

Ecole des Hautes Etudes en Sciences Sociales

## THESE

Pour l'obtention du grade de docteur en Sciences Économiques  
de l'École des Hautes Études en Sciences Sociales

*Présentée et soutenue publiquement le 3 juin 2020 par :*

Aurélie SOTURA

# Politiques du logement, inégalités territoriales et marchés de l'immobilier

Directeur de thèse : M. Thomas PIKETTY

### Composition du jury :

<i>Rapporteurs :</i>	Pierre-Philippe COMBES	Université de Lyon et Sciences Po
	Benoit SCHMUTZ	École Polytechnique
<i>Directeur :</i>	Thomas PIKETTY	EHESS - PSE
<i>Examineurs :</i>	Gilles DURANTON	Université de Pennsylvanie
	Miren LAFOURCADE	Université Paris-Sud et PSE

Ecole des Hautes Etudes en Sciences Sociales

## PhD Thesis

Submitted to École des Hautes Études en Sciences Sociales  
For the Degree of Doctor in Economics

*Prepared and defended at Paris School of Economics on June 3rd, 2020 by:*

Aurélie SOTURA

# Housing Public Policies, Spatial Inequalities, and Housing Markets

Thesis Advisor: M. Thomas PIKETTY

### Jury:

<i>Rapporteurs :</i>	Pierre-Philippe COMBES	University of Lyon and Sciences Po
	Benoit SCHMUTZ	École Polytechnique
<i>Directeur :</i>	Thomas PIKETTY	EHESS - PSE
<i>Examineurs :</i>	Gilles DURANTON	University of Pennsylvania
	Miren LAFOURCADE	University Paris-Sud and PSE

# Remerciements

Cette thèse doit énormément à tous ceux qui m'ont accompagné. Avant toute chose je souhaite remercier Thomas Piketty, un directeur de thèse exceptionnel. Sa hauteur de vue, son ouverture et ses intuitions sur les sciences humaines sont remarquables. Je lui suis particulièrement reconnaissante d'avoir dirigé ma thèse alors que j'avais mis entre parenthèses la recherche pendant quelques années.

Je tiens ensuite à remercier vivement les rapporteurs et les membres du jury qui m'ont fait l'honneur d'accepter cette tâche. Leurs commentaires et suggestions contribuent grandement à enrichir mon travail.

Je suis également reconnaissante à toute l'équipe du WID à PSE qui a égayé et nourri mon temps passé à Jourdan. Je pense particulièrement à celles et ceux dont j'ai partagé le bureau et à nos discussions économiques et politiques ! Je n'oublie pas non plus mes camarades du troisième étage !

Je remercie aussi chaleureusement l'ensemble de mes co-auteurs avec lesquels j'ai eu plaisir à collaborer pour l'écriture de certains chapitres de cette thèse.

Je remercie aussi avec gratitude Laurent Girometti, Clémentine Pesret, Hakim Lahlou du Ministère du Logement d'avoir rendu possible cette thèse. Un grand merci à Eva qui a brillamment repris le flambeau et m'a encouragé dans mes moments de doute. Je me remémore également avec chaleur mes camarades du bureau de FE5, bureau geek et pourtant féminin. Je remercie également les membres des autres bureaux de l'étage avec lesquels j'ai eu plaisir à échanger. Je n'oublie pas les encouragements de l'équipe de grimpeuses du midi.

Cette thèse a aussi été possible parce que j'ai pu la préparer durant mon dernier poste à Bercy, c'est donc l'occasion de remercier ma dernière équipe au Ministère des Finances. J'ai aussi terminé cette thèse dans d'excellentes conditions à la Banque de France et je me réjouis de continuer à partager mes journées avec mes nouveaux collègues.

Je suis aussi reconnaissante à tous mes amis non économistes avec qui j'ai partagé de bons moments en grim pant, skiant, ou tout simplement autour d'un verre. Un merci en particulier à Marine pour sa grande bienveillance, et vous tous qui avez contribué à me donner du cœur à l'ouvrage.

Enfin, je souhaite remercier ma famille sans qui rien ne serait possible, mes parents bienveillants qui m'ont donné le goût des maths, du débat et de la politique, mélange incongru qui m'a mené à l'économie, mes grand-parents chéris, et aussi, Alexis, Christelle et Pascal.

*A André, Jacqueline, Jeanne et Pierre*



# Table of contents

<b>Introduction générale</b>	<b>1</b>
<b>Summary</b>	<b>5</b>
<b>1 Welfare effects of the no-interest loans in France</b>	<b>11</b>
1.1 Literature on place based policies and borrowing constraints . . . . .	13
1.2 The no-interest loan policy . . . . .	15
1.2.1 Institutional background . . . . .	15
1.2.2 Descriptive statistics . . . . .	19
1.3 Model . . . . .	25
1.3.1 The setting . . . . .	25
1.3.2 The housing market . . . . .	25
1.3.3 Individual choice . . . . .	26
1.3.4 Housing supply . . . . .	27
1.3.5 Welfare implications . . . . .	28
1.3.6 Extensions . . . . .	34
1.4 Taking our theoretical model to the data . . . . .	35
1.4.1 Estimation of utility parameters . . . . .	36
1.4.2 Quantities involved in welfare effects . . . . .	37
1.5 Conclusion . . . . .	39
Appendices . . . . .	42
A Production of consumption goods . . . . .	42
B Extended model . . . . .	43
B.1 Extensions to new/old housing types and family size . . . . .	43
B.2 Welfare comparative statics . . . . .	45
C Borrowing constraints . . . . .	53
D NIL subsidy . . . . .	55
<b>2 Intraregional inequality in France in the last fifty years</b>	<b>57</b>
1 Introduction . . . . .	59
2 Data, Concepts and Methods . . . . .	62
2.1 Data Sources . . . . .	62
2.2 Methodology . . . . .	64
2.3 Concepts . . . . .	67
3 Results . . . . .	69
4 Conclusion . . . . .	88
Appendices . . . . .	90

A	Methodology for the 1986-2017 Period . . . . .	90
A.1	Cleaning Fiscal Tabulations . . . . .	90
A.2	Taxable Income Distributions . . . . .	91
A.3	Fiscal and Pretax Income Distributions . . . . .	91
A.4	Consistency with National Estimates . . . . .	92
B	Methodology for the 1966-1969 Period . . . . .	92
B.1	Recovering the Number of Non Taxable Tax Units . . . . .	93
B.2	Recovering the Average Income of the Non Taxable Tax Units . . . . .	94
<b>3</b>	<b>Les inégalités de revenu entre les départements français depuis cent ans</b>	<b>103</b>
1	Données et méthodes . . . . .	108
2	Variables et indicateurs . . . . .	112
3	Résultats . . . . .	114
4	Conclusion . . . . .	128
	Appendices . . . . .	131
<b>4</b>	<b>This Town Ain't Big Enough? Quantifying Local Public Goods Spillovers</b>	<b>139</b>
1	Introduction . . . . .	141
2	Institutional Background . . . . .	146
3	Theoretical Model . . . . .	150
3.1	Preferences . . . . .	151
3.2	Conditional Housing Demand . . . . .	153
3.3	Demand for Jurisdictions . . . . .	154
3.4	Housing Supply . . . . .	155
3.5	Labor Demand . . . . .	155
3.6	Public Good Supply . . . . .	156
3.7	Equilibrium . . . . .	157
3.8	Residential Amenities . . . . .	158
4	Data . . . . .	160
5	Reduced-Form Evidence . . . . .	164
5.1	DiD Framework . . . . .	164
5.2	Results . . . . .	166
5.3	Robustness Checks . . . . .	171
6	Structural Estimation . . . . .	173
6.1	Generalized Method of Moments . . . . .	173
6.2	Non-Parametric Evidence . . . . .	174
6.3	Moment Conditions . . . . .	175
6.4	Estimation Results . . . . .	177
7	Welfare Implications . . . . .	179
8	Conclusion . . . . .	182
	Appendices . . . . .	205
A	Jurisdiction Appeal . . . . .	205
B	Local Public Goods . . . . .	205
B.1	Extension of The Model . . . . .	205
B.2	Time Consistency and No Inherited Wealth . . . . .	206
B.3	Taking the Extended Model to the Data . . . . .	209

C	DiD Robustness Checks . . . . .	210
D	Welfare . . . . .	211
	D.1 Before The Merger . . . . .	211
	D.2 After the merger . . . . .	214
	D.3 Welfare comparison . . . . .	216
	D.4 Welfare estimation . . . . .	217
<b>5</b>	<b>Les étrangers font-ils monter les prix de l'immobilier? Estimation à partir de la base de la chambre des Notaires de Paris, 1993-2008.</b>	<b>221</b>
1	Stock et flux d'appartements à Paris . . . . .	225
	1.1 Stock d'appartements . . . . .	226
	1.2 Flux d'appartements. . . . .	238
	1.3 Flux d'appartements vendus par décile de prix. . . . .	242
2	Stratégie d'estimation . . . . .	247
	2.1 Quelques considérations théoriques . . . . .	247
	2.2 Méthodologie . . . . .	249
3	Résultats des estimations . . . . .	252
	3.1 Impact sur les prix des acheteurs résidant à l'étranger ou de nationalité étrangère . . . . .	252
4	Critiques et pistes d'améliorations . . . . .	281
	4.1 Robustesse des résultats . . . . .	281
	4.2 Impact global des étrangers sur les prix de l'immobilier . . . . .	281
5	Conclusion . . . . .	286
	Appendices . . . . .	289
A	Prix moyen des appartements vendus par catégorie de prix . . . . .	289
B	Définition des groupes de pays . . . . .	290
C	Exemple de régression hédonique . . . . .	292
D	Robustesse des résultats . . . . .	294
	D.1 Régressions avec pour variable explicative le pourcentage d'acheteurs étrangers net du pourcentage de vendeurs étrangers . . . . .	294
	D.2 Régressions où l'on n'a pas exclu les sociétés civiles immobilières et les acheteurs de type inconnu . . . . .	301
	<b>References</b>	<b>301</b>
	<i>List of tables</i>	<b>318</b>
	<i>List of figures</i>	<b>321</b>



# Introduction Générale

Le lien entre inégalités et territoires est devenu particulièrement prégnant ces dernières années en France lorsqu'ont éclaté des mouvements de protestations populaires et spontanées. Les émeutes qui ont secoué les banlieues des grandes villes françaises en 2005 avaient mis en lumière le sentiment d'exclusion d'une partie de la population en marge des métropoles. Le mouvement des Gilets Jaunes qui a commencé en novembre 2018 provient aussi d'un sentiment de déclassement de parties entières de la population, mais, cette fois, il s'agit des habitants des zones rurales et périurbaines. Ces mouvements de protestation sont indissociables des lieux de résidence des populations qui expriment leur mécontentement. Mieux comprendre le niveau et l'évolution des inégalités territoriales apparaît essentiel dans ce contexte : quel accès aux services publics selon les territoires ? Quelle répartition spatiale des plus hauts revenus et des plus modestes ?

Les inégalités territoriales sont elles-mêmes reliées aux marchés du logement et aux politiques publiques dans ce secteur. En effet, les biens immobiliers composent la majorité du patrimoine des ménages, mais sont inégalement répartis entre les ménages (d'après [Ferrante et al. \(2016\)](#)). D'autre part, le fonctionnement des marchés immobiliers exclut de fait les ménages modestes des logements les plus agréables et surtout les mieux localisés, tout simplement car les loyers et les prix de l'immobilier élevés les leur rendent inabordables. L'Etat cherche, à travers certaines de ses politiques, à rendre les logements plus accessibles aux ménages modestes et médians. Mais, ce faisant, il influe sur les prix de l'immobilier, les loyers, et aussi, sur la localisation des ménages. Avec quel impact sur les inégalités territoriales

et les marchés de l'immobilier?

Cette thèse porte sur les politiques du logement, les inégalités territoriales et les marchés de l'immobilier. Elle éclaire quelques aspects de l'entrelacs de liens entre ces sujets.

Le premier chapitre est issu d'un travail commun avec Gabrielle Fack et Laurent Gobillon. Nous étudions l'impact du prêt à taux zéro sur l'accession à la propriété des ménages, leur localisation, et leur bien-être. Le prêt à taux zéro est une politique publique française, créée en 1995, qui vise à améliorer l'accession à la propriété des ménages de revenus modeste et moyen. Elle consiste à octroyer un prêt sans intérêt aux ménages primo-accédants. Nous construisons un modèle d'équilibre spatial où les choix d'accession et de localisation des ménages sont contraints par leur niveaux de revenu et de patrimoine du fait de contraintes de crédit sur le marché de l'immobilier. Ce cadre théorique nous permet d'exprimer l'impact du prêt à taux zéro sous la forme de formules analytiques. Nous distinguons les effets de la politique suivant le niveau de ressources des ménages. Le modèle permet également de distinguer (i) les effets dits d'aubaine qui correspondent au gain d'utilité des ménages qui bénéficient du prêt à taux zéro sans que cela change leur choix d'accession et de localisation, (ii) les gains d'utilité des ménages qui accèdent à la propriété grâce au PTZ et (iii) les effets d'équilibre généraux résultant des variations des prix de l'immobilier. Nous proposons également une méthode afin d'estimer empiriquement les paramètres du modèle grâce à un ensemble de données administratives exhaustives. Une fois les paramètres estimés il est possible de simuler des scénarios contrefactuels et ainsi estimer dans quelle mesure le prêt à taux zéro a influé sur le pourcentage de propriétaires, par catégories de revenu, et sur la localisation des ménages.

Le deuxième chapitre a été réalisé avec Florian Bonnet. Il porte sur l'évolution des inégalités infra-départementales depuis cinquante ans. Tout d'abord, nous avons reconstitué la distribution des revenus de chaque département de France métropolitaine pour les périodes 1966-1969, 1986-2017, grâce à la collecte et à la numérisation de très nombreuses archives du

ministère des finances. Nous constatons que l'inégalité totale entre les foyers fiscaux provient en quasi majorité de l'inégalité au sein de chaque département. Plus précisément, une grande partie de l'inégalité totale vient de l'inégalité au sein des départements où sont situés les plus grandes villes françaises et dans une moindre mesure des départements près des côtes atlantiques, méditerranéennes ou des Alpes. Cette forte inégalité au sein de ces départements s'explique en grande partie par une surreprésentation des foyers fiscaux les plus riches. Cette surreprésentation s'est accrue. Ainsi un tiers des 1% des foyers fiscaux français les plus riches habitent dans l'un des trois départements suivants : Paris, Hauts-de-Seine, Yvelines. Pour autant il n'y a pas d'homogénéité en Ile de France comme en témoigne le cas de la Seine-Saint-Denis où les foyers fiscaux les plus riches sont de moins en moins représentés. Le troisième chapitre a été réalisé avec Florian Bonnet et Hippolyte d'Albis. Il porte sur l'évolution des inégalités interdépartementales de revenu depuis un siècle. La contribution la plus notable est la reconstruction du revenu moyen par département avant et après impôt sur le revenu grâce à la collecte et la numérisation inédite d'archives du ministère des Finances. Nous observons les faits stylisés suivants: (i) une très forte baisse des inégalités interdépartementales de revenu fiscal par adulte sur un siècle, avec deux périodes de baisse continue : entre 1922 et 1939 et depuis 1948 ; (ii) l'impôt sur le revenu réduit significativement les inégalités interdépartementales mais son effet a fortement varié au cours de la période considérée ; (iii) depuis 1948, tous les départements au sud de la ligne allant du Calvados au Gard ont vu leur situation relative s'améliorer.

Pour résumer le lien entre ces deux chapitres, l'inégalité entre départements est faible lorsque l'on compare les revenus moyens. Cette relative homogénéité masque cependant l'inégale répartition spatiale des plus hauts revenus. Ces derniers se concentrent fortement dans les départements où sont localisées les métropoles. Ces départements métropoles sont par conséquent bien plus inégalitaires que les autres.

Le quatrième chapitre a été réalisé avec Nicolas Jannin. Il s'intéresse à la question, classique en économie, mais peu étudiée empiriquement, du niveau optimal de provision des biens

publics locaux. Ce chapitre propose une méthode simple mais flexible pour tester si les ménages d'une commune valorisent non seulement les biens publics de leur commune mais aussi les biens publics des communes environnantes. Ces effets sont désignés par le terme spillover effect en anglais et jouent un rôle essentiel pour déterminer le niveau optimal de la provision de bien public. Nous construisons un modèle structurel d'équilibre spatial dans lequel des ménages mobiles choisissent leur commune de résidence, les communes étant caractérisées notamment par les biens publics qu'elles proposent. Ces biens publics sont susceptibles d'être également utilisés par les ménages des communes environnantes. Nous montrons, dans le cas de la France et de ses presque 36 000 communes, comment il est possible d'exploiter les réponses comportementales des ménages à des variations des biens publics locaux, pour estimer l'intensité des effets de spillover. Nous instrumentons les variations de biens publics locaux par des variations plausiblement exogènes des subventions d'investissement aux communes. Nous trouvons que les ménages valorisent à hauteur de 4 à 11% les biens publics de leurs propres communes et à hauteur de 3.2 à 3.5% les biens publics de chacune des communes qui font partie de leur établissement public de coopération intercommunale.

Le dernier chapitre s'intéresse au lien entre prix de l'immobilier à Paris et pourcentage d'acheteurs étrangers. Les prix de l'immobilier ont augmenté de façon spectaculaire à Paris durant les vingt dernières années et les étrangers sont souvent accusés par les médias d'être responsables des prix exorbitants qui ont cours dans la capitale. Je montre au contraire que l'impact des étrangers sur les prix, bien que statistiquement significatif, a été négligeable comparé à la hausse globale qui a eu lieu. Selon mes estimations, la proportion d'acheteurs étrangers a bien eu un impact positif et significatif sur les prix : une augmentation d'un point de % de la proportion d'étrangers parmi l'ensemble des acheteurs entraîne une augmentation des prix de 0,370%. Entre 1993 et 2008, la part des étrangers dans les achats a cru : elle est passée de 6% à 8% pour les acheteurs résidant à l'étranger et de 3% à 9% pour les acheteurs de nationalité étrangère. Les étrangers seraient donc responsables au plus d'une augmentation des prix de 3%, soit 3% de l'augmentation qui a effectivement eu lieu. Bien qu'ils achètent des



biens jusqu'à 30% plus chers, les étrangers ne sont pas assez nombreux pour avoir entraîné une hausse substantielle des prix. D'ailleurs, plus on considère des gammes de prix élevés, plus l'augmentation de la proportion d'acheteurs étrangers a été importante et plus l'augmentation de prix a été, elle, mesurée. C'est pourquoi malgré une forte augmentation de la proportion d'acheteurs étrangers sur le segment de l'immobilier haut de gamme, l'impact sur les prix de ce segment de marché n'aurait pas dépassé 3%.



# Summary

This PhD dissertation gathers five essays on housing policies, spatial inequalities and housing markets.

The first chapter, co-authored with Gabrielle Fack et Laurent Gobillon, studies how the no-interest loans affects the welfare of households using a spatial equilibrium model. We build a model where housing tenure, location choice, and borrowing constraints on the credit market are jointly modelled. In our model, households heterogeneous with respect to income and preferences can be renters or homeowners, and choose their location within a city. Borrowing constraints prevent low-income households from accessing home-ownership in locations where housing prices are too high. We derive closed-form formulas for the welfare effects at the intensive and extensive margins. We also propose an estimation strategy to take the model to the data.

The second chapter is co-authored with Florian Bonnet. We combine data and methods with a view to analyse the evolution of intra-regional inequality in France over half a century. Most notably, using numerous unexplored archives of the French Ministry of Finance, we reconstruct the income distribution of each French *département* since 1966. We find that the total inequality within tax units stems mostly from inequalities within *départements*. A big part of the total inequality comes from inequalities within *départements* where big cities are located, in particular in the Paris region. This arises mostly from the upper end of the French national income distribution as the richest tax units are getting increasingly over-represented

in *départements* where big cities are located. Three *départements* of the Paris region, Paris itself, Hauts de Seine et Yvelines accounts for one third of French top 1% tax units – based on income distribution -. Seine-Saint-Denis, another *département* of the Paris region, has an opposite trajectory than the rest of its region, as there are increasingly less rich tax units living there.

The third chapter is co-authored with Florian Bonnet and Hippolyte d’Albis. We combine different sources and methods in order to analyse the long run evolution of between regional inequality in France. Most importantly we reconstruct the average – pre and post income tax – income of each French *département* since 1922. We collect and digitize new archives of the French Ministry of Finance. We uncover the following stylized facts: (i) a strong decrease of between *département* inequality in the last century with two sub-periods: 1922-1939 and 1948 onwards; (ii) the income tax reduces significantly between *département* inequality, however its effect has greatly varied over the period; (iii) since 1948, the relative situation of all *départements* situated south of a line running from Calvados to Gard, has improved.

To sum up the link between these two chapters, the inequality between *départements* is low when we compare average income. However, this relative homogeneity conceals the unequal spatial location of top incomes. Top incomes are concentrated in *départements* where the biggest French cities are located. By consequences, these *département* are more unequal than the others.

The fourth chapter is co-authored with Nicolas Jannin. Despite long-standing theoretical interest, empirical attempts at investigating the appropriate level of decentralization remain scarce. This paper develops a simple and flexible framework to test for the presence of public good spillovers between fiscally autonomous jurisdictions and to investigate potential welfare gains from marginal fiscal integration. We build a quantitative spatial equilibrium model with many local jurisdictions, mobile households and endogenous local public goods causing spillovers across jurisdictional boundaries. We show how one can exploit migration

and house price responses to shocks in local public goods at different geographic scales to reveal the intensity of spillovers. Applying our framework to the particularly fragmented French institutional setting, we structurally estimate the model using a unique combination of jurisdiction-level administrative panel datasets. Estimation relies on plausibly exogenous variations in government subsidies to instrument changes in the supply of local public goods. We find that public spending in a municipality accounts for 4–11% of the local public good bundle enjoyed by its residents, and that public spending in each neighbor municipality accounts for an average 3.2–3.5% of this bundle. Finally, we simulate the effect of a reform increasing fiscal integration and find substantial welfare gains.

The fifth chapter investigates the link between foreign buyers and housing prices in Paris. Housing prices have increased dramatically in Paris in the last twenty years – circa 100% in constant euros between 1993 and 2008. The mass media regularly accuses foreign buyers to be responsible for this rise. I show that, on the contrary, foreign buyers impact on housing prices in Paris is, albeit statistically significant, negligible. I estimate that an increase of one percentage point of the share of foreign buyers leads to an increase of housing price of around 0,370%. Between 1993 and 2008, the share of foreign buyers has increased from 6% to 8%. Hence, foreign buyers are responsible of less than 1% of the housing prices increase in Paris.



# Chapter 1

## Welfare effects of the no-interest loans in France<sup>1</sup>

### Abstract

We study how the no-interest loans affects the welfare of households using a spatial equilibrium model. We build a model where housing tenure, location choice, and borrowing constraints on the credit market are jointly modelled. In our model, households heterogeneous with respect to income and preferences can be renters or homeowners, and choose their location within a city. Borrowing constraints prevent low-income households from accessing home-ownership in locations where housing prices are too high. We derive closed-form formulas for the welfare effects at the intensive and extensive margins. We also propose an estimation strategy to take the model to the data.

**JEL Classification:** R13, R21, R31

**Keywords:** no-interest loan, home-ownership, housing, construction, housing prices

---

1. This chapter is based on a joint work with Gabrielle Fack et Laurent Gobillon.

## Introduction

In many large cities, the steep increase in housing prices has made dwellings in attractive locations unaffordable for low and medium-income households. The efficiency of subsidized loans to access home-ownership in such locations is arguable. Those loans may rather increase the demand for poor-quality dwellings in places lacking amenities and job opportunities. Ultimately, they could even generate sprawl and segregation due to the construction of entire neighborhoods of affordable dwellings at the periphery of cities.<sup>2</sup> These effects might be avoided if subsidized loans were spatially targeted, but their net effect on welfare might be limited by local capitalization in housing prices.

In this paper, we propose a cost-benefit analysis of subsidized loans for households with heterogeneous income and preferences. We are particularly interested in the arbitrage between access to home-ownership and attractive location, and the extent to which subsidized loans can generate sprawl. We study the effects of the no-interest loan (NIL) in France that was introduced in 1995 and is still the main subsidized loan nowadays. Over the 1995-2017 period, this loan benefited yearly to 140 thousand households accessing home-ownership for an average yearly cost of 1.04 billion euros for the government.<sup>3</sup> NILs target renters, and access conditions and granted amount of loan depend on income, family size, location and dwelling type, ie. whether dwellings are new or second-hand. There are important variations in the NILs scheme during our period of study. In particular, they are initially restricted mostly to the purchase of new dwellings but the restriction on dwelling type is suppressed in 2005 before being reintroduced in 2012. There have also been significant changes in the geographical areas determining the subsidy amount such that initially only the Paris region was distinguished from the rest of the territory, whereas now medium and large cities are singled out.

We develop a spatial equilibrium model of tenure and location choices within a city.

---

2. In the case of France, see <https://www.lesechos.fr/idees-debats/editos-analyses/le-logement-nest-pas-quune-affaire-de-taux-1146057>.

3. The figure on the yearly number of loans is computed from SGFGAS data presented below and the yearly cost is computed from information available at the address: <https://www.data.gouv.fr/fr/datasets/compte-satellite-du-logement/>.



Households differ in their income and preferences, and they can be homeowners or renters. Borrowing constraints restrict access to credit and thus home-ownership. Locations within the city are characterized by different housing prices and amenities, and households may become homeowners only in a subset of locations. The no-interest loan loosens the constraints for targeted households and changes the subset of locations accessible for home-ownership. We obtain closed formulas for the effects of NILs on welfare both at the extensive margin, for households able to access home-ownership thanks to NILs, and at the intensive margin for households who would access home-ownership even in absence of NILs but still benefit from NILs as they may locate in a more attractive location.

We propose an estimation strategy to quantify welfare effects that can be decomposed into several steps. Estimating the income thresholds to access home-ownership is the first step. We derive computable structural formulas for income thresholds from our model. The second step is to determine the distribution of individuals across locations, tenure, dwelling types, family sizes and income brackets. The third step is to recover the model parameters by maximizing the likelihood of individual choices. In an extended version of this paper we will take our model to the data following this strategy.

The rest of this chapter proceeds as follows. Section 2 relates the paper to the existing literature. Section 3 gives some institutional background and descriptive statistics on no-interest loans in France. Section 4 details the theoretical model. Section 5 explains how to take the model to the data. Section 6 concludes.

## 1.1 Literature on place based policies and borrowing constraints

In our analysis, no-interest loans are modelled as a place-based policy. Our theoretical approach is thus close to equilibrium approaches used to study enterprise zones [Busso et al. \(2013\)](#) and transportation [Tsivanidis \(2018\)](#). Nevertheless, we develop further the housing market by allowing for two tenure choices and introducing imperfections on the mortgage

market in the form of borrowing constraints. The no-interest loan then constitutes a way to alleviate these imperfections and has thus a theoretical justification for its use to increase welfare (see [Kline and Moretti \(2014\)](#) for a general discussion). The welfare analysis is made more complex though by borrowing constraints since it is not possible to use a sufficient statistic approach as proposed by [Chetty \(2009\)](#).

Our model considers both sides of the housing market in any location. Characteristics of local housing supply functions are identified from changes in local housing demand due to space-and-time variations in the design of the no-interest loan. These variations are considered to be exogenous as decisions to change the design are mechanical rather than based on precise economic information. We thus depart from [Baum-Snow and Han \(2019\)](#) who identify local housing supply from local productivity shocks that affect local housing demand. Whereas their identification strategy requires the modelling of local labor markets, it is not the case here and we thus abstract from it.

As we argue, the no-interest loan can affect the stratification of space according to income as low and medium households purchasing a home thanks to NILs may need to move to less attractive places at the periphery of cities. Our work complements recent studies studying spatial income resorting between central cities and suburbs. In particular, [Couture and Handbury \(2017\)](#), [Baum-Snow and Hartley \(2019\)](#) and [Couture et al. \(2019\)](#) show that US city centers gentrified due to an increased demand for consumption amenities located downtown, and that poor households moved out. The important role of consumption amenities in shaping income sorting is quantified by [Gaigne et al. \(2020\)](#).

The restrictive role of borrowing constraints has been analyzed theoretically by [Ortalo-Magne and Rady \(2006\)](#) who show how workers adjust their tenure and dwelling type over the life-cycle depending on income shocks and wealth. In our setting, adjustments can only occur for tenure and location, since dwellings are homogeneous. [Gobillon and Le Blanc \(2008\)](#) estimate an a-spatial model of mobility and tenure choice in presence of borrowing constraints over the period preceding the introduction of NIL. They then simulate NIL as a down-payment subsidy and show that it generates a large windfall effect.

Our setting considers spatial variations of housing prices within city when assessing the local access to home-ownership depending on income. Such variations have already been quantified by [Albouy et al. \(2018\)](#) for the US and [Combes et al. \(2019\)](#) for France. They show that there are steeper land and housing price gradients in large cities, but at the same time large variations in these gradients across cities. [Landoigt et al. \(2015\)](#) estimate a quantitative assignment model in presence of borrowing constraints to understand spatial variations in housing prices within San Diego. The housing market is more elaborate than in our setting since dwellings are heterogeneous and households sort across dwelling types, but space is not explicitly modelled.

## 1.2 The no-interest loan policy

### 1.2.1 Institutional background

The no-interest loan scheme was created in 1995 in France with the aim to ease access to home-ownership for low to moderate income households. The scheme is means tested and since 1997, restricted to first-time buyers. It has become the most prominent policy scheme to help low-income households to buy a home in France: in 2017, the implied subsidy paid by the State to banks for no-interest loans amounted to around 1.7 billion euros, to be compared to 0.1 billion euros for other types of State-guaranteed loans, and 0.8 billion euros for housing subsidies directly distributed to home-owners (*Comptes du Logement 2017*).<sup>4</sup>

The no-interest loan can be distributed by all banks, which receive a subsidy or a tax credit from the State to compensate for the foregone interest payments. The main criteria that determine eligibility and amounts are: household income, family size, location, previous tenure status and the type of property (new or existing).

In addition to its zero-percent interest rate, the no-interest loan usually includes an initial grace period, allowing households to defer loan repayment or reduce the amount paid for

---

4. Direct housing subsidies initially represented the largest subsidy, amounting to around 2 billion euros in 1995, but they have been gradually reduced and are scheduled to completely disappear by 2020.

several years. The loan is equivalent to a subsidy for households equal to the difference between the loan amount and the net present value of all loan repayments. The loan profile (overall term, grace period) is a function of the borrower's income: the lower the income, the longer the term and the grace period, the bigger the subsidy.<sup>5</sup>

Since 1995, the program has gone through a number of policy reforms, which induced substantial variations in the subsidy over time and across space. In particular, while newly built properties have been the main priority of the program and are eligible since its inception, this was not always the case for existing units, and the amount of subsidy evolved differently by type of property. Figure 1.1 shows the evolution of the maximum amount that the subsidy represents for an eligible household, as a percentage of average household income, for new and existing units. Initially, households who purchased an existing property were not eligible for a no-interest loan, except if a major refurbishment was planned, in which case the implied subsidy was the same as for the purchase of a new unit, and represented on average 28% of one year's household income – the subsidy is received once –. Between 2005 and 2011, program eligibility was extended to all types of units, either existing or new, but the subsidy for existing units was continuously lowered, while it was substantially reevaluated for newly built properties after 2007. The subsidy for newly-built units peaked in 2009, when its amount reached almost 65% of one year's household income, before reverting to around 35%. After a three-years period where existing property stopped being eligible, no-interest loans can again be distributed for the purchase of existing property with major refurbishment since 2015, but the implied subsidy remains much lower for existing properties than for new units.<sup>6</sup> These dramatic changes in the size of the subsidy amount and the type of eligible units can be exploited as exogenous variations of the policy.

In addition, another source of variation comes from changes in the income eligibility threshold over time, as shown in Figure 1.2. While 80% of households were initially below the income eligibility thresholds, this share gradually declined to 70% until 2011, when the

---

5. See Appendix D for a detailed description of the policy parameters and calculation of the subsidy

6. In order to be eligible, the costs of renovation of an existing property had to represent at least 35% of the total costs, including renovation and the housing price before 2005, and 25% since 2015. Since 2013, units also have to satisfy energy efficiency requirements.

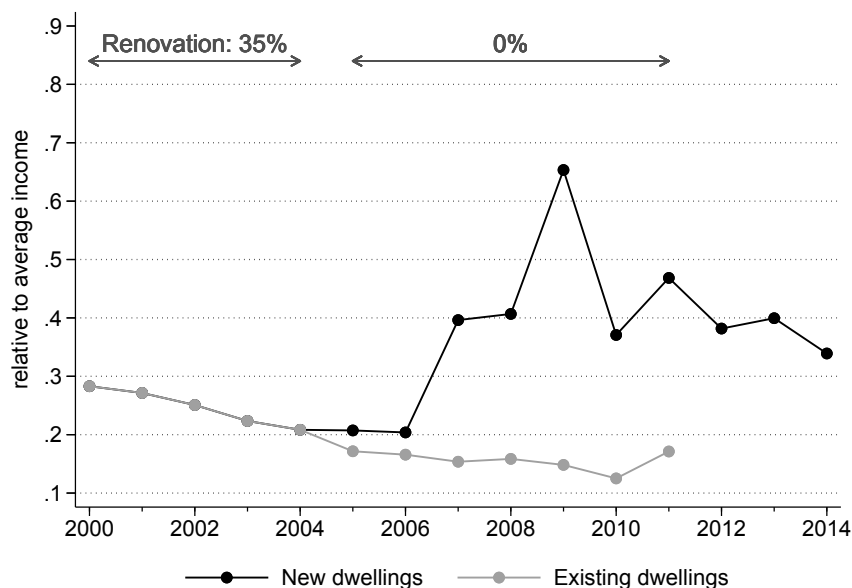


FIGURE 1.1 – EVOLUTION OF THE NO-INTEREST LOAN SUBSIDY FOR NEW AND EXISTING UNITS

*Notes:* This figure shows the evolution of the theoretical amount of subvention for the purchase of new or existing units by households eligible for a no-interest loan. This theoretical amount is calculated taking the average of the maximum amount that each eligible tenant household could claim if he decided to use a no-interest loan to buy a property in his municipality of residence (according to income and family size), for all tenants living in France. The amount is calculated as a share of household income. In 2000, the subsidy represented 28% of the average income of households that year. The subsidy is received once. Sources: FILOCOM.

income eligibility threshold was temporarily lifted. These variations in the income threshold significantly affect the number of distributed loans over time. Importantly only first-time buyers are eligible which reduce significantly the effective share of eligible households. Figure 1.3 represents the share of eligible households among renters and among all households. In 2000, around 90% of renters were below the income eligibility threshold and hence were eligible to the policy. As homeowners are not eligible, only 40% of all households were eligible in 2000. This share has slightly decreased over time if we exclude the particular case of 2011.

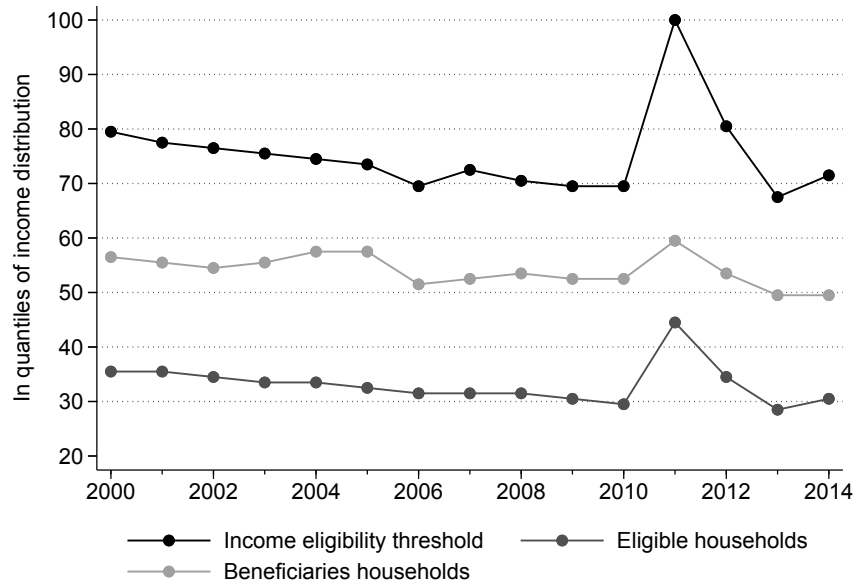
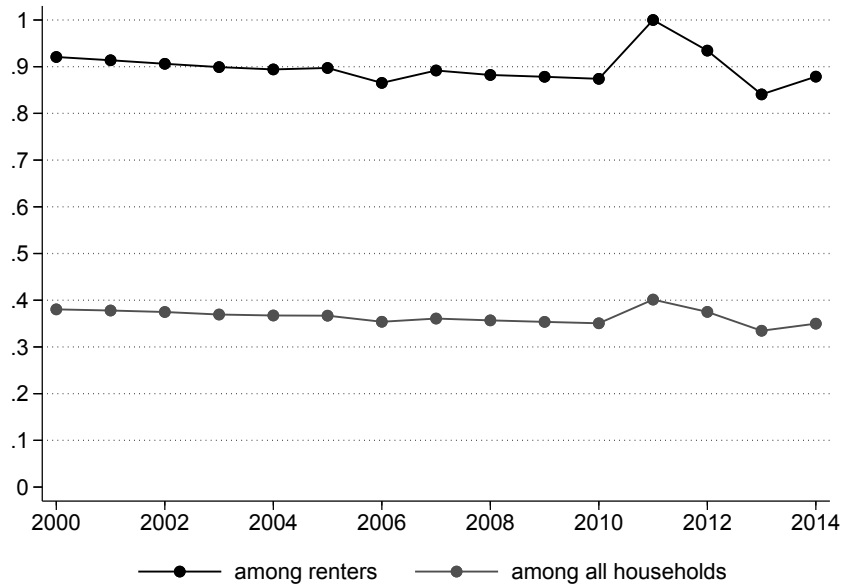


FIGURE 1.2 – EVOLUTION OF THE INCOME ELIGIBILITY THRESHOLD, INCOME OF ELIGIBLE HOUSEHOLDS AND BENEFICIARIES FOR A NO-INTEREST LOAN TO FINANCE THE PURCHASE OF A NEW UNIT

*Notes:* This figure shows the evolution of the income eligibility threshold, the average income of eligible households and the average income of beneficiaries of the no-interest loan in France. All measured in quantiles of the household income distribution (both renters and homeowners). In 2000, 80% of all households were below the income eligibility threshold. 36% of households were below the average income of eligible households. 57% of households were below the average income of beneficiaries. Sources: FILOCOM, SGFGAS.

Finally, the amount of subsidy varies with geographical location. A specific classification of areas was created in 1995, to allow for higher subsidies for purchases of dwellings in tight housing markets. This classification was substantially modified over time, as shown in Figure 1.4. When the program was created, only two categories were defined: the Paris region, where the subsidy was highest, and the rest of France. Over time, the classification was refined several times (in 2005, 2006, 2007, 2011, 2015), and the current classification, implemented in 2011, has four different categories: the subsidy is highest for units located in urban areas with tight housing markets (classified as “A”) and lowest for small urban units and rural areas (classified as “C”). This creates exogenous variations in subsidy amounts across and within



*Notes* This figure shows the evolution of the share of eligible households among all households and among renters. Only first-time buyers are eligible to the no-interest loan NIL. We assume here that all renters do not own a dwelling. In 2000 around 90% of renters and 40% of all households were eligible to the NIL. Sources: FILOCOM.

FIGURE 1.3 – SHARE OF ELIGIBLE HOUSEHOLDS AMONG ALL HOUSEHOLDS AND AMONG RENTERS

each metropolitan area over time.

## 1.2.2 Descriptive statistics

We now present some descriptive statistics that motivate this chapter.

Table 1.1 and 1.2 present descriptive statistics on respectively new and existing dwelling NIL beneficiaries during the 2002-2006 period and the 2009-2013 period. Importantly 38% of new dwelling buyers are NIL beneficiaries. This sizeable share is likely to impact new dwelling housing markets. The share of NIL beneficiaries among existing dwelling buyers is smaller, around 7% between 2002-2006 and 15% between 2009-2013. Even though this share is smaller it is still significant. On average NIL beneficiaries are poorer than no NIL beneficiaries – their income is on average 15% smaller –. However the difference is bigger for existing dwelling

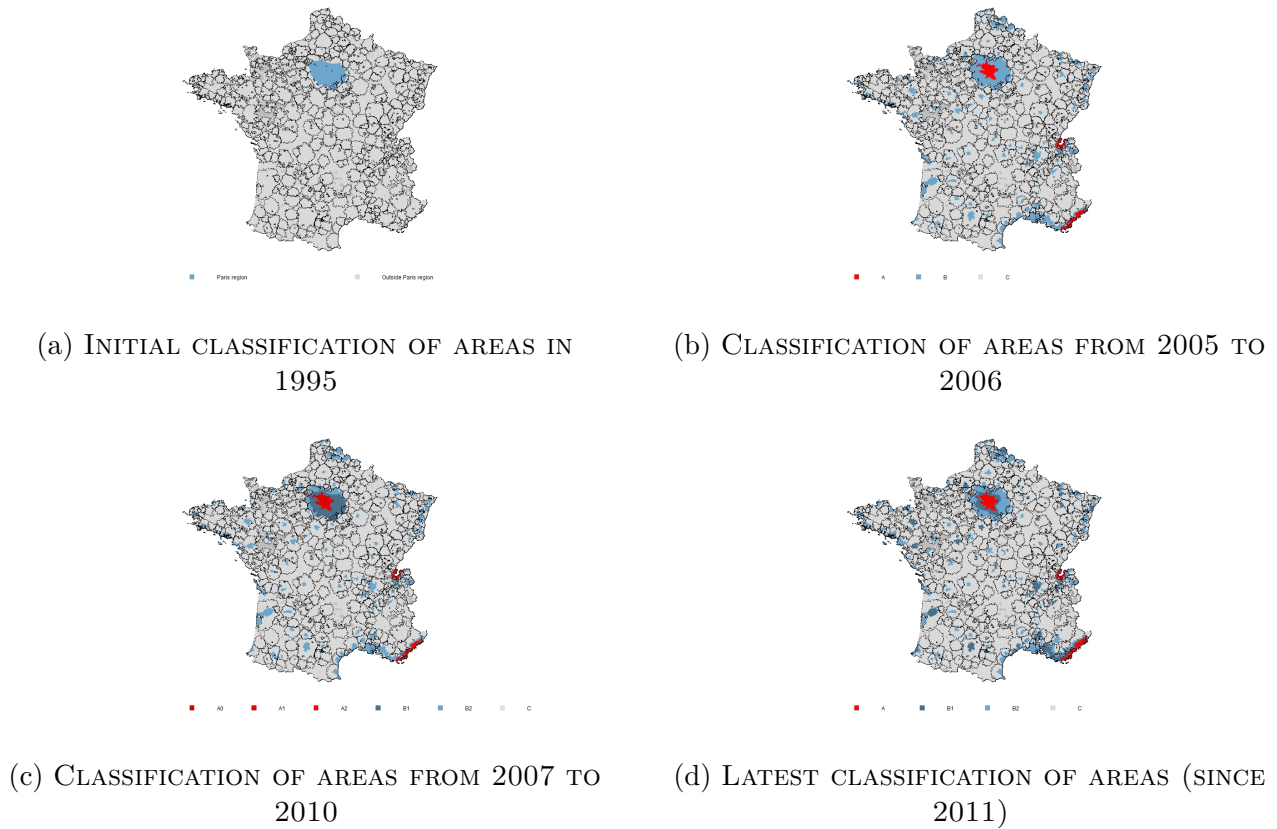


FIGURE 1.4 – EVOLUTION OF THE DEFINITION OF GEOGRAPHICAL AREAS FOR THE SUBSIDY

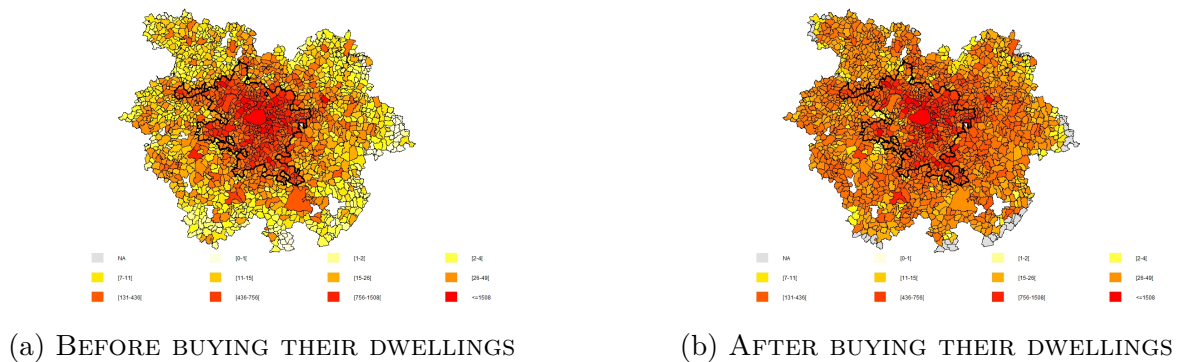
buyers and was bigger during the 2002-2006 period. It is not surprising as in 2011 income eligibility thresholds were lifted for a year. NIL beneficiaries buy on average cheaper dwelling – more than 25% less–. They buy smaller dwelling – around 10% less–. They also pay less in terms of price per square m2. However the difference is bigger for new dwelling – around 25% in 2002-2006, 15% in 2009-2013 versus 12% for existing dwellings. They are indebted more than other buyers to buy their dwellings – the average total amount of their loan is around 6 to 21% bigger than the other buyers.<sup>7</sup> On average between 2002 to 2006 the no interest loan amounted to 14% of the total amount of loans of new dwelling NIL beneficiaries. This share increased to 27% between 2009 to 2013. This share is smaller for existing dwelling NIL beneficiaries, around 12%.

7. This is due to the fact that there is a significant share of no NIL buyers that buy a dwelling without contracting any loan. On average no NIL buyers that do contract a loan contract bigger loan than NIL beneficiaries.



Households face borrowing constraints when they buy a dwelling. These constraints are location dependant as they are function of housing prices, the subsidy and the type of property (new or existing). Figure 1.6 represents the location of the NIL beneficiaries that live in the Paris urban area before and after buying their dwelling. Without any causal implication we notice that NIL beneficiaries relocate away from the center of the urban area.

Borrowing constraints may impact households differentially according to their income. High housing price may prevent low and middle income households to access home-ownership in urban area centers. At the same time low and middle income households benefit from bigger subsidies. As these subsidies primarily target new dwellings, and, as there is more space in urban area outskirts, it might also encourage low and middle income households to live outside urban area centers. Because of general equilibrium effects it is hard to know the effective impact of the policy. In the next section we built a model that conceptualize aforementioned channels.



*Note:* This figure shows the location of NIL beneficiaries that live in the Paris urban area before and after their purchase. Units of observations are French municipalities in the Paris urban area. We use an extended definition of the Paris urban area in which we include municipalities at the fringe of the urban area. Colors represent the number of NIL beneficiaries that live in municipalities before or after their relocation. The black line represents the border of the urban area center. We restrict to NIL beneficiaries that live in the Paris urban area before and after their purchase, over the period 1997 - 2017. 58 233 NIL beneficiaries lived in Paris municipality before their purchase. 26 105 NIL beneficiaries lived in Paris municipality after their purchase. Sources: SGFGAS.

FIGURE 1.6 – LOCATION OF NIL BENEFICIARIES THAT LIVE IN PARIS URBAN AREA

TABLE 1.1 – DESCRIPTIVE STATISTICS ON NIL NEW DWELLING BUYERS

	2002-2006	(sd)	2009-2013	(sd)
% buyers	38		38	
Income relative to no NIL	0.82*		0.94	
Price	147896	(2627)	193304	(7484)
Price relative to no NIL	0.71*		0.73*	
Surface	114	(2)	108	(3)
Surface relative to no NIL	0.93*		0.89*	
Price per m <sup>2</sup>	1361	(25)	1854	(97)
Price per m <sup>2</sup> relative to no NIL	0.76*		0.85*	
Loan	114509	(1623)	159740	(3564)
Loan relative to no NIL	1.21*		1.17*	
NIL amount	17475	(191)	37173	(1798)
NIL amount over price ratio	0.14	(0.005)	0.27	(0.07)

*Note:* This table gives descriptive statistics on new dwelling NIL beneficiaries. % buyers stands for the share of new dwellings buyers that are NIL beneficiaries. Income relative to no NIL stands for the average income of new dwelling NIL buyers relative to the average income of new dwelling buyers that are not NIL beneficiaries. Price stands for the average dwelling transaction amount of new dwelling NIL beneficiaries. Loan stands for the average total loan amount of NIL new dwelling beneficiaries. NIL amount over price ratio stands for the average NIL amount over dwelling price of new dwelling NIL beneficiaries. During the period 2002-2006, around 38% of the new dwelling buyers were NIL beneficiaries. \* means that the two quantities that are being compared are statistically significant at the 95% level. (sd) are standard deviations. Source: 2006 and 2013 French National Housing Surveys

TABLE 1.2 – DESCRIPTIVE STATISTICS ON NIL EXISTING DWELLING BUYERS

	2002-2006	(sd)	2009-2013	(sd)
% buyers NIL	7		18	
Income relative to no NIL	0.75*		0.83*	
Price	138200	(5794)	191083	(7336)
Price relative to no NIL	0.71*		0.74*	
Surface	89	(4)	91	(3)
Surface relative to no NIL	0.87*		0.89*	
Price per m <sup>2</sup>	1812	(94)	2366	(95)
Price per m <sup>2</sup> relative to no NIL	0.87*		0.88*	
Loan	111889	(4933)	148244	(4632)
Loan relative to no NIL	1.06		1.13*	
NIL amount	14559	(568)	20865	(984)
NIL amount over price ratio	0.12	(0.01)	0.12	(0.01)

*Note:* This table gives descriptive statistics on existing dwelling NIL beneficiaries. % buyers stands for the share of existing dwellings buyers that are NIL beneficiaries. Income relative to no NIL stands for the average income of existing dwelling NIL buyers relative to the average income of existing dwelling buyers that are not NIL beneficiaries. Price stands for the average dwelling transaction amount of existing dwelling NIL beneficiaries. Loan stands for the average total loan amount of NIL existing dwelling beneficiaries. NIL amount over price ratio stands for the average NIL amount over dwelling price of existing dwelling NIL beneficiaries. During the period 2002-2006, around 7% of the existing dwelling buyers were NIL beneficiaries. \* means that the two quantities that are being compared are statistically significant at the 95% level. (sd) are standard deviations. Source: 2006 and 2013 French National Housing Surveys

## 1.3 Model

We now present our spatial equilibrium model of tenure and location choices. Our goal is to derive a relationship between welfare and no-interest loan than can be brought to the data. We abstract from the choice between old and new dwellings in this section for simplicity, but we will re-introduce it in the empirical section.

### 1.3.1 The setting

We consider that cities are segmented labor markets and we propose a model for one given city. There is a finite number of locations  $L$  within the city and each individual chooses a location  $l$  where to live. For simplicity, we do not consider the workplace location as it is not central to our topic. Individual income is denoted  $I$ , and it has a distribution density  $f(I)$ .

### 1.3.2 The housing market

Each individual chooses their housing market  $m$  between rental ( $m = r$ ) and home-ownership ( $m = o$ ), and markets are characterized by location-specific user costs of the dwelling. For simplicity, individuals consume one unit of housing and, in a given location  $l$ , user costs of rental and home-ownership are supposed to be the same and equal to  $\pi_l$ .

In a given location, dwellings can be bought at a fixed unit price  $p_l$ . Individuals have borrowing constraints on the credit market when purchasing a dwelling. Those who are constrained cannot purchase a dwelling. We consider that an individual with a given income  $I$  who wants to purchase a dwelling may benefit from a state-subsidized zero-interest loan which is equivalent to a rebate on downpayment that depends on the location and income,  $R_l(I)$ .<sup>8</sup> In the sequel, we will consider a rebate scheme that is step-wise constant such that

---

8. Alternatively, individuals may be characterized by a skill level and income could be derived from it according to some given returns. In fact, if returns do not depend on location within the city, and income is strictly increasing in the skill level, there is a one-to-one mapping between income and skills. It is then equivalent to use one or the other in the model.

thresholds do not depend on location (but rebate does), consistently with the French subsidy scheme:  $R_l(I) = \sum_{k=1}^K 1_{\theta_{k-1} \leq I < \theta_k} R_l^k$  with  $\theta_0 = 0$  and  $\theta_K = +\infty$  (and possibly  $R_l^k = 0$  for some  $k$  if there is no subsidy in an income bracket). It is shown in appendix that the usual income and downpayment constraints on the credit market can be summarized with the following single equation provided that downpayment is proportional to income (ie. there is no wealth coming from sources other than income, such as inheritance):

$$y(p_l, R_l(I)) \leq I \tag{1.1}$$

with  $y(\cdot)$  a function that is positive, increasing with dwelling price but decreasing with the rebate. The left-hand side term can be interpreted as the income threshold above which it is possible to access home-ownership. Note that for income  $I \in [\theta_{k-1}, \theta_k[$ , the left-hand side is constant since  $R_l(I) = R_l^k$ , hence the income threshold is piece-wise constant and we denote  $y_l^k = y(p_l, R_l^k)$ .

### 1.3.3 Individual choice

A consumption good used as a numeraire is produced with labour and capital by a representative firm under constant returns to scales. Individuals are employed by this representative firm and get a wage such that their overall income  $I$  is the sum of non-labour income and wage. Wages are pinned down by the city productivity level as shown in Appendix A and do not depend on location.

Income is used to consume the numeraire and pay the housing user cost. The indirect utility of an individual  $i$  with income  $I$  living in location  $l$  in a dwelling in market  $m$  is given by:

$$u_{ilm}(I) = v_{lm}(I, R_l(I), p_l) + \epsilon_{ilm} \tag{1.2}$$

where we write the utility as a function of price instead of user cost as we assume that  $\pi_l = rp_l$

with  $r$  the interest rate,<sup>9</sup> with  $v_{lm}(\cdot)$  a function increasing with  $I$  and  $R_l(I)$  but decreasing with  $p_l$ , and  $\epsilon_{ilm}$  is an idiosyncratic term that captures heterogeneity in tastes. Idiosyncratic terms are assumed to be independently and identically distributed across individuals.

An individual chooses her residential location and market by maximizing her utility. Note that she does not have access to dwellings in home-ownership in locations where the borrowing constraint 1.1 is not verified. Hence, her choice set is constrained. The choice set for an individual with income  $I$  is given by:

$$\mathcal{B}_I = \{(l, r) \text{ for all } l\} \cup \{(l, o) \text{ for } l \text{ such that } y(p_l, R_l(I)) \leq I\} \quad (1.3)$$

The average utility reached by individuals with income  $I$  is:

$$V(I) = E_i \left[ \max_{l, m \in \mathcal{B}_I} u_{ilm}(I) \mid I_i = I \right] \quad (1.4)$$

where  $I_i$  denotes the income of individual  $i$ .

### 1.3.4 Housing supply

Finally, each location has a continuum of absentee risk-neutral land owners distributed on the unit interval. Each land owner may provide a dwelling on the market for a cost which is distributed according to the cumulative distribution function  $H_l(\cdot)$  with strictly positive support. If the land owner does not provide a dwelling, she does not receive any pay-off. If she does, she can then rent her dwelling or sell it (and then invest the money in riskless asset before purchasing the dwelling back at the end of the period). Pay-offs associated to the two options are the same and equal to  $rp_l$  such that land owners are indifferent between selling and renting.

Let  $n_l$  denote the number of dwellings that are rented or sold in location  $l$ . Optimization

---

9. This assumption is consistent with the user cost being the yearly foregone revenues for not buying riskless asset with return  $r$  in quantity  $p_l$  when the housing price does not vary over time.

implies that the marginal land owner in each location breaks even so that:

$$H_l^{-1}(n_l) = \pi_l = rp_l \tag{1.5}$$

To close the model, we assume that housing markets clear such that:

$$n_l = N_l \tag{1.6}$$

where  $N_l$  is the measure of individuals living in city  $l$ .

### 1.3.5 Welfare implications

Total social welfare for the economy can be decomposed into the aggregate welfare of individuals  $V$  and the aggregate profit of land owners for the whole economy  $\Pi$ :

$$W = V + \Pi \tag{1.7}$$

We want to assess how welfare varies when the rebate changes in a given location. We first study variations of the aggregate welfare of individuals. For computational tractability, we make the assumption that idiosyncratic shocks  $\epsilon_{ilm}$  follow an extreme value law with unitary variance. According to Small and Rosen (1981), the aggregate welfare of individuals is then given by:

$$V = \gamma + E_I \left[ \ln \left( \sum_{(l,m) \in \mathcal{B}_I} \exp [v_{lm}(I)] \right) \right] \tag{1.8}$$

where  $\gamma$  is the Euler constant and  $v_{lm}(I) \equiv v_{lm}(I, R_l(I), p_l)$ . The income threshold for access to home-ownership is rewritten as  $y_{lo}(I) \equiv y(p_l, R_l(I))$  and we use the convention that the income threshold for access to the rental sector is  $y_{lr}(I) = -\infty$  since that sector is accessible



in every location to every worker whatever her income. Equation (1.8) can be rewritten as:

$$V = \gamma + E_I \left[ \ln \left( \sum_{(l,m)} \exp [v_{lm} (I)] 1\{y_{lm} (I) \leq I\} \right) \right] \quad (1.9)$$

We now assess how this average utility varies when the no-interest loan rebate increases at the margin for the single option  $(l, o)$  for workers with income  $I \in [\theta_{k-1}, \theta_k[$ , ie.  $dR_l^k = \Delta > 0$ . All proofs are relegated to Appendix B.2. Denote by  $\frac{\partial Y}{\partial X} \Big|_{\chi}$  the partial derivative of  $Y$  with respect to  $X$  holding constant the parameters in the set  $\chi$ , and consider the following sets:

$$\begin{aligned} \Phi_l^k &= \{R_{l'}^{k'} \text{ for } (l', k') \neq (l, k), p_{l'} \text{ for all } l'\} \\ \Psi_l &= \{p_{l'} \text{ for } l' \neq l, R_{l'}^{k'} \text{ for all } l'\} \end{aligned}$$

Here,  $\Phi_l^k$  includes prices and rebates for all locations and income brackets except for the rebate in location  $l$  for income bracket  $k$ ;  $\Psi_l$  includes prices for all locations except for the one of location  $l$ . We consider the fact that the average utility of individuals with a given income can be expressed as a function of prices and rebates for all locations and income brackets. The variation in the aggregate welfare of individuals related to the increase in the rebate of interest is then given by:

$$\frac{\partial V}{\partial R_l^k} = \frac{\partial V}{\partial R_l^k} \Big|_{\Phi_l^k} + \sum_{l'} \frac{\partial p_{l'}}{\partial R_l^k} \frac{\partial V}{\partial p_{l'}} \Big|_{\Psi_{l'}} \quad (1.10)$$

The increase in the no-interest loan rebate has two effects. First, it affects the average utility in location  $l$  because it makes it easier to access home-ownership in that location and yields an increase in utility for home-ownership due to the money flow related to the rebate (direct effects). Second, it changes housing prices in all locations at the equilibrium and this affects access to home-ownership and utility of individuals in all locations (indirect effects). Note that equation (1.10) does not involve any derivatives of incomes with respect to change in rebate because incomes are pinned down by the city productivity level which is fixed.

We want to make more explicit the effects of the increase in rebate on utilities of individuals

who are able to access home-ownership thanks to this increase (extensive margin) and on utilities for individuals who would have accessed home-ownership even in the absence of this increase (intensive margin). For that purpose, we denote the income threshold in location  $l$  when the rebate is that of the  $k^{th}$  income bracket as  $y_l^k = y(p_l, R_l^k)$  and we introduce the following sets:

$$\begin{aligned}\Theta_{lm} &= \left\{ v_{l'm'} \text{ for } (l', m') \neq (l, m), y_{l'}^{k'} \text{ for all } (l', k') \right\} \\ \Omega_l^k &= \left\{ \left( y_{l'}^{k'} \right) \text{ for } (l', k') \neq (l, k), v_{l'm'} \text{ for all } (l', m') \right\}\end{aligned}$$

Here,  $\Theta_{lm}$  includes fixed components of utility and income thresholds to access home-ownership for all locations, housing tenure and income bracket except for the fixed component of utility for location  $l$  and housing tenure  $m$ ;  $\Omega_l^k$  includes the income thresholds to accept home-ownership except that for location  $l$  and income bracket  $k$ . We decompose further the right-hand side terms in equation (1.10) by considering now that the average utility of individuals with income  $I$  can be expressed as a function of fixed components  $v_{lm}(I)$  and income thresholds to access home-ownership  $y_{lo}(I)$ , instead of housing prices and rebates. The first term of equation (1.10) can be decomposed as follows:

$$\left. \frac{\partial V}{\partial R_l^k} \right|_{\Phi_l^k} = \left. \frac{\partial y_l^k}{\partial R_l^k} \frac{\partial V}{\partial y_l^k} \right|_{\Omega_l^k} + E_I \left[ \left. \frac{\partial v_{lo}(I)}{\partial R_l^k} \frac{\partial V(I)}{\partial v_{lo}} \right|_{\Theta_{lo}} \right] \quad (1.11)$$

The first right-hand term captures the jump in utility for individuals in the  $k^{th}$  income bracket who are able to become homeowners in location  $l$  thanks to the rebate increase that allows them to borrow enough money on the credit market (extensive margin). The second one corresponds to the windfall effect such that individuals in the  $k^{th}$  income bracket who can purchase a dwelling in location  $l$  in absence of the rebate increase experience a rise in their utility thanks to the rebate (intensive margin). The second term of equation (1.10) can be

decomposed as follows:

$$\left. \frac{\partial V}{\partial p_{l'}} \right|_{\Psi_{l'}} = \sum_k \frac{\partial y_{l'}^k}{\partial p_{l'}} \left. \frac{\partial V}{\partial y_{l'}^k} \right|_{\Omega_{l'}^k} + \sum_m E_I \left[ \left. \frac{\partial v_{l'm}(I)}{\partial p_{l'}} \frac{\partial V(I)}{\partial v_{l'm}} \right|_{\Theta_{l'm}} \right] \text{ for all } l' \quad (1.12)$$

The first right-hand side term corresponds to the variation in utility in any given location  $l'$  due to the change in the number of individuals able to access home-ownership when housing prices vary due to the rebate increase in location  $l$  (extensive margin). The second one captures the change in utility on the two markets in location  $l'$  because of the variation of housing prices due to the rebate increase holding constant the tenure choice of individuals (intensive margin).

It is possible to compute all the derivatives of the average utility appearing on the right-hand side of equations (1.11 and 1.12). The variation of average utility when changing the income threshold to access home-ownership of the  $k^{th}$  income bracket in location  $l$  can be written as:

$$\left. \frac{\partial V}{\partial y_l^k} \right|_{\Omega_l^k} = 1_{\{y_l^k \in [\theta_{k-1}, \theta_k]\}} f(y_l^k) \ln [1 - P(l, o | I = y_l^k)] \quad (1.13)$$

where  $P(l, o | I)$  is the probability for individual with income  $I$  of choosing a dwelling in home-ownership in location  $l$ . It depends on three quantities. The first one is a dummy for the income threshold  $y_l^k$  to be in the  $k^{th}$  income bracket which determines whether the access of individuals to the credit market is affected by a change in the income threshold. Indeed, if the income threshold was strictly below (resp. above) bracket  $[\theta_{k-1}, \theta_k[$ , individuals concerned by the threshold (ie. those with income  $I \in [\theta_{k-1}, \theta_k[$ ) would (resp. would not) have access to home-ownership whether or not the threshold changes marginally. It is only when the income threshold is in the bracket that individuals with income close to it are affected by its change. The second quantity is the income density of individuals  $f(\cdot)$  computed at the income threshold which captures the number of individuals who are triggered by the change in the income threshold. The last quantity reflects the propensity for triggered individuals to actually choose home-ownership in location  $l$ .

The change in average utility related to an increase in the utility component of home-ownership in location  $l$  due the variation of the rebate is given by:

$$E_I \left[ \frac{\partial v_{lo}(I)}{\partial R_l^k} \frac{\partial V(I)}{\partial v_{lo}} \Big|_{\Theta_{lo}} \right] = E_I \left[ 1_{\{I|I \in [\max(\theta_{k-1}, y_l^k), \theta_k]\}} P(l, o | I) \frac{\partial v_{lo}(I)}{\partial R_l^k} \right] \quad (1.14)$$

It is the average across all income levels of the product of three terms. A dummy determines whether an individual with a given income is in the income bracket such that she is both concerned by the change of rebate and has access to home-ownership. The probability for an individual with that income of choosing home-ownership in location  $l$  captures the quantity of individuals experiencing a change in their utility. Finally, the derivative of the fixed component of utility for home-ownership with respect to the rebate is the variation in utility provided that home-ownership is chosen.

Finally, the variation of average utility related to an increase in the utility component of market  $m$  in location  $l'$  due to the change in the housing price in that location can be rewritten as:

$$E_I \left[ \frac{\partial v_{l'm}(I)}{\partial p_{l'}} \frac{\partial V(I)}{\partial v_{l'm}} \Big|_{\Theta_{l'm}} \right] = E_I \left[ 1_{\{I|y_{l'm}(I) \leq I\}} P(l', m | I) \frac{\partial v_{l'm}(I)}{\partial p_{l'}} \right] \quad (1.15)$$

It is also the average across all income levels of the product of three terms. A dummy determines whether an individual with a given income has access to the market. The probability for an individual with that income of choosing market  $m$  in location  $l'$  captures the quantity of individuals experiencing a change in their utility. Finally, the derivative of the fixed component of utility with respect to the housing price in location  $l'$  is the variation in utility provided that market  $m$  is chosen.

The aggregate profit of land owners is the sum of land-owner profits across locations and is thus given by:

$$\Pi = \sum_{l'} \pi_{l'} N_{l'} - \int_0^{N_{l'}} H_{l'}^{-1}(x) dx \quad (1.16)$$

Variation in this aggregate profit is obtained by deriving this equation with respect to the change of rebate:

$$\frac{\partial \Pi}{\partial R_l^k} = \sum_{l'} \frac{\partial \pi_{l'}}{\partial R_l^k} N_{l'} + \sum_{l'} [\pi_{l'} - H_{l'}^{-1}(N_{l'})] \frac{\partial N_{l'}}{\partial R_l^k} \quad (1.17)$$

It comes from changes in profits in all locations, as well as variations in the measures of individuals in all locations. Using the zero profit condition (1.5) for the marginal land owners, the second right-hand side term is zero and we finally get:

$$\frac{\partial \Pi}{\partial R_l^k} = r \sum_{l'} \frac{\partial p_{l'}}{\partial R_l^k} N_{l'} \quad (1.18)$$

The change in profit comes from variations of housing prices in all locations and the numbers of housing units affected by those variations. Finally, it is possible to recover the change of overall welfare from those of average utility and profit.

In practice, changes in state-subsidized zero-interest loan are discrete and usually occur simultaneously in different places and income brackets. To compute the variation of welfare, we first introduce the following set:

$$\Upsilon_l^k = \left\{ R_{l'}^{k'} \text{ for } (l', k') \neq (l, k) \right\} \quad (1.19)$$

This set contains all the rebates except the one in location  $l$  for  $k^{th}$  income bracket. We can then consider the total differential of the welfare:

$$dW = \sum_{l,k} \left. \frac{\partial W}{\partial R_l^k} \right|_{\Upsilon_l^k} dR_l^k$$

The variation of welfare when changing the set of rebates from  $\{R_l^k\}_{l,k}$  to  $\{\tilde{R}_l^k\}_{l,k}$  can then be computed as the sum of infinitesimal incrementations which formula has been established:

$$\Delta W = \sum_{l,k} \int_{R_l^k}^{\tilde{R}_l^k} \left. \frac{\partial W}{\partial R_l^k} \right|_{\Upsilon_l^k} dR_l^k$$

### 1.3.6 Extensions

#### 1.3.6.1 Second-hand dwellings

The French no-interest loan system involves different profiles of rebates for new ( $h = 1$ ) and second-hand ( $h = 2$ ) dwellings. It is possible to extend our model to distinguish between these two types of dwellings. For that purpose, denote  $t$  the current period,  $t - 1$  the previous period, and  $n_{lh\tau}$  and  $p_{lh\tau}$  respectively the measure and price of type- $h$  dwellings in location  $l$  at date  $\tau$ .

At a given period, each landowner may provide a new dwelling on the market for a cost which is distributed according to the cumulative distribution function  $H_{l1t}(\cdot)$ . New dwellings are constructed as long as cost is lower than profit and we have  $n_{l1t} = H_{l1t}(rp_{l1t})$ . Dwellings that are new at a given period become second-hand at the next period, and a proportion  $\delta_t$  of second-hand dwellings do not remain on the market because they are not maintained. The dynamic equation determining the measure of second-hand dwellings at current period from measures of old and new dwellings at previous period is:

$$n_{l2t} = (1 - \delta_t)(n_{l2t-1} + n_{l1t-1}) \quad (1.20)$$

In fact, second-hand dwellings are maintained only if maintenance costs are such that it is profitable for landowners to keep these dwellings on the market. Denoting  $H_{l2t}(\cdot)$  the cumulative distribution function of maintenance costs, the proportion of second-hand dwellings exiting the market  $\delta_t$  is then determined from the relationship  $n_{l2t} = H_{l2t}(rp_{l2t})$ . To close the model, we assume that the housing market of each type clears such that  $n_{lht} = N_{lht}$  with  $N_{lht}$  the measure of households choosing a type- $h$  dwelling at period  $t$ . In that setting, the writing of landlord profit for each dwelling type  $\Pi_{ht}$  is similar to (1.16). Considering that rebate can be different for the two types of dwellings in accordance with the French no-interest loan policy and denoting it  $R_{lh}^k$ , the derivative of type- $h$  profit with respect to the corresponding

rebate is similar to (1.18) and verifies:

$$\frac{\partial \Pi_{ht}}{\partial R_{lh}^k} = r \sum_{l'} \frac{\partial p_{l'ht}}{\partial R_{lh}^k} N_{l'ht}$$

Note that the income thresholds to access home-ownership are not the same for the two types of dwellings because of the differences in housing prices and rebates of the no-interest loan.

### 1.3.6.2 Previous housing tenure and family size

In practice, the no-interest loan is restricted to renters accessing home-ownership. We can take into account this specific feature at the individual level by keeping track of individuals' previous tenure that we denote  $m_{t-1}$ . In that case, individuals who are previous homeowners ( $m_{t-1} = o$ ) can borrow on the credit market at the current period to purchase again a dwelling but cannot benefit from a no-interest loan. Finally, the French no-interest loan system also takes into account the family size in the computation of the rebate. Family size may then affect access to home-ownership and location choices.

The full model taking into account previous tenure and allowing for different dwelling types and family sizes is presented in Appendix B. We use its prediction when explaining our empirical strategy below.

## 1.4 Taking our theoretical model to the data

In this section we explain how to bring the full model to the data. We now consider that individuals are characterized by income, family size and previous tenure  $(I, f, m_{t-1})$ . From now on, we also index quantities with time  $t$  since we are going to exploit time variations for identification.

We denote by  $R_{lht}(I, f)$  the rebate at time  $t$  for a dwelling type  $h$  in location  $l$  for individuals with family size  $f$  and income  $I$ , and by  $R_{lht}^{kf}$  its specific value when income is in the  $k^{th}$  bracket. The income threshold  $y_{lht}^{kf m_{t-1}}$  depends on the same indexes but also on previous tenure  $m_{t-1}$  since only previous renters ( $m_{t-1} = r$ ) can benefit from a NIL. To make

tractable the expressions involved in the computation of the derivatives of the average utility, we consider that the utility function is linear such that:

$$u_{ilmht}(I, f, m_{t-1}) = v_{lmht}(I, f, m_{t-1}) + \epsilon_{ilmht}$$

with

$$v_{lmht}(I, f, m_{t-1}) = \mu R_{lht}(I, f) 1_{\{m_{t-1}=r, m=0\}} + \delta p_{lht} + \varsigma(I) + \xi(f) + \lambda_l + \varphi_m + \nu_h + \tau_t$$

where  $\varsigma(I)$  and  $\xi(f)$  are respectively income and family size functions that, as we will see, will not need to be specified;  $\lambda_l$ ,  $\varphi_m$ ,  $\nu_h$  and  $\tau_t$  are respectively location, market, dwelling type and time fixed effects;  $\epsilon_{ilmht}$  are error terms assumed to be identically and independently distributed, and to follow an extreme value law with mean zero and variance equal to one.

### 1.4.1 Estimation of utility parameters

We now explain how to estimate parameters of the utility function, and in particular  $\mu$  and  $\delta$  which will be needed to compute welfare effects. We adopt a likelihood approach for the choice of a location  $l$ , market  $m$  and dwelling type  $h$  conditionally on income  $I$  and family size  $f$  and previous tenure.

When borrowing constraints involve a downpayment constraint as in our case, income thresholds determining access to home-ownership in the different locations depend on the wealth-to-income ratio  $\beta$ . We assume here that  $\beta$  is not observed. The probabilities of choosing  $l, m, h$  option conditional on income, family size and previous tenure  $I, f, m_{t-1}$  are then given by:

$$\begin{aligned} P(l, m, h | I, f, m_{t-1}, t, \phi) &= E_\beta [P(l, m, h | I, f, m_{t-1}, t, \phi, \beta)] \\ &= E_\beta \left[ \frac{\exp [\bar{v}_{lmht}(I, f, m_{t-1})]}{\sum_{(l', m', h')} \exp (\bar{v}_{l'm'h't}(I, f, m_{t-1})) 1\{y_{l'm'h't}(I, f, m_{t-1}, \beta) \leq I\}} \right] \end{aligned} \tag{1.21}$$



where:

$$\bar{v}_{lmht}(I, f, m_{t-1}) = \mu R_{lht}(I, f) 1_{\{m=0\}} 1_{\{m_{t-1}=r\}} + \delta p_{lht} + \lambda_l + \varphi_m + \nu_h$$

Note that income, family size and period terms do not enter expression (1.21) because they are present at the numerator and denominator, and thus cancel out. Note also that, whereas utility terms related to income and family size do not enter choice probabilities, income and family size still influence the amount of the local rebate and the choice probabilities depend on them through the amounts of rebate obtained in every locations.

We can estimate parameters  $\phi$  by maximizing the log-likelihood formed from expression (1.21). Identification comes from variations of housing prices and rebates in the trivariate dimension  $(l, h, t)$  (with the normalization that, for each of the three series of fixed effects for location, market and dwelling type, one fixed effect is fixed to zero). Indeed, if variations were only univariate in the location or dwelling type dimension, the effect of rebate and prices could not be identified separately from those of respectively location and dwelling type fixed effects. If variations were only univariate in the time dimension, rebate and housing price terms would cancel out in the expressions of choice probabilities.

### 1.4.2 Quantities involved in welfare effects

We now explain how we compute the different components intervening in the variations of welfare for the full model. Appendix B.2 shows that expression (1.13) can be rewritten for our full model as:

$$\frac{\partial V}{\partial y_{lh}^{kfm_{t-1}}} \Bigg|_{\Omega_{lmh}^{kf}} = \ln \left[ 1 - P \left( l, o, h \mid f, m_{t-1}, I = y_{lht}^{kfm_{t-1}} \right) \right] P \left( f, m_{t-1}, y_{lht}^{kfr} \right) 1_{\{y_{lht}^{kfm_{t-1}} \in [\theta_{k-1,t}, \theta_{k,t}]\}} \quad (1.22)$$

where  $P(f, m_{t-1}, I) = P(f, m_{t-1} | I) f(I)$  is the proportion of individuals with family size  $f$ , previous tenure  $m_{t-1}$  and current income  $I$ , which can be decomposed multiplicatively

into the income density and the proportion of individuals with these family size and previous tenure conditionally on income,  $P(f, m_{t-1} | I)$ ; and  $P(l, o, h | f, m_{t-1}, I)$  is the probability for individuals with income  $I$ , family size  $f$  and previous tenure  $m_{t-1}$  of choosing a dwelling of type  $h$  in home-ownership in location  $l$ . Income thresholds can be computed under the assumption that there are downpayment and income constraints from expression (1.57) in Appendix (C), and parameters  $\theta_{k-1,t}$  and  $\theta_{k,t}$  are fixed by public authorities and are observed.

Expression (1.14) can be rewritten for our full model as:

$$E_{I',f',m'_{t-1}} \left[ \frac{\partial v_{loh}(I', f', m'_{t-1})}{\partial R_{lh}^{kf}} \frac{\partial V_t(I', f', m'_{t-1})}{\partial v_{loh}} \Bigg|_{\Theta_{loht}} \right] = \mu P_{lht}^{kfr} \quad (1.23)$$

where the expression of the aggregate utility  $V_t$  is given by equation (1.32) and that of the conditioning set on the left-hand side  $\Theta_{loht}$  is given by equation (1.38) in Appendix B.1, and:

$$P_{lht}^{kfm_{t-1}} = P\left(l, o, h, f, m_{t-1}, I \in \left[\max\left(\theta_{k-1,t}, y_{lht}^{kfm_{t-1}}\right), \theta_{k,t}\right]\right)$$

is the proportion of individuals with family size  $f$  and previous tenure  $m_{t-1}$ , who have an income in the given bracket and own a dwelling of type  $h$  in location  $l$  at current period. Finally, expression (1.15) can be rewritten for our full model respectively when owning or renting as:

$$E_{I',f',m'_{t-1}} \left[ \frac{\partial v_{loht}(I', f', m'_{t-1})}{\partial p_{lh}} \frac{\partial V_t(I', f', m'_{t-1})}{\partial v_{loht}} \Bigg|_{\Theta_{loht}} \right] = \delta \sum_{k,f',m'_{t-1}} P_{lht}^{k,f',m'_{t-1}} \quad (1.24)$$

$$E_{I',f',m'_{t-1}} \left[ \frac{\partial v_{lrht}(I', f', m'_{t-1})}{\partial p_{lh}} \frac{\partial V_t(I', f', m'_{t-1})}{\partial v_{lrht}} \Bigg|_{\Theta_{lrht}} \right] = \delta P(l, r, h) \quad (1.25)$$

where  $P(l, r, h)$  is the proportion of individuals who are previous renters and choose a dwelling of type  $h$  in location  $l$ . Note that the right-hand side of equation (1.24) involves the sum of terms which are similar to that intervening in equation (1.23) except that terms are

considers for all values of family size, previous tenure and income brackets.

## 1.5 Conclusion

In this paper, we provide a theoretical framework to analyze to what extent the no-interest loan affects location choices and welfare of households. We build a model where housing tenure, location choice, and borrowing constraints on the credit market are jointly modelled. We also propose an estimation strategy to take the model to the data. In an extended version of this chapter we will apply this empirical strategy in the case of France. Fortunately we have information on individuals income and housing choices: FILOCOM is an exhaustive administrative database on household housing stock which provide joint exhaustive information on households income, family size, tenure status (renters versus homeowners) and on type of dwellings (second hand versus new dwellings) and location. We also have information on housing prices as we use notaries databases.

Our model gives avenues for research. With the data and the estimated parameters it is possible to assess whether the policy generates sprawl. One might suspect the policy to drive the households with medium income accessing home-ownership to the city periphery where housing prices are lower. It would also be possible to propose counterfactuals in which eligibility conditions and the amount of the subsidy vary.

There are several aspects that have not been studied in this version of our work. We do not model how no-interest loans are financed by the government. We may consider in the model that households incur a uniform tax that is used to cover the costs of the subsidy. This makes it possible to conduct a cost-benefit analysis whereas we have restricted our attention so far to welfare effects. One can also wonder what it is the optimal policy and its consequences in terms of home-ownership rate and location choices of households.

In this version of our work, we have not modelled job location choices and commuting. We take into account commuting implicitly through a local component that affects the utility of households and in particular captures the effect of distance to the city center where a large share of jobs are located. It could be worth going further in the modelling of jobs to better characterize to what extent sprawl implies longer commutes. This seems important to relate sprawl to its environmental impact caused by gas emissions.



# Appendix

## A Production of consumption goods

We consider that income is generated from non-labour income  $I_0$  and wage  $w$  such that  $I = I_0 + w$ . We also assume that there is a distribution of skills such that for individual with skills  $s$  we have  $w = w(s)$  where  $w(\cdot)$  is an increasing function that maps skills into wages. Numeraire is produced in the city with a constant returns-to-scale technology  $G(K, N) = Ng(K/N)$  where  $K$  is capital input,  $N = \int_s \theta(s) f_S(s) ds$  is the aggregate effective labor (with the different skills being perfectly substitutable) where  $f_S(\cdot)$  denotes the density of skills, and  $\theta(s)$  is the productivity of an efficient unit of skill- $s$  labour.

We assume that capital is supplied at fixed rental rate  $\rho$  and that the output price and rental rate are determined on the international market. Without loss of generality, we fix output price to one. Firms maximize their profit, which yields that the marginal product is equal to the marginal cost for each factor of production (ie. capital and measures of workers with every skill level). For capital, we thus have:

$$g'(K_l/N_l) = \rho \tag{1.26}$$

and for every skill level  $s$ :

$$\theta(s) [g(K_l/N_l) - g'(K_l/N_l)K_l/N_l] = w(s) \tag{1.27}$$

The first-order condition for capital (1.26) can be rewritten as  $K_l/N_l = m(\rho)$  with  $m' \leq 0$  and first-order conditions on every skill level (1.27) then become:

$$w(s) = R(\rho)\theta(s) \tag{1.28}$$

with  $R(\rho) = g(h(\rho)) - h(\rho)\rho$ . Wages thus depend only on the exogenous city productivity levels associated to skills. There are two reasons for that. First, they do not depend on the local level of housing prices because land does not enter the production function. Second, they are not affected by labor supply and demand because the production function is constant returns to scales.

## B Extended model

In this appendix, we present the full model that takes into account previous tenure and differences in the rebate depending on family size, and whether the no-interest loan is for new or second-hand dwellings. We then establish results for the comparative statics of the welfare analysis.

### B.1 Extensions to new/old housing types and family size

We now consider an individual  $i$  with income  $I$ , family structure  $f$  and past tenure  $m_{t-1}$ . The utility of this individual when living in location  $l$  in a dwelling in market  $m$  of type  $h$  is given by:

$$u_{ilmh}(I, f, m_{t-1}) = v_{lmh}(I, f, R_{lh}(I, f) 1_{\{m_{t-1}=r\}}, p_{lh}) + \epsilon_{ilmhf} \quad (1.29)$$

Note that this utility depends on past tenure since it determines whether or not the individual can benefit from a no-interest loan. We restrict prices to depend only on location and dwelling type since, in practice, they are measured with price indices net of the effects of dwelling characteristics that are potentially correlated with family size. The credit constraint to access home-ownership is now given by:

$$y(p_{lh}, R_{lh}(I, f) 1_{\{m_{t-1}=r\}}) \leq I \quad (1.30)$$

In particular, individuals who are homeowners at date  $t - 1$  have a rebate that is fixed to zero. Finally, the cumulative of costs for land owners of type- $h$  dwellings is given by  $H_{lh}(n_{lh})$  with  $n_{lh}$  the number of type- $h$  dwellings in location  $l$ . The marginal land owner for dwellings of each type breaks even so that:

$$H_{lh}^{-1}(n_{lh}) = \pi_{lh} = rp_{lh} \quad (1.31)$$

We consider that housing markets clear so that  $n_{lh} = N_{lh}$ , which yields  $p_{lh} = H_{lh}^{-1}(N_{lh})/r$ .

The average utility of individuals with income  $I$ , family size  $f$  and past tenure  $m$  is given by:

$$V(I, f, m_{t-1}) = E_i \left[ \max_{(l,m,h) \in \mathcal{B}_{I,f,m_{t-1}}} u_{ilmh}(I, f, m_{t-1}) \mid I_i = I, f_i = f, m_{it-1} = m_{t-1} \right] \quad (1.32)$$

where:

$$\mathcal{B}_{I,f,m_{t-1}} = \{(l, r, h) \text{ for all } l, h\} \cup \{(l, o, h) \text{ for } l, h \text{ such that } y(p_{lh}, R_{lh}(I, f) 1_{\{m_{t-1}=r\}}) \leq I\} \quad (1.33)$$

Under the assumption that residuals  $\epsilon_{ilmhf}$  are identically and independently distributed, it becomes:

$$V(I, f, m_{t-1}) = \gamma + \ln \left( \sum_{(l,m,h) \in \mathcal{B}_{I,f,m_{t-1}}} \exp[v_{ilmh}(I, f, m_{t-1})] \right)$$

and the aggregate welfare of individuals is then:

$$V = \gamma + E_{I,f,m_{t-1}} \left[ \ln \left( \sum_{(l,m,h) \in \mathcal{B}_{I,f,m_{t-1}}} \exp[v_{ilmh}(I, f, m_{t-1})] \right) \right] \quad (1.34)$$

with  $v_{ilmh}(I, f, m_{t-1}) \equiv v_{ilmhf}(I, R_{lh}(I, f) 1_{\{m_{t-1}=r\}}, p_{lh})$ .



We denote

$$y_{loh}(I, f, m_{t-1}) \equiv y(p_{lh}, R_{lh}(I, f) 1_{\{m_{t-1}=r\}})$$

the income threshold for access to home-ownership in type- $h$  dwellings and we use the convention that the income threshold for access to the rental sector is  $y_{lrh}(I, f, m_{t-1}) = -\infty$  since that sector is accessible to all individuals. Equation (1.34) can be rewritten as:

$$V = \gamma + E_{I,f,m_{t-1}} \left[ \ln \left( \sum_{l,m,h} \exp[v_{lmh}(I, f, m_{t-1})] 1_{\{y_{lmh}(I, f, m_{t-1}) \leq I\}} \right) \right] \quad (1.35)$$

## B.2 Welfare comparative statics

We now derive results on comparative statistics for welfare when changing the amount of the no-interest loan rebate. We are going to make extensive use of the dominated convergence theorem that can be stated as follows:

**Lemma B.1** (Dominated convergence theorem). *Let  $X$  be an open subset of  $\mathbb{R}$  and  $\Omega$  be a measure space. Let  $x$  be a real belonging to the closure of  $X$ . Suppose  $f : X \times \Omega \rightarrow \mathbb{R}$  satisfies the following conditions:*

1.  $f(z, \mu)$  is a Lebesgue-integrable function of  $\mu$  for each  $z \in X$
2. For almost all  $\mu \in \Omega$ ,  $\lim_{z \rightarrow x} f(z, \mu) = \psi(\mu)$
3. There is an integrable function  $\theta : \Omega \rightarrow \mathbb{R}$  such that  $|f(z, \mu)| \leq \theta(\mu)$  for all  $z \in X$  and almost every  $\mu \in \Omega$ .

Then

$$\lim_{z \rightarrow x} \int_{\Omega} f(z, \mu) d\mu = \int_{\Omega} \psi(\mu) d\mu \quad (1.36)$$

The rebate in the general case is given by:  $R_{lh}(I, f) = \sum_{k=1}^K 1_{\{\theta_{k-1} \leq I < \theta_k\}} R_{lh}^{kf}$  with  $\theta_0 = 0$  and  $\theta_K = +\infty$  (and possibly  $R_{lh}^{kf} = 0$  for some  $k$ ). We consider the increase in the no-interest loan rebate at the margin for the single option  $(l, o, h)$  for individuals with income  $I \in [\theta_{k-1}, \theta_k[$  and family size  $f$ , ie.  $dR_{lh}^{kf} = \Delta > 0$ . We are going to decompose the change in welfare in

parts that can be more easily interpreted and computed. For that purpose, denote by  $\frac{\partial Y}{\partial X}\Big|_{\chi}$  the partial derivative of  $Y$  with respect to  $X$  holding constant the parameters in the set  $\chi$  and consider the following sets:

$$\begin{aligned}\Phi_{lh}^{kf} &= \left\{ R_{l'h'}^{k'f'} \text{ for } (l', h', k', f') \neq (l, h, k, f), p_{l'h'} \text{ for all } l', h' \right\} \\ \Psi_{lh} &= \left\{ p_{l'h'} \text{ for } (l', h') \neq (l, h), R_{l'h'}^{k'f'} \text{ for all } l', h', k', f' \right\}\end{aligned}$$

We have:

$$\frac{\partial V(I, f, m_{t-1})}{\partial R_{lh}^{kf}} = \frac{\partial V(I, f, m_{t-1})}{\partial R_{lh}^{kf}} \Big|_{\Phi_{lh}^{kf}} + \sum_{l', h'} \frac{\partial p_{l'h'}}{\partial R_{lh}^{kf}} \frac{\partial V(I, f, m_{t-1})}{\partial p_{l'h'}} \Big|_{\Psi_{l'h'}} \quad (1.37)$$

and thus:

$$\frac{\partial V}{\partial R_{lh}^{kf}} = E_{I', f', m'_{t-1}} \left[ \frac{\partial V(I', f', m'_{t-1})}{\partial R_{lh}^{kf}} \right] = \frac{\partial V}{\partial R_{lh}^{kf}} \Big|_{\Phi_{lh}^{kf}} + \sum_{l', h'} \frac{\partial p_{l'h'}}{\partial R_{lh}^{kf}} \frac{\partial V}{\partial p_{l'h'}} \Big|_{\Psi_{l'h'}}$$

This expression does not involve any derivative of incomes with respect to change in rebate because incomes are pinned down by the skill-specific productivity which is fixed (see Appendix A). We can decompose further this expression considering that utility can be expressed as a function of fixed components  $v_{lmh}$  and income thresholds  $y_{lh}^{kfm_{t-1}} = y(p_{lh}, R_{lh}^{kf} 1_{\{m_{t-1}=r\}})$  to access home-ownership instead of prices and rebates. Indeed, the changes of prices and no-interest loan rebates affect the fixed utility components of the housing local markets and the local income thresholds to access home-ownership. Introduce the sets:

$$\Theta_{lmh} = \left\{ v_{l'm'h'} \text{ for } (l', m', h') \neq (l, m, h), y_{l'h'}^{k'f'm'_{t-1}} \text{ for all } l', h', k', f', m'_{t-1} \right\} \quad (1.38)$$

$$\Omega_{lh}^{kfm_{t-1}} = \left\{ \left( y_{l'h'}^{k'f'm'_{t-1}} \right) \text{ for } (l', h', k', f', m'_{t-1}) \neq (l, h, k, f, m_{t-1}), v_{l'm'h'} \text{ for all } l', h', k', f', m'_{t-1}, m' \right\} \quad (1.39)$$

We then have:

$$\left. \frac{\partial V}{\partial R_{lh}^{kf}} \right|_{\Phi^{kf}} = \left. \frac{\partial y_{lh}^{kfr}}{\partial R_{lh}^{kf}} \frac{\partial V}{\partial y_{lh}^{kfr}} \right|_{\Omega^{kfr}} + E_{I',f',m'_{t-1}} \left[ \left. \frac{\partial v_{loh}(I',f',m'_{t-1})}{\partial R_{lh}^{kf}} \frac{\partial V(I',f',m'_{t-1})}{\partial v_{loh}} \right|_{\Theta_{loh}} \right] \quad (1.40)$$

$$\begin{aligned} \left. \frac{\partial V}{\partial p_{lh}} \right|_{\Psi_{lh}} &= \sum_{k,f',m'_{t-1}} \left. \frac{\partial y_{lh}^{kf'm'_{t-1}}}{\partial p_{lh}} \frac{\partial V}{\partial y_{lh}^{kf'm'_{t-1}}} \right|_{\Omega_{lh}^{kf'm'_{t-1}}} \\ &+ \sum_m E_{I',f',m'_{t-1}} \left[ \left. \frac{\partial v_{lmh}(I',f',m'_{t-1})}{\partial p_{lh}} \frac{\partial V(I',f',m'_{t-1})}{\partial v_{lmh}} \right|_{\Theta_{lmh}} \right] \end{aligned} \quad (1.41)$$

Note that in the right-hand terms of (1.40), there is no term when previous tenure is home-ownership. Indeed, for previous homeowners, there is no change in the income threshold and fixed utility component since they cannot benefit from the rebate. We are going to compute the right-hand side terms involving partial derivatives of the utility. We start with the computation of  $\left. \frac{\partial V}{\partial y_{lh}^{kfm_{t-1}}} \right|_{\Omega_{lh}^{kfm_{t-1}}}$  for any given  $l, h, k, f$  and  $m_{t-1}$ . Consider a small increase  $\Delta$  of the income threshold to access home-ownership for the income bracket  $[\theta_{k-1}, \theta_k[ : y_{lh}^{kfm_{t-1}} \rightarrow y_{lh}^{kfm_{t-1}} + \Delta$ . The sum of individuals' utilities when the income threshold is increased by  $\Delta$  is given by:

$$\tilde{V}(\Delta) = E_{I',f',m'_{t-1}} \left[ \tilde{V}(I',f',m'_{t-1},\Delta) \right]$$

where  $\tilde{V}(I',f',m'_{t-1},\Delta)$  is the utility of an individual with income  $I'$ , family size  $f'$  and past tenure  $m'_{t-1}$  when there is the change in the income threshold  $\Delta$ . It verifies:

$$\tilde{V}(I',f',m'_{t-1},\Delta) \quad (1.42)$$

$$= \gamma + \ln \left[ g(I',f',m'_{t-1}) + \exp[v_{loh}(I',f')] 1_{\{y_{loh}(I',f',m'_{t-1}) + \Delta 1_{\{I' \in [\theta_{k-1}, \theta_k[ \} \leq I'\}} \}} \right] \quad (1.43)$$

where:

$$g(I',f',m'_{t-1}) = \sum_{(l',m',h') \neq (l,o,h)} \exp[v_{l'm'h'}(I',f')] 1_{\{y_{l'm'h'}(I',f',m'_{t-1}) \leq I'\}} \quad (1.44)$$

where  $y_{lth}(I', f', m'_{t-1}) = -\infty$ . Indeed, quantities for  $(l', m', h') \neq (l, o, h)$  do not vary with the change in the income threshold  $\Delta$  when holding fixed  $\Phi_{lh}^{kf}$ . We then have:

$$\frac{\widetilde{V}(\Delta) - \widetilde{V}(0)}{\Delta} = E_{I', f', m'_{t-1}} \left[ \frac{\widetilde{V}(I', f', m'_{t-1}, \Delta) - \widetilde{V}(I', f', m'_{t-1}, 0)}{\Delta} \right] \quad (1.45)$$

Individuals with family size different from  $f$ , those with previous tenure different from  $m_{t-1}$ , and those with  $I' \notin [\theta_{k-1}, \theta_k[$  are not affected by the change in the income threshold and we have for them:  $\widetilde{V}(I', f', m'_{t-1}, \Delta) - \widetilde{V}(I', f', m'_{t-1}, 0) = 0$ . For individuals with family size  $f$  whose previous tenure is  $m_{t-1}$ , and whose income  $I' \in [\theta_{k-1}, \theta_k[$ , we have the following three cases:

- For  $I' < y_{loh}(I', f, m_{t-1})$ , option  $(l, o, h)$  is not available whether the income threshold takes the initial value or has been incremented by  $\Delta$ , and we have  $\widetilde{V}(I', f, m_{t-1}, \Delta) - \widetilde{V}(I', f, m_{t-1}, 0) = 0$ .
- For  $I' \geq y_{loh}(I', f, m_{t-1}) + \Delta$ , option  $(l, o, h)$  is available whether or not the income threshold is incremented by  $\Delta$ , and we have again  $\widetilde{V}(I', f, m_{t-1}, \Delta) - \widetilde{V}(I', f, m_{t-1}, 0) = 0$ .
- Finally, for  $I' \in [y_{loh}(I', f, m_{t-1}), y_{loh}(I', f, m_{t-1}) + \Delta[$ , option  $(l, o, h)$  is available only when the income threshold is at its initial value and we have:

$$\begin{aligned} \widetilde{V}(I', f, m_{t-1}, \Delta) - \widetilde{V}(I', f, m_{t-1}, 0) &= \ln [g(I', f, m_{t-1})] - \ln [g(I', f, m_{t-1}) + \exp [v_{loh}(I', f)]] \\ &\equiv -\widetilde{g}(I', f, m_{t-1}) \end{aligned}$$

that does not depend on  $\Delta$ .

Equation (1.45) can then be rewritten as:

$$\frac{\widetilde{V}(\Delta) - \widetilde{V}(0)}{\Delta} = -\frac{1}{\Delta} E_{I', f', m'_{t-1}} \left[ \widetilde{g}(I', f, m_{t-1}) \mathbf{1}_{\{I' | I' \in [y_{lh}^{kf m_{t-1}}, y_{lh}^{kf m_{t-1}} + \Delta] \cap [\theta_{k-1}, \theta_k[ \}} \mathbf{1}_{\{f' = f, m'_{t-1} = m_{t-1}\}} \right] \quad (1.46)$$

as for  $I'$  such that  $I' \in [\theta_{k-1}, \theta_k[$ , we have  $y_{loh}(I', f, m_{t-1}) = y_{lh}^{kf m_{t-1}}$ . We are going to compute

the limit of the right-hand side term as  $\Delta$  tends to zero. Having  $\Delta \rightarrow 0$ , we get:

$$\frac{\partial V}{\partial y_{lh}^{kfm_{t-1}}} \Big|_{\Omega_{lh}^{kfm_{t-1}}} = -\tilde{g}(y_{lh}^{kfm_{t-1}}, f, m_{t-1}) P(f, m_{t-1} | I = y_{lh}^{kfm_{t-1}}) f(y_{lh}^{kfm_{t-1}}) \mathbf{1}_{\{y_{lh}^{kfm_{t-1}} \in [\theta_{k-1}, \theta_k]\}} \quad (1.47)$$

where  $P(f, m_{t-1} | I)$  is the probability of having family since  $f$  and previous tenure  $m_{t-1}$  for individuals with income  $I$  and with:

$$\begin{aligned} \tilde{g}(I, f, m_{t-1}) &= \ln [g(I, f, m_{t-1}) + \exp[v_{loh}(I, f, m_{t-1})]] - \ln [g(I, f, m_{t-1})] \\ &= -\ln \left[ 1 - \frac{\exp[v_{loh}(I, f, m_{t-1})]}{g(I, f, m_{t-1})} \right] \\ &= -\ln \left[ 1 - \frac{\exp[v_{loh}(I, f, m_{t-1})]}{\sum_{(l', m', h') \neq (l, o, h)} \exp[v_{l'm'h'}(I, f)] \mathbf{1}_{\{y_{l'm'h'}(I, f, m_{t-1}) \leq I\}}} \right] \\ &= -\ln [1 - P(l, o, h | I, f, m_{t-1})] \end{aligned}$$

with  $P(l, o, h | I, f, m_{t-1})$  the probability of choosing location  $l$ , tenure  $o$  and housing type  $h$  for individuals characterized by  $(I, f, m_{t-1})$ . Hence:

$$\begin{aligned} &\frac{\partial V}{\partial y_{lh}^{kfm_{t-1}}} \Big|_{\Omega_{lh}^{kfm_{t-1}}} \\ &= \ln \left[ 1 - P(l, o, h | I = y_{lh}^{kfm_{t-1}}, f, m_{t-1}) \right] P(f, m_{t-1} | I = y_{lh}^{kfr}) f(y_{lh}^{kfm_{t-1}}) \mathbf{1}_{\{y_{lh}^{kfm_{t-1}} \in [\theta_{k-1}, \theta_k]\}} \end{aligned}$$

We then compute  $E_{I', f', m'_{t-1}} \left[ \frac{\partial v_{loh}(I', f', m'_{t-1})}{\partial R_{lh}^{kf}} \frac{\partial V(I', f', m'_{t-1})}{\partial v_{loh}} \Big|_{\Theta_{loh}} \right]$  for any given  $l, h$  and  $k$ . Individuals such that  $f' \neq f$ ,  $m'_{t-1} \neq r$  or  $I' \notin [\theta_{k-1}, \theta_k]$  are not affected by the change in rebate since, for them, we have  $\frac{\partial v_{loh}(I', f', m'_{t-1})}{\partial R_{lh}^{kf}} = 0$ . For the others, consider now a change of the fixed utility component in case of home-ownership  $v_{loh}$  by the amount  $\Delta$  and index utility by this amount as before. The average utility of individuals with income  $I'$ , family size  $f$  and

previous tenure  $r$  when there is such an increase verifies:

$$V(I', f, r, \Delta) = \gamma + \ln [g(I', f, r) + \exp [v_{loh}(I', f, r) + \Delta]] 1_{\{y_{loh}(I', f, r) \leq I'\}}$$

We are going to compute the limit when  $\Delta$  tends to zero of

$$E_{I', f', m'_{t-1}} \left[ \frac{\partial v_{loh}(I', f, r)}{\partial R_{lh}^{kf}} \frac{V(I', f, r, \Delta) - V(I', f, r, 0)}{\Delta} 1_{\{f'=f, m'_{t-1}=r, I' \in [\theta_{k-1}, \theta_k]\}} \right]$$

holding  $\Theta_{loh}$  fixed. Individuals with income  $I'$  such that  $y_{loh}(I', f, r) > I'$  are not affected by the change in the income threshold and we have  $V(I', f, r, \Delta) - V(I', f, r, 0) = 0$ . For individuals with income such that  $y_{loh}(I', f, r) \leq I'$ , we have:

$$\begin{aligned} & V(I', f, r, \Delta) - V(I', f, r, 0) \\ &= \ln [g(I', f, r) + \exp [v_{loh}(I', f, r) + \Delta]] - \ln [g(I', f, r) + \exp [v_{loh}(I', f, r)]] \\ &\equiv z(I', f, r, \Delta) \end{aligned}$$

This yields:

$$\begin{aligned} & E_{I', f', m'_{t-1}} \left[ \frac{\partial v_{loh}(I', f, m_{t-1})}{\partial R_{lh}^{kf}} \frac{V(I', f, m_{t-1}, \Delta) - V(I', f, m_{t-1}, 0)}{\Delta} 1_{\{f'=f, m'_{t-1}=r, I' \in [\theta_{k-1}, \theta_k]\}} \right] \\ &= E_{I', f', m'_{t-1}} \left[ 1_{\{I' | I' \in [\theta_{k-1}, \theta_k]\} \cap \{I' | y_{loh}(I, f, r) \leq I'\}} 1_{\{f'=f, m'_{t-1}=r\}} \frac{z(I', f, r, \Delta)}{\Delta} \frac{\partial v_{loh}(I', f, r)}{\partial R_{lh}^{kf}} \right] \\ &= E_{I', f', m'_{t-1}} \left[ 1_{\{I' | I' \in [\max(\theta_{k-1}, y_{lh}^{kfr}), \theta_k]\}} 1_{\{f'=f, m'_{t-1}=r\}} \frac{z(I', f, r, \Delta)}{\Delta} \frac{\partial v_{loh}(I', f, r)}{\partial R_{lh}^{kf}} \right] \end{aligned}$$

It is possible to establish the limit of the right-hand side expectation using the dominated

convergence theorem. We have:

$$\begin{aligned} \frac{z(I', f, r, \Delta)}{\Delta} &\rightarrow \frac{\partial}{\partial v_{loh}} \ln [g(I', f, r) + \exp [v_{loh}(I', f, r)]] \\ &= \frac{\exp [v_{loh}(I', f, r)]}{g(I', f, r) + \exp [v_{loh}(I', f, r)]} \\ &= P(l, o, h | I', f, r) \end{aligned}$$

where  $P(l, o, h | I', f, r)$  is the probability of choosing location  $l$ , tenure  $o$  and housing type  $h$  for individuals with income  $I'$ , family size  $f$  and previous tenure  $r$ . The function  $\Delta \mapsto \frac{z(I', f, m_{t-1}, \Delta)}{\Delta}$  is continuous in  $\Delta$  on an interval  $(0, \bar{\Delta}]$  for  $\bar{\Delta}$  small enough and has a limit when  $\Delta \rightarrow 0$ . It is thus bounded on the interval  $(0, \bar{\Delta}]$ . We can then apply the dominated convergence theorem and we get:

$$\begin{aligned} E_{I', f', m'_{t-1}} \left[ \frac{\partial v_{loh}(I', f', m'_{t-1})}{\partial R_{lh}^{kf}} \frac{\partial V(I', f', m'_{t-1})}{\partial v_{loh}} \Big|_{\Theta_{loh}} \right] & \quad (1.48) \\ = E_{I', f', m'_{t-1}} \left[ 1_{\{I' | I' \in [\max(\theta_{k-1}, y_{lh}^{kfr}), \theta_k]\}} 1_{\{f' = f, m'_{t-1} = r\}} P(l, o, h | I', f, r) \frac{\partial v_{loh}(I', f, r)}{\partial R_{lh}^{kf}} \right] \end{aligned}$$

Note in particular that if  $v_{loh}$  is linear in  $R_{lh}(I, f) 1_{\{m_{t-1}=r\}}$  such that  $\frac{\partial v_{loh}(I, f, r)}{\partial R_{lh}^{kf}} = \mu$  is a constant, we have:

$$\begin{aligned} E_{I', f', m'_{t-1}} \left[ \frac{\partial v_{loh}(I', f', m'_{t-1})}{\partial R_{lh}^{kf}} \frac{\partial V(I', f', m'_{t-1})}{\partial v_{loh}} \Big|_{\Theta_{loh}} \right] & \\ = \mu E_{I', f', m'_{t-1}} \left[ 1_{\{I' | I' \in [\max(\theta_{k-1}, y_{lh}^{kfr}), \theta_k]\}} 1_{\{f' = f, m'_{t-1} = r\}} P(l, o, h | I', f, r) \right] & \\ = \mu P(l, o, h, f, r, I \in [\max(\theta_{k-1}, y_{lh}^{kfr}), \theta_k]) & \end{aligned}$$

where  $P(l, o, h, f, r, I \in [\max(\theta_{k-1}, y_{lh}^{kfr}), \theta_k])$  is the probability of individuals having family size  $f$ , previous tenure  $r$ , income being in the bracket  $I \in [\max(\theta_{k-1}, y_{lh}^{kfr}), \theta_k]$  and choosing location  $l$ , tenure  $o$  and housing type  $h$ .

Finally, we compute  $E_{I', f', m'_{t-1}} \left[ \frac{\partial v_{lmh}(I', f', m'_{t-1})}{\partial p_{lh}} \frac{\partial V(I', f', m'_{t-1})}{\partial v_{lmh}} \Big|_{\Theta_{lmh}} \right]$  for any given  $l, m$  and  $h$ . In fact, the proof is similar to the previous one. There is one difference though: All individuals

are affected by the change of prices in  $l$ , ie. we have generically  $\frac{\partial v_{lmh}(I', f', m'_{t-1})}{\partial p_{lh}} \neq 0$  for all  $I'$ ,  $f'$  and  $m'_{t-1}$ . In particular, this translates into the fact that in all expressions, we do not restrict  $I'$  to the set  $\{I' | I' \in [\theta_{k-1}, \theta_k[ ]\}$ . By analogy with equation (1.48), we get:

$$\begin{aligned} & E_{I', f', m'_{t-1}} \left[ \left. \frac{\partial v_{lmh}(I', f', m'_{t-1})}{\partial p_{lh}} \frac{\partial V(I', f', m'_{t-1})}{\partial v_{lmh}} \right|_{\Theta_{lmh}} \right] \\ &= E_{I', f', m'_{t-1}} \left[ \frac{\partial v_{lmh}(I', f', m'_{t-1})}{\partial p_{lh}} P(l, m, h | I', f', m'_{t-1}) \mathbf{1}_{\{I' | y_{lmh}(I', f', m'_{t-1}) \leq I'\}} \right] \end{aligned}$$

and if  $v_{lmh}(I', f', m'_{t-1})$  is linear in  $p_l$  such that  $\frac{\partial v_{lmh}(I', f', m'_{t-1})}{\partial p_{lh}} = \delta$  is a constant, we get by analogy with equation (1.48):

$$E_{I', f', m'_{t-1}} \left[ \left. \frac{\partial v_{lmh}(I', f', m'_{t-1})}{\partial p_{lh}} \frac{\partial V(I', f', m'_{t-1})}{\partial v_{lmh}} \right|_{\Theta_{lmh}} \right] = \delta P(l, m, h, \{I' | y_{lmh}(I', f', m'_{t-1}) \leq I'\}) \quad (1.49)$$

Note that in the case of rental market ( $m = r$ ), we have  $y_{lmh}(I', f', m'_{t-1}) = -\infty$ , and the right-hand side simplifies further into  $\delta P(l, r, h)$ . This result is similar to that of Busso, Gregory and Kline (2013) when deriving the expected utility with respect to a local wage. In the case of the home-ownership market ( $m = o$ ), the expression is made more complicated by the fact that a change in  $v_{loh}$  only affects individuals that have access to home-ownership in location  $l$ . We rewrite expression (1.49) if  $m = o$  such that it will be easier to compute from the data. We have:

$$\begin{aligned} & E_{I', f', m'_{t-1}} \left[ \left. \frac{\partial v_{lmh}(I', f', m'_{t-1})}{\partial p_{lh}} \frac{\partial V(I', f', m'_{t-1})}{\partial v_{loh}} \right|_{\Theta_{loh}} \right] \\ &= \delta \sum_{k, f', m'_{t-1}} P\left(l, o, h, f', m'_{t-1}, I' \in \left[ \max\left(\theta_{k-1}, y_{lh}^{k f' m'_{t-1}}\right), \theta_k \right] \right) \end{aligned}$$

We also compute the variation of aggregate profit of land owners due to the change in the



no-interest loan rebate. The aggregate profit is given by:

$$\Pi = \sum_{l,h} \pi_{lh} N_{lh} - \int_0^{N_{lh}} H_{lh}^{-1}(x) dx$$

Variation in this aggregate profit is given by:

$$\frac{\partial \Pi}{\partial R_{lh}^{kf}} = \sum_{l',h'} \frac{\partial \pi_{l'h'}}{\partial R_{lh}^{kf}} N_{l'h'} + \sum_{l',h'} [\pi_{l'h'} - H_{l'h'}^{-1}(N_{l'h'})] \frac{\partial N_{l'h'}}{\partial R_{lh}^{kf}} \quad (1.50)$$

Using the zero-profit condition, the second right-hand side term is zero and we finally get:

$$\frac{\partial \Pi}{\partial R_{lh}^{kf}} = \sum_{l',h'} \frac{\partial \pi_{l'h'}}{\partial R_{lh}^{kf}} N_{l'h'} = r \sum_{l',h'} \frac{\partial p_{l'h'}}{\partial R_{lh}^{kf}} N_{l'h'} \quad (1.51)$$

## C Borrowing constraints

In this appendix, we provide a  $T$ -period framework to expose the rationale for the borrowing constraint introduced in our model (even if it is not completely consistent with our model).

Individuals who want to purchase a home at price  $p$  face two constraints imposed by lenders on the credit market. There is a downpayment constraint such that downpayment  $D$  must be at least a fraction  $a$  of the dwelling value:  $D \geq ap$ . Denoting by  $W$  the wealth of the individual and considering that all of it can be used to purchase the dwelling, we have  $D = W + R$  as the rebate due to the state-subsidized zero-interest loan is considered to be downpayment. The downpayment constraint can be rewritten as:

$$ap \leq W + R \quad (1.52)$$

There is also an income constraint such that the annual repayment  $A$  must be less than a given fraction  $e$  of income  $I$ :  $A/I \leq e$ . Consider that the loan has a fixed rate  $r$  and duration

$T$ . Denoting by  $M$  the value of the loan, we have:

$$M = \sum_{t=1}^T \frac{A}{(1+r)^t} = A(1+r)^{-T} \sum_{t=1}^T (1+r)^{T-t} \quad (1.53)$$

$$= A(1+r)^{-T} \sum_{t=0}^{T-1} (1+r)^t = A(1+r)^{-T} \frac{(1+r)^T - 1}{(1+r) - 1} \quad (1.54)$$

or equivalently  $A = \tilde{r}M$  with  $\tilde{r} = r(1+r)^T / [(1+r)^T - 1]$ . Considering also that we have the relationship  $p = W + R + M$ , the income constraint can be rewritten as:

$$p - \frac{e}{\tilde{r}}I \leq W + R \quad (1.55)$$

Denote the wealth-to-income ratio for a given household as  $\beta = \frac{W}{I}$ . Equations (1.52) and (1.55) can be rewritten for an individual with realization  $\beta$  as:

$$\begin{aligned} ap &\leq \beta I + R \\ p &\leq \left( \beta + \frac{e}{\tilde{r}} \right) I + R \end{aligned}$$

or equivalently:

$$\begin{aligned} \beta^{-1}(ap - R) &\leq I \\ \left( \beta + \frac{e}{\tilde{r}} \right)^{-1}(p - R) &\leq I \end{aligned}$$

or equivalently:

$$\max \left[ \beta^{-1}(ap - R), \left( \beta + \frac{e}{\tilde{r}} \right)^{-1}(p - R) \right] \leq I \quad (1.56)$$

and it has the form given for the borrowing constraints in the text (see equation 1.1). From this expression, we deduce that income threshold to access home-ownership in a dwelling of type  $h$  in location  $l$  for an individual with family size  $f$ , income in bracket  $[\theta_{k-1}, \theta_k[$ , previous

tenure  $m_{t-1}$  and realization of the wealth-to-income ratio  $\beta$  is given by:

$$y_{lh}^{kfm_{t-1}}(R_{lh}^{kf}, \beta) = \max \left[ \beta^{-1} (ap_{lht} - R_{lh}^{kf} 1_{\{m_{t-1}=r\}}), \left( \beta + \frac{e}{\tilde{r}} \right)^{-1} (p_{lht} - R_{lh}^{kf} 1_{\{m_{t-1}=r\}}) \right] \quad (1.57)$$

It can be computed provided that values for  $a$ ,  $e$ ,  $r$  and  $\beta$  are available.

## D NIL subsidy

In addition to its zero-percent interest rate, the no-interest loan usually include an initial grace period, allowing households to defer loan repayment or reduce the amount paid for several years. The loan is equivalent to a subsidy for households equal to the difference between the loan amount and the net present value of all loan repayments. The following formula calculates the subsidy that the loan represents for an eligible household :

$$R = M - \sum_{t=1}^G \frac{xM}{G} (1+r)^{-t} - \sum_{t=G+1}^L \frac{(1-x)M}{L-G} (1+r)^{-t}$$

where  $R$  is the subsidy,  $M$  the loan amount,  $L$  the loan period  $G$  the grace period, during which households only have to reimburse a share  $x$  of the loan, the remaining  $1-x$  being paid later (between  $G$  and  $L$ ) and  $r$  is the discount rate. The subsidy corresponds to the difference between the loan amount (first term of the formula) and the net present value of all loan repayments, which can be decomposed into repayments during the grace period (second term) and remaining repayments (third term). The values for  $L$ ,  $G$  and  $x$  vary with income, family size, location and property type, according specified policy parameters. The loan profile (overall term, grace period) is also a function of the borrower's income: the lower the income, the longer the term and the grace period. For our analysis the discount rate is set to be equal to the interest rate on standard bank mortgage loans to households. It is different from the discount rate used for reimbursing banks <sup>10</sup>

---

10. The yearly standard rate on standard bank mortgage loans is calculated using data on mortgage loans from the French Housing Survey.



## Chapter 2

# Intraregional inequality in France in the last fifty years <sup>1</sup>

### Abstract

This paper combines data and methods with a view to analyse the evolution of intraregional inequality in France over half a century. Most notably, using numerous unexplored archives of the French Ministry of Finance, we reconstruct the income distribution of each French *département* since 1966. We find that the total inequality within tax units stems mostly from inequalities within *départements*. A big part of the total inequality comes from inequalities within *départements* where big cities are located, in particular in the Paris region. This arises mostly from the upper end of the French national income distribution as the richest tax units are getting increasingly over-represented in *départements* where big cities are located. Three *départements* of the Paris region, Paris itself, Hauts de Seine et Yvelines accounts for one third of French top 1% tax units – based on income distribution –. Seine-Saint-Denis, another *département* of the Paris region, has an opposite trajectory than the rest of its region, as there are increasingly less rich tax units living there. Also, the French top 1% tax units

---

1. This work is a joint work with Florian Bonnet. We are particularly indebted to Thomas Piketty for his constant and excellent guidance. We are also grateful to Antoine Bozio, Vincent Bignon, Anne Duquerroy, Malka Guillot, Bertrand Garbinti, Jonathan Goupille-Lebret, Miren Lafourcade, for helpful comments and discussions and to numerous seminar participants. All remaining errors are our own.

are getting increasingly over-represented in Haute Savoie, a *département* in the French Alps bordering Switzerland, with no big city, a pattern likely caused by cross-borders workers.

**Keywords:** Intraregional Inequalities, Income distribution, Economic geography, Economic history

# 1 Introduction

The late political outcomes and social movements in developed economies have raised awareness on the question of spatial inequalities. In particular, the *Yellow vests* movement, that started in November 2018 in France, was partly the expression of a rising anger vis-à-vis unequal access to local public services and jobs. It was especially buoyant in rural and periurbans areas. Fifteen years ago, the 2005 French riots shed light on the increasing feeling of disconnection of Paris and other big French cities suburbs inhabitants. Each French election enlightens us on the increasing spatial polarization of the political landscape. This leads to a growing interest for the study of spatial inequalities, perceived both as a symptom and a vector of inequality. A symptom because they are the spatial materialization of inequalities between individuals. A vector because they may endogenously produce even more inequalities. A question of particular interest in this context is whether power and money are evenly spread over the country. Executives and top executives might be used as a proxy to answer this question. For instance, [Godechot \(2013\)](#) analyses the spatial segregation of top executives in the case of France. He uses individual data on wages to show that firms top earners are increasingly segregated in terms of workplace. They are increasingly working in Paris and in the business district of La Défense. However, if the perception of spatial inequality may depend on the spatial segregation of top executives, it may also depend more generally on the spatial distribution of all individuals and not only workers. In particular, the location of retirees and of capital owners may also matter. Using income data will allow us to study the location of the richest individuals, may they be top wage earners, top capital owners or rich retirees.

In this paper we document the evolution of intra-regional inequalities in France since 1966. Our first contribution is to reconstruct the income distribution, for the periods 1966-1969, 1986-1998, 2000-2017, of each French *département*. We use numerous unexplored archives of the French Ministry of Finance that we have painstakingly digitized. For the 1966-1969 period we propose a novel methodology to reconstruct the total number and income of non-taxable

tax units i.e. not filling a tax return. For the remaining part of the estimation we draw on [Garbinti et al. \(2018\)](#) methodology and on [Blanchet et al. \(2017\)](#) algorithm to interpolate income distribution from fiscal tabulations. Our main concept is fiscal income. That is the income reported on fiscal declarations, before any further deduction. It does not take into account taxes and transfers except the operation of the pension and unemployment systems.<sup>2</sup> Our income concept does not disentangle labor and capital income, nor wages and pensions. The present paper belongs to a broad international project aimed at improving inequality measurement, namely the WID.world project.

Our second contribution is to uncover novel empirical regularities regarding the intraregional evolution of inequalities in France in the last fifty years. First, we find that the total inequality within tax units almost exclusively comes from the inequality within *départements*. The inequality between *département* only accounts for 3.8% of the total inequality and has strongly decreased over time. In that sense, there has been some spatial convergence in terms of income. However, we show that this convergence only comes from the poorest tax units and is concealing a spatial divergence of the richest tax units. We show that while the average income of the bottom of the income distribution of each *département* has converged, the average income of the top of the income distribution of each *département* has diverged. This is because the richest tax units of the *départements* where big cities are located, are getting increasingly rich. This is also because an increasing share of the richest tax units live in these *départements*. In the end, these *départements* increasingly contribute to the total inequality. Importantly, Paris plays a major role, along with two other *départements* of the Paris region, Hauts de Seine, where La Défense business district is located, and Yvelines. These three *départements* are responsible for 20% of the total inequality while their tax units only represent 8.4% of the total number of tax units in 2017. In other words the contribution of Paris, Hauts de Seine and Yvelines to the total inequality comes more from the inequality within them than from the difference between their average income and the average income of the other *départements*. One third of the top 1% French metropolitan tax units live in

---

2. To avoid double counting, wages are net of pension contributions (as well as other social contributions, as defined by SNA 2008 national accounts guidelines).



these three *départements* in 2017. As a result, the top 1% French metropolitan tax units are over-represented there. In 1969, they accounted for 2.8% of the number of tax units of these three *départements*, today they account for 3.8%. Seine-Saint-Denis, another *département* of the Paris region, moves in the opposite direction, as less and less rich tax units are living there. The French top 1% tax units are also getting increasingly over-represented in Haute Savoie, a *département* in the French Alp, bordering Switzerland, in which there is no big city. It is likely due to an increasing share of cross-borders workers.

Our work relates to three streams of the literature. It relates to the literature studying national distribution of income starting with the seminal work by [Piketty \(2001\)](#). Numerous scholars have followed such as [Garbinti et al. \(2018\)](#) for France, [Atkinson \(2005\)](#) for UK, [Atkinson \(2005\)](#), [Roine et al. \(2009\)](#) [Alvaredo and Saez \(2009\)](#). Our paper provides sub-national insights on this literature.

Our paper is also related to the literature studying regional convergence and divergence. [Williamson \(1965\)](#) combines regional income data of several countries showing spatial inequality have followed a inverted U shape curve up to the middle of the 20<sup>th</sup> century. Broadly, economic growth first appears in a given region, increasing regional inequality in the first stages. Then factors mobility reduces or eliminates regional differences. The new economic geography literature, initiated by [Krugman \(1991\)](#), grounds this. However, by contrast with [Williamson \(1965\)](#), the empirical literature that has followed has used regional valued added data, constructed following [Geary and Stark \(2002\)](#) methodology. Scholars interested in the evolution of spatial inequalities in France developed another methodology than [Geary and Stark \(2002\)](#) for reconstructing regional valued added in the long run. They did so only for a limited number of years. [Combes et al. \(2011\)](#) use [Toutain \(1993\)](#) data for 1860, 1890 and 1930 and INSEE data for 1982 and 2000; [Bazot \(2014\)](#) reconstructs regional value added data every two years from 1840 to 1911; [Caruana-Galizia \(2013\)](#) develops an econometric model based on sectoral decomposition of *départements*; [Sanchis et al. \(2015\)](#) add to [Combes et al. \(2011\)](#) data INSEE data for the period 2000-2014. They acknowledge a great divergence in value added per capita per *département* since 2000. Conversely, we document a convergence

of income per capita.

Eventually, our paper is related to the urban economic literature. This literature has documented the growing polarization of labor markets in the US, a phenomenon coined as "the great divergence" by [Moretti \(2012\)](#) – see also [Hsieh and Moretti \(2015\)](#) and [Diamond \(2016\)](#) – as college graduates increasingly sorted into high wages, high rents cities. As a result [Giannone \(2017\)](#) shows that between 1940 and 1980 the average wage in US cities converged and that after 1980 a reverse phenomenon took place. Our work also relates to [Chen and Rosenthal \(2008\)](#) that study life-cycle migration and amenities. They show that US cities with improving business environments acquire increasing shares of workers and that cities with improving consumer amenities become relatively more populated by retirees.

The rest of the paper proceeds as follows. We present our data, methods and concepts in Section 2. We present our findings in Section 3. Section 4 concludes.

## 2 Data, Concepts and Methods

The first contribution of this paper is to reconstruct the income distribution of each French *département*.

### 2.1 Data Sources

**Sample** We restrict on French metropolitan *départements*. Corsica is treated as a sole *département* throughout our period of study. Before 1966, Paris region was split differently in terms of *département*. They were three *départements* only, Seine, Seine et Oise, et Seine et Marne. After 1966, Seine was split into four *départements*: *Paris*, *Hauts-de-Seine*, *Seine-Saint-Denis*, *Val de Marne*, et *Seine et Oise* was split into three *départements* *Essone*, *Val d’Oise* and *Yvelines*. We restrict our sample from 1966 onwards and use the more recent classification.

**Sources** Our main data source comes from more than 4000 newly digitized fiscal tabulations

that we have uncovered in the archives of the French Ministry of Finance. The French administration has published the national sum of these tabulations. [Piketty \(2001\)](#) has first used these published national tabulations to reconstruct the French income distribution. We have had access and digitized the income tabulations made by the French administration *at the département* level. These fiscal tabulations report the number of tax units and the total income, by income brackets, for each French *département*. We also use national income distribution produced by [Garbinti et al. \(2018\)](#) and populations by age per *département* computed by [Bonnet \(2018b\)](#).

**Time Period** We have discovered and digitized fiscal tabulations data for the periods 1960-1969 and 1986 onwards. For the periods before 1960 and between 1970 and 1985 no fiscal tabulations have been found in the Archives.

**Types of Income Series** There are several ways to represent income inequality depending on the income concept and population units used. Following [Garbinti et al. \(2018\)](#) we provide three series of income: (i) fiscal income without capital gains, (ii) fiscal income with capital gains and (iii) pretax income. Fiscal tabulations report taxable income. Taxable income is subject to a number of specific deductions instituted by the tax law. We are interested in fiscal income that is income reported on fiscal declarations, before any further deduction. We also produce, for future researches, pretax income that is the sum of all income flows going to labor and capital, after taking into account the operation of the pension system, but before taking into account other taxes and transfers. Pretax income hence includes pensions and excludes contributions. Interestingly, the sum of all pretax income at the national level is equal to the national income. Both fiscal income and pretax income excludes social benefits (except unemployment benefits). By contrast with [Garbinti et al. \(2018\)](#) we only produce series by tax units<sup>3</sup>. Each series provide complementary information on income inequality. In the analysis part of this article we focus on fiscal income excluding capital gains. Results are

---

3. [Garbinti et al. \(2018\)](#) produce income distribution for several population units (tax units, equal-split adults, individualistic adults) at the French national level. Because of data constraint we only provide income by tax units.

unchanged when using the two other concepts of income.

## 2.2 Methodology

There is two different periods in terms of methodology. We present in full length our methodology in Appendix A and B. Here we give a summarized explanation of what we do.

**Methodology for the 1986-2017 Period** From 1986 onwards we have fiscal tabulations of both taxable and non taxable tax units. We interpolate income distributions from income tabulations using [Blanchet et al. \(2017\)](#) methodology. We transform taxable income distribution into pretax and fiscal income distributions using enhancement rate taken from [Garbinti et al. \(2018\)](#)<sup>4</sup>. We make sure our estimates are consistent with theirs at the national level.

**Methodology for the 1966-1969 Period** For the 1966-1969 period we only have fiscal tabulations of the taxable tax units. Hence we need to reconstruct the number of non taxable tax units and the income distribution of the non taxable tax units of each *département*. To recover the total number of tax units by *département* we use an econometric method. We run the following regression for years 1986–2010:

$$N_t^d = \alpha_0 + \sum_{a=1}^6 \alpha_a P_t^{d,a} + \beta M_t^d + \gamma N_n + \delta_d + \varepsilon_t^d \quad (2.1)$$

$N_d$  is the total number of tax units of *département*  $d$ ,  $N_n$  the total number of tax units at the national level,  $M^d$  the number of married couples in *département*  $d$ ,  $P^{d,a}$  the number of individuals in the age group  $a$  in *département*  $d$  – there are six age groups  $a$  –.  $\delta_d$  is a *département* fixed effect.  $\varepsilon_d$  is an error term. We do not use the years 2011 to 2014 to fit the model because the definition of tax unit changed after 2011. Since 2011, married and divorced couples no longer complete three returns the year of their union or divorce. Overall,

---

4. These enhancement rates map the taxable income percentiles to fiscal income percentiles – and also to fiscal income with capital gains and pretax income percentiles –.

the  $R^2$  is equal to 0.9996. To test the consistency of our estimates, we sum up the estimated departmental tax units and compared them to the national tax units estimated by [Garbinti et al. \(2018\)](#). National estimates slightly differ from our own estimates because of overseas *départements*. If we assume that the share of overseas *départements* in the total number of tax units is 1%, then we overestimate the total number of metropolitan tax units by 1.8% points in 1966 and 1.25% point in 1969. We homogeneously adjust our tax number estimates to correct this discrepancy. We also test the fit of our model by comparing predicted values and observed values for the years 1986 to 2010. Our method has a good fit on average for the period 1995 and 2010 even though the difference can be up to more than 5% in a few cases. For the previous period, values are more overestimated, especially in 1986 when non taxable tax units were required to file a tax return for the first time.

We also need to estimate the income distribution of the non taxable tax units. According to [Piketty \(2001\)](#) the share of taxable tax units among all tax units was around 44% in 1966 and increased gradually up to 51% in 1969. Taxable tax units are typically richer than non taxable tax units.<sup>5</sup> We make strong assumptions to impute the income distribution of the non taxable tax units. We test two specifications with radically different assumptions. We provide evidence that the magnitude of our results are unchanged with these two specifications. We believe that we provide reliable estimations of the location and income distribution of French top 40 % tax units. We do not use the data on the bottom 60 tax units for the period 1966-1969.

We define  $Y^d$  as the average income of all tax units of *département*  $d$ ,  $Y_1^d$  the average income of the taxable tax units in *département*  $d$ ,  $Y_0^d$  the average income of the non taxable tax units in *département*  $d$ ,  $N_1^d$  the number of taxable tax units,  $N_0^d$  the number of non taxable tax units,  $L_t^d(x_d)$  the income share of the  $x_d$  poorest tax units in year  $t$  and *département*  $d$ .  $L_t^d$  is the Lorenz curve of the income distribution of *département*  $d$  in year  $t$ . We alternatively

---

5. This it is not always the case. Some rich and large families may be non taxable because the French tax income schedule depends of the size of the family. These situations are rare however. [Piketty \(2001\)](#) discusses this matter extensively.

define  $\tilde{Y}_1^d$  the average income of the taxable tax units in *département*  $d$  that belong to the top 40% French tax units,  $\tilde{Y}_0^d$  the average income of remaining tax units in *département*  $d$ ,  $\tilde{N}_1^d$  the number of taxable tax units in *département*  $d$  that belongs to the top 40% French tax units,  $\tilde{N}_0^d$  the number of the remaining tax units. We define  $L_t^N(x)$  the income share of the  $x$  poorest tax units in the *national* income distribution. We get the Lorenz curve  $L_t^N$  from national estimates by [Garbinti et al. \(2018\)](#).

In the first specification we make two assumptions. First we assume that the non taxable units are the tax units at the bottom of the income distribution of their *département*. We further assume that the income share of the non taxable tax units during the 1960-1969 period is equal to the income share of the  $x_d$  poorest tax units in 1986 where  $x_d$  is equal to the share of non taxable tax units in 1960-1969. In mathematical terms it means:  $L_t^d\left(\frac{N_{0,t}^d}{N_{0,t}^d+N_{1,t}^d}\right) = L_{1986}^d\left(\frac{N_{0,t}^d}{N_{0,t}^d+N_{1,t}^d}\right)$  for any  $t \in \{1960 - 1969\}$  and  $d$ . We deduce the average income of tax units living in *département*  $d$  with the following formula:

$$Y^d = \frac{Y_1^d}{1 - \tilde{L}\left(\frac{N_0^d}{N_0^d+N_1^d}\right)} \frac{N_1^d}{N_1^d + N_0^d}$$

In the second specification we impute not only the income distribution of the non taxable tax units but also the one of the taxable tax units that are below the P60 national fractile. We choose the P60 fractile as we know that at least 40% of all tax units are taxable since 1966. Here we assume that the tax units that are either non taxable or don't belong to the top 40% French tax units are at the bottom of the national income distribution. In mathematical terms we assume that their income share in the national income is equal to  $L_t^N\left(\frac{\sum_{d'} \tilde{N}_0^{d'}}{\sum_{d'} \tilde{N}_0^{d'} + \tilde{N}_1^{d'}}\right)$ . We further assume that this income is distributed over *départements* according to their share of below P60 non taxable tax units. In mathematical terms it means that the average income

of the below P60 non taxable tax units in any *département*  $d$  is:

$$\frac{L_t^N \left( \frac{\sum_{d'} \tilde{N}_0^{d'}}{\sum_{d'} \tilde{N}_0^{d'} + \tilde{N}_1^{d'}} \right) \sum_{d'} \tilde{Y}_1^{d'} N_1^{d'}}{1 - L_t^N \left( \frac{\sum_{d'} \tilde{N}_0^{d'}}{\sum_{d'} \tilde{N}_0^{d'} + \tilde{N}_1^{d'}} \right)} \frac{1}{\sum_{d'} \tilde{N}_0^{d'}}$$

with  $\frac{L_t^N \left( \frac{\sum_{d'} \tilde{N}_0^{d'}}{\sum_{d'} \tilde{N}_0^{d'} + \tilde{N}_1^{d'}} \right) \sum_{d'} \tilde{Y}_1^{d'} N_1^{d'}}{1 - L_t^N \left( \frac{\sum_{d'} \tilde{N}_0^{d'}}{\sum_{d'} \tilde{N}_0^{d'} + \tilde{N}_1^{d'}} \right)}$  the total income of the below P60 non taxable tax units and  $\frac{1}{\sum_{d'} \tilde{N}_0^{d'}}$  the total number of below P60 non taxable tax units. Then we can deduce the average income in *département*  $d$ .

Once imputed the average income of each *département* we are back to the 1986-2017 methodology. In a nutshell, in the first specification we use 1986 information and assume some stability in the income share of the bottom of the distribution. In the second specification we use information on the national income distribution at the year of estimation. We assume that the income distribution of the tax units below the national P60 fractile is identical across *départements*. We detail in Appendix B that the two specifications give similar results. We choose specification I because it gives the closest results to national estimates by [Garbinti et al. \(2018\)](#).

## 2.3 Concepts

To carry on our analysis we rely on four types of concepts.

**Theil Index** We start with the Theil indicator, a dispersion index, that has the advantage of taking into account the entire distribution and being independent of the average. Formally we define  $\mathbb{T}$  the total inequality between tax units as  $\mathbb{T} = \int_0^\infty f(\tilde{y}) \frac{\tilde{y}}{y} \ln \left( \frac{\tilde{y}}{y} \right) d\tilde{y}$  with  $f$  the income density and  $\tilde{y}$  the income. Interestingly the Theil index has a nice additive property. We can decompose  $\mathbb{T}$  the total inequality between tax units into two components. One is the between-*département* inequality, that is the inequality that would prevail if everybody within each *département* had the same income level, keeping the between-*département* inequality

unchanged. The other one is the within-*département* inequality, the inequality that would prevail if all *départements* had exactly the same average income keeping the shape of each *département* income distribution unchanged. In mathematical terms:

$$\begin{aligned}\mathbb{T} &= \mathcal{T} + \sum_{d=1}^D \frac{y_d n_d}{y \sum_{d=1}^D n_d} \mathbb{T}_d \\ &= \sum_{d=1}^D \frac{y_d n_d}{y \sum_{d=1}^D n_d} \log \frac{y_d}{y} + \sum_{d=1}^D \frac{y_d n_d}{y \sum_{d=1}^D n_d} \mathbb{T}_d\end{aligned}$$

with  $\mathcal{T}$  the inequality between *départements*,  $\mathbb{T}_d$  the total inequality between tax units of *département*  $d$ ,  $n_d$  the number of tax units in *département*  $d$ ,  $y_d$  the average income in *département*  $d$ ,  $D$  the total number of *départements* and  $y$  the average income of all tax units.

Importantly here when we compute  $\mathcal{T}$  the Theil index that represents the between *départements* inequality, units of observation are *départements* and they are *weighted* by their population. By weighting variables by population we adopt the point of view of people rather than territories. Doing so we depart from the strategy adopted by [Combes et al. \(2011\)](#).

**Income Distribution Partition** We split any income distribution into four income groups: (i) the top 1% tax units, (ii) the P90P99 tax units that is the tax units that belongs to the top 10% tax units but don't belong to the top 1% tax units, (iii) the middle 40 tax units, that is tax units whose income is above the median income but below the P90 fractile and (iv) the bottom 50 tax units, that is tax units whose income is below the median.

**Location of National Income Groups** We also study the location of different income groups of the *national* income distribution. For example we study the spatial distribution of the national top 1% richest tax units. We also study their segregation. We hence define the following quantities:

$$\Gamma(\mathbf{n}, x, y) = \sum_{d=1}^D \frac{n_d P(d, x, y)}{\sum_{d=1}^D n_d P(d, x, y)} \log \frac{n_d P(d, x)}{\sum_{d=1}^D n_d P(d, x, y)}$$

where  $P(d, x)$  is the share of *département*  $d$  tax units that belong to the  $PxPy$  percentiles of



national tax units.  $\mathbf{n}$  is the vector whose  $d$ -th element is  $n_d$ . Mathematically  $\Gamma(\mathbf{n}, x, y)$  is the unweighted Theil index of the number of tax units that belong to  $PxPy$  in each *département*. Importantly this index is invariant with respect to multiplication, hence it is comparable across income groups of different sizes.

**Distorsion Index** We also analyze the income distribution of each *département* in terms of fractiles of the national income distribution. We define distortion indices  $D(x, y, d) = \frac{P(d, x, y)}{y-x}$  where  $x$  and  $y$  are quantiles of the national income distribution and where  $P(d, x, y)$  is the share of *département*  $d$  tax units that belong to the  $PxPy$  percentiles of French metropolitan tax units. If the distorsion index is above one, the  $PxPy$  percentiles of French metropolitan tax units is overrepresented in  $d$ .

### 3 Results

We present here new empirical regularities on the evolution of spatial inequality.

Figure 2.1 represents the evolution of the between and within *département* inequality since 1966. The between *département* inequality represents the overall income Theil index that would prevail if there were no within *département* inequality. In that situation the income of a tax unit would be the its *département* average income. The within *département* inequality represents the overall Theil index that would prevail if there were no inequality between *départements*. In that situation the average income of a *département* would be the national average income. Using mathematical notations of Section 2, the between inequality is  $\sum_{d=1}^D \frac{y_d n_d}{y \sum_{d=1}^D n_d} \log \frac{y_d}{y}$ , the within inequality is  $\sum_{d=1}^D \frac{y_d n_d}{y \sum_{d=1}^D n_d} \mathbb{T}_d$ , and the sum of the two is equal to the total inequality between tax units  $\mathbb{T}$ . Both the between and within *départements* inequality went down between 1966 to 2017, decreasing from 0.038 to 0.018 for the inter *département* inequality and from 0.52 to 0.44 for the intra *département* inequality. The figure shows that the quasi totality of inequality comes from the inequality within *département*, the inequality between *départements* accounts only for 3.8% of the total inequality in 2017.

We provide evidence that the decrease in the inequality between *département* comes from

a convergence of the poorest tax units that conceal a divergence between the richest tax units. The Figure 2.2 represents the inequality between tax units from different income groups. We first look at the P0P50 between Theil index. This index compares the average income of the 50% poorest tax units of each *département*. In mathematical terms the P0P50 between Theil index is

$$\sum_{d=1}^D \frac{y_d^{P0P50} n_d^{P0P50}}{y^{P0P50} \sum_{d=1}^D n_d^{P0P50}} \log \frac{y_d^{P0P50}}{y^{P0P50}}$$

where  $y_d^{P0P50}$  is the average income of the 50% poorest tax units of  $d$ ,  $n_d^{P0P50}$  the number of 50% poorest tax units in  $d$ , also equals to  $\frac{n_d}{2}$ , and  $y^{P0P50} = \frac{\sum_{d=1}^D y_d^{P0P50} n_d^{P0P50}}{\sum_{d=1}^D n_d^{P0P50}}$  is the average income of all the tax units that are among the 50% poorest of their *département*. We do the same for the average income of the tax units that belongs to (i) the P50P90 of their *département*, (ii) the P9099 and the (iii) P99P100. The figure shows that in 2017, the poorer the income group, the smaller the Theil index is. The P0P50 between Theil index has slightly decreased, while the P99100 between Theil index has strongly increased. The P9099 between Theil index has also increased, though on a lesser extent than the P99100. It means that difference between poor tax units have decreased while differences between rich tax units are bigger and have increased. The decrease in the inequality between *départements* conceals these two simultaneous phenomenons.

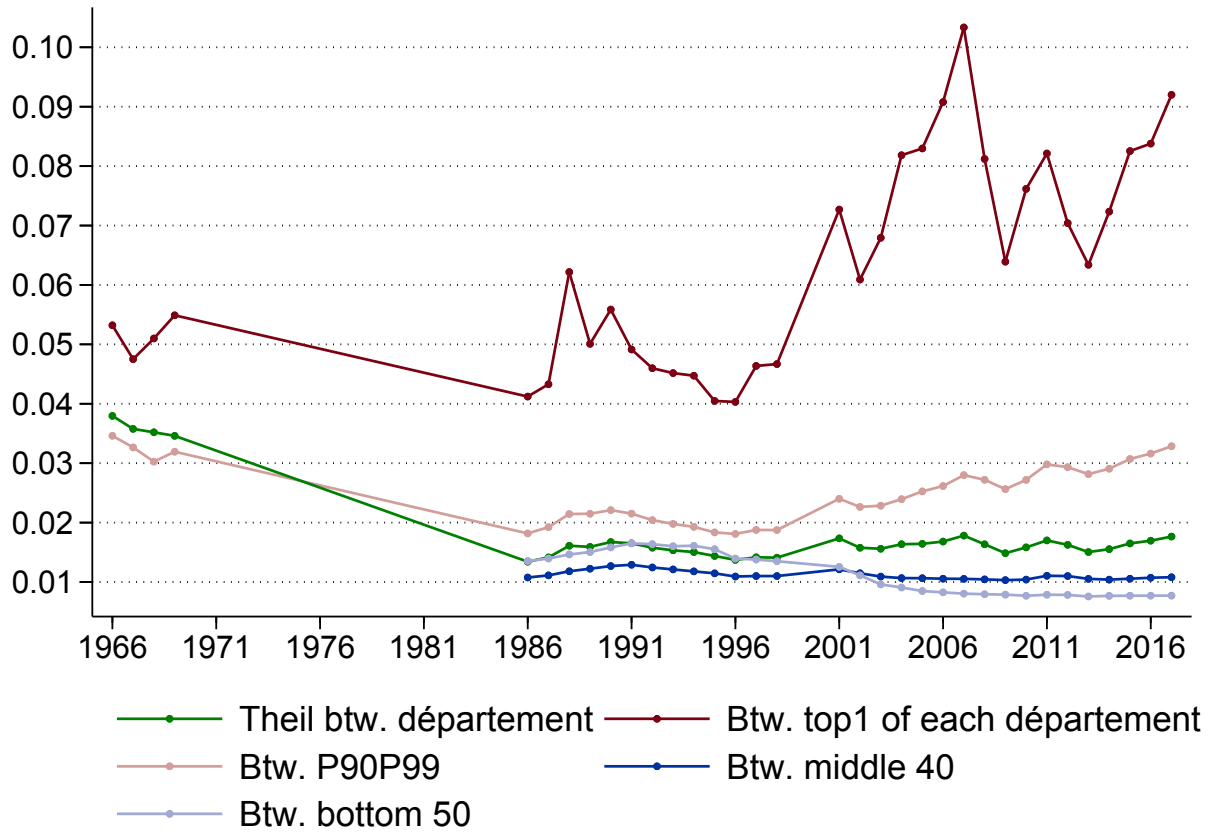
We further investigate the divergence of the richest tax units average income between *départements*. Richest tax units refers to the top 1% tax units of each *département*. Figure 2.3 maps the average income of the top 1% tax units of each *département* – we normalize by the average income of all tax units that belongs to the top 1% of their *départements*. In the interest of readability, our map zooms on the top left on Paris and its inner ring. The Paris region inner ring is made up of three *départements*, Hauts-De-Seine, Seine-Saint-Denis and Val de Marne. Figure 2.3 shows a clear pattern: in the center of France the tax units that belong to the top 1% of their *département* are poorer than the one of the Paris region and of some *départements* near the coasts or in the Alps. Importantly most of the *départements* where the

tax units that belong to the top 1% of their *département* are relatively richer, are *départements* where the biggest cities are located – *départements* in the Paris region, Rhône where Lyon is located, Bouches du Rhône where Aix-Marseille is located, Haute Garonne where Toulouse is located, Alpes Maritimes where Nice is located, Bas-Rhin where Strasbourg is located, Gironde where Bordeaux is located. There is also Haute Savoie, a *département* in the Alps, bordering Switzerland, where there is a lot of cross-border workers. Most importantly the Paris region plays a major role. The top 1% tax units of Paris are 3.2 times richer than their counterparts. In Hauts de Seine, in which is located la Défense, top 1% tax units are 2 times richer and in Yvelines 1.6 times richer. In addition, Seine-Saint-Denis in the Paris region, has a very different pattern than its region, as its top 1% tax units are 0.6 times poorer than their national counterparts. Figure 2.3 also maps the evolution between 1969 and 2017. Most French North Eastern *départements* experienced a relative decrease of their top 1% tax units average income. The *départements* where the relative income of the top 1% tax units increased are mainly the *départements* where big cities are located, but also more generally *départements* near the Atlantic coast, the south of France, or in the Northern Alps. The top 1% tax units in Paris are richer today than what they were relatively in 1969: 2.3 against 3.2 today; in Hauts de Seine 1.7 against 2 today, and in Yvelines 1.3 against 1.6 today. In Seine-Saint-Denis the opposite has happened as the top 1% tax units were 0.8 times poorer than their counterparts in 1969 and are 0.6 times poorer today.



Note: The table represents (i) the income inequality among all tax units, (ii) the inequality that would prevail if everybody within *département* had the same income level, keeping the between-*département* inequality unchanged, and (iii) the inequality that would prevail if all *départements* had exactly the same average income keeping the shape of each *département* income distribution unchanged. We use Theil index to measure inequality. The sum of this two latest quantities equals the first quantity. In 1966 the Theil on income among all individuals was equal to 0.56.

FIGURE 2.1 – BETWEEN AND WITHIN DÉPARTEMENT INEQUALITY MEASURED BY THE THEIL INDEX

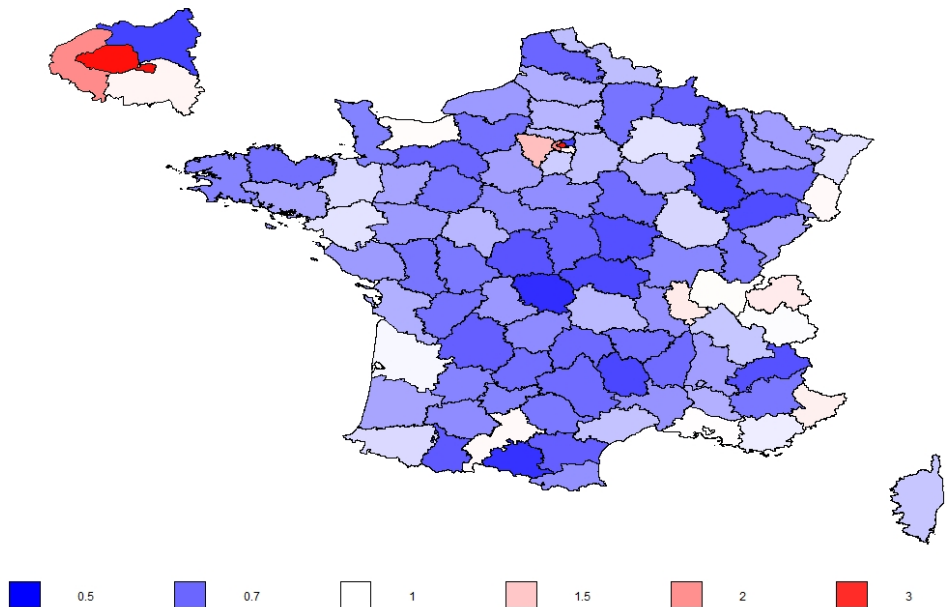


*Note:* The Figure represents the inequality between tax units that belong to different income groups. We first look at the P0P50 between Theil index. This index compares the average income of the 50% poorest tax units of each département. In mathematical terms the P0P50 between Theil index is

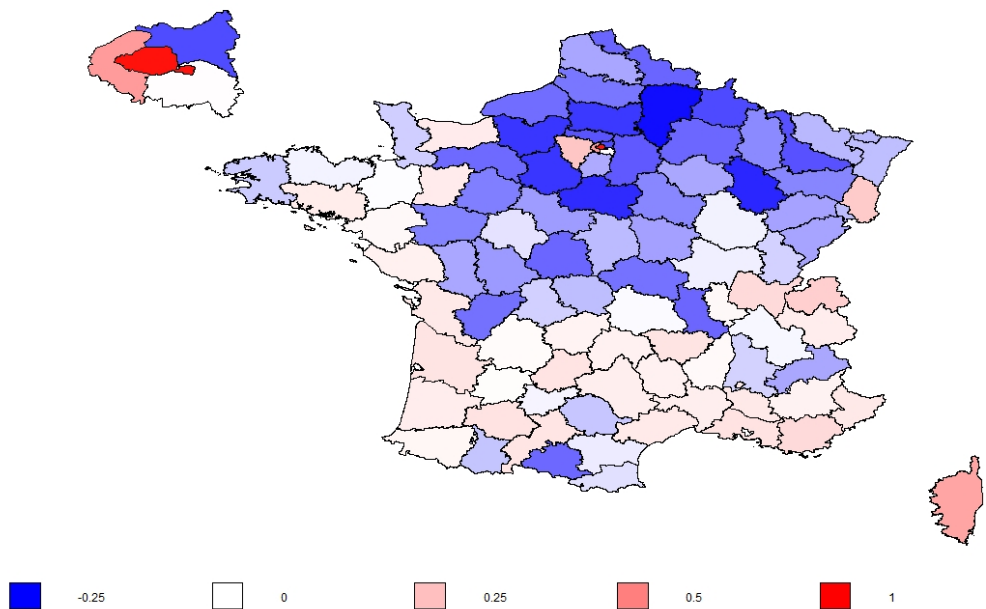
$$\sum_{d=1}^D \frac{y_d^{P0P50} n_d^{P0P50}}{y^{P0P50} \sum_{d=1}^D n_d^{P0P50}} \log \frac{y_d^{P0P50}}{y^{P0P50}}$$

where  $y_d^{P0P50}$  is the average income of the 50% poorest tax units of  $d$ ,  $n_d^{P0P50}$  the number of 50% poorest tax units in  $d$ , also equals to  $\frac{n_d}{2}$ ,  $y^{P0P50} = \frac{\sum_{d=1}^D y_d^{P0P50} n_d^{P0P50}}{\sum_{d=1}^D n_d^{P0P50}}$ . We do the same for the average income of the tax units that belongs to (i) the P50P90 of their département, (ii) the P90P99 and the (iii) P99P100, (iv) all income groups. In 2017 the P0P50 Theil index is equal to 0.007.

FIGURE 2.2 – BETWEEN DÉPARTEMENT INEQUALITY BY INCOME GROUPS



IN 2017



1969-2017 VARIATION

*Note:* In the interest of readability, the map zooms at the top left on Paris and its inner ring. The first map represents the average income of top 1% tax units of their *département*. The average income is normalized by the average income of all the tax units that belong to the top 1% of their *département*. The second maps represents the variation between 1969 and 2017. In 2017 the top 1% tax units of Paris are 3 times richer than the top 1% tax units of the other *départements*. In 1969 they were only 2 times richer.

FIGURE 2.3 – AVERAGE INCOME OF TOP 1% TAX UNITS OF THEIR DÉPARTEMENT

We investigate the contribution of each *département* to the total inequality. Figure 2.5 maps the contribution of each *département* to the total inequality in 2017. Using notations of Section 2 the contribution of a *département* to the total inequality is:

$$\frac{\frac{n_d}{\sum_{d=1}^D n_d} \frac{y_d}{y} \mathbb{T}_d}{\mathbb{T}}$$

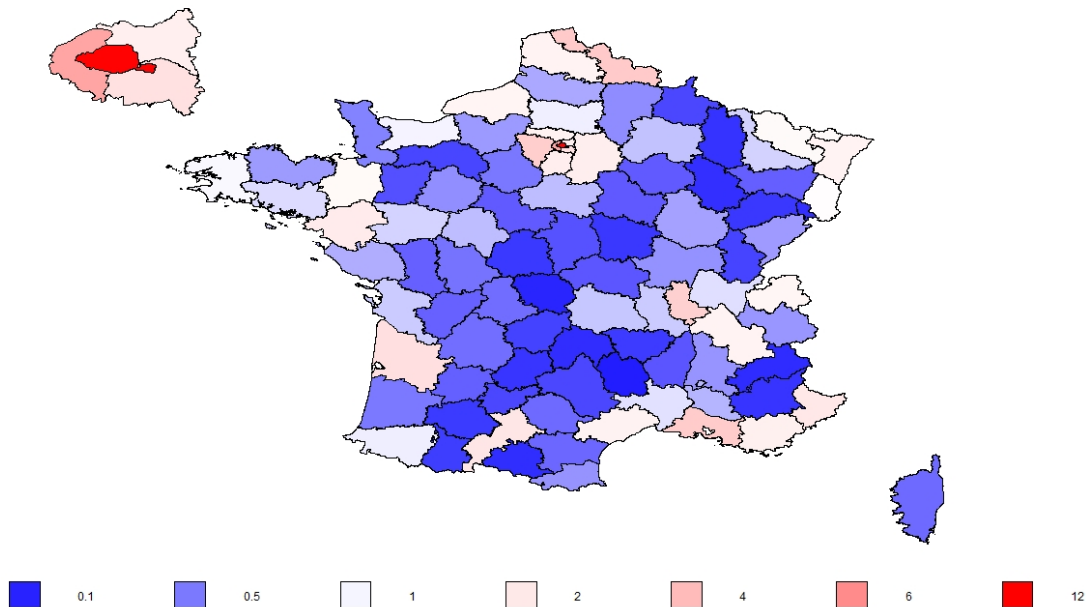
If all *départements* were equally contributing to the total inequality in 2017, their contribution would be  $(100 - 3.8)/95 \approx 1.01\%$ , with 3.8% the between *département* inequality in 2017 and 95 the number of *départements* in our study. In 2017 the majority of the *départements* contribute by less than 1.01%. The biggest contributors are *départements* where the biggest cities are located – *départements* in the Paris region, Rhône where is Lyon, Bouches du Rhône where is Aix-Marseille, Haute Garonne where is Toulouse, Alpes Maritimes where is Nice, Bas-Rhin where is Strasbourg, Nord where is Lille, Gironde where is Bordeaux, Loire Atlantique where is Nantes, Isère where is Grenoble. We suspect that the inequality in these *départements* are important because of their labor market. Some other *départements* are also above the 1.01% threshold, they are *départements* located either on the Atlantic Coast, the French Riviera or the French Alps. We suspect that the inequality in these *département* is important because they attract rich retirees. Most importantly the Paris region plays a major role in the total inequality, it accounts for circa 30% of the total inequality. Three *départements* are particularly responsible of the total inequality: Paris, Hauts de Seine and les Yvelines. Together they account for 20% of the total inequality while their tax units only represent 8.4% of the total number of tax units. In itself Paris accounts for approximately 12% of the total inequality, that is 3 times higher than the inequality between *départements*. In other words Paris contribution to the total inequality merely comes from the inequality within Paris rather than the difference of average income with other *départements*. Figure 2.5 also maps the evolution of the contribution of each *département* to the total inequality between 1969 and 2017. The majority of *département* saw their contribution decreased. The *départements* where the contribution increased are either *départements* where big cities are

located or *départements* near the coasts or in the Alps. In particular the contribution of Haute Garonne, Hauts de Seine, Yvelines, Bouches du Rhône, and Haute Savoie increased by more than 0.7% points each. The case of Haute Savoie is interesting because we suspect its evolution stems from cross-borders workers rather than a thriving labor market or attractiveness for retirees.<sup>6</sup>

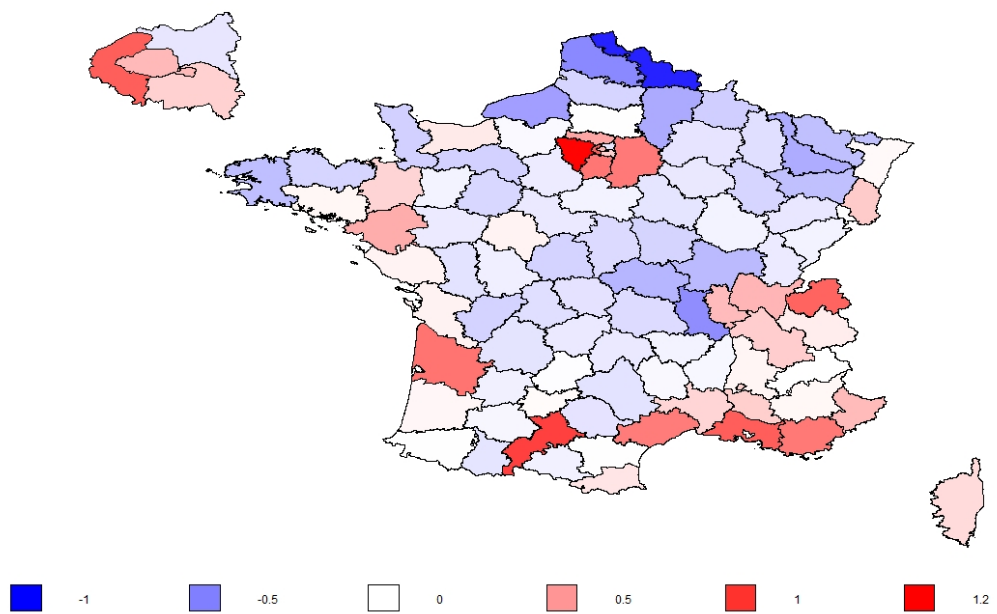
---

6. Without any causal implication, Figure 2.20 in the Appendix represents the share of the population living in Haute Savoie, aged between 20 and 64 years old, that is commuting to Switzerland. This share increased from 5.3% in 1968 to almost 19.4% in 2013.





IN 2017 (CONTRIBUTION IN %)



1969-2017 CONTRIBUTION'S VARIATION (IN % PTS)

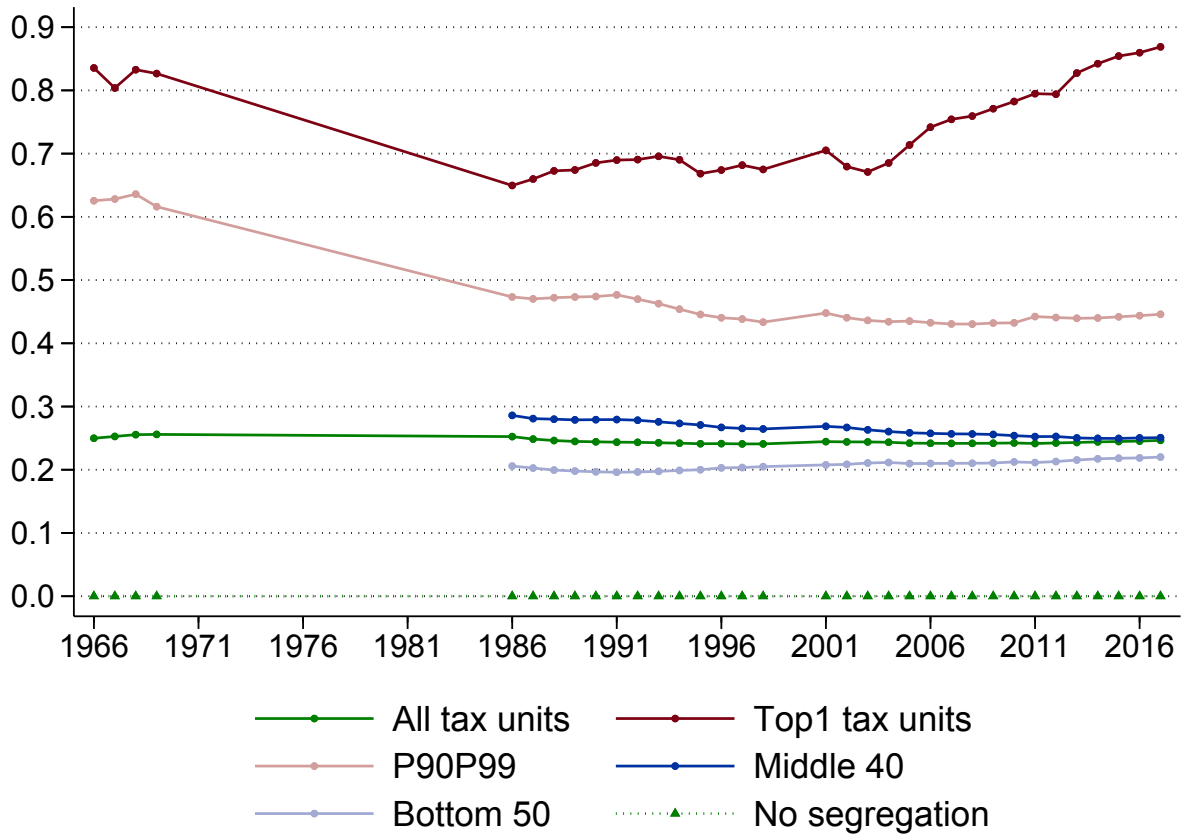
Notes: The first map shows the contribution of each *département* to the total inequality measured as the Theil index. Formally this contribution is  $\frac{\sum_{d=1}^D n_d \frac{y_d}{y} \mathbb{T}_d}{\mathbb{T}}$  using notations of Section 2. The second figure maps the evolution of this contribution between 1969 and 2017 in % pts. In 2017 Paris contributes by 12% to the total inequality. This share increased by 0.3 % pts between 1969 and 2017.

FIGURE 2.5 – CONTRIBUTION TO THE TOTAL INEQUALITY

We investigate the reason why some *départements* contribute more to the total inequality. We partition the national income distribution into four different groups, the national bottom 50, the national middle 40, the national P90P99 and the national top 1%. Figure 2.6 plots segregation measures of these four quantities. It also plots two benchmarks: (i) the degree of segregation of all tax units – which reflects the unequal spatial distribution of tax units across *départements* –, (ii) the case with no segregation, if tax units were equally spread, that is if they were  $\frac{100}{95}\%$  in each *département*. Interestingly, Figure 2.6 shows that the richer the income group, the more segregated it is. The national top 1% are much more segregated than the P9099 tax units, which are more segregated than the middle 40, which are more segregated than the bottom 50 tax units. Also importantly the top 1% tax units have first desegregated from the the 60's to the 80's and since the 2000's they have strongly segregated. At the same time both the P9099 and the middle 40 have desegregated. The bottom 50 have slightly segregated but it remains the less segregated income group.

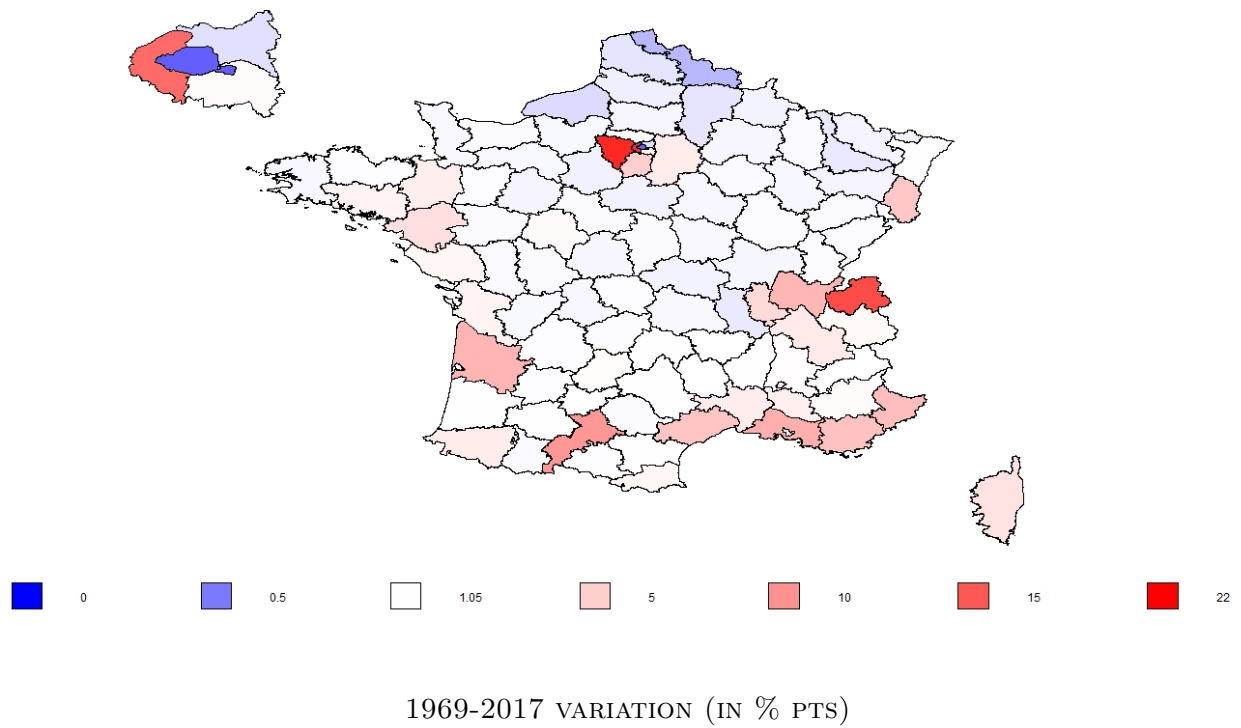
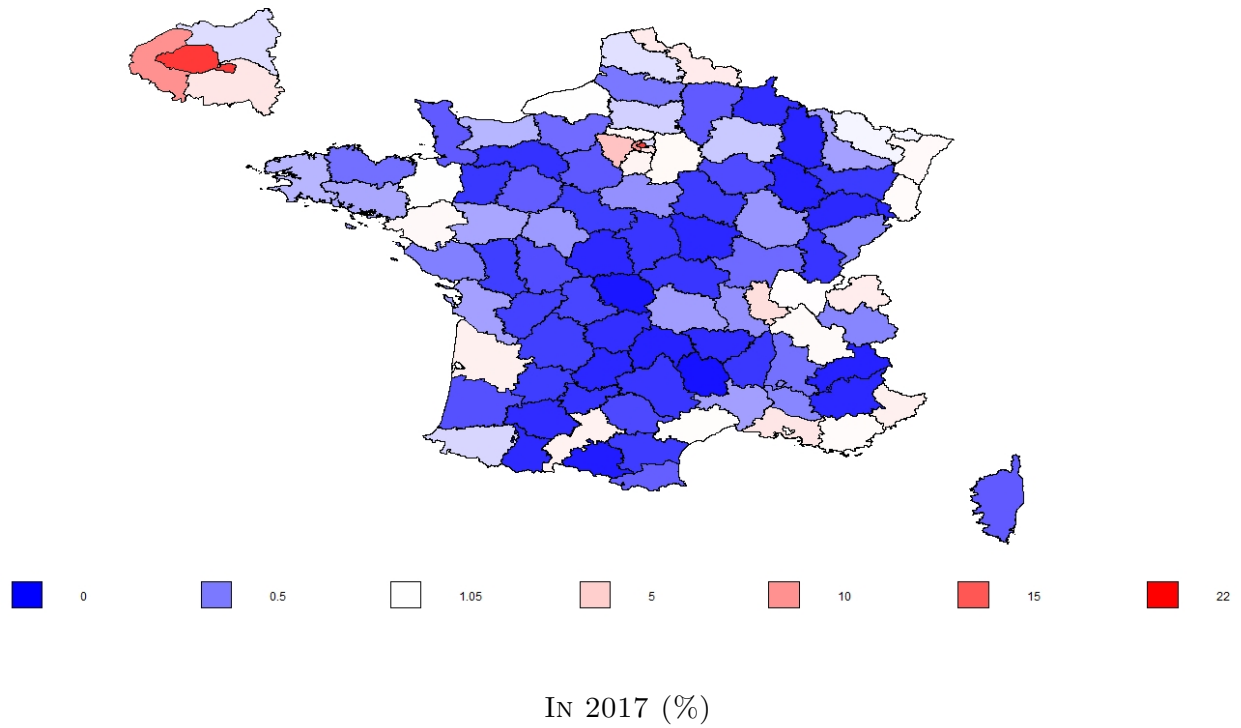
We know that the French top 1% tax units have strongly segregated since 1966. We know that the *départements* where big cities are increasingly contribute to the total inequality between national tax units. Now, we investigate the link between these two simultaneous patterns. We look at the location of the four different income groups across *départements*. Figure 2.7 maps the location of the national top 1% tax units. We represent the share of that income group living in each *département*: we plot  $\frac{n_d P(d,99,100)}{\sum_{d'} n_{d'} P(d',99,100)}$  for each *département*  $d$ . If national top 1% tax units were equally scattered across *département*, their share in each *département* would be  $\frac{100}{95}\%$ . On the contrary, Figure 2.7 shows that they are very segregated. In the vast majority of *département* the share of national top 1% tax units is well below 1.01%. The below 1.01% *départements* are all located in the center of France. Few *départements* account for the lion' share of the French top 1% tax units. Paris, Hauts de Seine, Yvelines, Val de Marne – four *départements* of the Paris region –, plus Rhône, Bouches du Rhône, Haute Garonne, Alpes Maritimes, Gironde, Nord and Haute Savoie account for 57% of the

top 1% tax units while they account for only 30% of the total number of tax units. Again these are the *départements* where the biggest French cities are located – Haute Savoie being an exception. In terms of evolution since 1969, most of the *départements* where big cities are located have seen an increase in their share of national top 1% tax units. Paris, Hauts de Seine, Yvelines, Val de Marne, Rhône, Bouches du Rhône, Haute Garonne, Alpes Maritimes, Gironde, Nord and Haute Savoie accounted only for 52% of the top 1% tax units in 1969. Paris is an exception. Its share of top 1% tax units was 21% in 1969 and is now 17%. Apart from the *départements* where big cities are located, other *départements* near the coasts or in the Alps have seen their national top 1% tax units share increased by more than 1% points. The case of Haute Savoie is particularly striking. Its share of top 1% tax units was 0.07% in 1969. It is 2.5 % in 2017 while Haute-Savoie only represents 1.2 % of the total number of tax units.



Notes: The Figure represents segregation indices of different incomes groups: (i) the national top 1% tax units, (ii) the national P90P99 tax units, (iii) the national middle 40 tax units, (iv) the national bottom 50 tax units. It also represents two benchmarks: (i) the degree of segregation of all tax units, and (ii) the degree of segregation that would prevail if tax units were equally split across *département*.

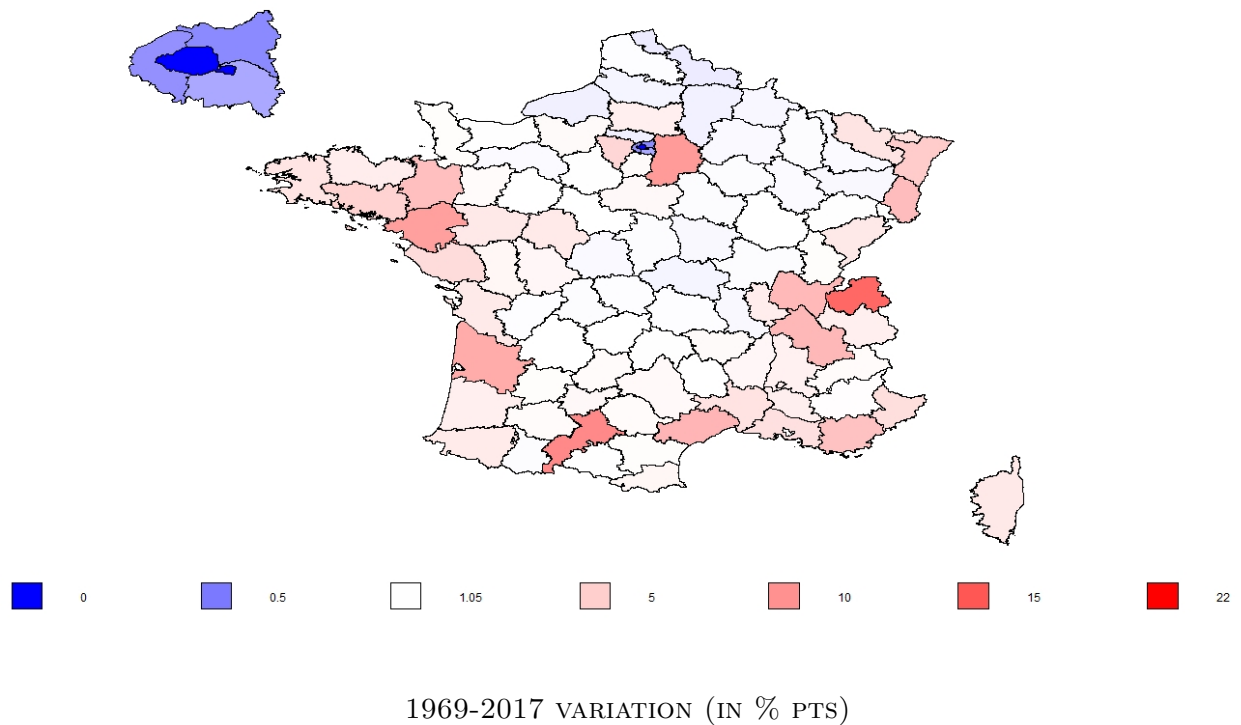
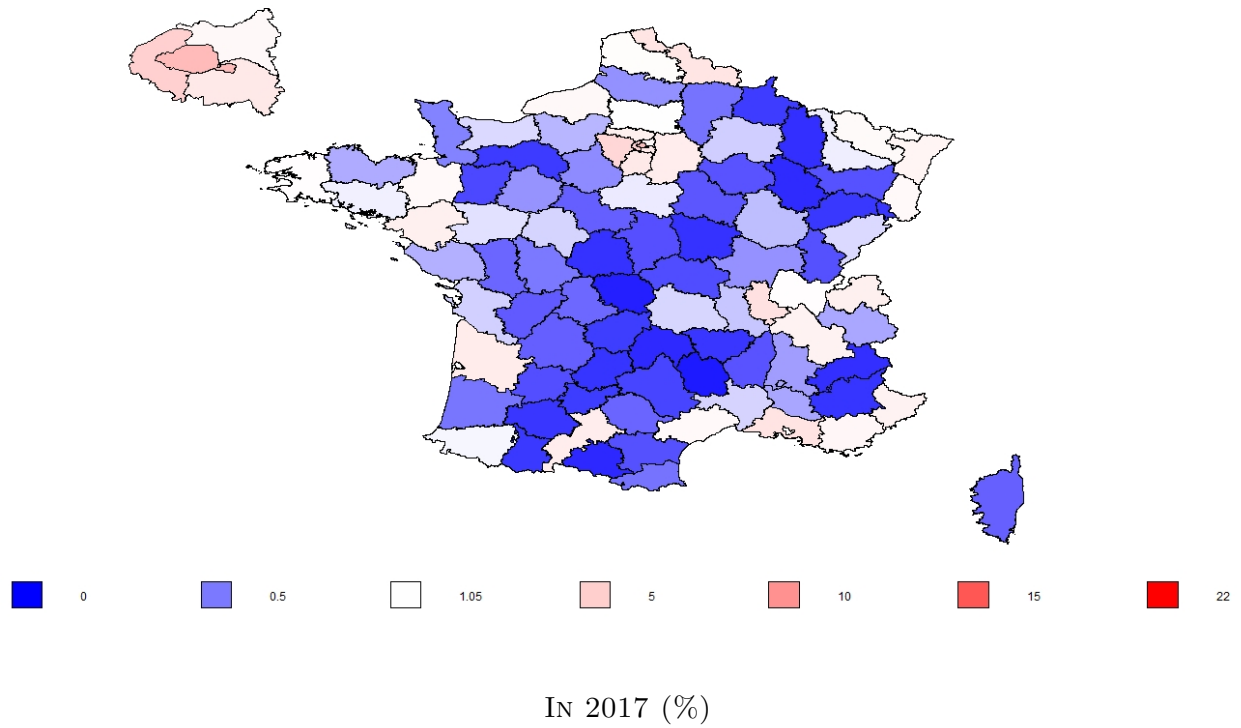
FIGURE 2.6 – NATIONAL INCOME GROUPS SEGREGATION



*Notes:* The first map shows the share of the national top 1% tax units living in each *département* in 2017. The second map show the variation between 1969 and 2017 in % pts. In 2017, 17% of the national top 1% tax units were living in Paris . Between 1969 and 2017 the share of national top 1% tax units living in Paris decreased by 3.5 % pts.

FIGURE 2.7 – NATIONAL TOP 1% TAX UNITS LOCATION (% LIVING IN EACH DÉPARTEMENT)

We present a similar map for the national P90P99 tax units in Figures 2.9. The national P90P99 are segregated in *départements* where big cities are located. However this segregation is less pronounced than for the national top 1% tax units. Paris, Hauts de Seine, Yvelines, Val de Marne, Rhône, Bouches du Rhône, Haute Garonne, Alpes Maritimes, Gironde, Nord and Haute Savoie accounts for 43% of the top P90P99 tax units in 2017 – against 57% for the top 1% –. Also importantly, since 1969, this share has decreased from 47% to 43 %.

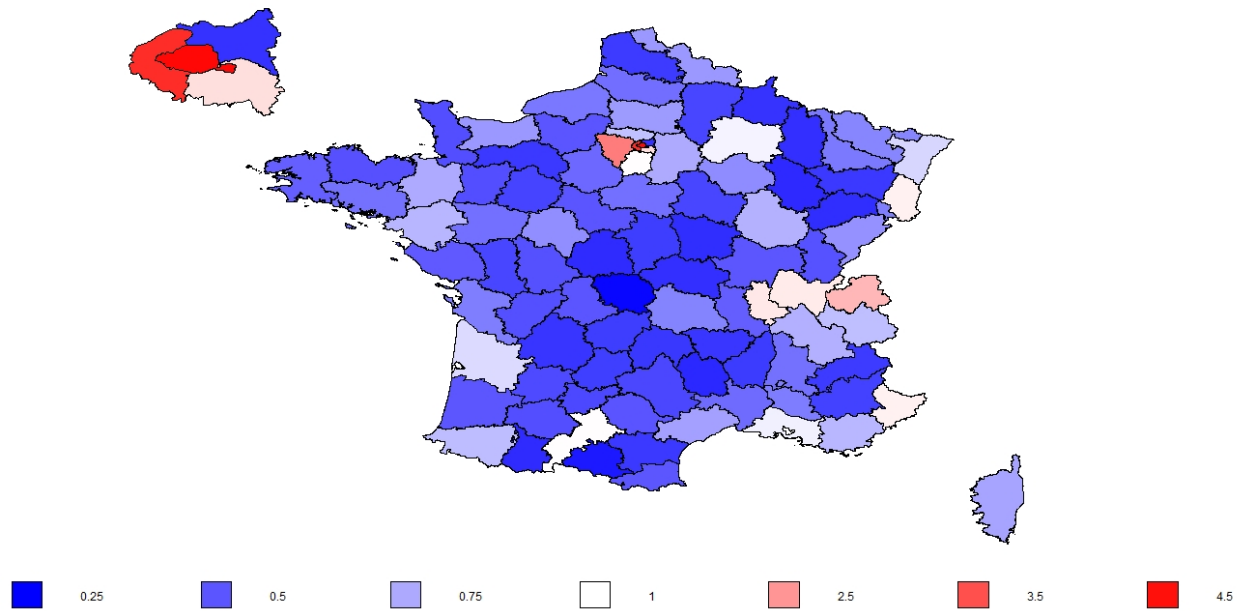


*Notes:* The first map shows the share of the national P90-P99 tax units living in each *département* in 2017. The second map shows the variation between 1969 and 2017 in % pts of the share of the national P90-P99 tax units living in each *département*. In 2017, 12% of the national P90-P99 tax units were living in Paris . Between 1969 and 2017 this share decreased by 5 % pts.

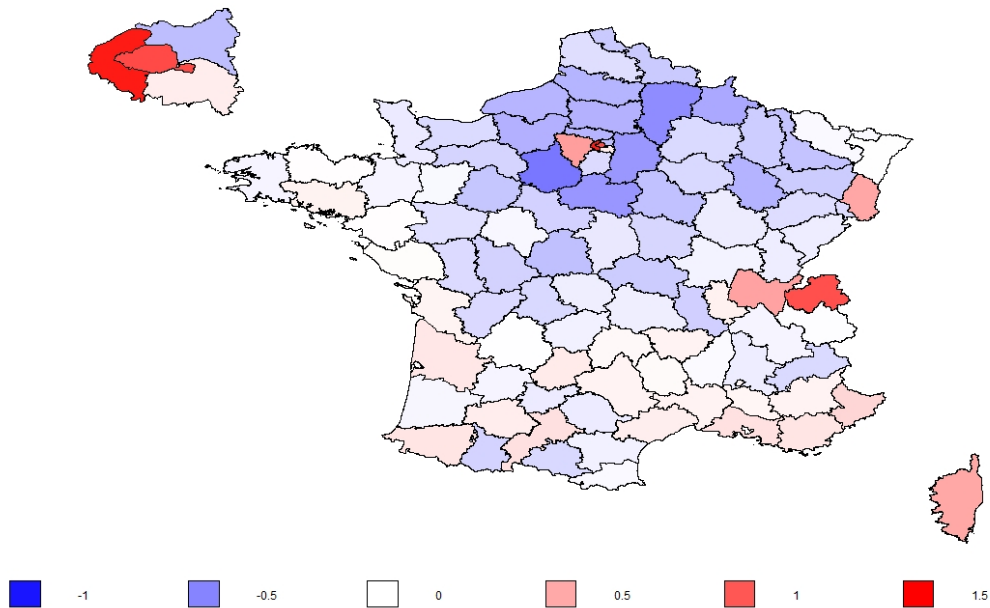
FIGURE 2.9 – NATIONAL P90-P99 TAX UNITS LOCATION (% LIVING IN EACH DÉPARTEMENT)

We have given insights into the location of the different French income groups. The fact that national top 1% tax units are getting more segregated, especially in the Paris region might explain part of the increasing feeling of disconnection between Paris and the rest of France. On the other hand what also matters for the perception of inequality might also be the level of inequality locally. Hence we compare the level of inequality within French *départements*. We keep our former partition of the French income distribution. We analyse if some income groups are over-represented in some *départements*. Figure 2.11 maps the over or the under-representation of the the national top 1% tax units. Mathematically we plot the distorsion index of the national top 1% tax units that we have defined in Section 2. If the distorsion index is above one in département  $d$ , it means that there are more French top 1% tax units than in the homogeneous case. In the homogeneous case, they would be 1% in each *département*. Figure 2.11 shows that the national top 1% tax units are over-represented in 11 départements. In Haute Garonne and Essone top 1% French are 1.02 times more numerous than if the homogeneous case. In Haut-Rhin et Alpes Maritimes they are 1.2 times more numerous, in Ain et Rhône 1.3, in Val de Marne 1.45. It increased up to Haute Savoie where they are 2 times more numerous, Yvelines 4.1, Hauts de Seine 4.1 and eventually Paris 4.5. Importantly, these *départements* have experienced the greatest increase in the over-representation of the French top 1%. The increase has been especially strong in Paris, Hauts de Seine, Haute Savoie. At the same time national top 1% are strongly under-represented in Seine-Saint-Denis. French top 1% tax units are 0.38 less numerous than what they should be in the homogeneous case. They were already under-represented in 1969 but to a lesser extent. At that time they were 0.65 times less numerous than in the homogeneous case.





NATIONAL TOP 1% TAX UNITS DISTORSION INDEX IN 2017

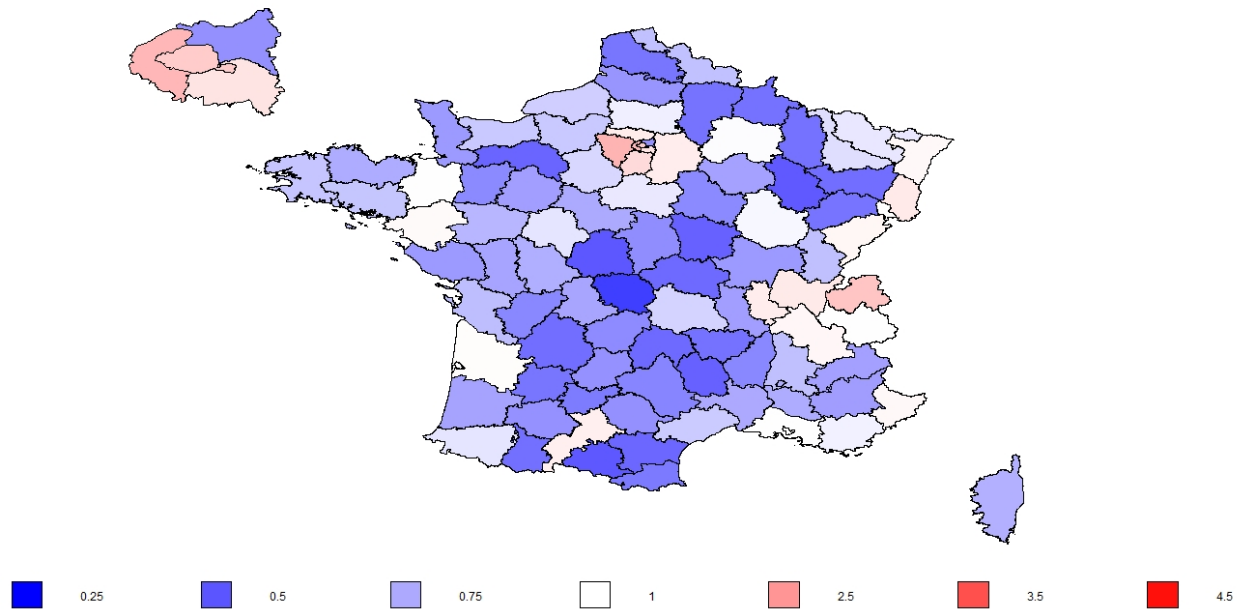


1969-2017 VARIATION

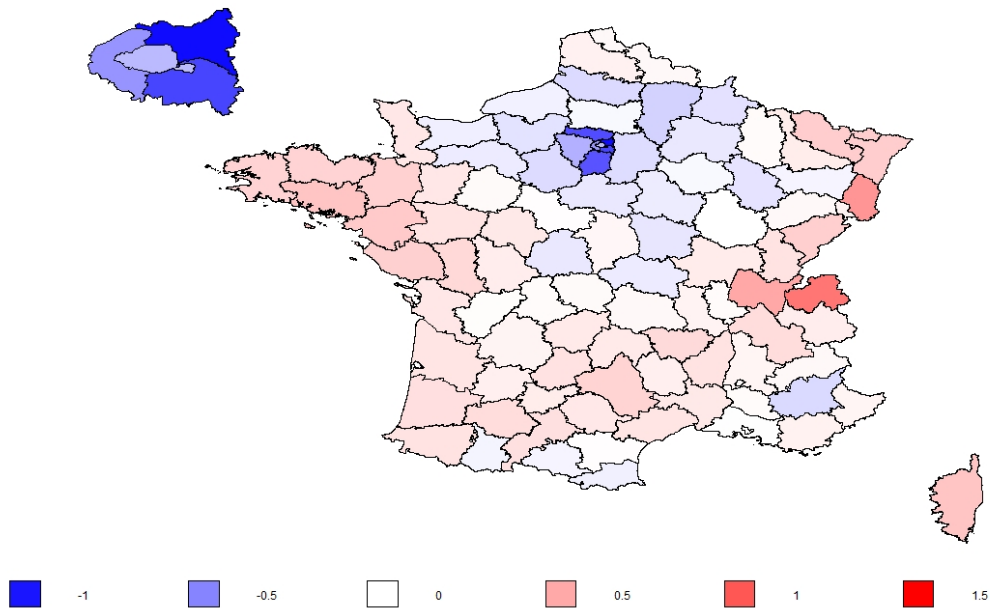
*Notes:* The first map shows the distorsion index of the national top 1% tax units living in each *département* in 2017. The second map shows the variation between 1969 and 2017 of the distorsion index. In 2017, there are 4.6 times more national top 1% tax units than if they were 1%. Between 1969 and 2017 the distorsion index of the top 1% tax units in Paris municipality increased from 3.5 to 4.6.

FIGURE 2.11 – OVER/UNDER-REPRESENTATION OF THE NATIONAL TOP 1% TAX UNITS

Figure 2.13 maps the over or under-representation of the national P90P99 tax units. There is less distortion for this income group. National P90P99 are over-represented in *départements* where big cities are located and in Haute Savoie. However the phenomenon is less acute than for the national top 1% tax units. Also very interestingly national P90P99 are disappearing of the Paris region. In 1969 they were twice as numerous than what they would have been in the homogeneous case. In 2017 they are only 1.48 times more numerous than the homogeneous case. The decrease has been particularly strong in Seine-Saint-Denis. In 1969 the national P90P99 were 1.71 times more numerous than in the homogeneous case. Now they are 0.68 times less numerous.



NATIONAL P0P99 TAX UNITS DISTORSION INDEX IN 2017



1969-2017 VARIATION

*Notes:* The first map shows the distorsion index of the national P90P99 tax units living in each *département* in 2017. The distorsion index compares the number of national P90P99 tax units living in a *département* to the number they would be if they were homogeneously scattered – in terms of number of tax units – across *départements*. In 2017, there are 1.7 times more national P90P99 tax units living in Paris municipality than what they should be in the homogenous case. Between 1969 and 2017 the distorsion index of the P90P99 tax units in Paris municipality decreased from 2.0 to 1.7.

FIGURE 2.13 – OVER/UNDER-REPRESENTATION OF THE P0P99 TAX UNITS

## 4 Conclusion

In this paper we have reconstructed the income distribution of each French metropolitan *département* at the end of the 60's and since 1986. We have uncovered new regularities on the evolution of spatial inequality in France. The total inequality in France merely derives from inequality within *départements*. The between *département* inequality is low and has strongly decreased. This decrease comes from a convergence of the poorest tax units that conceal a divergence between the richest tax units. The richest tax units are getting increasingly segregated in a few *départements* where biggest French cities are located. The richest tax units are over-represented in these *départements*. Importantly Paris, along with Hauts de Seine and Yvelines, is contributing massively to the total inequality. Right next to Paris, Seine-Saint-Denis is following a very different path, its richest tax units being increasingly and strongly under-represented.

This paper gives avenues for research. Our paper uncovers new empirical regularities on the location of different income groups. These regularities might be explained in part by a polarization in local labor markets. They could also be explained by retirees location choices. Disentangling and estimating the magnitude of each potential channel would be of interest.

Also one could investigate how our results would be changed if we were considering within and between French municipality inequality.

Lastly future research could investigate the link between the increasing over-representation of richest tax units in a few *départements* and political outcomes or social resentments.



# Appendix

## A Methodology for the 1986-2017 Period

We present here in full length our methodology to reconstruct the income distribution of each *département* for the 1986-2017 period. Between 1986 and 2017, fiscal tabulations include all tax units, when they are taxable or not.

### A.1 Cleaning Fiscal Tabulations

Blanchet et al. (2017) provides an algorithm that reconstruct very accurately the income distribution from fiscal tabulation. The algorithm requires the lower threshold of the highest bracket to be high enough to accurately estimate the top of the income distribution. Similarly the upper threshold of the lowest bracket must be low enough to accurately estimate the bottom of the income distribution. On top of that, the total number of brackets should not be too small to accurately estimate the middle of the distribution. Apart from 2001 and 2002, the number of brackets and the position of the thresholds in the income distribution meet these conditions.

In 2001 and 2002 they were only 6 brackets in the fiscal tabulations at the *département* level. We correct it by creating additional brackets. We transform fiscal tabulations at the *département* level for years 2001 and 2002 into tabulations with 12 brackets applying the following rule<sup>7</sup>: we define  $p_j^d$  as the share of tax units belonging to bracket  $j$  among all tax units of *département*  $d$ . We similarly define  $p_j$  as the share of tax units belonging to bracket  $j$  among all French tax units. We impose that

$$\left(\frac{p_j^d}{p_j}\right)_t = \left(\frac{p_j^d}{p_j}\right)_{2003} \quad \forall t \in \{2001, 2002\}$$

---

7. In 2003 to 2005 the fiscal administration kept the bracket thresholds of the previous years but doubled the number of brackets by splitting brackets into smaller brackets. For example, the 0–9,000 bracket in 2001-2002 was divided into two brackets in 2003: 0–7500 and 7500–9000. Besides, at the national level, the fiscal tabulations for 2001 and 2002 was already made of 12 brackets.

In a few isolated cases the number of tax units for a specific bracket is too small in a given *département* for a given year to interpolate the income distribution – always in rural and sparsely populated *départements* and for brackets at the top of the income distribution –. In those cases, we add up consecutive brackets.

## A.2 Taxable Income Distributions

We recover the taxable income distributions of each French metropolitan *département* by applying the algorithm developed by [Blanchet et al. \(2017\)](#). It is a non parametric method that allows us to recover with great precision the entire income distribution. We use the *gpinter* package provided by the authors.

## A.3 Fiscal and Pretax Income Distributions

Taxable income refers to the income earned by tax unit minus specific deductions imposed by the income tax law. We need to correct for these deductions because they have varied greatly over the years. We correct for it by transforming our taxable income distribution into (i) fiscal income without capital gains distribution, (ii) fiscal income with capital gains distribution and (iii) pretax income distribution. We apply the following rule:

$$Y_{x,t}^{H,d} = Y_{x,t}^{T,d} \frac{Y_{x,t}^{H,F}}{Y_{x,t}^{T,F}}$$

where  $Y_{x,t}^{H,d}$  is the average income in quantile  $x$  of the  $H$  income distribution of *département*  $d$  in year  $t$  and  $H$  stands either for (i) fiscal without, (ii) fiscal with or (iii) pretax.  $Y_{x,t}^{T,d}$  is the average income in quantile  $x$  of the taxable income distribution of *département*  $d$  in year  $t$ .  $Y_{x,t}^{H,F}$  is the average income in quantile  $x$  of the  $H$  income national distribution in year  $t$ ,  $Y_{x,t}^{T,F}$  the average income in quantile  $x$  of the taxable income national distribution in year  $t$ . We derive  $Y_{x,t}^{H,F}$  and  $Y_{x,t}^{T,F}$  from [Garbinti et al. \(2018\)](#). The underlying assumption is that the rank of a tax unit in the national taxable income distribution is equal to its rank in either the (i) fiscal without, (ii) fiscal with or (iii) pretax income distribution.

## A.4 Consistency with National Estimates

We make sure our results are consistent at the national level with [Garbinti et al. \(2018\)](#) estimates. When we add up our *département* estimates we find a slight differences of around 1% with national estimates. We correct our own national average income estimate to make it match with [Garbinti et al. \(2018\)](#). However [Garbinti et al. \(2018\)](#) include overseas *départements* when computing national estimates whereas we work only with metropolitan *départements*. We define  $Y^{T,n}$  as the average national taxable income. We similarly define  $Y^{T,m}$  as the average taxable income of metropolitan *départements*. We can get  $Y^{T,n}$  and  $Y^{T,m}$  from fiscal tabulations. We define  $Y^{F,n}$  the average national pretax income estimate by [Garbinti et al. \(2018\)](#). We define  $Y^{F,m}$  our estimate of the average pretax income of metropolitan *départements* before adjustment. We homogeneously correct the income distribution of each *département*. We multiply the income of each tax unit by:

$$\frac{Y_t^{F,n}}{Y_t^{F,m}} \times \frac{Y_t^{T,m}}{Y_t^{T,n}} \quad \forall t = \{1986, \dots, 1998, 2001, \dots, 2014\},$$

We assure that the latter is consistent with national estimates only in terms of average. We need to make sure that the income share of each quantile of the national distribution is close to the same quantile of the summed distribution. We test the fit of our estimates by comparing the top 0.1% share from [Garbinti et al. \(2018\)](#) with our adjusted estimates. The figure 2.14 presents the results of this comparison for the bottom 90 income share, the top 10 without the top 1 income share and the top 1 income share. Differences are deemed small for the period.

## B Methodology for the 1966-1969 Period

We present here in full length our methodology to reconstruct the income distribution of each *département* for the 1966-1969 period. We only have fiscal tabulations of the taxable tax units for this period. Hence we need to reconstruct the number of non taxable tax units and



the income distribution of the non taxable tax units of each *département*.

## B.1 Recovering the Number of Non Taxable Tax Units

First we recover the total number of tax units by *département* based on an econometric method. We run the following regression for years 1986–2010:

$$N_t^d = \alpha_0 + \sum_{a=1}^6 \alpha_a P_t^{d,a} + \beta M_t^d + \gamma N_n + \delta_d + \varepsilon_t^d$$

$N_d$  is the total number of tax units of *département*  $d$ ,  $N_n$  the total number of tax units at the national level,  $M^d$  the number of married couples in *département*  $d$ ,  $P^{d,a}$  the number of individuals in the age group  $a$  in *département*  $d$  – there is six age groups  $a$  –.  $\delta_d$  is a *département* fixed effect.  $\varepsilon_d$  is an error term. We do not use the years 2011 to 2014 to fit the model because the definition of tax unit changed after 2011. Since 2011, married and divorced couples no longer complete three returns the year of their union or divorce. Overall, the  $R^2$  is equal to 0.9996. To test the consistency of our estimates, we sum up the estimated departmental tax units and compared them to the national tax units estimated by [Garbinti et al. \(2018\)](#). National estimates slightly differ from our own estimates because of overseas *départements*. If we assume that the share of overseas *départements* in the total number of tax units is 1%, then we overestimate the total number of metropolitan tax units by 1.8% points in 1966 and 1.25% point in 1969. We adjust homogeneously our tax number estimates to correct for this discrepancy. We also test the fit of this model by comparing predicted values and observed values for the years 1986 to 2010. Our method has a good fit on average for 1995-2010 period even though the difference can be up to more than 5% in a few cases. For the previous period, values are more overestimated, especially in 1986 when non taxable tax units are required for the first time to file a tax return. We do not use the years 2011 to 2014 to fit the model since the definition of tax unit changed at that date. Married and divorced couples no longer complete three returns in the year of their union or divorce (one for each spouse and one for the couple), which is no longer true.

We also compare our method to another methodology to estimate the total number of fiscal units. In this second method we assume that all individuals over age 20 complete a tax return unless they are married. Hence we compute total fiscal units by subtracting the total of married couples to the total of individuals aged 20 or more. Table 2.1 presents the distribution of the difference in % between estimated and observed values with the two methods. On average, our preferred method estimates the number of tax units very properly between 1995 and 2010. Before 1994 values are slightly more overestimated. The second method is doing less good.

## B.2 Recovering the Average Income of the Non Taxable Tax Units

We also need to estimate the income distribution of the non taxable tax units. According to Piketty (2001) the share of taxable tax units among all tax units was around 44% in 1966 and increased gradually up to 51% in 1969. Taxable tax units are typically richer than non taxable tax units but it is not always the case. Some rich and large family may be non taxable because the French tax income schedule depends of the size of the family. These situations are rare however. Piketty (2001) discusses this matter extensively. We make strong assumptions to impute the income distribution of the non taxable tax units. We test two specifications with radically different assumptions. We provide evidence that the magnitude of our results are unchanged with these two specifications. We believe that we provide reliable estimations of the location and income distribution of French top 40 % tax units. We do not use the data on the bottom 60 tax units for the period 1966-1969.

We define  $Y^d$  as the average income of all tax units in *département*  $d$ ,  $Y_1^d$  the average income of the taxable tax units in *département*  $d$ ,  $Y_0^d$  the average income of the non taxable tax units in *département*  $d$ ,  $N_1^d$  the number of taxable tax units,  $N_0^d$  the number of non taxable tax units,  $L_t^d(x_d)$  the income share of the  $x_d$  poorest tax units in year  $t$  and *département*  $d$ .  $L_t^d$  is the Lorenz curve of the income distribution of *département*  $d$  in year  $t$ . We alternatively define  $\tilde{Y}_1^d$  the average income of the taxable tax units in *département*  $d$  that belongs to the top 40% French tax units,  $\tilde{Y}_0^d$  the average income of remaining tax units in *département*  $d$ ,

$\tilde{N}_1^d$  the number of taxable tax units in *département*  $d$  that belongs to the top 40% French tax units,  $\tilde{N}_0^d$  the number of the remaining tax units. We define  $L_t^N(x)$  the income share of the  $x$  poorest tax units in the *national* income distribution. We get the Lorenz curve  $L_t^N$  from national estimates by [Garbinti et al. \(2018\)](#).

In the first specification we make two assumptions. First we assume that the non taxable units are the tax units at the bottom of the income distribution of their *département*. We further assume that the income share of the non taxable tax units during the 1960-1969 period is equal to the income share of the  $x_d$  poorest tax units in 1986 where  $x_d$  is equal to the share of non taxable tax units in 1960-1969. In mathematical terms it means:

$$L_t^d\left(\frac{N_{0,t}^d}{N_{0,t}^d+N_{1,t}^d}\right) = L_{1986}^d\left(\frac{N_{0,t}^d}{N_{0,t}^d+N_{1,t}^d}\right) \text{ for any } t \in \{1960 - 1969\} \text{ and } d.$$

We deduce the average income of tax units living in *département*  $d$  with the following formula in specification I:

$$Y^d = \frac{Y_1^d}{1 - \tilde{L}\left(\frac{N_0^d}{N_0^d+N_1^d}\right)} \frac{N_1^d}{N_1^d + N_0^d}$$

In the second specification we impute not only the income distribution of the non taxable tax units but also the one of the taxable tax units that are below the P60 national fractile. We choose the P60 fractile as we know that at least 40% of all tax units are taxable since 1966. Here we assume that the tax units that are either non taxable or don't belong to the top 40% French tax units are at the bottom of the national income distribution. In mathematical terms we assume that their income share in the national income is equal to  $L_t^N\left(\frac{\sum_{d'} \tilde{N}_0^{d'}}{\sum_{d'} \tilde{N}_0^{d'} + \tilde{N}_1^{d'}}\right)$ . We further assume that this income is distributed over *départements* according to their share of below P60 non taxable tax units. In mathematical terms it means that the average income of the below P60 non taxable tax units in any *département*  $d$  is:

$$\frac{L_t^N\left(\frac{\sum_{d'} \tilde{N}_0^{d'}}{\sum_{d'} \tilde{N}_0^{d'} + \tilde{N}_1^{d'}}\right) \sum_{d'} \tilde{Y}_1^{d'} N_1^{d'}}{1 - L_t^N\left(\frac{\sum_{d'} \tilde{N}_0^{d'}}{\sum_{d'} \tilde{N}_0^{d'} + \tilde{N}_1^{d'}}\right)} \frac{1}{\sum_{d'} \tilde{N}_0^{d'}}$$

with  $\frac{L_t^N \left( \frac{\sum_{d'} \tilde{N}_0^{d'}}{\sum_{d'} \tilde{N}_0^{d'} + \tilde{N}_1^{d'}} \right) \sum_{d'} \tilde{Y}_1^{d'} N_1^{d'}}{1 - L_t^N \left( \frac{\sum_{d'} \tilde{N}_0^{d'}}{\sum_{d'} \tilde{N}_0^{d'} + \tilde{N}_1^{d'}} \right)}$  the total income of the below P60 non taxable tax units and  $\frac{1}{\sum_{d'} \tilde{N}_0^{d'}}$  the total number of below P60 non taxable tax units. Then we can deduce the average income in *département*  $d$ .

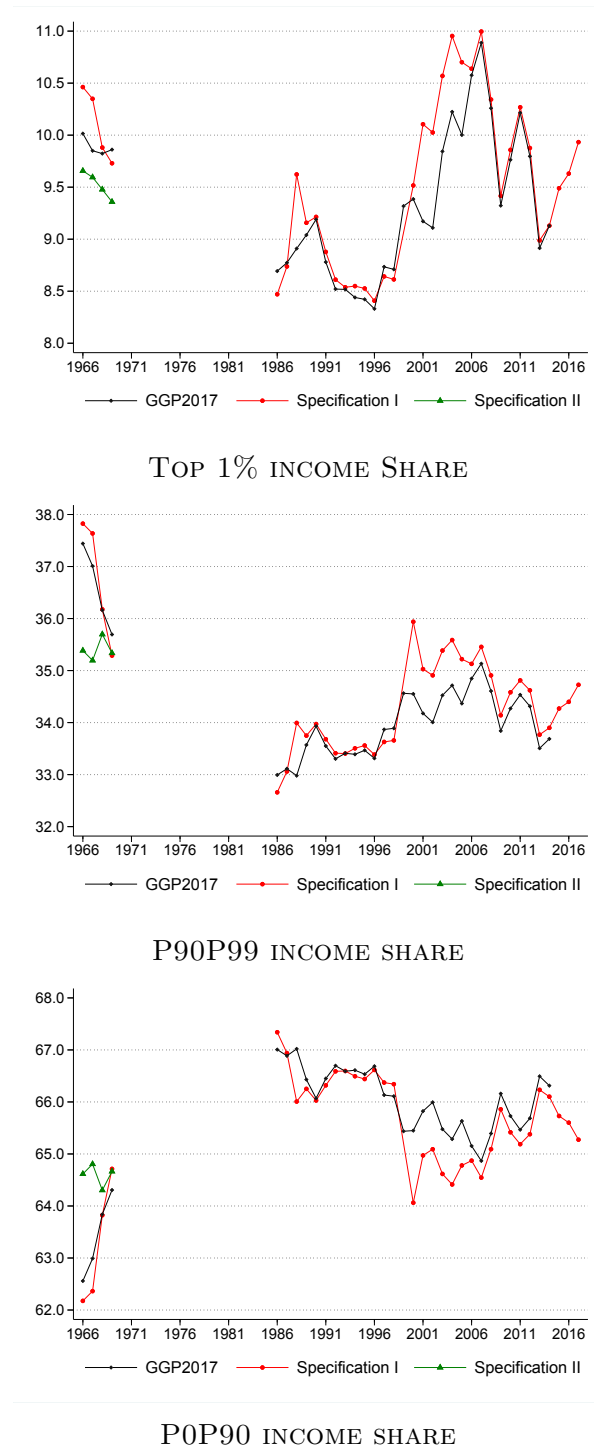
Once imputed the average income of each *département* we are back to the 1986-2017 methodology.

In a nutshell, in the first specification we use 1986 information and assume some stability in the income share of the bottom of the distribution. In the second specification we use information on the national income distribution at the year of estimation. We assume that the income distribution of the tax units below the national P60 fractile is identical across *départements*. We assume that the income distribution of the the tax units below the national P60 fractile is identical whatever the *département*.

Figure 2.14 shows the income shares of the national distribution and the sum of *département* distributions with the two specifications. Both specification are closed to national estimates but underestimates national top 1% and P90PP99 income shares. However difference are much bigger for specification II – 0.5 versus more than 2 % points of difference. Hence we choose specification I because it gives the closest results results to national estimates by Garbinti et al. (2018).

We also compare the distributions obtained with the two specifications. First we compare the predicted numbers of tax units that belong to the national top 1% tax units, the national P90P90 tax units and the national P090 tax units in each *département*. Figure 2.18 presents the results. Differences are negligible. Then we compare the predicted income share of the top 1%, P90P99 and P0P90 tax units of each *département* obtained with the two specifications. We also compare predicted average income in each *département*. Figure 2.18 presents the results. On average specification II underestimates top 1% and P90PP99 income share. It is unsurprising as specification II also underestimates the national top 1% income shares and

the national P90P99 income share.



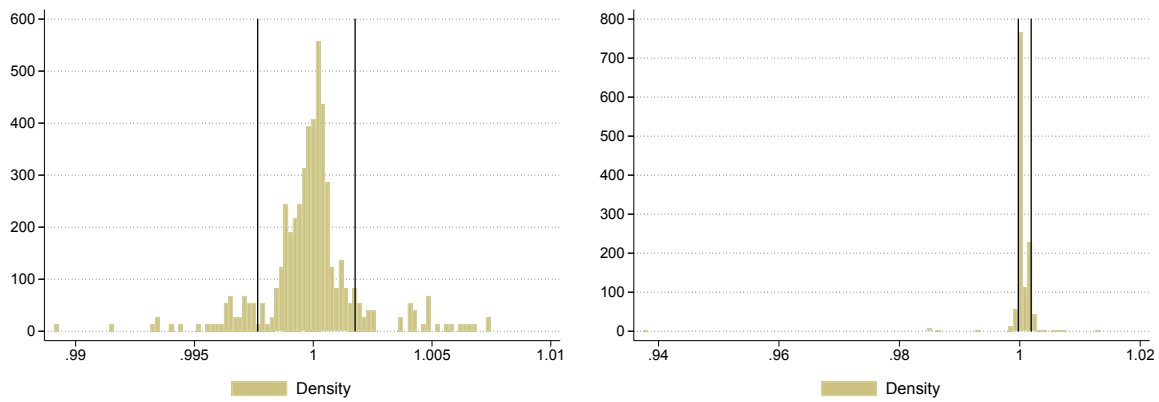
*Notes:* These figures compare three income distribution income shares: (i) the national distribution, (ii) the sum of the *département* income distribution produced with specification I, (iii) the sum of the *département* income distribution produced with specification II. The first figure presents the income shares of the top 1% tax units. The second presents the income share of the P90P99 tax units. The third figure presents the income share of the P0P90 tax units. In 1966 the top 1% income share in the national distribution was 10%.

FIGURE 2.14 – INCOME SHARES OF THE NATIONAL DISTRIBUTION AND THE SUM OF DÉPARTEMENTS DISTRIBUTIONS

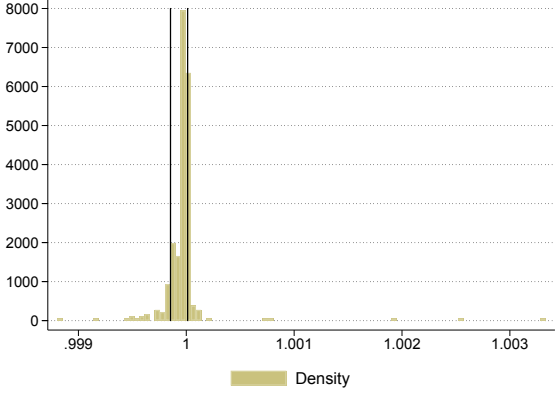
TABLE 2.1 – ANNUAL DISTRIBUTIONS OF DEPARTMENTAL DIFFERENCES BETWEEN TAX UNITS RECORDED AND ESTIMATED ACCORDING TO THE TWO METHODS (IN %)

Year	Preferred Method					Second Method				
	Min.	Quart. 1	Méd.	Quart. 3	Max.	Min.	Quart. 1	Méd.	Quart. 3	Max.
1986	-3.6	-0.8	0.5	3.1	20.1	-0.2	2.65	4.7	9.95	46
1990	-2.9	-1.35	-0.7	0.3	6.1	-1.5	0.75	2	4.5	24.2
1994	-2.4	-0.3	0.1	0.6	2.9	-1.3	1.1	1.8	3.4	17.2
1998	-3.8	-0.8	-0.2	0.25	1.7	-4.4	-1.1	-0.3	0.9	6.3
2002	-4	-0.55	0.3	0.65	1.7	-7.3	-1.7	-0.6	0.25	4.2
2006	-3.8	-0.4	0	0.65	3.1	-7.4	-2.35	-1.4	-0.7	3.2
2010	-3	-0.2	0.7	1.45	5	-7	-2.45	-1.2	-0.2	4.1

*Note:* Our preferred method estimates tax units by an econometric method based on demographic data. The second method estimates tax units by subtracting married couples from adult population. “Quart. 1” means that 25% of observations have a difference lower than the threshold. “Med.” is the usual median. Sample includes 95 *départements*. Values in % are ratio between tax units estimated and recorded.



NUMBER OF TAX UNITS IN NATIONAL TOP 1% NUMBER OF TAX UNITS IN NATIONAL P90P99

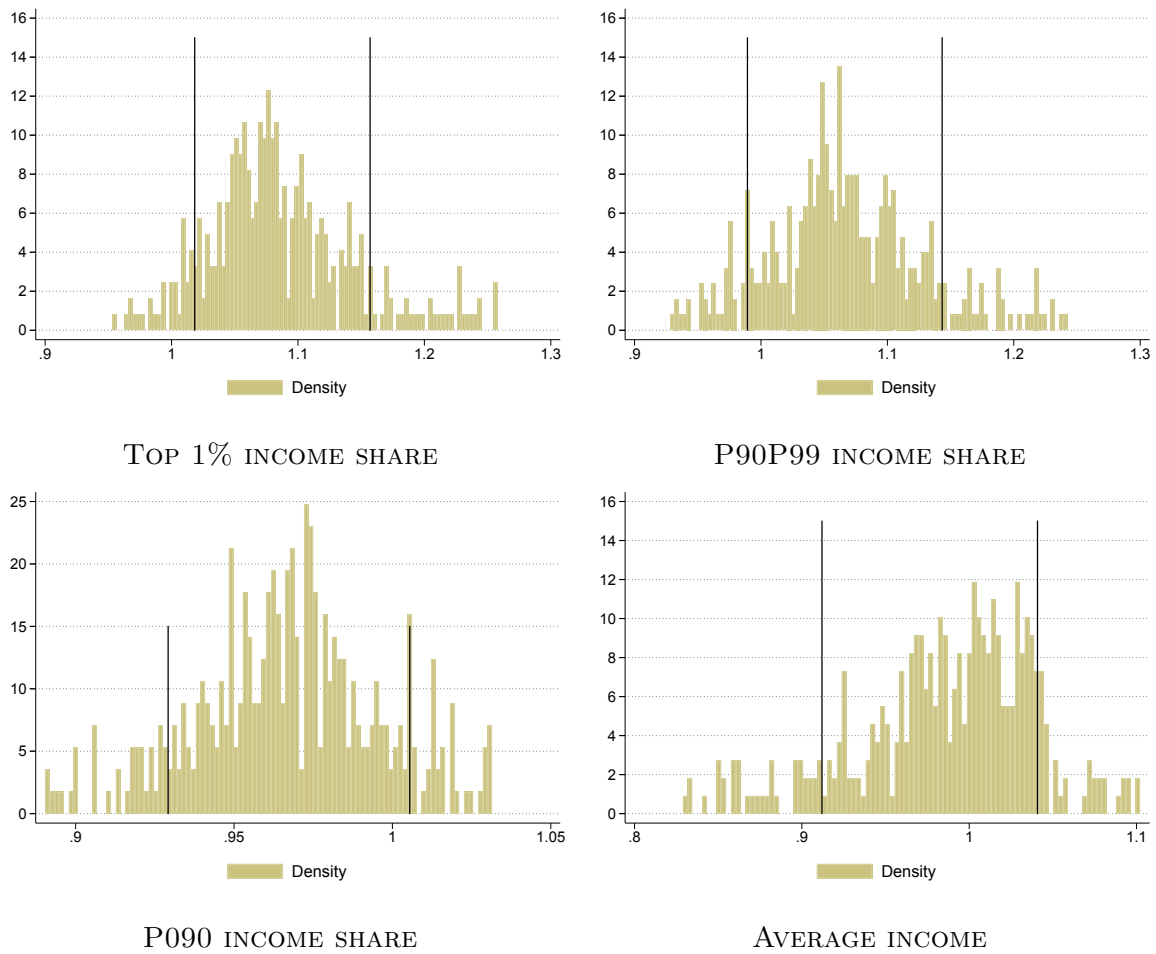


NUMBER OF TAX UNITS IN NATIONAL P0P90

*Notes:* These figures compares the income distribution estimated with the two specifications. Units of observations are *départements* for years in {1966-1969}. We compare the predicted numbers of tax units that belong to the national top 1% tax units, the national P90P90 tax units and the national P090 tax units in each *département*. We plot the distribution of the ratio of the predicted values in specification I over predicted values in specification II. The black lines represent the 10th and 90th percentiles.

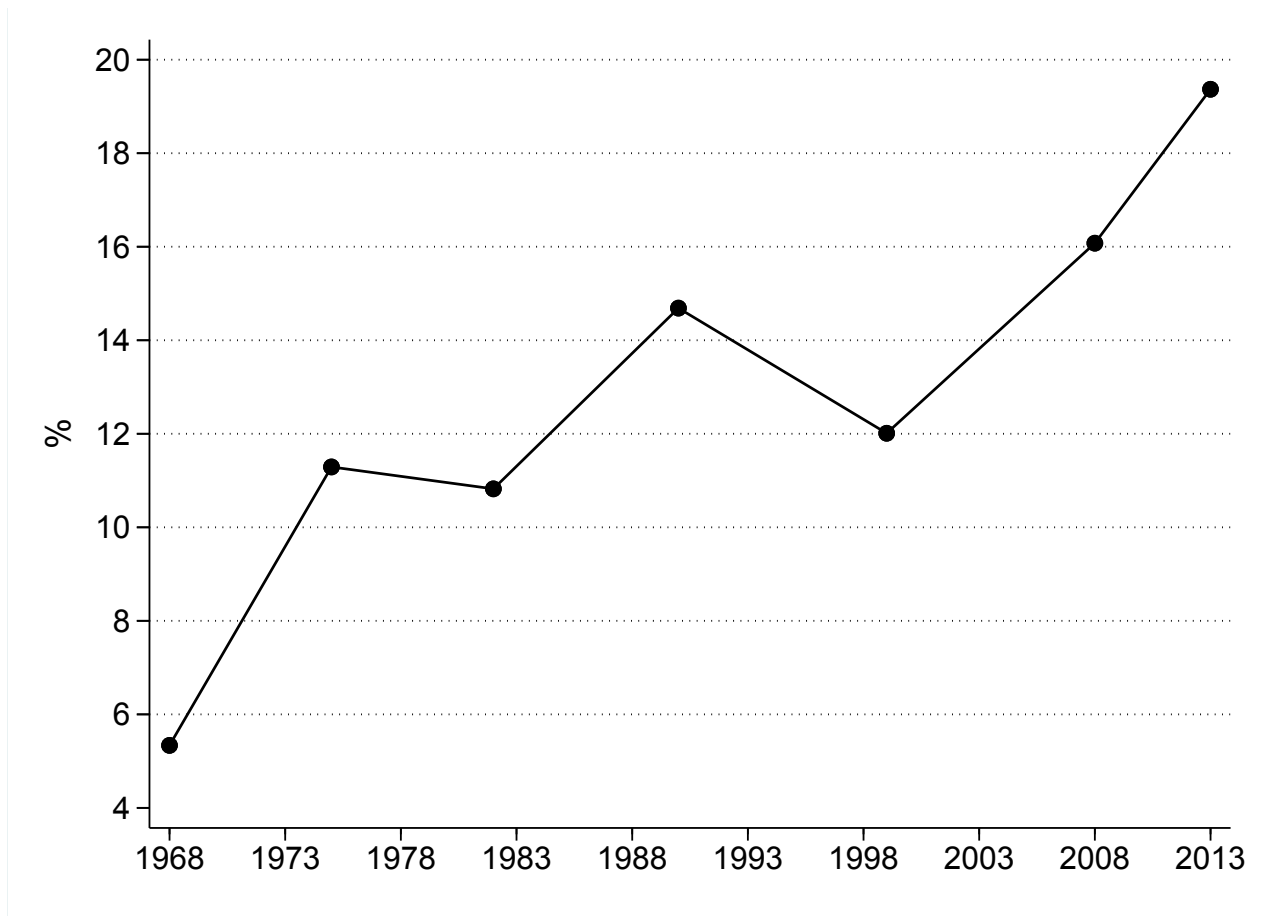
FIGURE 2.16 – RATIO OF PREDICTED VALUES IN SPECIFICATION I OVER SPECIFICATION II  
(1)





*Notes:* These figures compares the income distribution estimated with the two specifications. Units of observations are *départements* for years in {1966-1969}. We compare the top 1% income share, the P90P90 income share, the P0P90income share and the average income. We plot the distribution of the ratio of the predicted values in specification I over predicted values in specification II. The black lines represent the 10th and 90th percentiles.

FIGURE 2.18 – RATIO OF PREDICTED VALUES IN SPECIFICATION I OVER SPECIFICATION II  
(2)



*Notes:* The figure represents the share of the population of Haute Savoie, aged 20 to 64 years old, that is working in Switzerland. In 1968 this share was 5.3%. Source: French Census

FIGURE 2.20 – SHARE OF THE WORKING AGE POPULATION LIVING IN HAUTE SAVOIE THAT ARE COMMUTING TO SWITZERLAND (IN %)

## Chapter 3

# Les inégalités de revenu entre les départements français depuis cent ans<sup>1</sup>

### Résumé

Cet article utilise différentes sources et méthodologies pour analyser l'évolution des inégalités spatiales de revenu entre départements depuis 1922. La contribution la plus notable est la reconstruction du revenu moyen par département avant et après impôt sur le revenu grâce à la collecte et la numérisation inédite d'archives du ministère des Finances. Nous observons les faits stylisés suivants: (i) une très forte baisse des inégalités interdépartementales de revenu fiscal par adulte sur un siècle, avec deux périodes de baisse continue : entre 1922 et 1939 et depuis 1948 ; (ii) l'impôt sur le revenu réduit significativement les inégalités interdépartementales mais son effet a fortement varié au cours de la période considérée ; (iii) depuis 1948, tous les départements au sud de la ligne allant du Calvados au Gard ont vu leur situation relative s'améliorer.

**JEL codes:** N34 ; N94.

**Keywords:** Inégalités de revenus, départements français, impôt sur le revenu.

---

1. Ce chapitre reprend un article coécrit avec Florian Bonnet et Hippolyte d'Albis.

## Introduction

Les résultats électoraux et mouvements sociaux récents de plusieurs pays développés ont suscité un intérêt grandissant pour la question des inégalités spatiales. Les métropoles semblent s'être bien remises de la crise financière de 2008 tandis que les territoires moins denses continuent de souffrir des conséquences de la récession. Ces territoires, qui souffrent d'une double érosion de leur population et de leurs services publics et de proximité, semblent toujours plus éloignés des riches métropoles. L'objet de cet article est de mettre cette évolution en perspective grâce à une reconstitution inédite des inégalités interdépartementales de revenu depuis 1922.

L'échelle du département est utile et pertinente. Utile, car leur périmètre est relativement stable depuis leur création en 1789, ce qui facilite la comparaison historique des données statistiques qui leur sont associées. Pertinente, car c'est à cet échelon que se déploient les politiques de développement territorial en lien avec la solidarité nationale, telles que les politiques sanitaires et sociales, éducative et d'aménagement. Pertinente également du fait de l'attachement des populations à leurs départements, comme l'illustrent diverses anecdotes : forte opposition en 2008 au projet de suppression du numéro de département des plaques d'immatriculation, forte proportion de groupes Facebook de Gilets Jaunes faisant référence au département (20% selon [Boyer et al. \(2020\)](#)), etc. Des reconstitutions historiques des revenus à des échelles plus petites, telles que la commune, sont envisageables mais elles se heurteraient aux épineuses questions de l'évolution spatiale des villes, à la prise en compte du revenu des ruraux et à la déconnexion entre les lieux d'habitation et de travail. L'échelle du canton serait également intéressante mais leur périmètre a beaucoup trop varié au cours du XXème siècle pour que l'on puisse faire une analyse statistique temporellement cohérente.

Dans cet article, nous documentons l'évolution des inégalités interdépartementales de revenus avant et après paiement de l'impôt sur le revenu. Notre première contribution est d'avoir reconstitué ces revenus moyens pour chacun des 90 départements français (nous excluons les départements d'outre-mer et les nouveaux départements issus des scissions des

années 1960 et 1970). L'étude des inégalités spatiales en France n'est évidemment pas nouvelle, mais jusqu'à présent les historiens et chercheurs en sciences sociales ne disposaient que de mesures indirectes et parcellaires des niveaux de vie locaux sur longue période. Plusieurs travaux ([Combes et al. \(2011\)](#) ; [Bazot \(2014\)](#)) se sont dès lors appuyés sur des reconstitutions départementales de la valeur ajoutée par habitant. Si cette dernière permet d'étudier l'évolution de la localisation des activités et les différences spatiales de productivité, elle risque néanmoins d'offrir une mesure biaisée des inégalités de niveaux de vie du fait des transferts monétaires qui interviennent entre les territoires après la phase de production.

Nous utilisons ici des données fiscales relatives à l'impôt sur le revenu, ce qui nous permet de mesurer d'une façon nouvelle et plus directe les niveaux de vie au sein de chaque département. Nous calculons, en outre, les revenus moyens avant et après impôt sur le revenu pour étudier l'évolution de la redistribution interdépartementale opérée par cet impôt pendant près d'un siècle. Pour construire notre base de données sur les revenus, nous avons collecté un grand nombre d'informations dans les archives du Ministère des Finances, et avons numérisé en particulier, pour chaque département français, le nombre d'unités sujettes à l'impôt sur le revenu, le revenu imposable qu'elles ont déclaré ainsi que le montant total d'impôt qu'elles ont payé. Nous utilisons également la base de données de [Bonnet and Sotura \(2020\)](#) sur la distribution des revenus de chaque département pour les périodes 1960-1969, 1986-1998 et 2001-2017 et la base de données de [Bonnet \(2018b\)](#) sur la population par âge de chaque département depuis 1901. Grâce à une méthode originale, nous en déduisons le revenu moyen par département avant et après impôt depuis 1922.

La construction d'indicateurs d'inégalités interdépartementales nous permet alors d'analyser l'évolution qui a pris place depuis un siècle. Nous observons globalement une très forte baisse des inégalités de revenu fiscal par adulte, et particulièrement aux cours des périodes 1922-1939 et 1948-2015. La période entre 1922 et 1948 est marquée par le déclin d'un arc nord-est de départements entourant Paris, ainsi que par la détérioration relative de la situation de la plupart des départements de la façade atlantique, de la vallée de la Loire, et dans une moindre mesure de l'Ile-de-France. La baisse des inégalités depuis 1948 s'apparente quant à

elle à un phénomène de convergence qui touche la plupart des départements. Ainsi, sous une ligne allant du Calvados au Gard, tous les départements ont connu une amélioration de leur situation relative : le revenu moyen d'un lozérien était égal à 83% du revenu moyen d'un adulte en 2015, contre seulement 50% en 1948. Les habitants des départements de Haute-Garonne et Loire-Atlantique (où l'on retrouve Toulouse et Nantes) profitent aujourd'hui d'un revenu supérieur à la moyenne nationale, alors qu'il était 15% inférieur en 1948. A l'inverse, les départements des trois grandes métropoles du « Paris-Lyon-Marseille » ont vu leur situation relative se détériorer. Ce dernier phénomène est particulièrement marquant car il se distingue du phénomène connu de hausse de la productivité dans les grandes métropoles et illustre la déconnexion entre la valeur ajoutée par habitant d'un département et son revenu moyen. Outre les trois grandes métropoles, la dégradation du revenu relatif concerne essentiellement les départements situés au nord-est d'une ligne allant du Calvados au Jura, à l'exception notable des départements alsaciens.

Nous constatons également que l'impôt sur le revenu réduit significativement les inégalités interdépartementales avec trois périodes distinctes: jusqu'au début des années 1950, la réduction engendrée est faible ; elle prend ensuite progressivement de l'ampleur jusqu'à la fin des années 1980 pour atteindre un maximum en 1989 au moment où le taux d'imposition au niveau national est au plus haut. Depuis 1989, la baisse des inégalités engendrée est plus faible. Enfin, en termes de répartition spatiale du revenu en France, on observe un très fort mouvement de concentration entre les années 1940 et 1950, qui a été effacé au cours des deux décennies qui ont suivi. En revanche, depuis la fin des années 1970, la concentration des revenus reste inchangée. Ce phénomène est en opposition avec celui de la concentration spatiale accrue de la valeur ajoutée qui a eu lieu dans les dernières décennies et réaffirme l'intérêt de disposer de données de revenu (voir [Sanchis et al. \(2015\)](#)).

Cet article est principalement lié à deux champs de la littérature économique. Tout d'abord, il s'inscrit dans celui des inégalités de revenus, où l'on retrouve l'article fondateur de Piketty [Piketty \(2001\)](#) sur les hauts revenus en France, suivi par de nombreux travaux qui ont étudié les inégalités de revenu dans d'autres pays, tels que [Atkinson \(2005\)](#) pour le

Royaume-Uni, [Atkinson and Salverda \(2005\)](#) pour les Pays-Bas, ou [Alvaredo \(2009\)](#) pour le Portugal, ou encore récemment par [Garbinti et al. \(2018\)](#) pour la France. Nous étendons donc cette littérature aux dynamiques locales, qui par leur diversité permettent aussi de mieux comprendre l'évolution nationale.

Notre article est également lié à la littérature étudiant les processus de convergence et divergence régionales. [Williamson \(1965\)](#), à partir d'une analyse des revenus régionaux d'un ensemble de pays dont la France, montre que l'évolution spatiale des inégalités a suivi une courbe en U inversée jusqu'au milieu du 20ème siècle. Dans les premières phases de développement d'un pays les différences régionales s'accroissent car les régions les plus avancées bénéficient le plus du développement : elles deviennent relativement plus productives et concentrent une part croissante de la production. Puis, dans un second temps, la mobilité des facteurs de production et les rendements décroissants engendrent un processus de convergence. L'existence d'une courbe en U inversée de la concentration spatiale est aussi présente dans la littérature de la nouvelle économie géographique initiée par [Krugman \(1991\)](#). La littérature empirique qui a suivi a confirmé cette analyse en ayant recours à des données de valeurs ajoutées régionales reconstruites en utilisant la méthodologie de [Geary and Stark \(2002\)](#).

Pour la France, les économistes intéressés ont eu recours à un autre type de méthode pour reconstruire la valeur ajoutée par département et n'ont pu construire des données que pour un nombre limité d'années. [Combes et al. \(2011\)](#) utilisent des données de Toutain (1992) pour les années 1860 et 1930 et des données de l'INSEE pour les années 1980 et 2000 ; [Bazot \(2014\)](#) reconstruit des données de valeurs ajoutées tous les dix ans entre 1840 et 1911 ; [Caruana-Galizia \(2013\)](#) développe un modèle économétrique basé sur la composition sectorielle des départements ; enfin, [Sanchis et al. \(2015\)](#) complètent les données de [Combes et al. \(2011\)](#) par des données de l'INSEE sur la période 2000-2014. Ces auteurs constatent une forte divergence entre départements de la valeur ajoutée par habitant depuis les années 2000. Cette divergence se retrouve dans la littérature en économie urbaine qui analyse les marchés du travail locaux depuis les années 1980 : [Moretti \(2012\)](#) et [Diamond \(2016\)](#) pour les métropoles aux États-Unis, mais aussi [Lessmann \(2014\)](#) et [Lessmann and Seidel \(2017\)](#) qui

trouvent que les inégalités entre territoires augmentent graduellement.

La suite de cet article est structurée comme suit : nous présentons tout d’abord les choix méthodologiques effectués lors de la réalisation de notre base de données, ainsi que les variables et indicateurs que nous utilisons. Nous présentons ensuite l’évolution des inégalités interdépartementales entre 1922 et 2015.

## 1 Données et méthodes

La construction de la base de données utilisée repose sur l’utilisation de deux bases de données départementales récentes, sur la collecte de données originales pour l’ensemble du territoire métropolitain et sur une procédure d’estimation pour prédire les données manquantes. La variable d’intérêt est le revenu fiscal d’un département rapporté à sa population d’adultes (20 ans et plus). Le revenu fiscal correspond au revenu déclaré par les ménages avant la prise en compte des divers abattements auxquels ils ont droit. Nous nous intéressons également au revenu après impôt, soit la différence entre le revenu fiscal et le montant de l’impôt sur le revenu acquitté dans chaque département. Nous déterminons ces revenus par adulte pour chaque département et pour chaque année entre 1922 et 2015. Le choix de la date initiale est lié à celle du premier recensement effectué après la Première Guerre Mondiale pendant laquelle l’impôt sur le revenu fut créé.

Le périmètre concerne 90 départements français. Afin de conserver une structure géographique stable tout au long de la période, nous avons procédé à plusieurs choix méthodologiques. Premièrement, nous avons conservé les périmètres qui prévalaient avant les scissions opérées dans les années 1960 et 1970. En 1964, il est décidé de réorganiser la région parisienne et de scinder (à partir de 1968) le département de la Seine en quatre départements (nommés Hauts-de-Seine, Paris, Seine-Saint-Denis et Val-de-Marne) et celui de la Seine-et-Oise en trois (Essonne, Val-d’Oise et Yvelines) . En 1975, la Corse scinde son département en deux en créant la Corse-du-Sud et la Haute-Corse. À partir de ces dates, nous avons utilisé les données relatives aux départements nouvellement créés pour reconstruire celles relatives aux



populations des départements initiaux. Deuxièmement, les départements d’outre-mer ne sont pas inclus dans l’analyse car les séries statistiques ne sont disponibles que depuis une période trop récente. La liste et la carte des 90 départements étudiés sont disponibles dans l’annexe (Graphique 3.11).

Notre première source statistique est la base de données de [Bonnet and Sotura \(2020\)](#), que dont l’on présente la construction dans le chapitre précédent. Nous avons estimé, à partir d’archives administratives produites par les services fiscaux, la distribution des revenus fiscaux dans chaque département et pour chaque année des périodes suivantes : 1960-1969, 1986-1998 et 2001-2017. Nous utilisons le revenu fiscal total (hors gains en capitaux) de chaque département. La seconde source statistique est la base de données issue de [Bonnet \(2018b\)](#). L’auteur estime annuellement la population par âge de chaque département. En recoupant ces deux sources, on dispose donc du revenu fiscal par adulte pour chaque département pour les années 1960-1969, 1986-1998 et 2001-2017.

Nous avons en outre collecté pour notre estimation de nouvelles données. Il s’agit, pour chacun des 90 départements et pour toutes les années allant de 1922 à 2015, d’une base contenant les variables suivantes : le revenu imposable déclaré par unité taxable, le nombre d’unités sujettes à l’impôt sur le revenu et le montant total des impôts de ces unités. L’année pertinente est celle du revenu pour laquelle les impôts sont collectés (et non celle de la collecte). Nous avons numérisé des tableaux statistiques contenus dans les archives du Ministère des Finances conservées sur le site de Savigny-le-Temple. Pour la période allant de 1922 à 1974, nous avons numérisé les tables provenant des Renseignements Statistiques Relatifs aux Impôts Directs (RSRID, volumes édités entre 1930 et 1975). Pour l’année 1975 et la période 1978-2000, nous avons numérisé les tables provenant de l’Annuaire Statistique de la Direction Générale des Impôts (ASDGI, volumes 1976, 1979-2002). Nous avons également récupéré les données des tabulations fiscales numérisées par [Bonnet et Sotura \(2020b\)](#) ; on note que les données sur le revenu imposable sont indisponibles pour les années 1978-1985 et que celles sur le nombre d’unités sujettes à l’impôt le sont pour les années 1986-1989. Pour toutes les années à partir de 2003, nous utilisons les ASDGI disponibles en ligne sur le site de la Direction

Générale des Finances Publiques (volumes 2004 et suivants). Outre les années manquantes, plusieurs années posent problème. Entre 1939 et 1945, il n’y a pas de données pour les trois départements occupés (Bas-Rhin, Haut-Rhin et Moselle). Les données ont été imputées de la façon suivante : pour chacun des trois départements et pour chacune des trois variables pertinentes, nous avons calculé le ratio entre la variable du département et celle des Vosges en 1938 et en 1946 et l’avons interpolé linéairement. En 1954, les données présentes dans les RSRID sont peu crédibles car les impôts collectés sont ridiculement faibles, ce qui est peut-être dû à une sous-évaluation volontaire décidée en réaction aux diverses révoltes fiscales ayant eu lieu à ce moment-là.

Les valeurs des années manquantes sont obtenues à partir d’un modèle économétrique simple ayant pour variable expliquée  $y_{it}$ , le revenu fiscal par adulte d’un département  $i$  à l’année  $t$  rapporté au revenu fiscal par adulte de l’ensemble des 90 départements au cours de la même année. La série de valeurs du numérateur provient de [Bonnet and Sotura \(2020\)](#) tandis que celle du dénominateur est construite à partir de [Garbinti et al. \(2019\)](#). Nous régressons le revenu fiscal par adulte d’un département par les variables fiscales collectées et par des variables démographiques permettant de tenir compte de l’évolution du revenu (et de l’impôt acquitté) au cours du cycle de vie.

Pour déterminer les revenus fiscaux des périodes 1922-1959 et 1970-1975, le modèle estimé s’écrit :

$$y_{it} = \sum_{a=1}^7 \alpha_a p_{ait} + \beta r_{it} + \gamma r_{it} s_t + \delta_i + \theta + \epsilon_{it} \quad (3.1)$$

où  $p_{ait}$  représente la part de la classe d’âge  $a$  dans la population du département  $i$  (les sept classes d’âge considérées sont 0-19, 20-29, 30-39, 40-59, 60-64, 65-79, 80+) rapportée à la part de la classe d’âge  $a$  dans la population des 90 départements,  $r_{it}$  est le revenu par adulte des foyers fiscaux imposables du département  $i$  rapporté au revenu imposable par adulte de l’ensemble des 90 départements,  $s_t$  est le ratio entre le total des revenus déclarés par les foyers fiscaux imposables et le revenu fiscal de l’ensemble des départements,  $\delta_i$  est l’effet fixe départemental,  $\theta$  est la constante, et  $\epsilon_{it}$  est le terme d’erreur. Nous interagissons le revenu

imposable par adulte avec un ratio calculé au niveau national pour tenir compte du fait que la part du revenu fiscal soumis à l'impôt sur le revenu change au cours du temps et qu'elle est susceptible d'affecter la valeur de  $\beta$ . Le modèle (3.1) est estimé pour différentes périodes en fonction de la date pour laquelle on souhaite déterminer les revenus fiscaux. Ainsi les revenus de la période 1922-1944 sont prédits à partir de l'équation estimée avec valeurs observées au cours de la période 1960-1969, tandis que ceux de 1945-1959 et de 1970-1975 le sont avec les valeurs de 1960-1969 et de 1986-1998.

Pour déterminer les revenus fiscaux des périodes 1978-1985 et 1999-2000, pour lesquelles les données sur le revenu imposable sont indisponibles, le modèle estimé s'écrit :

$$y_{it} = \sum_{a=1}^7 \alpha_a p_{ait} + \beta t_{it} + \gamma tu_{it} + \delta_i + \theta + \epsilon_{it} \quad (3.2)$$

où  $t_{it}$  représente la somme des impôts par adulte du département  $i$  rapporté à la somme des impôts par adulte de l'ensemble des 90 départements, et où  $tu_{it}$  est le nombre d'unités taxables par adulte du département  $i$  rapporté au nombre d'unités taxables par adulte de l'ensemble des 90 départements. Le modèle (3.2) est estimé à partir des valeurs de 1960-1969 et de 1986-1998 pour prédire les revenus manquants de 1978 à 1985 et à partir des valeurs de 1986-1998 pour prédire les revenus de 1999 et 2000. On remarque que les données du chapitre [Bonnet and Sotura \(2020\)](#) pour les années 2001-2017 ne sont pas utilisées pour prédire les années antérieures, du fait d'une rupture dans les données causée par un changement de la législation.

Le Tableau 3.1 renseigne pour chaque période considérée le  $R^2$  des estimations réalisées. Au total, quatre estimations ont été effectuées, et dans chaque cas, le modèle estime de façon quasi-exacte le ratio entre le revenu fiscal d'un département et le revenu des 90 départements.

Trois années n'ont pas été prédites par les modèles du fait de l'absence d'informations fiables dans les archives du Ministère des Finances : 1954, 1976 et 1977. Nous avons alors procédé à une interpolation linéaire du ratio entre le revenu fiscal par adulte de chaque

TABLE 3.1 –  $R^2$  DES DIFFÉRENTES ESTIMATIONS

Estimation	1	2	3	4
Données	1960-69	1960-69 ; 1986-98	1960-69 ; 1986-98	1986-98
Modèle	1	1	2	2
$R^2$	0,993	0,989	0,984	0,993

département et le revenu fiscal par adulte des 90 départements, avant de procéder à un recalage uniforme de la variable obtenue pour le numérateur afin que la somme obtenue pour les 90 départements corresponde aux valeurs données dans [Garbinti et al. \(2018\)](#). Le Tableau 3.2 résume pour chaque année de la période de référence la provenance de notre évaluation du revenu fiscal des 90 départements considérés.

## 2 Variables et indicateurs

La variable de référence de cet article est le revenu fiscal d'un département rapporté au nombre d'adultes du département. Comme nous disposons d'une série quasi-complète des impôts sur le revenu collecté dans chaque département, nous avons pu calculer une seconde variable, le revenu fiscal après impôt sur le revenu. Cette variable nous permet d'appréhender l'effet sur les inégalités spatiales de l'impôt sur le revenu. Il serait utile de compléter l'analyse en évaluant l'effet des autres impôts payés par les ménages même si leur effet est a priori plus faible. Le caractère progressif au niveau individuel de l'impôt sur le revenu implique qu'au total certains départements sont plus taxés proportionnellement que d'autres. Ce mode de calcul engendre une forme de redistribution spatiale. Comme nous n'avons pas d'information sur la façon dont cet impôt est dépensé et alloué entre les différents départements, nous ne pouvons néanmoins pas faire une analyse globale de son effet redistributif.

Nous avons également rapporté nos variables de revenu à la surface en km<sup>2</sup> de chaque département pour évaluer une sorte de rendement de la terre. Les différences entre les départements quant à ce rendement sont susceptibles de refléter une inégalité de concentration

TABLE 3.2 – DONNÉES UTILISÉES PAR PÉRIODE TEMPORELLE

Années	
1922-44	Prédit à partir de l'estimation 1
1945-53	Prédit à partir de l'estimation 2
1954	Interpolation
1955-59	Prédit à partir de l'estimation 2
1960-69	<a href="#">Bonnet and Sotura (2020)</a>
1970-75	Prédit à partir de l'estimation 2
1976-77	Interpolation
1978-85	Prédit à partir de l'estimation 3
1986-98	<a href="#">Bonnet and Sotura (2020)</a>
1999-2000	Prédit à partir de l'estimation 4
2001-15	<a href="#">Bonnet and Sotura (2020)</a>

de l'activité.

Nous utilisons plusieurs indicateurs d'inégalité. Nous débutons avec l'indicateur de Gini qui a l'avantage de prendre en compte toute la distribution et d'être indépendant de la moyenne. En complément, nous évaluons la part de différents quantiles dans la distribution tels que la part du premier décile (noté P90-100), ou des deux derniers déciles (P0-20). L'analyse de l'évolution du coefficient de Gini permet de comprendre s'il y a eu convergence des revenus entre les départements; cette « sigma-convergence » est plus robuste qu'une analyse régressant les taux de croissance sur les conditions initiales. L'analyse de l'évolution de la part de divers quantiles permet d'évaluer la déformation de la distribution.

Lorsqu'ils agrègent des revenus par adultes, nos indicateurs d'inégalité ne sont pas pondérés. Ceci permet de nous focaliser sur les différences entre les entités pertinentes, à notre sens pour ce travail, que sont les départements. Ceci permet en outre de ne pas supposer implicitement que les revenus sont également distribués au sein d'un département, ce qui serait une hypothèse très éloignée de la réalité ([Bonnet and Sotura \(2020\)](#)). En revanche, dans le cas des revenus

rapportés à la surface du département, nous pondérons nos observations par la superficie (qui est, évidemment, fixe dans le temps) afin de porter le regard sur l'évolution des écarts de concentration de revenus en France métropolitaine.

### 3 Résultats

Le graphique 3.1 représente l'évolution des inégalités interdépartementales de revenu fiscal par adulte. La tendance depuis un siècle est clairement à la baisse. L'indicateur de Gini était supérieur à 0,14 en début de période tandis qu'il est inférieur à 0,06 aujourd'hui. On constate deux périodes de baisse quasi continue : entre 1922 et 1939 et entre 1948 et 2015. Depuis 1948, la baisse est quasi-linéaire et le taux de croissance de l'indicateur est en moyenne de -1,3% par an. On constate néanmoins un ralentissement en fin de période, le taux de croissance annuel étant en moyenne de -0,3% depuis 2000.

La période couvrant la seconde guerre mondiale et les quelques années qui ont suivi est plus heurtée. Elle débute par une forte baisse, qui est liée à la désorganisation des départements urbains et à l'accroissement relatif du poids de l'agriculture dans l'économie française pendant la guerre. Cette baisse est compensée par une très forte reprise des inégalités entre 1944 et 1948, conduisant à un retour aux niveaux constatés au milieu des années 1920. Les années de guerre ont entraîné de forts mouvements de population entre les départements de la frontière est et le reste de la France [Bonnet \(2018a\)](#) ce qui a également conduit à une importante redistribution spatiale des revenus. En période de guerre, les données statistiques, tant sur les revenus que sur la population, sont néanmoins beaucoup plus fragiles ; ceci nous incite à les considérer avec davantage de distance et de concentrer nos analyses sur les périodes 1922-1939 et 1948-2015.

Le Graphique 3.1 représente également l'évolution des inégalités de revenu fiscal après impôt sur le revenu (IR). Ce dernier réduit significativement les inégalités interdépartementales, mais l'ampleur de la réduction a considérablement varié au cours de la période considérée. On peut définir trois périodes distinctes. Jusqu'au début des années 1950, la réduction des

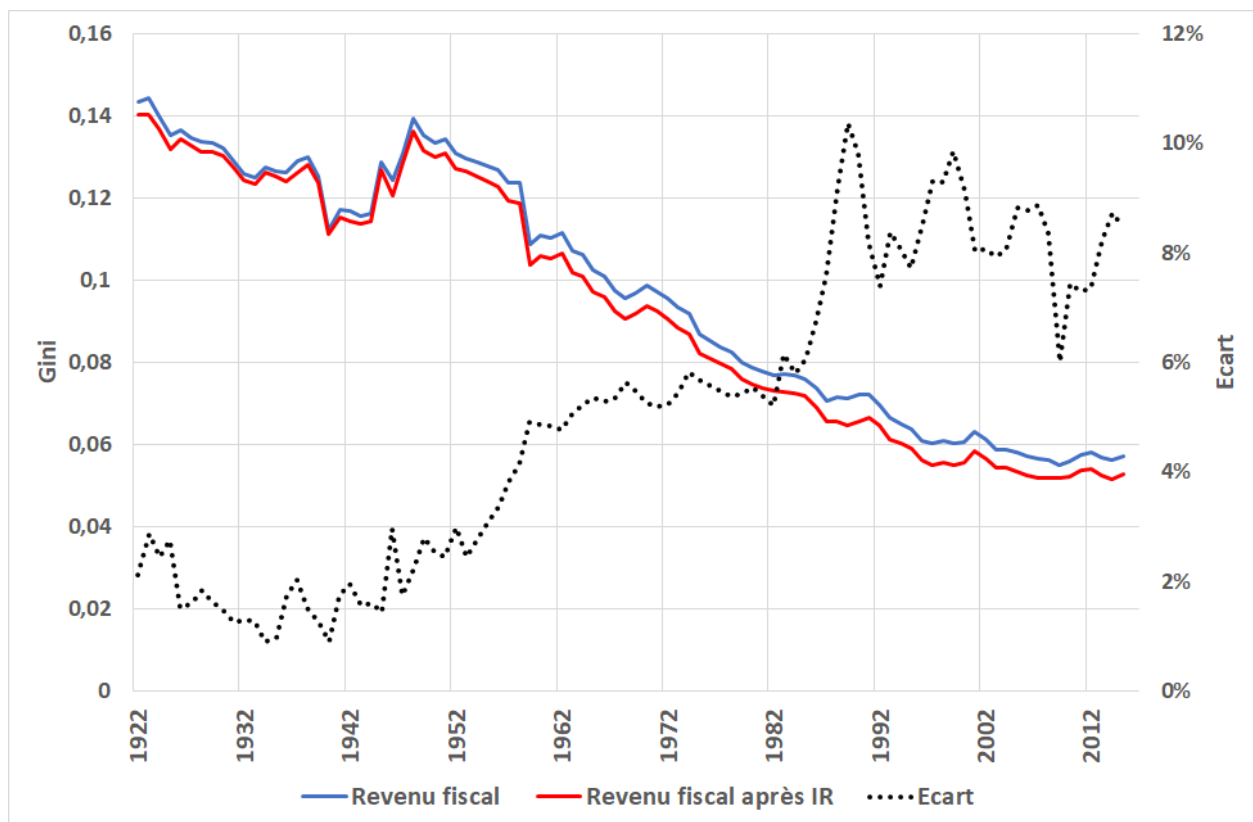


FIGURE 3.1 – COEFFICIENT DE GINI DU REVENU PAR ADULTE ENTRE 90 DÉPARTEMENTS, 1922-2015

inégalités interdépartementales engendrée par l'impôt sur le revenu est faible, et toujours inférieure à 3% des inégalités initiales. La réduction va ensuite progressivement prendre de l'ampleur jusqu'à la fin des années 1980. Elle atteindra un maximum en 1989 (cf. Graphique 3.12 en annexe). Depuis 1989, la baisse des inégalités interdépartementales engendrée par l'impôt oscille dans un tunnel allant de 6 à 9%, avec une légère tendance à la baisse.

L'évolution des inégalités interdépartementales de revenus est très différente de celle constatée pour des indicateurs d'activité économique tels que la valeur ajoutée mesurée à l'échelle des départements. Le Tableau 3.3 renseigne pour trois dates clefs, le Gini des revenus fiscaux par adulte après paiement de l'impôt sur le revenu tel que présenté dans le Graphique 1 et le Gini des valeurs ajoutées par adulte départementales que nous avons calculé à partir des données de Combes et al. (2011) et de l'INSEE pour notre classification des départements et, à des fins de robustesse, pour celle retenue dans Combes et al. (2011). Le Tableau 3.3

indique que le Gini des valeurs ajoutées a connu une évolution en forme de U : la stabilisation des deux dernières décennies du XX<sup>ème</sup> siècle a été suivie par une remontée des inégalités depuis 2000. Nous en déduisons que l'ensemble des transferts sociaux-fiscaux, qui représentent une part importante du revenu fiscal, contribuent fortement à l'atténuation des inégalités engendrées par la concentration des activités économiques. Aujourd'hui, le Gini des revenus est deux fois plus faible que le Gini des valeurs ajoutées.

L'évolution décrite par l'indicateur de Gini peut être complétée par les indicateurs spécifiques à certaines parties de la distribution. Le Graphique 3.2 représente l'évolution de la part de la somme des revenus moyens qui est détenue par les neuf départements les plus riches dans la somme des revenus moyens des 90 départements. Le profil temporel de l'indicateur est assez proche de celui de l'indicateur de Gini ; on constate néanmoins un net ralentissement de la baisse à partir des années 1970, et une stabilisation des inégalités depuis 2007. Le rôle de l'impôt sur le revenu a également évolué au cours de la période. Entre 1954 et 1988, il a réduit de façon croissante le P90-100, avec un maximum de 3,2% en 1989. Depuis, l'effet de l'impôt sur le revenu s'est considérablement réduit. Il est revenu en 2015 aux niveaux qui prévalaient à la fin des années 1970 (soit 1,9%).

Les départements les plus pauvres ont en revanche connu un fort processus de rattrapage. Le Graphique 3.3 montre que la part de la somme des revenus moyens détenue par les dix-huit départements les plus pauvres a continûment augmenté depuis la fin de la seconde guerre

TABLE 3.3 – INDICATEURS DE GINI, 90 DÉPARTEMENTS

	1930	1980	2000	2014
Revenus après IR	0,140	0,075	0,061	0,052
Valeurs ajoutées:				
Notre classification	0,121	0,103	0,097	0,111
Classification de <a href="#">Combes et al. (2011)</a>	0,118	0,104	0,098	0,113



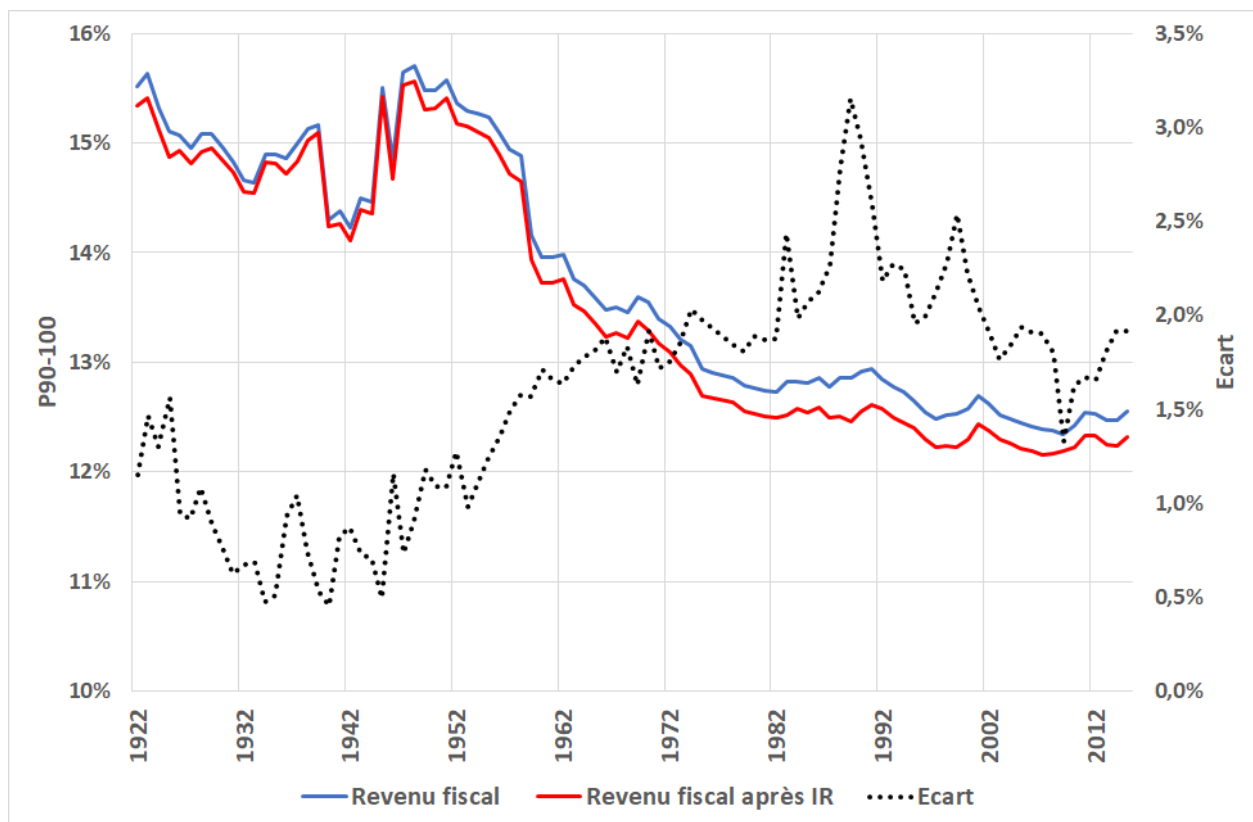


FIGURE 3.2 – P90-100 DU REVENU PAR ADULTE ENTRE 90 DÉPARTEMENTS, 1922-2015

mondiale. Comme pour P90-100, le niveau d'inégalité se stabilise en fin de période (à partir de 2003). De même, après avoir été haussière, la contribution de l'impôt sur le revenu à l'augmentation de la part des dix-huit départements les plus pauvres ne cesse de diminuer depuis la fin des années 1970.

Au total, les Graphiques 3.2 et 3.3 révèlent une situation qui est aujourd'hui beaucoup moins inégalitaire que dans le passé. On trouve que les 10% les plus riches gagnent 12,5% du total tandis que les 20% les plus pauvres gagnent près de 18% du total. Ces chiffres peuvent être également présentés de la façon suivante : les 10% les plus riches gagnent 25,5% de plus que ce qu'ils gagneraient si la distribution était égalitaire tandis que les 20% les plus pauvres gagnent 11,3% de moins que dans le cas égalitaire. Après paiement de l'impôt sur le revenu, ces pourcentages tombent à 23,1% et 10,6%.

Il est important de noter que l'absence de pondération des départements dans le calcul

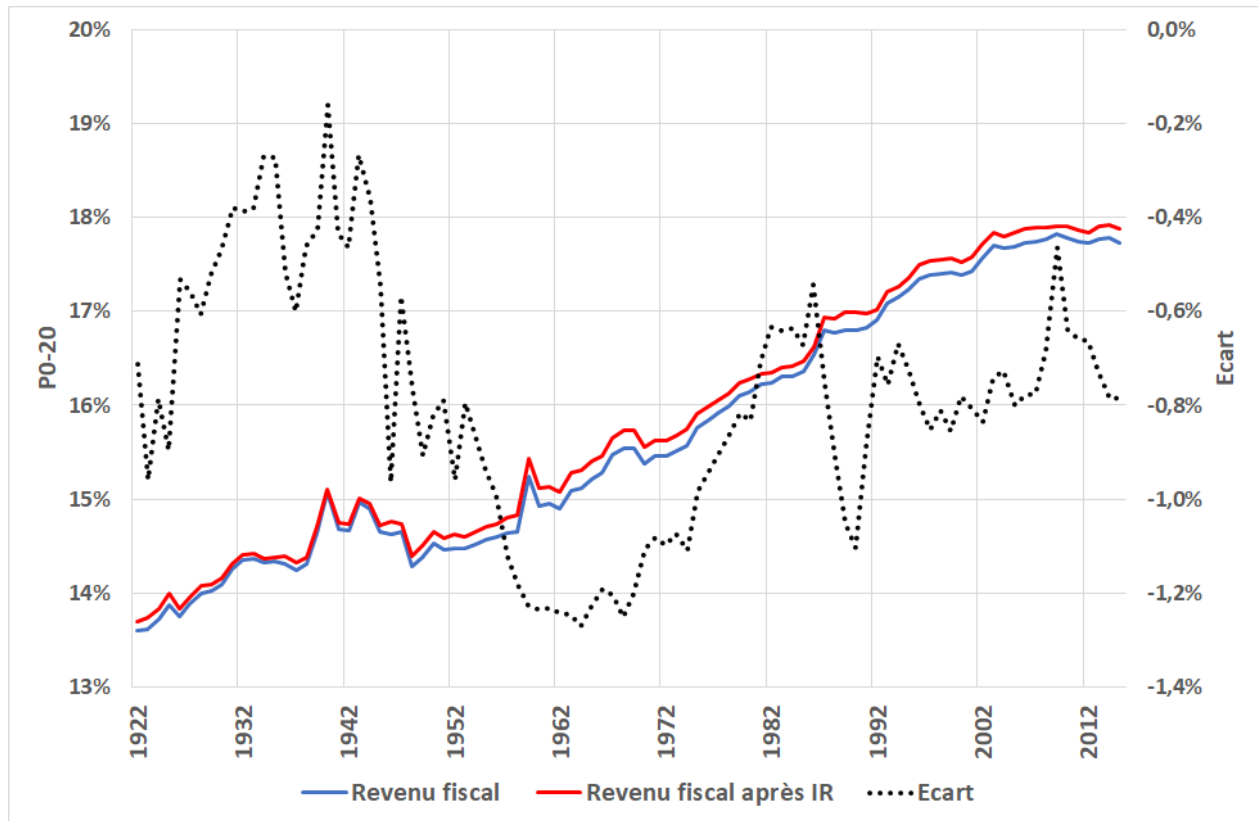


FIGURE 3.3 – P0-20 DU REVENU PAR ADULTE ENTRE 90 DÉPARTEMENTS, 1922-2015

des indicateurs d'inégalité est ici sans conséquence. Si l'on pondérait les départements par la population adulte, les parts des différents quantiles seraient relativement peu modifiées, avec un écart d'environ un point de pourcentage en fin de période, à la hausse pour P90-100 et à la baisse pour P0-20 (cf. les Graphiques 3.13 et 3.14 reportés en annexe). Cette pondération ignore néanmoins les inégalités infra-départementales, qui ont évolué au cours de la période (Bonnet and Sotura (2020)).

De façon équivalente, la non prise en compte des départements de l'actuelle région Ile-de-France, qui concentre aujourd'hui environ 30% du PIB français, ne modifie que peu l'évolution de la dynamique des inégalités. Le profil temporel de l'indicateur de Gini est similaire à celui représenté dans le Graphique 3.1, si ce n'est qu'il est translaté vers le bas (cf. Graphique 3.15 en Annexe). La relation à ces départements est néanmoins éclairante. Ainsi, une manière alternative de représenter la convergence des départements est de mesurer leur écart à celui

dont le revenu est le plus haut. La fin de période mise à part, c'est dans le département de la Seine que les revenus sont les plus élevés. Nous avons donc calculé, pour chaque année, la part des départements dont le revenu moyen était supérieur à 60% de celui de la Seine. Le Graphique 3.4 représente cette évolution. Au début des années 1950, moins de 10% des départements franchissaient le seuil tandis que depuis les années 2000, plus de 90% le franchissent. Si on considère le revenu après impôt, c'est dès le début des années 1990 que 90% des départements avaient un revenu supérieur à 60% de celui de la Seine.

Cette évolution s'explique par une baisse du revenu relatif de la Seine par rapport aux 89 autres départements. En 1950, le revenu par adulte du département était supérieur de 80% à la moyenne nationale, contre seulement 35% en 2015. Après paiement de l'impôt sur le revenu, cet écart était seulement de 27% à la même date. A titre de comparaison, la valeur ajoutée par adulte dans le département de la Seine était supérieure de 114% à la valeur ajoutée par adulte au niveau national en 2014.

Lorsque l'on rapporte les revenus des départements, non pas au nombre d'adultes, mais à la surface du département, on mesure alors les inégalités de concentration géographique du revenu entre les départements. Le graphique 3.5 représente l'évolution de l'indicateur de Gini au cours de la période considérée. Les niveaux d'inégalités apparaissent tout d'abord beaucoup plus importants. Par ailleurs, même si la forme générale des profils est similaire à partir des années d'après-guerre, on constate un très fort mouvement de concentration du revenu dans les années 1940 et 1950, qui a ensuite été effacé au cours des deux décennies qui ont suivi. En revanche, depuis la fin des années 1970, la concentration reste inchangée.

Les indicateurs précédents agrègent les dynamiques des différents départements et masquent leurs évolutions propres. Pour analyser la dispersion géographique des revenus en France métropolitaine, ainsi que sa recomposition depuis cent ans, nous avons identifié trois dates clef —1922, 1948 et 2015— et présentons, successivement, la géographie des écarts de revenus à ces dates et leurs évolutions entre ces dates.

Le Graphique 3.6 présente les rapports entre les revenus fiscaux par adulte de chaque département et la moyenne des revenus fiscaux par adulte de l'ensemble des départements

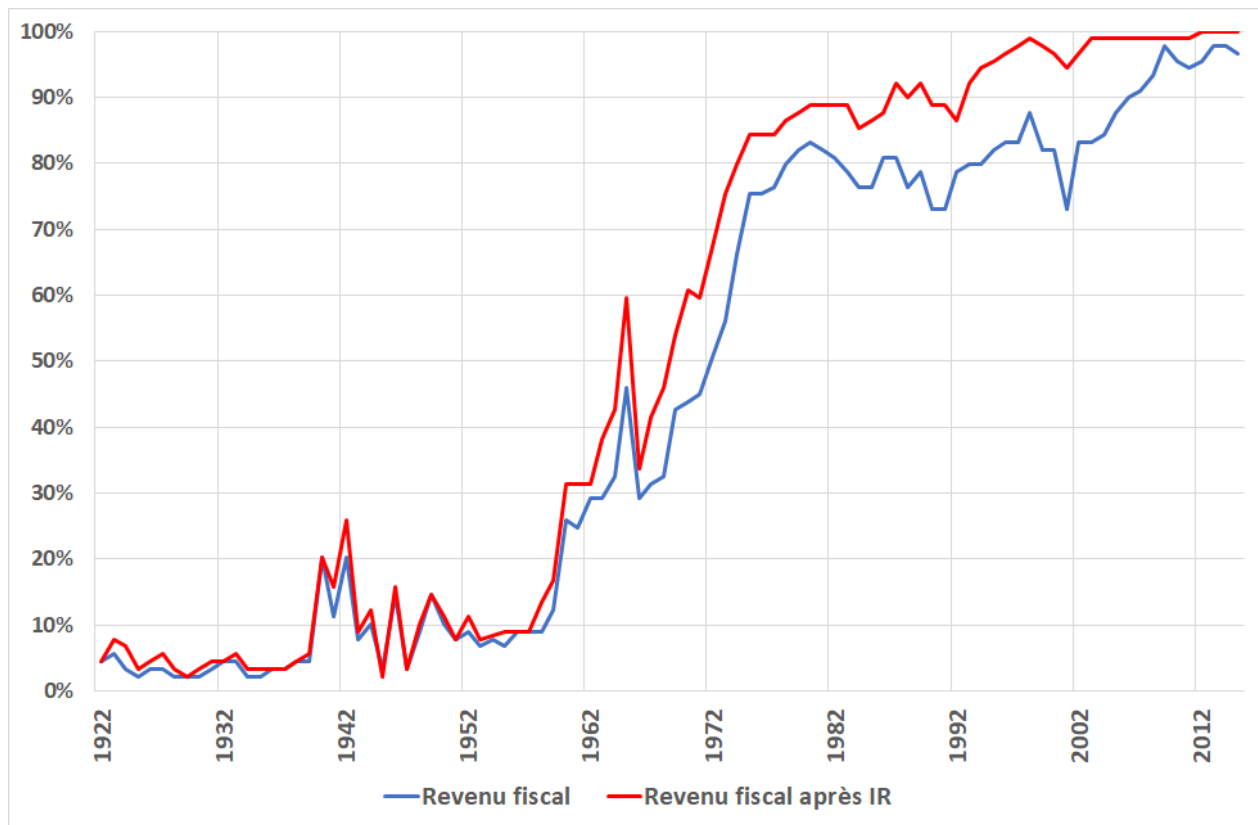


FIGURE 3.4 – PART DES DÉPARTEMENTS DONT LE REVENU MOYEN EST SUPÉRIEUR À 60% DU CELUI DE LA SEINE. 1922-2015

en 1922. On observe que le nord de la France était particulièrement riche : à l'exception du Pas-de-Calais, tous les départements profitaient d'un revenu par adulte supérieur à la moyenne nationale (en bleu sur le graphique), avec un maximum atteint dans les départements de la Seine et de la Seine-et-Oise (en bleu foncé). Les départements limitrophes (Eure, Eure-et-Loir, Loiret, Meuse, Haute-Marne et Côte-d'Or) présentaient un revenu par adulte proche de la moyenne nationale (en orange). Au sud de cette zone, la quasi-totalité des départements souffraient d'un revenu par adulte inférieur à 90 % du revenu par adulte moyen. Les zones géographiques où la faiblesse des revenus était la plus marquée (soit un revenu par adulte inférieur à 75% du revenu par adulte moyen ; identifiées en noir sur le graphique) se trouvaient en Bretagne, dans le Sud-Ouest ainsi que dans les Alpes du Sud et la Corse. Au sud, les exceptions significatives étaient le Rhône et les Bouches-du-Rhône, abritant de grandes métropoles régionales, ainsi que les Alpes-Maritimes.

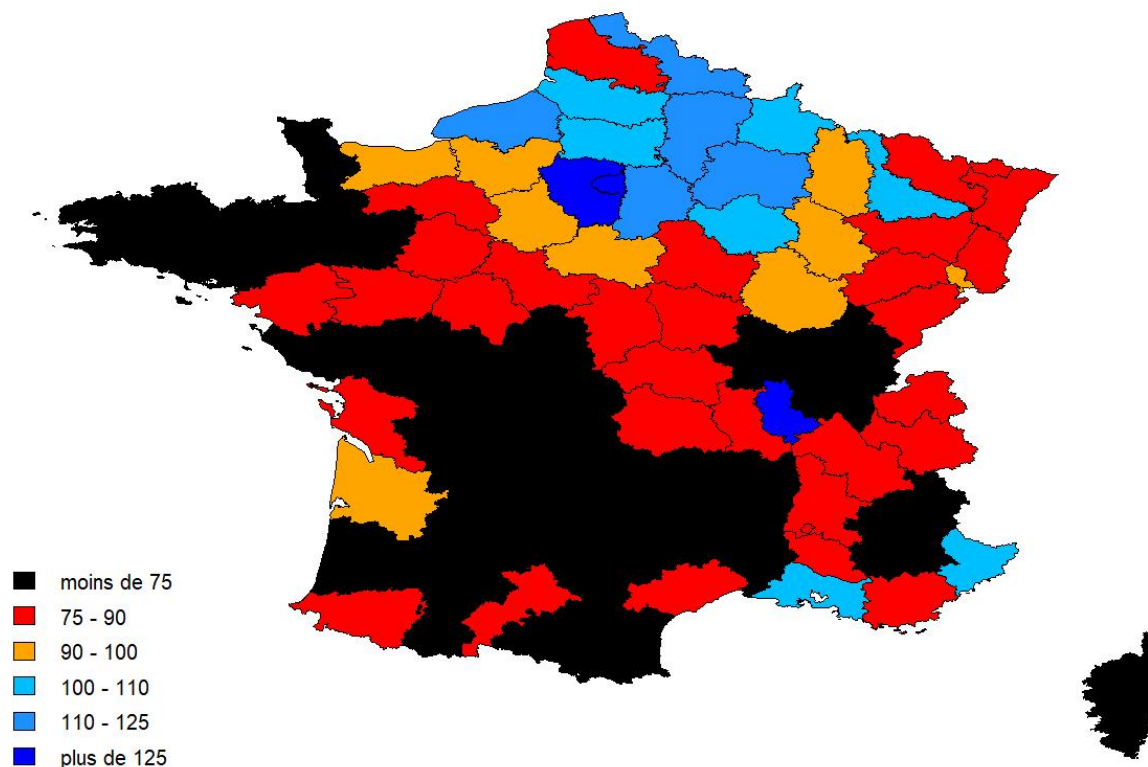


FIGURE 3.5 – COEFFICIENT DE GINI DU REVENU PAR KM<sup>2</sup> ENTRE 90 DÉPARTEMENTS, 1922-2015.

En 1948, la géographie des revenus en France n'a pas tellement changée (Graphique 3.7). Globalement, les départements avec un revenu par adulte supérieur à la moyenne nationale se trouvent toujours au nord de la France. Cette zone est néanmoins beaucoup moins homogène qu'en 1922 ; des départements tels que la Somme, l'Aisne, la Marne ou l'Aube souffraient d'un revenu par adulte inférieur à la moyenne nationale, tandis que des départements frontaliers de la Suisse tels que le Doubs ou le Haut-Rhin profitaient d'un revenu supérieur à la moyenne nationale. Au sud-ouest, tous les départements ou presque présentaient un revenu par adulte inférieur à la moyenne nationale, avec une vaste zone homogène s'étirant de la Bretagne jusqu'au sud des Cévennes accusant un fort retard en termes de revenus. Les Bouches-du-Rhône et le Rhône sont toujours des exceptions, et on note l'émergence de la Loire, qui abrite Saint-Etienne et ses industries.

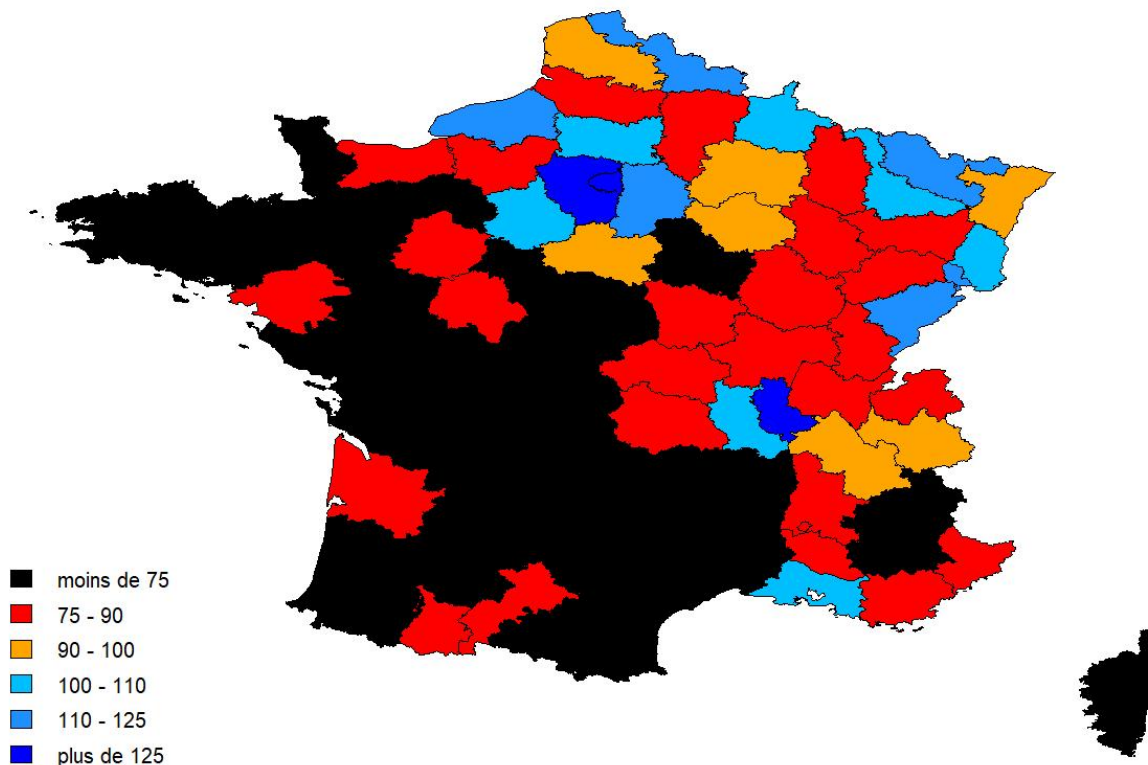


FIGURE 3.6 – RAPPORTS ENTRE LES REVENUS FISCAUX PAR ADULTE DE CHAQUE DÉPARTEMENT ET LA MOYENNE DES REVENUS FISCAUX PAR ADULTE DES 90 DÉPARTEMENTS EN 1922

Le Graphique 3.8 présente la carte en 2015. On constate, en premier lieu, la disparition des zones de revenus relativement faibles : il n'existe plus de départements pour lesquels le revenu par adulte est inférieur à 75% du revenu moyen, ce qui corrobore la baisse des inégalités documentée plus haut. Par ailleurs, il apparaît que les départements profitant d'un revenu par adulte supérieur à la moyenne nationale ne se trouvent plus dans le nord de la France, mais près de la frontière suisse, en région parisienne ainsi que dans les départements abritant certaines métropoles régionales telles que Toulouse et Nantes. On observe aussi que les départements dont le revenu par adulte est compris entre 75 et 90% du revenu moyen se trouvent le long d'une diagonale reliant la frontière espagnole et la frontière belge, avec deux excroissances dans le nord du pays et dans l'intérieur normand. Ces départements sont ceux de la « diagonale du vide ». A l'inverse, la façade atlantique est devenue une zone homogène

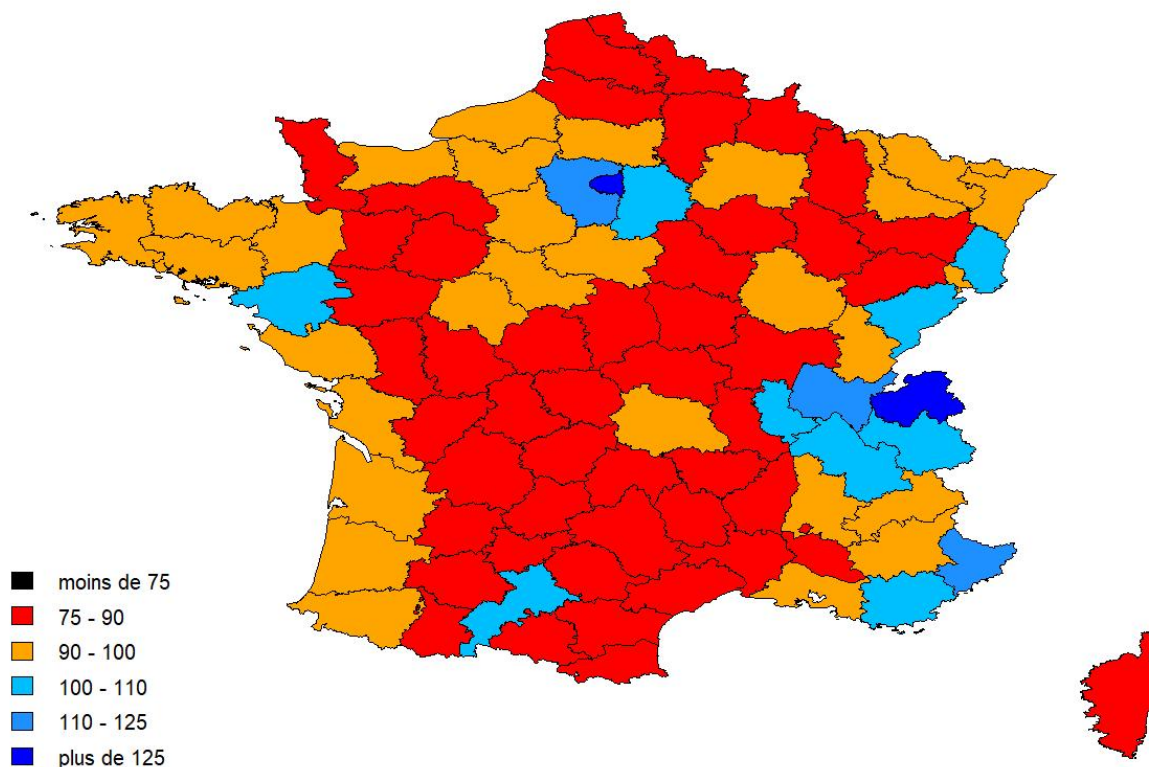


FIGURE 3.7 – RAPPORTS ENTRE LES REVENUS FISCAUX PAR ADULTE DE CHAQUE DÉPARTEMENT ET LA MOYENNE DES REVENUS FISCAUX PAR ADULTE DES 90 DÉPARTEMENTS EN 1948

où les niveaux de vie sont proches de ceux de la moyenne nationale. Les 3 cartes représentées dans les Graphiques 6, 7 et 8 montrent clairement le mouvement de bascule de la diagonale du vide, qui autrefois allait du nord-ouest au sud-est et qui aujourd'hui s'étend du nord-est au sud-ouest.

Pour représenter l'hétérogénéité de cette dynamique, nous avons classifié l'évolution des revenus fiscaux par adulte des départements en six catégories. Au cours du temps, certains départements ont connu une baisse du ratio entre leur revenu moyen et celui de l'ensemble des départements. Ce faisant, certains départements se sont rapprochés de la moyenne sans l'atteindre (cas i) ; un cas typique est le département de l'Oise, dont le revenu par adulte relatif était de 110% en 1922 et qui atteint 100% aujourd'hui. D'autres départements étaient déjà en dessous de la moyenne et s'en sont écartés (cas ii) ; ainsi, la Haute-Marne avait un



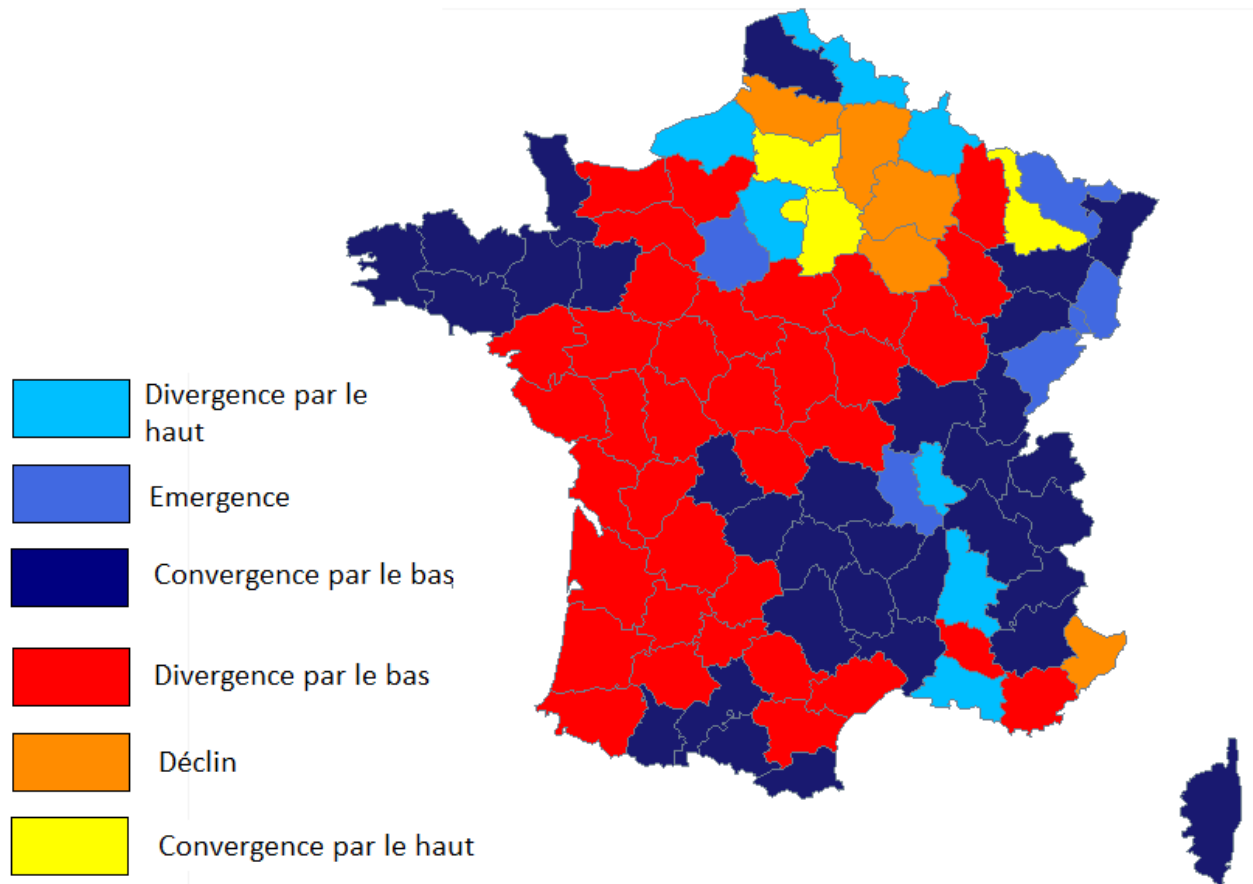


FIGURE 3.8 – RAPPORTS ENTRE LES REVENUS FISCAUX PAR ADULTE DE CHAQUE DÉPARTEMENT ET LA MOYENNE DES REVENUS FISCAUX PAR ADULTE DES 90 DÉPARTEMENTS EN 2015

revenu relatif de 90% en 1922 et de 82% en 2015. Enfin, certains départements sont passés au-dessous de la moyenne (cas iii) telle que la Meurthe-et-Moselle, qui a vu son revenu moyen relatif passer de plus de 110% à 90%. Nous qualifions ces départements avec les termes suivants : (i) convergence par le haut, (ii) divergence par le bas, (iii) déclin. Par symétrie, les départements qui ont connu une hausse du ratio entre leur revenu moyen et celui de l'ensemble des départements, peuvent être qualifiés ainsi : (i) convergence par le bas, l'exemple typique étant celui de la Lozère dont le revenu fiscal par adulte correspondait à 45% de la situation égalitaire en 1922 contre 83% aujourd'hui ; (ii) divergence par le haut, avec par exemple les Alpes-Maritimes dont le revenu relatif est passé de 101% en 1922 à 110% en 2015 ; (iii)



émergence, telle la Haute-Savoie, dont le revenu par adulte relatif est passé de 75% à 130%. Cette classification est similaire à celle employée dans la littérature étudiant la dynamique des populations locales (cf. [Oswalt and Rieniets \(2006\)](#), et la littérature qui a suivi).

Le Graphique 3.9 représente les départements ainsi catégorisés entre 1922 et 1948 tandis que le Graphique 3.10 les présente entre 1948 et 2015. La première période permet de comparer les départements sur un quart de siècle à l'issue des deux guerres mondiales. La baisse de la situation relative concerne la plupart des départements de la façade atlantique et de la basse vallée de la Loire ; elle est moins marquée, mais tout de même présente pour l'Île-de-France, tandis qu'un arc nord-est de départements entourant Paris subit un déclin.

Entre 1948 et 2015, les 90 départements que nous considérons ont globalement connu un processus continu de convergence. La carte est cette fois radicalement différente. Sous une ligne allant du Calvados au Gard, tous les départements ont connu une amélioration de leur situation relative. Sous cette ligne « Caen-Nîmes », les départements de Toulouse et Nantes émergent particulièrement. A l'inverse les départements des trois grandes métropoles du « Paris-Lyon-Marseille » ont vu leur situation relative se détériorer. Pour Lyon et Marseille, c'est plutôt la situation des départements qui leur sont situés à l'Est qui s'améliore. Outre les trois grandes métropoles, la dégradation concerne essentiellement les départements situés au nord-est d'une ligne allant du Calvados au Jura, à l'exception notable des départements alsaciens.

Catégorisation des départements selon leur évolution relative entre 1922 et 1948

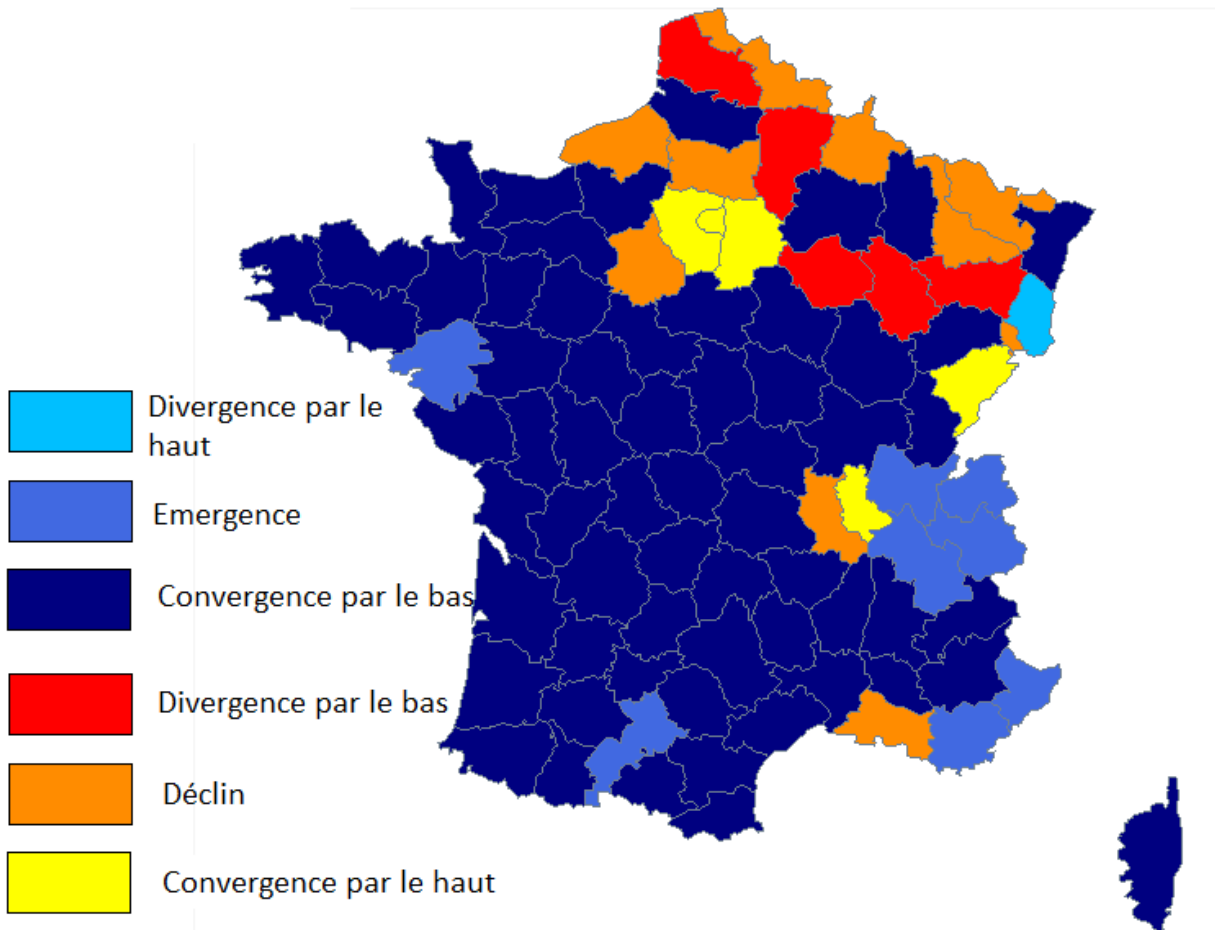


FIGURE 3.9 – CATÉGORISATION DÉPARTEMENTS SELON LEUR ÉVOLUTION RELATIVE ENTRE 1922 ET 1948

Catégorisation des départements selon leur évolution relative entre 1948 et 2015

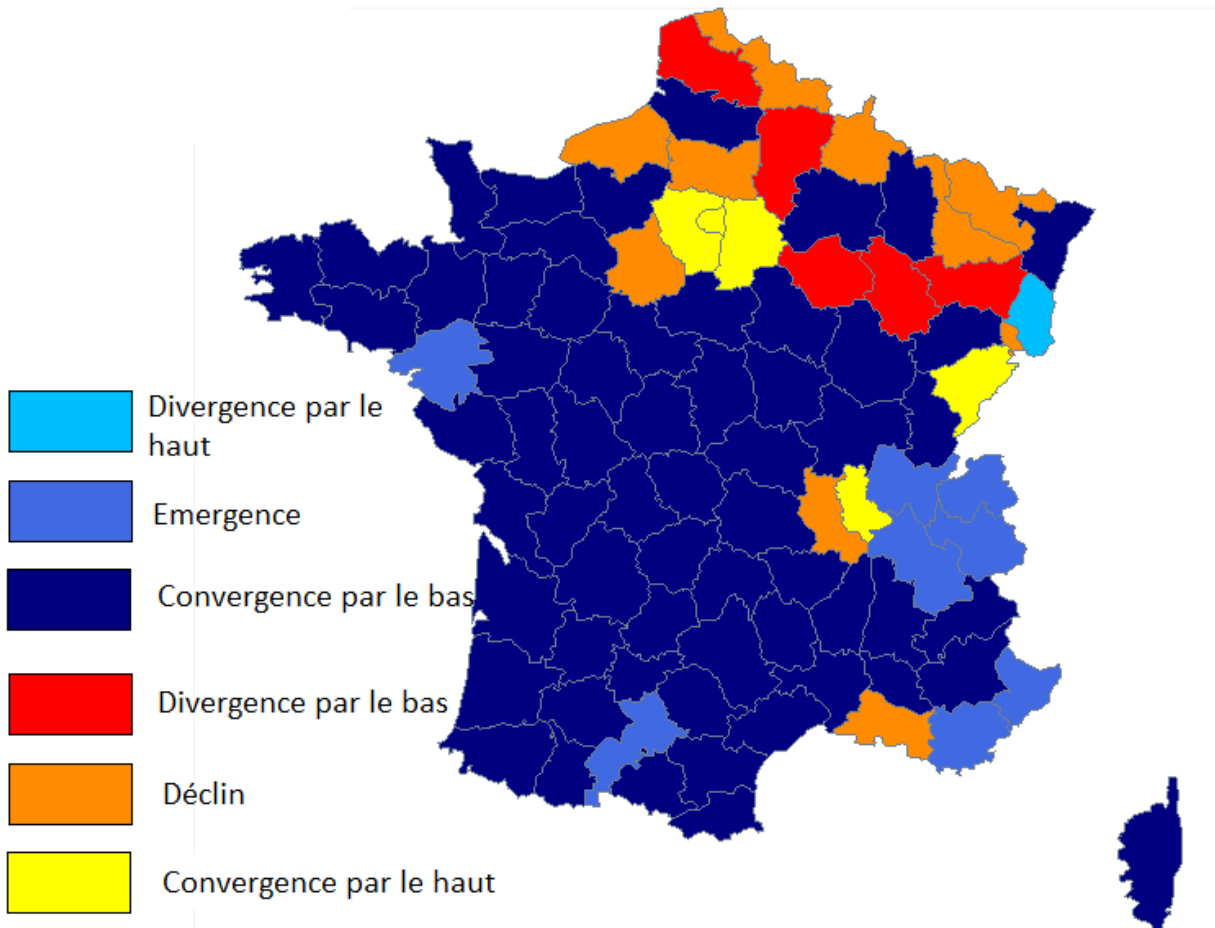


FIGURE 3.10 – CATÉGORISATION DÉPARTEMENTS SELON LEUR ÉVOLUTION RELATIVE ENTRE 1948 ET 2015

## 4 Conclusion

Cet article présente une base de données inédite sur les revenus moyens dans chaque département de la France métropolitaine depuis 1922, développée à partir de la numérisation des archives du Ministère des Finances. Plusieurs indicateurs d'inégalités entre les départements ont été construits et ont révélé une très forte convergence des revenus au cours de la période considérée. La diminution des inégalités est particulièrement frappante depuis 1948 et, même si l'on relève une réduction du rythme de décroissance depuis les années 2000, les inégalités ont atteint en 2015 leur plus faible niveau depuis cent ans. Aujourd'hui, tous les départements de métropole ont un revenu fiscal moyen après impôt sur le revenu supérieur à 60% de celui de la Seine – pour rappel, il s'agit de Paris et sa petite couronne. En 1950, seuls 10% d'entre eux franchissaient un tel seuil. Cette convergence départementale est similaire au processus analysé par [Bonnet and d'Albis \(2018\)](#) pour les espérances de vie mais contraste avec le processus décrit par [Combes et al. \(2011\)](#) pour les valeurs ajoutées. Ceci éclaire le rôle joué par les transferts publics dans l'égalisation des niveaux des vie, qui vient plus que compenser la force de divergence créée par la concentration des activités économiques sur certaines zones du territoire, et en particulier dans les grandes métropoles. Le rôle de l'impôt sur le revenu est à cet égard significatif. Il réduit significativement les inégalités entre les territoires : les neuf départements les plus riches ont un revenu fiscal par adulte supérieur de 25,5% au revenu fiscal par adulte national ; cet avantage relatif tombe à 23,1% après paiement de l'impôt sur le revenu. Dans la continuité de cet article, il serait intéressant d'évaluer la contribution des autres transferts sociaux-fiscaux, et en particulier du système de retraite. Le processus de convergence a bénéficié à de nombreux départements qui ont vu leur situation relative s'améliorer, mais il ne doit pas cacher la détérioration de la situation relative d'autres départements. Très schématiquement, la France est coupée en deux par une diagonale allant du Calvados au Gard : depuis la seconde guerre mondiale, les « gagnants » sont souvent au sud-ouest de cette ligne, tandis que les « perdants » se trouvent souvent au nord-est. Le déclin

vécu par certains est potentiellement susceptible d'engendrer un sentiment de mal-être dans les populations et de créer une rupture dans la cohésion nationale, et ce malgré la convergence des revenus. Là encore, davantage de données collectées au niveau local sont nécessaires pour emprunter cette voie de recherche prometteuse.

Notre travail peut être complété dans deux directions. La première est de considérer les inégalités infra-départementales et de décomposer l'évolution des inégalités nationales entre les inégalités intra-départementales et les inégalités inter-départementales. La seconde consiste à comparer l'évolution des inégalités territoriales entre les pays. Il n'existe malheureusement à ce jour pratiquement pas de bases de données équivalentes à la nôtre. L'exception concerne les 51 Etats américains dont les revenus ont été reconstitués depuis 1917 par [Frank \(2015\)](#). Le Graphique [3.16](#), reporté en annexe, compare l'évolution de notre coefficient de Gini pour les revenus fiscaux avec celui que nous avons calculé à partir des données de [Frank \(2015\)](#). La diminution des inégalités a débuté beaucoup plus tôt aux Etats-Unis, vers 1933, mais elle s'est achevée au milieu des années 1990 et a considérablement augmenté depuis. Les inégalités entre les états américains ont retrouvé aujourd'hui le niveau qui prévalait au début des années 1960. Les inégalités y sont, par ailleurs, considérablement plus élevées car le coefficient de Gini était égal à 0,11 en 2014. Il serait pertinent et intéressant de poursuivre cette comparaison avec les pays européens.



# Appendix

Le graphique 3.12 représente l'évolution du taux d'imposition moyen national (soit le rapport entre le montant de l'impôt sur le revenu et celui du revenu fiscal, calculés pour l'ensemble des 90 départements), le taux d'imposition minimum (le taux d'imposition moyen du département qui a le plus faible taux d'imposition), le taux d'imposition maximum et l'écart entre le taux maximum et minimum. Tout comme le taux d'imposition moyen, l'écart entre les taux minimum et maximum atteint son point haut en 1989 : il est alors de 11,5 points de pourcentage alors qu'il était égal à 3,7 points de pourcentage en 1922 et qu'il est égal à 8,3 points de pourcentage en 2015.



FIGURE 3.11 – CARTE ET LISTE DES 90 DÉPARTEMENTS FRANÇAIS UTILISÉS DANS L'ARTICLE



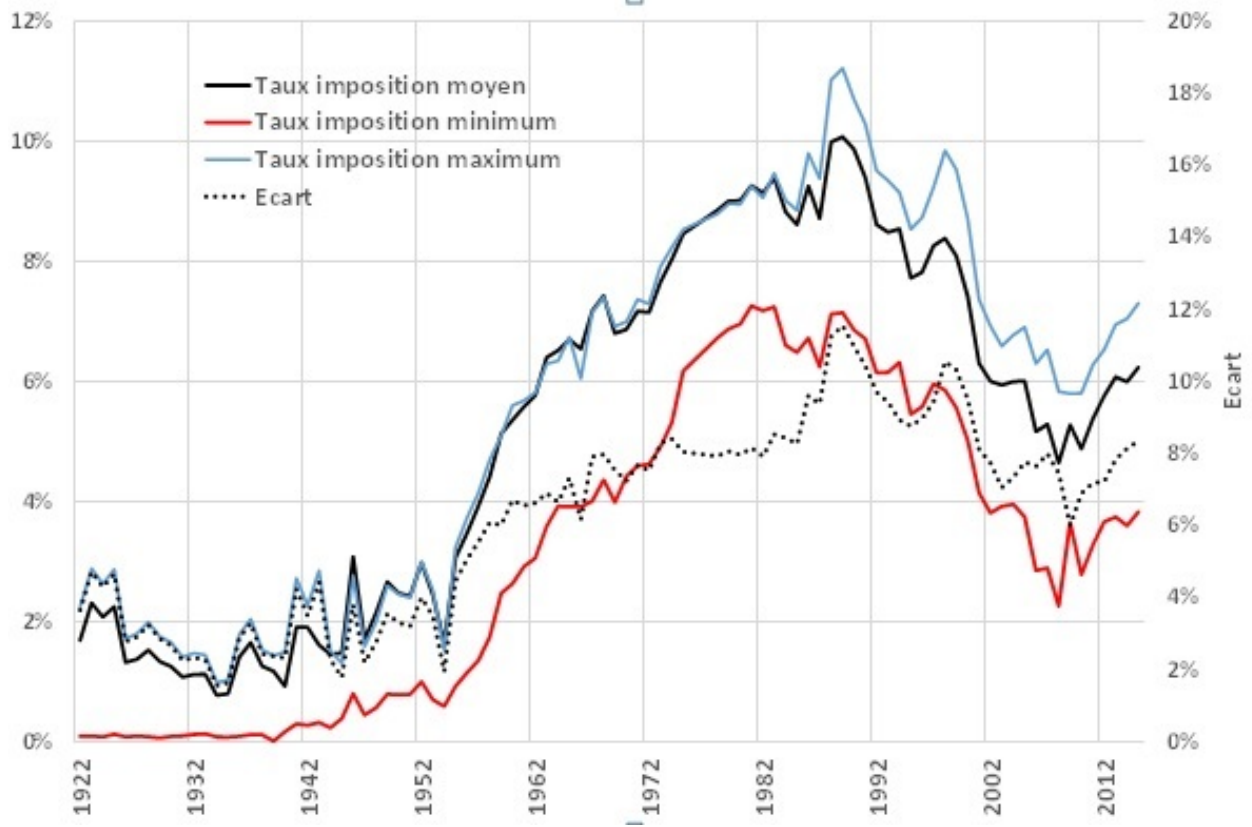


FIGURE 3.12 – TAUX D'IMPOSITION MOYEN, MINIMUM ET MAXIMUM, 1922-2015

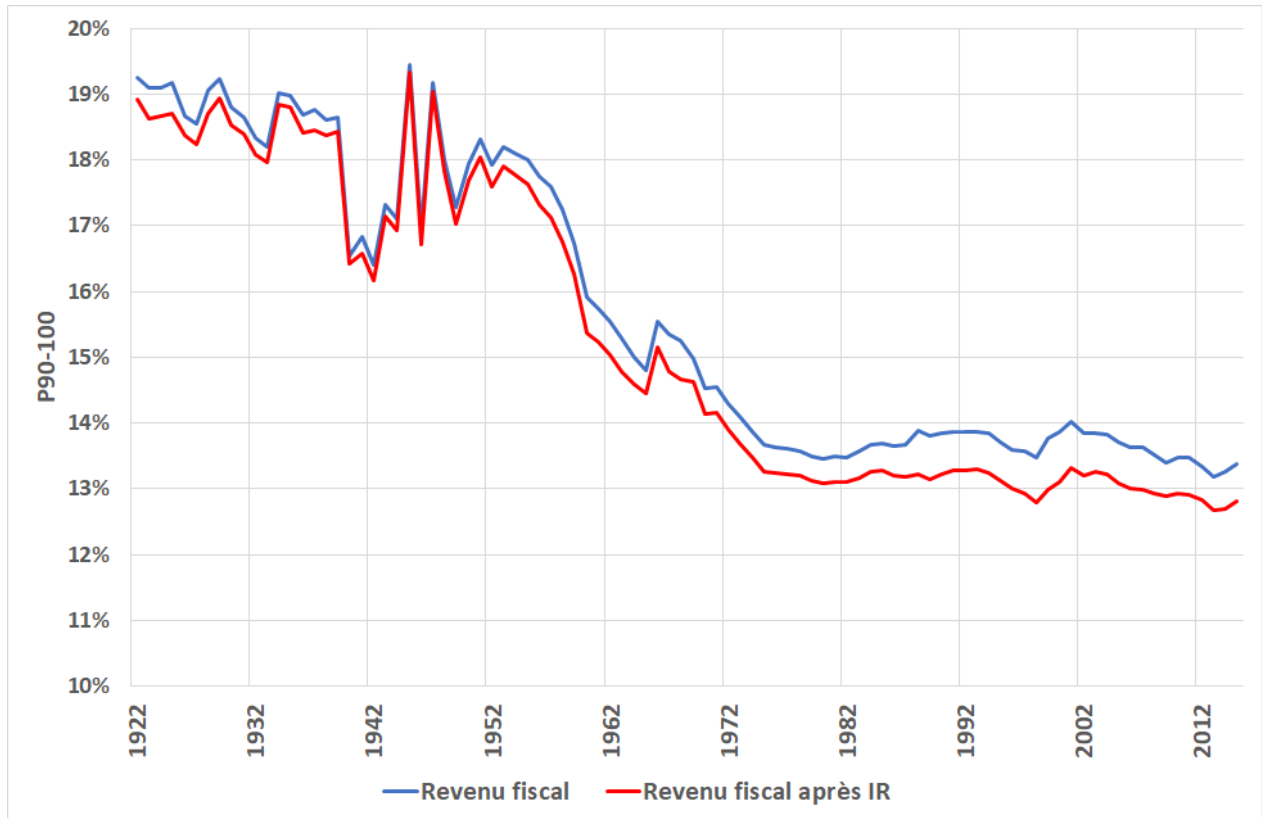


FIGURE 3.13 – P90-100 DU REVENU PAR ADULTE, 90 DÉPARTEMENTS PONDÉRÉS PAR LEUR POPULATION ADULTE, 1922-2015

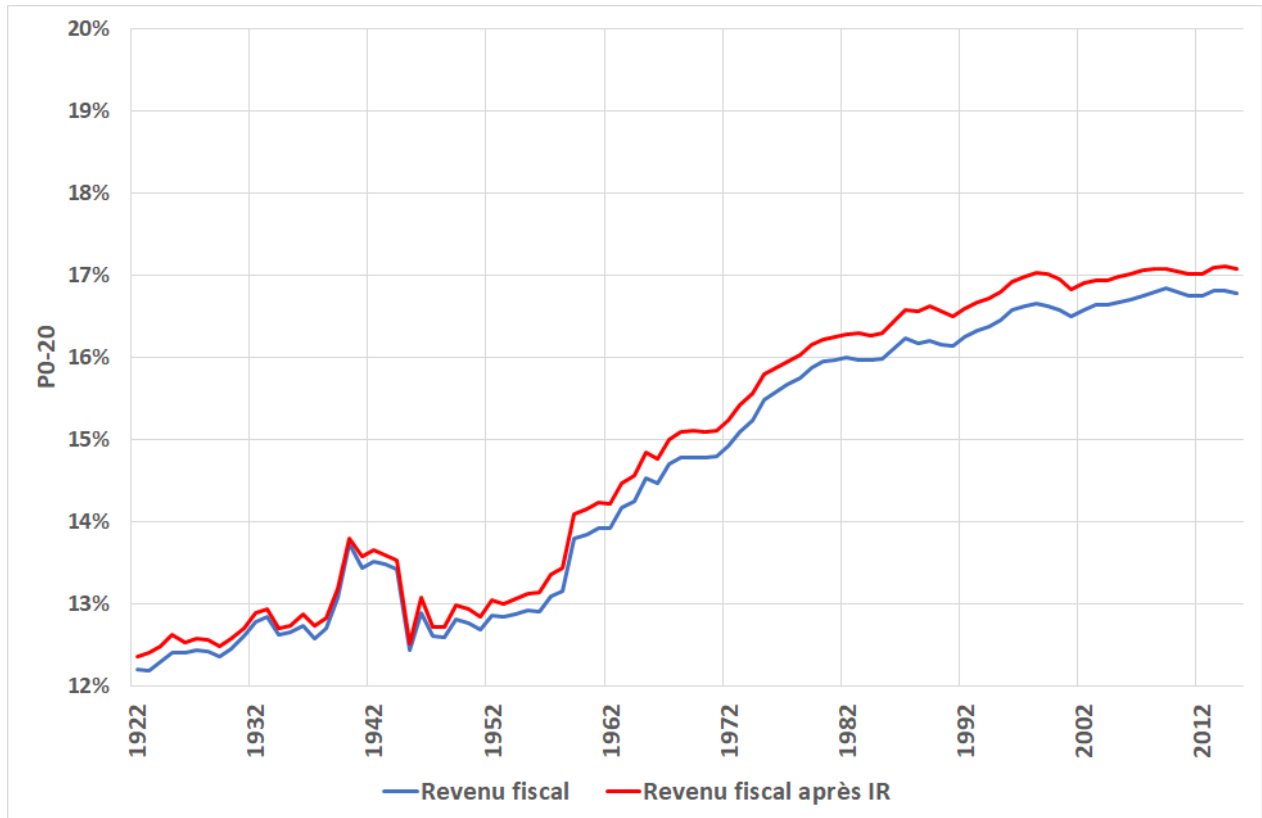


FIGURE 3.14 – P0-20 DU REVENU PAR ADULTE, 90 DÉPARTEMENTS PONDÉRÉS PAR LEUR POPULATION ADULTE, 1922-2015

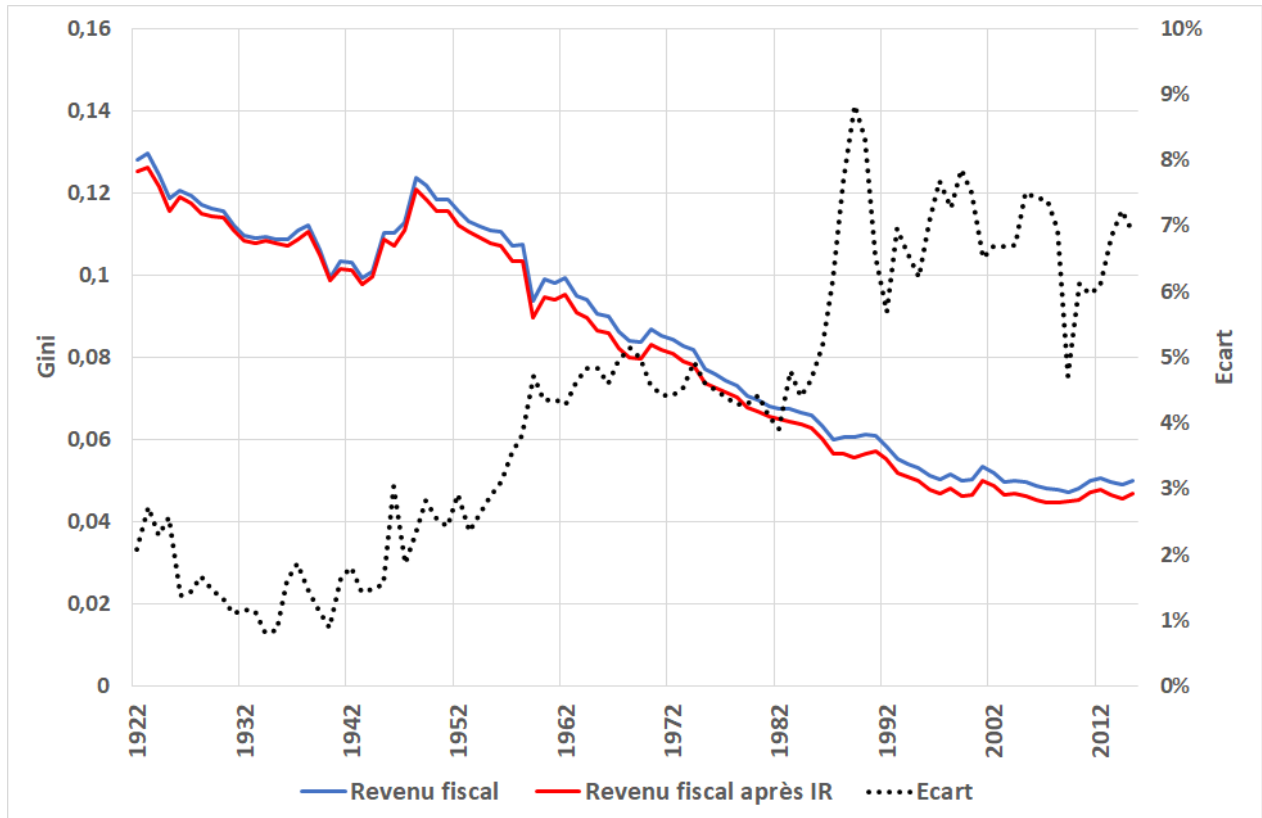


FIGURE 3.15 – COEFFICIENT DE GINI DU REVENU PAR ADULTE ENTRE 87 DÉPARTEMENTS HORS CEUX D'ILE-DE-FRANCE (SEINE, SEINE-ET-MARNE ET SEINE-ET-OISE), 1922-2015

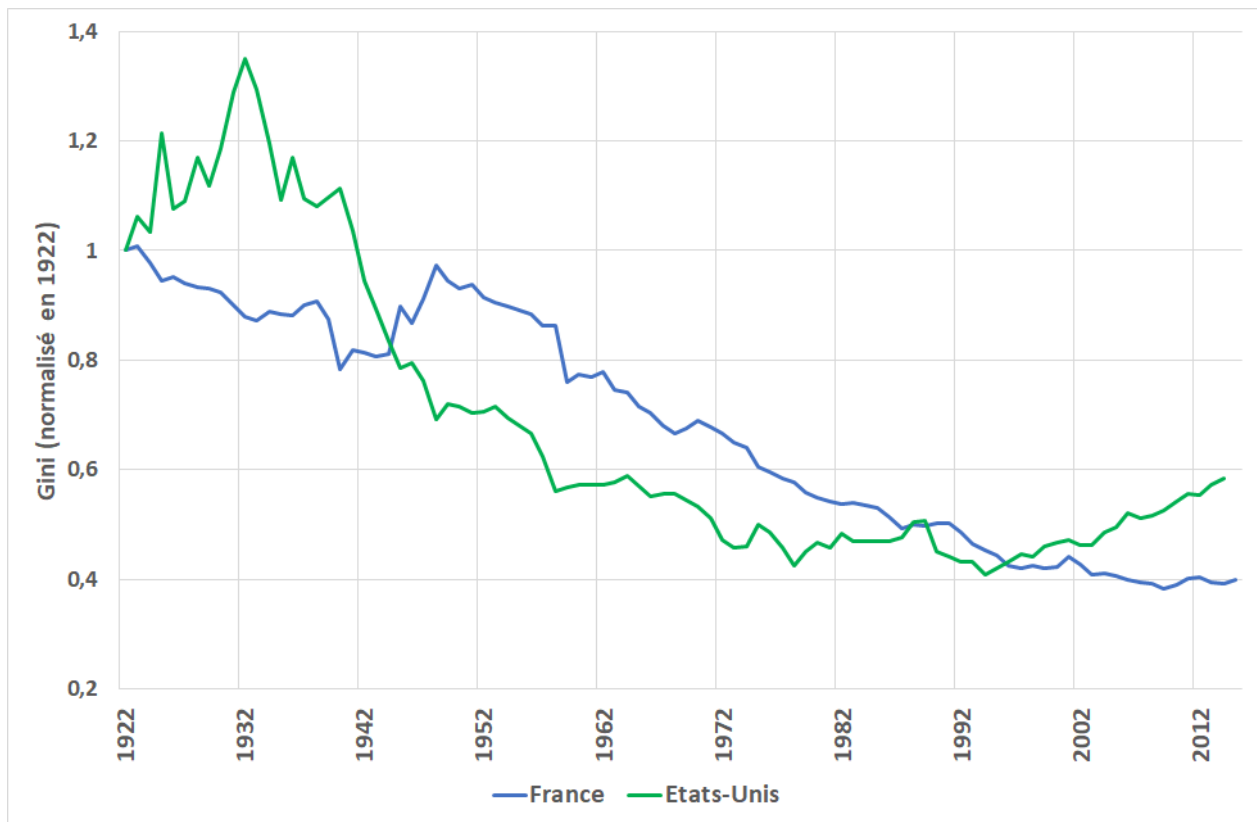


FIGURE 3.16 – COEFFICIENT DE GINI DU REVENU FISCAL PAR ADULTE EN FRANCE ET AUX ETATS-UNIS. 1922=1



# Chapter 4

## This Town Ain't Big Enough?

### Quantifying Local Public Goods

### Spillovers <sup>1</sup>

#### Abstract

Despite long-standing theoretical interest, empirical attempts at investigating the appropriate level of decentralization remain scarce. This paper develops a simple and flexible framework to test for the presence of public good spillovers between fiscally autonomous jurisdictions and to investigate potential welfare gains from marginal fiscal integration. We build a quantitative spatial equilibrium model with many local jurisdictions, mobile households and endogenous local public goods causing spillovers across jurisdictional boundaries. We show how one can exploit migration and house price responses to shocks in local public goods at different geographic scales to reveal the intensity of spillovers. Applying our framework to the particularly fragmented French institutional setting, we structurally estimate the model using a unique combination of jurisdiction-level administrative panel datasets. Estimation

---

1. This chapter is a joint work with Nicolas Jannin. We thank Antoine Bozio, Jan K. Brueckner, Pierre-Philippe Combes, Gabrielle Fack, Stéphane Gauthier, Laurent Gobillon, Stéphane Guibaud, Philipp Ketz, Thomas Piketty, Emmanuel Saez, Daniel Sturm, Etienne Wasmer and seminar participants at Amsterdam 2019 UEA European Meeting, Philadelphia 2019 UEA Meeting, UC Berkeley, Panthéon-Assas CRED and PSE for helpful comments and discussions. All remaining errors are our own.

THIS TOWN AIN'T BIG ENOUGH?  
QUANTIFYING LOCAL PUBLIC GOODS SPILLOVERS

---

relies on plausibly exogenous variations in government subsidies to instrument changes in the supply of local public goods. We find that public spending in a municipality accounts for 4–11% of the local public good bundle enjoyed by its residents, and that public spending in each neighbor municipality accounts for an average 3.2–3.5% of this bundle. Finally, we simulate the effect of a reform increasing fiscal integration and find substantial welfare gains.

**JEL Classification:** D16, H41, H71, H72, H73, R13

**Keywords:** Local Public Service, Spillover Effect, Spatial General Equilibrium, Tiebout, Welfare Economics, State Government Subsidies.



## 1 Introduction

Take an economy divided into geographically distinct jurisdictions. Who should be providing local public services? Local governments or the central government? If local jurisdictions are tasked with providing public goods, what should be their boundaries? Since the seminal works of [Tiebout \(1956\)](#) and [Oates \(1972\)](#), academics have investigated the optimal balance of power between local and more centralized forms of government. Simply put, the political economy of centralized decision-making misallocates local public services. On the other hand, decentralization may be inefficient because of spatial spillovers, i.e., the extent to which a jurisdiction's local public goods also benefit its neighbors.<sup>2</sup> When spillovers are strong, local jurisdictions may under-provide local public services as they do not internalize their benefits to neighboring jurisdictions. In addition, jurisdictions may actively free-ride on neighboring jurisdictions' public goods, worsening the under-provision problem.

In practice, there seems to be no consensus about the optimal size and autonomy of local jurisdictions. There remains substantial variation in institutional geography across western countries. Average town population is 4,300 in the EU compared to 16,000 in the US. Within the EU itself, there are large disparities between otherwise comparable countries. Average town density is similar in Germany (179 inh. per  $km^2$ ), France (154 inh. per  $km^2$ ) and Spain (177 inh. per  $km^2$ ). However, with mean jurisdiction population of respectively 7,100 and 5,800, German and Spanish towns are much larger than the average French town only home to 1,753 inhabitants.

Depending on the magnitude of decentralization and centralization inefficiency costs, jurisdictional fragmentation may have important welfare consequences. Policy and institutional solutions to remedy spillover inefficiencies typically include Pigovian subsidies and boundary redefinition. While the former solution requires the tailoring of subsidies to spillover intensity, one simply needs knowledge of the presence of spillovers between jurisdictions of designated

---

2. Decentralization may be inefficient for other reasons, e.g., tax competition between local governments (see for example [Brueckner 2000](#) or [Brueckner 2003](#) and [Boadway and Tremblay 2012](#) for reviews). In this paper, we focus only on local public good spillovers in the tradition of the early literature on decentralization. Our estimates are however robust to the existence of tax competition, as explained in Section 5.

geographic areas to implement the latter. However, while a rich theoretical literature studies efficient fiscal federalism (see [Oates 2005](#) for a review), there are few empirical attempts at structurally quantifying local public good spillovers. Indeed, there are empirical and theoretical challenges to disentangling public good spillovers from other general equilibrium mechanisms. For example, spillovers may worsen migration externalities when public goods are subject to agglomeration or scale economies.<sup>3</sup>

Our paper revisits the local public good provision debate in a quantitative spatial equilibrium model. Our contribution is twofold. First, we develop a simple framework to test for potential welfare gains from arbitrary increases in the level of centralization of local public goods provision. We ground it on a location choice model borrowed from the urban economics literature that allows us to isolate public good spillovers from other mechanisms at play in equilibrium. The key ingredient is the nesting of fiscally autonomous jurisdictions in geographic areas within which one suspects there are cross-border spillovers. By studying migration and housing price responses to local public good shocks first between jurisdictions within nests, then between such nests, one can reveal the presence of spillovers. Because these nesting areas can be made to encompass an arbitrary number of jurisdictions, our framework allows to repeatedly test for the presence of spatial spillovers until finding the jurisdictional fragmentation that cancels them. To the best of our knowledge, this paper is the first attempt at taking a structural approach to fiscal decentralization. Second, we apply our framework to French data and provide new estimates for structural parameters that are key to the local public finance debate. In the much fragmented French context, we estimate strong spatial spillovers and public good scale economies.

Our approach has two limitations. First, assessing the inefficiency cost of centralization is beyond the scope of the present paper. These costs typically stem from the interaction between centralized political frictions that tend to create winners and losers, and heterogeneous local needs.<sup>4</sup> As such, we abstract from the taste heterogeneity motivating the standard

---

3. When public goods are not fully rival, the costs of providing residents with a given level of public good benefits increase less than one for one with population. Denser jurisdictions typically provide more public goods for less taxes, which creates an agglomeration force.

4. See for example [Carbonnier et al. \(2008\)](#) for an attempt at assessing both the costs of centralization

Tiebout literature and do not account for interactions between local and central governments. Second, structural parameters may themselves be endogenous to local public goods, e.g., to the development of regional transit.

We first develop a spatial equilibrium model that draws on the seminal framework of [Rosen \(1974\)](#) and [Roback \(1982\)](#) and allows for endogenous wages, rents and local public good provision. We let households be potentially infra-marginal in their migration choices by introducing heterogeneous preferences for jurisdictions. Our agents are otherwise homogeneous in skills and have identical preferences for public goods over private consumption. They vote for local taxes and a level of public goods that will in turn affect equilibrium demand for jurisdictions. Most importantly, we allow for cross-border public good spillovers in a simple flexible structure that keeps the model amenable to reduced-form empirical analysis. Our model pinpoints key structural parameters related to spillovers, local public good agglomeration economies, preference for public services, household mobility and housing supply elasticity.

Home to around 35,000 autonomous local jurisdictions accounting for 38% of EU's total, France is a natural context in which to apply our framework. We first provide difference-in-differences (DiD) evidence on the impact of public good supply shocks on migration, housing consumption and housing prices using comprehensive administrative datasets on French municipalities. We combine data covering local taxes and public spending, population, housing consumption, wages and housing prices from 2000 to 2016. Our identification strategy exploits plausibly exogenous variation in investment-targeted subsidies to instrument changes in local public goods. Overall, we find significant migration responses and housing price capitalization which we interpret as evidence that households are mobile, enjoy local public services and that housing supply is not inelastic in the medium run. However, we find that migration responses to public good shocks within municipal federations — which are administrative groups of close-by municipalities — are smaller than when comparing these groups between them. We also estimate significant housing price capitalization in the latter case and not in the former. In line with our theoretical framework, a candidate mechanism to explain these reduced-form

---

and decentralization.

results is the presence of spatial spillovers, rendering location decisions less relevant for the enjoyment of publicly provided amenities within groups of close-knit jurisdictions.

We take our theoretical framework to the data and use a Generalized Method of Moments (GMM) to estimate the model's parameters. We express moments conditions in terms of the model's structural residuals and investment subsidy shocks. We combine moment conditions related to demand for municipalities and average demand for municipal federations. Importantly, the model is identified using the same subsidy shocks that we exploit in our DiD analysis. Different variations in subsidy shocks differently affect demand for municipalities and local housing supply and demand, in particular as we allow for local housing supply elasticities to vary across municipalities.

We find significant cross-border spillovers. Our results suggest that for the average municipal federation size, public spending in a French municipality accounts for 4–11% of the bundle of local public goods enjoyed by its residents, and that public spending in its average neighbor municipality — belonging to the same municipal federation — accounts for 3.2–3.5% of this bundle. These high estimates for cross-border spillovers suggest that residents in a given municipality benefit from neighboring municipalities' public goods almost as much as their own municipality's. This may be because French municipalities are in general very small, and because there is almost no legal restriction such as zoning to benefiting from other municipalities' public goods. In a simple application of our method, we simulate the impact of redefining French jurisdictions' boundaries along pre-existing administrative lines and find strong welfare gains.

Our paper is related to the vast literature on fiscal decentralization. In his seminal paper, [Tiebout \(1956\)](#) argues that decentralized public good provision is efficient because people “vote with their feet” to choose their optimal bundle of taxes and public goods. However, [Bewley \(1981\)](#) provides a formal treatment of Tiebout's ideas and concludes that this efficiency result only holds when assuming away interesting features such as spatial spillovers and fiscal externalities. Following [Oates \(1972\)](#), a rich theoretical literature investigates the consequences of spillovers on local public good provision and efficient federalism (e.g., [Gordon 1983](#); [Wellisch](#)

1994; Conley and Dix 1999; Lockwood 2002; Besley and Coate 2003; Bloch and Zenginobuz 2007; Cheikbossian 2008; Bloch and Zenginobuz 2015).

Some empirical work tests the presence of spatial spillovers. Solé-Ollé (2006) investigates benefits spillovers — when households enjoy public goods of neighboring jurisdictions — and congestion spillovers — when households congest such public goods — in the case of local public spending in Spain. The author finds significant evidence of both in equal magnitude. Case et al. (1993) offer a test when local public goods of neighboring jurisdictions are complements and conclude to the existence of spillovers.

On the methodological side, our paper relates to the canonical spatial equilibrium framework of Rosen (1974) and Roback (1982). Drawing on the seminal logit choice setup of McFadden (1973), this workhorse model has since been extended to account for heterogeneous mobility frictions both for households and firms (e.g., Fajgelbaum et al. 2015; Suárez Serrato and Zidar 2016). We also relate to the large literature studying Tiebout type models with endogenous public good provision (Konishi 1996; Epple and Sieg 1999; Brueckner 2000; Bloch and Zenginobuz 2006). Our work is also related to recent research in urban economics modelling endogenous amenities such as Ahlfeldt et al. (2015), Diamond (2016) or Fajgelbaum and Gaubert (2018). In our setup, amenities take the form of public goods and taxes that are endogenous both because of household mobility and the local voting process.

Related empirical work has investigated Tiebout type drivers of migration decisions. Early work such as Oates (1969) studies the impact of local fiscal amenities on housing price capitalization. His estimates show that property values are positively affected by public spending on schools and negatively affected by local taxes. Banzhaf and Walsh (2008) look at the impact of a particular residential amenity, air quality, on jurisdiction density using large plants openings. The authors find that location choices are environmentally motivated. Lutz (2015) estimates significant effects of lower property taxation on residential investment and housing prices, with magnitudes depending on the elasticity of housing supply.

The rest of the paper proceeds as follows. Section 2 provides background on French local public finance and presents some empirical regularities. In Section 3, we develop our spatial

equilibrium model of jurisdictions with endogenous fiscal amenities. Section 4 describes our data. Section 5 presents reduced-form evidence on the impact of local public good supply shocks on different economic outcomes. In Section 6, we use these shocks to structurally estimate our model with GMM. Section 7 presents welfare implications. Section 8 concludes.

## 2 Institutional Background

In this section we provide background on the French local institutional context. We give some historical elements on the early acknowledgement by politicians of the existence of cross-border public good spillovers and of the potential gains from coordination. We then present stylized facts to highlight the prevalence of municipalities in France in the provision of local public services and lay the foundation for our empirical analysis by discussing municipalities' finances.

French local institutions belong to a four-tier system. As of 2016, the territory is divided into 35,352 municipalities (*communes*), nested in 2,000 municipal federations (*intercommunalités*), henceforth MF, 100 counties (*départements*) and 13 provinces (*régions*). Following a series of decentralization laws starting in the early 1980's, France's local authorities increasingly gained autonomy regarding local public services. Each layer of local government has the autonomy to levy taxes and is in charge of providing some public goods. The 35,352 French municipalities represent around 38% of EU's total.<sup>5</sup> This large number of municipalities means that most of them are home to very few residents—383 in the median municipality.

The French government long acknowledged that this large number of jurisdictions may be a source of inefficiency in the provision of local public goods because of unexploited economies of scale, tax competition and poor public service coordination.<sup>6</sup> Central authorities hence created financial incentives to encourage municipalities to merge into larger jurisdictions (see [Leprince and Guengant 2002](#)). Local officials however, supported by their constituents, have

---

5. See data on local administrative areas by Eurostat <https://ec.europa.eu/eurostat/fr/web/nuts/local-administrative-units>.

6. See report by the French Senate <https://www.senat.fr/rap/r05-193/r05-1931.html>.

traditionally opposed such mergers. As a result, the number of municipalities has been fairly stable over time. There were around 38,000 municipalities in the late 18<sup>th</sup> century, compared to roughly 35,000 in 2016. To bypass political obstacles to mergers, central authorities introduced the possibility for neighbouring municipalities to group into municipal federations. This new tier of local government, made up of elected officials from member jurisdictions, would allow municipalities to coordinate without losing autonomy. Initially optional, being part of a municipal federation became compulsory in 2013. As a result, the share of federated municipalities jumped from 74% in 2002 to 100% in 2016, with an average 27 member municipalities per federation.

Despite the creation of this new layer, municipalities are still the largest providers of local public goods and local cooperation beyond basic services remains limited. In 2015, the local public sector — municipalities, counties and provinces — spent €229 billion, representing 9.4% of GDP. Municipalities accounted for 41% of the total, followed by counties (35%), municipal federations (14%) and provinces (11%). When looking at local infrastructure investments alone, totalling €47 billion in 2015, the relative weight of municipalities is even greater. Municipalities' investments represented 41% of the total, twice as much as counties (21%), provinces (20%) or municipal federations (18%).

Municipalities are responsible for urban planning, transport, and environmental policies (19% of their budget in 2009). They are also responsible for kindergartens and primary schools (13% of the budget) which includes school creation, location decision, management and financing, i.e., everything except teachers' hiring and payroll. Finally, they are also in charge of sports, tourism and culture (17% of their budget) including the construction and maintenance of public libraries, museums, music schools, public theaters, sports and tourist facilities, and financial support to sports or cultural events and associations. Municipal federations are mostly in charge of economic development and they may cooperate on the provision of local public goods that are of the responsibility of municipalities, without any obligation.<sup>7</sup>

---

7. They all will be in charge of water provision and waste management in 2020.

Importantly, almost all local public goods provided by municipalities are enjoyable by any household regardless of municipality of residence. This is clearly the case for urban planning, transport and environmental policies. However, it is also the case for sports or cultural facilities and events. In the case of swimming pools and public libraries for instance, there are two kinds of preferential access rules: a special fee for residents or a special fee for residents and residents of close-by jurisdictions. While fees might differ depending on jurisdiction of residence, access can never be denied to anyone. Access to primary schools and kindergartens is subject to zoning which in principle prevents residents of a municipality to benefit from the school quality of a neighboring municipality. However, there are ways to circumvent school zoning in practice. Furthermore, small municipalities have been facing population decline for decades. In order to meet their legal obligation to provide primary and kindergarten schooling, municipalities in rural areas are allowed to form groups regarding schooling provision.<sup>8</sup>

To finance local public goods, municipalities have the autonomy to levy taxes.<sup>9</sup> In 2015, they raised around €50 billion or 2% of GDP in direct and indirect taxes of which 33% were from the tax on resident households, 28% from the tax on property owners, 20% from the local business tax, 2% from the land tax and the remaining 17% from various small taxes (house transaction taxes, waste management tax, etc.).<sup>10</sup> As an alternative source of funding, municipalities receive €17 billion in operating subsidies to cover operating (i.e., non-investment) expenditure. These subsidies are formula-based—loosely speaking increasing with population and decreasing with mean income—and mostly coming from the central government. Finally, municipalities receive around €13 billion from other smaller sources such as various user fees. Together, these sources of funding cover 116% of municipalities' annual operating expenses (around €69 billion in 2015), the 16% surplus being invested in infrastructure.

In 2015, French municipalities invested €19 billion in local infrastructure. Funding in the

---

8. In any case, primary school and kindergarten only account for 13% of municipalities' budget

9. This autonomy, however, is constrained by the presence of several rules limiting year-on-year variations in tax rates.

10. Source: Ministry of Finance.



form of general endowments and investment subsidies accounted for 42% of total investment, the remaining 58% being financed by operating surpluses and additional debt. General endowments can take the form of in-kind gifts from the central government or non-targeted and automatic transfers such as VAT refunds on infrastructure expenditure. Investment-targeted subsidies however are more specific and aimed at financing well-defined investment projects. These are awarded by the boards of counties and provinces as well as by the central government to municipalities that were successful in their grant application. While we do not observe jurisdictions' applications to investment grants, we will argue that these subsidies are plausibly exogenous shocks to local public good supply.

There is substantial cross-sectional variation in investment subsidies received each year. For the purpose of our paper, we define a jurisdiction's *investment subsidy stock* in year  $t$  as the sum of subsidies ever received by the jurisdiction up to  $t$ . This investment subsidy stock concept captures how much of a jurisdiction's public capital is being financed by public funds coming from a higher layer of government.<sup>11</sup> To make yearly subsidies more comparable between municipalities, we operate some normalization. We divide each subsidy amount received in year  $t$  by the subsidy stock in year  $t - 1$ . The normalized yearly subsidy can then be interpreted as the growth in the municipal subsidy stock. We then subtract to each municipal-level observation the national or own MF average. In Figure 4.1 we report the resulting distributions pooling subsidy stocks' yearly growth in 2007, 2009 and 2010.<sup>12</sup> Panel A shows the pooled distribution of this percentage change in excess of the national percentage change. It exhibits substantial variation with the 1<sup>st</sup> percentile being at  $-11\%$  and the 99<sup>th</sup> percentile at  $+59\%$ . Panel B shows a slightly modified distribution, where each jurisdiction subsidy percentage change is considered relatively to the mean change in the jurisdiction's MF. Again, it exhibits substantial dispersion with the 1<sup>st</sup> percentile being at  $-19\%$  and the 99<sup>th</sup> percentile at  $+51\%$ .

---

11. Subsidies financing durable investments are recorded as a liability stock in the municipal accounts. They are depreciated at the same speed as the investment they help financing to keep reflecting their current contribution to local assets.

12. Our data are from the French municipal financial accounts (*Balance Comptable des Communes*) that we present in more details in Section 4.

### 3 Theoretical Model

This section develops our spatial equilibrium model. It extends the seminal framework of [Rosen \(1974\)](#) and [Roback \(1982\)](#) to account for heterogeneous preferences for jurisdictions in the spirit of the recent urban economics literature. Most importantly, we allow for endogenous rents, wages, local public goods and taxes. A distinguishing feature of our model is its focus on local public goods. Endogenous fiscal amenities — taxes and public spending — are central in households' location decisions and are determined through an elementary voting mechanism.

The model shares some methodological features with the urban or economic geography frameworks of [Busso et al. \(2013\)](#), [Ahlfeldt et al. \(2015\)](#), [Suárez Serrato and Zidar \(2016\)](#), [Diamond \(2016\)](#) and [Fajgelbaum and Gaubert \(2018\)](#). There is a finite collection of  $J$  jurisdictions indexed by  $j$  with fixed boundaries, as well as a finite collection of  $A$  mutually exclusive geographic areas indexed by  $a$  in which the  $J$  jurisdictions are nested. Because these nesting areas can be made arbitrarily large, our framework is flexible enough to accommodate many institutional settings. We note  $a_j$  the area  $j$  belongs to.

There is a continuum of imperfectly mobile households of measure 1,  $N_j$  being the share of households living in jurisdiction  $j$ . Households inelastically supply one unit of labor in their jurisdiction of residence. Jurisdiction  $j$  is characterized by a vector of endogenous observables — wage  $w_j$ , rental price  $r_j$ , aggregate public good measure  $\mathcal{G}_j$ , ad-valorem local housing tax  $\tau_j^h$  and local business tax  $\tau_j^k$  — as well as unobserved residential amenities. Local public goods are financed by local housing and business taxes as well as subsidies coming from the central government. A national proportional income tax  $\tau^w$  finances government subsidies to the  $J$  jurisdictions. The sections below describe how demand for jurisdictions, housing supply, wages and local public goods are endogenously set in equilibrium.

### 3.1 Preferences

In order to easily connect theory and empirical analysis, we develop our conceptual framework in a Cobb-Douglas environment. Utility of household  $i$  living in jurisdiction  $j$  is

$$U(\mathcal{C}, \mathcal{G}, i, j) = \mathcal{C}^{1-\phi} \mathcal{G}^\phi \mathcal{E}_j^A \exp(\mu_{ij}) \quad (4.1)$$

where  $\mathcal{G}$  is the aggregate local public good and  $\mathcal{C}$  is a measure of aggregate private consumption,  $\mu_{ij}$  represents the idiosyncratic taste of individual  $i$  for jurisdiction  $j$  and  $\mathcal{E}_j^A$  captures jurisdiction  $j$ 's exogenous residential amenities. Parameter  $\phi \in [0, 1]$  is the taste for public good relative to private consumption, that we assume homogeneous across households.<sup>13</sup>

**Consumption** Agents enjoy aggregate private consumption  $\mathcal{C}$  defined by

$$\mathcal{C} = c^{1-\alpha} h^\alpha \quad (4.2)$$

where  $c$  is consumption of the nationally traded good taken as the numéraire and  $h$  is  $m^2$  housing consumption. Parameter  $\alpha \in [0, 1]$  is the housing consumption share which we assume to be constant across households. Given post-tax rental prices  $r(1 + \tau^h)$  and net income  $(1 - \tau^w)w$ , consumption of the numéraire good is

$$c = (1 - \tau^w)w - r(1 + \tau^h)h \quad (4.3)$$

**Public good** The aggregate public good measure  $\mathcal{G}_j$  depends on the congested local public good of the jurisdiction households live in and on that of all other jurisdictions belonging to the same area  $a_j$  due to the presence of cross-boundary spillovers. For a jurisdiction  $j$ , we assume that congested public goods of other jurisdictions belonging to area  $a_j$  enter  $\mathcal{G}_j$  with equal spillover weights while spillovers coming from jurisdictions outside of  $a_j$  are zero. This binary structure for spillovers' spatial decay is simplistic yet allows to flexibly test for their

---

<sup>13</sup> We hence depart from the Tiebout framework with people sorting according to their taste for public good.

presence in different settings. We define  $\mathcal{G}_j$  as the geometric average of  $j$ 's own congested public good and that of other municipalities of  $a_j$ :

$$\mathcal{G}_j = \mathbb{G}_j^\delta \prod_{j' \in a_j} \mathbb{G}_{j'}^{(1-\delta)/|a_j|} \quad (4.4)$$

where  $\mathbb{G}_j$  represents the congested public good in  $j$  and  $|a_j|$  is the cardinal of  $a_j$ , i.e., the number of jurisdictions in the geographic area  $j$  belongs to. Parameter  $\delta \in [0, 1]$  controls the intensity of spillovers, that is, the extent to which households benefit from local public goods of neighbouring jurisdictions. When  $\delta = 1$  there are no spillovers and residents only enjoy the public goods provided in their jurisdiction. When  $\delta = 0$  there are full spillovers within any area  $a$  so that conditional on living in  $a$ , jurisdiction of residence does not matter for the enjoyment of public goods present in jurisdictions of  $a$ . We adopt a symmetric approach for modelling congestion. We model  $\mathbb{G}_j$ , the congested public good of jurisdiction  $j$  as

$$\mathbb{G}_j = \frac{G_j}{\left( N_j^\delta \prod_{j' \in a_j} N_{j'}^{(1-\delta)/|a_j|} \right)^\kappa} \quad (4.5)$$

where  $G_j$  is the quantity of local public goods provided by jurisdiction  $j$ , which will be endogenously determined along with local taxes. Parameter  $\kappa \in [0, 1]$  controls the intensity of congestion. When  $\kappa = 1$ , public good is fully rival and public good benefits are appropriately measured by per capita spending. When  $\kappa = 0$ , public good is fully non-rival and public good benefits are appropriately measured by absolute public spending. As such,  $\kappa$  is a parameter central to agglomeration economies. The spillover parameter  $\delta$  is also involved in determining the amount of congestion deteriorating the benefits from  $G_j$ . It controls how much of public service congestion is coming from neighbouring jurisdictions as a direct consequence of symmetric benefit spillovers. Absent spillovers,  $\delta = 1$  and public good in  $j$  is only congested by residents of  $j$ . When  $\delta = 0$ , public good in  $j$  is equally enjoyed and congested by all residents of  $a_j$ .

**Residential amenities**    Jurisdiction  $j$  is further characterized by the unobserved residential amenities  $\mathcal{E}_j^A$ . They capture the mean appeal of the jurisdiction's fixed characteristics across individuals and include traditional amenities such as weather, geographic location, etc. They also capture time-varying amenities other than those explicitly modeled. These amenities are equally valued by all residents of  $j$ .

**Idiosyncratic tastes**    Each individual  $i$  draws a vector  $\{\mu_{i1}, \dots, \mu_{iJ}\}$  of idiosyncratic tastes. These  $\mu_{ij}$ 's are assumed to be i.i.d. across and among individuals and distributed Extreme Value Type-I with parameters  $(0, \sigma)$ . They represent individual-city specific utility premia and notably capture heterogeneity in mobility costs and in the valuation of jurisdictions' fixed amenities.<sup>14</sup> Parameter  $\sigma$  controls the dispersion of these idiosyncratic draws and is inversely related to household mobility. When  $\sigma$  is higher, density around the indifference threshold between any two jurisdictions is thinner as more households are infra-marginal. As a consequence, the migration response to a marginal change in the appeal of one jurisdiction relative to the other gets smaller.

Model parameters to be estimated so far are  $\{\sigma, \phi, \kappa, \delta\}$  capturing inverse household mobility, taste for the public good, public good congestion and cross-boundary spillovers. The housing consumption share  $\alpha$  will be calibrated from the literature.

### 3.2 Conditional Housing Demand

Conditional on living in  $j$ , agent  $i$  decides how much housing to consume while being net wage, rental price and tax taker. Given the constant share assumption, conditional individual

---

14. An interpretation of a higher value for  $\mu_{ij}$  relative to any other  $\mu_{ij'}$  is  $j$  being the jurisdiction in which  $i$  was born, educated and socialized. Another interpretation is heterogeneity in preferences for local exogenous amenities (e.g., weather, natural amenities, etc.).

housing demand and numéraire consumption equal

$$\begin{aligned} h_j^D &= \alpha \frac{(1 - \tau^w) w_j}{r_j (1 + \tau_j^h)} \\ c_j &= (1 - \alpha) (1 - \tau^w) w_j \end{aligned} \tag{4.6}$$

and do not depend on  $i$ . Per capita housing and numéraire consumption will hence be treated as endogenous jurisdiction amenities.

### 3.3 Demand for Jurisdictions

Agent  $i$  chooses to live in the jurisdiction that maximizes  $U(\mathcal{C}, \mathcal{G}, i, j)$ . We can write:

$$\ln U(\mathcal{C}, \mathcal{G}, i, j) = v_j + \mu_{ij} \tag{4.7}$$

where

$$v_j = (1 - \phi) \ln(\mathcal{C}_j) + \phi \ln(\mathcal{G}_j) + \ln(\mathcal{E}_j^A) \tag{4.8}$$

Households first solve for optimal housing and numéraire good consumption conditional on jurisdiction of residence according to (4.6). Based on their idiosyncratic tastes  $\mu_{ij}$ , they make the extensive margin choice of where to live upon observing local aggregate private consumption  $\{\mathcal{C}_j\}$ , aggregate local public goods  $\{\mathcal{G}_j\}$  as well as exogenous amenities  $\{\mathcal{E}_j^A\}$ . When comparing different jurisdictions, households are thus  $v_j$  takers. This conditional logit setup was first introduced by [McFadden \(1973\)](#) in a broader context of discrete choices. Demand for jurisdiction  $j$  then equals the expected set of households for which  $j$  yields the highest utility, i.e.,  $N_j = \mathbb{E} [\mathbf{1}_{\{u_{ij} > u_{iq} \forall q \neq j\}}]$ . Because idiosyncratic shocks are distributed Extreme Value Type-I and enter utility separately from other components, population in jurisdiction  $j$  is equal to

$$N_j = \frac{\exp(v_j/\sigma)}{\sum_{j'} \exp(v_{j'}/\sigma)} \tag{4.9}$$

Loosely speaking, demand for jurisdiction  $j$  is the ratio between how attractive the jurisdiction is and the mean jurisdiction appeal in the country. Equation (4.9) notably implies that no jurisdiction is empty and that the market for jurisdictions clears:

$$\sum_j N_j = 1 \tag{4.10}$$

Total housing demand in jurisdiction  $j$  is then the result of intensive margin consumption and extensive margin in-migration:

$$H_j^D = h_j^D N_j \tag{4.11}$$

### 3.4 Housing Supply

We assume that a representative absentee landlord has the opportunity to put existing homes on the market or to develop new ones, rented at a price  $r$  per  $m^2$  of housing. The marginal opportunity or development cost is increasing in the quantity of housing already on the market and decreasing in the jurisdiction area  $T$ . Formally, the cost of providing housing is  $\left(\frac{H}{T}\right)^{1+\frac{1}{\eta_j}} \mathcal{E}_j^C$  where  $\eta_j$  is the housing supply elasticity. Further differences in local housing supply determinants are captured by the cost shifter  $\mathcal{E}_j^C$ . Since it enters the housing supply cost function in a multiplicative way, it is isomorphic to a reduction in available land according to  $T/\mathcal{E}_j^C$ . Profit maximization yields the inverse housing supply equation

$$\ln(r_j) = \frac{1}{\eta_j} \ln\left(\frac{H_j^S}{T_j}\right) + \ln(\mathcal{E}_j^C) \tag{4.12}$$

### 3.5 Labor Demand

Local labor markets are not the focus of this paper. However, our analysis needs to account for endogenous wages as they are potentially affected by endogenous business taxation and productivity, e.g, through public investments. We make two standard assumptions. First, each jurisdiction produces a quantity  $Y_j$  of the traded good with a constant return to scale

technology using local labour and capital. Second, absentee capital owners supply capital in all jurisdictions at a fixed interest rate  $R$ . We further assume that local business taxation is proportional to the outflow of local interest payments accruing to capital owners. With these assumptions, one easily shows that the inverse labor demand equation has the following form:

$$w_j = \theta_j^Y (\{G\}_{j'}) f ((1 - \tau_j^k) R) \quad (4.13)$$

with  $w_j$  the local wage,  $\tau_j^k$  the local business tax,  $f$  any function and  $\theta_j^Y (\{G\}_{j'})$  the local labour productivity that may be endogenous to the quantity of public goods  $G_{j'}$  of any jurisdiction  $j'$  in  $a_j$ . With this formulation, local wages are negatively affected by an increase in the local business tax and are positively affected by productivity shocks. Our empirical analysis would not be affected if we allowed for a more general production function with decreasing returns to scale (e.g., because of land use), productivity agglomeration gains or productivity spillovers as long as we econometricians observe the equilibrium wage. We would simply replace our inverse labor demand equation by a reduced-form expression such as

$$w_j = f_j (R, \{N, \tau^k, \theta^Y\}_{j'})$$

### 3.6 Public Good Supply

As mentioned in Section 3.1,  $G_j$  is the quantity of public goods provided by jurisdiction  $j$ .<sup>15</sup> Local taxes — ad-valorem housing tax  $\tau_j^h$  and business tax  $\tau_j^k$  — and subsidies  $F_j$  from upper layers of government fund the production of  $G_j$ . In our setup,  $F_j$  is exogenous. Residents of jurisdiction  $j$  choose the triplet  $\{G_j, \tau_j^h, \tau_j^k\}$  that maximizes their utility subject to the municipal budget constraint. Because we assume homogeneous preferences for private consumption, the voting mechanism is akin to a maximization problem by a local social planner.

---

15. Section B details how we construct  $G$  from available data. In our empirical application,  $G$  includes both the flow of public services stemming from the use of public infrastructure and the flow of local public services that are typically staff expenditure, maintenance costs, subsidies to associations, etc.



Residents' preferred policy is found by maximizing

$$(1 - \alpha) \ln \left( (1 - \tau^w) w - r (1 + \tau^h) h \right) + \alpha \ln h + \frac{\phi}{1 - \phi} \delta \ln G \quad (4.14)$$

over  $\{G, \tau^h, \tau^k\}$  subject to jurisdiction budget constraint

$$\Gamma_j(G) = \tau^h r H + \tau^k R K + F \quad (4.15)$$

where  $\Gamma_j$  maps public good quantities into the required public funds. Subsidies are financed by a national income tax  $\tau^w$  that endogenously adjusts so that national budget is balanced:

$$\tau^w = \frac{\sum_j F_j}{\sum_j N_j w_j} \quad (4.16)$$

Although we do not solve for local policies in the comprehensive case, this framework will be useful in our welfare application in Section 7.

### 3.7 Equilibrium

We close the model by imposing market clearing conditions for the traded good and housing:

$$\begin{aligned} \sum_j Y_j &= \sum_j N_j c_j \\ H_j^S &= H_j^D \quad \text{for all } j \end{aligned} \quad (4.17)$$

**Definition 1.** *Given the model's parameters  $\{\sigma, \phi, \kappa, \delta, \alpha\}$  and  $\{\eta_j, T_j\}$ , local fundamentals  $\{\mathcal{E}_j^A, \mathcal{E}_j^C\}$ , functional forms  $\{f, \Gamma_j, \theta_j^Y\}$  and national subsidies  $\{F_j\}$ , an equilibrium is defined by jurisdiction population  $\{N_j\}$ , individual consumption  $\{\mathcal{C}_j, c_j, h_j\}$ , housing supply  $\{H_j\}$ , local prices  $\{w_j, r_j\}$ , public goods  $\{G_j, \mathbb{G}_j, \mathcal{G}_j\}$ , local taxes  $\{\tau_j^h, \tau_j^k\}$  and the income tax  $\tau^w$ , which are determined by the following system of equations: aggregate and congested local public good definitions (4.4) and (4.5); local planner's optimization problem (4.14) and (4.15); jurisdiction utility (4.8); demand for jurisdictions (4.9); conditional housing demand (4.6);*

*housing supply* (4.12); *labor demand* (4.13); *market clearing conditions* (4.17) and *national budget constraint* (4.16).

Because of agglomeration forces in the model, Definition 1 may not characterize a unique equilibrium. We argue in the following Section 6 that structural parameters can be uniquely identified nonetheless.

### 3.8 Residential Amenities

We use the equilibrium equations to express changes in unobserved residential amenities as a function of structural parameters and changes in observable endogenous variables. This step provides theoretical support to our difference-in-differences strategy in Section 5 and is the foundation of our GMM structural estimation in Section 6.

We first take the log of the demand for jurisdictions equation (4.9) and plug in jurisdiction utility equation (4.8), consumption equations (4.2) and (4.6), and public good definitions (4.4) and (4.5). We find an expression for residential amenities  $\ln(\mathcal{E}_j^A)$ :

$$\begin{aligned}
 \ln(\mathcal{E}_j^A) = & -(1 - \phi)(1 - \alpha) \ln(1 - \tau^w) - (1 - \phi)(1 - \alpha) \ln(w_j) \\
 & - (1 - \phi)\alpha \ln(h_j) - \phi\delta \ln(G_j) - \phi(1 - \delta) \frac{1}{|a_j|} \sum_{j' \in a_j} \ln(G_{j'}) \\
 & + (\sigma + \kappa\phi\delta^2) \ln(N_j) + \phi(1 - \delta^2) \kappa \frac{1}{|a_j|} \sum_{j' \in a_j} \ln(N_{j'}) \\
 & + \sigma \ln \sum_{j'} \exp(v_{j'}/\sigma) + \text{constant}
 \end{aligned} \tag{4.18}$$

While our model is set up in a static framework, we introduce time here as our empirical strategy in the following sections relies on time variation. In this new setup, residential amenities  $\mathcal{E}_{jt}^A$  may contain time fixed effects in addition to jurisdiction fixed effects. Using (4.18) we derive two expressions describing how residential amenities change over time. First, take a jurisdiction  $j$  nested in the larger geographic area  $a_j$ . We find an expression for how jurisdiction  $j$ 's residential amenities change over time relatively to residential amenities of all jurisdictions belonging to  $a_j$  (*within-a amenity changes*). Then, consider any geographic area

*a*. We find an expression for how the average residential amenity in *a* changes relatively to the average amenity in all geographic areas in the country (*between-a amenity changes*).

Looking at within-*a* and between-*a* relative changes bears two advantages. First, it provides a theoretical framework for our DiD and GMM estimations that does not require to fix the utility of one jurisdiction or the denominator of (4.9).<sup>16</sup> Second and given the assumption on the structure of spillovers, it allows us to isolate the spillover parameter  $\delta$ .

**Within-*a* amenity changes** We subtract to equation (4.18) its within- $a_j$  average. This differences out all terms that are common across jurisdictions of  $a_j$ , including any  $a_j$ -specific time fixed effect. We then first-difference the resulting equation which absorbs any jurisdiction fixed effect. The final expression for amenity changes is

$$\begin{aligned} \Delta \ln \overline{\mathcal{E}}_{jt}^A &= (\sigma + \kappa\phi\delta^2) \Delta \ln \overline{N}_{jt} - (1 - \alpha)(1 - \phi) \Delta \ln \overline{w}_{jt} \\ &\quad - \alpha(1 - \phi) \Delta \ln \overline{h}_{jt} - \delta\phi \Delta \ln \overline{G}_{jt} \end{aligned} \tag{4.19}$$

where

$$\overline{X}_j = \frac{X_j}{\prod_{j' \in a_j} X_{j'}^{\frac{1}{|a_j|}}}$$

for any variable  $X$  and  $\Delta$  is the first-difference operator between any two arbitrary periods. Note that with full spillovers ( $\delta = 0$ ), within-*a* relative public good changes have no effect on within-*a* relative migration responses as location does not matter for the enjoyment of public goods conditional on living in *a*.<sup>17</sup>

**Between-*a* amenity changes** Going back to equation (4.18), we take its average across all jurisdictions of any geographic area *a*. Treating *a* as the new level of observation, we then

---

16. See for instance the seminal work of [McFadden \(1973\)](#) and more recently [Diamond \(2016\)](#) or [Fajgelbaum et al. \(2015\)](#) for examples of how this technical point is dealt with. We instead absorb the unobserved denominator of equation (4.9) by expressing residuals in relative terms.

17. Alternatively, one can plug in the housing consumption expression from equation (4.6):

$$\begin{aligned} \Delta \ln \overline{\mathcal{E}}_{jt}^A &= (\sigma + \kappa\phi\delta^2) \Delta \ln \overline{N}_{jt} - (1 - \phi) \Delta \ln \overline{w}_{jt} \\ &\quad + \alpha(1 - \phi) \Delta \ln \overline{r}_{jt} + \alpha(1 - \phi) \Delta \ln \overline{1 + \tau_{jt}^h} - \delta\phi \Delta \ln \overline{G}_{jt} \end{aligned}$$

subtract the average across all  $a$ 's. This step absorbs all fixed effects common across  $a$ 's, including common time fixed effect. We then first-difference the resulting equation which absorbs any  $a$  fixed effect. The final expression is

$$\begin{aligned} \Delta \ln \widehat{\mathcal{E}}_{at}^A &= (\sigma + \kappa\phi) \Delta \ln \widehat{N}_{at} - (1 - \alpha)(1 - \phi) \Delta \ln \widehat{w}_{at} \\ &\quad - \alpha(1 - \phi) \Delta \ln \widehat{h}_{at} - \phi \Delta \ln \widehat{G}_{at} \end{aligned} \tag{4.20}$$

where

$$\widehat{X}_a = \frac{\prod_{j \in a} X_j^{\frac{1}{|a|}}}{\prod_{a'} \left( \prod_{j \in a'} X_j^{\frac{1}{|a'|}} \right)^{\frac{1}{A}}}$$

for any variable  $X$  and  $A$  is the number of areas  $a$ . Importantly, these between- $a$  amenity changes do not feature parameter  $\delta$ . Indeed, public good spillovers are contained within each  $a$  while equation (4.20) is at the  $a$  level.<sup>18</sup>

## 4 Data

We take advantage of a unique combination of panel administrative datasets that we obtained from the French Ministry of Finance and Ministry for the Environment as well as publicly available datasets at the jurisdiction level, overall covering the period 2000–2016. In this section we introduce the data and describe how we construct the variables present in the model.

**Sample** Our sample is the universe of mainland French municipalities that experience no boundary changes between 1999 and 2016. Not dropping municipalities experiencing boundary changes—such as municipalities merging or acquiring land from others—may lead

---

18. Alternatively, one can plug in the housing consumption expression from equation (4.6):

$$\begin{aligned} \Delta \ln \widehat{\mathcal{E}}_{at}^A &= (\sigma + \kappa\phi) \Delta \ln \widehat{N}_{at} - (1 - \phi) \Delta \ln \widehat{w}_{at} \\ &\quad + \alpha(1 - \phi) \Delta \ln \widehat{r}_{at} + \alpha(1 - \phi) \Delta \ln \widehat{1 + \tau_{at}^h} - \phi \Delta \ln \widehat{G}_{at} \end{aligned}$$

to artificial variation in their supply of local public goods, population and economic outcomes. Fortunately, very few jurisdictions experience such changes and this manipulation leaves us with 34,835 jurisdictions, i.e., 96% of them, and an almost-complete partition of the French mainland territory.

**Municipal financial accounts** We use detailed municipal financial accounts (*Balance Comptable des Communes*) obtained from the French Ministry of Finance for every year between 2002 and 2016. Financial accounts contain information on municipalities' detailed asset and liability position as well as the composition of their yearly budget. Although it would be of great interest to disentangle the various components of local public goods, we are unable to classify the various items by their nature as the existing categories follow accounting definitions. This is of little consequence, however, as this paper is more general and seeks to address the broad inefficiencies in the provision of local public goods and not the effect of a particular type of public good.

We construct the public good quantity  $G$  from municipal accounts by assuming that  $G$  is a Cobb-Douglas public good index capturing both durable facilities and operating expenditure that benefit residents (see Appendix B for detailed methodology). We construct durable facilities as the sum of all public assets minus the raw value of the land and financial assets such as cash.<sup>19</sup> As mentioned in Section 2, investments notably include schools, transportation infrastructure, parks improvements, sports facilities, museums, art collections, etc. They also include investment subsidies to other parties such as local clubs and associations. Importantly, they do not contain social housing units.<sup>20</sup> We construct operating expenditure as the sum of staff expenditure, maintenance spending, payments for external services and operating subsidies to third parties.<sup>21</sup>

In our model, local public budgets are partly financed by public subsidies  $F_j$  coming from

---

19. Taking out the raw value of the land seems natural as residents are unlikely to value it. Given that land is mostly a gift from the central government, erasing this asset as well as the corresponding liability from the balance sheet is neutral in our analysis. Cash and other liquid assets can be considered negative debt and are accounted for in our theoretical framework in the form of future taxes. They are recorded at book value and account for investment depreciation.

20. Social housing units, when publicly owned, are held by ad hoc entities and not by jurisdictions.

21. We exclude interests payments as they do not correspond to consumable services.

the central government and upper layers of local government. In our empirical application, we instrument local public goods using a subset of these subsidies, namely investment-targeted subsidies, that we call  $S_j$ . We define  $S_j$  as the sum of all investment subsidies ever received by municipality  $j$  from counties, provinces and the central government minus all associated depreciation. In 2016, investment subsidies represent 23% of the financing of public capital. Endowments account for 28% while the remaining 41% are from local contributions—past (reserves) or future (debt). As mentioned in Section 2,  $S_j$  depreciates together with the public capital it helps financing. We will see that  $S_j$  proves a good basis for instrumenting public goods  $G_j$ .

**Population and housing** We use FILOCOM (*Fichier des Logements à la Commune*) which is an exhaustive database on household housing stock. It provides information on each non-commercial dwelling every two years between 1994 and 2014 such as its location, surface, vacancy status, whether it is a main or a secondary home, rented or owner-occupied, and its number of residents. We use FILOCOM to construct our municipality-level population and housing database which contains the stock of rented or owner-occupied main homes per jurisdiction and the total and per capita housing  $m^2$  surface. Population  $N_j$  from the model is all inhabitants of housing units for which the head of household is aged 20–65 in jurisdiction  $j$ . Total housing consumption  $H_j$  is how much  $m^2$  are consumed by the  $N_j$  residents, and per capita housing consumption is  $h_j = H_j/N_j$ .

**Housing prices** We use data on housing prices instead of rents as there is no exhaustive database on rents at the jurisdiction level prior to 2015. We combine housing transactions database from the notary offices (named BIEN for the Parisian region and PERVAL for the rest of France). We construct a database on house prices per  $m^2$  for every two years between 2000 and 2014 at the municipality level. We assume that housing transaction prices from the notary data are the net present value of unobserved rents  $r$ . Proportional changes in  $r$  thus equal proportional changes in housing prices. This method bears two caveats, one methodological and the other regarding the quality of the price data. First, although tenants

pay rent every year—explicitly or implicitly—not all jurisdictions record house transactions every year. As such, we can infer the rent growth of a jurisdiction only when it experiences a transaction in two consecutive periods. Second, the coverage of housing transactions in the notary data is not exhaustive.

**Local tax data** We use detailed local tax data (*Recensement des Éléments d'Imposition à la Fiscalité Directe Locale*—REI) for every year from 2002 to 2016. REI features all tax bases and rates at the jurisdiction level. We compute the ad-valorem local residence tax of the model from observed residence tax revenues  $\text{Rev}_h$ , total housing stock  $H$  and house prices  $p$  at the jurisdiction level:  $\tau^h = \frac{\text{Rev}_h}{Hr}$  where house prices are the net present value of rents  $r = \frac{R}{1+R} \times p$ .

**Wages** We use labor income data from the IRCOM (*Impôt sur le Revenu par Commune*) dataset. IRCOM summarizes labor and social security total income at the municipal level from 2002 to 2016. We construct our jurisdiction wage measure  $w$  as total labor income divided by the number of tax units reporting positive labor income.

**Socio-demographics and geographic data** We use census data to get jurisdiction-level information on total population and socio-demographic characteristics for years 1990, 1999, 2008 and 2013. We also use publicly provided data from the National Statistical Institute (INSEE) on jurisdictions' geography (municipal federation they belong to, distance to center of urban area center, area, etc.). Although our analysis does not study the public goods provided by municipal federations, we use these geographic groupings for the definition of the  $a$ 's. From 2000 to 2016, municipalities gradually joined MFs. In 2016, all municipalities belonged to municipal federations. We assign each municipality to its 2016 MF for the whole duration of our panel so that these geographic groupings are constant over time.

Table 4.1 gives elementary descriptive statistics on jurisdictions summarizing some of the above constructed variables.

## 5 Reduced-Form Evidence

This section builds on the theoretical framework set out in Section 3 and presents the results of DiD models looking at the impact of changes in local public goods on changes in a range of economic outcomes—population, housing consumption, house prices and wages.

In our empirical application, the  $a$ 's are the municipal federations in their 2016 form introduced in Section 2. Investigating whether these geographic areas are relevant for further fiscal integration makes economic and historical sense. Indeed, as argued in Section 2, jurisdictions historically grouped into MFs to partly internalize public good externalities and rationalize costs. We expect a relatively higher public good supply shock to increase relative migration towards targeted jurisdictions, increase total housing consumption, bid up rents in the housing market and consequently lower per capita housing  $m^2$  consumption. Because of public good spillovers within MFs, we expect within-MF changes in public good supply to have a smaller impact than between-MF changes. Although labor markets are not the focus of this paper, we can conjecture that potential productivity increases would increase wages.<sup>22</sup>

### 5.1 DiD Framework

Changes in public goods are likely endogenous to changes in the other variables of our model. We instrument changes in local public goods using the large variation in investment-targeted subsidies for which we give descriptive evidence in Sections 2 and 4. This identification strategy would be undermined if investment subsidies were correlated with changes in jurisdictions' characteristics. Such correlation would arise, for instance, if investment subsidy were systematically granted based on jurisdictions' population, housing or labor market outcomes. We follow [Suárez Serrato and Zidar \(2016\)](#) and [Fuest et al. \(2018\)](#) and use DiD models to look at cumulative changes in a range of economic outcomes around investment subsidy shocks while relying on the absence of pre-trends to make the case for the validity of our instrument.

In line with our theoretical framework, we look at both within-MF and between-MF

---

22. Note that although it is not modeled here, a public good supply shock may come with a public labor demand shock that would increase wages.



variations in public goods and other endogenous variables. We instrument within-MF changes in public goods with within-MF yearly changes in jurisdictions' subsidy stock  $\Delta \ln \bar{S}_j$ . Similarly, we instrument between-MF changes in public goods with between-MF yearly changes in jurisdictions' subsidy stock  $\Delta \ln \widehat{S}_a$ .

We pool shocks happening in 2007, 2009 and 2010 and keep a  $[-4, 4]$  distance windows around shock years. We limit ourselves to these shocks for two reasons. First, it allows us to have a balanced panel in terms of distance to the shocks and sufficient temporal depth—i.e., 4 years pre-shock—to inspect the pre-trends. This mechanically reduces the scope to shocks happening around the middle of our sample. Second, we do not include the 2008 subsidies since 2008 is a municipal election year and we suspect (and empirically confirm) that subsidies that year were much more endogenous to jurisdiction conditions.

For both within-MF and between-MF regressions, we investigate responses from population, housing per capita, total housing, wage, rents and the public good index.

**Within-MF regressions**      We first run the following regressions:

$$\Delta \ln \bar{X}_{j,d,-6} = \beta_d^W \Delta \ln \bar{S}_{j,0,-1} + \sum_{k=-6}^{\min\{d,-1\}} \beta_k^W \Delta \ln \bar{S}_{j,k,k-1} + \mathcal{I}_j \cdot \gamma_d^W + u_{j,d}^W \quad (4.21)$$

where  $X$  is any of the model's endogenous variables. We run a separate regression for each  $d \in [-4, 4]$  where  $d = 0$  is any shock year, i.e., 2007, 2009 or 2010. The main explanatory variable is the one-year relative subsidy change  $\Delta \ln \bar{S}_{j,0,-1}$ .  $\mathcal{I}_j$  is a vector of flexible dummies for 1999 baseline characteristics included to control for those changes in unobserved amenities, housing supply or productivity determinants that are correlated with these observed characteristics.<sup>23</sup>

We cluster standard errors at the MF level to account for spatial correlation of the error term as well as auto-correlation since our regressions pool shocks happening in different years.

While the instrument is the one-year log difference in our subsidy stock measure, we look at long differences in post-shock outcomes to account for construction delays, frictions in year-on-year responses by individuals and more generally for dynamic adjustments to

---

23. Controls include normalized baseline 1999 jurisdiction population and density.

the shock. Regression coefficients we obtain are interpreted as cumulative elasticities. To deal with potential serial correlation of the instrument we also control for  $\Delta \ln \bar{S}_{j,k,k-1}$  with  $k \leq \min\{d, -1\}$ , that is, all subsidy shocks happening before the shock year and before period  $d$ .

**Between-MF regressions** Similarly we run the following regressions:

$$\Delta \ln \widehat{X}_{a,d,-6} = \beta_d^B \Delta \ln \widehat{S}_{a,0,-1} + \sum_{k=-6}^{\min\{d,-1\}} \beta_k^B \Delta \ln \widehat{S}_{a,k,k-1} + \mathcal{I}_a \cdot \gamma_d^B + u_{a,d}^B \quad (4.22)$$

The reduced-form elasticities  $\beta_d$ 's represent how the subsidy shock in period 0 affects cumulative outcome growth between  $-6$  and  $d \in [-4, 4]$ . The absence of pre-trends, i.e.,  $\beta_d = 0$  for  $d < 0$  for all outcomes of our within-MF and between-MF regressions would strongly support the validity of our instrument.

## 5.2 Results

We present graphical evidence on the evolution of our different outcomes around subsidy shocks based on our within-MF and between-MF DiD models. These graphs have no immediate “treatment effect” interpretation but show how the gradient of different outcomes is affected by presumably exogenous changes in subsidy stocks. Their contribution, however, is twofold. They provide convincing evidence of absence of selection into treatment, as well as of significant behavioral responses to changes in local public goods. They also give support to our theoretical framework as we find stronger behavioral responses—in terms of migration, housing consumption and rents—in the between-MF setup than in the within-MF setup, which we interpret as evidence of stronger public good spillovers within municipal federations than between.

**Within-MF behavioural responses** Figure 4.3 shows the first stage result, i.e., how

the cumulative growth of our public good index evolves around a sudden investment subsidy shock. The graph reports our DiD coefficients  $\beta_d^W$ , that is, the effect of the shock in  $d = 0$  on public good's cumulative growth starting in  $d = -6$ . Visual inspection of the pre-trends confirms the exclusion restriction. Before  $d = 0$ , cumulative growth is flat and while it is significantly different from zero, it is not economically significant compared to the subsequent hike. After  $d = 0$ , growth in  $G$  becomes strongly positively correlated with the shock. A 1% change in the stock of investment-targeted subsidies leads to a 0.37% change in the public good index after five years. All changes are relative to the mean changes in the MF. This effect is significant at the 1% level.

We now turn to the reduced-form effect of subsidy shocks on population. Figure 4.4 shows that treatment intensity is not correlated with municipality migration dynamics prior to  $d = 0$ . However, municipalities which received a relatively higher subsidy shock subsequently experienced relatively higher in-migration. A 1% relative increase in investment-targeted subsidies leads to a 0.03% relative increase in population after five years. This effect is significant at the 1% level. Put together with the first stage result, the reduced-form elasticity of population to public spending within a municipal federation is approximately 0.07. In other words, when municipality public spending grows 10 percentage points more than (geometric) mean public spending in the MF, municipality population grows 0.7 percentage points more than (geometric) mean population in the MF.

It is useful to give a concrete example illustrating the intensity of migration responses within municipal federations. Consider a municipal federation  $a$  in its final 2016 form with mean number of member municipalities (27 municipality members) all having 2009 mean population (1,000 residents), mean investment level (€9,400,000) and mean operating expenditure (€1,540,000). Let us abstract from public good depreciation for simplicity. Assume that municipality  $j$  of  $a$  raises its investment level and yearly expenditure by 10%, that is, invests €940,000 in durable infrastructure and commits to increase yearly operating expenditure by €154,000. All other member municipalities of  $a$  decrease them by  $10\%/26 = 0.38\% \approx 0\%$ . Other municipalities in other MFs do not change their policies. Mean public good growth

in  $a$  is zero so that  $a$  will experience no in-migration according to the model's equations. Municipality  $j$  experiences a 10 percentage points growth in public services in excess of the zero mean MF growth, and will experience a 0.7 percentage points population growth. This represents in-migration of 7 additional residents coming from out-migration from other municipalities of  $a$  for an initial investment of €940,000 and an increase of €154,000 in yearly expenditure.

Figure 4.5 shows how housing consumption per capita and wages evolve around the subsidy shock. Pre-trends are flat and not significantly different from zero in all Panels. They make an even stronger case for investment subsidies as exogenous shocks to local public good supply. None of the endogenous outcomes—public goods, population, housing consumption and wages—exhibit pre-shock dynamics correlated with shock intensity. Panel A offers evidence that subsidy shocks are not correlated with increasing or decreasing pressure in local housing markets. Furthermore, we find that per capita housing consumption, which subsumes housing prices and residence tax responses, is not affected by the subsidy shocks. Panel B gives comfort that public investment shocks are not driven by favourable or unfavourable trends in local productivity. Wages seem to be unaffected by subsidy shocks. The distance window we consider might be too small to dissipate frictions in wage adjustments, however.<sup>24</sup> Within-MF subsidy shocks hence seem to induce migration responses through changes in local public goods that do not capture significant changes in other endogenous municipality characteristics.

Table 4.2 reports alternative estimates without the different control variables. Our preferred estimates in column (3) correspond to the endpoints of Figures 4.3, 4.4 and 4.5. They are largely unaffected when we experiment with different specifications.

**Between-MF behavioural responses** We now investigate how MF-level changes in investment subsidies affect changes in MF-level economic outcomes. Figure 4.7 is the symmetric of Figure 4.3 and shows the first stage of our between-MF regressions. Public good cumulative growth is not significantly different from zero before the shock but it adjusts sharply after the

---

24. We also abstract from the possibility that agents commute to neighbor municipalities to work, which would dilute effects on wages even further.

shock. A 1% increase in the subsidy stock leads to a 0.42% increase in the public good index after five years. All changes are relative to mean regional changes taking the geometric average across each municipal federation as the new level of observation. This effect is significant at the 1% level.

Figure 4.8 shows that MFs which received higher subsidy shocks also experienced relatively higher in-migration. A 1% relative shock leads to a 0.15% relative population increase after five years. This effect is significant at the 1% level. It represents a population elasticity with respect to public goods of approximately 0.37, that is, a 10 percentage points increase in public spending in excess of regional public spending growth is met with a subsequent 3.7 percentage points excess population growth. This estimate is roughly five times higher than the point estimate of within-MF regressions. As outlined in our theoretical framework, a candidate mechanism to explain this discrepancy is the presence of cross-boundary spillovers. Indeed, in the polar case of full spillovers, changes in municipality local public goods relative to mean changes in the MF should not affect within-MF migration. However, changes between MFs would still affect migration decisions as we assume spillovers abruptly die out at the MF frontier.

Again it is useful to illustrate the intensity of migration responses between municipal federations. Consider the same setting as the above example, but this time assume that *all* municipalities of *a* increase their investment and yearly expenditure by 10%, that is, invest €940,000 in durable infrastructure and commit to increase yearly expenditure by €154,000. All other municipalities in all other municipal federations decrease them by 10%/1266  $\approx$  0%. Municipalities of *a* experience a mean 10 percentage points growth in public services in excess of the zero mean growth, and will experience a mean 3.7 percentage points population growth. This represents in-migration of 37 additional residents coming to each municipality of *a* from out-migration *from other federations* for an initial investment of €940,000 and an increase of €154,000 in yearly services.

Figure 4.9 suggests that this reduced-form elasticity differs from the micro (i.e., absent general equilibrium adjustments) migration responses to public good changes since it does not

hold constant adjustments in other local amenities. Panel A shows again that subsidies were not awarded according to local trends in housing supply determinants. However, consumable housing per capita decreases in equilibrium in response to migration pressure and increasing marginal housing supply costs with a reduced-form elasticity of housing per capita with respect to public goods of  $-0.07$ . This also contrasts with our within-MF analysis. Panel B shows again that local productivity dynamics are not correlated with subsidy shocks. However, we find this time that wages are affected by local public goods. This result suggests that local public investment is also shifting the labor demand curve upwards by boosting local productivity. This is further evidenced by Panel C where we report the cumulative elasticities of the (geometric) average number of businesses with respect to the subsidy stock. Contrasting this result with the absence of within-MF wage variation hints at the presence of production spillovers or commuting within municipal federations. In any case, these effects are fully captured by the residential wage variation and do not bias our spillover estimates as long as we separately instrument wage changes in the GMM estimation.

Overall these results are consistent with weaker public good spillovers between municipal federations than within. Table 4.3 reports alternative estimates where we let the fixed effect in the residential amenities be at the national or county level, i.e., where we divide each observation by the national or county geometric mean instead of the regional mean. They are largely unaffected when we experiment with different specifications. While the migration response seems to be smaller when absorbing county fixed effects instead of national or regional ones, it mirrors a weaker first stage intensity so that the reduced-form elasticities are comparable. Our preferred estimates in column (2) correspond to the endpoints of Figures 4.7, 4.8 and 4.9.

**Housing supply and price capitalization** Figures 4.11 and 4.13 report the results of our housing supply regressions and show how total  $m^2$  housing consumption and housing prices (in  $\text{€}/m^2$ ) are affected by investment subsidies changes. Panel A of Figure 4.11 reports estimates for housing consumption changes conditional on municipal federation. A 1% increase in the subsidy stock relative to the MF average increase is met with a 0.02% relative increase

in total consumed  $m^2$  by 20–65 residents after five years. This estimate is significant at the 1% level. Panel B shows housing price capitalization results: a 1% increase in the subsidy stock is met with an insignificant 0.02% increase in house prices after five years.

Panel A of Figure 4.13 reports estimates for MF-level mean housing consumption changes conditional on region. A 1% increase in the subsidy stock relative to the regional average increase is met with a 0.13% relative increase in total consumed  $m^2$  by 20–65 residents after five years. This estimate is significant at the 1% level. Panel B again shows house price capitalization estimate. This time, a 1% increase in the subsidy stock is met with an 0.24% increase in house prices after five years significant a 1%. Panel B hence comforts our interpretation of subsidy shocks as additional public good amenities that are positively valued by residents. Most importantly, significant housing price capitalization in the between-MF setup and not in the within-MF setup is consistent with the presence of strong within-MF spillovers. It is worth noting that housing price response is an order of magnitude larger than the per capita housing drop in the between-MF case. We argue that housing price responses are more representative of the longer-term adjustment in rental prices facing newcomers as short- to medium-term housing per capita responses might be dampened by adjustment frictions. We will use house prices in our GMM to estimate the model's parameters.

Table 4.4 reports alternative estimates. Our preferred estimates in column (3) correspond to the endpoints of Figures 4.11 and 4.13. They are robust to different specifications.

Overall, these reduced-form results show that households value local public goods as evidenced by migration responses and capitalization in housing prices. They also highlight, in line with our theoretical framework, that the intensity of cross-boundary spillovers influences the magnitude of migration responses to local public good supply shocks.

### 5.3 Robustness Checks

We run a series of robustness checks investigating whether the observed migration patterns could be driven by mechanisms other than responses to changes in positively valued public

goods.

A first concern is that changes in local public spending may not be valued by residents in themselves, but may be correlated with changes in housing supply determinants through  $\mathcal{E}^C$ . This would be the case if public good shocks were land improvements—new roads, pathways, etc.—of no intrinsic value but destined to welcome social housing units or private housing developments following changes in land use regulation. Residents would migrate towards municipalities experiencing positive housing supply shocks because of lower posted rents. More generally, if our subsidy shocks are correlated with shocks in the determinants of housing supply it may bias the interpretation of our estimates. As a first test to alleviate this concern, we can look back at the rents results of Figure 4.13. A positive housing supply shock would have a negative effect on rents. Our results show the exact opposite suggesting that people indeed value local public goods beyond any correlated shift in the housing supply curve.

A second concern is that migration responses may entirely be driven by the inflow of public employees necessary to operate the new facilities or services financed by the subsidy shocks. Indeed, our model does not account for public employment. Public goods may be of little value in themselves, but workers may react to public labor demand shocks that would increase wages. It is not conceptually a problem as we could have modelled public good provision as taking public employment as input. The identification of the partial effect of public good supply would still be achieved in the GMM procedure provided that we separately instrument changes in local wages. We nevertheless assess the importance of this channel and show evidence that the public employment effect is marginal. We look at how the share of public employees in the population—crudely measured as public staff payroll divided by total local payroll—evolves around subsidy shocks. Combining this estimate with total population responses and the pre-shock shares of public employees, we conclude that observed behavioral responses coming from public employment only explain approximately 12% of total 20–65 population response (see derivations in appendix C).



## 6 Structural Estimation

This section carries out a GMM estimation of the model's parameters. We derive moments conditions using expressions for changes in residential amenities set out in Section 3. We show that a sufficient condition for an instrument to be valid in this setup is to be uncorrelated with the endogenous variables of our model—city population, housing consumption, rents and wages. Section 5 makes the case for subsidy shocks as a valid instrument in our framework. We provide additional evidence that the endogenous variables of our model are mean-independent of subsidy shocks and ground our GMM estimation strategy on these more stringent conditions.

### 6.1 Generalized Method of Moments

As mentioned, we structurally estimate the models' parameters using a non-linear generalized method of moments. We ground our GMM estimation on moment conditions of the form

$$\mathbf{E} [\Delta \ln \mathcal{E}^A \times Z] = 0 \tag{4.23}$$

where  $Z$  is an instrument uncorrelated with changes in unobserved residential amenities  $\Delta \ln \mathcal{E}^A$ . Section 5 gives support for the following two moment conditions:

$$\begin{aligned} \mathbf{E} [\Delta \ln \bar{\mathcal{E}}_j^A \times \bar{Z}_j] &= 0 \\ \mathbf{E} [\Delta \ln \widehat{\mathcal{E}}_a^A \times \widehat{Z}_a] &= 0 \end{aligned} \tag{4.24}$$

where  $\bar{Z}_j$  (resp.  $\widehat{Z}_a$ ) is a variable of residuals obtained from regressing 2007, 2009 and 2010 subsidy shocks  $\Delta \ln \bar{S}_j$  (resp.  $\Delta \ln \widehat{S}_a$ ) on a vector of flexible dummies for 1999 baseline characteristics and lagged shocks as in regression (4.21) (resp. (4.22)). To see why, notice that equations (4.19) and (4.20) express changes in residential amenities as a sum of linear functions of observables. A sufficient condition for (4.24) is then  $\mathbf{E} [\Delta \ln \bar{Y}_j \times \bar{Z}_j] = 0$  and  $\mathbf{E} [\Delta \ln \widehat{Y}_a \times \widehat{Z}_a] = 0$  for all observables  $Y$  of (4.19) and (4.20). DiD evidence of Section 5 precisely suggest that these conditions hold.

In practice, the GMM procedure looks for the set of parameters that minimizes the empirical counterparts of our moment conditions, keeping the endogenous variables of the model at their observed values in the data. As such, the two moment conditions in (4.24) are not enough on their own to identify all the parameters of our model. Our GMM estimation relies on more moments conditions. In particular, we investigate whether more restrictive conditions of the form  $\mathbf{E} [\Delta \ln \mathcal{E}^A | Z] = 0$  hold. It would similarly be supported by evidence that the empirical counterpart of  $\mathbf{E} [\Delta \ln Y | Z] = 0$  holds for all outcomes  $Y$ . The next section provides non-parametric evidence of such relationship between subsidy shock intensity and cumulative outcome changes.

## 6.2 Non-Parametric Evidence

We provide further evidence on the absence of pre-trends by looking at the non-parametric relationship between subsidy shocks on the one hand, and pre- or post-shock cumulative outcome changes on the other hand. We run kernel regressions where the dependent variable is alternatively  $\Delta \ln \bar{G}_{j,-6,-2}$ ,  $\Delta \ln \bar{G}_{j,-2,4}$ ,  $\Delta \ln \bar{N}_{j,-6,-2}$  and  $\Delta \ln \bar{N}_{j,-2,4}$ . The explanatory variable is  $\bar{Z}_{j,-1,0}$ , i.e., the variable of residuals obtained from regressing 2007, 2009 and 2010 shocks  $\Delta \ln \bar{S}_{j,-1,0}$  on the set of baseline controls and lagged shocks.<sup>25</sup> Figure 4.15 presents the results.

The dependent variable in Panel A (resp. in Panel C) is  $\Delta \ln \bar{G}_{-6,-2}$  (resp.  $\Delta \ln \bar{N}_{-6,-2}$ ). Panels A and C show again that the average relationship between subsidy shock and pre-shock outcome growth is close to zero. In addition, they offer evidence that municipality expected outcomes grow at the same rate as their MF geometric average *conditional on shock intensity*. In mathematical terms, this translates into  $E [\Delta \ln \bar{Y}_{-6,-2} | \bar{Z}_{j,-1,0}] = 0$  for  $Y \in \{G, N\}$ . This

---

25. The fitted values and confidence bands are computed from running kernel regressions of the dependent variable on these initial residuals and on 1000 additional samples of residuals. We generate synthetic residuals using the wild cluster bootstrap procedure proposed in Cameron et al. (2008). We assume that errors are correlated within clusters which we take to be counties. Each cluster randomly draws a  $+1/-1$  coefficient with probability 0.5 and all residuals of a same cluster are multiplied by the same coefficient. These synthetic residuals are then added back to the original fitted values. We run the first-step regression again on this pseudo-sample and store the coefficient estimates. We do this 1000 times. We then generate 1000 pseudo-samples of residuals by fitting each saved model on the original data and saving the residuals. The graphs report the pivotal bootstrap confidence band and the bias-corrected fitted values.

mean-independence property will be central to our GMM analysis.

In Panels B and D, the dependent variables are  $\Delta \ln \bar{G}_{j,-2,4}$  and  $\Delta \ln \bar{N}_{j,-2,4}$ . They offer reassurance that the average effects reported in the Figures of Section 5 are not driven by outliers. Public good and migration responses are visible across the full distribution of subsidy shocks.

Figure 4.17 provides similar evidence on how  $\Delta \ln \hat{G}_a$  and  $\Delta \ln \hat{N}_a$  change with  $\hat{Z}_{a,-1,0}$ . Panels A and C show the absence of correlation between subsidy shock and pre-shock normalized outcome growth which similarly translates into  $E[\Delta \ln \hat{Y}_{-6,-2} | \hat{Z}_{a,-1,0}] = 0$  for  $Y \in \{G, N\}$ . Panels B and D offer additional evidence that responses to subsidy shocks are observed all along the distribution of shock intensity. Finally, Figure 4.19 similarly shows that housing prices  $\Delta \ln \hat{r}_a$  evolve in a similar fashion in different cells of shock intensity in pre-shock periods and that the price response is coming from the full range of shock values.

These graphs also show that changes in jurisdictions' populations and rents are not necessarily proportional to changes in local public goods. Our model supports these non-proportional responses. For example, jurisdictions' proportional responses to changes in local public goods may differ because of heterogeneity in housing supply elasticity or production technology for the traded good. In addition, jurisdictions' budget constraints introduce non-linearity in endogenous variables' responses to subsidy shocks. Importantly, these non-proportional responses are precisely what enables us to add more moment conditions to (4.24).

### 6.3 Moment Conditions

The non-parametric evidence in Section 6.2 showed that subsidy shocks likely satisfy mean-independence conditions of the form  $\mathbf{E}[\Delta \ln \mathcal{E} | Z] = 0$ . These are more restrictive than the usual conditions of zero correlation between instrument and unobserved fundamentals  $\mathbf{E}[\Delta \ln \mathcal{E} \times Z] = 0$ . Indeed, with mean-independence of the unobserved fundamentals with respect to our original instrument  $Z$ , any function of  $Z$  may be used as an additional instrument (see Wooldridge 2010). We apply this property to a discrete number of indicator functions as

in Ahlfeldt et al. (2015): we partition the empirical distribution of  $\bar{Z}$  (or  $\widehat{Z}$ ) into subintervals of equal range. We define  $\mathbb{I}_m^W$  (resp.  $\mathbb{I}_m^B$ ) the indicator function equal to one if  $\bar{Z}$  (resp.  $\widehat{Z}$ ) belongs to the subinterval  $m$  of the partition of  $\bar{Z}$  (resp. of  $\widehat{Z}$ ).

We can now state all our moment conditions:

$$\begin{aligned} \mathbf{E} \left[ \Delta \ln \overline{\mathcal{E}}_j^A \times \mathbb{I}_m^W \right] &= 0 \quad \text{for all } m \in \{1, \dots, M\} \\ \mathbf{E} \left[ \Delta \ln \overline{\mathcal{E}}_j^A \times \bar{Z}_j \right] &= 0 \end{aligned} \tag{4.25}$$

using the within-MF expression (4.19) for residential amenities and

$$\begin{aligned} \mathbf{E} \left[ \Delta \ln \widehat{\mathcal{E}}_a^A \times \mathbb{I}_m^B \right] &= 0 \quad \text{for all } m \in \{1, \dots, M\} \\ \mathbf{E} \left[ \Delta \ln \widehat{\mathcal{E}}_a^A \times \widehat{Z}_a \right] &= 0 \end{aligned} \tag{4.26}$$

using the between-MF expression (4.20) for residential amenities. These additional moments conditions provided by indicator variables  $\{\mathbb{I}_k^m, \mathbb{I}_m^B\}$  bring additional information for the GMM estimation only if they are not collinear. Section 6.2 provides such evidence and discusses the sources of non-linearity driving the endogenous variables' non-proportional responses.

All outcome changes plugged in (4.25) and (4.26) are the eight-year relative cumulative growths between year  $-4$  and year  $4$  around the 2007, 2009 and 2010 subsidy shocks. The identifying assumption is that subsidy shocks and indicator variables  $\{\mathbb{I}_k^m, \mathbb{I}_m^B\}$  are uncorrelated with jurisdiction or MF changes in unobserved residential amenities. As mentioned, DiD results of Section 5.2 and their decomposition in Section 6.2 suggest that this is indeed the case.

We simultaneously estimate the model's parameters using a two-step non-linear GMM procedure. The GMM estimation looks for the set of parameters that minimizes the empirical counterpart of moment conditions (4.25) and (4.26), keeping the endogenous variables of the model at their observed values in the data. As in Ahlfeldt et al. (2015), our moment conditions may not uniquely identify the model's parameters as the objective function might not be globally concave. In practice, the GMM objective function is well behaved and we find

a global minimum.

To make sure that the arbitrarily chosen number of subintervals for the partition of  $\bar{Z}$  and  $\hat{Z}$  does not affect our results, we try alternative specifications with different numbers of moments. However, the estimation requires a sufficient number of subintervals to pick up enough variation in jurisdiction responses to subsidy shocks. We report parameter estimates for specifications varying between 40 and 60 subintervals for each calibrated value of the housing consumption share.

## 6.4 Estimation Results

We report estimates of our structural parameters for different calibrated values of the housing consumption share  $\alpha$ . Calibrating the consumption share allows us to improve the robustness of our estimates given the multiplicative fashion in which it interacts with other parameters. Tables 4.5, 4.6, 4.7 and 4.8 report estimates of the model's parameters for  $\alpha \in \{0.25, 0.30, 0.35, 0.40\}$ .

We estimate a taste parameter for public goods  $\phi$  that varies between 0.114 and 0.252 according to the calibrated housing share. For a given value of  $\alpha$ , our estimate is robust for different number of moments. Given the Cobb-Douglas specification of preferences, the interpretation is that communities spend between 11% and 25% of their total resources on public goods. This estimate is typically increasing with the calibrated  $\alpha$ . Indeed, a higher taste for public goods is needed to compensate for given house price changes when the taste for housing is higher, holding constant migration and public good changes. Parameter  $\phi$  is the only parameter which is directly comparable with estimates from the existing literature, which we report in Table 4.9. Our estimates fall in the range of existing estimates.

Estimates for inverse household mobility  $\sigma$  lie between 0.03 and 0.14. These estimates are much lower than those found for instance in [Serrato and Wingender \(2011\)](#) or [Diamond \(2016\)](#). First, notice that the size of the considered French jurisdictions is much smaller than the geographic unit of these studies, typically the MSA. Mobility between locations is then expected to be different. Second, it is interesting to note that given modelling assumptions,

parameter  $\sigma$  is isomorphic to any combination of parameters  $\sigma - \tilde{\kappa}$  where  $\tilde{\kappa}$  would capture positive agglomeration externalities from increased density beyond those transiting through increased public goods and potential changes in wages. Hence, our estimated  $\sigma$  is actually capturing mobility frictions net of all non-public and non-productive agglomeration effects, e.g., endogenous residential amenities such as in [Diamond \(2016\)](#).

We find substantial public good spillovers between jurisdictions of a same municipal federation. Estimates for  $\delta$  all lie between 0 and 0.08. While they are typically decreasing with the calibrated  $\alpha$ , the relationship is not trivial. When taste for housing increases, lower benefit spillovers are needed to explain observed house price increases, i.e., location must matter for the enjoyment of public goods. However, higher congestion spillovers can also explain observed price changes, because additional congestion brought by new residents matters less than congestion caused by residents of neighboring jurisdictions. Our estimates show that the second effect prevails.

Recall that  $\delta = 0$  means full spillovers within a municipal federation, while  $\delta = 1$  means no spillovers. In line with our previous empirical evidence, our estimates suggest that spillovers are very strong in the examined French setting. A back-of-the-envelope calculation for the average municipal federation size—27 member municipalities—implies that local public spending in a municipality accounts for 4–11% of the bundle of local public goods enjoyed by its residents, and that local public spending in its average neighbor municipality—belonging to the same municipal federation—accounts for 3.2–3.5% of this bundle.<sup>26</sup> These large estimates are unsurprising given the large number of French jurisdictions and the fact that the only restriction to benefiting from local public goods in other municipalities is travel distance—there is almost no legal restriction such as zoning.

Our estimates for public good congestion are the least robust across specifications. However, we find significant estimates for  $\kappa\phi$  that are typically below estimates for  $\phi$ . This points to local public services being not fully rival and subject to economies of scale, and to the existence of other sources of local inefficiencies due to fiscal agglomeration effects. Indeed,

---

26. The weight put on own public goods is  $\delta + \frac{1}{|a|}(1 - \delta)$  where  $|a|$  is the number of member jurisdictions in the MF. We take the mean number of member jurisdictions, i.e., 27 for this computation.

these estimates suggest that the cost of providing public services increases less than one for one with population, making denser jurisdictions more attractive from a public amenity perspective.

## 7 Welfare Implications

Our empirical analysis documented strong public good spillovers across jurisdiction borders. The decentralized provision of public goods is likely to be sub-optimally low, i.e., there may be welfare gains from coordination. While current subsidies from counties, regions and the central government may carry some Pigovian flavour, it is unlikely that they fully tackle spillover inefficiency. In this section, we focus on a simplified version of our model and simulate the welfare impact of an administrative reform merging all jurisdictions belonging to a same municipal federation. As mentioned in Section 2, these groupings were historically introduced precisely to deal with spillovers and economies of scale. However, member jurisdictions still maintained a high degree of independence and it is a natural check to investigate whether this residual autonomy is at the root of significant deadweight loss.

We first back out the fundamentals of the model  $\{\mathcal{E}_j^A, \mathcal{E}_j^C\}$  from (i) equilibrium endogenous variables of the model that we observe as econometricians, and (ii) parameters that we have estimated or calibrated from the literature. We then make several assumptions to ensure tractability of the equilibrium in our simplified setup. Motivated by the suppression of the French local business tax in 2011, we first assume that local public revenues only come from the residence tax. We make five others assumptions: (i) residents are myopic and do not anticipate migration responses when they vote for public goods and taxes; (ii) the public good production function is the identity, i.e.,  $\Gamma(x) = x$ ; (iii) housing supply elasticities  $\eta_j$  are constant across jurisdictions; (iv) local productivity  $\theta_j^Y$  does not depend on local public goods and (v) the geographic fundamentals that we have backed out still define the geographic landscape in our simplified model. This version of the model has a unique and tractable equilibrium given parameters  $\{\alpha, \kappa, \sigma, \phi, \delta, \eta\}$  and geographic fundamentals  $\{\mathcal{E}_j^A, \mathcal{E}_j^C\}$ .

We compare equilibrium situations before and after the merger. In both hypothetical situations, public subsidies are absent. While this setting does not enable us to evaluate a reform that applies to the current observed situation, it helps illustrate the magnitude of potential welfare gains.

Prior to the reform, residents vote for public goods as in Section 3.6 absent subsidies and business taxation. They choose the housing tax rate  $\tau_j^h$  and the level of public good  $G_j$  that maximize  $v_j$  under the budget constraint  $r_j N_j h_j \tau_j^h = G_j$ . Following the merger, residents vote for a new level of public spending pooling tax bases of all former member jurisdictions. To keep welfare comparable between pre- and post-reform situations, we assume that prior jurisdiction blocks still exist for the purpose of labor and housing markets. The only change is the degree of cooperation in the provision of local public goods, now decided at the  $a$  level for all former jurisdictions of  $a$ . Residents hence choose a housing tax  $\tau_a^h$  and a level of public good  $G_a$  that maximize  $v_j$  for  $j$  in  $a$  under the budget constraint  $G_a = \sum_{j' \in a} r_{j'} N_{j'} h_{j'} \tau_a^h$ .<sup>27</sup> We assume that the total amount of local public goods is then allocated to all jurisdictions of  $a$  in proportion to their pre-reform share of the total quantity of public good in the municipal federation. This allocation rule allows us to center our welfare analysis on efficiency gains and neutralizes the redistributive channel that would be present with, say, an egalitarian allocation. Our analysis would also work with alternative allocation rules but the welfare channels would be more intricate.

We note  $X^o$  all variables observed in 2014, that we assume are the equilibrium variables of the model with subsidies. In the model without public subsidies, we call  $X^n$  the pre-reform equilibrium variables and  $X^m$  the post-reform equilibrium variables. Absent subsidies, local public good supply before the reform equals

$$\begin{aligned}
 1 + \tau_j^n &= 1 + \frac{\phi(\delta + \frac{1-\delta}{|a_j|})}{(1-\phi)\alpha} \\
 G_j^m [1 + \tau_j^n] &= \frac{\varphi^\varphi (1-\varphi)^{1-\varphi}}{\zeta^\varphi} \tau_j^n N_j^n w_j^n \alpha (1 - \tau^w)
 \end{aligned} \tag{4.27}$$

---

27. The result of this optimization problem is the same for any jurisdiction  $j$  in  $a$ , see Appendix D.



After the merger, local public good supply is

$$1 + \tau_a^m = 1 + \frac{\phi}{(1 - \phi)\alpha}$$

$$G_j^m = \frac{\varphi^\varphi(1 - \varphi)^{1-\varphi}}{\zeta^\varphi} \frac{G_j^n}{\sum_{j' \in a_j} G_{j'}^n} \sum_{j' \in a_j} \frac{\tau_{j'}^m}{1 + \tau_{j'}^m} N_{j'}^m w_{j'} \alpha (1 - \tau^w) \quad (4.28)$$

Using backed out fundamentals, we can express all endogenous variables in these two simulated cases as functions of 2014 observables. We assume that social welfare is defined as <sup>28</sup>

$$W = \mathbb{E} \left[ \max_j \ln U(i, j) \right] \quad (4.29)$$

Because idiosyncratic preferences are distributed Extreme Value Type-I, the welfare change associated with the reform equals

$$\begin{aligned} \Delta W &= W^m - W^n \\ &= \sigma \ln \left( \frac{\sum_j e^{v_j^m/\sigma}}{\sum_j e^{v_j^n/\sigma}} \right) \end{aligned} \quad (4.30)$$

In the model, utility is homogeneous to euros so we interpret these welfare changes as percentage changes of a money metric. All detailed equations are given in Appendix D.

In our baseline scenario, we simulate the welfare change taking  $\alpha = 0.3$  for the housing consumption share and  $\eta = 0.2$  for the housing supply elasticity. These two values correspond to the literature's central estimates for France (see Table 4.9). For each parameter of  $\{\kappa, \sigma, \phi, \delta\}$ , we use its average estimation across specifications of Section 6. Because this simulation is based on estimated parameters and fundamentals, we report the associated Monte Carlo standard errors that we compute using the parameters' estimated variance-covariance structure.

In our central specification, we estimate that a reform that would fully merge jurisdictions at the existing federation level would increase welfare by 60%. This welfare gain is significant

---

28. The choice of a utilitarian welfare criterion is akin to considering that policy choices are made behind the "veil of ignorance". Increasing welfare is then equivalent to increasing the expected utility of ex-ante homogeneous agents. Evaluating the proposed reform through the lens of an ex-post Pareto criterion will likely change its desirability.

at the 1% level. Tables 4.10 and 4.11 report welfare change estimates for alternative values of  $\eta$  and  $\alpha$ . They are largely robust across calibration for  $\eta$ . However they vary between 40% and 137% when we plug alternative housing consumption shares holding  $\eta = 0.20$ . This is expected as our GMM estimates for spillovers and taste for public services are increasing with the calibrated  $\alpha$ , which makes welfare gains from a coordination reform increasing with  $\alpha$ .

This strong result is unsurprising given our spillover estimates, but should be interpreted as an upper bound. First, it is obtained in a hypothetical context without public subsidies. Because current subsidies may already be correcting some of the externalities, we expect that the welfare gains from a reform departing from the current situation and holding subsidies constant would be smaller. We also emphasize that by abstracting from the political deadweight loss that may arise under a more centralized regime, our analysis only investigated one side of the centralization efficiency trade-off. More work is needed to finely assess how much the inability to tailor policies to local needs as well as other potential political frictions would decrease overall welfare gains.

## 8 Conclusion

This paper develops a simple yet flexible framework to test for potential welfare gains from centralization of public goods provision. We first build a spatial equilibrium model of local jurisdictions with endogenous public goods. The binary structure for spillover spatial decay makes it easily amenable to empirical analysis. Our model shows how one can exploit differential behavioral responses to shocks in local public goods at different geographic levels to uncover the intensity of local spillovers.

We then bring new insights on spatial spillovers in the fragmented French institutional context by providing reduced-form evidence of migration and house price responses to changes in local public goods. We then estimate our model with GMM and document substantial public good spillovers, corroborating our reduced-form evidence. A jurisdiction's public goods account for approximately 4–11% of the local public good bundle enjoyed by its residents

and public spending in each neighbor municipality approximately accounts for an average 3.2–3.5% of this bundle.

In a final exercise, we simulate the effect of a reform redefining jurisdiction administrative boundaries in a simpler version of our model. Although we do not estimate the cost of centralization, our results suggest that increased coordination in the provision of public services may substantially improve welfare.

TABLE 4.1 – DESCRIPTIVE STATISTICS ON JURISDICTIONS

Variable	Mean	St. Dev.
Current spending <sup>a</sup>	1,517,088	23,022,598
Current spending per adult <sup>b</sup>	1,111	1,786
Investment stock <sup>a</sup>	9,383,611	98,714,400
Investment stock per adult <sup>b</sup>	13,481	24,130
Population <sup>c</sup>	1,746	14,614
Population 20–65 <sup>d</sup>	994	4,573
Housing price per square meter <sup>e</sup>	1,570	978
Net income per adult <sup>d</sup>	10,848	3,614
Housing service per jurisdiction <sup>d</sup>	16,247	83,086
Housing surface per adult <sup>d</sup>	39	6

*Note:* This table gives the averages and standard deviations across jurisdictions for some of section 5 constructed variables in 2009 or the closest available year.

- 
- a.* 2009 jurisdiction accounts, author’s calculations
  - b.* 2009 jurisdiction accounts, 2008 Census data, author’s calculations
  - b.* 2009 jurisdiction accounts, 2009 FILOCOM, author’s calculations
  - d.* 2009 FILOCOM, author’s calculations
  - c.* 2008 Census data, author’s calculations
  - e.* 2008 Notaries databases, author’s calculations

TABLE 4.2 – WITHIN-MF SENSITIVITY ANALYSIS

	(1)	(2)	(3)
	$\Delta \ln \bar{G}_j$	$\Delta \ln \bar{G}_j$	$\Delta \ln \bar{G}_j$
$\Delta \ln \bar{S}_j$	0.363*** (0.009)	0.372*** (0.009)	0.369*** (0.009)
	$\Delta \ln \bar{N}_j$	$\Delta \ln \bar{N}_j$	$\Delta \ln \bar{N}_j$
$\Delta \ln \bar{S}_j$	0.033*** (0.007)	0.032*** (0.008)	0.026*** (0.008)
	$\Delta \ln \bar{w}_j$	$\Delta \ln \bar{w}_j$	$\Delta \ln \bar{w}_j$
$\Delta \ln \bar{S}_j$	0.002 (0.003)	0.003 (0.003)	0.001 (0.003)
	$\Delta \ln \bar{h}_j$	$\Delta \ln \bar{h}_j$	$\Delta \ln \bar{h}_j$
$\Delta \ln \bar{S}_j$	-0.006 (0.003)	-0.006 (0.004)	-0.007 (0.004)
	$\Delta \ln \bar{\mathcal{T}}_j$	$\Delta \ln \bar{\mathcal{T}}_j$	$\Delta \ln \bar{\mathcal{T}}_j$
$\Delta \ln \bar{S}_j$	0.002 (0.003)	0.002 (0.003)	0.001 (0.003)
Lag $\Delta \ln \bar{S}_j$		Yes	Yes
Baseline Controls			Yes
Observations	99,593	99,593	99,593

*Note:* This Table reports estimates for  $\beta_{d=4}^W$  i.e. the effect of the shock happening in  $d = 0$  on cumulative outcome growth between  $d = -6$  and  $d = 4$ . Lag shocks include all past shocks up to  $d = -1$ . We flexibly control for baseline 1999 jurisdiction population and density. Standard errors are clustered at the MF level. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

TABLE 4.3 – BETWEEN-MF SENSITIVITY ANALYSIS

	(1)	(2)	(3)
	$\Delta \ln \widehat{G}_a$	$\Delta \ln \widehat{G}_a$	$\Delta \ln \widehat{G}_a$
$\Delta \ln \widehat{S}_a$	0.258*** (0.048)	0.416*** (0.055)	0.466*** (0.055)
	$\Delta \ln \widehat{N}_a$	$\Delta \ln \widehat{N}_a$	$\Delta \ln \widehat{N}_a$
$\Delta \ln \widehat{S}_a$	0.094** (0.038)	0.152*** (0.042)	0.148*** (0.045)
	$\Delta \ln \widehat{w}_a$	$\Delta \ln \widehat{w}_a$	$\Delta \ln \widehat{w}_a$
$\Delta \ln \widehat{S}_a$	0.008 (0.017)	0.055*** (0.019)	0.050** (0.020)
	$\Delta \ln \widehat{h}_a$	$\Delta \ln \widehat{h}_a$	$\Delta \ln \widehat{h}_a$
$\Delta \ln \widehat{S}_a$	-0.036** (0.016)	-0.031 (0.018)	-0.030 (0.019)
	$\Delta \ln \widehat{\mathcal{T}}_a$	$\Delta \ln \widehat{\mathcal{T}}_a$	$\Delta \ln \widehat{\mathcal{T}}_a$
$\Delta \ln \widehat{S}_a$	-0.017 (0.015)	-0.019 (0.016)	-0.023 (0.016)
Lag $\Delta \ln \widehat{S}_a$	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes
Time FE X	County	Region	Nation
Observations	5,428	5,428	5,428

*Note:* This Table reports estimates for  $\beta_{d=4}^B$  i.e. the effect of the shock happening in  $d = 0$  on cumulative outcome growth between  $d = -6$  and  $d = 4$ . Lag shocks include all past shocks up to  $d = -1$ . We flexibly control for baseline 1999 jurisdiction population and density. In the model, time fixed effects in the residuals are assumed to be uniform at the national level hence the division of endogenous variables by their national geometric mean. Here, we allow for division by either national, regional or county geometric mean to account for year, year X region or year X county fixed effects in the structural residuals. Standard errors are clustered at the MF level. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

TABLE 4.4 – HOUSING SUPPLY SENSITIVITY ANALYSIS

	(1)	(2)	(3)
	$\Delta \ln \bar{H}_j$	$\Delta \ln \bar{H}_j$	$\Delta \ln \bar{H}_j$
$\Delta \ln \bar{S}_j$	0.027*** (0.007)	0.026*** (0.007)	0.020*** (0.007)
	$\Delta \ln \bar{r}_j$	$\Delta \ln \bar{r}_j$	$\Delta \ln \bar{r}_j$
$\Delta \ln \bar{S}_j$	-0.026 (0.025)	0.015 (0.017)	0.016 (0.017)
Lag $\Delta \ln \bar{S}_j$		Yes	Yes
Baseline Controls			Yes
Observations	99,593	99,593	99,593
	$\Delta \ln \hat{H}_a$	$\Delta \ln \hat{H}_a$	$\Delta \ln \hat{H}_a$
$\Delta \ln \hat{S}_a$	0.073** (0.035)	0.125*** (0.039)	0.132*** (0.041)
	$\Delta \ln \hat{r}_a$	$\Delta \ln \hat{r}_a$	$\Delta \ln \hat{r}_a$
$\Delta \ln \hat{S}_a$	0.173** (0.072)	0.246*** (0.073)	0.239*** (0.075)
Lag $\Delta \ln \hat{S}_a$	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes
Time FE X	County	Region	Nation
Observations	5,428	5,428	5,428

*Note:* This Table reports estimates for  $\beta_{d=4}$  i.e. the effect of the shock happening in  $d = 0$  on cumulative outcome growth between  $d = -6$  and  $d = 4$ . Lag shocks include all past shocks up to  $d = -2$ . Controls include baseline 1999 jurisdiction population and density. Standard errors are clustered at the MF level. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

TABLE 4.5 – STRUCTURAL PARAMETERS FOR  $\alpha = 0.25$

# Subintervals	40	42	44	46	48	50	52	54	56	58	60
Spillovers ( $\delta$ )	0.077	0.051	0.071	0.024	0.029	0.034	0.063	0.056	0.031	0.063	0.042
	0.011	0.005	0.008	0.009	0.012	0.014	0.007	0.007	0.017	0.007	0.009
PG Taste ( $\phi$ )	0.114	0.150	0.139	0.142	0.147	0.138	0.155	0.175	0.153	0.163	0.148
	0.007	0.005	0.005	0.004	0.009	0.012	0.006	0.007	0.009	0.012	0.007
Mobility ( $\sigma$ )	0.113	0.103	0.142	0.069	0.065	0.079	0.125	0.124	0.080	0.118	0.099
	0.009	0.011	0.006	0.008	0.014	0.012	0.007	0.005	0.013	0.006	0.004
PG Congestion ( $\kappa\phi$ )	0.056	0.113	0.065	0.134	0.166	0.118	0.094	0.131	0.174	0.144	0.061
	0.013	0.021	0.012	0.012	0.022	0.027	0.017	0.010	0.024	0.016	0.038

*Note:* This Table reports our structural parameters estimation for  $\alpha = 0.25$ . Standard errors are computed from a wild cluster bootstrap procedure based on 1,000 replications.

TABLE 4.6 – STRUCTURAL PARAMETERS FOR  $\alpha = 0.30$

# Subintervals	40	42	44	46	48	50	52	54	56	58	60
Spillovers ( $\delta$ )	0.073	0.037	0.053	0.016	0.028	0.031	0.056	0.041	0.015	0.052	0.035
	0.009	0.005	0.004	0.007	0.013	0.009	0.008	0.004	0.014	0.008	0.006
PG Taste ( $\phi$ )	0.132	0.178	0.162	0.165	0.165	0.154	0.173	0.194	0.183	0.190	0.168
	0.004	0.009	0.009	0.005	0.007	0.011	0.009	0.007	0.012	0.012	0.007
Mobility ( $\sigma$ )	0.104	0.083	0.112	0.048	0.046	0.069	0.104	0.098	0.067	0.105	0.080
	0.007	0.013	0.006	0.008	0.013	0.009	0.009	0.007	0.016	0.010	0.006
PG Congestion ( $\kappa\phi$ )	0.063	0.152	0.100	0.160	0.168	0.138	0.109	0.166	0.197	0.161	0.064
	0.010	0.023	0.013	0.018	0.023	0.025	0.013	0.011	0.021	0.024	0.024

*Note:* This Table reports our structural parameters estimation for  $\alpha = 0.30$ . Standard errors are computed from a wild cluster bootstrap procedure based on 1,000 replications.



TABLE 4.7 – STRUCTURAL PARAMETERS FOR  $\alpha = 0.35$

# Subintervals	40	42	44	46	48	50	52	54	56	58	60
Spillovers ( $\delta$ )	0.047	0.027	0.043	0.000	0.000	0.025	0.037	0.040	0.012	0.025	0.015
	0.008	0.005	0.005	0.008	0.007	0.006	0.006	0.010	0.007	0.006	0.010
PG Taste ( $\phi$ )	0.155	0.188	0.179	0.200	0.196	0.177	0.194	0.211	0.193	0.219	0.188
	0.009	0.010	0.005	0.007	0.005	0.015	0.006	0.013	0.010	0.008	0.007
Mobility ( $\sigma$ )	0.092	0.071	0.088	0.031	0.014	0.063	0.092	0.086	0.045	0.071	0.065
	0.010	0.010	0.006	0.011	0.013	0.007	0.006	0.011	0.008	0.008	0.005
PG Congestion ( $\kappa\phi$ )	0.092	0.156	0.123	0.200	0.225	0.176	0.154	0.179	0.229	0.187	0.142
	0.023	0.020	0.017	0.013	0.009	0.043	0.019	0.019	0.015	0.012	0.033

*Note:* This Table reports our structural parameters estimation for  $\alpha = 0.35$ . Standard errors are computed from a wild cluster bootstrap procedure based on 1,000 replications.

TABLE 4.8 – STRUCTURAL PARAMETERS FOR  $\alpha = 0.40$

# Subintervals	40	42	44	46	48	50	52	54	56	58	60
Spillovers ( $\delta$ )	0.035	0.019	0.037	0.007	-0.001	0.014	0.025	0.027	0.017	0.023	0.017
	0.005	0.005	0.007	0.004	0.004	0.008	0.003	0.009	0.008	0.007	0.004
PG Taste ( $\phi$ )	0.181	0.221	0.183	0.214	0.221	0.205	0.230	0.254	0.232	0.252	0.228
	0.009	0.008	0.009	0.006	0.006	0.021	0.009	0.010	0.012	0.012	0.007
Mobility ( $\sigma$ )	0.070	0.050	0.074	0.033	-0.003	0.046	0.078	0.063	0.042	0.063	0.051
	0.007	0.014	0.005	0.007	0.012	0.015	0.005	0.008	0.015	0.011	0.010
PG Congestion ( $\kappa\phi$ )	0.126	0.217	0.140	0.230	0.252	0.156	0.193	0.226	0.262	0.171	0.163
	0.017	0.030	0.022	0.011	0.027	0.041	0.016	0.013	0.017	0.027	0.033

*Note:* This Table reports our structural parameters estimation for  $\alpha = 0.40$ . Standard errors are computed from a wild cluster bootstrap procedure based on 1,000 replications.

TABLE 4.9 – STRUCTURAL PARAMETERS CALIBRATION AND ESTIMATION

Structural parameters	Values from literature	Estimation/calibration													
$\alpha$	Diamond (2016) <sup>a</sup> 43% for US non college workers and 46% for US college workers	from literature													
	Combes et al. (2018) <sup>c</sup> : 0.314 for homeowners and 0.352 for renters, and more precisely														
	<table border="1"> <thead> <tr> <th></th> <th>Paris</th> <th>Lyon, Lille, Marseille</th> <th>Pop &gt; 200 000</th> <th>Pop ≤ 200 000</th> </tr> </thead> <tbody> <tr> <td>homeowners</td> <td>0.344</td> <td>0.344</td> <td>0.304</td> <td>0.293</td> </tr> <tr> <td>renters</td> <td>0.369</td> <td>0.367</td> <td>0.382</td> <td>0.285</td> </tr> </tbody> </table>			Paris	Lyon, Lille, Marseille	Pop > 200 000	Pop ≤ 200 000	homeowners	0.344	0.344	0.304	0.293	renters	0.369	0.367
	Paris	Lyon, Lille, Marseille	Pop > 200 000	Pop ≤ 200 000											
homeowners	0.344	0.344	0.304	0.293											
renters	0.369	0.367	0.382	0.285											
$\phi$	Fajgelbaum et al. (2015): on US data														
	Serrato and Wingender (2011) <sup>b</sup> : [0.391,0.502] for US non college workers and [0.228,0.267] for US college workers														
	Fajgelbaum et al. (2015) <sup>c</sup> : [0.11,0.23]														
	Serrato and Wingender (2016) <sup>d</sup> : 0.26														
$\sigma$	Diamond (2016) <sup>a</sup> : 0.03 for US non college workers, 0.32 for US college workers	range from literature													
	Serrato and Wingender (2011) <sup>b</sup> : US non college workers [0.342, 0.399] and for US college workers [0.350, 0.376]														
$\eta$	Diamond (2016) <sup>a</sup> : 0.21 with a standard deviation of 0.22	(1) GMM estimation or (2) range from literature													
	Serrato and Wingender (2011) <sup>b</sup> : [0.407,0.813]														
	France Combes et al. (2018) <sup>c</sup> : 0.208 with most alternative estimates being between 0.15 and 0.30 seminal work Saiz (2010) us data														
$\zeta$		calibration with our data													
$\varphi$		calibration with our data													
$\kappa$		estimation with GMM													
$\delta$		estimation with GMM													

a.  $\alpha$ : the author uses US CEX survey data.

$\phi$ : for unskilled workers  $1.012/2.116 = \frac{\phi}{1-\phi}$ , for skilled workers. This is a structural parameter for valuation by workers of all amenities compared to national good. It is not local public good specific.  $0.274/4.026 = \frac{\phi}{1-\phi}$

$\sigma$ : We take  $\sigma$  as the inverse of the structural wage coefficient in the favorite specification (3) of the author. More precisely, it gives for US non college workers  $1/0.4026=0.24$  and US college workers  $1/2.116=0.47$

$\eta$ : inverse housing supply elasticity of 0.21 with standard deviation of 0.22

b.  $\eta=0.813$  when using housing prices,  $=0.407$  when using rents. Author’s preferred specification is non linear.

c. No parameters for  $\sigma$  assume the distribution of idiosyncratic shocks to follow a Frechet distribution

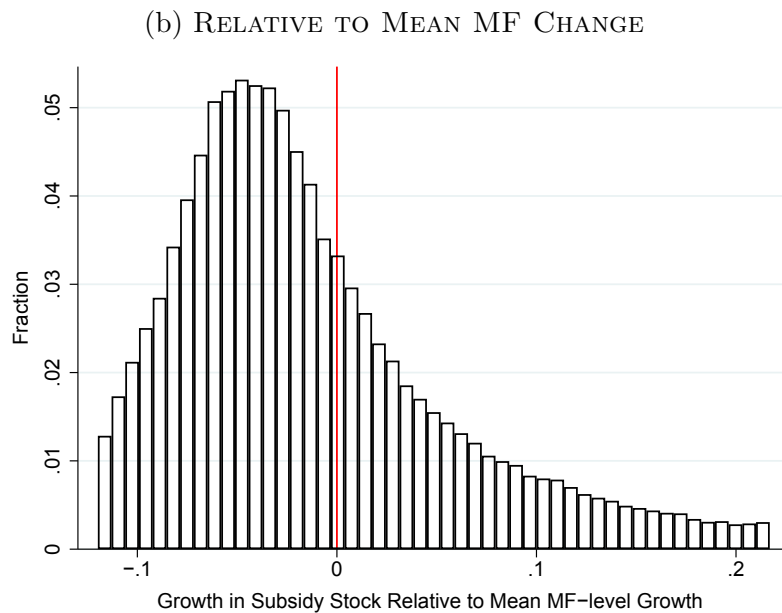
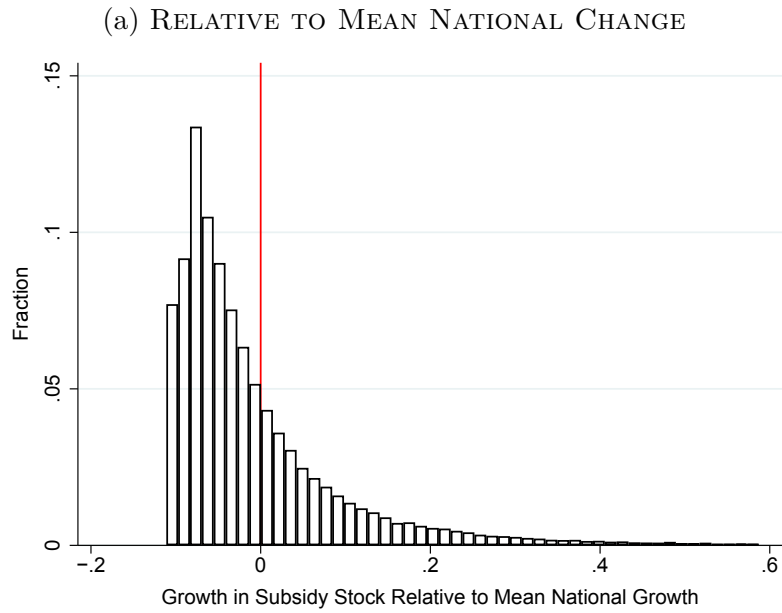
d. from Fajgelbaum et al. (2015) literature review on structural parameters. See their Table A.17.

e. Use Family expenditure survey for  $\alpha$  values

TABLE 4.10 – WELFARE CHANGE ESTIMATION FOR  $\alpha=0.30$

$\eta$	$\Delta W$ mean	$\Delta W$ sd
0.15	60%	19%
0.20	63%	21%
0.25	66%	23%

*Note:* This Table reports the welfare impact of a merger of all French jurisdictions at the municipal federation level in our simplified model version. Estimation for  $\alpha = 0.30$ . Standard errors are computed based on 1,000 Monte Carlo simulations.



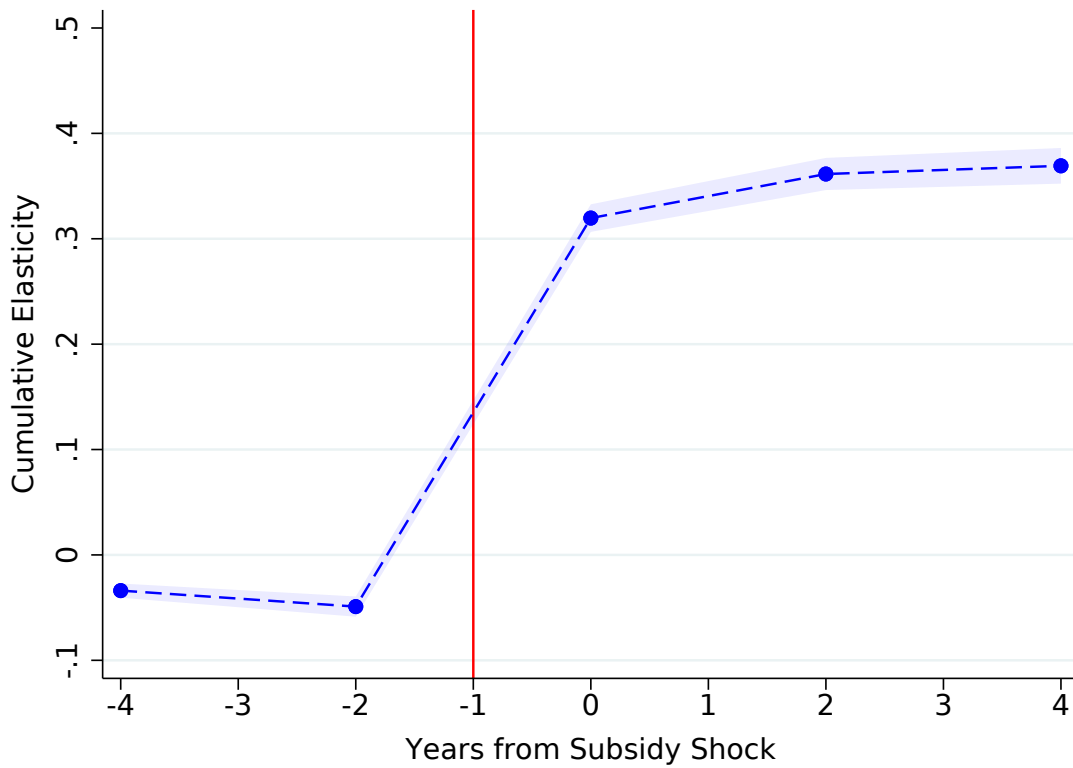
*Note:* These histograms plot the distribution of the yearly growth in normalized subsidy stocks pooling years 2007, 2009 and 2010. Panel A normalizes each city-level observation by the national geometric mean. Panel B normalizes each city-level observation by the geometric mean of all cities belonging to the city's MF. In each Panel, top and bottom 1% observations are censored for exposition purposes.

FIGURE 4.1 – DISTRIBUTION OF SUBSIDY STOCK YEARLY GROWTH

TABLE 4.11 – WELFARE CHANGE ESTIMATION FOR  $\eta = 0.20$

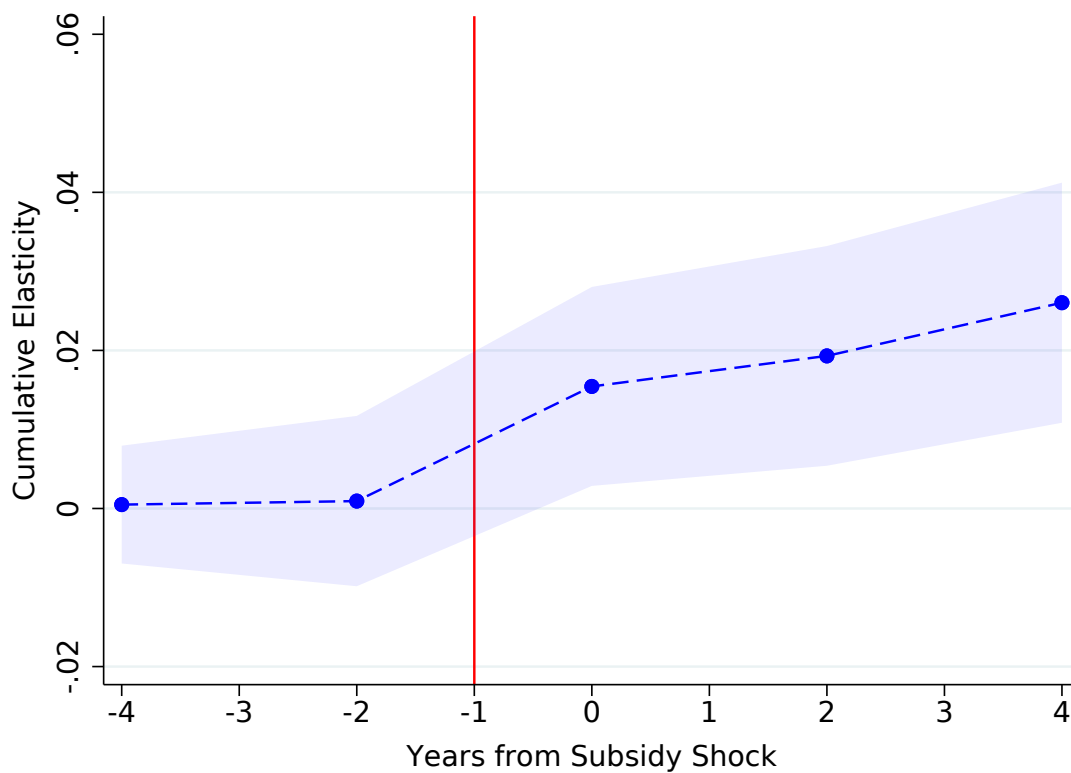
$\alpha$	$\Delta W$ mean	$\Delta W$ sd
0.25	40%	13%
0.30	63%	21%
0.35	100%	29%
0.40	137%	35%

*Note:* This Table reports the welfare impact of a merger of all French jurisdictions at the municipal federation level in our simplified model version. Estimation for  $\eta = 0.20$ . Standard errors are computed based on 1,000 Monte Carlo simulations.



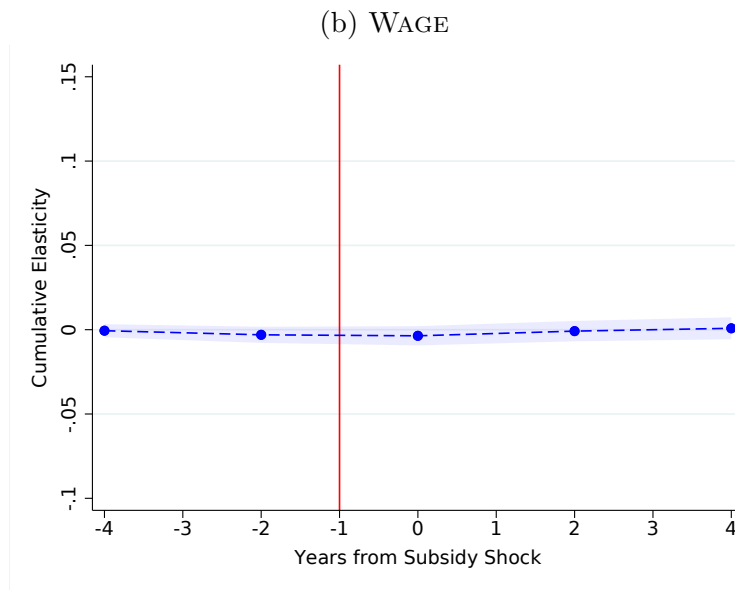
*Note:* This graph plots the coefficients  $\beta_d$  of Section 5 regressions. It shows the effect of changes in a city's subsidy stock in year 0 on cumulative changes in local public good starting 6 years prior to the shock year. All changes are relative to mean changes in the MF. Standard errors are clustered at the MF level. We report the 5% confidence bands.

FIGURE 4.3 – WITHIN-MF PUBLIC GOOD CHANGES



*Note:* This graph plots the coefficients  $\beta_d$  of Section 5 regressions. It shows the effect of changes in a city's subsidy stock in year 0 on cumulative changes in the number of residents aged 20–65 starting 6 years prior to the shock year. All changes are relative to mean changes in the MF. Standard errors are clustered at the MF level. We report the 5% confidence bands.

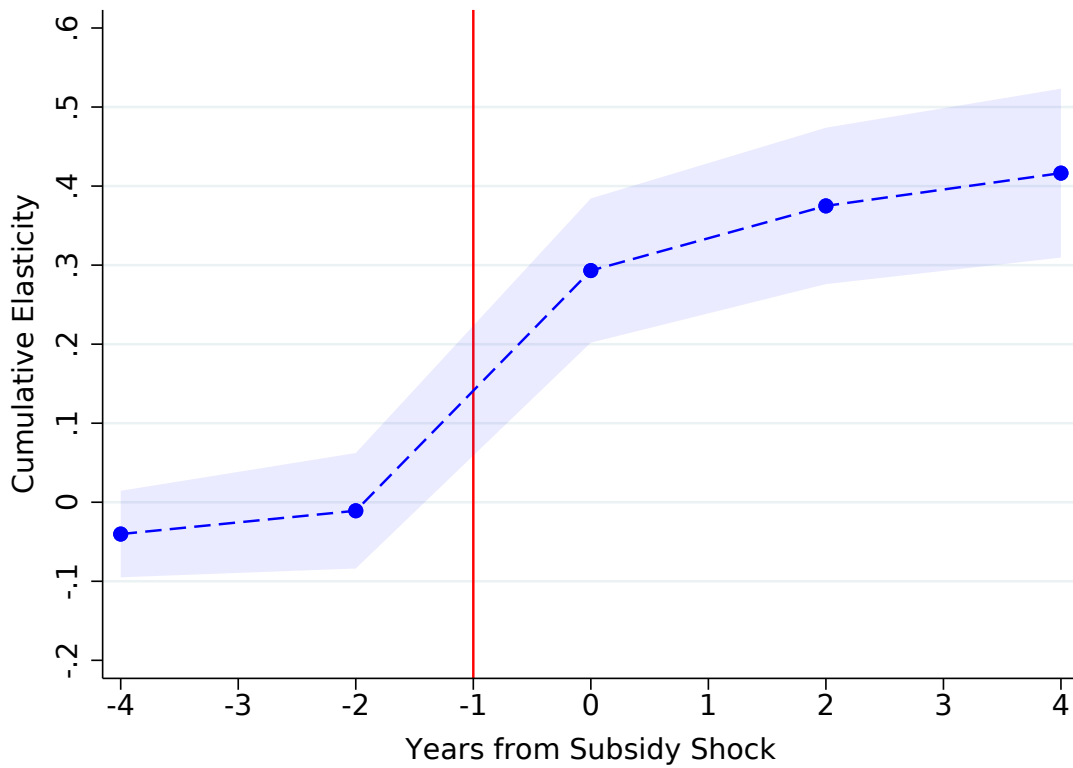
FIGURE 4.4 – WITHIN-MF MIGRATION RESPONSE



*Note:* These graphs plot the coefficients  $\beta_d$  of Section 5 regressions. They show the effect of changes in a city's subsidy stock in year 0 on cumulative changes in per capita housing consumption, wages and house prices starting 6 years prior to the shock year. All changes are relative to mean changes in the MF. Standard errors are clustered at the MF level. We report the 5% confidence bands.

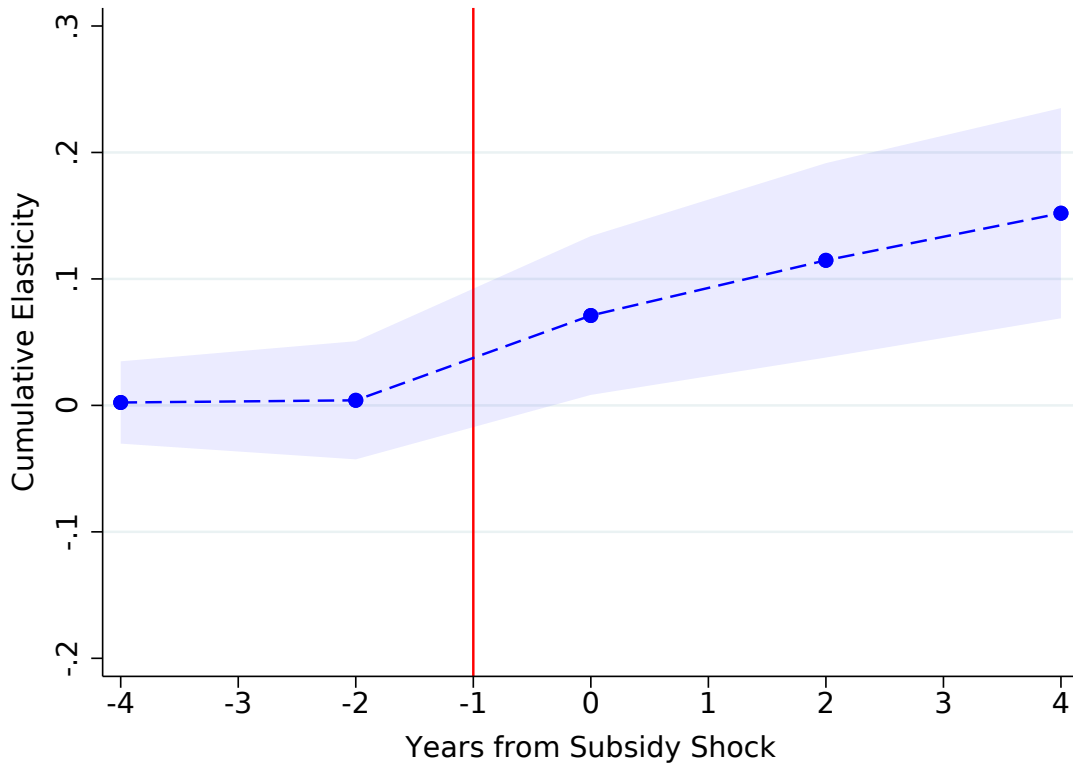
FIGURE 4.5 – WITHIN-MF PER CAPITA HOUSING AND WAGE





*Note:* This graph plots the coefficients  $\beta_d$  of Section 5 regressions. It shows the effect of changes in a MF subsidy stock in year 0 on cumulative changes in local public goods starting 6 years prior to the shock year. All changes are relative to regional mean changes. Standard errors are clustered at the MF level. We report the 5% confidence bands.

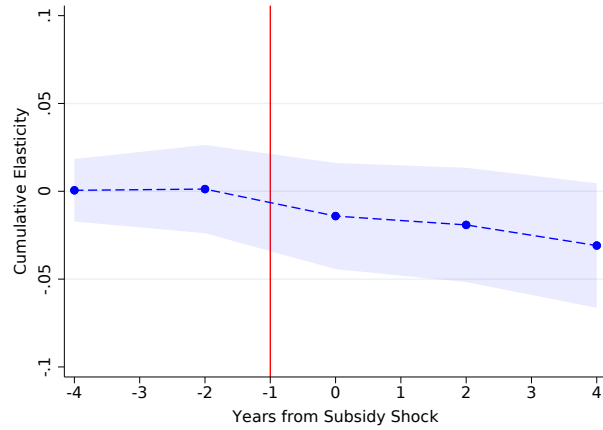
FIGURE 4.7 – BETWEEN-MF LOCAL PUBLIC GOOD CHANGES



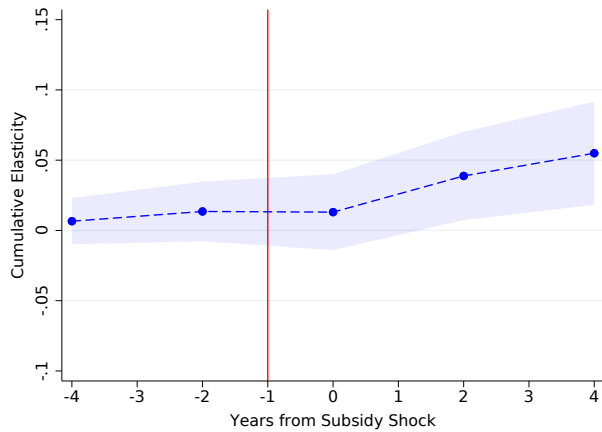
*Note:* This graph plots the coefficients  $\beta_d$  of Section 5 regressions. It shows the effect of changes in a MF subsidy stock in year 0 on cumulative changes in population starting 6 years prior to the shock year. All changes are relative to regional mean changes. Standard errors are clustered at the MF level. We report the 5% confidence bands.

FIGURE 4.8 – BETWEEN-MF MIGRATION RESPONSE

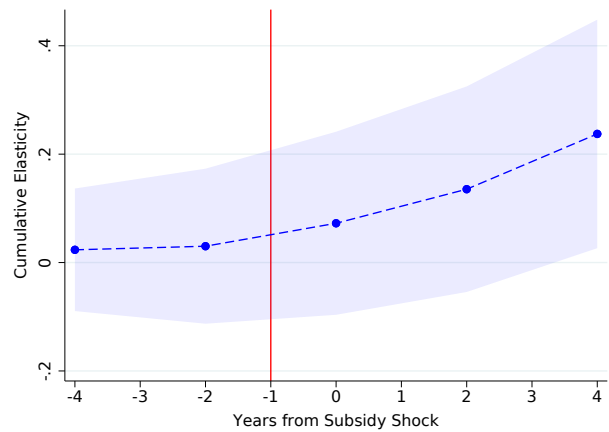
(a) PER CAPITA HOUSING CONSUMPTION (IN  $m^2$ )



(b) WAGE

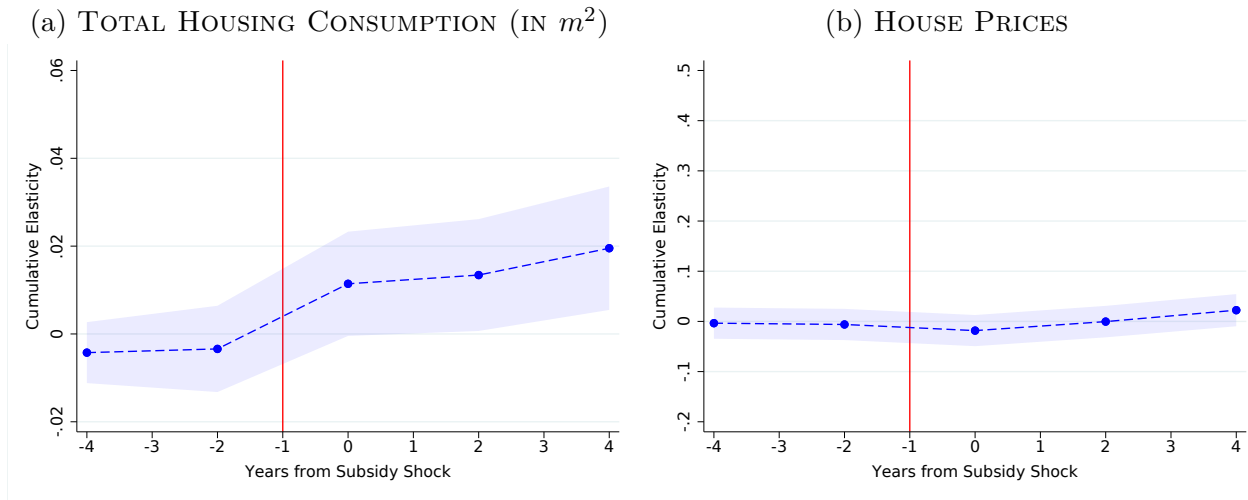


(c) RECORDED BUSINESSES



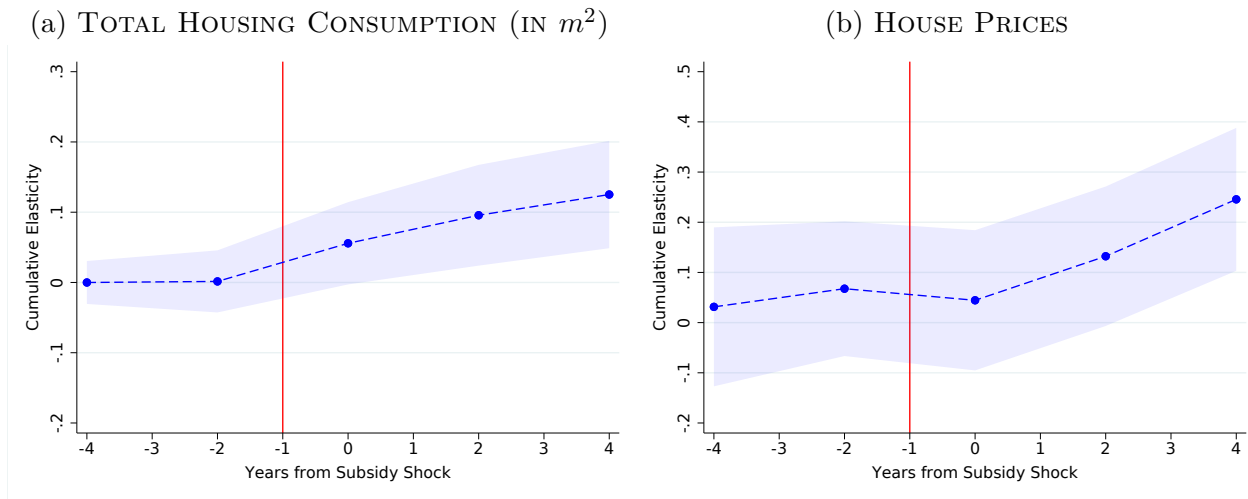
*Note:* These graphs plot the coefficients  $\beta_d$  of Section 5 regressions. They show the effect of changes in a MF subsidy stock in year 0 on cumulative changes in per capita housing consumption, wages and house prices starting 6 years prior to the shock year. All changes are relative to regional mean changes. Standard errors are clustered at the MF level. We report the 5% confidence bands.

FIGURE 4.9 – BETWEEN-MF PER CAPITA HOUSING AND WAGE



*Note:* These graphs plot the coefficients  $\beta_d$  of Section 5 regressions. They show the effect of changes in a city’s subsidy stock in year 0 on cumulative changes in total  $m^2$  of housing consumed starting 6 years prior to the shock year. All changes are relative to MF mean changes. Standard errors are clustered at the MF level. We report the 5% confidence bands.

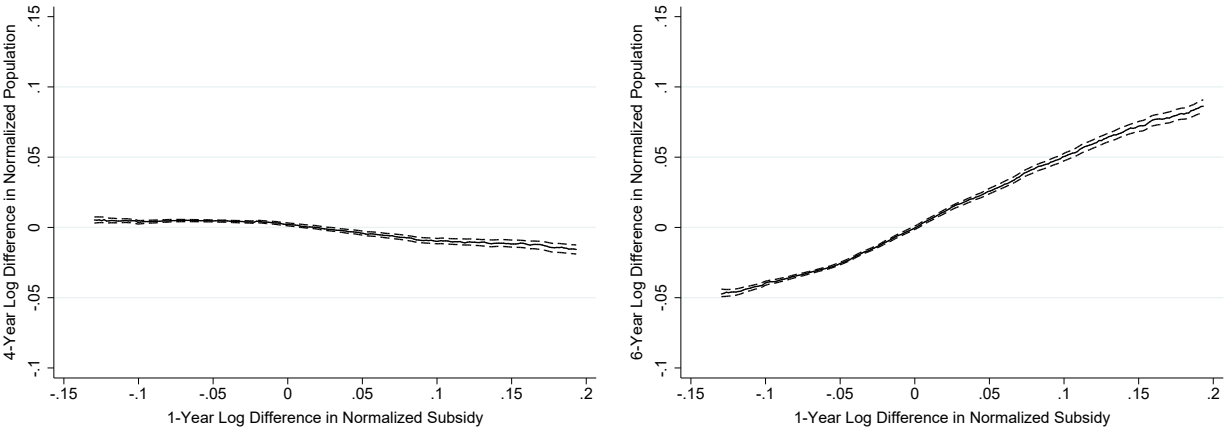
FIGURE 4.11 – WITHIN-MF HOUSING SUPPLY AND HOUSE PRICES



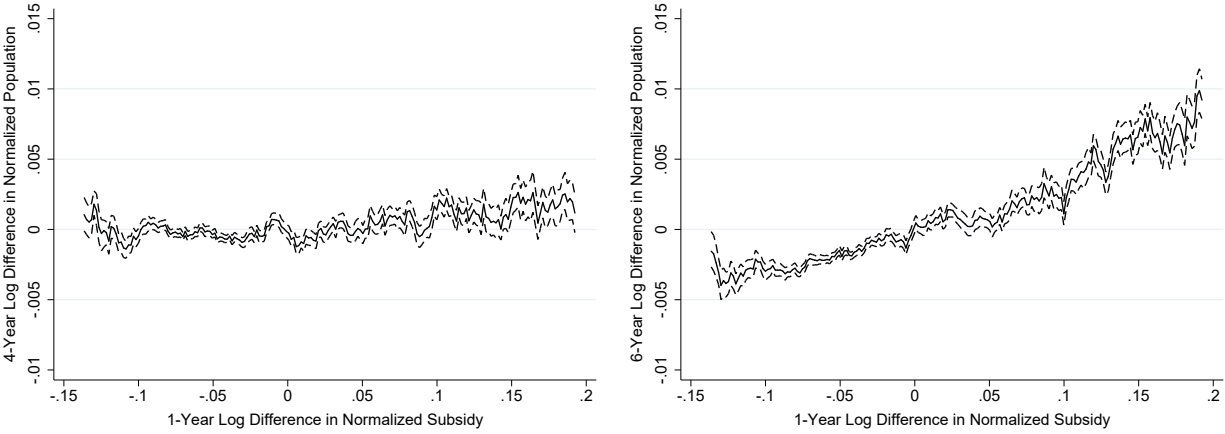
*Note:* These graphs plot the coefficients  $\beta_d$  of section 5 regressions. They show the effect of changes in a city’s subsidy stock in year 0 on cumulative changes in total  $m^2$  of housing consumed starting 6 years prior to the shock year. All changes are relative to regional mean change. Standard errors are clustered at the MF level. We report the 5% confidence bands.

FIGURE 4.13 – BETWEEN-MF HOUSING SUPPLY AND HOUSE PRICES

(a) *G* 4-YEAR RELATIVE GROWTH PRE-SHOCK (b) *G* 6-YEAR RELATIVE GROWTH POST-SHOCK



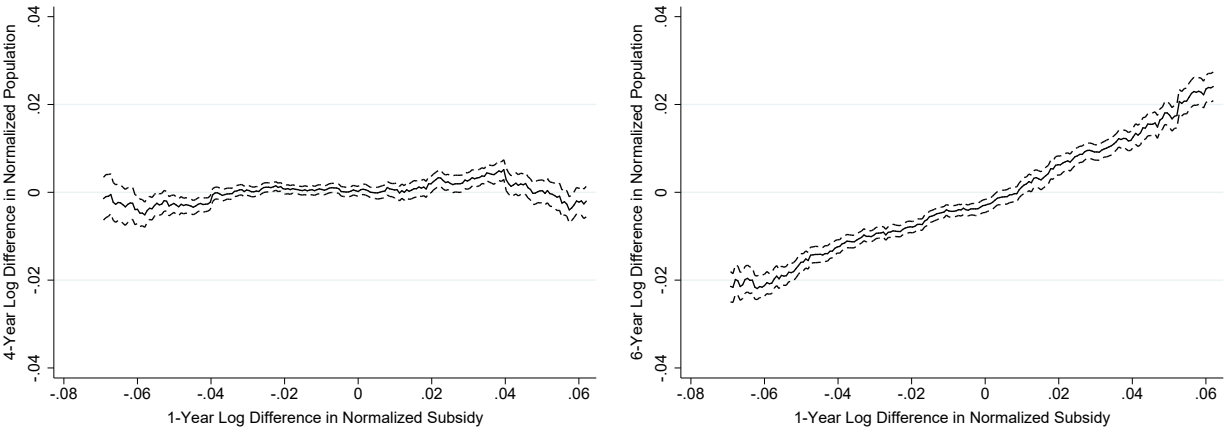
(c) *N* 4-YEAR RELATIVE GROWTH PRE-SHOCK (d) *N* 6-YEAR RELATIVE GROWTH POST-SHOCK



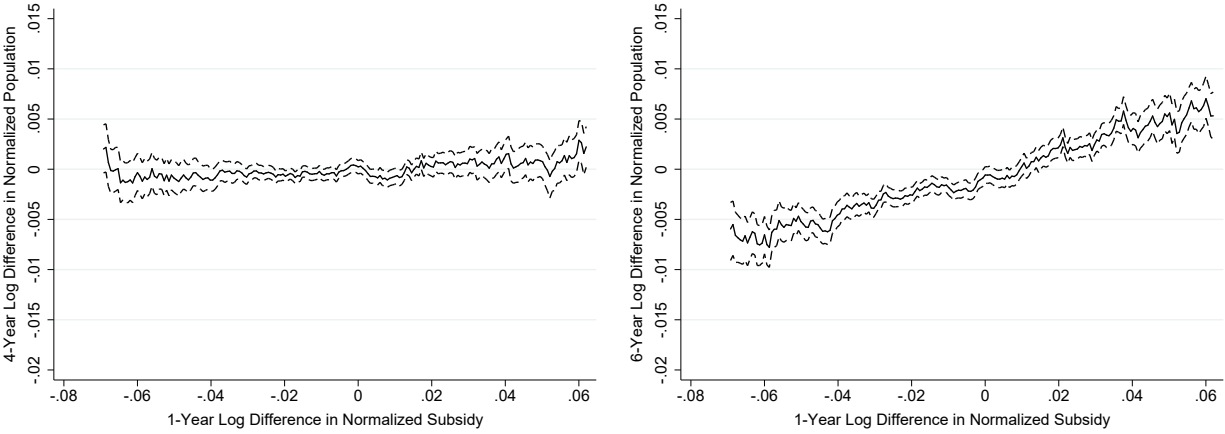
*Note:* This Figure shows the outcomes of non-parametric regressions where the explanatory variable is the within-MF relative subsidy shock. Dependent and explanatory variables are first residualized with respect to bins of fixed 1999 characteristics and lagged shocks. Bias-corrected fitted values and 1% confidence band are computed based on 1000 bootstrap replications.

FIGURE 4.15 – KERNEL REGRESSIONS: WITHIN-MF CHANGES

(a) *G* 4-YEAR RELATIVE GROWTH PRE-SHOCK (b) *G* 6-YEAR RELATIVE GROWTH POST-SHOCK



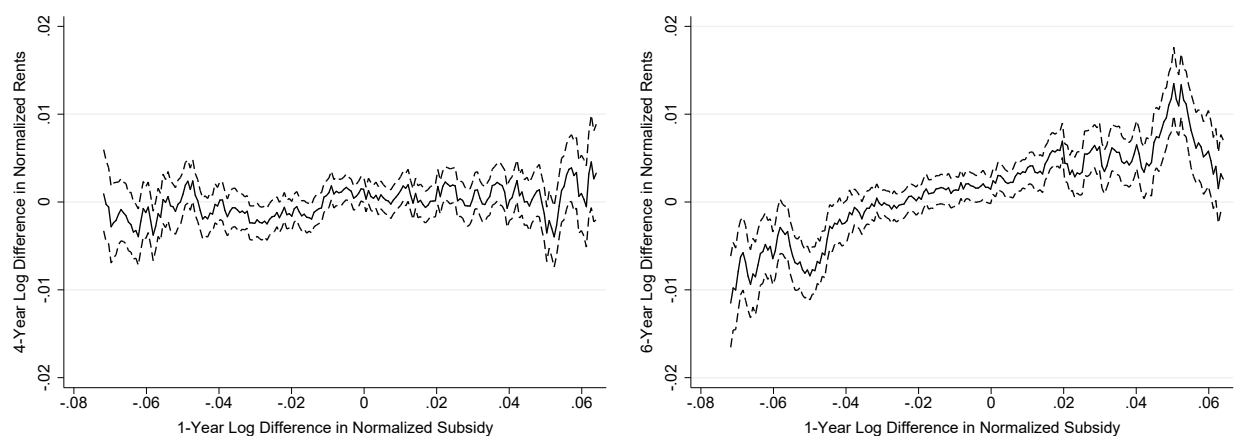
(c) *N* 4-YEAR RELATIVE GROWTH PRE-SHOCK (d) *N* 6-YEAR RELATIVE GROWTH POST-SHOCK



*Note:* This Figure shows the outcomes of non-parametric regressions where the explanatory variable is the between-MF relative subsidy shock. Dependent and explanatory variables are first residualized with respect to bins of fixed 1999 characteristics and lagged shocks. Bias-corrected fitted values and 1% confidence band are computed based on 1000 bootstrap replications.

FIGURE 4.17 – KERNEL REGRESSIONS: BETWEEN-MF CHANGES (1/3)

(a) *r* 4-YEAR RELATIVE GROWTH PRE-SHOCK (b) *r* 6-YEAR RELATIVE GROWTH POST-SHOCK



*Note:* This Figure shows the outcomes of non-parametric regressions where the explanatory variable is the between-MF relative subsidy shock. Dependent and explanatory variables are first residualized with respect to bins of fixed 1999 characteristics and lagged shocks. Bias-corrected fitted values and 1% confidence band are computed based on 1000 bootstrap replications.

FIGURE 4.19 – KERNEL REGRESSIONS: BETWEEN-MF CHANGES (1/3)





# Appendix

## A Jurisdiction Appeal

The full expression for utility in jurisdiction  $j$  is:

$$\begin{aligned}
 v_j = & (1 - \phi)(1 - \alpha) \ln(1 - \alpha) + (1 - \phi)(1 - \alpha) \ln(1 - \tau^w) + (1 - \phi)(1 - \alpha) \ln(w_j) \\
 & + (1 - \phi)\alpha \ln(h_j) + \phi\delta \ln(G_j) + \phi(1 - \delta) \frac{1}{\#a_j} \sum_{j' \in a_j} \ln(G_{j'}) - \phi\delta^2 \kappa \ln(N_j) \\
 & - 2\phi\delta(1 - \delta)\kappa \frac{1}{\#a_j} \sum_{j' \in a_j} \ln(N_{j'}) - \phi(1 - \delta)^2 \kappa \frac{1}{\#a_j} \sum_{j' \in a_j} \ln(N_{j'}) + \ln(\mathcal{E}_j^A)
 \end{aligned} \tag{4.31}$$

## B Local Public Goods

Our baseline model features a simple representation of local public spending. In Appendix B.1, we develop a more realistic version that we use in our empirical applications in Sections 5 and 6. Appendix B.2 gives conditions under which our framework is time-consistent. We take our extended version to the data in Appendix B.3.

### B.1 Extension of The Model

We assume that public good quantity  $G$  produced by a jurisdiction has a Cobb-Douglas form:

$$\ln G = \varphi \ln G^s + (1 - \varphi) \ln G^f \tag{4.32}$$

where  $G^s$  is the stock of public capital and  $G^f$  the flow of services annually consumed. Both are directly measured as spending in terms of the numéraire good and we abstract from differences in public good provision efficiency. Residents vote on the residence tax  $\tau^h$ , the business tax  $\tau^k$  and the amount of  $G^s$  and  $G^f$ . Because we assumed homogeneous preferences, the voting mechanism is akin to a maximization problem by a local social planner.

Residents commit to policy  $\{G^s, G^f, \tau^h, \tau^k\}$  for current and all future periods.<sup>29</sup> They

---

29. Appendix B.2 shows that policy choices are time-consistent if the environment stays constant. When

pay for  $G^f$  every year. Durable investments depreciate at the annual rate  $\rho$ . To maintain a constant level of infrastructure, residents hence have to pay  $G^s$  the first year and  $\rho G^s$  every following year.

Cities inherit zero net wealth from the past.<sup>30</sup> They anticipate a flow of future subsidies  $\{F_t\}_0^\infty$  and have access to international debt markets with fixed interest rate  $R$ . Residents' preferred policy is found by maximizing

$$(1 - \alpha) \ln((1 - \tau^w)w - r(1 + \tau^h)h) + \alpha \ln h + \frac{\phi}{1 - \phi} \delta (\varphi \ln G^s + (1 - \varphi) \ln G^f) \quad (4.33)$$

over  $\{G^s, G^f, \tau^h, \tau^k\}$  subject to jurisdiction inter-temporal budget constraint

$$\zeta G^s + G^f = \tau^h r H + \tau^k R K + \frac{R}{1 + R} \left[ \sum_{t=0}^{\infty} \left( \frac{1}{1 + R} \right)^t F_t \right] \quad (4.34)$$

where  $\zeta = (\rho + R)/(1 + R)$ .

## B.2 Time Consistency and No Inherited Wealth

We show that jurisdictions' decisions are time consistent. We prove it in the case without local business tax, but the argument remains valid in the presence of business taxes. We assume that jurisdictions actually inherit investment  $G_{init}^s$  and debt  $D_{init}$  from the past. Residents' preferred policy is given by

$$\max_{G^s, G^f, \tau} (1 - \alpha)(1 - \phi) \ln((1 - \tau^w)w - r(1 + \tau)h) + \alpha(1 - \phi) \ln(h) + \phi \left( \delta + \frac{1 - \delta}{|a_j|} \right) (\varphi \ln(G^s) + (1 - \varphi) \ln(G^f))$$

That is, after substituting in optimal housing and numéraire consumption:

$$\max_{G^s, G^f, \tau} -\alpha(1 - \phi) \ln(1 + \tau) + \phi \left( \delta + \frac{1 - \delta}{|a_j|} \right) (\varphi \ln(G^s) + (1 - \varphi) \ln(G^f))$$

---

there is a shock to the environment (e.g., in local amenities or public subsidies) jurisdictions change their equilibrium policy which is again time-consistent.

30. Appendix B.2 shows that under no-Ponzi conditions, jurisdictions inherit exactly zero net wealth from past periods. Hence, there is no path-dependency in local public goods choices even in the presence of durable investments and a seemingly dynamic problem collapses into a static one.

**Optimal policy from  $t=0$  (creation of the jurisdiction)** In period  $t = 0$  (e.g., the creation of the jurisdiction), jurisdictions inherit zero investments  $G_{init}^s = 0$  and debt  $D_{init} = 0$ . Residents choose and commit to a constant level of public good  $G_0$  and a constant tax rate  $\tau_0$  for current and all future periods  $\{t = 0, t = 1, \dots\}$ . They furthermore assume that population won't change in future periods. When population changes because of say an amenity shock, the problem is reinitialized at period 0, but this time with a history and an *a priori* non-zero initial jurisdiction net wealth. Public goods are durable but depreciate at speed  $\rho$ . To commit to the initially chosen  $G_0^s$  residents hence have to invest  $\rho G_0^s$  every period from  $t = 1$ . Residents receive a flow of subsidy equal to  $F$  and is assumed constant over time (again, if  $F$  changes the problem is reinitialized at period 0 with an *a priori* non-zero initial net wealth). jurisdictions can levy debt each year  $\{D_0, D_1, D_2, \dots\}$  that has to be repaid in full the next year, plus interests. Yearly budgets are given by

$$\begin{aligned}
 G_0^s + G_0^f &= D_0 + \tau_0 Nhr + F \\
 \rho G_0^s + G_0^f &= -(1 + R)D_0 + D_1 + \tau_0 Nhr + F \\
 \rho G_0^s + G_0^f &= -(1 + R)D_1 + D_2 + \tau_0 Nhr + F \\
 &\vdots \\
 \rho G_0^s + G_0^f &= -(1 + R)D_{t-1} + D_t + \tau_0 Nhr + F \\
 &\vdots
 \end{aligned} \tag{4.35}$$

Let's multiply each budget line by

$$p_t = \left( \frac{1}{1 + R} \right)^t \tag{4.36}$$

and sum all lines up to  $t = T$ :

$$G_0^s + \rho G_0^s \sum_{t=1}^T p_t + G_0^f \sum_{t=0}^T p_t = \sum_{t=0}^T p_t D_t - (1 + R) \sum_{t=1}^T p_t D_{t-1} + \tau_0 Nhr \sum_{t=0}^T p_t + F \sum_{t=0}^T p_t \tag{4.37}$$

That is:

$$G_0^s + G_0^f + (\rho G_0^s + G_0^f) \sum_{t=1}^T p_t = p_T D_T + \tau_0 Nhr \sum_{t=0}^T p_t + F \sum_{t=0}^T p_t \quad (4.38)$$

We assume that the following no-Ponzi scheme condition holds:

$$\lim_{T \rightarrow \infty} p_T D_T = 0 \quad (4.39)$$

and take the limit when  $T \rightarrow \infty$  to get the inter-temporal budget constraint

$$\zeta G_0^s + G_0^f = \tau_0 Nhr + F \quad (4.40)$$

Note that we can express the debt stock at all  $t$  by solving

$$D_t = (1 + R)D_{t-1} + \rho G_0^s + G_0^f - \tau_0 Nhr - F = (1 + R)D_{t-1} + (\rho - 1)G_0^s + D_0$$

that is (noticing an arithmetico-geometric sequence):

$$D_t = (1 + R)^t \left( D_0 \frac{1 + R}{R} + (\rho - 1)G_0^s \frac{1}{R} \right) - \frac{1}{R} ((\rho - 1)G_0^s + D_0)$$

That is:

$$p_t D_t = D_0 \frac{1 + R}{R} + (\rho - 1)G_0^s \frac{1}{R} - p_t \frac{1}{R} ((\rho - 1)G_0^s + D_0)$$

Taking the limit when  $t \rightarrow \infty$  using the no-Ponzi condition yields

$$0 = D_0 \frac{1 + R}{R} + (\rho - 1)G_0^s \frac{1}{R} \quad (4.41)$$

so that finally we show that debt is constant and that we are in a stationary setting (holding environment fixed):

$$D_t = D_0 = \frac{1 - \rho}{1 + R} G_0 \quad (4.42)$$

**Optimal policy from t=1 (keeping environment fixed)**      At period 1, jurisdictions

inherit debt (plus interests) and depreciated assets from period 0:

$$(1 - \rho)G_0^s - (1 + R)D_0$$

However, according to equation (4.41) this quantity is exactly zero. The optimization problem in period 1 is hence the same as in period 0, and policy choices from period 1 onward are unchanged compared to those in period 0. The proof by induction for any  $t$  follows straightforwardly.

**Optimal policy from t=1 (with a change in the environment)** Initial net jurisdiction wealth is zero so the problem is reinitialized at period zero following a shock to the environment. The environment has changed however, so policy choices may be different in the new equilibrium.

### B.3 Taking the Extended Model to the Data

We construct the public good index  $G$  using our data on municipal accounts. First, we construct  $G^s$  and  $G^f$ , i.e., municipal durable facilities and annual public services. Then, we estimate parameter  $\varphi$  of the Cobb-Douglas function.

**Construction of  $G^s$**  We construct our public investment variable  $G^s$  as the sum of all public assets minus the raw value of the land and financial assets such as cash.<sup>31</sup> They are recorded at book value and account for investment depreciation. As mentioned in Section 2, investments notably include schools, transportation infrastructure, parks improvements, sports facilities, museums, art collections, etc. They also include investment subsidies to other parties such as local clubs and associations. Importantly,  $G^s$  does not contain social housing units.<sup>32</sup>

---

31. Taking out the raw value of the land seems natural as residents are unlikely to value it. Given that land is mostly a gift from the central government, erasing this asset as well as the corresponding liability from the balance sheet is neutral in our analysis. Cash and other liquid assets can be considered negative debt and are accounted for in our theoretical framework in the form of future taxes.

32. Social housing units, when publicly owned, are held by ad hoc entities and not by jurisdictions.

**Construction of  $G^f$**  We construct  $G^f$  the flow of services annually consumed as the sum of staff expenditure, maintenance spending, payments for external services and operating subsidies to third parties. We exclude interests payments as they do not correspond to consumable services.

**Construction of  $G$**  To make sure that our constructed measures for  $G^s$  and  $G^f$  correctly reflect the amount of services chosen at the local level, we first check that they satisfy standard budget requirements. We then directly calibrate the ratio  $\frac{G^s}{G^f}$  to recover parameter  $\varphi$ . We first estimate parameter  $\rho$  by calibrating asset depreciation based on municipal financial accounts. Our central estimate is  $\rho = 0.010$ . Then we calibrate parameter  $R$  from interests paid as a share of the debt stock. Our central estimate is  $R = 0.041$ . Both estimates are robust across different calibration methods. We are then able to calibrate parameter  $\zeta = \frac{\rho+R}{1+R}$  with a central estimate of  $\zeta = 0.049$ . Finally, we estimate the ratio  $\frac{G^s}{G^f} = \frac{\varphi}{(1-\varphi)\zeta}$  with a log-log regression reproduced in Figure 4.21. The Figure substantiates our Cobb-Douglas specification as the slope is close to unity and residuals are quite small as indicated by the large  $R^2$ . Point estimate for the intercept is  $1/0.074 \approx 13.5$ , that is, capital investment approximately equals 13.5 years of operating expenditures. This leads to a central estimate  $\varphi = 0.378$ . We can then rewrite our public good index  $G$  as

$$G = G^s \left( \frac{1-\varphi}{\varphi} \zeta \right)^{1-\varphi} \approx 0.208 \times G^s$$

## C DiD Robustness Checks

Let jurisdiction total 20–65 population be

$$N = N_p + N_{\bar{p}}$$

where  $N_p$  is the public sector population. Let  $s_p = N_p/N$  be the pre-shock share of public employees in the population. Within-MF changes in  $N$  following a shock can be decomposed

as follows:

$$\begin{aligned}
 d \ln (N/\bar{N}) &= \frac{d(N/\bar{N})}{N/\bar{N}} \\
 &= \frac{d(N_{\bar{p}}/\bar{N})}{N/\bar{N}} + \frac{d(N_p/\bar{N})}{N/\bar{N}} \\
 &= \frac{d(N_{\bar{p}}/\bar{N})}{N/\bar{N}} + \frac{d(s_p N/\bar{N})}{N/\bar{N}} \\
 &= \frac{d(N_{\bar{p}}/\bar{N})}{N/\bar{N}} + s_p \left[ \frac{d(N/\bar{N})}{N/\bar{N}} + \frac{ds_p}{s_p} \right]
 \end{aligned}$$

The share of migration responses coming from public employment is then

$$\frac{\frac{d(N_p/\bar{N})}{N/\bar{N}}}{\frac{d(N/\bar{N})}{N/\bar{N}}} = s_p \left[ 1 + \frac{ds_p/s_p}{\frac{d(N/\bar{N})}{N/\bar{N}}} \right]$$

Figure 4.22 shows an estimate for  $\frac{ds_p/s_p}{\frac{d(N/\bar{N})}{N/\bar{N}}}$  of around 0.09. The estimate for  $\frac{\frac{d(N/\bar{N})}{N/\bar{N}}}{\frac{d(S/\bar{S})}{S/\bar{S}}}$  from Section 5.2 is around 0.03. The average estimate for the pre-shock ratio of public staff to total 20–65 population is  $s_p \approx 3\%$ . Estimate for the share of migration responses coming from public employment is hence  $\approx 3\% \times (1 + 0.09/0.03) = 12\%$ . We get similar estimates when including paid subsidies to third parties in addition to public staff payroll.

## D Welfare

We detail here the calculations for the welfare analysis.

### D.1 Before The Merger

We first back out the fundamentals of the model  $\{\mathcal{E}_j^A, \mathcal{E}_j^C\}$  from (i) equilibrium endogenous variables of the model that we observe as econometricians, and (ii) parameters that we have estimated or calibrated from the literature. We then make several assumptions to ensure tractability of the equilibrium in our simplified setup. Motivated by the suppression of the French local business tax in 2011, we first assume that local public revenues only come from the residence tax. We make four others assumptions: (i) residents are myopic and do not

anticipate migration responses when they vote for public goods and taxes; (ii) the public good production function is the identity, i.e.,  $\Gamma(x) = x$ ; (iii) housing supply elasticities  $\eta_j$  are constant across jurisdictions; (iv) local productivity  $\theta_j^Y$  does not depend on local public goods and (v) the geographic fundamentals that we have backed out are still relevant in our simplified model.

This version of the model has a unique and tractable equilibrium defined by the following equations:

$$N_j = \frac{\exp(v_j/\sigma)}{\sum_{j'} \exp(v_{j'}/\sigma)} \quad (4.43)$$

$$v_j = (1 - \phi)\alpha \ln \alpha + (1 - \phi)(1 - \alpha) \ln(1 - \alpha) \quad (4.44)$$

$$+ (1 - \phi) \ln((1 - \tau^w)w_j) - (1 - \phi)\alpha \ln(1 + \tau_j^h) - (1 - \phi)\alpha \ln(r_j) \quad (4.45)$$

$$+ \phi\delta \ln(G_j) + \phi \frac{1 - \delta}{\#a_j} \sum_{j' \in a_j} \ln G_{j'} \\ - \phi\kappa\delta^2 \ln N_j - \phi\kappa \frac{1 - \delta^2}{\#a_j} \sum_{j' \in a_j} \ln N_{j'} + \sigma\theta_j^A$$

$$\ln(r_j) = \frac{1}{\eta} \ln \frac{H_j}{T_j} + \theta_j^C \quad (4.46)$$

$$1 + \tau_j^h = 1 + \frac{\phi(\delta + \frac{1-\delta}{|a_j|})}{(1 - \phi)\alpha} \quad (4.47)$$

$$G_j [1 + \tau_j^h] = \frac{\varphi^\varphi(1 - \varphi)^{1-\varphi}}{\zeta^\varphi} \tau_j^h N_j w_j \alpha (1 - \tau^w) \quad (4.48)$$

Equations 4.47 and 4.48 are the vote conditions in our simplified version.<sup>33</sup> Equations 4.43, 4.45 and 4.46 are identical in the simplified case and in our baseline model with subsidies.

We note  $N^o, G^o, v^o, r^o, \tau^o$  all the equilibrium variables in the normal case and  $N^n, G^n, v^n, r^n, \tau^n$  all equilibrium variables in the simplified version. With our data we observe each year a set

---

33. The vote conditions derive from the following optimization problem:  $\max_{\tau_j^H, G_j^s, G_j^f, G_j} \left( G_j^\delta \prod_{j' \in a_j} G_{j'}^{\frac{1-\delta}{|a_j|}} \right)^\phi C_j^{1-\phi}$  under the constraints  $\zeta G_j^s + G_j^f = N_j h_j r_j \tau_j^H$ ,  $G_j = (G^s)^\varphi (G_j^f)^{1-\varphi}$ ,  $C_j = c_j^{1-\alpha} h_j^\alpha$ ,  $c_j = (1 - \alpha)w_j(1 - \tau^W)$ ,  $h_j = \frac{\alpha w_j(1 - \tau^W)}{r_j(1 + \tau_j^H)}$



of equilibrium variables  $N^o, G^o, v^o, r^o, \tau^o$ . We show that we can deduce equilibrium variables  $N^n, G^n, v^n, r^n, \tau^n$  from observed data. It is equivalent to infer residuals  $\theta_j^A, \theta_j^C$  from observed data.

We rewrite equilibrium conditions 4.43, 4.45 and 4.46:

$$N_j = \frac{\exp(v_j/\sigma)}{\sum_{j'} \exp(v_{j'}/\sigma)} \quad (4.49)$$

$$v_j = -\frac{\eta}{\eta+1}(1-\phi)\alpha \ln(1+\tau_{a_j}) + \phi\delta \ln(G_j) + \phi \frac{1-\delta}{\#a_j} \sum_{j \in a_j} \ln G_{j'} \quad (4.50)$$

$$- \left[ (1-\phi)\alpha \frac{1}{\eta+1} + \phi\kappa\delta^2 \right] \ln N_j - \phi\kappa \frac{1-\delta^2}{\#a_j} \sum_{j' \in a_j} \ln N_{j'} + \sigma\Theta_j^A$$

$$\ln r_j = \frac{1}{\eta+1} \ln N_j - \frac{1}{\eta+1} \ln(1+\tau_{a_j}) + \Theta_j^r \quad (4.51)$$

with the transformed residuals

$$\Theta_j^r = \frac{\eta}{\eta+1} \theta_j^C + \frac{1}{\eta+1} \ln \frac{\alpha w_j (1-\tau^W)}{T_j}$$

$$\sigma\Theta_j^A = (1-\phi)\alpha \ln \alpha + (1-\phi)(1-\alpha) \ln(1-\alpha) + (1-\phi) \ln((1-\tau^w)w_j) - (1-\phi)\alpha\Theta_j^r + \phi\delta\Theta_j^G$$

$$+ \phi \frac{1-\delta}{\#a_j} \sum_{j \in a_j} \Theta_{j'}^G + \theta_j^A$$

Let's define the following operators:

- $\underline{X}_j = X_j - \sum_{j \in a} \frac{X_j}{\#a_j}$
- $\bar{X}_j = \exp(\underline{\ln}(X_j))$
- $\widehat{X}_{a_j} = \sum_{j' \in a} \frac{X_{j'}}{\#a} - \sum_{a'} \sum_{j' \in a'} \frac{X_{j'}}{J\#a'}$
- $\widehat{X}_j = \exp(\widehat{\ln}(X_j))$
- $\hat{X}_j = X_j - \sum_{j'} \frac{X_{j'}}{J}$
- $\check{X}_j = \exp(\hat{\ln}(X_j))$

Applying the  $\underline{L}$  operator to equations 4.47, 4.48, 4.49, and 4.50 gives:

$$\left[ \sigma + (1 - \phi)\alpha \frac{1}{\eta + 1} + \phi\kappa\delta^2 - \phi\delta \right] \ln \frac{\overline{N}_j^n}{\overline{N}_j^o} = \phi\delta (\ln \overline{N}_j^o - \ln \overline{G}_j^o + \ln \overline{w}_j) + \frac{\eta}{\eta + 1}(1 - \phi)\alpha \ln \overline{1 + \tau}_j^o \quad (4.52)$$

Applying the  $\widehat{L}$  operator to equations 4.48, 4.49, and 4.50 we get:

$$\left[ \sigma + \frac{(1 - \phi)\alpha}{(\eta + 1)} + \phi\kappa - \phi \right] \ln \frac{\widehat{N}_a^n}{\widehat{N}_a^o} = \phi \left( \ln \widehat{N}_a^o - \ln \widehat{G}_a^o + \ln \widehat{w}_a + \ln \widehat{\tau}_a^n - \ln \widehat{1 + \tau}_a^n \right) - \frac{\eta}{\eta + 1}(1 - \phi)\alpha \ln \frac{\widehat{1 + \tau}_a^n}{\widehat{1 + \tau}_a^o} \quad (4.53)$$

Notice that we can write:

$$\frac{N_j^n}{N_j^o} = \frac{\widehat{N}_j^n \overline{N}_j^n \sum_{j'} \widehat{N}_{j'}^o \overline{N}_{j'}^o}{\widehat{N}_j^o \overline{N}_j^o \sum_{j'} \widehat{N}_{j'}^n \overline{N}_{j'}^n} \quad (4.54)$$

With equations 4.47, 4.52, 4.53, 4.54 and 4.48 — in this order — we get  $\tau_j^n, N_j^n, G_j^n$ .

## D.2 After the merger

Now we depart from our simplified version by assuming all jurisdictions within a municipal federation cooperate for the production of public goods. There is still no business tax and no state subsidies. Prior to the merger, residents vote for public goods in each jurisdiction. Following the merger, residents vote for a new level of public goods pooling resources of all former member jurisdictions. To keep welfare comparable between pre- and post-reform situations, we assume that prior jurisdictions still exist for the purpose of labor and housing markets. The only thing that changes is the level at which public good is supplied.

Housing tax rates are fixed by MF council whose preferences reflect exactly those of all inhabitants of the MF. For a given MF  $a$  the council chooses the housing tax rate and the level of local public good  $G$  which maximize  $v_j$  — with  $j$  a jurisdiction belonging to the MF — the budget constraint  $\sum_{j' \in a} r_{j'} N_{j'} h_{j'} \tau_{j'}^h = G$ . Importantly the optimization problem gives the

same  $G$  and  $\tau^H$  for all  $j$  belonging to  $a$ . Once  $G$  is chosen it is split and a share is allocated to each jurisdiction. We assume in our reform that the allocation rule keeps the previous share unchanged, that is, jurisdiction  $j$  gets the share  $\frac{G_j^n}{\sum_{j' \in a_j} G_{j'}^n}$  of  $G$ . We could, however, choose another rule, such as an egalitarian rule.

We note  $N^m, G^m, v^m, r^m, \tau^m$  all the equilibrium variable in the merger case. The new vote conditions are<sup>34</sup>

$$1 + \tau_a^m = 1 + \frac{\phi}{(1 - \phi)\alpha} \quad (4.55)$$

$$G_j^m = \frac{\varphi^\varphi (1 - \varphi)^{1 - \varphi}}{\zeta^\varphi} \frac{G_j^n}{\sum_{j' \in a_j} G_{j'}^n} \sum_{j' \in a_j} \frac{\tau_{j'}^m}{1 + \tau_{j'}^m} N_{j'}^m w_{j'} \alpha (1 - \tau^w) \quad (4.56)$$

where  $\tau_a^m$  and  $G_a^m$  are independant of the allocation rule chosen.

Equations 4.49, 4.50 and 4.51 still hold. We now rewrite  $N^m, G^m, v^m, r^m, \tau^m$  as function of  $N^n, G^n, v^n, r^n, \tau^n$ .

Applying the  $\underline{L}$  operator to equations 4.47, 4.48, 4.49, 4.50, 4.55, and 4.56 we get:

$$\overline{N_j^m} = \overline{N_j^n} \quad (4.57)$$

The allocation of population within any MF stays constant.

Note that

$$\begin{aligned} G_j^m &= \frac{G_j^n}{\sum_{j' \in a_j} G_{j'}^n} \sum_{j' \in a} G_{j'}^m \frac{\tau_{j'}^m}{1 + \tau_{j'}^m} \frac{1 + \tau_{j'}^n}{\tau_{j'}^n} \frac{N_{j'}^m}{N_{j'}^n} \\ &= \left[ \prod_{j' \in a_j} \frac{N_{j'}^m}{N_{j'}^n} \right]^{\frac{1}{\#a_j}} \frac{G_j^n}{\sum_{j' \in a_j} G_{j'}^n} \sum_{j' \in a_j} G_{j'}^m \frac{\tau_{j'}^m}{1 + \tau_{j'}^m} \frac{1 + \tau_{j'}^n}{\tau_{j'}^n} \frac{\overline{N_{j'}^m}}{\overline{N_{j'}^n}} \\ &= \left[ \prod_{j' \in a_j} \frac{N_{j'}^m}{N_{j'}^n} \right]^{\frac{1}{\#a_j}} \frac{G_j^n}{\sum_{j' \in a_j} G_{j'}^n} \sum_{j' \in a_j} G_{j'}^m \frac{\tau_{j'}^m}{1 + \tau_{j'}^m} \frac{1 + \tau_{j'}^n}{\tau_{j'}^n} \end{aligned} \quad (4.58)$$

---

34. The vote conditions derive from the following optimization problem:  $\max_{\tau_{a_j}^H, G_{a_j}, G_{a_j}^s, G_{a_j}^f} \left( [G_{a_j} w_j]^\delta \prod_{j' \in a_j} [G_{a_j} w_{j'}]^{\frac{1 - \delta}{|a_j|}} \right)^\phi C_j^{1 - \phi}$  under the constraints  $\zeta G_{a_j}^s + G_{a_j}^f = \sum_{j' \in a_j} N_j h_j r_j \tau_{a_j}^H$ ,  $G_{a_j} = (G_{a_j}^s)^\varphi (G_{a_j}^f)^{1 - \varphi}$ ,  $C_j = c_j^{1 - \alpha} h_j^\alpha$ ,  $c_j = (1 - \alpha) w_j (1 - \tau^w)$ ,  $h_j = \frac{\alpha w (1 - \tau^w)}{r_j (1 + \tau_{a_j}^H)}$ .

We define  $X_a$  such as

$$\frac{X_a}{G_j^n} = \frac{\sum_{j' \in a} G_{j'}^m \frac{\tau_{j'}^m}{1+\tau_{j'}^m} \frac{1+\tau_{j'}^n}{\tau_{j'}^n}}{\sum_{j' \in a_j} G_{j'}^n} \quad (4.59)$$

Applying the  $\widehat{L}$  operator to equations 4.48, 4.49, 4.50, 4.55, 4.56, 4.58 and 4.59 we get:

$$\left[ \sigma + (1 - \phi)\alpha \frac{1}{\eta + 1} + \phi\kappa - \phi \right] \ln \frac{\widehat{N}_j^m}{\widehat{N}_j^n} = \phi \ln \frac{\widehat{X}_a}{\widehat{G}_a^n} - \frac{\eta}{\eta + 1} (1 - \phi)\alpha \ln \frac{\widehat{1 + \tau_a^m}}{\widehat{1 + \tau_a^n}} \quad (4.60)$$

Note that we have

$$\frac{N_j^m}{N_j^n} = \frac{\widehat{N}_j^m \widehat{N}_j^m}{\widehat{N}_j^n \widehat{N}_j^n} \frac{\sum_{j'} \widehat{N}_{j'}^n \widehat{N}_{j'}^n}{\sum_{j'} \widehat{N}_{j'}^m \widehat{N}_{j'}^m} \quad (4.61)$$

With equations 4.55, 4.57, 4.58, 4.59, 4.60, 4.61 and 4.56 — in this order — we get  $\tau_j^m, N_j^m, G_j^m$ .

Finally:

$$\begin{aligned} v_j^m - v_j^n &= \phi\delta \ln \frac{G_j^m}{G_j^n} + \phi \frac{1 - \delta}{\#a_j} \sum_{j' \in a_j} \ln \frac{G_{j'}^m}{G_{j'}^n} - \left[ (1 - \phi)\alpha \frac{1}{\eta + 1} + \phi\kappa\delta^2 \right] \ln \frac{N_j^m}{N_j^n} \\ &\quad - \phi\kappa \frac{1 - \delta^2}{\#a_j} \sum_{j' \in a_j} \ln \frac{N_{j'}^m}{N_{j'}^n} - \frac{\eta}{\eta + 1} (1 - \phi)\alpha \ln \left( \frac{1 + \tau_{a_j}^m}{1 + \tau_{a_j}^n} \right) \end{aligned} \quad (4.62)$$

### D.3 Welfare comparison

We define social welfare as

$$W = \mathbb{E} \max_j u_{ij} = \mathbb{E} \max_j (v_j + \epsilon_{ij})$$

Since idiosyncratic shock follow a Gumbel law, we have<sup>35</sup>

$$W = \sigma\zeta + \sigma \ln \left( \sum_j e^{\frac{v_j}{\sigma}} \right)$$

---

35.  $\zeta$  is the Euler constant.

hence

$$W^o = \sigma\zeta + \sum_j \frac{v_j^o}{J} + \sigma \ln \left( \sum_j \exp \left( \frac{v_j - \sum_j \frac{v_j^o}{J}}{\sigma} \right) \right)$$

Note that  $\check{N}_j^n = \exp \left( \frac{\widehat{v}_j^n}{\sigma} \right)$  therefore

$$W^n = \sigma\zeta + \sum_j \frac{v_j^n}{J} + \sigma \ln \left( \sum_j \check{N}_j^n \right)$$

and

$$W^m = \sigma\zeta + \sum_j \frac{v_j^m}{J} + \sigma \ln \left[ \sum_j \check{N}_j^n \exp \frac{v_j^m - v_j^n}{\sigma} \right]$$

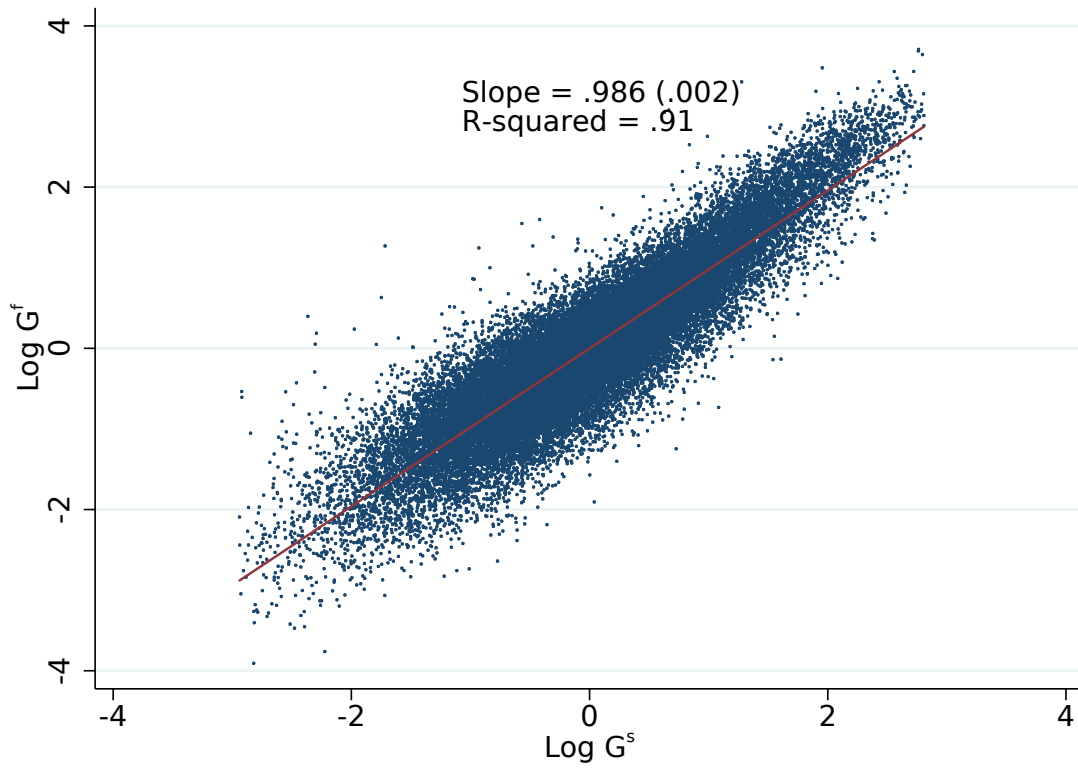
Eventually

$$W^m - W^n = \sigma \ln \left( \frac{\sum_j \check{N}_j^n \exp \frac{v_j^m - v_j^n}{\sigma}}{\sum_j \check{N}_j^n} \right) \tag{4.63}$$

## D.4 Welfare estimation

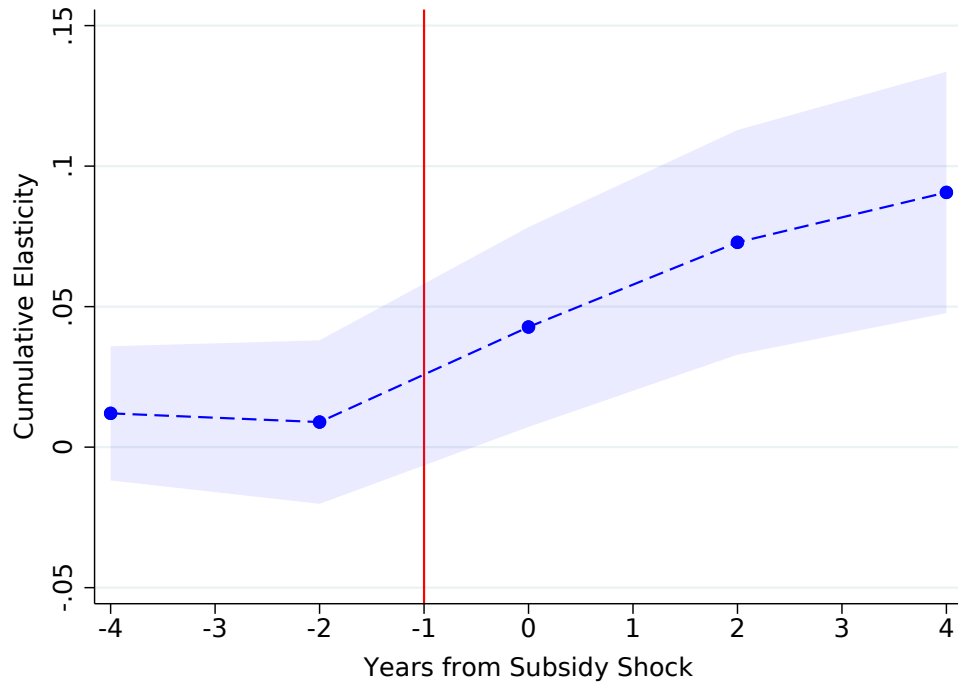
For the estimation of welfare change, we use parameters estimates of Section 6 and their estimated variance-covariance structure. As we have as many parameters estimates as we have IV specifications, we take the average across specifications for each parameter to compute the point estimate for welfare change. We compute Monte Carlo standard errors by simulating 10,000 new values for each parameter using the estimated variance-covariance matrix. We then get the mean estimate for  $\kappa$ ,  $\phi$ ,  $\delta$ ,  $\sigma$  across specifications. We then estimate the welfare impact of a reform for each of these simulated set of parameters and compute percentile-based standard errors. As for  $\eta$  and  $\alpha$ , we take standard values from the literature. In our baseline scenario we take  $\eta = 0.2$  and  $\alpha = 0.3$ .





*Note:* This Figure shows the calibration exercise we carry out for our Cobb-Douglas modeling of how public services enter utility.

FIGURE 4.21 – MODEL FIT:  $G^f$  vs  $G^s$



*Note:* This Figure plots the coefficient of regressions similar to Section 5 regressions. The dependent variable is the proportional change in the absolute share of public employees in city population  $\Delta \ln(s_p)$  and the explanatory variable is the within-MF subsidy shock  $\Delta \ln(S/\bar{S})$ . Standard errors are clustered at the MF level. We report the 5% confidence bands.

FIGURE 4.22 – CHANGES IN (PUBLIC EMPLOYEES)/ (20-65 POPULATION)



## Chapter 5

# Les étrangers font-ils monter les prix de l'immobilier? Estimation à partir de la base de la chambre des Notaires de Paris, 1993-2008.<sup>1</sup>

### Résumé

Les prix de l'immobilier ont augmenté de façon spectaculaire à Paris durant les vingt dernières années: hausse de 150% en euros courants et de 100% en euros constants. Les étrangers sont souvent accusés par les médias d'être responsables des prix exorbitants qui ont cours dans la capitale. En utilisant une base de données exhaustive sur le parc de logements à Paris<sup>2</sup> et une autre base de donnée, quasiment exhaustive, sur les transactions d'appartements à Paris<sup>3</sup>, nous démontrons que l'impact positif des étrangers sur les prix a été négligeable comparé à la hausse qui a eu lieu.

---

1. Ce chapitre reprend dans son intégralité le mémoire que j'ai réalisé en 2011 à l'Ecole d'Economie de Paris sous la direction de Thomas Piketty. Je le remercie vivement ainsi que Gabrielle Fack, Julien Grenet, Alain Jacquot, Jean-Michel Arnoux, Karl Even, Rémi Josnin et Ngoc Yen Tran Duy.

2. Il s'agit de la base FILOCOM / Fichier des LOGements par COMMunes

3. Il s'agit de la base BIEN/ Notaires de Paris d'Ile-de-France

Selon nos estimations, la proportion d'acheteurs étrangers a bien un impact positif et significatif sur les prix: une augmentation d'un point de % de la proportion d'étrangers parmi l'ensemble des acheteurs entraîne une augmentation des prix de 0,370%. Entre 1993 et 2008, la part des étrangers dans les achats a cru: elle est passée de 6% à 8% pour les acheteurs résidant à l'étranger et de 3% à 9% pour les acheteurs de nationalité étrangère. Les étrangers sont donc responsables au plus d'une augmentation des prix de 3%, soit 3% de l'augmentation qui a effectivement eu lieu. Bien qu'ils achètent des biens jusqu'à 30% plus chers, les étrangers ne sont pas assez nombreux pour avoir entraîné une hausse substantielle des prix, surtout que l'élasticité-prix au logement est très élevée à Paris au vu de nos estimations.

Toujours selon nos estimations, ce sont les acheteurs de nationalité française ou étrangère qui habitent à l'étranger plutôt que les acheteurs de nationalité étrangère habitant en France qui ont eu un impact significatif sur les prix. Les provinciaux et habitants de la petite et grande couronne, qui sont eux aussi des étrangers de la zone de Paris, n'ont eu, eux, aucun impact sur les prix. Et parmi les acheteurs de nationalité étrangère, ce sont les acheteurs originaires d'Europe -hors France- et d'Amérique qui ont eu un impact significatif sur les prix.

Et si les étrangers faisaient surtout monter les prix de l'immobilier haut de gamme? Plus on considère des gammes de prix élevés, plus l'augmentation de la proportion d'acheteurs étrangers a été importante et plus l'augmentation de prix a été, elle, mesurée. C'est pourquoi malgré une forte augmentation de la proportion d'acheteurs étrangers sur le segment de l'immobilier haut de gamme, l'impact sur les prix de ce segment de marché n'a pas dépassé 3%: selon nos estimations, une augmentation d'un point de % de la proportion d'acheteurs étrangers dans le secteur de l'immobilier haut de gamme entraîne une augmentation des prix d'au plus 0,232%.

**JEL Classification:** F22, R21, R31

**Keywords:** Marché de l'immobilier, Immigration, Prix des logements, Prix hédoniques, Paris

## Introduction

Les prix de l'immobilier en France ont subi une hausse spectaculaire durant les 20 dernières années, en particulier à Paris. De nombreux facteurs sont invoqués par les médias et les économistes pour expliquer un tel phénomène: forte baisse des taux d'intérêt, spéculation, augmentation du nombre de pied-à-terre. Tous les mois ou presque, un journal accuse les étrangers ou les provinciaux d'être responsable de la forte hausse des prix à Paris.

Grâce à une étude empirique approfondie, se basant sur des bases de données quasiment exhaustives sur le parc de logements parisien et son flux de transactions, nous sommes en mesure de réfuter cette accusation.

Avant de commencer cette étude, connaître l'impact de l'immigration sur les prix de l'immobilier nous a semblé intéressant pour deux raisons :

D'une part, cela nous permet de mieux appréhender les effets de l'immigration sur le niveau de vie des natifs. Toute augmentation des prix de l'immobilier a un impact important sur le niveau de vie des ménages car ces derniers dépensent une part importante de leurs revenus pour se loger. En cela, ce mémoire rejoint la littérature qui traite de l'impact économique de l'immigration.

D'autre part, depuis la crise financière récente, l'intérêt de comprendre le fonctionnement des marchés immobiliers n'est plus à démontrer. Pourquoi les prix de l'immobilier explosent en France, aux États-Unis et en Angleterre et stagnent au Japon et en Allemagne (voir [Adam et al. \(2011\)](#))?

Dans ce mémoire, nous allons montrer que les étrangers et les provinciaux ne peuvent être tenus pour responsable de la hausse spectaculaire des prix de l'immobilier à Paris. Pour des raisons pratiques, nous n'étudions que l'impact des étrangers sur les prix des appartements anciens à Paris entre 1993 et 2008. Nous n'étudions pas l'impact de l'immigration sur le niveau des loyers à Paris. De plus, nous utilisons au fil de ce mémoire plusieurs définitions du mot "étranger": les étrangers peuvent être les personnes qui résident à l'étranger, voire en province ou les personnes de nationalité étrangère.

Dans la fin de cette section, nous présentons brièvement la littérature relative à ce mémoire. Dans la deuxième section, nous présentons les caractéristiques essentielles du parc de logement et du flux de logements vendus chaque année à Paris et nous présentons les bases de données utilisées. Dans la troisième section, nous présentons quelques résultats théoriques puis nous détaillons la méthodologie adoptée pour nos régressions. Dans la quatrième section, nous présentons et analysons nos résultats. Dans la cinquième section, nous testons la robustesse de nos résultats et nous estimons l'impact global des étrangers sur les prix de l'immobilier. La dernière section conclut ce mémoire.

La littérature relative à notre sujet d'étude "Les étrangers font-ils monter les prix de l'immobilier?" est récente.

Une longue tradition empirique s'est d'abord intéressée à l'impact économique de l'immigration, en se concentrant sur les effets sur le marché du travail. Plusieurs papiers ([Card \(1989\)](#), [Card \(2001\)](#)) soutiennent que l'immigration n'a eu qu'un faible impact sur les salaires aux États-Unis, même lorsque la réponse migratoire des natifs est prise en compte ([Card and DiNardo \(2000\)](#)).

Cependant, depuis une dizaine d'années, certains économistes souhaitent adopter un point de vue d'équilibre général ([Ottaviano and Peri \(2007\)](#)) et s'intéressent aussi à l'impact de l'immigration sur les prix des logements. À priori, quel devrait être l'impact de l'immigration sur le marché de l'immobilier? Prenons l'exemple des États-Unis où les immigrants se concentrent dans quelques états et vivent principalement en zone urbaine. Si les afflux d'immigrants ne sont pas totalement compensés par un flux migratoire des natifs vers les autres états, les immigrants devraient augmenter la demande et le prix des logements. Sur le long terme, l'augmentation de la demande devrait conduire à une construction de logement accrue. Cependant, l'offre de logements est quasiment inélastique dans les villes du fait de la rareté des terrains et des régulations imposées par les autorités. Les flots d'immigration peuvent donc conduire à une augmentation des prix et à une baisse de la quantité de logement consommée par personne. Dans les faits, les travaux empiriques sont partagés. Des preuves empiriques suggèrent qu'aux États-Unis, les loyers des natifs ont augmenté de façon significative dans

les zones où les flux d'immigrations ont été les plus importants (Saiz (2003b), Saiz (2006), Saiz (2003a) et Ottaviano and Peri (2005)). Par ailleurs, en Espagne, les flux d'immigration massifs seraient responsables de près d'un tiers de l'augmentation des prix de l'immobilier et du nombre de constructions nouvelles entre 1998 et 2008 (Gonzalez and Ortega (2009)). Néanmoins, d'autres études aux États-Unis, qui se focalisent sur les groupes d'individus les moins riches et les moins éduqués, n'ont pas trouvé un impact significatif de l'immigration sur les loyers (Greulich et al. (2005)), suggérant que l'immigration est peut-être compensée par des flux migratoires des natifs vers d'autres états.

Dans ce mémoire, contrairement aux articles précédemment cités, nous ne nous restreignons pas aux immigrants, c'est-à-dire aux personnes de nationalité étrangère qui décident de s'installer dans un autre pays, mais également aux personnes qui achètent simplement un pied-à-terre sans changer de lieu de résidence principale.

## 1 Stock et flux d'appartements à Paris

Pour mener à bien notre étude, nous utilisons deux bases de données d'une grande richesse. La première, la base FILOCOM (Fichier des LOgements par COMmunes), est une base de données exhaustive sur le parc de logement en France. Elle est constituée du rapprochement du fichier de la taxe d'habitation (TH), du fichier foncier, du fichier des propriétaires et du fichier de l'impôt sur les revenus des personnes physiques (IRPP ou IR). Tous les locaux à usage d'habitation, assujettis à la taxe d'habitation (occupés et imposés ou occupés et exonérés) ou non assujettis (vacants ou meublés soumis à la taxe professionnelle) sont présents dans cette base. La base est une photographie du parc de logement en 1999, 2001, 2003, 2005, 2007, 2009 et 2010. Pour notre étude, les variables utilisées sont la surface habitable des logements, leur mode et statut d'occupation, le type de leur propriétaire et le département de résidence de ce dernier.

Nous utilisons par ailleurs la base BIEN/ Notaires de Paris D Ile-de-France. En France, toute vente de biens immobiliers doit être enregistrée par un notaire. La Chambre des Notaires

de Paris et d'Ile France a rassemblé les données associées à ces ventes depuis le début des années 1990 et a constitué la base d'informations économiques notariales. Cette base comprend la quasi-totalité des transactions immobilières ayant eu lieu en région parisienne depuis une quinzaine d'années - le taux de couverture est compris entre 80 et 90% depuis 1997. Nous ne nous intéressons qu'à certaines variables de cette base: caractéristiques détaillées du bien vendu (étage, surface, nombre de chambres, de salles de bain, présence d'un parking, d'un ascenseur, date de construction, localisation précise et montant de la transaction). Nous connaissons également la nationalité et le pays de résidence de l'acheteur et du vendeur. L'INSEE -l'Institut national de la statistique et des études économiques - utilise cette base pour produire les indices des prix de l'immobilier Insee - Notaires de la région parisienne. L'INSEE a d'ailleurs publié un document de travail ([Alain David \(2002\)](#)) qui fournit de nombreuses précisions sur la façon de construire ces indices et dont nous nous sommes largement inspirés pour calculer les prix hédoniques de chaque transaction (voir la section [2.2.1](#)).

Ces deux bases vont nous permettre de décrire succinctement le stock et le flux d'appartements à Paris.

## 1.1 Stock d'appartements

TABLE 5.1 – PARC DE LOGEMENTS PARISIEN. 1999-2010.

Année	Appartements	Maisons	(% du parc)
1999	1 379 428	5 817	(0,4%)
2001	1 377 119	5 873	(0,4%)
2003	1 378 854	5 979	(0,4%)
2005	1 379 189	6 007	(0,4%)
2007	1 378 871	6 043	(0,4%)
2009	1 380 520	6 211	(0,4%)
2010	1 380 879	6 251	(0,5%)

Source: Base FILOCOM

TABLE 5.2 – STATUT D'OCCUPATION DES LOGEMENTS. 1999-2010.

Année	Propriétaire	Locataires	Vacant	Occupés gratuitement	Meublés
1999	30,0%	53,8%	12,4%	3,6%	0,2%
2001	31,0%	54,3%	11,2%	3,4%	0,2%
2003	32,1%	54,1%	10,2%	3,2%	0,3%
2005	32,9%	53,4%	10,2%	3,2%	0,3%
2007	33,7%	53,0%	10,0%	3,0%	0,3%
2009	34,0%	53,0%	9,8%	2,8%	0,3%
2010	34,1%	53,3%	9,4%	2,8%	0,3%

*Note:* Les maisons sont exclues. Propriétaires signifie que le logement est occupé par son propriétaire, locataire par un locataire. Source: Base FILOCOM

TABLE 5.3 – MODE D'OCCUPATION DES LOGEMENTS. 1999-2010.

Année	Résidence principale	Résidence secondaire	Vacant
1999	80,2%	7,4%	12,4%
2001	81,5%	7,3%	11,2%
2003	81,8%	8,0%	10,2%
2005	81,7%	8,1%	10,2%
2007	81,5%	8,5%	10,0%
2009	81,5%	8,7%	9,8%
2010	81,6%	8,9%	9,4%

*Note:* Les maisons et les locations meublées sont exclues. Source: Base FILOCOM

TABLE 5.4 – TYPE DE PROPRIÉTAIRES DES LOGEMENTS. 1999-2010.

Année	Personne physique	Personne morale	HLM, SEM
1999	65,3%	19,9%	14,8%
2001	68,3%	20,5%	11,2%
2003	68,3%	20,0%	11,7%
2005	68,5%	19,7%	11,8%
2007	68,4%	18,8%	12,8%
2009	68,1%	17,5%	14,4%
2010	67,9%	17,4%	14,7%

*Note:* Les maisons et les locations meublées sont exclues. HLM, SEM comprend aussi les logements détenus par les collectivités publiques. Source: Base FILOCOM

TABLE 5.5 – MODE D'OCCUPATION DES LOGEMENTS SELON LE LIEU DE RÉSIDENCE DES PROPRIÉTAIRES. 2005-2010.

Mode d'occupation	Lieu de résidence du propriétaire	
	Étranger	France et étranger
Loué ou vacant	55,6%	57,7%
Occupé par propriétaire, résidence principale	6,9%	35,9%
Occupé par propriétaire, résidence secondaire	37,5%	6,4%
Total	100,0%	100,0%

*Note:* Les maisons, les locations meublées et le parc possédé par les sociétés HLM, SLM ou collectivités publiques sont exclues. Les logement occupés par leurs propriétaires incluent les logements occupés à titre à gratuit. Source: Base FILOCOM



TABLE 5.6 – PART DES PROPRIÉTAIRES ÉTRANGERS SELON LE MODE D'OCCUPATION.  
2005-2010.

Mode d'occupation	Proportion de propriétaires étrangers				
	2005	2007	2009	2010	2005-2010
Loué ou vacant	2,1%	2,4%	2,7%	3,2%	2,6%
Occupé par propriétaire, résidence principale	0,5%	0,5%	0,5%	0,6%	0,5%
Occupé par propriétaire, résidence secondaire	14,5%	15,3%	16,1%	16,8%	15,7%
Total	2,3%	2,5%	2,8%	3,2%	2,7%

*Note:* Les maisons, les locations meublées et le parc possédé par les sociétés HLM, SLM ou collectivités publiques sont exclues. Les logement occupés par leurs propriétaires incluent les logements occupés à titre à gratuit. Les propriétaires étrangers sont les propriétaires résidant à l'étranger. Source: Base FILOCOM

TABLE 5.7 – TYPE DE PROPRIÉTAIRE SELON LE LIEU DE RÉSIDENCE. 2005-2010.

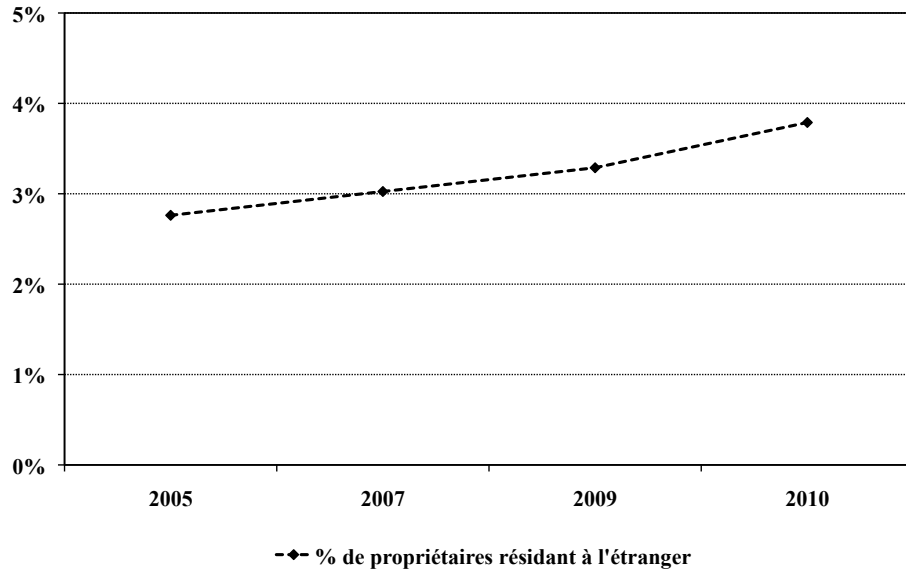
Type de propriétaire	Lieu de résidence du propriétaire	
	Étranger	France et étranger
Personne morale	5,4%	21,2%
Personne physique	94,6%	78,8%
Total	100,0%	100,0%

*Note:* Les maisons, les locations meublées et le parc possédé par les sociétés HLM, SLM ou collectivités publiques sont exclues. Les logement occupés par leurs propriétaires incluent les logements occupés à titre à gratuit. Les propriétaires étrangers sont les propriétaires résidant à l'étranger. Source: Base FILOCOM

TABLE 5.8 – PART DES PROPRIÉTAIRES ÉTRANGERS SELON LEUR TYPE. 2005-2010.

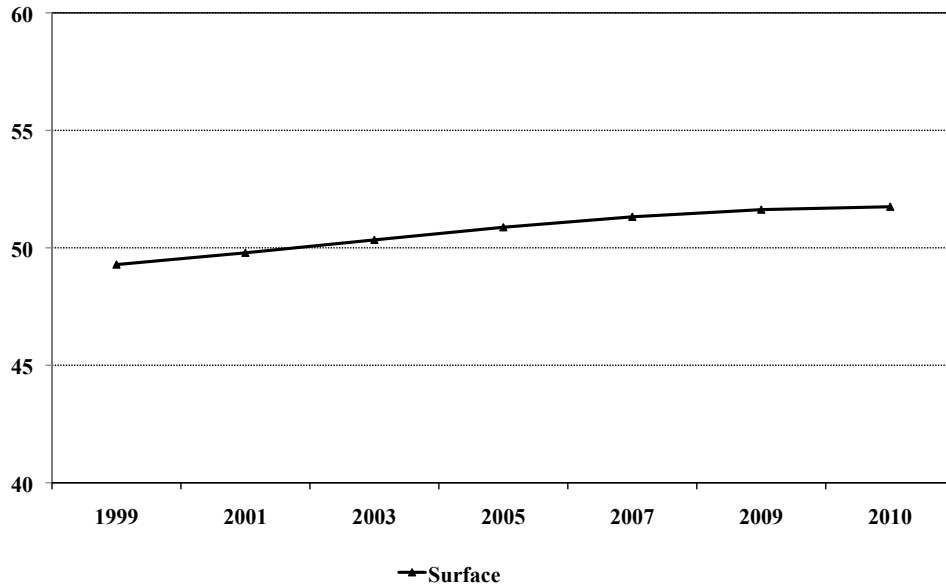
Type de propriétaire	Proportion de propriétaires étrangers				
	2005	2007	2009	2010	2005-2010
Personne morale	0,6%	0,6%	0,8%	0,8%	0,7%
Personne physique	2,8%	3,0%	3,3%	3,8%	3,2%
Total	2,3%	2,5%	2,8%	3,2%	2,7%

*Note:* Les maisons, les locations meublées et le parc possédé par les sociétés HLM, SLM ou collectivités publiques sont exclues. Les logement occupés par leurs propriétaires incluent les logements occupés à titre à gratuit. Les propriétaires étrangers sont les propriétaires résidant à l'étranger. Source: Base FILOCOM



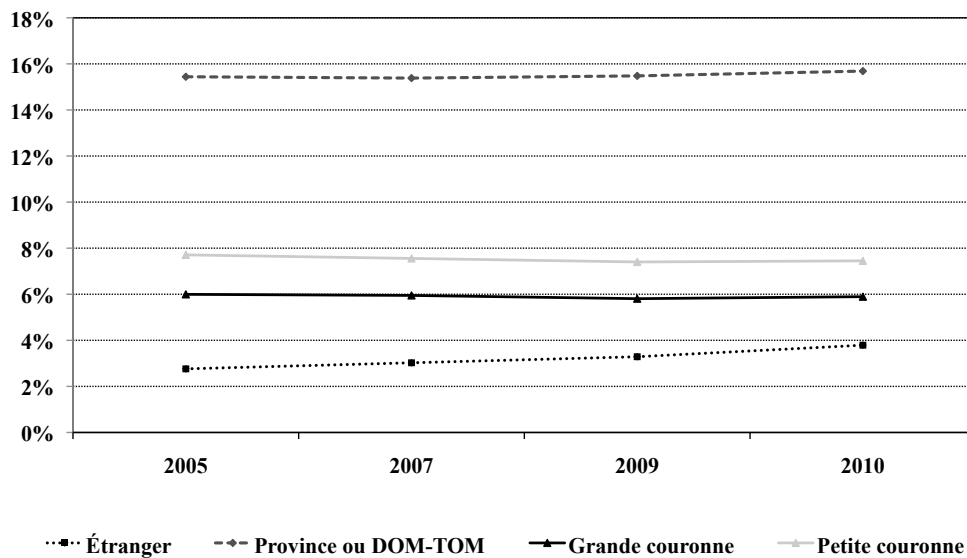
*Note:* Les maisons, les locations meublées sont exclues. Seul le parc des propriétaires physiques est considéré. Les propriétaires étrangers sont les propriétaires résidant à l'étranger. Source: Base FILOCOM.

FIGURE 5.1 – PART DES PROPRIÉTAIRES RÉSIDANT À L'ÉTRANGER. 2005-2010.



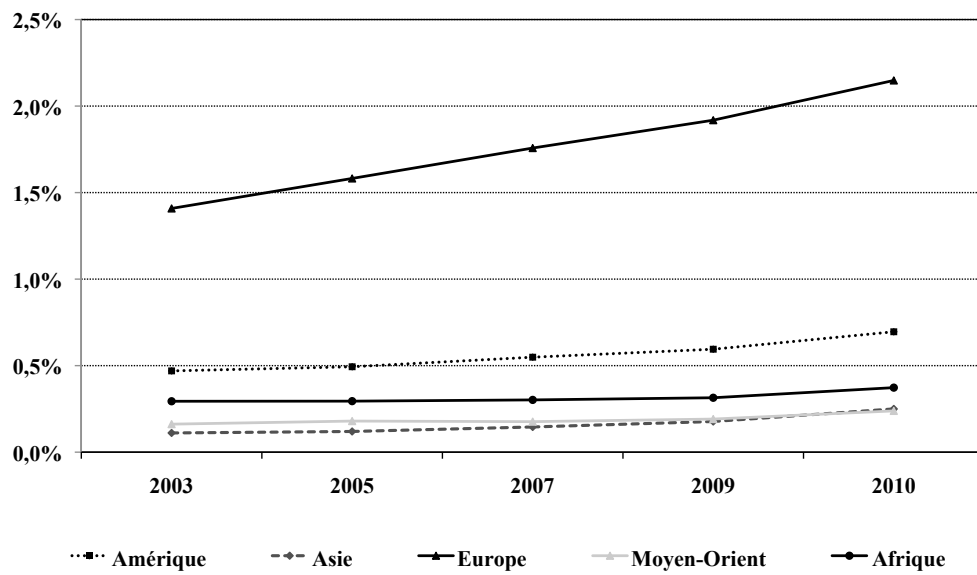
*Note:* Les maisons, les locations meublées sont exclues. Seul le parc des propriétaires physiques est considéré. Les propriétaires étrangers sont les propriétaires résidant à l'étranger. Source: Base FILOCOM.

FIGURE 5.2 – SURFACE MOYENNE DU PARC DE LOGEMENT. 1999-2010.



*Note:* Les maisons, les locations meublées sont exclues. Seul le parc des propriétaires physiques est considéré. Les propriétaires étrangers sont les propriétaires résidant à l'étranger. Source: Base FILOCOM.

FIGURE 5.3 – DÉPARTEMENT DE RÉSIDENCE DES PROPRIÉTAIRES. 2005-2010.



*Note:* Les maisons, les locations meublées sont exclues. Seul le parc des propriétaires physiques est considéré. Les propriétaires étrangers sont les propriétaires résidant à l'étranger. Source: Base FILOCOM.

FIGURE 5.4 – CONTINENT DE RÉSIDENCE DES PROPRIÉTAIRES ÉTRANGERS. 2003-2010.

TABLE 5.9 – NOMBRE D'APPARTEMENTS ET DE MAISONS VENDUS À PARIS. 1993-2008.

Année	Appartements	Maisons
1993	18 982	101
1994	24 338	123
1995	20 514	116
1996	29 593	225
1997	29 364	193
1998	33 469	242
1999	40 486	297
2000	39 953	218
2001	35 188	266
2002	37 236	200
2003	35 472	239
2004	40 409	185
2006	21 907	96
2008	28 066	144

*Note:* Le taux de couverture de la base a augmenté depuis 1990. Ces chiffres ne permettent pas tels quels de connaître l'évolution du volume d'appartements vendus. Ils nécessiteraient une correction - voir le document de travail [Alain David \(2002\)](#) pour plus de précisions. Source: Base BIEN

Le parc de logement à Paris est constituée en très grande majorité d'appartements. Les maisons représentent moins de 1% du parc, comme le montre le tableau [5.1](#).

Plus de 50% des logements sont occupés par des locataires, près d'un tiers des logements sont occupés par leurs propriétaires et environ 10% des logements sont vacants (voir le tableau [5.2](#)).

Une très grande majorité des logements (plus de 80%) servent de résidence principale, moins de 10% des logements sont utilisées comme résidence secondaire (voir le tableau [5.3](#)).

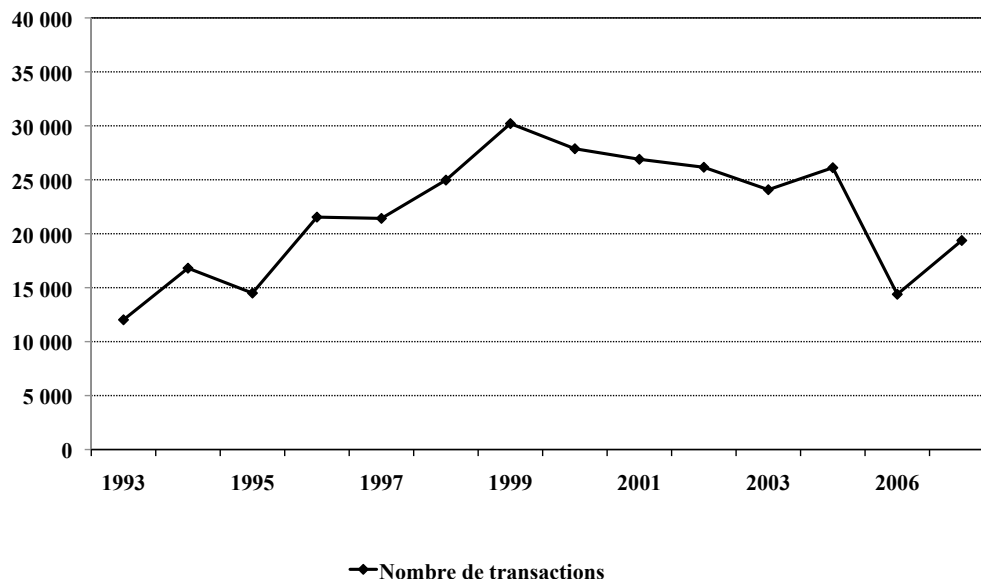
TABLE 5.10 – RÉPARTITION DES VENTES PAR TYPE D'ACHETEURS. 1993-2008.

Année	Particuliers	SCI	Marchands	Social	Inconnu
1993	78,8%	0,0%	3,8%	2,4%	15,0%
1994	82,7%	2,2%	1,5%	2,6%	11,0%
1995	84,0%	3,0%	1,3%	2,5%	9,2%
1996	87,5%	2,6%	1,1%	1,1%	7,7%
1997	86,4%	3,2%	1,3%	1,4%	7,8%
1998	85,9%	3,4%	1,4%	1,6%	7,7%
1999	85,9%	3,4%	1,1%	1,6%	8,0%
2000	85,3%	4,4%	1,2%	1,6%	7,5%
2001	86,1%	4,9%	1,3%	1,5%	6,2%
2002	86,4%	5,7%	1,4%	1,2%	5,4%
2003	84,6%	5,9%	1,8%	1,1%	6,6%
2004	84,3%	6,4%	1,9%	1,1%	6,3%
2006	83,8%	6,5%	2,1%	0,7%	6,9%
2008	83,9%	7,2%	2,5%	1,0%	5,5%

*Note:* Les maisons sont exclues. Les SCI sont les sociétés civiles immobilières. Le "social" englobe le service public, les entreprise et le secteur social. Les marchands sont les marchands de bien. Source: Base BIEN

Cependant, la part des résidences secondaires dans le parc de logements à Paris a augmenté de 1,5 points de % entre 1999 et 2010, tandis que la part des logements vacants dans le parc a, elle, diminué de 3 points de %. Ce dernier point suggère que le marché immobilier parisien est de plus en plus tendu. La part des logements qui sont loués meublés est très faible -environ 0,3% - et est stable dans le temps. Nous excluons ces logements dans le reste de notre description du stock de logements.

70% des appartements parisiens sont possédés par des personnes physiques, 20% par des personnes morales, les sociétés HLM ou SLM et les collectivités publiques possèdent le reste des appartements. La part des propriétaires personnes physiques a augmenté de 2,7 points



*Note:* Nous n'avons conservé que les transactions d'appartements vendus en pleine propriété, de gré à gré, sans viager, à usage d'habitation, libres au moment de la vente, dont le prix est renseigné, dont soit le nombre de pièces, soit la surface habitable sont également renseignés, et dont l'acheteur est un particulier. Nous avons également exclu les logements atypiques. Source: Base BIEN.

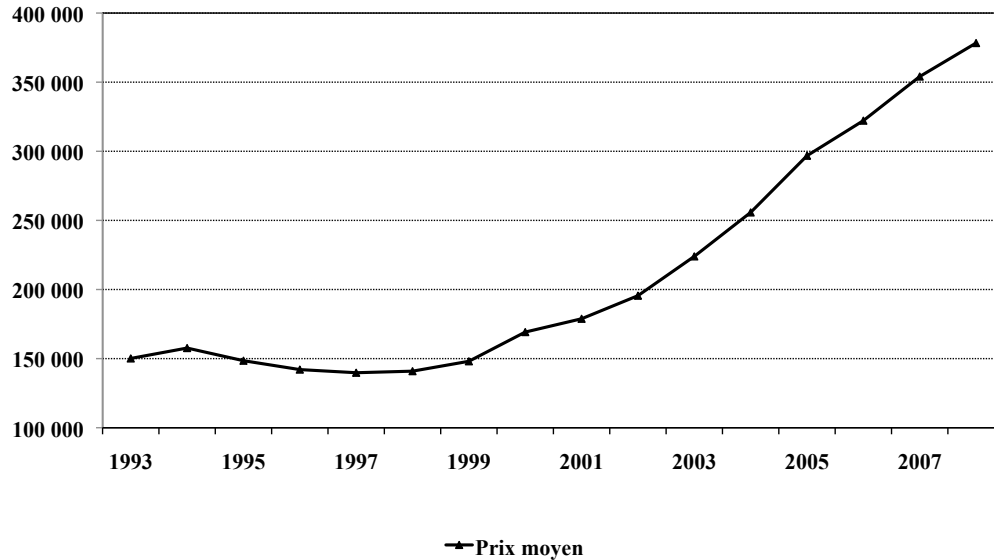
FIGURE 5.5 – NOMBRE DE TRANSACTIONS DANS NOTRE BASE. 1993-2008.

de % entre 1999 et 2010 tandis que celle des propriétaires personnes morales a diminué de 2,5 points de % (voir le tableau 5.4). Dans la suite, nous excluons les appartements possédés par des sociétés HLM, SEM ou des collectivités publiques car ils ne sont pas pertinents pour notre étude.

Le parc de logement est réparti entre les logements loués<sup>4</sup> (60% du parc), les logements occupés par leur propriétaire<sup>5</sup> au titre de leur résidence principale (36% du parc) et les logements occupés par leur propriétaire au titre de leur résidence secondaires (6,5% du parc). Si l'on se restreint aux logements possédés par des propriétaires physiques résidant à l'étranger, cette répartition change beaucoup: sur l'ensemble des logements possédés par des étrangers, près de 40% sont utilisés en tant que résidence secondaire (voir le tableau 5.5). D'ailleurs, les étrangers possèdent environ 16% de l'ensemble des résidences secondaires à Paris (voir le

4. Ils peuvent également être vacants.

5. Ils peuvent également être occupés à titre gratuit.



*Note:* Nous n'avons conservé que les transactions d'appartements vendus en pleine propriété, de gré à gré, sans viager, à usage d'habitation, libres au moment de la vente, dont le prix est renseigné, dont soit le nombre de pièces, soit la surface habitable sont également renseignés, et dont l'acheteur est un particulier. Nous avons également exclu les logements atypiques. Source: Base BIEN.

FIGURE 5.6 – ÉVOLUTION DU PRIX MOYEN DES APPARTEMENTS VENDUS. 1993-2008.

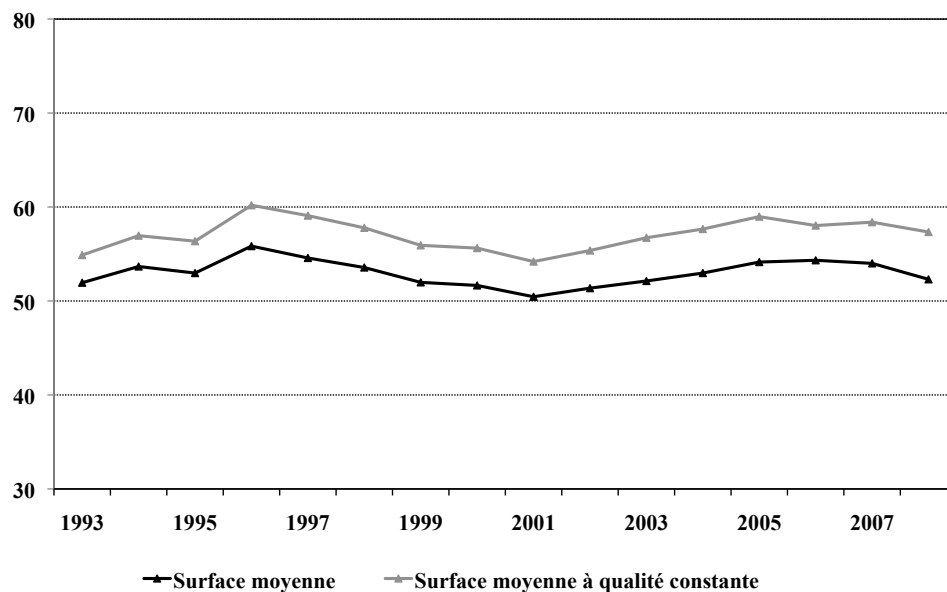
tableau 5.6). La proportion d'étrangers possédant une résidence secondaire a augmenté de 2,3 points de % entre 2005 et 2010.

Près de 80% du parc<sup>6</sup> est possédé par des propriétaires personnes physiques. Les propriétaires étrangers, eux, sont constitués à 95% de personnes physiques (voir le tableau 5.7). Parmi les propriétaires personnes morales, la part des propriétaires étrangers est très faible - environ 0,7%- et a très peu augmenté entre 2005 et 2010 (voir le tableau 5.4). Par opposition, la proportion de propriétaires étrangers parmi les propriétaires personnes physiques est plus importante - environ 3,2% - et a augmenté de 1 point de % entre 2005 et 2010 (voir également le graphique 5.1). Dans la suite de notre étude, nous ne considérons que les transactions de logements effectués par des acheteurs personnes physiques. Ce faisant, ne coure-t-on pas le risque de sous-estimer l'influence des propriétaires étrangers sur le marché immobilier

---

6. Nous excluons les maisons, les locations meublées et les logements loués par des collectivités publiques ou des sociétés HLM ou SEM.





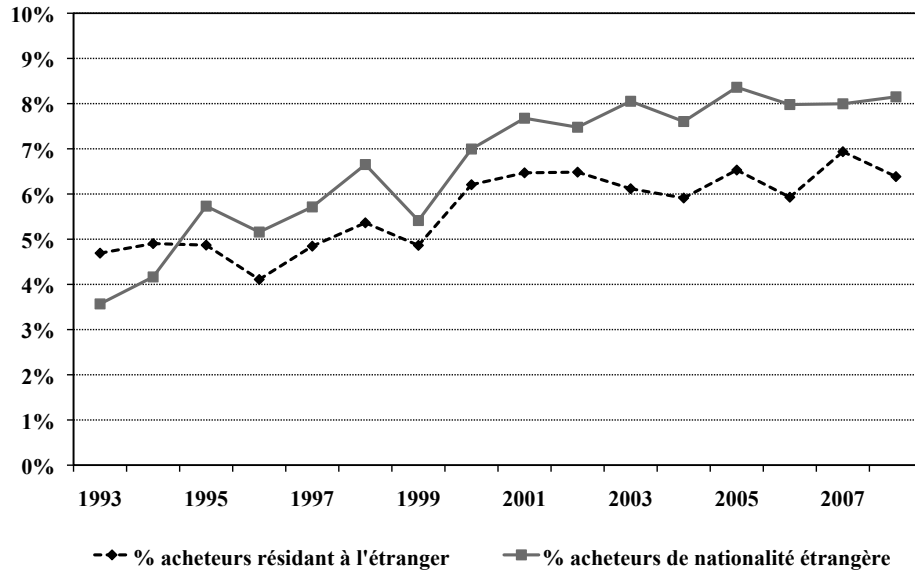
*Note:* La surface à qualité constante d'un appartement correspond à la surface que devrait avoir un appartement pour qu'il soit vendu au même prix, mais en ayant les caractéristiques du logement "typique". Se référer à la section 2.2.1 pour avoir plus de précisions sur la façon de calculer la surface à qualité constante d'un appartement et la définition du logement "typique".

Nous n'avons conservé que les transactions d'appartements vendus en pleine propriété, de gré à gré, sans viager, à usage d'habitation, libres au moment de la vente, dont le prix est renseigné, dont soit le nombre de pièces, soit la surface habitable sont également renseignés, et dont l'acheteur est un particulier. Nous avons également exclu les logements atypiques. Source: Base BIEN.

FIGURE 5.7 – SURFACE MOYENNE DES APPARTEMENTS VENDUS. 1993-2008.

parisien? Ces quelques chiffres nous assurent du contraire. Ils démontrent en effet que la part des propriétaires étrangers parmi les propriétaires personne morale est très faible, presque négligeable et que l'immense majorité des propriétaires étrangers est constituée de propriétaire personne physique.

Selon le graphique 5.1, la part des propriétaires résidant à l'étranger a augmenté entre 2005 et 2010, passant de moins de 3% à presque 4%. Plus généralement, selon le graphique 5.3, la part des propriétaires résidant en province, petite ou grande couronne, ou dans les DOM-TOM a diminué entre 2005 et 2010. Parmi les propriétaires ne résidant pas à Paris, seule la part des propriétaires étrangers a augmenté. On peut aller plus loin et décomposer les propriétaires



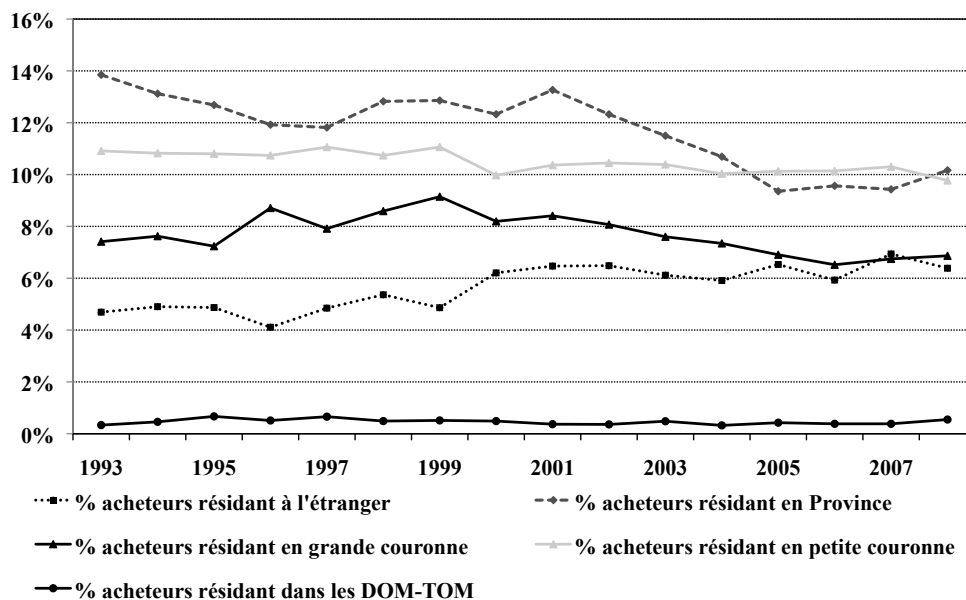
*Note:* Deux définitions du pourcentage d'acheteurs étrangers peuvent être utilisées: le pourcentage d'acheteurs de nationalité étrangère et le pourcentage d'acheteurs résidant à l'étranger. Source: Base BIEN.

FIGURE 5.8 – PART DES ACHETEURS ÉTRANGERS. 1993-2008.

étrangers en cinq catégories (Amérique, Asie, Europe, Moyen-Orient et Afrique) selon leur continent de résidence. Par commodité, nous dirons qu'un propriétaire vient d'Amérique si son pays de résidence est situé en Amérique ou en Australie. Nous invitons le lecteur à se référer à l'annexe B pour obtenir plus de précisions sur la façon dont les groupes de pays ont été constitués. Selon le graphique 5.4, la part des propriétaires de chacun de ces cinq groupes a augmenté mais l'augmentation a été plus importante pour les propriétaires en provenance d'Europe et d'Amérique. Les propriétaires étrangers vivent d'ailleurs en grande majorité en Europe puis dans une moindre mesure en Amérique.

## 1.2 Flux d'appartements.

Compte tenu de la très faible part des maisons dans le parc de logements à Paris, il n'est pas surprenant que le nombre de maison vendues comparé au nombre d'appartements vendus soit très faible (voir le tableau 5.9). Par la suite, nous pouvons sans regret exclure les maisons



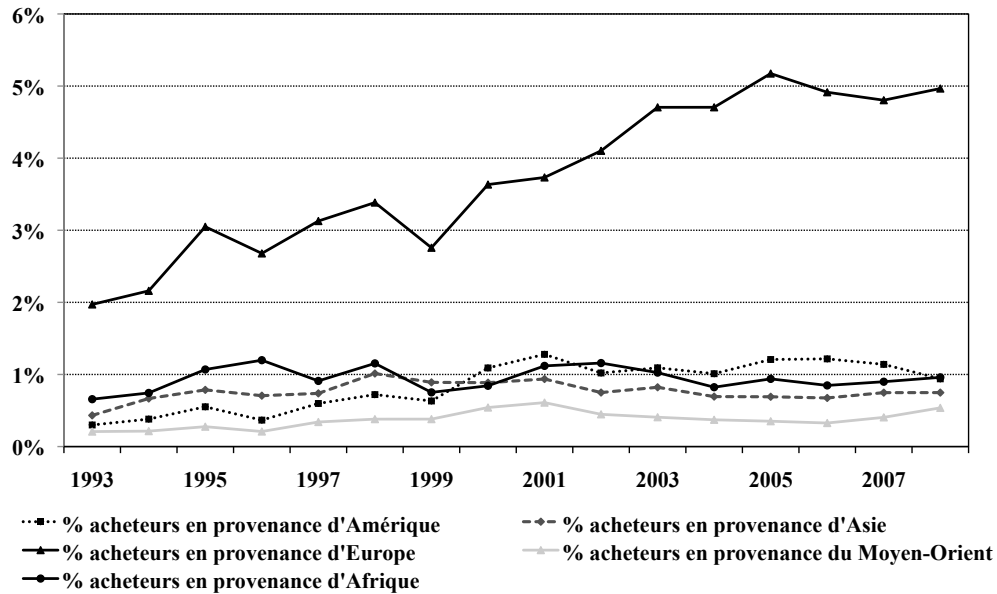
Note: La courbe "% acheteurs résidant à l'étranger" représente la part des acheteurs qui résident à l'étranger. En majorité, les acheteurs résident à Paris. La part des acheteurs résidant à Paris n'est pas représentée mais se déduit aisément de la somme des autres courbes. Source: Base BIEN.

FIGURE 5.9 – DÉPARTEMENT DE RÉSIDENCE DES ACHETEURS. 1993-2008.

de notre étude - leurs caractéristiques sont trop différentes de celles des appartements pour que l'on puisse les utiliser. Le tableau 5.9 ne nous permet pas de connaître précisément l'évolution du flux de transactions car le taux de couverture de la base BIEN a augmenté depuis 1990.

Le tableau 5.10 indique l'évolution de la répartition des acheteurs selon leur type. Les acheteurs peuvent soit être des personnes physiques soit des personnes morales. Les acheteurs personnes physiques sont constitués des particuliers et des sociétés civiles immobilières. Parmi les acheteurs personnes morales, on distingue les marchands de biens et des acheteurs du service public, du secteur social ou des entreprises. Il reste environ 10% des acheteurs dont le type est inconnu, c'est-à-dire que le type n'est pas renseigné dans la base BIEN. La proportion d'acheteurs de type inconnu est passée de 15 à 5,5% entre 1993 et 2008. Durant la même période, la proportion de particuliers a augmenté de 5,1 points de % et la proportion de SCI<sup>7</sup> de 7,2 points de %. La proportion de marchands de biens a, elle, baissé de 1,3 points de %.

7. Sociétés civiles immobilières



*Note:* La courbe "% acheteurs résidant en Amérique" représente la part des acheteurs dont le pays de résidence est situé en Amérique. En majorité, les acheteurs résident en France. La part des acheteurs résidant en France n'est pas représentée mais se déduit aisément de la somme des autres courbes. Source: Base BIEN.

FIGURE 5.10 – CONTINENT DE RÉSIDENCE DES ACHETEURS. 1993-2008.

Nous pouvons donc être sûrs que la proportion des marchands de biens a diminué.

En nous inspirant du document de travail [Alain David \(2002\)](#), nous ne conservons dans la base que les transactions d'appartements vendus en pleine propriété, de gré à gré, sans viager, à usage d'habitation, libres au moment de la vente, dont le prix est renseigné, et dont soit le nombre de pièces, soit la surface habitable sont également renseignés. Nous excluons également les logements atypiques comme les greniers, les chambres de service, les lofts, les loges de gardien, les chalets et les grandes propriétés. Toutes ces restrictions éliminent moins de 20% des transactions. Le graphique [5.5](#) représente le nombre de transactions dans la base nettoyée.

Par la suite, nous nous restreignons en outre aux acheteurs particuliers. Nous avons testé la robustesse de nos résultats lorsque nous considérons l'ensemble des transactions dont l'acheteur est un particulier, une société immobilière ou dont le type est inconnu. Ces résultats sont disponibles dans l'annexe [D.2](#).

Selon le graphique 5.6, le prix moyen des appartements vendus à Paris est passé de 150 000 euros en 1993 à plus de 375 000 euros en 2008, soit une augmentation de plus de 150%. Cette augmentation n'est pas due à une augmentation de la surface ou de la qualité des appartements vendus. En effet, selon le graphique 5.7, la surface moyenne des appartements vendus est restée relativement stable pendant la période considérée. Ce même graphique représente également la surface moyenne "à qualité constante" des appartements vendus. Qu'est ce que la surface "à qualité constante" d'un appartement?

Idéalement, nous souhaiterions nous assurer que la hausse des prix de l'immobilier n'est pas simplement due à une amélioration de la qualité des logements. Considérons une transaction d'un appartement au prix  $P$ . Cet appartement est par exemple situé au quatrième étage d'un immeuble sans ascenseur. Il mesure 50 m<sup>2</sup>. Toutes choses étant égales par ailleurs, nous cherchons à imaginer la surface fictive que devrait avoir ce même appartement, pour qu'il soit vendu au même prix, s'il était situé au rez-de-chaussée du même immeuble et non plus au quatrième, c'est la surface à qualité constante<sup>8</sup>.

Selon le graphique 5.7, la surface moyenne à qualité constante des appartements vendus est elle aussi restée relativement stable au cours de la période. ce n'est donc pas une augmentation substantielle de la surface ou de la qualité des appartements vendus qui est responsable de la forte hausse des prix.

Selon le graphique 5.8, le pourcentage d'acheteurs étrangers a cru entre 1993 et 2008. Le pourcentage d'acheteurs de nationalité étrangère est passé de moins de 4% à 8% et le pourcentage d'acheteurs résidant à l'étranger est passé de presque 5% à plus de 6%. Selon le graphique 5.9, la proportion d'acheteurs résidant en province, petite, grande couronne ou dans les DOM-TOM a baissé entre 1993 et 2008. Selon le graphique 5.10, la proportion d'acheteurs d'Europe - dont le pays d'origine est situé en Europe mais pas en France - a beaucoup augmenté sur la période, passant de 2 à 5%. La proportion d'acheteurs dont le pays d'origine est situé en Amérique a elle aussi augmenté, passant de moins 0,5 à 1%.

En résumé, sur la période 1993-2008, la proportion d'acheteurs étrangers a cru et les prix

---

8. En fait, la surface à qualité constante prend en compte toutes les caractéristiques observables des logements.

ont fortement augmenté. Nous verrons dans la partie 4 s'il y a un lien de causalité entre les deux phénomènes.

On pourrait craindre par exemple que les étrangers soient de plus en plus nombreux à acheter des appartements à Paris via des marchands de biens. Le tableau 5.10 nous assure du contraire.

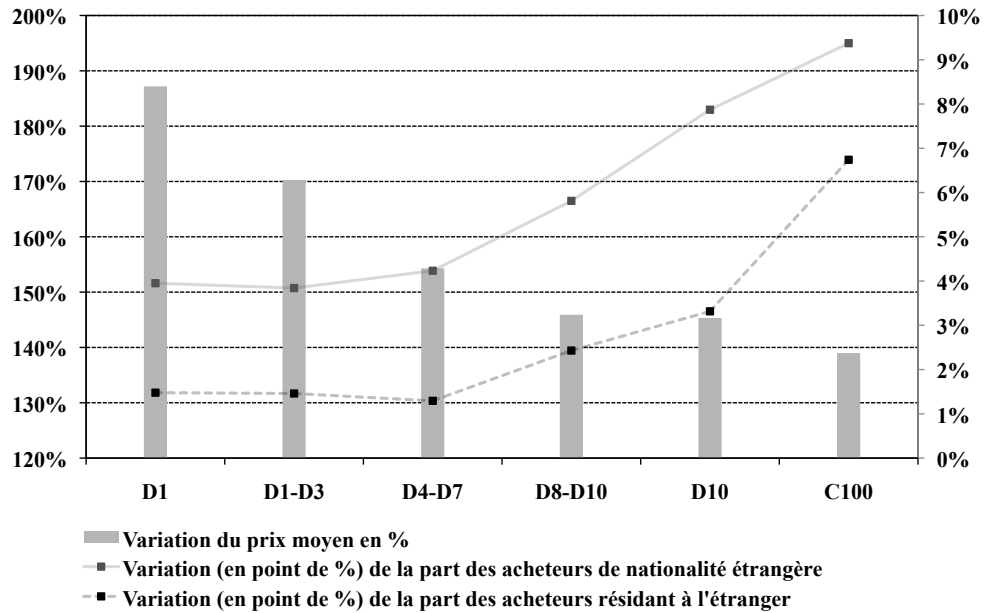
### 1.3 Flux d'appartements vendus par décile de prix.

Il est intéressant de décomposer les transactions par gamme de prix. Les augmentations de prix ont-elle été les mêmes pour les transactions les moins chères et les plus chères? Qu'en est-il du pourcentage d'acheteurs étrangers?

Selon le graphique 5.11, le prix moyen des 10% de transactions les moins chères a augmenté de presque 190% entre 1993 et 2008 alors que le prix moyen des 1% de transactions les plus chères a seulement augmenté de 140%. Plus généralement, plus on diminue dans les gammes de prix, plus les augmentations de prix ont été importantes. Toujours selon le graphique 5.11, le pourcentage de résidents étrangers parmi les acheteurs des 10% de transactions les moins chères a cru de moins d'1,5 points de % tandis que le pourcentage de résidents étrangers parmi les acheteurs des 1% de transactions les plus chères a cru de presque 7% sur la période. Plus généralement, plus on augmente dans les gammes de prix, plus l'augmentation du pourcentage d'acheteurs étrangers (que l'on considère les acheteurs de nationalité étrangère ou les acheteurs résidant à l'étranger) a été importante.

Étrangement, plus on considère des gammes de prix élevés, plus l'augmentation de la proportion d'acheteurs étrangers a été importante et plus l'augmentation de prix a été, elle, mesurée. Les étrangers n'ont pas l'air d'être responsable de la hausse très forte du prix des logements les moins chers (précisons qu'à priori, les logements les moins chers sont aussi les plus petits et sont occupés par les ménages les plus modestes). Cette hausse s'explique par un autre phénomène. Dans Fack (2007), Gabrielle Fack étudie l'effet qu'a eu l'extension des aides au logement (à la personne) au début des années 90 sur les prix des loyers des ménages les plus modestes. Elle démontre que l'extension des critères d'éligibilité aux aides à la personne

a augmenté la demande des ménages bénéficiaires de ces aides. Cette augmentation de la demande couplée à la très faible élasticité de l'offre de logements des bailleurs à court et moyen terme a entraîné une forte hausse des loyers des petits logements. L'auteur estime que les aides à la personne sont responsables de 58% de l'augmentation des loyers des ménages pauvres entre 1984 et 2002, sans que cette hausse se soit traduite par une forte amélioration de la qualité des logements. La hausse des loyers des logements les plus petits a augmenté leur rentabilité locative. Ceci explique en moins en partie la plus forte hausse du prix des logements les moins chers (relativement à l'ensemble des logements) que l'on observe sur le graphique 5.11.



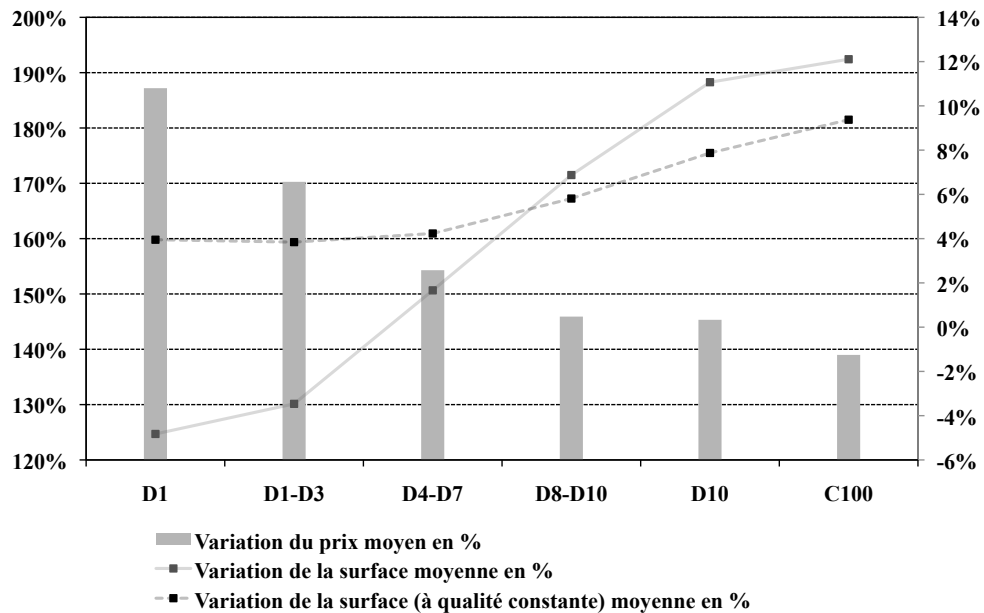
*Note:* D1 correspond aux 10% de transactions les moins chères. D1-D3 correspond aux 30% de transactions les moins chères. C100 correspond aux 1% de transactions les plus chères.

La définition de prix utilisée pour classer les transactions est le prix brut. Se référer au graphique 5.13 pour une classification des transactions par prix au m<sup>2</sup> équivalent. Se référer à la section 2.2.1 pour plus de précisions sur les prix au m<sup>2</sup> équivalents.

Les barres représentent la variation en % du prix moyen. Selon la première barre, le prix moyen des 10% de transactions les moins chères a augmenté de presque 190% entre 1993 et 2008. La courbe "Variation (en points de %) de la part des acheteurs de nationalité étrangère" représente la variation (en points de %) de la proportion d'acheteurs de nationalité étrangère. Selon le premier point de la courbe, le pourcentage d'acheteurs de nationalité étrangère parmi les 10% de transactions les moins chères a augmenté de 4 points entre 1993 et 2008. Source: Base BIEN.

FIGURE 5.11 – VARIATION DU PRIX MOYEN DES TRANSACTIONS ET DU POURCENTAGE D'ACHETEURS ÉTRANGERS PAR DÉCILE DE PRIX, 1993-2008.





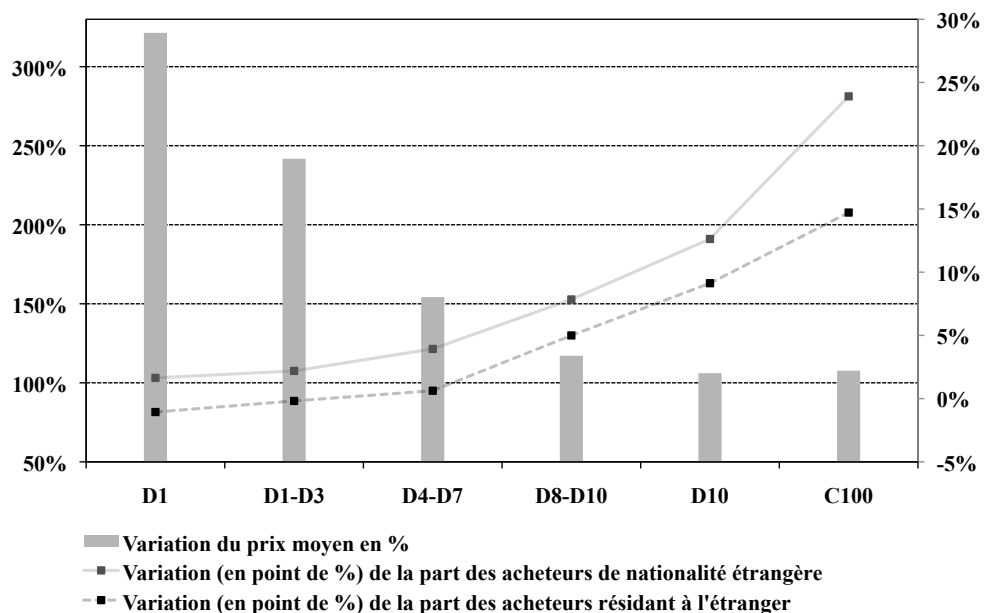
*Note:* D1 correspond aux 10% de transactions les moins chères. D1-D3 correspond aux 30% de transactions les moins chères. C100 correspond aux 1% de transactions les plus chères.

La définition de prix utilisée pour classer les transactions est le prix brut.

La surface à qualité constante d'un appartement correspond à la surface qu'il devrait avoir pour qu'il soit vendu au même prix, mais en ayant les caractéristiques du logement "typique". Se référer à la section 2.2.1 pour avoir plus de précisions sur la façon de calculer la surface à qualité constante et la définition du logement "typique".

Les barres représentent la variation en % du prix moyen. Selon la première barre, le prix moyen des 10% de transactions les moins chères a augmenté de presque 190% entre 1993 et 2008. Selon le premier point de la courbe "Variation de la surface moyenne en %", la surface moyenne des 10% de transactions les moins chères a baissé de 5% entre 1993 et 2008. Source: Base BIEN.

FIGURE 5.12 – VARIATION DU PRIX MOYEN ET DE LA SURFACE MOYENNE DES TRANSACTIONS PAR DÉCILE DE PRIX. 1993-2008.



*Note:* D1 correspond aux 10% de transactions les moins chères. D1-D3 correspond aux 30% de transactions les moins chères. C100 correspond aux 1% de transactions les plus chères.

La définition de prix utilisée pour classer les transactions est le prix au m<sup>2</sup> équivalent.

Les barres représentent la variation en % du prix moyen. Selon la première barre, le prix moyen des 10% de transactions les moins chères a augmenté de presque 190% entre 1993 et 2008. La courbe "Variation (en points de %) de la part des acheteurs de nationalité étrangère" représente la variation (en points de %) de la proportion d'acheteurs de nationalité étrangère. Selon le premier point de la courbe, le pourcentage d'acheteurs de nationalité étrangère parmi les 10% de transactions les moins chères a augmenté de 4 points entre 1993 et 2008. Source: Base BIEN.

FIGURE 5.13 – VARIATION DU PRIX MOYEN DES TRANSACTIONS ET DU POURCENTAGE D'ACHETEURS ÉTRANGERS PAR DÉCILE DE PRIX HÉDONIQUES. 1993-2008.

## 2 Stratégie d'estimation

### 2.1 Quelques considérations théoriques

On considère une ville où l'offre de logement est fixe, égale à  $\bar{S}$ . Chaque unité de  $\bar{S}$  correspond à une unité de logement de surface et qualité constante.  $P$  est le prix sur le marché d'une unité de logement. Le nombre d'habitants  $L$  qui résident dans cette ville est fixe. On note  $\alpha(P)$  ( $\alpha$  décroît avec  $P$ ) la fonction de demande, identique pour tous les agents. La fonction de demande totale  $D(L, P)$  est égale à  $\alpha(P)L$ . On suppose aussi que des étrangers, au nombre de  $F$ , souhaitent acheter des logements dans la ville. On suppose tout d'abord qu'ils ont la même fonction de demande que les résidents. On pose  $N = L + F$  et  $E = \frac{F}{N}$  le pourcentage d'étrangers. à l'équilibre,

$$\bar{S} = \alpha(P)N \quad (5.1)$$

On suppose que le pourcentage d'étrangers passe de  $E$  à  $E' = E + \Delta E$ . à l'équilibre, on a maintenant

$$\bar{S} = \alpha(P')N' \quad (5.2)$$

avec  $P'$  le nouveau prix d'équilibre. Il est facile de voir que  $N' = \frac{L}{1-E'}$ . En utilisant cette dernière équation, et l'équation 5.1 et 5.2, on obtient facilement que

$$\frac{1 - E'}{1 - E} = \frac{\alpha(P')}{\alpha(P)} \quad (5.3)$$

On note  $\epsilon$  l'élasticité-prix du bien logement, on obtient alors  $\frac{1-E'}{1-E} = 1 - \epsilon \frac{\Delta P}{P}$ . L'augmentation de la proportion des étrangers entraîne donc une augmentation de prix égale à

$$\frac{\Delta P}{P} = \frac{\Delta E}{\epsilon} \frac{1}{1 - E} \quad (5.4)$$

et on a aussi

$$\epsilon = \frac{E}{\frac{\Delta P}{P}} \frac{1}{1 - E} \quad (5.5)$$

Ici, nous avons supposé que chaque unité de bien logement était identique. Supposons maintenant que la qualité des unités de logements est variable et se répercute dans les prix. Comment l'équation 5.4 est-elle transformée? Si les étrangers achètent en moyenne le même type de logements que les locaux, alors l'équation est inchangée. Par contre, si les étrangers achètent en moyenne des logements de meilleure qualité que l'ensemble des acheteurs, l'équation 5.4 est transformée. On note  $1 + y$  le prix relatif payé par les étrangers - c'est-à-dire le prix moyen payé par les étrangers divisé par le prix moyen payé par l'ensemble des acheteurs.  $y$  est strictement positif car les étrangers achètent en moyenne des logements plus chers. On a alors:

$$\begin{aligned} \frac{\Delta P}{P} &= \frac{\Delta(E(1+y))}{\epsilon} \frac{1}{1 - E(1+y)} \\ &\simeq \frac{\Delta(E(1+y))}{\epsilon} \frac{1}{1 - E} \end{aligned} \quad (5.6)$$

$$(5.7)$$

Une augmentation de la proportion d'étrangers entraîne une hausse des prix et ce, d'autant plus qu'ils paient un prix relatif élevé.

Pour mener notre raisonnement, nous avons supposé que le nombre d'habitants est fixe et que l'offre de logements est inélastique. Cette deuxième hypothèse est réaliste à Paris intramuros. Par contre, la première hypothèse l'est moins. Rien ne permet d'affirmer que les habitants de Paris désirant devenir propriétaire ne sont pas sensibles au prix des logements et ne sont pas susceptibles de renoncer à un achat à Paris pour acheter un bien de substitution relativement moins cher (un logement en petite couronne par exemple). La hausse de la demande provoquée par l'arrivée d'acheteurs potentiels étrangers peut à la limite être totalement compensée par un déplacement de la demande des parisiens vers la petite ou grande couronne. Dans ce cas limite, l'arrivée d'acheteurs potentiels étrangers n'aura aucun

effet sur les prix.

## 2.2 Méthodologie

Pour estimer l'impact des étrangers sur les prix de l'immobilier, nous procédons en trois étapes.

1. Nous découpons Paris en zones homogènes en terme de comportement d'achats immobiliers par les étrangers. Ces zones correspondent aux 80 quartiers administratifs de Paris<sup>9</sup>.
2. Comparer le montant de deux transactions n'est pas pertinent en soi: peut-on vraiment comparer les prix au m<sup>2</sup> d'un appartement situé au rez-de-chaussée d'un vieil immeuble et d'un autre appartement situé au quatrième étage d'un immeuble récent avec ascenseur? Pour pallier à cette difficulté, nous définissons un logement type, qui correspond à une liste de caractéristiques fixées (deux-pièces, de 20 à 30 m<sup>2</sup> chacune, avec salle de bains, sans pièce de service ni parking, situé au rez-de-chaussée d'un immeuble construit entre 1850 et 1914). Nous ramenons ensuite le prix au m<sup>2</sup> de chacune des transactions à un prix fictif, correspondant au montant de la même transaction *si le logement vendu possédait les caractéristiques du logement type*.
3. On calcule chaque année le pourcentage d'acheteurs étrangers et le prix moyen au m<sup>2</sup>- en utilisant les prix fictifs qu'on a calculé auparavant - par quartiers administratifs. On régresse alors le logarithme du prix moyen par le pourcentage d'acheteurs étrangers.

### 2.2.1 Prix hédoniques

Comment définir de façon satisfaisante le niveau des prix des appartements à Paris? On ne peut comparer les prix de deux appartements sans travail préalable. En effet, deux logements (à fortiori deux appartements) ne sont jamais identiques. Selon sa localisation géographique, un logement donne acc<sup>s</sup> à des biens publics, un voisinage, un environnement différent. Le type

---

9. Paris est découpé en 80 quartiers administratifs. Un arrondissement correspond à quatre quartiers administratifs.

de construction, l'étage, le nombre de pièces et le niveau de confort varient également d'un appartement à un autre. Le prix d'un appartement lors d'une transaction reflète l'ensemble de ses caractéristiques. En règle générale, il n'est donc guère pertinent de comparer le prix de deux appartements, quand bien même seraient-ils vendus la même année, dans le même arrondissement.

Pour comparer les prix de ventes d'appartements différents au sein d'un même arrondissement, nous adoptons l'approche hédonique utilisée par l'INSEE pour construire les indices de prix des logements anciens INSEE-Notaires [Alain David \(2002\)](#). On considère que le prix de vente pendant l'année  $n$  et le trimestre  $q$  d'un appartement situé dans l'arrondissement  $A$  et dans le quartier administratif  $Q$ , de caractéristiques observables

$X_i = (X_{1,i}, \dots, X_{K,i})$  vérifie :

$$\log p_{i,A,Q,n,q} = \sum_k^K \alpha_{k,A} X_{k,i} + \sum_{t'} \beta_{t',A} \mathbf{1}_{t'=n} + \sum_{q'} \gamma_{q',A} \mathbf{1}_{q'=q} + \sum_{Q'=1}^4 \zeta_{Q',A} \mathbf{1}_{Q'=Q} + \epsilon_{i,A,Q,n,q} \quad (5.8)$$

avec les  $\epsilon_{i,A,Q,n,q}$  indépendants. On remarque que les coefficients  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\zeta$  peuvent dépendre de l'arrondissement  $A$ , c'est-à-dire que les régressions sont effectuées arrondissement par arrondissement. Un exemple de régression hédonique est donné dans l'annexe C. Nous définissons un logement "type" et grâce à nos régressions, nous pouvons maintenant ramener le prix de chacune des transactions à *un prix "équivalent", correspondant au prix de la même transaction si le logement vendu avait les mêmes caractéristiques que le logement "type"*. Le logement "type" possède les caractéristiques suivantes: deux-pièces, de 20 à 30 m<sup>2</sup> chacune, avec salle de bains, sans pièce de service ni parking, situé au rez-de-chaussée d'un immeuble construit entre 1850 et 1914, vendu durant le dernier trimestre. Le prix "équivalent" de chaque transaction est obtenu par la formule:

$$\log p_{i,A,Q,n}^{eq} = \log p_{i,A,Q,n,q} - \left( \sum_k^K \alpha_{k,A} X_{k,i} + \sum_{q'} \gamma_{q',A} \mathbf{1}_{q'=q} \right) \quad (5.9)$$

Les prix "équivalents" peuvent être comparés entre eux car ils sont corrigés des différences de prix structurelles dues aux caractéristiques observables variables des appartements.

### 2.2.2 Régressions

On estime l'équation suivante:

$$\log \left( \sum_{i=1}^{N_{Q,n}} \frac{1}{N_{Q,n}} p_{i,A,Q,n}^{eq} \right) = \theta E_{Q,n} + \sum_{n',A'} \eta_{n',A'} \mathbf{1}_{n'=n} \mathbf{1}_{A'=A} + \sum_{Q'} \vartheta_{Q'} \quad (5.10)$$

$N$  est égal au nombre de transactions dans le quartier  $Q$  (appartenant à l'arrondissement  $A$ ) pendant l'année  $n$ . Les  $p_{i,A,Q,n}^{eq}$  sont les prix "équivalents" de toutes ces transactions.  $E_{Q,n}$  correspond à la proportion d'acheteurs étrangers dans le quartier  $Q$  durant l'année  $n$ . Les  $e_{Q,n}$  sont supposés indépendants. Dans certains quartiers administratifs (par exemple les quartiers administratifs du premier arrondissement), il y a très peu de transactions par an, ce qui induit une très grande variabilité dans le calcul de la part des acheteurs étrangers. Pour cette raison, nous avons regroupé les années par deux. Nous avons ensuite régressé le logarithme de la moyenne des prix au m<sup>2</sup> "équivalents" par le pourcentage d'acheteurs étrangers. Nous avons inclus des indicatrices temporelles et géographiques dans toutes nos régressions.

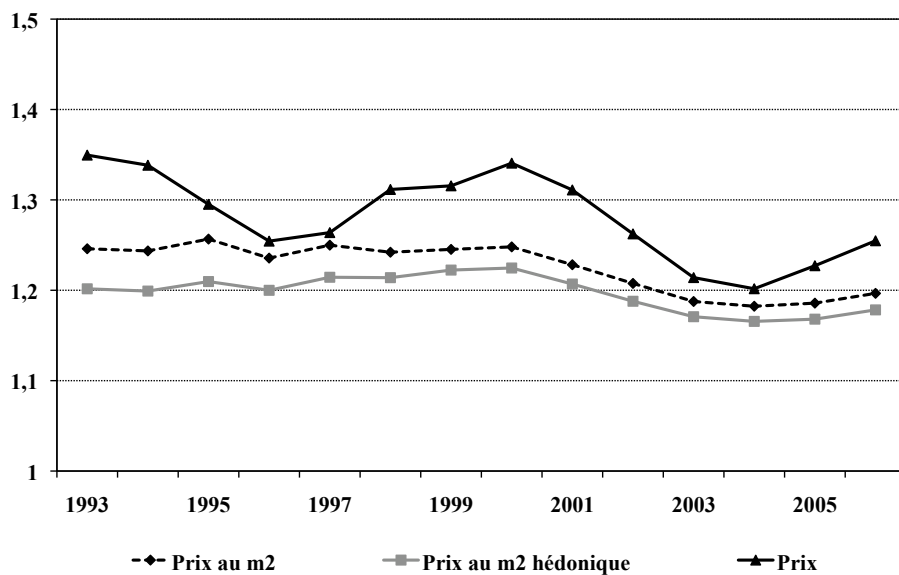
### 3 Résultats des estimations

Dans cette section, nous présentons le résultat de nos régressions.

Tout d'abord, nous analysons l'impact des acheteurs résidant à l'étranger ou de nationalité étrangère sur les prix, séparément puis simultanément. Nous décomposons ensuite les acheteurs de nationalité étrangère en cinq sous-groupes selon leur continent d'origine et étudions l'effet simultané de ces cinq groupes sur les prix de l'immobilier. De façon similaire, nous décomposons les acheteurs en six catégories selon leur lieu de résidence (Paris, petite couronne, grande couronne, province, DOM-TOM, étranger). Les acheteurs qui appartiennent aux cinq dernières catégories peuvent tous être considérés comme des étrangers car ils n'habitent pas à Paris. Nous étudions l'impact simultané de ces cinq catégories d'acheteurs sur les prix de l'immobilier à Paris. Enfin, nous approfondissons notre étude en nous restreignant au marché immobilier de luxe, c'est à dire aux 30% de transactions les plus chères.

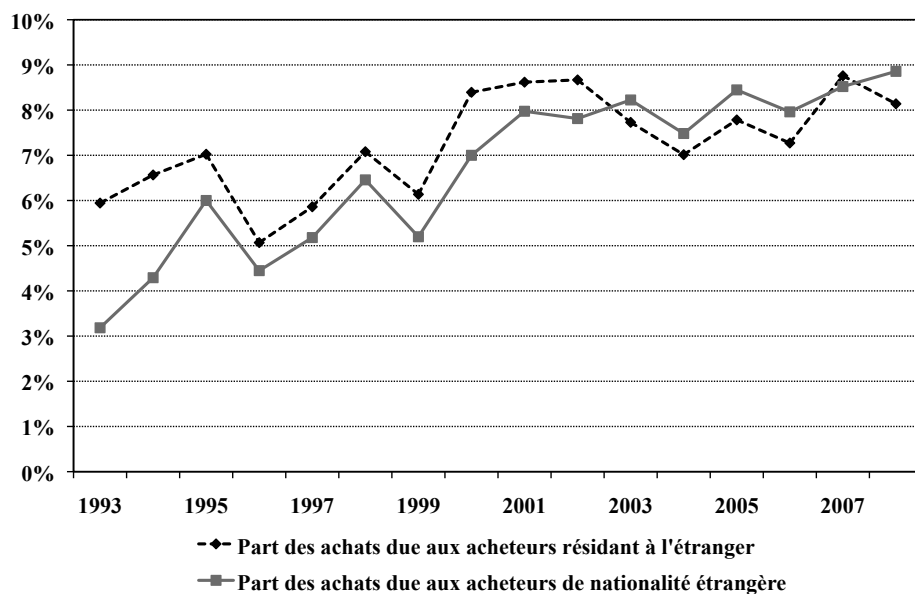
#### 3.1 Impact sur les prix des acheteurs résidant à l'étranger ou de nationalité étrangère





*Note:* La courbe "prix au m<sup>2</sup>" correspond au prix au m<sup>2</sup> moyen payé par les acheteurs résidant à l'étranger divisé par le prix au m<sup>2</sup> moyen payé par l'ensemble des acheteurs. Trois définitions du prix sont utilisées: le prix moyen, le prix moyen au m<sup>2</sup> et le prix moyen au m<sup>2</sup> hédonique. Pour une meilleure lisibilité du graphique, les courbes correspondent à des moyennes mobiles d'ordre 3. Source: Base BIEN.

FIGURE 5.14 – PRIX RELATIF PAYÉ PAR LES ACHETEURS RÉSIDANT À L'ÉTRANGER. 1993-2008.



*Note:* La courbe "% des achats due aux acheteurs résidant à l'étranger " correspond au produit de la part des acheteurs résidant à l'étranger parmi l'ensemble des acheteurs et du prix relatif payé par les acheteurs résidant à l'étranger. Source: Base BIEN.

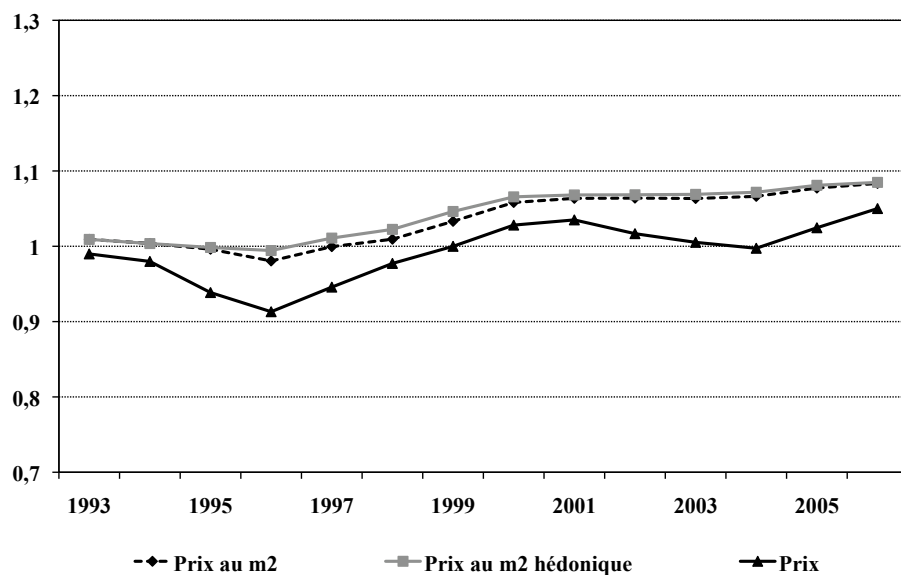
FIGURE 5.15 – PART DES ACHATS DUE AUX ACHETEURS DE NATIONALITÉ ÉTRANGÈRE OU RÉSIDANT À L'ÉTRANGER. 1993-2008.

TABLE 5.11 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LA VARIABLE EXPLICATIVE EST LE POURCENTAGE D'ACHETEURS RÉSIDANT À L'ÉTRANGER.

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>Part des acheteurs résidant à l'étranger</b>	1,904**	0,630**	0,663**	0,358**	0,370**
(Écart-type)	(0,068)	(0,055)	(0,052)	(0,070)	(0,046)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,826	0,956	0,965	0,977	0,993
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la part des acheteurs résidant à l'étranger de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,370% pour les acheteurs personne physique. Source: base BIEN.



*Note:* La courbe "prix au m<sup>2</sup>" correspond au prix au m<sup>2</sup> moyen payé par les acheteurs de nationalité étrangère divisé par le prix au m<sup>2</sup> moyen payé par l'ensemble des acheteurs. Trois définitions du prix sont utilisées: le prix moyen, le prix moyen au m<sup>2</sup> et le prix moyen au m<sup>2</sup> hédonique.

Pour une meilleure lisibilité du graphique, les courbes correspondent à des moyennes mobiles d'ordre 3.

Source: Base BIEN.

FIGURE 5.16 – PRIX RELATIF PAYÉ PAR LES ACHETEURS DE NATIONALITÉ ÉTRANGÈRE. 1993-2008.

TABLE 5.12 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LA VARIABLE EXPLICATIVE EST LE POURCENTAGE D'ACHETEURS DE NATIONALITÉ ÉTRANGÈRE.

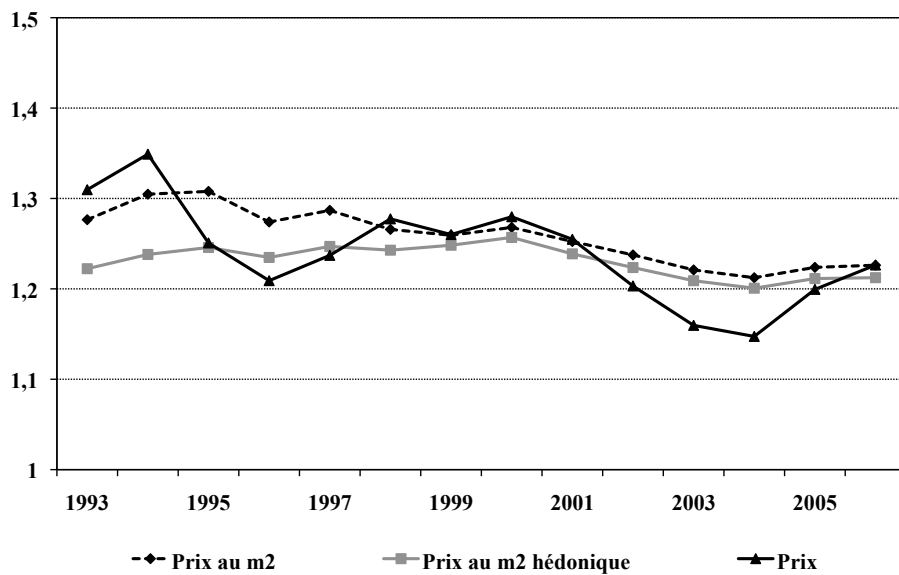
Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>Part des acheteurs de nationalité étrangère</b>	1,929**	0,486**	0,611**	0,132	0,300**
(Écart-type)	(0,124)	(0,069)	(0,069)	(0,073)	(0,054)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,721	0,951	0,959	0,976	0,992
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%.

Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans.

D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la part des acheteurs de nationalité étrangère parmi les acheteurs personnes physiques résidant à l'étranger de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,300% pour les acheteurs personne physique. Source: base BIEN.

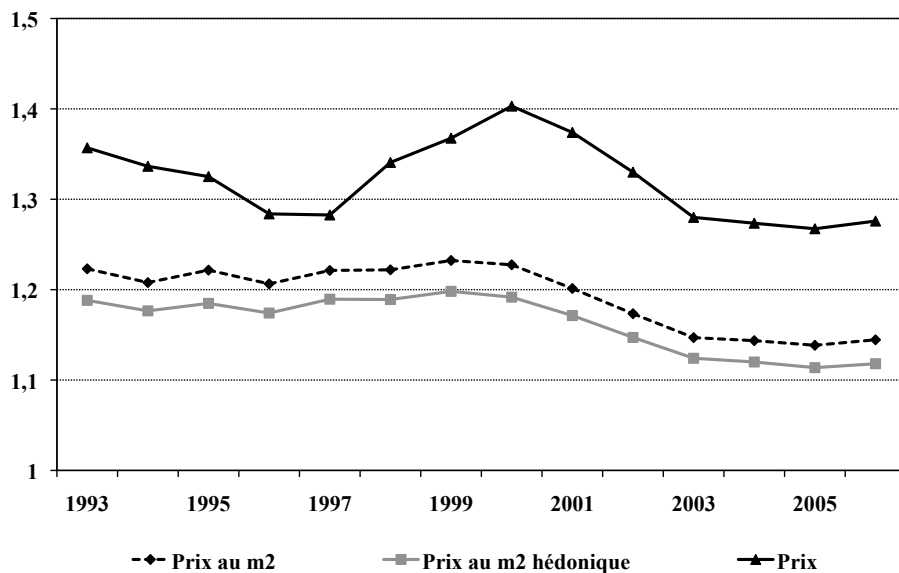


*Note:* La courbe "prix au m<sup>2</sup>" correspond au prix au m<sup>2</sup> moyen payé par les acheteurs étrangers résidant à l'étranger divisé par le prix au m<sup>2</sup> moyen payé par l'ensemble des acheteurs.

Pour une meilleure lisibilité du graphique, les courbes correspondent à des moyennes mobiles d'ordre 3.

Source: Base BIEN.

FIGURE 5.17 – PRIX RELATIF PAYÉ PAR LES ACHETEURS ÉTRANGERS RÉSIDANT À L'ÉTRANGER. 1993-2008.



*Note:* La courbe "prix au m<sup>2</sup>" correspond au prix au m<sup>2</sup> moyen payé par les acheteurs français résidant à l'étranger divisé par le prix au m<sup>2</sup> moyen payé par l'ensemble des acheteurs. Pour une meilleure lisibilité du graphique, les courbes correspondent à des moyennes mobiles d'ordre 3. Source: Base BIEN.

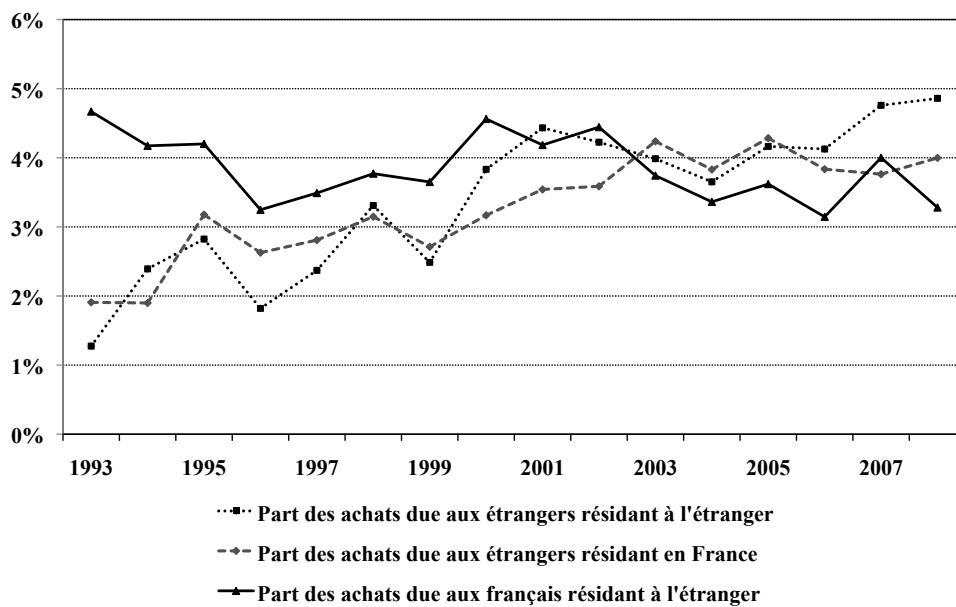
FIGURE 5.18 – PRIX RELATIF PAYÉ PAR LES ACHETEURS FRANÇAIS RÉSIDANT À L'ÉTRANGER. 1993-2008.



*Note:* La courbe "prix au m<sup>2</sup>" correspond au prix au m<sup>2</sup> moyen payé par les acheteurs étrangers résidant en France divisé par le prix au m<sup>2</sup> moyen payé par l'ensemble des acheteurs. Pour une meilleure lisibilité du graphique, les courbes correspondent à des moyennes mobiles d'ordre 3. Source: Base BIEN.

FIGURE 5.19 – PRIX RELATIF PAYÉ PAR LES ACHETEURS ÉTRANGERS RÉSIDANT EN FRANCE. 1993-2008.





*Note:* La courbe "% des achats due aux étrangers résidant à l'étranger " correspond au produit de la % des étrangers résidant à l'étranger parmi l'ensemble des acheteurs et du prix relatif payé par les étrangers résidant à l'étranger. Source: Base BIEN.

FIGURE 5.20 – PART DES ACHATS SELON LE LIEU DE RÉSIDENCE ET LA NATIONALITÉ DES ACHETEURS.

TABLE 5.13 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LES VARIABLES EXPLICATIVES SONT LES POURCENTAGES D'ACHETEURS ÉTRANGERS RÉSIDANT OU NON EN FRANCE ET LE POURCENTAGE D'ACHETEURS FRANÇAIS RÉSIDANT À L'ÉTRANGER.

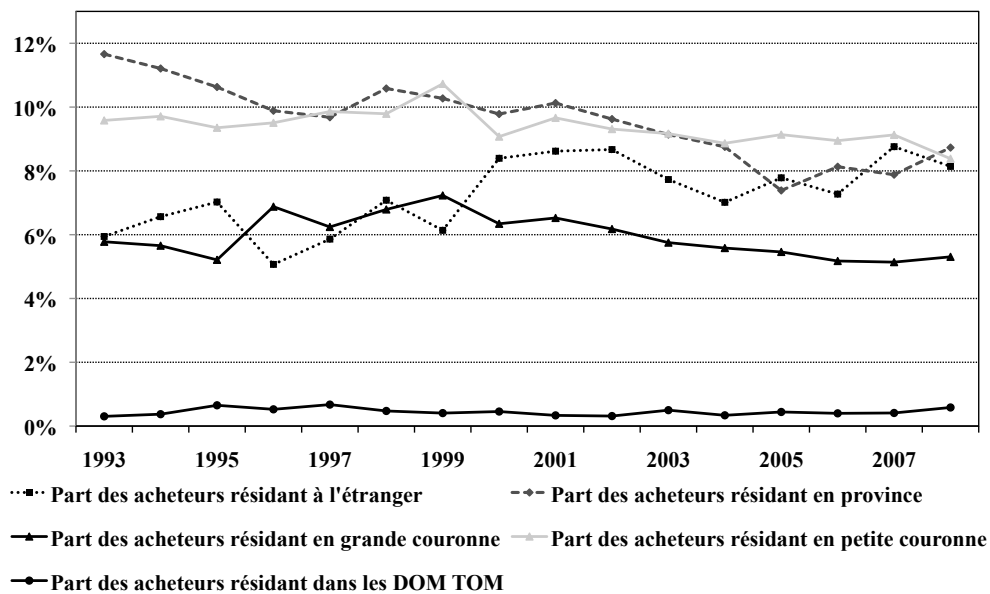
Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>Part des acheteurs de nationalité étrangère résidant à l'étranger</b>	1,038** (0,118)	0,595** (0,073)	0,663** (0,075)	0,325** (0,078)	0,398** (0,055)
<b>Part des acheteurs de nationalité étrangère résidant en France</b>	-2,537** (0,286)	-1,034** (0,189)	-0,878** (0,198)	-0,633** (0,166)	-0,156 (0,117)
<b>Part des acheteurs français résidant à l'étranger</b>	3,840** (0,235)	0,848** (0,162)	0,753** (0,165)	0,365** (0,139)	0,253** (0,088)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers administratifs	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,865	0,958	0,966	0,978	0,993
Nombre d'observations	640	640	640	640	640

Note: Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%.

Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans.

D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la part des acheteurs de nationalité étrangère résidant à l'étranger de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,398% pour les acheteurs personne physique. Source: base BIEN.



Note: La courbe "% des achats due aux acheteurs résidant à l'étranger " correspond au produit de la part des acheteurs résidant à l'étranger parmi l'ensemble des acheteurs et du prix relatif payé par ces mêmes acheteurs.  
 Source: Base BIEN.

FIGURE 5.21 – % DES ACHATS SELON LE DÉPARTEMENT DE RÉSIDENCE DES ACHETEURS.

TABLE 5.14 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LES VARIABLES EXPLICATIVES SONT LES POURCENTAGES D'ACHETEURS RÉSIDANT EN PETITE OU GRANDE COURONNE, EN PROVINCE, DANS LES DOM-TOM OU À L'ÉTRANGER.

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>Part acheteurs résidant à l'étranger</b>	1,470**	0,525**	0,610**	0,252**	0,378**
(Écart-type)	(0,077)	(0,059)	(0,056)	(0,073)	(0,050)
<b>Part acheteurs province</b>	1,813**	0,113	0,280*	-0,134	0,092
(Écart-type)	(0,171)	(0,114)	(0,118)	(0,098)	(0,066)
<b>Part acheteurs résidant en grande couronne</b>	-0,800**	-0,798**	-0,574**	-0,551**	-0,097
(Écart-type)	(0,287)	(0,173)	(0,177)	(0,140)	(0,092)
<b>Part acheteurs résidant en petite couronne</b>	-1,288**	-0,468**	-0,290*	-0,341**	-0,053
(Écart-type)	(0,207)	(0,129)	(0,134)	(0,103)	(0,069)
<b>Part acheteurs résidant dans DOM-TOMS</b>	0,223	-0,414	-0,147	-0,392	0,013
(Écart-type)	(0,966)	(0,563)	(0,558)	(0,445)	(0,280)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,867	0,959	0,966	0,978	0,993
Nombre d'observations	640	640	640	640	640

*Note:* D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la part des acheteurs résidant à l'étranger de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,378% pour les acheteurs personne physique.

Quelle sens donner à l'expression acheteur étranger? Les deux définitions les plus intuitives sont les acheteurs résidant à l'étranger et les acheteurs de nationalité étrangère.

Prenons la première définition et considérons le graphique 5.14. Il représente le prix relatif payé par les acheteurs résidant à l'étranger - c'est-à-dire le prix moyen payé par les acheteurs résidant à l'étranger divisé par le prix moyen payé par l'ensemble des acheteurs. Si ce prix relatif est égal à 1,2, alors les acheteurs résidant à l'étranger achètent en moyenne des logements 20% plus chers que l'ensemble des acheteurs. Selon le graphique 5.14, les acheteurs résidant à l'étranger achètent des logements 20% plus chers que l'ensemble des acheteurs, voire certaines années plus de 30% plus chers. Le graphique représente également le prix au m<sup>2</sup> relatif: en 1993, les acheteurs résidant à l'étranger payaient à surface égale, presque 25% de plus que la moyenne. En 2008, ils ne paient que 20% de plus que la moyenne. Le prix au m<sup>2</sup> équivalent<sup>10</sup> relatif est également représenté sur le graphique: les acheteurs étrangers, de 1993 à 2001, ont payé environ 20% de plus que la moyenne à surface et qualité égale. Depuis, 2001, ils paient environ 18% plus chers. Il y a deux raisons pour lesquelles les acheteurs résidant à l'étranger peuvent payer plus cher, à surface et qualité égale: ils achètent plus que la moyenne dans les arrondissements ou quartiers côtés où les prix au m<sup>2</sup> équivalent sont plus élevés<sup>11</sup> ou ils sont près, à qualité et surface égale, à payer plus que la moyenne.

Selon le graphique 5.14, le prix relatif payé par les acheteurs résidant à l'étranger est toujours supérieur au prix au m<sup>2</sup> relatif, lui-même étant toujours supérieur au prix au m<sup>2</sup> équivalent relatif: nous en déduisons que si les acheteurs résidant à l'étranger achètent des logements plus chers que la moyenne, c'est parce qu'ils achètent à la fois des logements plus grands, de meilleure qualité, et mieux situés.

Le graphique 5.15 représente l'évolution de la part des achats due aux acheteurs résidant à l'étranger. Cette part correspond tout simplement au produit du prix relatif payé par les acheteurs résidant à l'étranger par la proportion de ces acheteurs. Cette part a augmenté

---

10. Voir la section 2.2.1 pour avoir plus de détails sur les prix équivalents.

11. Rappelons que le calcul des prix au m<sup>2</sup> équivalents - ou hédoniques - est fait arrondissement par arrondissement.

entre 1993 et 2008, passant de 6% à 8%. En nous basant sur les prédictions théoriques de la section 2.1, nous nous attendons à ce que les acheteurs résidant à l'étranger aient eu un impact significatif et positif sur les prix de l'immobilier à Paris.

Cette prédiction est confirmée par la lecture du tableau 5.11. La colonne (1) de ce tableau représente les résultats de la régression du prix par le pourcentage d'acheteurs résidant à l'étranger avec uniquement des indicatrices temporelles. Le coefficient obtenu par cette régression est significatif à 5% et égal à 1,904. Il risque cependant d'être surestimé par rapport au "vrai" coefficient: les étrangers ont tendance à acheter naturellement dans les arrondissement les plus prisés où les prix sont les plus élevés. La colonne (2) du tableau représente les résultats de la régression lorsqu'on a ajouté des indicatrices des arrondissements. Le coefficient baisse et passe à 0,595, tout en restant significatif à 5%. Ce dernier coefficient peut être sous-estimé: si les étrangers sont - ne serait-ce qu'un peu - sensibles aux prix, alors le pourcentage d'acheteurs résidant à l'étranger va avoir tendance à baisser lorsque les prix d'un arrondissement augmentent. Cet effet est pris en compte lorsqu'on remplace les indicatrices des années et des arrondissements par des indicatrices croisées année par arrondissement. C'est le cas de la régression de la colonne (3). Dans ce cas, le coefficient est toujours significatif à 5% mais remonte légèrement à 0,663. Ce coefficient est peut-être encore biaisé vers le haut: les étrangers ont sans doute tendance à acheter dans les quartiers les plus prisés au sein de chaque arrondissement. La régression de la colonne (4) prend en compte cet effet. Les indicatrices croisées sont supprimées et remplacées par des indicatrices des quartiers administratifs. Le coefficient, toujours significatif à 5% tombe à 0,358. Là encore, cette estimation risque d'être légèrement biaisée vers la bas si on n'introduit pas d'indicatrices croisées arrondissement par année. Dans la régression de la colonne (5), on inclut à la fois des indicatrices des quartiers et des indicatrices croisées année par arrondissement. Le coefficient remonte alors légèrement par rapport à celui de la colonne (4), étant égal à 0,370 et toujours significatif à 5%.

Le résultat des régressions reste stable lorsqu'on ajoute des indicatrices. La façon dont varie le coefficient à mesure qu'on ajoute ou retire des indicatrices est en accord avec les prédictions de la théorie économique. Les résultats nous semblent donc crédibles. Pour résumer,

les acheteurs résidant à l'étranger achètent des appartements plus grands, de meilleure qualité et mieux situés que la moyenne. Selon nos régressions, une augmentation d'1 point de % de leur proportion entraîne une augmentation de 0,370% des prix au m<sup>2</sup> équivalents.

Intéressons nous maintenant à la deuxième définition des acheteurs étrangers: les acheteurs de nationalité étrangère. Selon les graphiques 5.15 et 5.16 et le tableau 5.12, jusqu'en 1999, les acheteurs de nationalité étrangère achetaient des logements plutôt moins chers que la moyenne. Depuis, ils achètent des logements légèrement plus chers - pas plus de 5% plus chers néanmoins. Jusqu'en 1997, ils ne payaient pas plus cher à surface et qualité égale. Depuis, ils achètent légèrement plus cher à surface et qualité égale. Ce phénomène a tendance à s'accroître: en 2008, ils ont acheté des logements environ 10% plus chers que la moyenne pour la même surface et la même qualité. Grâce à la disposition des courbes du graphique 5.16, nous savons que jusqu'en 1997, les acheteurs de nationalité étrangère achetaient des logements plus petits que la moyenne, mais d'aussi bonne qualité. Depuis, ils achètent des logements légèrement plus grands, d'aussi bonne qualité, et mieux situés. Selon le graphique 5.15, leur part dans les achats de l'ensemble des acheteurs a beaucoup augmenté entre 1993 et 2008, elle est passée de 3 à 9%. En outre, selon le tableau 5.12, une augmentation de 1 point de % des acheteurs de nationalité étrangère entraîne une augmentation de 0,300% des prix au m<sup>2</sup> équivalent.

Quel est le lien entre ces deux définitions des acheteurs étrangers? Nous décomposons les acheteurs en quatre catégories: les français qui résident en France, les français qui résident à l'étranger, les étrangers qui résident en France et les étrangers qui résident à l'étranger. Nous allons considérer que les acheteurs des trois dernières catégories sont tous des "étrangers". Nous cherchons à savoir si ces trois catégories ont joué le même rôle sur les prix de l'immobilier à Paris. Selon le graphique 5.17, les acheteurs étrangers qui résident à l'étranger achètent des logements au moins 20% plus chers que la moyenne pour la même surface et la même qualité. En début de période, les français qui résident à l'étranger achetaient des logements 20% plus chers que la moyenne à surface et qualité égale. Ils achètent maintenant des logements 10% plus chers. Selon le graphique 5.18, ils achètent des logements plus grands, de meilleurs

qualité et mieux situé que la moyenne. È contrario, selon le graphique 5.19, les acheteurs étrangers qui habitent en France achètent des logements plus petits, de moins bonne qualité et moins bien situés que la moyenne. Ce phénomène a néanmoins diminué entre 1993 et 2008. È priori, ce sont donc les acheteurs qui résident à l'étranger, et particulièrement les acheteurs de nationalité étrangère vivant à l'étranger, qui doivent avoir un impact sur les prix.

D'après le graphique 5.20, la part des achats due aux acheteurs étrangers résidant à l'étranger est passée de 1 à 5% entre 1993 et 2008, celle des français qui résident à l'étranger est passée de 4,6% à 3,2% et celle des étrangers qui résident en France est passée de 2 à 4%. Selon le tableau 5.13, seuls les acheteurs qui résident à l'étranger ont un impact sur les prix: une augmentation de 1 point de % des étrangers résidant à l'étranger entraîne une augmentation de 0,398% des prix, tandis qu'une augmentation de 1 point de % des français résidant à l'étranger entraîne une augmentation de 0,253%. D'après la colonne (5) du tableau 5.20, les acheteurs étrangers qui résident en France n'ont pas eu d'impact significatif sur les prix<sup>12</sup>.

Nous savons maintenant que seuls les étrangers, au sens de "qui résident hors de France", ont eu un impact significatif sur les prix. Ce sont les résidents étrangers de nationalité étrangère qui ont eu le plus fort impact sur les prix.

### 3.1.1 Impact des acheteurs sur les prix selon leur département de résidence

Si les personnes qui résident à l'étranger ont fait monter les prix, on peut légitimement se demander si les provinciaux qui s'achètent des pied-à-terre à Paris n'ont pas eu aussi un impact sur les prix.

Nous décomposons les acheteurs en six sous-groupes: les acheteurs qui habitent à Paris, ceux qui habitent en petite couronne, en grande couronne, en province, dans les DOM-TOM et enfin ceux qui habitent à l'étranger. Les acheteurs qui appartiennent aux cinq dernières catégories peuvent tous être considérés comme des étrangers, dans le sens où ils n'habitent pas

---

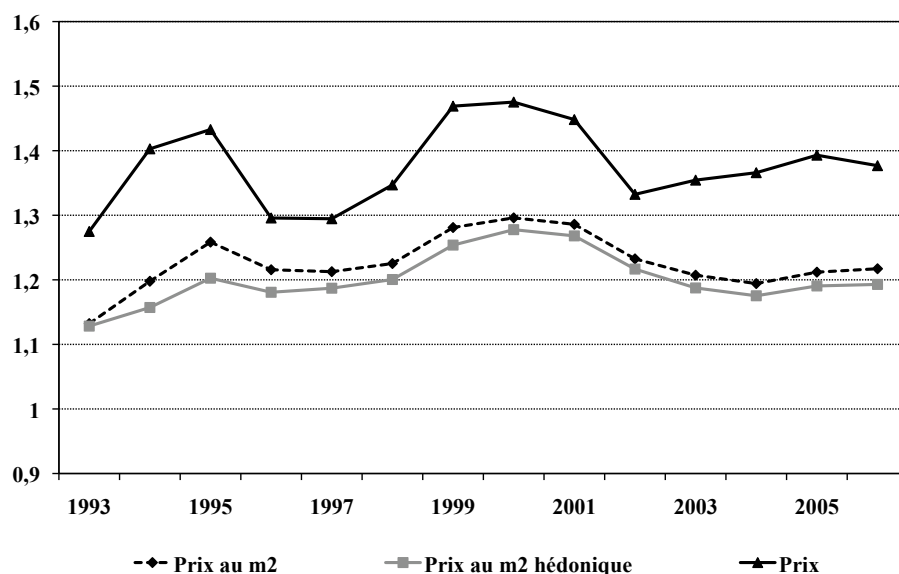
12. Le coefficient pour les acheteurs de nationalité étrangère résidant en France est négatif et significatif dans les régressions des colonnes (1) à (4) parce qu'en moyenne ces acheteurs achètent, à surface et qualité égale, plutôt moins cher que la moyenne. Ce point est démontré par le graphique 5.19. Le coefficient redevient non significatif lorsque toutes les indicatrices sont ajoutées et que tous les biais sont éliminés.



à Paris. Ces cinq catégories d'acheteurs ont-elles joué le même rôle sur les prix de l'immobilier à Paris?

On a déjà vu précédemment que la part des achats due aux acheteurs résidant à l'étranger a augmenté, ce que le graphique 5.21 nous rappelle. C'est la seule de nos cinq catégories. La part dans les achats de chacune de ces quatre autres catégories est soit restée stable, soit a baissé. On s'attend donc à ce que seuls les étrangers aient eu un impact positif sur les prix. Cette prédiction est confirmée par nos régressions. Selon le tableau 5.14, seuls les acheteurs résidant à l'étranger ont eu impact significatif sur les prix.

### 3.1.2 Impact des acheteurs sur les prix selon leur continent d'origine



*Note:* La courbe "prix au m<sup>2</sup>" correspond au prix au m<sup>2</sup> moyen payé par les acheteurs dont le pays d'origine est situé en Amérique divisé par le prix au m<sup>2</sup> moyen payé par l'ensemble des acheteurs.

Pour une meilleure lisibilité du graphique, les courbes correspondent à des moyennes mobiles d'ordre 3.

Source: Base BIEN.

FIGURE 5.22 – PRIX RELATIF PAYÉ PAR LES ACHETEURS EN PROVENANCE D'AMÉRIQUE. 1993-2008.

TABLE 5.15 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LES VARIABLES EXPLICATIVES SONT LES POURCENTAGES D'ACHETEURS EN PROVENANCE D'AMÉRIQUE, D'ASIE, D'EUROPE (HORS FRANCE), DU MOYEN-ORIENT ET D'AFRIQUE.

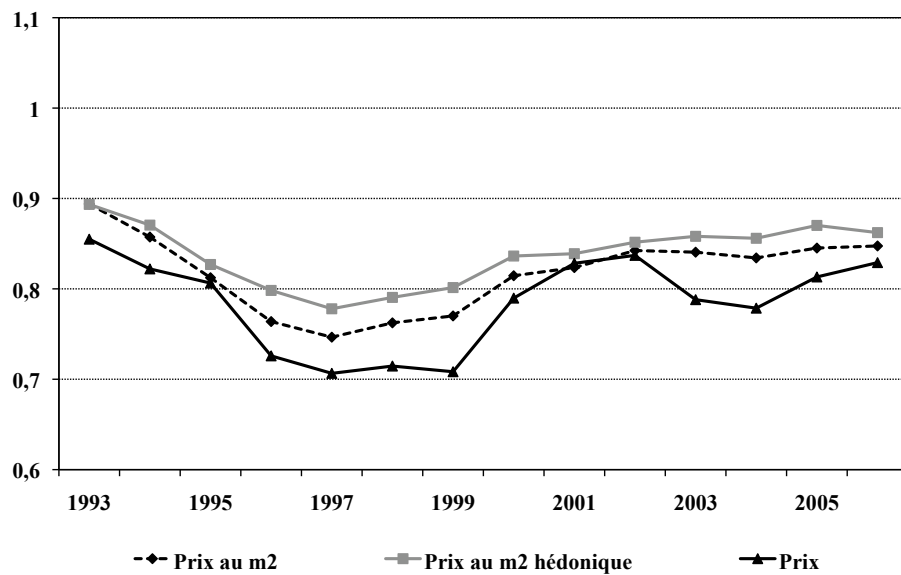
Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>Part des acheteurs d'Amérique</b>	4,460**	1,182**	1,103**	0,743**	0,506**
(Écart-type)	(0,404)	(0,235)	(0,258)	(0,205)	(0,166)
<b>Part des acheteurs d'Asie</b>	-4,967**	-1,467**	-1,486**	-0,970**	-0,206
(Écart-type)	(0,625)	(0,361)	(0,364)	(0,353)	(0,250)
<b>Part des acheteurs d'Europe</b>	1,357**	0,428**	0,499**	0,248**	0,343**
(Écart-type)	(0,159)	(0,089)	(0,091)	(0,082)	(0,059)
<b>Part des acheteurs du Moyen-Orient</b>	3,518**	2,126**	2,255**	-0,574	0,014
(Écart-type)	(0,720)	(0,393)	(0,391)	(0,421)	(0,288)
<b>Part des acheteurs d'Afrique</b>	-4,454**	-1,812**	-0,742	-1,757**	-0,140
(Écart-type)	(0,748)	(0,409)	(0,426)	(0,326)	(0,228)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,820	0,957	0,964	0,978	0,992
Nombre d'observations	640	640	640	640	640

Note: Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%.

Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans.

D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la part des acheteurs en provenance d'Amérique de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,506% pour les acheteurs personne physique. Source: base BIEN.

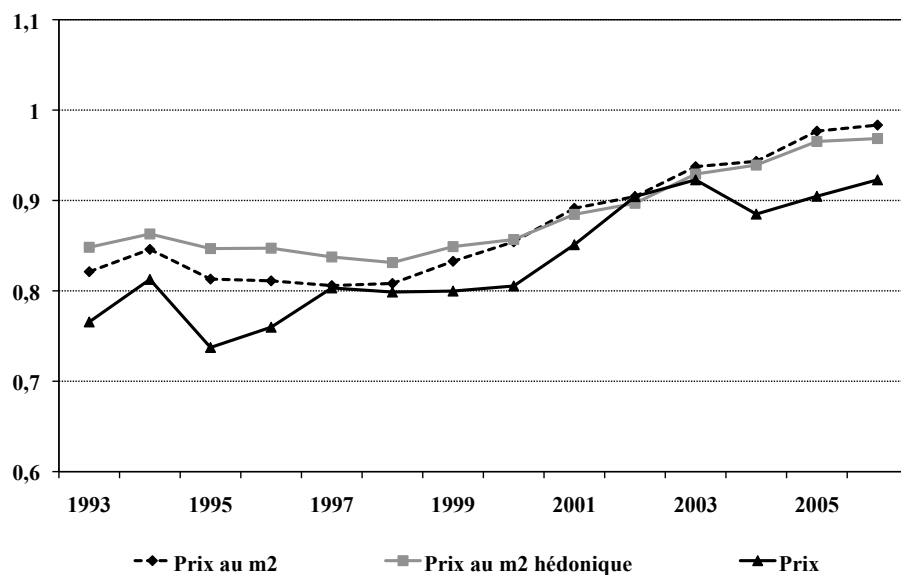


*Note:* La courbe "prix au m<sup>2</sup>" correspond au prix au m<sup>2</sup> moyen payé par les acheteurs dont le pays d'origine est situé en Amérique divisé par le prix au m<sup>2</sup> moyen payé par l'ensemble des acheteurs.

Pour une meilleure lisibilité du graphique, les courbes correspondent à des moyennes mobiles d'ordre 3.

Source: Base BIEN.

FIGURE 5.23 – PRIX RELATIF PAYÉ PAR LES ACHETEURS EN PROVENANCE D'ASIE.  
1993-2008.

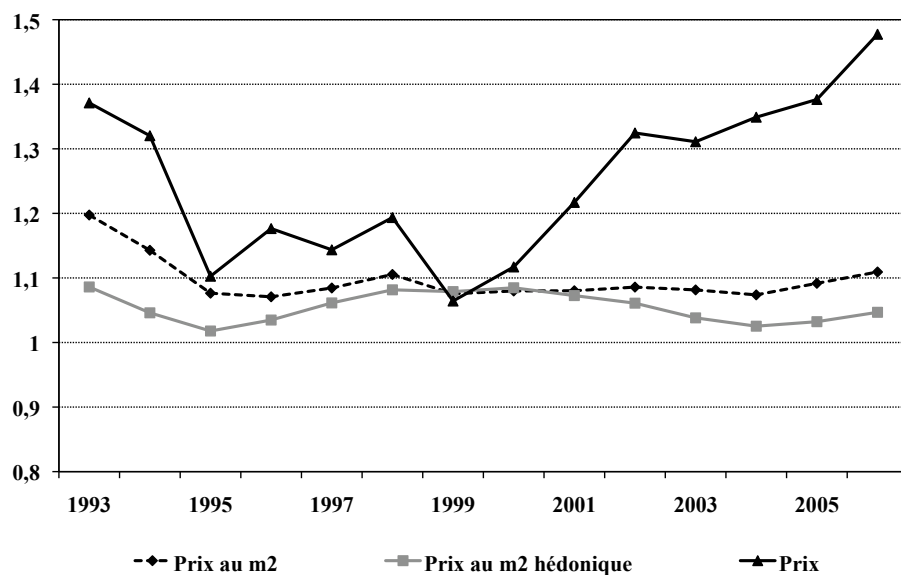


*Note:* La courbe "prix au m<sup>2</sup>" correspond au prix au m<sup>2</sup> moyen payé par les acheteurs dont le pays d'origine est situé en Afrique divisé par le prix au m<sup>2</sup> moyen payé par l'ensemble des acheteurs.

Pour une meilleure lisibilité du graphique, les courbes correspondent à des moyennes mobiles d'ordre 3.

Source: Base BIEN.

FIGURE 5.24 – PRIX RELATIF PAYÉ PAR LES ACHETEURS EN PROVENANCE D'AFRIQUE. 1993-2008.

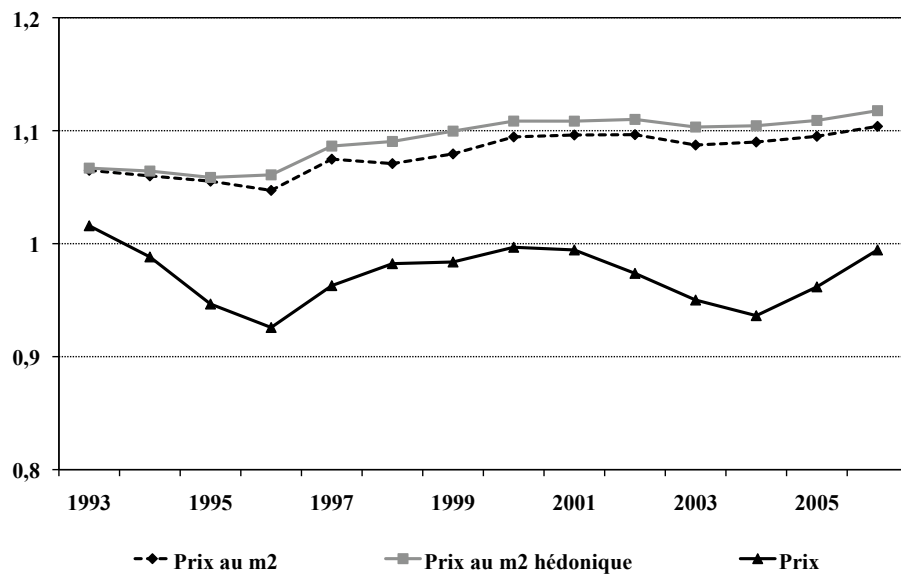


*Note:* La courbe "prix au m<sup>2</sup>" correspond au prix au m<sup>2</sup> moyen payé par les acheteurs dont le pays d'origine est situé du Moyen-Orient divisé par le prix au m<sup>2</sup> moyen payé par l'ensemble des acheteurs.

Pour une meilleure lisibilité du graphique, les courbes correspondent à des moyennes mobiles d'ordre 3.

Source: Base BIEN.

FIGURE 5.25 – PRIX RELATIF PAYÉ PAR LES ACHETEURS EN PROVENANCE DU MOYEN-ORIENT. 1993-2008.

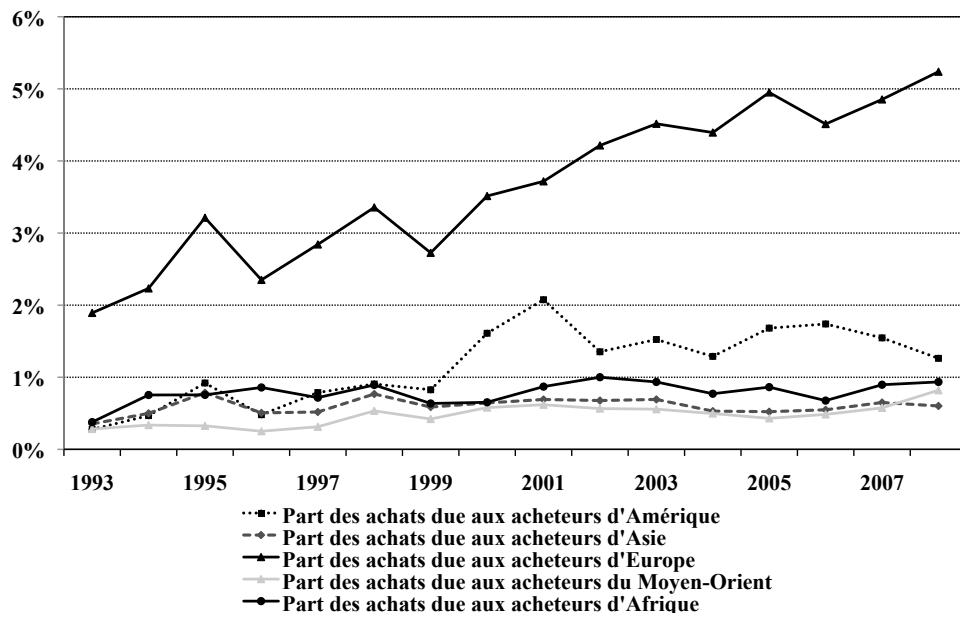


*Note:* La courbe "prix au m<sup>2</sup>" correspond au prix au m<sup>2</sup> moyen payé par les acheteurs dont le pays d'origine est situé en Amérique divisé par le prix au m<sup>2</sup> moyen payé par l'ensemble des acheteurs.

Pour une meilleure lisibilité du graphique, les courbes correspondent à des moyennes mobiles d'ordre 3.

Source: Base BIEN.

FIGURE 5.26 – PRIX RELATIF PAYÉ PAR LES ACHETEURS EN PROVENANCE D'EUROPE.  
1993-2008.



*Note:* La courbe "Part des achats due aux acheteurs d'Amérique " correspond au produit de la part des acheteurs dont le pays d'origine est situé en Amérique et du prix relatif payé par ces mêmes acheteurs. Source: Base BIEN.

FIGURE 5.27 – PART DES ACHATS SELON LE CONTINENT D'ORIGINE DES ACHETEURS. 1993-2008.

TABLE 5.16 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LA VARIABLE EXPLICATIVE EST LE POURCENTAGE D'ACHETEURS DE NATIONALITÉ ÉTRANGÈRE OU RÉSIDANT À L'ÉTRANGER. **On se restreint aux 30% des transactions les plus chères.**

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>Part des acheteurs de nationalité étrangère résidant à l'étranger</b>	0,702** (0,047)	0,361** (0,028)	0,368** (0,027)	0,251** (0,032)	0,232** (0,027)
<b>Part des acheteurs de nationalité étrangère résidant en France</b>	0,171 (0,136)	-0,027 (0,072)	-0,002 (0,072)	-0,063 (0,065)	-0,051 (0,054)
<b>Part des acheteurs français résidant à l'étranger</b>	1,105** (0,089)	0,194** (0,051)	0,197** (0,052)	0,086 (0,048)	0,069 (0,040)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,928	0,981	0,985	0,987	0,993
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. Pour définir les 30% de transactions les plus chères, la définition de prix utilisée est le prix au m<sup>2</sup> hédonique. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la part des acheteurs de nationalité étrangère résidant à l'étranger parmi les acheteurs des 30% des transactions les plus chères de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,232% pour ces mêmes transactions. Source: base BIEN.



Nous nous sommes également demandés si les étrangers ont le même impact sur les prix de l'immobilier à Paris quelque soit leur pays d'origine. Nous ne pouvons pas décomposer les acheteurs d'origine étrangère pays par pays, ils ne sont pas suffisamment nombreux. Nous avons donc choisi de décomposer les acheteurs en six groupes selon leur pays d'origine: les acheteurs de nationalité française, les acheteurs dont le pays d'origine est situé en Amérique, en Asie, en Afrique, au Moyen-Orient et en Europe (hors France bien sûr)<sup>13</sup>.

Selon les graphiques 5.22, 5.23, 5.24, 5.25 et 5.26, les acheteurs en provenance d'Amérique achètent des logements beaucoup plus chers que la moyenne - plus de 30% plus chers. Ils achètent des logements plus grands, de meilleure qualité et mieux situés que la moyenne. Les acheteurs en provenance d'Asie achètent, eux, des logements globalement moins chers que la moyenne - toujours au moins 10% moins chers et jusqu'à 30% moins chers. Ils achètent des logements plus petits, de moins bonne qualité et moins bien situés que la moyenne. Les acheteurs en provenance d'Afrique achètent eux aussi des logements plus petits, de moins bonne qualité et moins bien situés que la moyenne. Cependant, ce phénomène s'est atténué entre 1993 et 2008. En 1993, ces acheteurs achetaient des logements 20% moins chers que la moyenne, alors qu'en 2008, ils ont acheté des logements seulement 10% moins chers que la moyenne. Les acheteurs en provenance du Moyen-Orient achètent des logements beaucoup plus grand que la moyenne<sup>14</sup>, légèrement de meilleure qualité et mieux situés que la moyenne. Les acheteurs en provenance d'Europe - hors France - achètent des logements plus petits que la moyenne mais de meilleure qualité et mieux situés.

Selon les prédictions de la section 2.1, nous nous attendons à ce qu'une augmentation de la proportion d'acheteurs en provenance d'Amérique, d'Europe et du Moyen-Orient ait un impact positif sur les prix. Par contre, nous ne pouvons pas prédire l'impact que peut avoir une augmentation de la proportion d'acheteurs en provenance d'Asie et d'Afrique.

Selon le graphique 5.27, la part des acheteurs d'Europe dans les achats a beaucoup augmenté

---

13. Pour plus de précisions sur la façon dont les groupes de pays ont été constitués, voir l'Annexe B

14. Le prix relatif est très variable, sans doute à cause du petit nombre de ce type d'acheteurs.

sur la période considérée, passant de moins de 2 à plus de 5%. La part des acheteurs d'Amérique dans les achats a elle aussi augmenté, passant de moins de 0,5 à plus de 1% avec un pic à 2%. Les parts des acheteurs d'Afrique, d'Asie, et du Moyen-Orient sont restées plus stables ou ont très légèrement augmenté. Nous nous attendons donc à ce que seuls les acheteurs en provenance d'Amérique et d'Europe aient un impact significatif et positif sur les prix. Cette prédiction est validée par les résultats du tableau 5.15: seuls les proportions d'acheteurs en provenance d'Europe et d'Amérique ont eu un impact sur les prix. Une augmentation de 1 point de % des acheteurs en provenance d'Europe (respectivement d'Amérique) fait varier les prix au m<sup>2</sup> équivalents de 0,343% (resp. 0,506). Si les acheteurs en provenance d'Amérique ont un impact plus fort, c'est qu'ils achètent des logements très chers. Les acheteurs en provenance du Moyen-Orient, ont un impact positif mais non significatif. C'est le signe d'un problème d'identification: la proportion de ces acheteurs est trop faible et a trop peu augmenté pour que leur impact sur les prix soit détecté.

Seuls les acheteurs européens ou américains ont eu un impact significatif et positif sur les prix.

### 3.1.3 Impact sur les prix de l'immobilier haut de gamme des acheteurs étrangers

Les acheteurs résidant à l'étranger achètent des logements en moyenne plus chers que l'ensemble des acheteurs. Nous avons vu dans la section 1.3 que l'augmentation de la proportion d'acheteurs étrangers a été d'autant plus forte dans les gammes de prix élevés. Les étrangers ont-ils eu un impact plus fort sur les prix de l'immobilier haut de gamme?

Nous avons refait nos régressions en nous restreignant aux 30% de transactions les plus chères. Les résultats obtenus sont présentés dans le tableau 5.16. Seuls les acheteurs étrangers résidant à l'étranger ont eu un impact significatif sur les prix de l'immobilier haut de gamme. Cet impact est plus faible que lorsqu'on considère toutes les transactions: le coefficient est égal à 0,232 contre 0,370. Cela ne nous surprend qu'à moitié. Dans la section 1.3, nous avons vu que, plus les gammes de prix étaient élevées, plus l'augmentation du pourcentage d'acheteurs étrangers avait été forte, et moins l'augmentation des prix de l'immobilier avait, elle, été

importante.

Selon nos estimations, la proportion d'acheteurs étrangers a bien un impact significatif et positif sur les prix de l'immobilier. Une augmentation d'un point de % de la proportion d'acheteurs étrangers entraîne une augmentation d'environ 0,370% des prix de l'immobilier. Parmi les étrangers, nous distinguons les personnes qui résident à l'étranger, qu'ils soient français ou de nationalité étrangère, et les personnes qui vivent en France mais qui ne sont pas de nationalité française. Selon nos estimations, ce sont les personnes qui résident à l'étranger qui ont eu un impact significatif sur les prix de l'immobilier, les étrangers vivant en France, ont eu, eux, un impact non significatif sur les prix. Les personnes qui vivent en province ou dans la petite ou la grande couronne et qui ont acheté un pied-à-terre à Paris, ont eu, eux aussi, un impact non significatif sur les prix. Ce sont en particulier les acheteurs d'origine européenne ou américaine qui ont eu un impact significatif et positif sur les prix. Enfin, les étrangers ont bien eu un impact positif et significatif sur les prix de l'immobilier haut de gamme, mais cet impact a été plus faible que leur impact global sur les prix. Une augmentation d'un point de % de la proportion d'acheteurs étrangers dans le secteur de l'immobilier haut de gamme entraîne une augmentation de 0,232% des prix de ce secteur.



## 4 Critiques et pistes d'améliorations

Dans cette section, nous discutons la robustesse de nos résultats puis nous estimons l'impact global qu'ont eu les acheteurs étrangers sur les prix de l'immobilier à Paris.

### 4.1 Robustesse des résultats

Nous avons d'abord testé la robustesse de nos estimations en prenant une définition alternative de nos variables explicatives: on pourrait en effet objecter qu'il ne suffit pas de considérer le pourcentage d'acheteurs étrangers mais qu'il faut aussi prendre en compte le pourcentage de vendeurs étrangers. Nous avons donc refait toutes nos régressions en prenant comme variable explicative non plus le pourcentage d'acheteurs étrangers mais la différence entre le pourcentage d'acheteurs étrangers et le pourcentage de vendeurs étrangers. Les résultats que nous obtenons sont très similaires à ceux que nous avons présenté dans la section précédente. Ils sont tous disponibles dans l'annexe [D.1](#).

Nous avons ensuite refait toutes nos régressions en n'excluant plus les transactions où l'acheteur est soit une SCI soit de type inconnu. Les résultats sont disponibles dans l'annexe [D.2](#). Les résultats sont là aussi très similaires à ceux de la section précédente.

### 4.2 Impact global des étrangers sur les prix de l'immobilier

Nous cherchons à calculer une borne maximale de l'impact que les étrangers ont eu sur les prix de l'immobilier à Paris. L'ensemble des acheteurs est divisé en cinq catégories: les acheteurs qui résident en France et qui sont français, les acheteurs de type inconnu, les acheteurs qui résident à l'étranger et qui sont français, les acheteurs qui résident à l'étranger et qui ne sont pas français et les acheteurs qui résident en France et qui ne sont pas français. Les trois dernières catégories correspondent à des acheteurs "étrangers". Les acheteur de type inconnu sont les acheteurs dont soit la nationalité, soit le lieu de résidence, soit la qualité

TABLE 5.17 – IMPACT DES ÉTRANGERS SUR LES PRIX. 1993-2008.

Part des achats	Type d'acheteurs			
	Année	Étrangers à l'étranger	Étrangers en France	Français à l'étranger
1993	1,0%	1,9%	2,8%	18,8%
1994	1,5%	2,2%	2,8%	13,6%
1995	1,8%	3,4%	2,6%	11,5%
1996	1,5%	3,2%	2,3%	9,4%
1997	1,9%	3,3%	2,5%	9,8%
1998	2,3%	3,8%	2,6%	9,2%
1999	1,8%	3,1%	2,6%	9,4%
2000	2,8%	3,7%	2,9%	8,6%
2001	3,3%	3,9%	2,8%	7,3%
2002	3,0%	3,9%	3,0%	8,7%
2003	2,9%	4,2%	2,4%	13,3%
2004	2,8%	3,8%	2,3%	12,5%
2006	3,2%	3,9%	2,1%	11,6%
2008	3,3%	3,9%	2,3%	11,5%
Variation (en points de %) entre 1993 et 2008:				
	3,1%	1,9%	-0,7%	-10,5%
Ce qui entraîne au grand maximum une augmentation des prix de (en %):				
	1,1%	0,7%	0,0%	0,0%

L'augmentation des prix a été de 150 en euros courants et de 100% en euros constants. Les étrangers sont responsables au plus d'une augmentation des prix de moins de 2%, soit 1/50 de cette hausse.

*Note:* Les acheteur de type inconnu sont les acheteurs dont soit la nationalité, soit le lieu de résidence, soit la qualité (particulier, SCI, marchand de bien) est inconnu. Tous les acheteurs de type inconnu sont considérés comme étrangers. Source: Base BIEN.

TABLE 5.18 – IMPACT DES ÉTRANGERS SUR LES PRIX DES 30% DE TRANSACTIONS LES PLUS CHÈRES. 1993-2008.

Année	Type d'acheteurs			
	Étrangers à l'étranger	Étrangers en France	Français à l'étranger	Type inconnu
1993	1,6%	1,5%	4,4%	20,1%
1994	2,7%	1,4%	4,8%	15,5%
1995	3,3%	2,0%	4,4%	13,4%
1996	2,5%	1,7%	3,6%	10,7%
1997	3,7%	1,8%	4,2%	11,7%
1998	4,2%	2,0%	4,3%	10,7%
1999	3,3%	1,9%	4,5%	11,2%
2000	5,5%	2,6%	5,0%	9,3%
2001	6,1%	3,0%	4,9%	8,7%
2002	5,8%	3,2%	5,2%	9,1%
2003	5,4%	3,6%	3,8%	14,0%
2004	5,5%	3,5%	3,7%	12,8%
2006	6,6%	3,9%	3,5%	12,9%
2008	7,0%	3,0%	3,8%	11,1%
Variation (en points de %) entre 1993 et 2008:				
	5,7%	1,8%	-0,7%	-13,2%
Ce qui entraîne au grand maximum une augmentation des prix de (en %):				
	1,3%	0,4%	0,0%	0,0%

L'augmentation des prix a été de 145 en euros courants et de 90% en euros constants. Les étrangers sont responsables au plus d'une augmentation des prix de moins de 2%, soit moins d'1/50 de cette hausse.

*Note:* Les acheteur de type inconnu sont les acheteurs dont soit la nationalité, soit le lieu de résidence, soit la qualité (particulier, SCI, marchand de bien) est inconnu. Tous les acheteurs de type inconnu sont considérés comme étrangers. Source: Base BIEN.

(particulier, SCI, marchand de bien..) est inconnu <sup>15</sup>

Nous nous plaçons dans un scénario extrême, peu réaliste. Nous considérons qu'une augmentation de la proportion d'étrangers (quelque soit la définition que l'on donnera au terme "étranger") entraîne une augmentation de 0,370% des prix. Le coefficient 0,370 est une borne maximale des coefficients obtenus dans nos régressions lorsque toutes les indicatrices sont incluses. Nous considérons que les trois catégories d'étrangers peuvent avoir un impact sur les prix. Nous ne tenons donc pas compte de nos résultats antérieurs: nous avons vu que les acheteurs étrangers résidant en France ont eu un impact non significatif sur les prix. Nous allons même jusqu'à considérer que tous les acheteurs de type inconnu sont des étrangers. Nous avons donc quatre catégories d'étrangers. Lorsque que la proportion d'une de ces catégories a augmenté entre 1993 et 2008, nous calculons l'impact que cela a eu sur les prix. Lorsque la proportion d'une de ces catégorie a baissé, nous considérons que l'impact de cette dernière sur les prix a été nul. L'impact global est la somme de tous ces impacts. Ces calculs sont résumés dans le tableau 5.17.

Les étrangers sont donc responsables d'au plus 1/50 de la hausse des prix qui a eu lieu à Paris entre 1993 et 2008. Leur rôle a donc été négligeable.

Nous appliquons le même raisonnement à l'immobilier haut de gamme, c'est-à-dire aux 30% de transactions les plus chères. Nos calculs sont résumés dans le tableau 5.18. Le coefficient que nous prenons n'est plus égal à 0,370 mais à 0,232. Selon le tableau 5.18, les étrangers sont responsables tout au plus d'1/50 de la hausse des prix de l'immobilier haut de gamme. Là encore, leur rôle a été négligeable.

Selon nos estimations, une augmentation de la proportion d'acheteurs étrangers d'un point de 1% se traduit par une augmentation des prix de 0,37%. Si l'on suppose que l'offre de logements à Paris est fixe, on en déduit que l'élasticité-prix au logement est environ égale à 2,7. Cette valeur est élevée comparée aux estimations de [Saiz \(2006\)](#) et [Ottaviano and Peri \(2007\)](#). [Saiz \(2006\)](#) a estimé l'impact de l'immigration sur les prix des logements et sur les loyers dans les grandes villes des Etats-Unis : un flot d'immigration équivalent à 1% de la

---

15. Pour faire ce calcul, nous avons donc gardé toutes les transactions et pas seulement celles dont l'acheteur est un particulier.



population initiale entraîne une augmentation des loyers de 1% et également une augmentation des prix des logements de 3%. [Ottaviano and Peri \(2007\)](#) étudient l'effet de l'immigration sur le marché du travail et le marché immobilier en utilisant des données américaine. Selon leurs estimations, l'élasticité des loyers au flux d'immigration est proche de 0,7 et l'élasticité des prix des logements à l'immigration est comprise entre 1 et 2.

Selon ces deux articles, l'élasticité-prix au logement dans les villes américaines est comprise entre 0,3 et 1. Il semble donc que l'élasticité-prix au logement soit très élevée à Paris intramuros, ce qui peut peut-être s'expliquer par l'existence d'un bien de substitution très proche et relativement moins cher : les logements en petite couronne. Cette forte élasticité-prix explique en partie le faible impact qu'ont eu les acheteurs étrangers sur les prix à Paris. Si l'élasticité-prix au logement à Paris était similaire à celles des villes américaines (0,3 à 1), l'augmentation de la proportion d'acheteurs étrangers aurait entraîné une augmentation des prix de 5 à 15%, soit 5% à 15% de la hausse qui a eu lieu. Pour que l'augmentation de la proportion d'acheteurs étrangers ait été responsable de la moitié de la hausse qui a eu lieu (soit une augmentation de 50%), il aurait fallu que l'élasticité-prix au logement soit égale à 0,1 à Paris. Il aurait donc fallu que l'élasticité-prix au logement soit 27 fois inférieure à sa valeur estimée.

## 5 Conclusion

Les étrangers ne sont pas responsables de la forte hausse des prix de l'immobilier à Paris. Nos régressions prouvent qu'une augmentation de la proportion d'acheteurs étrangers a bien un impact positif et significatif sur les prix. Plus précisément, une augmentation d'un point de % de la proportion d'acheteurs étrangers entraîne une variation des prix de 0,370%. Ce sont surtout les acheteurs qui résident à l'étranger qui ont un impact positif et significatif sur les prix plutôt que les acheteurs de nationalité étrangère qui vivent en France. En effet, les acheteurs qui résident à l'étranger ont tendance à acheter des logements au moins 20% plus chers que l'ensemble des acheteurs alors que les acheteurs de nationalité étrangère qui vivent en France, ont eux plutôt tendance à acheter des logements moins chers que l'ensemble des acheteurs. Les provinciaux ou habitants de la petite ou grande couronne n'ont eu aucun impact sur les prix de l'immobilier à Paris sur la période considérée. La raison en est simple: leur proportion est restée très stable. Seuls les acheteurs résidant à l'étranger ont eu un impact sur les prix car leur proportion a augmenté entre 1993 et 2008.

L'impact des étrangers sur les prix de l'immobilier haut de gamme est plus faible que leur impact sur l'ensemble des transactions. Une augmentation d'un point de % de la proportion d'acheteurs étrangers dans le secteur de l'immobilier haut de gamme entraîne une variation des prix de ce secteur de 0,232%.

La proportion d'acheteurs étrangers a bien eu une influence sur les prix de l'immobilier à Paris. Néanmoins, l'impact des acheteurs étrangers sur les prix a été négligeable comparé à la hausse qui a eu lieu car les étrangers ne sont tout simplement pas assez nombreux. Ils ont beau acheter des logements jusqu'à 30% plus chers que le reste des acheteurs, ils sont en nombre insuffisant pour entraîner des augmentations substantielles des prix. Ils ne sont responsables que d'1/50 de la hausse qui a eu lieu. Autant dire que leur impact global a été négligeable..





# Appendix

## A Prix moyen des appartements vendus par catégorie de prix

TABLE 5.19 – PRIX MOYEN DES APPARTEMENTS VENDUS PAR DÉCILE DE PRIX. 1993-2008.

Année	Prix moyen des appartements vendus selon la catégorie de prix (en euros courants)						
	Ensemble	D1	D1-D3	D4-D7	D8-D10	D10	C100
1993	150 129	37 239	54 567	111 987	300 303	477 182	1029 896
1994	157 645	38 519	54 922	115 083	322 907	509 817	1010 748
1995	148 497	35 487	50 586	107 869	301 564	476 837	946 742
1996	142 064	34 763	49 747	105 962	286 033	447 441	860 625
1997	139 786	32 325	47 980	104 061	280 615	439 130	857 042
1998	140 921	31 491	46 611	101 839	287 523	457 233	909 075
1999	148 075	33 418	48 923	106 538	303 292	484 674	981 542
2000	169 203	36 034	52 874	117 926	356 964	587 045	1226 557
2001	178 832	38 369	56 936	125 888	371 993	610 012	1335 978
2002	195 527	42 556	64 065	141 564	399 039	642 024	1410 443
2003	223 955	51 495	76 548	165 320	450 583	713 041	1450 896
2004	255 754	62 131	91 618	193 188	505 667	785 736	1517 638
2005	296 779	81 928	114 687	228 432	572 614	871 318	1657 510
2006	322 165	82 010	118 147	245 262	630 405	979 513	1940 711
2007	354 073	103 168	141 183	271 481	679 111	1052 795	2000 609
2008	378 306	106 948	147 477	284 775	738 466	1170 650	2461 276

*Note:* D1 correspond aux 10% de transactions les moins chères. D1-D3 correspond aux 30% de transactions les moins chères. C100 correspond aux 1% de transactions les plus chères. La définition de prix utilisée pour classer les transactions est le prix brut. En 1993, le prix moyen des appartements vendus était égal à 150 129 euros. Le prix moyen des 10% d'appartements les moins chers était égal à 37 239 euros. Source: Base BIEN.

## B Définition des groupes de pays

**Europe:** Albanie, Allemagne, Andorre, Autriche, Belgique, Bielorussie, Bosnie-herzegovine, Bouvet (ile), Bulgarie, Croatie, Danemark, Espagne, Estonie, Ex-republique yougoslave de macedoine, Feroe (iles), Finlande, Gibraltar, Grece, Guernesey, Hongrie, Irlande, ou eire, Islande, Italie, Jersey, Kosovo, Lettonie, Liechtenstein, Lituanie, Luxembourg, Malte, Man (ile), Moldavie, Monaco, Montenegro, Norvege, Pays-bas, Pologne, Portugal, Republique democratique allemande, Republique federale d'Allemagne, Roumanie, Royaume-uni, Russie, Saint-marin, Serbie, Slovaquie, Slovenie, Suede, Suisse, Svalbard et ile jan mayen, Tchechoslovaquie, Tcheque (republique), Turquie d'Europe, Ukraine, Vatican, ou saint-siege.

**Afrique:** Amérique: Alaska, Anguilla, Antigua-et-barbuda, Antilles neerlandaises, Argentine, Aruba, Bahamas, Barbade, Belize, Bermudes, Bolivie, Bresil, Caimanes (iles), Canada, Chili, Colombie, Costa rica, Cuba, Dominicaine (republique), Dominique, El salvador, Equateur, Etats-unis, Georgie du sud et les iles sandwich du sud, Grenade, Groenland, Guatemala, Guyana, Haiti, Honduras, Jamaïque, Labrador, Malouines, ou falkland (iles), Mexique, Montserrat, Nicaragua, Panama, Paraguay, Perou, Porto rico, Saint-christophe-et-nieves, Sainte-lucie, Saint-vincent-et-les grenadines, Suriname, Terr. des etats-unis d'amerique en amerique, Terr. du royaume-uni dans l'atlantique sud, Terre-neuve, Territoire des pays-bas, Territoires du royaume-uni aux antilles, Trinite-et-tobago, Turks et caïques (iles), Uruguay, Venezuela, Vierges britanniques (iles), Vierges des etats-unis (iles), Australie, Christmas (ile), Cocos ou keeling (iles), Cook (iles), Fidji, Guam, Hawaii (iles), Heard et macdonald (iles), Kiribati, Mariannes du nord (iles), Marshall (iles), Micronesie (etats federes de), Nauru, Niue, Norfolk (ile), Nouvelle-zelande, Palaos (iles), Papouasie-nouvelle-guinee, Pitcairn (ile), Salomon (iles), Samoa americaines, Samoa occidentales, Terr. des etats-unis d'amerique en oceanie, Tokelau, Tonga, Tuvalu, Vanuatu.

**Asie:** Afghanistan, Armenie, Azerbaïdjan, Bangladesh, Bhoutan, Birmanie, Brunei, Cambodge, Chine, Coree, Coree (republique de), Coree (republique populaire democratique de),

Etats malais non federes, Georgie, Goa, Hong-kong, Inde, Indonesie, Japon, Kamtchatka, Kazakhstan, Kirghizistan, Laos, Macao, Malaisie, Maldives, Mandchourie, Mongolie, Nepal, Ouzbekistan, Pakistan, Palestine, Philippines, Possessions britanniques au proche-orient, Siberie, Singapour, Sri lanka, Tadjikistan, Taiwan, Thaïlande, Timor oriental, Turkestan russe, Turkmenistan, Vietnam, Vietnam du nord, Vietnam du sud.

**Moyen-Orient:** Arabie saoudite, Bahrein, Chypre, Egypte, Emirats arabes unis, Iran, Iraq, Israel, Jordanie, Koweït, Liban, Oman, Qatar, Syrie, Turquie, Yemen, Yemen democratique, Yemen (republique arabe du).

**Afrique:** Acores, madere, Afrique du sud, Algerie, Angola, Benin, Botswana, Burkina, Burundi, Cameroun, Cameroun et togo, Canaries (iles), Cap-vert, Centrafricaine (republique), Comores, Congo, Congo (republique democratique), Cote d'ivoire, Djibouti, Erythree, Ethiopie, Gabon, Gambie, Ghana, Guinee, Guinee equatoriale, Guinee-bissau, Iles portugaises de l'ocean indien, Kenya, Lesotho, Liberia, Libye, Madagascar, Malawi, Mali, Maroc, Maurice, Mauritanie, Mozambique, Namibie, Niger, Nigeria, Ocean indien (territoire britannique de l'), Ouganda, Presides, Provinces espagnoles d'afrique, Rwanda, Sahara occidental, Sainte-hélène, ascension et tristan da cunh, Sao tome-et-principe, Senegal, Seychelles, Sierra leone, Somalie, Soudan, Soudan anglo-egyptien, kenya, ouganda, Swaziland, Tanger, Tanzanie, Tchad, Togo, Tunisie, Zambie, Zanzibar, Zimbabwe.

## C Exemple de régression hédonique

TABLE 5.20 – RÉGRESSION HÉDONIQUE, XV<sup>e</sup> ARRONDISSEMENT. 1993 À 2008.

Variable dépendante: log prix au m <sup>2</sup> <sup>(1)</sup>	Coefficient	(Écart-type)
Constante	8,702	(0,005)
<b>Époque de construction</b>		
Inconnue	0,061	(0,003)
Avant 1850	-0,004	(0,010)
1850-1913	ref	
1914-1947	0,016	(0,002)
1948-1969	0,046	(0,003)
1970-1980	0,103	(0,003)
Après 1981	0,335	(0,004)
<b>Salle de bains</b>		
Pas de salle de bains	-0,115	(0,002)
Une salle de bains	ref	
Deux salles de bains ou plus	0,014	(0,004)
<b>Garage</b>		
Pas de garage	ref	
Un garage	0,128	(0,003)
Deux garages ou plus	0,186	(0,007)
<b>Étage</b>		
Rez-de-chaussée	ref	
Premier étage	-0,042	(0,003)
Deuxième étage	-0,000	(0,003)
Troisième étage	0,003	(0,003)
Quatrième étage ou plus avec ascenseur	0,022	(0,002)
Quatrième étage ou plus sans ascenseur	-0,015	(0,006)
<b>Nombre de pièces</b>		
Une	-0,004	(0,002)
Deux	ref	
Trois	0,024	(0,002)
Quatre	0,041	(0,003)
Cinq ou plus	0,024	(0,005)
<b>Taille des pièces</b>		
Petite	0,019	(0,002)
Moyenne	ref	
Grande	-0,023	(0,002)
<b>Quartiers administratifs</b>		
Saint-Lambert	ref	
Grenelle	0,093	(0,002)
Necker	0,095	(0,002)
Javel	0,031	(0,002)
<b>Chambre de service</b>	0,151	(0,007)

<sup>(1)</sup> La régression comprend des effets fixes temporelles et trimestrielles.





## D Robustesse des résultats

### D.1 Régressions avec pour variable explicative le pourcentage d'acheteurs étrangers net du pourcentage de vendeurs étrangers

TABLE 5.21 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LA VARIABLE EXPLICATIVE EST LE POURCENTAGE D'ACHETEURS - NET DU POURCENTAGE DE VENDEURS - RÉSIDANT À L'ÉTRANGER.

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>% acheteurs -net des vendeurs- résidant à l'étranger</b>	2,303** (0,108)	0,663** (0,069)	0,693** (0,068)	0,341** (0,062)	0,309** (0,038)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,775	0,954	0,961	0,977	0,993
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la part des acheteurs nette des vendeurs résidant à l'étranger de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,309% pour les acheteurs personne physique. Source: Régressions à partir de la base BIEN.

TABLE 5.22 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LA VARIABLE EXPLICATIVE EST LE POURCENTAGE D'ACHETEURS - NET DU POURCENTAGE DE VENDEURS - DE NATIONALITÉ ÉTRANGÈRE.

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>% acheteurs -net des vendeurs- de nationalité étrangère</b>	2,138** (0,168)	0,556** (0,081)	0,690** (0,083)	0,224** (0,070)	0,324** (0,046)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,693	0,951	0,959	0,976	0,992
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la % acheteurs de nationalité étrangère parmi les acheteurs personnes physiques résidant à l'étranger de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,324% pour les acheteurs personne physique. Source: Régressions à partir de la base BIEN.

TABLE 5.23 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LES VARIABLES EXPLICATIVES SONT LES POURCENTAGES D'ACHETEURS - NETS DES VENDEURS - ÉTRANGERS RÉSIDANT OU NON EN FRANCE ET LE POURCENTAGE D'ACHETEURS FRANÇAIS RÉSIDANT À L'ÉTRANGER.

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>% acheteurs -net des vendeurs- de</b>	2,134**	0,805**	0,840**	0,424**	0,399**
<b>nationalité étrangère résidant à l'étranger</b>	(0,162)	(0,087)	(0,087)	(0,077)	(0,049)
<b>% acheteurs -nette des vendeurs- de</b>	-1,628**	-0,568**	-0,205	-0,436**	0,047
<b>nationalité étrangère résidant en France</b>	(0,375)	(0,187)	(0,202)	(0,143)	(0,097)
<b>% acheteurs -nette des vendeurs- français</b>	2,801**	0,320*	0,366*	0,126	0,166*
<b>résidant à l'étranger</b>	(0,292)	(0,148)	(0,152)	(0,111)	(0,069)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,782	0,955	0,961	0,978	0,993
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la part des acheteurs nette des vendeurs de nationalité étrangère résidant à l'étranger de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,399% pour les acheteurs personne physique.

TABLE 5.24 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LA VARIABLE EXPLICATIVE EST LE POURCENTAGE D'ACHETEURS - NET DES VENDEURS - RÉSIDANT À L'ÉTRANGER OU DE NATIONALITÉ ÉTRANGÈRE. **On se restreint aux 30% des transactions les plus chères.**

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
% acheteurs -net des vendeurs- de nationalité étrangère résidant à l'étranger	0,923**	0,380**	0,379**	0,252**	0,214**
	(0,060)	(0,031)	(0,030)	(0,030)	(0,024)
% acheteurs -net des vendeurs- de nationalité étrangère résidant en France	0,396**	-0,000	0,081	-0,063	0,000
	(0,125)	(0,056)	(0,057)	(0,047)	(0,039)
% acheteurs -net des vendeurs- français résidant à l'étranger	0,749**	0,084	0,093*	0,051	0,052
	(0,098)	(0,045)	(0,047)	(0,039)	(0,032)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,898	0,980	0,983	0,987	0,993
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la part des acheteurs nette des vendeurs de nationalité étrangère résidant à l'étranger parmi les acheteurs des 30% des transactions les plus chères de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,214% pour ces mêmes transactions.

TABLE 5.25 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LES VARIABLES EXPLICATIVES SONT LES POURCENTAGES D'ACHETEURS NETS DES VENDEURS EN PROVENANCE D'AMÉRIQUE, D'ASIE, D'EUROPE (HORS FRANCE), DU MOYEN-ORIENT ET D'AFRIQUE.

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
% acheteurs -net des vendeurs- d'Amérique	5,072** (0,478)	1,304** (0,248)	0,956** (0,280)	1,038** (0,196)	0,451** (0,150)
% acheteurs -net des vendeurs- d'Asie	-5,927** (0,639)	-1,272** (0,334)	-1,105** (0,345)	-0,665* (0,277)	0,023 (0,192)
% acheteurs -net des vendeurs- d'Europe	1,951** (0,177)	0,558** (0,092)	0,654** (0,094)	0,299** (0,077)	0,349** (0,052)
% acheteurs -net des vendeurs- du Moyen-Orient	2,760** (1,004)	2,214** (0,477)	2,704** (0,472)	-0,400 (0,434)	0,233 (0,288)
% acheteurs -net des vendeurs- d'Afrique	-0,878 (0,770)	-0,534 (0,376)	0,383 (0,396)	-0,939** (0,282)	0,104 (0,193)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,780	0,955	0,962	0,978	0,992
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la % acheteurs en provenance d'Amérique de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,451% pour les acheteurs personne physique.

TABLE 5.26 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LES VARIABLES EXPLICATIVES SONT LES POURCENTAGES D'ACHETEURS NETS DES VENDEURS RÉSIDANT EN PETITE OU GRANDE COURONNE, EN PROVINCE, DANS LES DOM-TOM OU À L'ÉTRANGER.

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
% acheteurs -net des vendeurs- résident à l'étranger	2,282** (0,105)	0,691** (0,071)	0,754** (0,070)	0,318** (0,064)	0,308** (0,040)
% acheteurs -net des vendeurs- résident en province	1,094** (0,144)	0,161* (0,074)	0,197** (0,074)	0,032 (0,057)	0,036 (0,035)
% acheteurs -net des vendeurs- résident en grande couronne	0,193 (0,266)	-0,094 (0,133)	0,000 (0,135)	-0,152 (0,101)	-0,022 (0,064)
% acheteurs -net des vendeurs- résident en petite couronne	0,581** (0,173)	0,071 (0,084)	0,191* (0,086)	-0,133* (0,062)	-0,064 (0,039)
% acheteurs -net des vendeurs- dans les DOM-TOMS	0,408 (0,862)	-0,105 (0,415)	0,138 (0,422)	-0,211 (0,305)	-0,003 (0,192)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,795	0,954	0,962	0,977	0,993
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la part des acheteurs nette des vendeurs résident à l'étranger de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,308% pour les acheteurs personne physique.





## D.2 Régressions où l'on n'a pas exclu les sociétés civiles immobilières et les acheteurs de type inconnu

TABLE 5.27 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LA VARIABLE EXPLICATIVE EST LE POURCENTAGE D'ACHETEURS RÉSIDANT À L'ÉTRANGER.

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>Part des acheteurs résidant à l'étranger</b>	2,425**	0,799**	0,859**	0,368**	0,379**
(Écart-type)	(0,083)	(0,070)	(0,067)	(0,083)	(0,055)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,831	0,955	0,964	0,977	0,993
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la proportion d'acheteurs de nationalité étrangère parmi les acheteurs personnes physiques résidant à l'étranger de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,379% pour les acheteurs personne physique.

TABLE 5.28 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LA VARIABLE EXPLICATIVE EST LE POURCENTAGE D'ACHETEURS DE NATIONALITÉ ÉTRANGÈRE.

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>Part des acheteurs de nationalité étrangère</b>	2,379**	0,562**	0,733**	0,100	0,275**
(Écart-type)	(0,157)	(0,087)	(0,089)	(0,087)	(0,065)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,713	0,949	0,957	0,976	0,992
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la proportion d'acheteurs de nationalité étrangère parmi les acheteurs personnes physiques résidant à l'étranger de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,275% pour les acheteurs personne physique.

TABLE 5.29 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LES VARIABLES EXPLICATIVES SONT LES POURCENTAGES D'ACHETEURS ÉTRANGERS RÉSIDANT OU NON EN FRANCE ET LE POURCENTAGE D'ACHETEURS FRANÇAIS RÉSIDANT À L'ÉTRANGER.

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>Part des acheteurs de nationalité étrangère résidant à l'étranger</b>	1,297** (0,140)	0,727** (0,090)	0,844** (0,093)	0,310** (0,093)	0,390** (0,067)
<b>Part des acheteurs de nationalité étrangère résidant en France</b>	-2,942** (0,320)	-1,274** (0,218)	-1,163** (0,228)	-0,708** (0,192)	-0,233 (0,135)
<b>Part des acheteurs français résidant à l'étranger</b>	4,840** (0,281)	1,116** (0,202)	0,961** (0,205)	0,488** (0,169)	0,313** (0,106)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,872	0,958	0,965	0,978	0,993
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la proportion d'acheteurs de nationalité étrangère parmi les acheteurs personnes physiques résidant à l'étranger de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,390% pour les acheteurs personne physique.

TABLE 5.30 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LA VARIABLE EXPLICATIVE EST LE POURCENTAGE D'ACHETEURS FRANÇAIS OU ÉTRANGER RÉSIDANT À L'ÉTRANGER ET LE POURCENTAGE D'ACHETEURS DE NATIONALITÉ FRANÇAISE RÉSIDANT EN FRANCE. **On se restreint aux 30% des transactions les plus chères.**

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>Part des acheteurs de nationalité étrangère</b>	0,870**	0,405**	0,435**	0,206**	0,190**
<b>résidant à l'étranger</b>	(0,058)	(0,034)	(0,035)	(0,036)	(0,030)
<b>Part des acheteurs de nationalité étrangère</b>	0,105	-0,057	-0,044	-0,071	-0,071
<b>résidant en France</b>	(0,159)	(0,084)	(0,087)	(0,071)	(0,057)
<b>Part des acheteurs français</b>	1,460**	0,288**	0,270**	0,131*	0,101*
<b>résidant à l'étranger</b>	(0,111)	(0,064)	(0,066)	(0,056)	(0,045)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,927	0,981	0,984	0,988	0,994
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la proportion d'acheteurs de nationalité étrangère parmi les acheteurs personnes physiques résidant à l'étranger de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,190% pour les acheteurs personne physique.

TABLE 5.31 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LES VARIABLES EXPLICATIVES SONT LES POURCENTAGES D'ACHETEURS EN PROVENANCE D'AMÉRIQUE, D'ASIE, D'EUROPE (HORS FRANCE), DU MOYEN-ORIENT ET D'AFRIQUE.

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>Part des acheteurs d'Amérique</b>	5,576**	1,510**	1,519**	0,871**	0,613**
(Écart-type)	(0,485)	(0,287)	(0,318)	(0,243)	(0,198)
<b>Part des acheteurs d'Asie</b>	-5,452**	-1,607**	-1,646**	-1,089**	-0,268
(Écart-type)	(0,709)	(0,409)	(0,412)	(0,396)	(0,281)
<b>Part des acheteurs d'Europe</b>	1,587**	0,504**	0,596**	0,243*	0,332**
(Écart-type)	(0,190)	(0,108)	(0,110)	(0,097)	(0,070)
<b>Part des acheteurs du Moyen-Orient</b>	5,027**	2,798**	3,106**	-0,897	-0,168
(Écart-type)	(0,917)	(0,508)	(0,508)	(0,521)	(0,361)
<b>Part des acheteurs d'Afrique</b>	-5,348**	-2,391**	-1,180*	-2,293**	-0,456
(Écart-type)	(0,882)	(0,488)	(0,512)	(0,385)	(0,271)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,819	0,956	0,963	0,979	0,992
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la proportion d'acheteurs en provenance d'Amérique de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,613% pour les acheteurs personne physique.

TABLE 5.32 – RÉSULTATS DES RÉGRESSIONS. PÉRIODE: 1993 À 2008. LES VARIABLES EXPLICATIVES SONT LES POURCENTAGES D'ACHETEURS RÉSIDANT EN PETITE OU GRANDE COURONNE, EN PROVINCE, DANS LES DOM-TOM OU À L'ÉTRANGER.

Variable dépendante: moyenne biannuelle par quartier administratif  
du logarithme du prix "équivalent" au m<sup>2</sup> en euros 2008

	(1)	(2)	(3)	(4)	(5)
<b>Part des acheteurs résidant à l'étranger</b>	1,800**	0,635**	0,746**	0,255**	0,363**
(Écart-type)	(0,101)	(0,076)	(0,075)	(0,086)	(0,059)
<b>Part des acheteurs résidant en province</b>	1,916**	0,092	0,161	-0,137	-0,041
(Écart-type)	(0,180)	(0,125)	(0,128)	(0,107)	(0,072)
<b>Part des acheteurs résidant en grande couronne</b>	-1,114**	-0,969**	-0,739**	-0,569**	-0,058
(Écart-type)	(0,312)	(0,196)	(0,201)	(0,160)	(0,105)
<b>Part des acheteurs résidant en petite couronne</b>	-1,352**	-0,537**	-0,389*	-0,318*	-0,024
(Écart-type)	(0,234)	(0,150)	(0,156)	(0,123)	(0,082)
<b>Part des acheteurs dans les DOM-TOMS</b>	0,347	-0,570	-0,370	-0,370	-0,058
(Écart-type)	(1,053)	(0,628)	(0,627)	(0,493)	(0,309)
Années	OUI	OUI	OUI	OUI	OUI
Arrondissements	NON	OUI	OUI	OUI	OUI
Arrondissements et années croisées	NON	NON	OUI	NON	OUI
Quartiers	NON	NON	NON	OUI	OUI
R <sup>2</sup> corrigé	0,870	0,958	0,965	0,978	0,993
Nombre d'observations	640	640	640	640	640

*Note:* Les écarts-types sont indiqués entre parenthèses. \*\*\* : significatif à 5%, \*\* : significatif à 10%. Les observations sont au niveau du quartier administratif et correspondent à des tranches de deux ans. D'après la régression avec les contrôles de la cinquième colonne, une augmentation de la proportion d'acheteurs résidant à l'étranger de 1 point de pourcentage entraîne une augmentation du prix au m<sup>2</sup> équivalent de 0,255% pour les acheteurs personne physique.

# References

- Adam, K., P. Kuang, and A. Marcet (2011). House price booms and the current account. In *NBER Macroeconomics Annual 2011, Volume 26*, NBER Chapters. National Bureau of Economic Research, Inc.
- Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015). The economics of density: Evidence from the berlin wall. *Econometrica* 83(6), 2127–2189.
- Alain David, François Dubujet, C. G. A. L. (2002). *Les indices de prix des logements anciens*, Volume 98. Insee Méthode.
- Albouy, D., G. Ehrlich, and M. Shin (2018). Metropolitan land values. *Review of Economics and Statistics* 100(3), 454–466.
- Alvaredo, F. (2009). Top incomes and earnings in portugal 1936–2005. *Explorations in economic history* 46(4), 404–417.
- Alvaredo, F. and E. Saez (2009). Income and wealth concentration in spain from a historical and fiscal perspective. *Journal of the european economic Association* 7(5), 1140–1167.
- Atkinson, A. B. (2005). Top incomes in the uk over the 20th century. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 168(2), 325–343.
- Atkinson, A. B. and W. Salverda (2005). Top incomes in the netherlands and the united kingdom over the 20th century. *Journal of the European Economic Association* 3(4), 883–913.

- Banzhaf, H. S. and R. P. Walsh (2008). Do people vote with their feet? an empirical test of tiebout. *American Economic Review* 98(3), 843–63.
- Baum-Snow, N. and L. Han (2019). The microgeography of housing supply. Technical report, mimeo.
- Baum-Snow, N. and D. Hartley (2019). Accounting for central neighborhood change, 1980-2010. *Journal of Urban Economics*, 103228.
- Bazot, G. (2014). Interregional inequalities, convergence, and growth in france from 1840 to 1911. *Annals of Economics and Statistics/Annales d'Économie et de Statistique* (113/114), 309–345.
- Besley, T. and S. Coate (2003). Centralized versus decentralized provision of local public goods: a political economy approach. *Journal of public economics* 87(12), 2611–2637.
- Bewley, T. F. (1981). A critique of tiebout's theory of local public expenditures. *Econometrica: Journal of the Econometric Society*, 713–740.
- Blanchet, T., J. Fournier, and T. Piketty (2017). Generalized pareto curves: Theory and applications.
- Bloch, F. and E. Ü. Zenginobuz (2006). Tiebout equilibria in local public good economies with spillovers. *Journal of Public Economics* 90(8-9), 1745–1763.
- Bloch, F. and U. Zenginobuz (2007). The effect of spillovers on the provision of local public goods. *Review of Economic Design* 11(3), 199–216.
- Bloch, F. and Ü. Zenginobuz (2015). Oates' decentralization theorem with imperfect household mobility. *International Tax and Public Finance* 22(3), 353–375.
- Boadway, R. and J.-F. Tremblay (2012). Reassessment of the tiebout model. *Journal of public economics* 96(11-12), 1063–1078.



- 
- Bonnet, F. (2018a). Beyond the exodus of may-june 1940: Internal migrations in france during the second world war. Technical report, mimeo.
- Bonnet, F. (2018b). Spatial distribution of populations by age, 1851–2014. *mimeo*.
- Bonnet, F. and H. d’Albis (2018). Spatial inequality in mortality in france over the past two centuries.
- Bonnet, F. and A. Sotura (2020). Intra-regional inequality in france in the last fifty years.
- Boyer, P. C., T. Delemotte, G. Gauthier, V. Rollet, and B. Schmutz (2020). Les déterminants de la mobilisation des gilets jaunes. *Revue économique* 71(1), 109–138.
- Brueckner, J. K. (2000). A tiebout/tax-competition model. *Journal of Public Economics* 77(2), 285–306.
- Brueckner, J. K. (2003). Strategic interaction among governments: An overview of empirical studies. *International regional science review* 26(2), 175–188.
- Busso, M., J. Gregory, and P. Kline (2013). Assessing the incidence and efficiency of a prominent place based policy. *American Economic Review* 103(2), 897–947.
- Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics* 90(3), 414–427.
- Carbonnier, C. et al. (2008). Fiscal competition between decentralized jurisdictions, theoretical and empirical evidence. Technical report, THEMA (THéorie Economique, Modélisation et Applications), Université de . . . .
- Card, D. (1989, August). The impact of the mariel boatlift on the miami labor market. Working Paper 3069, National Bureau of Economic Research.
- Card, D. and J. E. DiNardo (2000, March). Do immigrant inflows lead to native outflows? Working Paper 7578, National Bureau of Economic Research.

- Card, D. E. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics* 19(1), 22–64.
- Caruana-Galizia, P. (2013). Indian regional income inequality: estimates of provincial gdp, 1875–1911. *Economic History of Developing Regions* 28(1), 1–27.
- Case, A. C., H. S. Rosen, and J. R. Hines Jr (1993). Budget spillovers and fiscal policy interdependence: Evidence from the states. *Journal of public economics* 52(3), 285–307.
- Cheikbossian, G. (2008). Rent-seeking, spillovers and the benefits of decentralization. *Journal of Urban Economics* 63(1), 217–228.
- Chen, Y. and S. S. Rosenthal (2008). Local amenities and life-cycle migration: Do people move for jobs or fun? *Journal of Urban Economics* 64(3), 519–537.
- Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annu. Rev. Econ.* 1(1), 451–488.
- Combes, P.-P., G. Duranton, and L. Gobillon (2019). The costs of agglomeration: House and land prices in french cities. *The Review of Economic Studies* 86(4), 1556–1589.
- Combes, P.-P., G. Duranton, L. Gobillon, et al. (2018). The costs of agglomeration: House and land prices in french cities. *CEPR Discussion Paper No. DP9240 (updated)*.
- Combes, P.-P., M. Lafourcade, J.-F. Thisse, and J.-C. Toutain (2011). The rise and fall of spatial inequalities in france: A long-run perspective. *Explorations in Economic History* 48(2), 243–271.
- Conley, J. and M. Dix (1999). Optimal and equilibrium membership in clubs in the presence of spillovers. *Journal of Urban Economics* 46(2), 215–229.
- Couture, V., C. Gaubert, J. Handbury, and E. Hurst (2019). Income growth and the distributional effects of urban spatial sorting. Technical report, National Bureau of Economic Research.

- 
- Couture, V. and J. Handbury (2017). Urban revival in america, 2000 to 2010. Technical report, National Bureau of Economic Research.
- Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review* 106(3), 479–524.
- Epple, D. and H. Sieg (1999). Estimating equilibrium models of local jurisdictions. *Journal of political economy* 107(4), 645–681.
- Fack, G. (2007). *Formation des Inégalités, Politiques du Logement et Ségrégation résidentielle*. Ph. D. thesis, Ecole des Hautes Etudes en Sciences Sociales.
- Fajgelbaum, P. and C. Gaubert (2018). Optimal spatial policies, geography and sorting. Technical report, National Bureau of Economic Research.
- Fajgelbaum, P. D., E. Morales, J. C. S. Serrato, and O. M. Zidar (2015). State taxes and spatial misallocation. Technical report, National Bureau of Economic Research.
- Ferrante, A., D. Guillas, and R. Solotareff (2016). Entre 2010 et 2015, les inégalités de patrimoine se réduisent légèrement.
- Frank, M. (2015). Frank-sommeiller-price series for top income shares by us states since 1917. *WTID Methodological Notes*.
- Fuest, C., A. Peichl, and S. Siegloch (2018, February). Do higher corporate taxes reduce wages? micro evidence from germany. *American Economic Review* 108(2), 393–418.
- Gaigne, C., H. R. Koster, F. Moizeau, J.-F. Thisse, et al. (2020). Income sorting across space: The role of amenities and commuting costs. Technical report.
- Garbinti, B., J. Goupille-Lebret, et al. (2019). Income and wealth inequality in france: Developments and links over the long term. *Economie et Statistique/Economics and Statistics* (510-511-512), 69–87.

- Garbinti, B., J. Goupille-Lebret, and T. Piketty (2018). Income inequality in france, 1900–2014: Evidence from distributional national accounts (dina). *Journal of Public Economics* 162, 63–77.
- Geary, F. and T. Stark (2002). Examining ireland’s post-famine economic growth performance. *The Economic Journal* 112(482), 919–935.
- Giannone, E. (2017). Skilled-biased technical change and regional convergence. Technical report, Technical report, University of Chicago Working Paper, available at: [http . . .](http://www.uchicago.edu/~giannone)
- Gobillon, L. and D. Le Blanc (2008). Economic effects of upfront subsidies to ownership: The case of the prêt à taux zéro in france. *Journal of Housing Economics* 17(1), 1–33.
- Godechot, O. (2013). Financiarisation et fractures socio-spatiales. *L’année sociologique* 63(1), 17–50.
- Gonzalez, L. and F. Ortega (2009, July). Immigration and housing booms: Evidence from spain. CReAM Discussion Paper Series 0919, Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London.
- Gordon, R. H. (1983). An optimal taxation approach to fiscal federalism. *The Quarterly Journal of Economics* 98(4), 567–586.
- Greulich, E., J. M. Quigley, and S. Raphael (2005). The anatomy of rent burdens: Immigration, growth and rental housing. Berkeley program on housing and urban policy, working paper series, Berkeley Program on Housing and Urban Policy.
- Hsieh, C.-T. and E. Moretti (2015). Why do cities matter? local growth and aggregate growth.
- Kline, P. and E. Moretti (2014). People, places, and public policy: Some simple welfare economics of local economic development programs.
- Konishi, H. (1996). Voting with ballots and feet: Existence of equilibrium in a local public good economy. *Journal of Economic Theory* 68(2), 480–509.

- 
- Krugman, P. (1991). Increasing returns and economic geography. *Journal of political economy* 99(3), 483–499.
- Landvoigt, T., M. Piazzesi, and M. Schneider (2015). The housing market (s) of san diego. *American Economic Review* 105(4), 1371–1407.
- Leprince, M. and A. Guengant (2002). Interactions fiscales verticales et réaction des communes à la coopération intercommunale. *Revue économique* 53(3), 525–535.
- Lessmann, C. (2014). Spatial inequality and development—is there an inverted-u relationship? *Journal of development economics* 106, 35–51.
- Lessmann, C. and A. Seidel (2017). Regional inequality, convergence, and its determinants—a view from outer space. *European Economic Review* 92, 110–132.
- Lockwood, B. (2002). Distributive politics and the costs of centralization. *The Review of Economic Studies* 69(2), 313–337.
- Lutz, B. (2015). Quasi-experimental evidence on the connection between property taxes and residential capital investment. *American Economic Journal: Economic Policy* 7(1), 300–330.
- McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*.
- Moretti, E. (2012). *The new geography of jobs*. Houghton Mifflin Harcourt.
- Oates, W. E. (1969). The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the tiebout hypothesis. *Journal of political economy* 77(6), 957–971.
- Oates, W. E. (1972). *Fiscal Federalism*. Number 14708 in Books. Edward Elgar Publishing.
- Oates, W. E. (2005). Toward a second-generation theory of fiscal federalism. *International tax and public finance* 12(4), 349–373.

- Ortalo-Magne, F. and S. Rady (2006). Housing market dynamics: On the contribution of income shocks and credit constraints. *The Review of Economic Studies* 73(2), 459–485.
- Oswalt, P. and T. Rieniets (2006). *Atlas of shrinking cities*. Hatje Cantz.
- Ottaviano, G. I. and G. Peri (2005). Rethinking the gains from immigration: Theory and evidence from the u.s. NBER Working Papers 11672, National Bureau of Economic Research, Inc.
- Ottaviano, G. I. P. and G. Peri (2007, November). The effects of immigration on us wages and rents: A general equilibrium approach. CEPR Discussion Papers 6551, C.E.P.R. Discussion Papers.
- Piketty, T. (2001). *Les hauts revenus en France au XXe siècle*. Hachette.
- Roback, J. (1982). Wages, rents, and the quality of life. *Journal of political Economy* 90(6), 1257–1278.
- Roine, J., J. Vlachos, and D. Waldenström (2009). The long-run determinants of inequality: What can we learn from top income data? *Journal of Public Economics* 93(7-8), 974–988.
- Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy* 82(1), 34–55.
- Saiz, A. (2003a). Immigration and housing rents in american cities. Working Papers 03-12, Federal Reserve Bank of Philadelphia.
- Saiz, A. (2003b, 05). Room in the kitchen for the melting pot: Immigration and rental prices. *The Review of Economics and Statistics* 85(3), 502–521.
- Saiz, A. (2006). Immigration and housing rents in american cities. IZA Discussion Papers 2189, Institute for the Study of Labor (IZA).
- Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics* 125(3), 1253–1296.

- Sanchis, M. T., J. R. Rosés, and A. Díez (2015). Regional inequality in France 1860–2010: Structural change dynamics. In *International Conference on Regional Science, Universitat Rogira i Virgili (Spain), November*, pp. 18–20.
- Serrato, J. C. S. and P. Wingender (2011). Estimating the incidence of government spending. Technical report, Mimeo, 2011. 1.
- Serrato, J. C. S. and P. Wingender (2016). Estimating local fiscal multipliers. Technical report, National Bureau of Economic Research.
- Solé-Ollé, A. (2006). Expenditure spillovers and fiscal interactions: Empirical evidence from local governments in Spain. *Journal of Urban Economics* 59(1), 32–53.
- Suárez Serrato, J. C. and O. Zidar (2016). Who benefits from state corporate tax cuts? a local labor markets approach with heterogeneous firms. *American Economic Review* 106(9), 2582–2624.
- Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of political economy* 64(5), 416–424.
- Toutain, J.-C. (1993). *La production agricole de la France de 1810 à 1990: départements et régions: croissance, productivité, structures*. Number 17. Presses universitaires de Grenoble.
- Tsivanidis, N. (2018). The aggregate and distributional effects of urban transit infrastructure: Evidence from Bogotá’s TransMilenio. *Job Market Paper*.
- Wellisch, D. (1994). Interregional spillovers in the presence of perfect and imperfect household mobility. *Journal of Public Economics* 55(2), 167–184.
- Williamson, J. G. (1965). Regional inequality and the process of national development: a description of the patterns. *Economic development and cultural change* 13(4, Part 2), 1–84.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.





# List of tables

1.1	Descriptive statistics on NIL new dwelling buyers . . . . .	23
1.2	Descriptive statistics on NIL existing dwelling buyers . . . . .	24
2.1	Annual distributions of departmental differences between tax units recorded and estimated according to the two methods (in %) . . . . .	99
3.1	R <sup>2</sup> des différentes estimations . . . . .	112
3.2	Données utilisées par période temporelle . . . . .	113
3.3	Indicateurs de Gini, 90 départements . . . . .	116
4.1	Descriptive Statistics on jurisdictions . . . . .	184
4.2	Within-MF Sensitivity Analysis . . . . .	185
4.3	Between-MF Sensitivity Analysis . . . . .	186
4.4	Housing Supply Sensitivity Analysis . . . . .	187
4.5	Structural parameters for $\alpha = 0.25$ . . . . .	188
4.6	Structural parameters for $\alpha = 0.30$ . . . . .	188
4.7	Structural parameters for $\alpha = 0.35$ . . . . .	189
4.8	Structural parameters for $\alpha = 0.40$ . . . . .	189
4.9	Structural Parameters Calibration and Estimation . . . . .	190
4.10	Welfare change estimation for $\alpha=0.30$ . . . . .	191
4.11	Welfare change estimation for $\eta = 0.20$ . . . . .	193
5.1	Parc de logements parisien . . . . .	226
5.2	Statut d'occupation des logements . . . . .	227
5.3	Mode d'occupation des logements . . . . .	227
5.4	Type de propriétaires des logements . . . . .	228
5.5	Mode d'occupation des logements selon le lieu de résidence des propriétaires . . . . .	228
5.6	Part des propriétaires étrangers selon le mode d'occupation . . . . .	229
5.7	Type de propriétaire selon le lieu de résidence . . . . .	229
5.8	Part des propriétaires étrangers selon leur type . . . . .	230
5.9	Nombre d'appartements et de maisons vendus à Paris . . . . .	233
5.10	Répartition du type d'acheteurs . . . . .	234
5.11	Effet sur les prix de la part des acheteurs résidant à l'étranger . . . . .	255
5.12	Effet sur les prix de la part des acheteurs de nationalité étrangère . . . . .	257
5.13	Effet sur les prix de la part des acheteurs selon leur lieu de résidence et leur nationalité . . . . .	262
5.14	Effet sur les prix de la part des acheteurs selon leur département de résidence . . . . .	264
5.15	Effet sur les prix de la part des acheteurs selon leur continent d'origine . . . . .	270

5.16	Effet sur les prix de la part des acheteurs étrangers lorsqu'on se restreint aux 30% de transactions les plus chères . . . . .	276
5.17	Impact des étrangers sur les prix . . . . .	282
5.18	Impact des étrangers sur les prix de l'immobilier haut de gamme . . . . .	283
5.19	Prix moyen par décile de prix . . . . .	289
5.20	Exemple de régression hédonique . . . . .	292
5.21	Effet sur les prix de la % acheteurs <i>nette des vendeurs</i> résidant à l'étranger .	294
5.22	Effet sur les prix de la % acheteurs <i>nette des vendeurs</i> de nationalité étrangère	295
5.23	Effet sur les prix de la % acheteurs <i>nette des vendeurs</i> selon leur lieu de résidence et leur nationalité . . . . .	296
5.24	Effet sur les prix de la % acheteurs <i>nette des vendeurs</i> de nationalité étrangère ou résidant à l'étranger lorsqu'on se restreint à l'immobilier haut de gamme .	297
5.25	Effet sur les prix de la % acheteurs <i>nette des vendeurs</i> selon leur continent d'origine . . . . .	298
5.26	Effet sur les prix de la % acheteurs <i>nette des vendeurs</i> selon leur département de résidence . . . . .	299
5.27	Effet sur les prix de la proportion d'acheteurs résidant à l'étranger . . . . .	301
5.28	Effet sur les prix de la proportion d'acheteurs de nationalité étrangère . . . .	302
5.29	Effet sur les prix de la proportion d'acheteurs selon leur lieu de résidence et leur nationalité . . . . .	303
5.30	Effet sur les prix de la proportion d'acheteurs de nationalité étrangère ou résidant à l'étranger lorsqu'on se restreint à l'immobilier haut de gamme . .	304
5.31	Effet sur les prix de la proportion d'acheteurs selon leur continent d'origine .	305
5.32	Effet sur les prix de la proportion d'acheteurs selon leur département de résidence	306

# List of figures

1.1	Evolution of the no-interest loan subsidy for new and existing units . . . . .	17
1.2	Evolution of the income eligibility threshold, income of eligible households and beneficiaries for a no-interest loan to finance the purchase of a new unit . . . . .	18
1.3	Share of eligible households among all households and among renters . . . . .	19
1.4	Evolution of the definition of geographical areas for the subsidy . . . . .	20
1.6	Location of NIL beneficiaries that live in Paris urban area . . . . .	22
2.1	Between and within département inequality measured by the Theil index . . . . .	72
2.2	Between département inequality by income groups . . . . .	73
2.3	Average income of top 1% tax units of their département . . . . .	74
2.5	Contribution to the total inequality . . . . .	77
2.6	National income groups segregation . . . . .	80
2.7	National top 1% tax units location (% living in each département) . . . . .	81
2.9	National P90-P99 tax units location (% living in each département) . . . . .	83
2.11	Over/under-representation of the national top 1% tax units . . . . .	85
2.13	Over/under-representation of the P0P99 tax units . . . . .	87
2.14	Income shares of the national distribution and the sum of départements distributions . . . . .	98
2.16	Ratio of predicted values in specification I over specification II (1) . . . . .	100
2.18	Ratio of predicted values in specification I over specification II (2) . . . . .	101
2.20	Share of the working age population living in Haute Savoie that are commuting to Switzerland (in %) . . . . .	102
3.1	Coefficient de Gini du revenu par adulte entre 90 départements, 1922-2015 . . . . .	115
3.2	P90-100 du revenu par adulte entre 90 départements, 1922-2015 . . . . .	117
3.3	P0-20 du revenu par adulte entre 90 départements, 1922-2015 . . . . .	118
3.4	Part des départements dont le revenu moyen est supérieur à 60% du celui de la Seine. 1922-2015 . . . . .	120
3.5	Coefficient de Gini du revenu par km2 entre 90 départements, 1922-2015. . . . .	121
3.6	Rapports entre les revenus fiscaux par adulte de chaque département et la moyenne des revenus fiscaux par adulte des 90 départements en 1922 . . . . .	122
3.7	Rapports entre les revenus fiscaux par adulte de chaque département et la moyenne des revenus fiscaux par adulte des 90 départements en 1948 . . . . .	123
3.8	Rapports entre les revenus fiscaux par adulte de chaque département et la moyenne des revenus fiscaux par adulte des 90 départements en 2015 . . . . .	124
3.9	Catégorisation départements selon leur évolution relative entre 1922 et 1948 . . . . .	126
3.10	Catégorisation départements selon leur évolution relative entre 1948 et 2015 . . . . .	127

3.11	Carte et liste des 90 départements français utilisés dans l'article . . . . .	132
3.12	Taux d'imposition moyen, minimum et maximum, 1922-2015 . . . . .	133
3.13	P90-100 du revenu par adulte, 90 départements pondérés par leur population adulte, 1922-2015 . . . . .	134
3.14	P0-20 du revenu par adulte, 90 départements pondérés par leur population adulte, 1922-2015 . . . . .	135
3.15	Coefficient de Gini du revenu par adulte entre 87 départements hors ceux d'Ile-de-France (Seine, Seine-et-Marne et Seine-et-Oise), 1922-2015 . . . . .	136
3.16	Coefficient de Gini du revenu fiscal par adulte en France et aux Etats-Unis. 1922=1 . . . . .	137
4.1	Distribution of Subsidy Stock Yearly Growth . . . . .	192
4.3	Within-MF Public Good Changes . . . . .	194
4.4	Within-MF Migration Response . . . . .	195
4.5	Within-MF Per Capita Housing and Wage . . . . .	196
4.7	Between-MF Local Public Good Changes . . . . .	197
4.8	Between-MF Migration Response . . . . .	198
4.9	Between-MF Per Capita Housing and Wage . . . . .	199
4.11	Within-MF Housing Supply and House Prices . . . . .	200
4.13	Between-MF Housing Supply and House Prices . . . . .	200
4.15	Kernel Regressions: Within-MF Changes . . . . .	201
4.17	Kernel Regressions: Between-MF Changes (1/3) . . . . .	202
4.19	Kernel Regressions: Between-MF Changes (1/3) . . . . .	203
4.21	Model Fit: $G^f$ vs $G^s$ . . . . .	219
4.22	Changes in (Public Employees)/ (20-65 Population) . . . . .	220
5.1	Part des propriétaires résidant à l'étranger . . . . .	230
5.2	Surface moyenne du parc de logement . . . . .	231
5.3	Département de résidence des propriétaires . . . . .	231
5.4	Continent de résidence des propriétaires étrangers . . . . .	232
5.5	Nombre de transactions dans notre base . . . . .	235
5.6	Prix moyen des appartements vendus . . . . .	236
5.7	Surface moyenne des appartements vendus . . . . .	237
5.8	Part des acheteurs étrangers . . . . .	238
5.9	Département de résidence des acheteurs . . . . .	239
5.10	Continent de résidence des acheteurs . . . . .	240
5.11	Variation du prix moyen des transactions et du pourcentage d'acheteurs étrangers par décile de prix . . . . .	244
5.12	Variation du prix moyen et de la surface moyenne des transactions par décile de prix . . . . .	245
5.13	Variation du prix moyen des transactions et du pourcentage d'acheteurs étrangers par décile de prix hédoniques . . . . .	246
5.14	Prix relatif payé par les acheteurs résidant à l'étranger . . . . .	253
5.15	Part des achats due aux acheteurs de nationalité étrangère ou résidant à l'étranger	254
5.16	Prix relatif payé par les acheteurs de nationalité étrangère . . . . .	256
5.17	Prix relatif payé par les acheteurs étrangers résidant à l'étranger . . . . .	258
5.18	Prix relatif payé par les acheteurs français résidant à l'étranger . . . . .	259

5.19	Prix relatif payé par les acheteurs étrangers résidant en France . . . . .	260
5.20	Part des achats selon le lieu de résidence et la nationalité des acheteurs . . .	261
5.21	% des achats selon le département de résidence des acheteurs . . . . .	263
5.22	Prix relatif payé par les acheteurs en provenance d'Amérique . . . . .	269
5.23	Prix relatif payé par les acheteurs en provenance d'Asie . . . . .	271
5.24	Prix relatif payé par les acheteurs en provenance d'Afrique . . . . .	272
5.25	Prix relatif payé par les acheteurs en provenance du Moyen-Orient . . . . .	273
5.26	Prix relatif payé par les acheteurs en provenance d'Europe . . . . .	274
5.27	Part des achats selon le continent d'origine des acheteurs . . . . .	275