsidized Rental Housing: New Evidence from the LIHTC Program”. *Journal of Public Economics* 94.11-12, pp. 953–966.
- Etheridge, Ben (2019), “House Prices and Consumption Inequality”. *International Economic Review* 60.4, pp. 1781–1822.
- Fack, Gabrielle (2006), “Are Housing Benefit an Effective Way to Redistribute Income? Evidence from a Natural Experiment in France”. *Labour Economics* 13.6, pp. 747–771.
- Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri (2020), “Heterogeneity and Persistence in Returns to Wealth”. *Econometrica* 88.1, pp. 115–170.
- Finkelstein, Amy and Nathaniel Hendren (2020), “Welfare Analysis Meets Causal Inference”. *Journal of Economic Perspectives* 34.4, pp. 146–167.
- Fisher, Jonas D. M. and Martin Gervais (2011), “Why Has Home Ownership Fallen Among the Young?” *International Economic Review* 52.3, pp. 883–912.
- Fotheringham, A. Stewart, Chris Brunsdon, and Martin Charlton (2002), *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester, England ; Hoboken, NJ, USA: Wiley.
- Gabriel, Stuart and Gary Painter (2020), “Why Affordability Matters”. *Regional Science and Urban Economics* 80.103378.
- Garbinti, Bertrand, Jonathan Goupille-Lebret, and Thomas Piketty (2021), “Accounting for Wealth-Inequality Dynamics: Methods, Estimates, and Simulations for France”. *Journal of the European Economic Association* 19.1, pp. 620–663.
- Gibbons, Stephen and Stephen Machin (2005), “Valuing Rail Access Using Transport Innovations”. *Journal of Urban Economics* 57.1, pp. 148–169.
- Gilbert, Brian, Abhirup Datta, Joan A. Casey, and Elizabeth L. Ogburn (2023), “A Causal Inference Framework for Spatial Confounding”. *arXiv:2112.14946*.
- Glaeser, Edward L, Joseph Gyourko, and Raven E Saks (2005), “Why Have Housing Prices Gone Up?” *American Economic Review* 95.2, pp. 329–333.

- Glaeser, Edward L. and Joseph Gyourko (2018), “The Economic Implications of Housing Supply”. *Journal of Economic Perspectives* 32.1, pp. 3–30.
- Glaeser, Edward L. and Jesse M. Shapiro (2003), “The Benefits of the Home Mortgage Interest Deduction”. NBER Chapters. National Bureau of Economic Research, Inc, pp. 37–82.
- Goffette-Nagot, Florence and Modibo Sidibé (2016), “Housing Wealth Accumulation: The Role of Public Housing”. *Regional Science and Urban Economics* 57, pp. 12–22.
- Goodman, Allen C. and Thomas G. Thibodeau (1998), “Housing Market Segmentation”. *Journal of Housing Economics* 7.2, pp. 121–143.
- Goodman, Allen C. and Thomas G. Thibodeau (2007), “The Spatial Proximity of Metropolitan Area Housing Submarkets”. *Real Estate Economics* 35.2, pp. 209–232.
- Goodman-Bacon, Andrew (2021), “Difference-in-Differences with Variation in Treatment Timing”. *Journal of Econometrics* 225.2, pp. 254–277.
- Goupille-Lebret, Jonathan and Jose Infante (2018), “Behavioral Responses to Inheritance Tax: Evidence from Notches in France”. *Journal of Public Economics* 168, pp. 21–34.
- Green, Richard K. and Patric H. Hendershott (2001), “Home-Ownership and Unemployment in the US”. *Urban Studies* 38.9, pp. 1509–1520.
- Green, Richard K. and Kerry D. Vandell (1999), “Giving Households Credit: How Changes in the U.S. Tax Code Could Promote Homeownership”. *Regional Science and Urban Economics* 29.4, pp. 419–444.
- Green, Richard K. and Michelle J. White (1997), “Measuring the Benefits of Homeowning: Effects on Children”. *Journal of Urban Economics* 41.3, pp. 441–461.
- Grossman, Sanford T and Guy Laroque (1987), “Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumer Goods”. *Econometrica : journal of the Econometric Society* 58, pp. 25–21.
- Gruber, Jonathan, Amalie Jensen, and Henrik Kleven (2021), “Do People Respond to the Mortgage Interest Deduction? Quasi-experimental Evidence from Denmark”. *American Economic Journal: Economic Policy* 13.2, pp. 273–303.
- Guerrieri, Veronica, Daniel Hartley, and Erik Hurst (2013), “Endogenous Gentrification and Housing Price Dynamics”. *Journal of Public Economics* 100, pp. 45–60.
- Gyourko, Joseph and Raven Molloy (2015), “Regulation and Housing Supply”. *Handbook of Regional and Urban Economics*, Vol 5b. Boston, MA: Elsevier.
- Ha, Sejeong, Christian A.L. Hilber, and Olivier Schöni (2021), “Do Long-Distance Moves Discourage Homeownership? Evidence from England”. *Journal of Housing Economics* 53.101766.
- Hahn, Jinyong (1998), “On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects”. *Econometrica* 66.2, pp. 315–331.

- Hahn, Jinyong, Petra Todd, and Wilbert Klaauw (2001), “Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design”. *Econometrica* 69.1, pp. 201–209.
- Hall, Andrew B. and Jesse Yoder (2022), “Does Homeownership Influence Political Behavior? Evidence from Administrative Data”. *The Journal of Politics* 84.1, pp. 351–366.
- Hanson, Andrew (2012), “Size of Home, Homeownership, and the Mortgage Interest Deduction”. *Journal of Housing Economics* 21.3, pp. 195–210.
- Harding, John, Thomas J. Miceli, and C.F. Sirmans (2000), “Do Owners Take Better Care of Their Housing Than Renters?” *Real Estate Economics* 28.4, pp. 663–681.
- Harkness, Joseph M and Sandra J Newman (2003), “Effects of Homeownership on Children: The Role of Neighborhood Characteristics and Family Income”. *Economic Policy Review* 9.2, pp. 87–107.
- Hembre, Erik (2018), “An Examination of the First-Time Homebuyer Tax Credit”. *Regional Science and Urban Economics* 73, pp. 196–216.
- Henderson, J. Vernon and Yannis M. Ioannides (1983), “A Model of Housing Tenure Choice”. *American Economic Review* 73.1, pp. 98–113.
- Henderson, J. Vernon and Yannis M. Ioannides (1986), “Tenure Choice and the Demand for Housing”. *Economica* 53.210, pp. 231–246.
- Hendren, Nathaniel and Ben Sprung-Keyser (2020), “A Unified Welfare Analysis of Government Policies”. *Quarterly Journal of Economics* 135.3, pp. 1209–1318.
- Hilber, Christian A. L. and Tracy M. Turner (2014), “The Mortgage Interest Deduction and Its Impact on Homeownership Decisions”. *Review of Economics and Statistics* 96.4, pp. 618–637.
- Hilber, Christian A. L. and Wouter Vermeulen (2016), “The Impact of Supply Constraints on House Prices in England”. *The Economic Journal* 126.591, pp. 358–405.
- Hilber, Christian A.L. and Christopher Mayer (2009), “Why Do Households without Children Support Local Public Schools? Linking House Price Capitalization to School Spending”. *Journal of Urban Economics* 65.1, pp. 74–90.
- Hilber, Christian A.L. and Andreas Mense (2021), “Why Have House Prices Risen so Much More than Rents in Superstar Cities?” *Geography and Environment Discussion Paper Series* 30, p. 51.
- Hoebeeck, Annelies and Koen Inghelbrecht (2017), “The Impact of the Mortgage Interest and Capital Deduction Scheme on the Belgian Mortgage Market”. *Working Paper Research: Bank of Belgium* 327.
- Huggett, Mark, Gustavo Ventura, and Amir Yaron (2011), “Sources of Lifetime Inequality”. *American Economic Review* 101.7, pp. 2923–2954.

- Hwang, Min and John M. Quigley (2006), “Economic Fundamentals in Local Housing Markets: Evidence from U.S. Metropolitan Regions”. *Journal of Regional Science* 46.3, pp. 425–453.
- Imbens, Guido W. (2000), “The Role of the Propensity Score in Estimating Dose-Response Functions”. *Biometrika* 87.3, pp. 706–710.
- Jakobsen, Katrine, Kristian Jakobsen, Henrik Kleven, and Gabriel Zucman (2020), “Wealth Taxation and Wealth Accumulation: Theory and Evidence From Denmark”. *Quarterly Journal of Economics* 135.1, pp. 329–388.
- Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M Taylor (2019), “The Rate of Return on Everything, 1870–2015”. *The Quarterly Journal of Economics* 134.3, pp. 1225–1298.
- Keele, Luke and Rocío Titiunik (2015), “Geographic Boundaries as Regression Discontinuities”. *Political Analysis* 23.1, pp. 127–155.
- Kiel, Katherine A. and Jeffrey E. Zabel (2008), “Location, Location, Location: The 3L Approach to House Price Determination”. *Journal of Housing Economics* 17.2, pp. 175–190.
- Kleven, Henrik J. (2021), “Sufficient Statistics Revisited”. *Annual Review of Economics* 13.1, pp. 515–538.
- Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins (2013), “The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets”. *Journal of Economic Literature* 51.4, pp. 1007–1062.
- Kunovac, Davor and Ivan Zilic (2022), “The Effect of Housing Loan Subsidies on Affordability: Evidence from Croatia”. *Journal of Housing Economics* 55.101808.
- Labonne, Claire and Cécile Welter-Nicol (2015), “Cheap Credit, Unaffordable Houses?” *Working Paper*.
- Lee, Hyojung, Dowell Myers, Gary Painter, Johanna Thunell, and Julie Zissimopoulos (2020), “The Role of Parental Financial Assistance in the Transition to Homeownership by Young Adults”. *Journal of Housing Economics* 47, p. 101597.
- Lee, Sanghoon and Jeffrey Lin (2018), “Natural Amenities, Neighbourhood Dynamics, and Persistence in the Spatial Distribution of Income”. *The Review of Economic Studies* 85.1, pp. 663–694.
- Lee, Seong Woo and Dowell Myers (2003), “Local Housing-Market Effects on Tenure Choice”. *Journal of Housing and the Built Environment* 18.2, pp. 129–157.
- Li, Ziqi and A. Stewart Fotheringham (2020), “Computational Improvements to Multi-Scale Geographically Weighted Regression”. *International Journal of Geographical Information Science* 34.7, pp. 1378–1397.

- Li, Ziqi, A. Stewart Fotheringham, Wenwen Li, and Taylor Oshan (2019), “Fast Geographically Weighted Regression (FastGWR): A Scalable Algorithm to Investigate Spatial Process Heterogeneity in Millions of Observations”. *International Journal of Geographical Information Science* 33.1, pp. 155–175.
- Madec, Pierre (2018), “La territorialisation de la politique du logement : un enjeu d’égalité”. *Hérodote* 3.170, pp. 175–194.
- Malpezzi, Stephen and Kerry Vandell (2002), “Does the Low-Income Housing Tax Credit Increase the Supply of Housing?” *Journal of Housing Economics* 11.4, pp. 360–380.
- Martin, Hal and Andrew Hanson (2016), “Metropolitan Area Home Prices and the Mortgage Interest Deduction: Estimates and Simulations from Policy Change”. *Regional Science and Urban Economics* 59, pp. 12–23.
- McMillen, Daniel P. (2010), “Issues in Spatial Data Analysis”. *Journal of Regional Science* 50.1, pp. 119–141.
- Molloy, Raven (2020), “The Effect of Housing Supply Regulation on Housing Affordability: A Review”. *Regional Science and Urban Economics* 80.103350.
- Murakami, Daisuke, Narumasa Tsutsumida, Takahiro Yoshida, Tomoki Nakaya, and Binbin Lu (2020), “Scalable GWR: A Linear-Time Algorithm for Large-Scale Geographically Weighted Regression with Polynomial Kernels”. *Annals of the American Association of Geographers*, pp. 1–22.
- Ortalo-Magné, François and Andrea Prat (2014), “On the Political Economy of Urban Growth: Homeownership versus Affordability”. *American Economic Journal: Microeconomics* 6.1, pp. 154–181.
- Ortalo-Magné, François and Andrea Prat (2016), “Spatial Asset Pricing: A First Step”. *Economica* 83.329, pp. 130–171.
- Pfeffer, Fabian T and Nora Waitkus (2021), “The Wealth Inequality of Nations”. *American Sociological Review* 86.4, pp. 567–602.
- Piazzesi, Monika, Martin Schneider, and Johannes Stroebe (2020), “Segmented Housing Search”. *American Economic Review* 110.3, pp. 720–759.
- Piketty, Thomas and Gabriel Zucman (2014), “Capital Is Back: Wealth-income Ratios in Rich Countries 1700–2010”. *Quarterly Journal of Economics* 129.3, pp. 1255–1310.
- Roback, Jennifer (1982), “Wages, Rents, and the Quality of Life”. *Journal of political Economy* 90.6, pp. 1257–1278.
- Robins, James M. and Andrea Rotnitzky (1995), “Semiparametric Efficiency in Multivariate Regression Models with Missing Data”. *Journal of the American Statistical Association* 90.429, pp. 122–129.
- Rosen, Sherwin (1974), “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition”. *Journal of Political Economy* 82.1, pp. 34–55.

- Roth, Jonathan (2022), “Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends”. *American Economic Review: Insights* 4.3, pp. 305–322.
- Rothenberg, Jerome, ed. (1991), *The Maze of Urban Housing Markets: Theory, Evidence, and Policy*. Chicago: University of Chicago Press.
- Rubin, Donald B. (1974), “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.” *Journal of Educational Psychology* 66.5, pp. 688–701.
- Saberi, Meead, Hongzhi Wu, Richard Amoh-Gyimah, Jonathan Smith, and Dharmalingam Arunachalam (2017), “Measuring Housing and Transportation Affordability: A Case Study of Melbourne, Australia”. *Journal of Transport Geography* 65, pp. 134–146.
- Saez, Emmanuel and Gabriel Zucman (2016), “Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data”. *Quarterly Journal of Economics* 131.2, pp. 519–578.
- Saiz, Albert (2010), “The Geographic Determinants of Housing Supply”. *Quarterly Journal of Economics* 125.3, pp. 1253–1296.
- Sinai, Todd and Nicholas S. Souleles (2005), “Owner-Occupied Housing as a Hedge Against Rent Risk”. *Quarterly Journal of Economics* 120.2, pp. 763–789.
- Słoczyński, Tymon and Jeffrey M. Wooldridge (2018), “A General Double Robustness Result for Estimating Average Treatment Effects”. *Econometric Theory* 34.1, pp. 112–133.
- Sommer, Kamila and Paul Sullivan (2018), “Implications of US Tax Policy for House Prices, Rents, and Homeownership”. *American Economic Review* 108.2, pp. 241–274.
- Sotura, Aurélie (2020), “Housing Policies, Spatial Inequalities and Housing Markets”. PhD thesis. Paris: EHESS.
- Szumilo, Nikodem (2019), “The Spatial Consequences of the Housing Affordability Crisis in England”. *Environment and Planning A: Economy and Space* 51.6, pp. 1264–1286.
- Tur-Sinai, Aviad, Larisa Fleishman, and Dmitri Romanov (2020), “The Accuracy of Self-Reported Dwelling Valuation”. *Journal of Housing Economics* 48.101660.
- Turner, Matthew, Andrew Haughwout, and Wilbert Van der Klauw (2014), “Land Use Regulation and Welfare”. *Econometrica* 82.4, pp. 1341–1403.
- Van Ommeren, Jos and Michiel Van Leuvensteijn (2005), “New Evidence of the Effect of Transaction Costs on Residential Mobility”. *Journal of Regional Science* 45.4, pp. 681–702.
- Waxman, Andrew, Yuanning Liang, Shanjun Li, Panle Jia Barwick, and Meng Zhao (2020), “Tightening Belts to Buy a Home: Consumption Responses to Rising Housing Prices in Urban China”. *Journal of Urban Economics* 115.103190.
- Wheaton, William C. (1990), “Vacancy, Search, and Prices in a Housing Market Matching Model”. *Journal of Political Economy* 98.6, pp. 1270–1292.

- Wijburg, Gertjan (2022), “Tax-Incentivized Housing Production and the Affordability Crisis: International Lessons from the Low-Income Housing Tax Credit Program in the United States”. *Housing Studies* 0.0, pp. 1–26.
- Wood, Simon N. (2017), *Generalized Additive Models: An Introduction with R*. Second. Chapman and Hall/CRC.
- Wood, Simon N., Natalya Pya, and Benjamin Säfken (2016), “Smoothing Parameter and Model Selection for General Smooth Models”. *Journal of the American Statistical Association* 111.516, pp. 1548–1563.
- Wooldridge, Jeffrey M. (2016), *Introductory Econometrics: A Modern Approach*. Sixth edition. Boston, MA: Cengage Learning.
- Zabel, Jeffrey E. (2012), “Migration, Housing Market, and Labor Market Responses to Employment Shocks”. *Journal of Urban Economics* 72.2-3, pp. 267–284.
- Zucman, Gabriel (2019), “Global Wealth Inequality”. *Annual Review of Economics* 11.1, pp. 109–138.

List of Figures

1.1	Housing Wealth Persistence between Commuting Zones of homeowners' main residences	46
1.2	Partial Correlation Between Housing Wealth Variation, Initial Level of Wealth and Location of Main Residence in 2018	48
1.3	Partial Correlation of Initial Level of Wealth with Housing Decisions within Commuting Zones for Owner-Occupiers for 2017	51
1.4	Partial Correlation of Initial Level of Wealth and Individual Locations with Transfer Share in 2017	53
1.5	Correlation of Annual Gross Rates Between Wealth and Housing Market Dynamics for Local and Distant Investors	55
1.A.1	Share of Developed Land According to Commuting Area	65
1.A.2	Relationship for 1D Variables with Unitary Housing Price	73
1.A.4	Descriptive Statistics about Appraised Values According to Location	78
1.A.5	Age of New and Former Owners According to Change Type	80
1.B.1	Partial Correlation Between Housing Wealth Variation, Initial Level of Wealth and Location of Main Residence (2011-2014)	81
1.B.2	Partial Correlation Between Housing Wealth Variation, Initial Level of Wealth and Location of Main Residence (2015-2018)	82
1.B.3	Correlation between Previous Wealth, Location Choices and Transaction Price for Owner-Occupiers (2012-2014)	83
1.B.4	Correlation between Previous Wealth, Location Choices and Transaction Price for Owner-Occupiers (2015-2017)	84
1.B.5	Mean Housing Transfers per Year per Individuals	85
1.B.6	Partial Correlation of Initial Level of Wealth and Residence Place with Housing Transfers (2012–2014)	86
1.B.7	Partial Correlation of Initial Level of Wealth and Residence Place with Housing Transfers (2015–2017)	87
2.5.1	Overlap between predictions of tenseness between the different ABC zones .	106
2.5.2	Naive and policy-relevant treatment effects used in placebo analysis	112

2.A.1	Number of IFL Policy and Average Subsidy	119
2.A.2	Current ABC Zoning	122
2.C.1	Contribution for 1D-variable in the GPS Estimation	124
2.C.2	Spatial Smoothing Function for the GPS Estimation Based On Municipality Coordinates	125
2.C.3	Estimated Municipalities Classification	126
2.D.1	Marginal Contribution for Spatial Coordinates (Second Step)	130
2.D.2	Dose-Response Plots	133
2.D.3	Dose-Response Plots	134
2.D.4	Dose-Response Plots	135
2.D.5	Dose-Response Plots	136
2.E.1	Difference on Homeownership Cost Induced by the IFL	139
3.2.1	Price Adjustment According to Market Segmentation	146
3.2.2	Price Adjustment According to Market Segmentation	147
3.3.1	Number of Transactions According to Defined Segmentation	149
3.3.2	Share of Narrow Purchasers According to Housing Market Segmentation . .	149
3.4.1	Temporal and Spatial Distribution of Exception Acceptance	152
3.4.2	Descriptive Statistics About B ₂ Municipalities	153
3.A.1	Commuting Area Fixed Effects for Positive Demand Shock to Explain Uni- tary Housing Price	171
3.A.2	Time Fixed Effects for Positive Demand Shock to Explain Unitary Housing Price	172
3.A.3	Commuting Area Fixed Effects for Negative Demand Shock to Explain Uni- tary Housing Price	173
3.A.4	Time Fixed Effects for Negative Demand Shock to Explain Unitary Housing Price	174
3.A.5	Commuting Area Fixed Effects for Positive Demand Shock to Explain Prob- ability to be Purchased by Local Investors	175
3.A.6	Time Fixed Effects for Positive Demand Shock to Explain Probability to be Purchased by Local Investors	176
3.A.7	Commuting Area Fixed Effects for Negative Demand Shock to Explain Prob- ability to be Purchased by Local Investors	177
3.A.8	Time Fixed Effects for Negative Demand Shock to Explain Probability to be Purchased by Local Investors	178

List of Tables

1.A.1	Descriptive Statistics for Housing Stock	63
1.A.2	Descriptive Statistics for the DV3F database	64
1.A.3	Potential National ID Based on Civil State and Name Truncating	66
1.A.4	Overall Results for Potential Errors Assessments according to Potential ID .	67
1.A.5	Summary of Methodological Choices	69
1.A.6	Distribution of Relative Errors According to the Estimation Procedure . . .	71
1.A.7	Distribution of Relative Errors According to the Estimation Procedure and Housing Locations	71
1.A.8	Joint Significance of Covariates in Pooled Models for Auto Valuation Method	74
1.A.9	Results from Joint Significance of Smoothing Splines Transformation	75
1.A.10	Descriptive Statistics for Housing Value Imputation	76
1.A.11	Missing Rate According to Owner Type and Temporal Assumption	79
2.2.1	Main variables and primitives for municipalities along the ABC zoning . . .	96
2.5.1	Covariates' joint significance from first-stage ordered GAMs	105
2.5.2	Covariates' joint significance from second-stage pooled GAMs	107
2.5.3	OLS coefficients for policy primitives from dose-response functions	109
2.5.4	Measures of cost-effectiveness for different marginal values of externalities .	111
2.A.1	Pre-Treatment Variables About Municipality Exploit for GPS Specification .	120
2.A.2	Descriptive Statistics about Outcomes Exploit for Econometric Model	121
2.C.1	Overlap	127
2.D.1	Joint Significance for Control Variables (Second Step)	129
2.D.2	Policy Relevant Treatment Effects (ATE)	131
2.D.3	Policy Relevant Treatment Effects (ATT)	132
2.D.4	Measures of cost-effectiveness for different marginal values of externalities (Alternative GPS Specification with $df = 50$ for Spatial Smoothing)	137
2.D.5	Measures of cost-effectiveness for different marginal values of externalities (Alternative GPS Specification with No Spatial Smoothing)	138
3.4.1	Composition of Groups According to the Nature of the Demand Shock . . .	156

3.5.1	Results for Price Adjustment (OLS Estimation) and Probability to be Purchased by Local Investors (Logit) Under Homogeneous Assumption	158
3.5.2	Results from TWFE Regressions According to Demand Shock Nature and Housing Market Segments	159
3.5.3	Results from Logistic Regression to Explain Probability to Be Purchased by Local Investors	160
3.5.4	Results From TWFE With Falsified Introduction Timing	162
3.5.5	Robustness Checks From TWFE Results	163
3.A.1	Requirement for Municipalities to Obtain Exception	168
3.A.2	Descriptive Statistics for Transactions Selected for Repeal Natural Experiments	169
3.A.3	Descriptive Statistics for Transactions Selected for Introduction Natural Experiments	170
3.A.4	Treatment Effects According to Time Range From Policy Introduction	179
3.A.5	Treatment Effects for Logistic Regression According to Time Range From Policy Introduction	180

Titre : Dynamique de la propriété immobilière : entre marchés et politiques spatialisés

Mots clés : Propriété Immobilière ; Inégalités Spatiales ; Politiques Publiques ; Marchés Locaux ; Hétérogénéité Spatiale

Résumé : Mettant à profit les données fiscales, nous proposons trois analyses économiques à propos de la propriété immobilière et les politiques publiques. Le premier chapitre illustre la persistance de la richesse immobilière (2011-2019), résultant principalement des choix patrimoniaux. Le second chapitre est une évaluation d'une politique de soutien à la propriété occupante. Nous distinguons les effets aux marges extensive et intensive, dérivant l'efficacité à partir d'une fonction de bien-être social. L'augmentation de l'intensité de traitement apparaît inefficace. Le troisième chapitre est focalisé sur les effets de capitalisation résultant de politiques publiques dans un contexte de marché segmenté.

La capitalisation des aides pour l'investissement locatif affectent uniquement le segment locatif. Ces travaux confirment l'importance de la propriété immobilière dans la construction des inégalités, tout en précisant la faible efficacité des aides publiques.

Title : The Dynamics of Property Ownership with Spatial Markets and Policies

Keywords : Property Ownership ; Spatial Inequalities ; Public Policies ; Local Markets ; Spatial Heterogeneity

Abstract : Making the best of fiscal data, we bring three economic studies about the property ownership and housing policies. The first chapter illustrate the housing wealth persistence (2011-2019) that mainly results from portfolios choices. The second chapter assesses French housing policies that support transition to homeownership. We disentangle effect at the extensive and intensive margins, and derive cost-efficiency from social welfare functions. Increasing subsidy appears inefficient. The third chapter focuses on price capitalisation yielding by housing policies under potential housing market segmentation.

Price capitalisation concerns solely rental segments, leaving owner-occupier unaffected. These studies confirm the role of property ownership to inequalities, and precise the low-efficiency of public policies.