Return on Investment for Promoting Homeownership: The case of the French Interest-Free Loan Policy

Rémi Lei*
remi.lei.univ.acad@gmail.com

Jean-Sauveur Ay*
jean-sauveur.ay@inrae.fr

Julie Le Gallo*
julie.le-gallo@institut-agro.fr

September 2024

Abstract

This article estimates the impact of subsidised loans for new homeownership on the number of new homeowners (extensive margin), housing choices (intensive margin), and dwelling prices (capitalization effect). Our identification relies on the spatial and temporal variation of the French interest-free loan policy over the last decade, controlling for confounding assignment through a spatial semi-parametric propensity score. Our doubly robust results cannot rule out that the policy has no effect at the extensive margin, while it has significant intensive margin and capitalization effects. By considering a wide range of different policy objectives in terms of extensive and intensive margins, we compute the returns to government spending for counterfactual policy schemes and credit market conditions. Our simulations suggest that for reasonable values of policy objectives, increasing public spending has a return on investment lower than one, and may even be negative in some situations.

JEL classification: H81, R21, R38

Keywords: Housing policy ; subsidised mortgage ; unconfoundedness ; generalized additive model ; spatial smoothing.

*UMR CESAER, INRAE, Institut Agro, Univ. Bourgogne-Franche-Comté, F-21000, Dijon, France.

1 Introduction

Access to homeownership is increasingly challenging in the context of the housing affordability crisis. The divergence between dwelling prices and households' incomes represents the primary obstacle, not compensated by the decrease in interest rates in the 2010s (Barone et al., 2021). Not only does the difficulty in accessing homeownership raise concerns with respect to inequality between households in terms of wealth accumulation (Sodini et al., 2023), but the slowdown in homeownership may also be detrimental from a social perspective. The positive externalities associated with homeownership, including improved housing maintenance (Harding, Miceli and Sirmans, 2000), school performance (Green and White, 1997; Harkness and Newman, 2003), or self-employment opportunities (Harding and Rosenthal, 2017) are found to outweigh the negative ones, such as reduced residential mobility (Green and Hendershott, 2001), or the risk of capital loss (Cunningham and Reed, 2013). Thus, homeowners generate positive externalities that amount to approximately 1,300$ per year (Coulson and Li, 2013). The slowdown of homeownership reduce these externalities, possibly leading to a loss for society. Hence, most developed countries support homeownership to increase welfare in addition to tenants' aspirations.

The return on investment of public spending to support homeownership depends on its ability to increase the homeownership, which is commonly defined as the extensive margin. Most policies leverage the financial channel in order to make tenants solvable. Whereas the direct effect on tenure decision is not clearcut, the stimulation of housing demand in the context of low elasticity of housing supply may inadvertently exacerbate affordability issues through price capitalisation, and thus strengthen difficulties to homeownership. In addition to the effect at the extensive margin, making tenants solvable may also change their housing choices because of these policies (variation in size, location or characteristics), commonly defined as the intensive margin. While this effect is not inherently negative from a political point of view, it does not encourage the production of positive externalities associated with homeownership. The most widespread policy supporting homeownership (the Mortgage Interest Deduction, MID) appears inefficient (Valentin, 2023) as it produces most of these effects at the intensive margin (Hanson, 2012), whereas reinforcing affordability issues (Martin and Hanson, 2016) that are detrimental for homeownership development (Sommer and Sullivan, 2018). Hence, the policy return on investment depends on the balance between extensive and intensive margin effects in regards with policy cost.

However, a number of countries, such as France, have implemented alternative policy schemes to the MID where recipients benefit from a complementary loan with no interest charge rather than income tax reduction. In these cases, policymakers determine both the proportion of the loan that will be subsidized (the *covering share*) and the maximum loan value (the *ceiling value*). Despite the fact that this type of policy has been implemented in various countries, policy assessments are scarcer than the MID in the literature with mixed results about impact on homeownership(Gobillon and Le Blanc, 2008; Hembre, 2018; Carozzi, Hilber and Yu, 2024). Yet, these results are context-dependent,[1] which make the comparison difficult. In this paper, we estimate both extensive and intensive margin responses to publicly supported loan by computing return on investment of public spending, which overcomes context dependency. We leverage the French Interest-Free-Loans (IFL) policy from 2015 to 2019,[2] which provides financial support to potential first-time owners purchasing newly built housing units.[3]

We estimate the extensive and intensive margins responses to treatment intensity by leveraging spatio-temporal variations of the IFL policy at the French national scale. This French scheme employs a four-level spatial delineation, which defines the value of both the covering share and ceiling value for each year, as set by policymakers. However, the policy support is of greater importance in areas with the highest price per square metres. The four-level design is called the ABC zoning, which is shaped to reflect the tenseness of the local housing market. It is assumed that the term "tenseness" relates to housing affordability.[4] Hence, spatio-temporal variations of treatment are endogenous to the municipality characteristics. To handle this endogeneity, we exploit a selection-on-observables approach with a spatial smoothing specification to control for unobserved spatial confounders (Gilbert et al., 2024). Our approach is close, in terms of identifying conditions, to a geographic regression discontinuity approach. Yet, we leverage all spatial heterogeneity in treatment intensity, which avoids the estimation of local treatment effects hindering the applicability of our results to other contexts.

We proceed as follows. In the initial stage, we estimate the probability at the municipality level to receive each level of treatment from a large set of pre-treatment variables and a smooth bivariate function of geographical coordinates. The estimation procedure is carried out using an ordinal logistic regression accounting for the ordered structure of the treatment intensity and reflects, as in the policy scheme, that tenseness is a latent variable. In the second stage, we employ an inverse probability weighting mechanism derived from the Generalised Propensity Score (GPS, Imbens (2000)) to regress outcomes of the extensive and the intensive margins (respectively new homeowners and the housing choices of recipient) and capitalisation effects (the average housing price). The

[1] Hembre (2018) find positive effect on the number of homeowners in the context of the Great Recession. In contrast, Carozzi, Hilber and Yu (2024) finds no significant effect for the Greater London.

[2] This period contains three important reforms about treatment intensity leaving eligibility conditions unaffected. The policy also restricts access based on income cap, but the income ceiling is not restrictive (Sotura, 2020).

[3] The French housing policy defined potential first-time owners as individuals that rent their main residence for at least two years.

[4] According to the definition provided by the administration in charge of housing policies, tenseness is defined as "the imbalance between the housing supply and the housing demand".

assumption allowing to interpret these results as policy effects is that conditional on the set of pre-treatment variables and location, treatment variations are exogenous to municipality characteristics. The weighting scheme aims to weight observations according to their counterfactual power to estimate potential outcomes if municipality would have received another treatment level. Finally, we regress the estimated treatment effects for different treatment intensities on the policy schemes in place (the covering share and the ceiling value) and mortgage conditions (interest rate and loan maturity). For extensive and intensive margin effects, we also account for the indirect effect resulting from the price capitalisation. Based on these estimations, we compute the return of investment of additional public spending that account for policymaker objectives.

The validity of our results depends on the specification of the treatment assignment. To ensure the good specification of the GPS, our approach is fourfold. Firstly, a large set of pre-treatment variables are collected, which are expected to relate to housing affordability. These variables encompass both demand-side (e.g. population density, income, socio-economic status) and supply-side (e.g. housing construction, past housing prices, past neighbourhood prices). Secondly, unobserved variables may still violate the assumption of unconfoundedness (Lewbel and Yang, 2016). We account for spatial confounding variables (Gilbert et al., 2024) by introducing smoothing functions of the spatial coordinates of the location of the municipalities in a semiparametric Generalised Additive Model (herafter GAM, Wood, 2017). For instance, geographical constraints (Saiz, 2010), housing supply elasticity (Accetturo et al., 2021), or demand for particular amenities (Bayer et al., 2016) might vary smoothly in space contrary to the discontinuities introduced by the ABC zoning. Thirdly, these variables are included in the outcome regression, thus providing a doubly robust estimation (for applications of the doubly robust estimator see e.g. Brei et al., 2023; Deng et al., 2023; Hoang-Duc et al., 2024). The consistency of policy-relevant treatment effects can be ensured if either the first or second step of the identification strategy is well-defined (Robins and Rotnitzky, 1995; Słoczyński and Wooldridge, 2018). Finally, our main findings are also consistent with alternative results obtained from a quasi-experimental approach. Leveraging the 2018-reform using difference-in-difference approach, we retrieve similar results that some obtained through our selection-on-observables approach. Nevertheless, our results are more complete by disentangling policy effects according to the source of variation (ceiling values, covering share and mortgage conditions).

Our contribution to the literature is threefold. First, we estimate the policy effects of subsidized loan on both extensive and intensive margins that overcome context-dependent issues. We cannot rule out that raising public spending has no effect at the extensive margin, whereas it distorts recipients' housing choices and induces price capitalisation on the local housing market. Second, we disentangle the impact of each treatment variable on policy effects. The price capitalisation is mainly caused by an increase in the covering share, while recipients housing choices are sensitive to the ceiling value.[5] Third, accounting for policy effects and public cost, we estimate return on investment for one additional public spending. Increasing the public support for homeownership is not positive for reasonable valuation of extensive and intensive effects, while it might be counter-productive in some situations.

The remainder of the paper is structured as follows. Section 2 presents the institutional context of the French IFL policy and the datasets we gather. In Section 3, we introduce the return on investment function accounting for the effects from the extensive and the intensive margins. Our identification restriction to tackle endogenous treatment intensity and the doubly robust estimation procedure are presented in Section 4. The Section 5 provides the empirical results from the two-step approach and Section 6 concludes.

2 Context and Data

2.1 The French IFL Policy

The IFL policy was introduced in 1995 in France to encourage first-time homeownership. Recipients benefit from a complementary loan being publicly supported with no interest to pay for a given share of their overall housing credit (the covering share) up to a maximum value (the ceiling value). The opportunity cost for public finance amounts to the interests not at charge for recipients, as the government consents a tax reduction to banks supplying such contracts. As shown in Appendix A.1 of Online Appendix (OA), the total number of recipients since the beginning of the policy is about 3.2 millions, for an overall budget cost of about 26.1 billions euros (about 8,000 euros by recipient). The policy excludes a small share of high-income households (about 10% of richer tenants according to Sotura, 2020) and specifically targets the newly built housings (although existing housings are eligible conditionally on renovation from 2016). Any IFL contract is associated with a classical loan with interests to pay, so that recipients must comply with the usual conditions to access to the credit market.[6] Modifications are made at the discretion of the policymakers on a yearly basis. Recipients cannot refinance through IFL if conditions are more or less favourable after the policy reforms. Hence, the IFL conditions are those in effect at the time of the housing purchase. Similarly, recipients lose benefit from the policy if they sell their house to achieve residential mobility as they must reimburse remaining IFL capital.

[5]The covering share reinforces the treatment intensity for all subsidised operations, while the ceiling value increase the treatment intensity for the most expensive operations.

[6]For the French market, an usual condition to access to a classical loan is that reimbursement payments cannot be higher than 1/3 of income. Consequently, this condition also restricts access to the IFL policy.

The four parameters that shape IFL intensity have remained stable over time. For each period and each group of municipalities (to which we will return below), policymakers fix s, the maximum coverage that the IFL can represent in the total loan amount $\widetilde{V}$ (ranging from 10% to 40%) and $\overline{V}$ a maximum ceiling value on the loan (ranging from 100 to 150 thousand euros for one-person household). These are the two parameters at the disposal of the government to implement the IFL policy.[7] These two variables are not expected to have the same impact on the subsidy with potential differentiated effect on extensive and intensive margins. On the one hand, increasing the covering share increases the subsidy for all policy recipients. On the other hand, increasing the ceiling value increases the subsidy for the most expensive housings, while leaving unaffected the less expensive ones.[8] The budget cost for the policy also depends on two parameters that are not chosen by the government: the interest rate r and the loan maturity m. We therefore assume that they are exogenous to the IFL policy. Considering the budget cost c of a IFL contract as the subsidy-equivalent for the recipient, Appendix B.1 of OA shows that the following equality holds:

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

The budget cost of a IFL contract weakly increases with the four previous parameters, which indicates that increasing one of them is equivalent to increase the financial support for recipients. This defines the treatment intensity of the policy both as an increase of one of the four considered parameters or an increase of the budget cost for the government. We restrict our studied period on the last three IFL reforms of the 2015–2019 period, as eligibility conditions remains unaffected and the ABC classification of municipalities (i.e. the zoning for the two parameters under policy control) does not change.

2.2 The ABC Zoning

The two policy parameters (covering share and ceiling value) depend on the location of IFL contracts, from an exhaustive and mutually exclusive classification of French municipalities based on the *tenseness* of the housing markets. The municipality is the smallest jurisdiction for France with 34,970 municipalities.[9] Whereas the administration does not provide a clear definition of housing market tenseness and fuzzy classification rules, we consider that it relates to the level of housing affordability .[10] Areas with the less afford-

[7] We do not model income cap as it only excludes a small fraction of the tenants.

[8] More precisely, it affects home purchases that are censored, i.e. home purchases with purchase price higher than the previous ceiling value (from a reform perspective).

[9] Half of municipalities have less than 457 inhabitants in 2019.

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

able housing are considered as the most tense local housing market. This ABC zoning introduces four ordered degrees of tenseness, from C the lowest level to A the highest level, with B_2 and B_1 as intermediate levels. This official zoning was updated four times since its introduction in 2003, the latest update of October 2014 is stable for the 2015–2019 period under study. The choice of municipality as the spatial unit to implement the ABC zoning is consistent with other public policies design. Indeed, most housing policies (including social housing) and land-use planning are implemented at the municipality level.

Most French municipalities are rural and belong to zone C, the lowest level of the zoning (Panel A of Table 1). Table 1 also shows that the distribution of the ABC classification is consistent with expectations, as municipalities with higher population densities and higher housing prices per living area (unitary prices hereafter) are higher in the hierarchy. Despite the correlations between the ABC hierarchy and reported pre-treatment variables, it is well recognized that similar municipalities might be assigned to different tier classification: the French administration in charge of monitoring public expense noted in 2012 the lack of transparency of the ABC zoning (Cour des Comptes, 2012). It concluded that the zoning does not depend exclusively on objective characteristics, suggesting potential subjectivity in the assignment. Most existing quasi-experimental approaches dealing with the endogeneity of the housing policy assignment rely on the arbitrariness of this zoning (Labonne and Welter-Nicol, 2015; Beaubrun-Diant and Maury, 2021; Chareyron, Ly and Trouvé-Sargison, 2021). Our identification method aims to leverage the zoning arbitrariness using a selection-on-observables approach to estimate policy effects.

2.3 Data

We aggregate three exhaustive individual[11] data sources at the municipal level ($N = 34{,}970$) and match them with demographic data. We remove municipalities on the Corsican island due to the absence of spatial contiguity[12] (360 observations) and municipalities of the *Alsace-Moselle* region (1,605 observations) as housing price data are missing for administrative reasons. We filter missing observations or data inconsistencies to obtain a final sample of 26,819 municipalities. We report descriptive statistics in Appendix A.2 for the relevant variables used in the empirical analysis. Most observations with missing variables concern median income as the French secrecy rule imposes to have at least 11 observations to provide statistical information. Removed observations mostly belong to the C-tier, which are significantly different from the B_2 municipalities on observable vari-

[11] The data sources to which we have access are either at the recipients level (IFL database), the transaction level (the transaction database) or the housing unit for first-time owners (the fiscal database). However, we cannot merge these data.

[12] The introduction of spatial smoothing functions in the specification of the propensity score assumes that the probability is spatially continuous. Yet, strong geographical constraints such as the Mediterranean Sea violate such assumption.

ables.[13] Consequently, potential selection bias on our estimation is at worse marginal as these observations have small impact due to low counterfactual power.

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

Tax data. To circumvent the problem of having only subsidized new homeowners from the IFL files, we use exhaustive tax files about French homeownership (*Fichiers Fonciers*) to determine the total number of new homeowners (subsidized or not). IFL database only contains policy recipients, whereas the objective is to increase overall homeownership rate, not only recipients. Using the temporal dimension of these administrative data, we identify first-time homeowners as defined by the IFL policy. Hence, we can disentangle an increase in the number of homeowners from an increase restricted to recipients only. We obtain for each municipality the number of first-time owners, which were eligible to the IFL, by counting the number of homeowners that were absent from the tax file in the previous two years.[14] Although the tax files and IFL files are independent databases, the total number of first-time homeowners estimated from tax files is always higher than the numbers of contracts from IFL files (except for two municipalities that are removed from our sample). Finally, we recover the total number of newly built housing over the 2010–2013 period based on the construction year reported in the tax files.

Transaction data. We use a third exhaustive individual dataset on housing transactions (DV3F) to compute, at the municipal level, the average unitary price of housings over the pre-treatment period 2010–2013. In order to mitigate border effects in the delineation of housing markets and tenseness, we also compute the averages of the unitary prices over the same pre-treatment period for neighbouring municipalities using spatial contiguity definition. These data also allow building variables related to the post-treatment out-

[13]Considering a selection model within municipalities C with binary variable as dependent one that indicate whether variables are fully filled, we find that municipalities with missing variables on median income have lower density and lower housing prices. Thus, they are likely to have a low counterfactual power.

[14]The policy defines first-time owners as individuals being renter of their primary residence for at least two years purchasing a housing unit for residence.

comes for the overall housing markets, including average housing prices, average surfaces, and average unitary prices for the three periods.

Socio-demographic data. For each municipality considered, we obtain from the *French National Institute of Statistics and Economic Studies* (INSEE) (year 2013) data prior to the ABC reform in 2014 on population density, median income, and socio-professional categories.

Sample. Our sample is composed of all French municipalities. We assign each municipality to its ABC tier and report descriptive statistics about IFL treatment intensity, mortgage conditions, and pre-treatment variables in Table 1.

The IFL treatment variables are ordered according to the ABC zoning. Indeed, independently from the period of interest, the A-tier benefit from the highest maximum of ILF, whereas the C-tier benefit from the lowest. Moreover, according to pre-treatment variable (Panel A), A-tier exhibit the highest housing price, unitary housing price, inhabitants density or number of new housing units whereas the C-tier from the lowest. Hence, the treatment intensity is not endogenously defined.

Table 1. Main variables and treatment variables for municipalities along the ABC zoning

Variable	Period	Country	*ABC Zoning Areas* A	B_1	B_2	C
			A - Pre-treatment variables			
Number of Municipalities	2013	34.970	0.724	1.535	3.828	28.883
(*thousand of units*)		100%	2.07%	4.39%	10.95%	82.59%
Housing Price	2010–2013	153.1	284.0	234.9	188.6	139.0
(*thousand euros*)		*(68.0)*	*(124.8)*	*(74.8)*	*(59.6)*	*(54.9)*
Unit. Housing Price	2010–2013	1,608.7	3,558.7	2,597.0	2,003.5	1,430.9
(*euros by squared meter*)		*(691.8)*	*(1,054.5)*	*(557.5)*	*(569.3)*	*(502.6)*
Unit. Price of Neighbors	2010–2013	1,561.0	3,654.1	2,618.6	1,975.2	1,371.6
(*euro per squared meter*)		*(724.0)*	*(1,099.1)*	*(587.2)*	*(582.8)*	*(517.1)*
Population Density	2013	1.9	26.4	6.8	3.1	0.7
(*inhabitants by hectare*)		*(8.1)*	*(38.2)*	*(9.4)*	*(4.6)*	*(1.0)*
Median Household Income	2013	20.0	25.3	24.0	22.1	19.2
(*thousand euros by year*)		*(3.4)*	*(6.0)*	*(4.8)*	*(3.7)*	*(2.6)*
Number of New Housings	2010–2013	41.9	364.5	226.8	72.4	15.9
(*number of units*)		*(226.7)*	*(918.7)*	*(648.3)*	*(151.7)*	*(28.7)*
			B - IFL Policy Values			
Maximum Ceiling Value	2015	-	150.0	135.0	110.0	100.0
(*thousand of euros*)	2016–2017	-	150.0	135.0	110.0	100.0
	2018–2019	-	150.0	135.0	110.0	100.0
Maximum Coverage Share	2015	-	26.0	26.0	21.0	18.0
(*percent*)	2016–2017	-	40.0	40.0	40.0	40.0
	2018–2019	-	40.0	40.0	20.0	20.0
Maximum IFL Amount	2015	-	39.0	35.1	23.1	18.0
(*thousand of euros*)	2016–2017	-	60.0	54.0	44.0	40.0
	2018–2019	-	60.0	54.0	22.0	20.0
Average Subsidy	2015	5.21	10.31	9.41	5.81	4.02
(*thousand euros*)	2016–2017	9.63	13.96	12.35	10.45	8.99
	2018–2019	5.03	12.44	10.89	4.64	3.98
			C - Mortgage Market Conditions			
Mortgage Maturity	2015	228	244	254	238	221
(*percent*)	2016–2017	260	268	269	265	258
	2018–2019	258	267	267	262	255
Annual Interest Rate	2015	2.51	2.46	2.45	2.52	2.52
(*number of months*)	2016–2017	1.87	1.77	1.77	1.89	1.88
	2018–2019	1.62	1.53	1.52	1.65	1.63

Notes: French municipalities are classified according to the ABC zoning in columns. Panel A reports the average and standard deviation of pre-treatment variables used to control the endogenous treatment assignment. The first three variables of panel B correspond to the IFL parameters for each period (constant between municipalities) with a Maximum IFL Amount that equals the maximum ceiling value times the Maximum Covering Share. The Average Subsidy is computed from IFL data and Equation 1. Panel C reports the average of loan maturities and interest rates, also extracted from IFL files.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

3 Empirical Framework

We consider the objective of the IFL policy from a policymaker perspective in terms of extensive and intensive margins. Policymaker objectives are proxied through the marginal valuation of respectively one additional homeowner and recipients purchase value. We detail our extensive and intensive margin definitions related to the public spending according to a return on investment function. We then present first-order response to treatment variables variation and highlight the needed statistics to compute return on investment with respect to the maximization of specified policy objectives.

3.1 Outcome variables

We develop a framework for evaluating IFL policy that derives from the objectives set by policymakers. On the one hand, the number of new homeowners (denoted N) is likely to be the most relevant variable according to the policy design. We also consider the number of IFL recipients $\widetilde{N}$, although it does not correspond to the full extensive margin, as increasing $\widetilde{N}$ without changing N cannot be considered as favouring homeownership. Indeed, the number of homeowners N estimated through fiscal files is the main variable to discuss extensive margin effects, as increasing only the number of recipients would be considered only as an opportunistic behaviour. On the other hand, the intensive margin captures the distortion of recipients' housing choices. The recipients could use the subsidy to increase their purchase price (noted $\widetilde{V}$), either by increasing the area (noted $\widetilde{S}$) or price per area (noted $\widetilde{P}$). Finally, policy-induced demand could affect the average housing price from capitalization effects, especially in the context of low supply elasticity. Hence, the increase in housing demand (that could be proxied by the number of recipients $\widetilde{N}$) might affect the entire local population through the average housing price V. We then consider that the potential capitalisation effect could in return, affect both extensive and intensive margin effects.

Beyond the extensive and intensive margin effects, variation in the IFL intensity may affect the cost for the policymaker. The cost at charge for the administration depends on the number of recipients (denoted $\widetilde{N}$) and the average cost per recipients denoted $\widetilde{C}$). Hence, the effect of treatment intensity variation on the cost for the administration is not clear-cut as it depends both on first-time owners behaviour (whether they benefit from the policy) in addition to the treatment intensity per recipients. Finally, we assume that policymakers aim to achieve their policy objectives (support homeownership development) while minimising the overall cost of the policy.

3.2 Policy objectives

Considering that policy objectives are reflected through the monetary valuation of both extensive and intensive margin effects, we define return on investment function from the perspective of the policymaker, denoted E, as

$$E = \phi N + \psi \widetilde{V} - \widetilde{N}\widetilde{C} \tag{2}$$

with ϕ and ψ representing respectively the policymaker's valuation of an additional homeowner and the distortion of recipients' housing choice. E represents the monetary valuation of the IFL policy effects, taking into account the total costs defined by the product of the number of recipients ($\widetilde{N}$) and the average cost per recipients ($\widetilde{C}$). However, we consider the case where policymakers seek to improve return on investment by varying the policy scheme. Following the definition of efficiency, we derive first order Equation 2 according to the variation for variable a that define treatment intensity (covering share, ceiling value, and mortgage conditions). The sign of e_a is the main criterion to determine whether increasing public spending by treatment variable a is positive according to the policy objectives. Indeed, if e_a is positive, the policymaker would consider the benefits to be greater than the drawbacks and public expenditure, leading to positive return on investment.

$N = N[a, V(a)]$ and $\widetilde{V} = \widetilde{V}[a, V(a)]$

$$e_a = \phi \left(\frac{\partial N}{\partial a} + \frac{\partial N}{\partial V} \cdot \frac{\partial V}{\partial a} \right) + \psi \left(\frac{\partial \widetilde{V}}{\partial a} + \frac{\partial \widetilde{V}}{\partial V} \cdot \frac{\partial V}{\partial a} \right) - \left(\widetilde{N} \frac{d\widetilde{C}}{da} + \widetilde{C} \frac{d\widetilde{N}}{da} \right) \tag{3}$$

The return on investment function is composed of a direct and an indirect effect at both extensive and intensive margins. On the one hand, either the number of homeowners or the recipients' housing choices might vary according to treatment intensity variation (the direct effect). On the other hand, potential price capitalisation resulting from treatment variation might in return reduce either the likelihood to transition into homeownership or affect recipients' housing choices (the indirect effect).[15] Hence, the return on investment for additional public spending depends at both margin on the direct (treatment response, noted $\frac{\partial \cdot}{\partial a}$) and the indirect (adaptation to potential price capitalisation, noted $\frac{\partial \cdot}{\partial V} \cdot \frac{\partial V}{\partial a}$) effects. Finally, we consider that the policy cost only depends on variation in average cost per recipients and the number of recipients, which motivate the simple derivate framework.

We distinguish two sets of values from Equation 3. On the one hand, it requires to observe the monetary values set by policymakers that reflects policy objectives. Although some

[15] For instance, see Waxman et al. (2020) for housing choices responses to housing affordability issues.

papers address the valuation of externalities associated with homeownership (Coulson and Li, 2013), it is not clear how policymakers value intensive margin effects relative to extensive margin ones. We then discuss the return on investment of increasing public spending through IFL subsidies, following scenarios that differ according to valuation of extensive and intensive margins effects. In reasonable scenarios, additional homeowner is positively valued as it is the policy stated objective. The valuation of modification of recipients' housing choices is less straightforward, leading us to introduce more variation in the credible valuation range. On the other hand, first-order response to treatment variable variation for main outcomes such as number of homeowners, housing market price or recipients' purchase price are needed. It composes our building blocks to estimate the return on investment of raising IFL subsidy through each treatment variable. Moreover, we must estimate similar quantity for average spending per recipient and number of recipients are needed to assess the impact on policy cost. To estimate these first-order responses, we leverage differences in treatment intensity following the ABC zoning to obtain dose-response functions.

3.3 Dose-response Functions

We recover the marginal effects at both extensive and intensive margins for the IFL policy from the counterfactual framework (Rubin, 1974), through dose-response functions relating policy-relevant treatment effects to the four outcomes variables of interest. Variations of the IFL policy across the four ABC zones and the three periods define a multi-valued treatment taking $G = 12$ levels. Let g denote a level of treatment and T_g a dummy variable that indicates whether the municipality receives this level. Then, we have:

$$Y = \sum_{g=1}^{G} T_g Y_g, \tag{4}$$

where Y is the observed outcome, equals to its potential value Y_g only if a municipality receives treatment g. The main outcomes of interest are $Y = N$ for the extensive margin, $Y = V$ for the intensive margin, and $Y = \widetilde{C}$ for budget costs, while $Y = \widetilde{V}$ informs about unintended effects on consumption. Each bilateral combination of different treatment levels g and g' corresponds to a variation of at least one policy treatment variable. We exploit this structure of the IFL policy to map policy-relevant treatment effects to treatment variable variations. Considering the requirement of first-order derivates to estimate Equation 3, we retain a set of linear dose-response functions for each outcome Y with:

$$\mathbb{E}(Y_g - Y_{g'}) = \beta_0^Y + \sum\nolimits_a \beta_a^Y (a_g - a_{g'}) + \xi. \tag{5}$$

with mean-independent errors ξ. The Ordinary Least Squares (OLS) coefficients β_a^Y provide a summary of the effects of treatment variables a on the heterogeneity of treatment effects $\mathbb{E}(Y_g - Y_{g'})$ and allow recovering the derivatives of efficiency measures from Equation 3. For outcomes concerning the whole population, $Y \in \{N, V, P, S\}$, the average treatment effects (ATEs) are clearly policy relevant as they appear in the left-hand side of Equation 5. In effect, ATEs represent the change of Y caused by the policy g relatively to g' for the whole population and β_a^Y summarizes how these changes can be attributed to differences between a_g and $a_{g'}$. For outcomes affecting only recipients, $Y \in \{\widetilde{N}, \widetilde{V}, \widetilde{P}, \widetilde{S}, \widetilde{C}\}$, the policy-relevant treatment effects concern recipients (ATT). The left-hand side of the dose-response function (Equation 5) is then $\mathbb{E}(Y_g - Y_{g'} \mid T = g)$. As we study bilateral combinations *within* three distinct periods, this gives $4 \times (4-1) \times 3 = 36$ policy-relevant treatment effects. Therefore, each set of dose-response functions is estimated based on 36 observations for each of the nine outcomes. This allows us to recover marginal effect at both margins and estimate the return on investment for additional public spending. Yet, to estimate dose-response functions, we need to first compute the left-hand side of the Equation 5, which represents outcome counterfactual differences when switching treatment intensity from one level to another (e.g. A to B_1).

4 Identification strategy

4.1 Identifying Assumptions

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

Assumption 1 *Weak Unconfoundedness.*

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

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

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

Assumption 2 *Overlap*

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

Under the two previous assumptions, Słoczyński and Wooldridge (2018, Lemma 3.2) show that counter-factual treatment effects can be identified from the data at hand. The average outcome $Y_{g'}$ for a counter-factual treatment level g', respectively for the whole population and for municipalities that actually receive the treatment level g, are respectively:

$$\mathbb{E}(Y_{g'}) = \mathbb{E}\left[\frac{T_{g'}}{p_{g'}(\mathbf{X})} Y\right] \quad \text{and} \quad \mathbb{E}(Y_{g'} \,|\, T = g) = \frac{1}{\mathbb{P}(T = g)} \cdot \mathbb{E}\left[\frac{p_g(\mathbf{X})}{p_{g'}(\mathbf{X})} T_{g'}\, Y\right]. \tag{6}$$

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

4.2 Specification of the Propensity Score

In accordance with the notion of housing market tenseness, we define a unobserved latent variable η_i^* crossing thresholds to model the classification of municipalities. The propensity for a municipality i to be high in the hierarchy depends on the J pre-treatment variables x_{ji} used to proxy the political decision, a bivariate smooth function of the geographical coordinates of its centroid $\mathbf{z}_i$ (longitude and latitude, Gilbert et al., 2024), and a random term ε_i representing the arbitrary part of the zoning explained in the Section 2.2. This latter term is assumed to be logistically distributed to produce an ordered logit model. The latent variable describing the tightness of the housing market η_i^* is then:

$$\eta_i^* = \alpha + \sum_{j=1}^{J} f_j(x_{ji}) + h(\mathbf{z}_i) + \varepsilon_i. \tag{7}$$

The J univariate functions f_j are specified as additive spline transformations of pre-treatment variables, in accordance with the generalized additive model framework (GAM, Wood, 2017). The spline coefficients are shrunk endogenously by penalized iterated weighted least squares while the smoothing parameters are estimated using a separate criterion from the restricted maximum likelihood (Wood, Pya and Säfken, 2016). The same estimation procedure is used simultaneously for the bivariate smooth function h of coordinates, the main difference is the *a priori* specification of the spline that is bivariate thin plate. Whereas the geographic regression discontinuity restricts the sample to observations close to the boundary to assume that the treatment is quasi-random between treatment and control groups, our approach explicitly models the spatial contribution in the treatment assignment to control for proximity and recover the arbitrary part of treatment assignment, making the best of the full sample.

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

$$p_g(\eta_i) = \Lambda(\mu_g - \eta_i) - \Lambda(\mu_{g-1} - \eta_i). \tag{8}$$

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

4.3 Specification of the Outcomes

The outcomes are specified using the same semi-parametric GAM framework. Recall that we introduce as control variables in the outcome specification pre-treatment variables introduced in the GPS specification to provide a doubly robust estimator, which reinforces the credibility of our results. The main difference is that each outcome Y is modelled separately for each subsample defined from the treatments g received by the municipalities. The smooth functions f_j and h are now indexed by the outcome y and the treatment g such that:

$$y_{gi} = \alpha_g^y + \sum_{j=1}^{J} f_{gj}^y(x_{ji}) + h_g^y(\mathbf{z}_i) + \varepsilon_{gi}^y. \tag{9}$$

The same pre-treatment variables and geographical coordinates are used, with different smoothing parameters. As we have nine outcomes, four treatment levels and three periods, Equation 7 corresponds to 108 GAM estimations in order to estimate the full set of functions f_{gj}^y and h_g^y for a given GPS. From the quasi-loglikelihood arguments of Słoczyński and Wooldridge (2018), the double robustness property requires that outcome regressions are weighted according to GPS ratios as in Equation 6. To recover the average counter-factual outcome for the treatment g' for the municipalities actually receiving g, each municipality is weighted by $\widehat{p}_g(\eta_i)/\widehat{p}_{g'}(\eta_i)$ predicted from the first stage. As generally advised in the literature, we use normalized weights by dividing them by their sum within each treatment subsample.

We close this section with the formulas that we use to assess the return on investment of the IFL policy. The counter-factual building blocks of Equation 6 are recovered from the regression of the outcome Y on the sub-sample of municipalities with treatment g' using respectively $1/p_{g'}(\eta_i)$ and $p_g(\eta_i)/p_{g'}(\eta_i)$ as weights. Under assumptions 1 and 2, noting $\mu_g \equiv \mathbb{P}(T = g)$ the share of municipalities that receive the treatment g, averaging the fitted values provides a consistent estimations as:

$$\mathbb{E}(Y_{g'}) = N^{-1} \times \sum_{\ell=1}^{N} \hat{y}_\ell(g') \quad \text{and} \quad \mathbb{E}(Y_{g'} \,|\, T = g) = \mu_g^{-1} \times \sum_{\ell=1}^{N} T_{g\ell} \times \hat{y}_\ell(g') \tag{10}$$

where $\hat{y}_\ell(g') \equiv \hat{\alpha}_{g'}^Y + \sum_{j=1}^J \hat{f}_{g'j}^Y(x_{j\ell}) + \hat{h}_{g'}^Y(z_\ell)$ comes from the estimation of the outcome Y for the subset of municipalities that receive treatment g'. It is the predicted outcome values for the whole population of municipalities with $\ell = 1, \dots, N$, that will be used as outcomes in the Equation 5.

5 Results

We first present the estimation results for the first-stage models, followed by second-stage models, the estimation of policy-relevant treatment effects and dose-response functions. We close this section with our measures of return on investments resulting from IFL subsidy variations according to each treatment variable, i.e. variables that affect treatment intensity such as the ceiling value, the covering share and the mortgage conditions.

5.1 First-stage Models from ABC Zoning

We include the geographical coordinates of the centroids of each municipality through bivariate smoothing splines to control for spatial confounders (Gilbert et al., 2024). As allowed by the GAM framework, all variables enter semi-parametrically with a degree of smoothing that is endogenously shrunk by the penalised estimation procedure. Table 2 provides the joint significance of the spline transformations of each variable according to different specifications and maximum degree of spatial smoothing, while detailed results on GPS estimation are reported from Appendix C.1 to Appendix C.4.

Table 2. Covariates' joint significance from first-stage ordered GAMs

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

Notes: The top panel reports χ^2 statistics of joint significance for each covariate of the first-stage GPS. Professional Occupations are coded as population shares of eight categories according to the one-digit French *Catégories Socio-Professionelles.* The effective degrees of freedom reported in brackets indicate the smoothing intensity, low values correspond to more smoothing. The unit of observation is the French municipality, columns reports different specifications with different covariates and different maximum spatial smoothing. Estimations come from the `gam` function of the `mgcv` R package (Wood, Pya and Säfken, 2016).
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.
*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

The specification with the lowest spatial smoothing (that allows the consideration of spatial heterogeneity on a fine scale, reported in the last column) yields 90.1% of correct predictions of the ABC classification for municipalities. This indicates the relevance of the ordered framework for modelling the ABC classification and increases the likelihood of having a well-specified GPS. Although models with higher maximum degrees of freedom allowed for spatial coordinates yield better predictions (91.2% for df=400), their computational cost and the issues of dimensionality for our smallest sample lead us to prefer more parsimonious specification. Nevertheless, our main results regarding the effects of IFL policies are robust to the specification of the maximum degrees of freedom for spatial coordinates, albeit increasing the maximum degree of freedom for spatial smoothing significantly reduces standard errors. In particular, the introduction of spatial coordinates affects the joint significance associated with the unitary housing price being consistent with the local characteristics of the housing market. The joint significance is the highest of the pre-treatment variables, confirming our expectation about the presence of unobservable spatial variables. In our preferred specification ($df = 200$), the contribution of the unitary housing price is linear and increasing, consistent with the ABC perimeter definition.

Although prediction errors are limited, they concern municipalities commonly used for geographic regression discontinuity design, as 77.2% of errors have at least one neighbouring municipalities being classified in another treatment level. In line with the regression discontinuity assumptions, for these municipalities the treatment assignment is quasi-random, which increase the likelihood of wrong predictions for these municipalities. In addition, our prediction errors concerns more importantly municipalities that change of ABC-tier during the 2014 reform. Similarly, we might expect the treatment assignment to be less predictable.

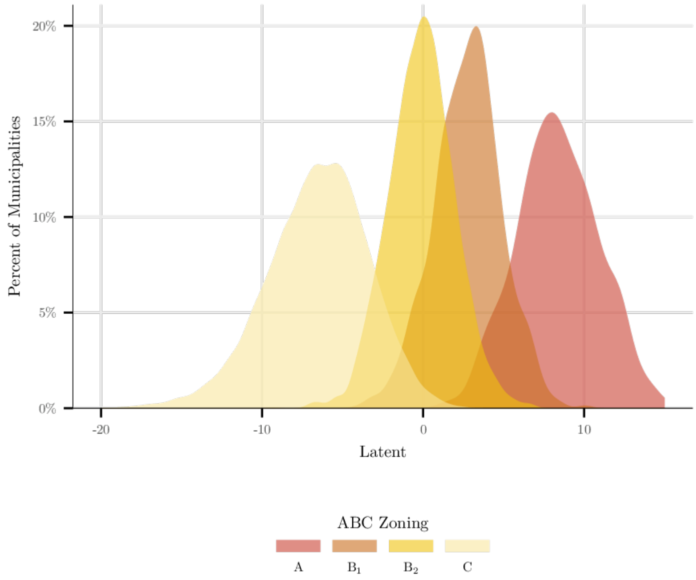
Since overlap is crucial to recover consistent effects and likely to be reduced for high-dimension model variables (D'Amour et al., 2021), we compare the distribution of the latent variable underlying the classification process (Fig. 1). Latent distributions follow the ordered structure of the ABC classification, as consecutive treatment levels have greater common support than non-consecutive ones. Although the overlap is reduced, there is still common support for extreme levels. This is probably due to the spatial proximity of some A- and C-tier municipalities. However, although treatment assignment is based on the characteristics of the municipality, it still contains some arbitrariness, which we exploit for identification.

5.2 Second-stage Models for the Outcomes

We now assess the relevance of our control variables in the outcome specification using pooled models for our doubly robust estimator. Considering the large set of pre-treatment variables and our nine outcomes, we only report joint significance of each pre-treatment variables in Table 3 to assess the statistical power of pre-treatment variables. We report spatial smoothing splines functional forms in Appendix C.5.

Pre-treatment variables introduced as regression adjustment in the pooled models explain more than 74% of the observed variance in the number of first-time owners. The develop-

Figure 1. Overlap between predictions of tenseness between the different ABC zones



Notes: The distributions of the latent tenseness variable (x-axis) are predicted from the first stage GPS with a maximum degree of freedom sets to 200 (6^{th} column of Table 2). As a latent variable, $\hat{\eta}^*$ is unit-less and is displayed between municipalities according to the ABC classification. We report the distribution *within* each classification level (rather than the distribution of the all population) for clarity reasons.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

ment of housing supply as measured by the number of new housing is highly significant in explaining both the number of first-time owners and the number of recipients. In addition, the local housing market price and median income are significant in explaining the number of transitions to homeownership, highlighting the importance of affordable housing for transitions to homeownership. It supports our approach to the introduction of pre-treatment variables in the regression adjustment to control for potential heterogeneity.

5.3 Treatment Effects and Dose-Response Functions

From the second step, we estimate combinations of bilateral effects (g, g') to infer required statistics for the estimation of efficiency first-order response to treatment variable variations (Equation 3). We report these bilateral estimations and standard errors, using a bootstrap approach with 500 iterations, respectively, in Appendix C.7 for ATE and Appendix C.8 for ATT. They constitute our building blocks to estimate dose-response functions according to each treatment variable for both extensive and intensive margin outcomes.

We estimate each β_a^Y from Equation 5 by regressing the bilateral combinations of the treatment level on the differences in treatment variable values between treatment level g

Table 3. Covariates' joint significance from second-stage pooled GAMs

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

Notes: For the nine outcomes of interest (in columns), the table reports the F statistics for the joint significance of each covariates (in rows). N accounts for the number of new homeowners, V for housing value, S for surface, and P for unitary housing price. The variables with a $\sim$ are the same variables computed for IFL recipients, $\widetilde{C}$ is the IFL cost. We report pooled GAMs on all treatment levels for the sake of clarity, different GAMs are estimated for each treatment level in the policy-relevant treatment effects reported in the text. Professional Occupations are coded as population shares of eight categories according to the one-digit French *Catégories Socio-Professionelles*. The effective degrees of freedom reported in brackets indicate the smoothing intensity, low values correspond to more smoothing. The unit of observation is the French municipality and the maximum degree of freedom we allow for the spatial coordinates is 200.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.
*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

and g'. Since we consider small treatment variable variations, our dose-response functional form is linear. In addition, our weighting scheme to estimate the relevant dose-response parameters depends on the nature of the estimands. We weight bilateral combinations according to the number of municipalities that currently received treatment level g for ATT estimands. We do not introduce weights for the ATE as it concerns the entire population, unlike ATT. We report our results from linear dose-response specifications in Table 4 estimated by WLS (for ATT) and OLS (for ATE). Standard errors are estimated using bootstrap with 500 iterations.

Despite a significant effect on policy costs for the covering share (column 9, Table 4), which is the main treatment variable used for policy reforms, it has no significant effect on the number of homeowners (column 1, Table 4). Increasing the covering share is unlikely to achieve the main policy objective, as the number of homeowners does not increase significantly with this variable. Meanwhile, increasing the ceiling value (the second treatment variable that policymakers can control) has no significant effect on the number of homeowners, while it increases the number of policy recipients (+9.1%). Given the

Table 4. OLS coefficients for policy treatment variables from dose-response functions

	Outcome variables from ...								
	Tax	Transaction Data			IFL files				
	N	V	S	P	$\widetilde{N}$	$\widetilde{V}$	$\widetilde{S}$	$\widetilde{P}$	$\widetilde{C}$
Covering Share	-0.007	0.032**	0.004	0.007	-0.033	0.001	0.010	-0.000	0.041***
	(0.025)	(0.015)	(0.006)	(0.015)	(0.046)	(0.007)	(0.011)	(0.011)	(0.011)
Ceiling Value	-0.024	-0.003	-0.006	-0.004	0.091**	-0.013	-0.047***	0.026***	-0.006
	(0.026)	(0.015)	(0.005)	(0.014)	(0.035)	(0.009)	(0.010)	(0.008)	(0.010)
Interest Rate	-0.004	0.001	-0.001	-0.003	0.021**	-0.005***	-0.009***	0.004	-0.001
	(0.007)	(0.004)	(0.001)	(0.004)	(0.010)	(0.002)	(0.003)	(0.002)	(0.003)
Loan Maturity	0.009	-0.002	-0.001	-0.004	-0.025	0.004	0.025***	-0.016***	0.021***
	(0.012)	(0.011)	(0.004)	(0.009)	(0.021)	(0.003)	(0.005)	(0.004)	(0.006)
Price	0.003					-0.001			
	(0.005)					(0.003)			
Constant	-0.038	-0.000	-0.000	-0.000	-0.130	0.009	-0.039	0.035*	-0.014
	(0.052)	(0.000)	(0.000)	(0.000)	(0.077)	(0.021)	(0.023)	(0.020)	(0.023)
R^2	0.162	0.569	0.505	0.649	0.522	0.334	0.456	0.434	0.906
Adj. R^2	0.022	0.513	0.441	0.603	0.460	0.223	0.386	0.361	0.894
N	36	36	36	36	36	36	36	36	36

Notes: For the nine outcomes Y in columns, the table reports the β_a^Y coefficients associated to each treatment variable in rows. They are estimated from dose-response functions of Equation 5. N accounts for the number of new homeowners, V for housing value, S for surface, and P for unitary housing price. The variables with a $\widetilde{}$ are the same variables computed for IFL recipients, $\widetilde{C}$ is the IFL cost per recipient. The interest rates is expressed in hundredth of percent. The unit of observation is the bilateral combination of four ABC zones for the three periods of interest, the full set of policy-relevant treatment effects is reported in the Appendix C.7 and Appendix C.8 of OA. Standards errors in parenthesis are estimated using bootstrap with 500 iterations accounting for the uncertainty of treatment effects. ATEs for tax and transaction data are weighted according to the inverse of their bootstrapped standard errors, ATT for IFL variables are additionally weighted according to the number of municipalities receiving the considered treatment levels.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.
*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

joint effect of the ceiling on the number of homeowners and recipients, raising the ceiling value is more likely to induce opportunistic behaviour than to cause homeownership. In addition, it should be noted that the impact of IFL on tenure decision is also likely to be independent from the characteristics of the credit market, while opportunistic behaviour is more likely to be favoured in a context where credit conditions are less favourable for households (higher interest rates).

Whereas the covering share has at best a weak effect on the number of homeowners, it increases housing price for the entire local housing market. A one-unit increase in the covering share raises the price of all transactions by 3.2%, leading to a significant unintended effect on the housing market. The variation in subsidy resulting from the covering share might increase housing demand whereas it does not distort recipients' housing choices. In contrast, raising the ceiling value has no significant effect on the local housing price, while the effect on recipients' housing choices is more contrasted. Indeed, while it has a negative impact on housing size (columns 6 and 7, Table 4), the price per area unit covaries positively with the ceiling value (column 8). As raising the ceiling value increase treatment intensity only for the more expensive operations (for more details, see Appendix B), the distortion of housing choices might correspond either to an increase in location quality, construction quality or both. Most of the effects for the treatment variable variations controlled by policymakers concern the intensive margin, with differentiated effects on recipients' housing choices. Similarly, recipients housing choices are sensitive to mortgage conditions. On the one hand, recipients decrease housing size when interest rates increase. We assume that the increase in treatment intensity due to interest rates is outweighed by the additional interests they must pay (recall that the IFL is a complementary loan that cannot overcome 40% of the overall borrowed amount). On the other hand, longer loan maturity increase the housing unit recipients purchase, whereas it decreases the average price per square metres decrease.

These results are consistent with alternative results from natural experiments. Using difference-in-differences (refer to Appendix E for the alternative results), we focus on the 2018 reform, which change the coverage share for two ABC tiers, while leaving this policy unchanged for two others. The findings imply that reducing the coverage share has a significant effect on the cost of the policy but no significant effect on the number of homeowners. Our study extends these results by accounting for the treatment variable source of variation.

5.4 Simulation According to Policy Objectives

We finally provide counterfactual simulations that define return on investment for increasing public spending according to each treatment variable. For comparison purpose, we choose to simulate treatment variable variations (ceiling value, covering share and mortgage conditions) that have a similar impact on the total cost of the policy and provide normalised measures for one additional euro of public spending. The statistics needed to estimate Equation 3 are obtained from the dose-response functions (Table 4). Note that the effect of the treatment variable variation a on the total cost of the policy aggregates the effect on the cost per recipient ($\partial\widetilde{C}/\partial a$) and the effect on the number of recipients ($\partial\widetilde{N}/\partial a$). Then, despite a non-significant effect on the cost per recipient of raising the ceiling value (Table 4), it has a positive effect on the total cost of the policy, given the strong effect on the number of recipients. We report the total cost effect by decomposing the price per recipient and the number of recipients effects in Appendix C.16. It is noteworthy that the effect on total cost is positive regardless of the treatment variable, although it is not statistically significant for some treatment variables.

As we do not observe the marginal valuation of both extensive and intensive margin effects from a policymaker perspective, we make different simulations to assess in which conditions increasing the subsidy provides positive return on investment. Yet, we only consider cases where increasing the number of homeowners is positively valued, as this is the main objective in order to generate additional homeowners. For the intensive margin valuation, the valuation is not straightforward from a policymaker perspective. On the one hand, policymakers might aim increasing the recipients' utility derived from more expensive housing purchase. On the other hand, considering raising issues about land consumption especially in France, policymakers might aim to level down recipients' housing choices, especially through the housing size. Hence, we assess policy efficiency in both situations (positively and negatively valued), while we add a situation with a zero valuation of the intensive margin effect (denoted as the indifferent situation). We report the results in Table 5.

Table 5. Return on Investment Measures According to Different Valuation of Extensive and Intensive Margin Effects from a Policymaker Perspective

	$\phi = 5$k				$\phi = 1$k			
ψ	-2k	-1k	0k	1k	-2k	-1k	0k	1k
Covering Share	-0.62	2.23	5.08	7.93	-5.49	-2.64	0.22	3.07
	(-26.1)	(-22.0)	(-20.7)	(-22.8)	(-17.2)	(-9.2)	(-4.1)	(-9.6)
Ceiling Value	-1.16	-2.33	-3.50	-4.67	0.84	-0.33	-1.50***	-2.67*
	(4.2)	(3.2)	(2.7)	(3.1)	(3.2)	(1.7)	(0.5)	(1.6)
Interest Rate	0.29	-0.72	-1.74	-2.75*	0.88	-0.13	-1.15***	-2.16***
	(2.1)	(1.7)	(1.5)	(1.6)	(1.3)	(0.7)	(0.3)	(0.7)
Maturity	-0.95	-0.48	-0.02	0.44	-1.73	-1.27*	-0.80***	-0.34
	(1.9)	(1.5)	(1.3)	(1.4)	(1.4)	(0.8)	(0.3)	(0.7)

Notes: Exploiting coefficients derived from the dose-response function (Table 4), we calculate return on investment for a cost-normalised increase of the overall IFL budget using Equation 3 for the four sources of treatment variables. Our results can be interpreted as the monetary benefits from a 1 euro increase of the IFL budget cost according from a policymaker perspective. As our return on investment measure depends on the marginal valuation of extensive margin effects (ϕ), price capitalisation (ψ) and distortion of housing choices, we simulate different scenarios. The top panel (resp. bottom panel) corresponds to the situation in which an additional homeowner is valued at 10k (resp. 1k). Although price capitalisation is likely to be negatively valued, we also assess scenarios with opposite sign according to our expectations. We report between parentheses the standard errors using 500-iterations bootstrap procedure.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.
*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

Regardless of the valuation of extensive and intensive margins in credible scenarios, return on investment of additional spending is not positive (Table 5). The return of investment is indeed at best non significant, while it might be negative in some situations, that correspond to a counter-productive effect of public spending. When valuing additional homeowners to 1k, one additional euro through raising the ceiling value induces a 2.7 euros loss when intensive margin effects are also sought by policymakers ($\phi = 1$k). The loss is however less important when policymakers are indifferent to the intensive margin effects ($\phi = 0$k). Remark that we find similar results when the additional public spending is caused by interest rates variation.

In conclusion, raising the subsidy regardless the channel is at best inefficient, while it might be counter-productive in specific situations as the return on investment is negative.

The policy mostly modifies recipients housing choices, favours opportunistic behaviour, while it does not strengthen homeownership development.

6 Conclusion

The French IFL policy aims to increase the number of new homeowners through interest cuts for eligible households. We leverage spatial variation of treatment using selection-on-observables approach to assess the effect of subsidy variation on policy objectives that relates either to extensive margin or intensive margin. Our GPS specification and regression adjustment involve, among other variables, spatial coordinates to prevent for omitted spatial variables. From the linear dose-response functions, we discuss the policy results for different simulations that differ according to the policymaker valuation of extensive and intensive margins effects to derive conclusions about policy return on investment.

We cannot reject the possibility that increasing policy expenditure on the IFL has no effect on the number of homeowners. Indeed, based on our identification strategy, we cannot exclude that increasing both policy control treatment variables affects tenure decisions at the individual level. However, we precisely estimate that the intensive margin effects exceed potential ones at the extensive margin despite the fact that it is the latter that is being targeted by policymaker. In addition, increasing the IFL subsidy causes demand to shift from existing to new housing, resulting from opportunistic behaviour. It turns out that the relevance of the IFL mainly depends on the valuation of intensive margin effects. In credible situations with positive valuation of the extensive margin, the return of investment from policymaker perspective is negative, or at best null. Hence, regardless the treatment variable under consideration, increasing public spendings to support homeownership through the IFL appears to be inefficient.

Our paper leaves open questions for further research on the assessment of public support to homeownership. As housing market capitalisation is related to housing supply and land availability, externalities produced by interest cuts are likely to depend on local characteristics. Assessing the structure of such heterogeneity is crucial for more precisely assessing the IFL policy according to the areas policymakers aim to favour. Finally, as supporting homeownership affects recipients housing choices, it raises concerns about the impact of interest cuts on land consumption and policy interaction with land constraints for housing development.

References

Accetturo, A, A. R Lamorgese, S Mocetti and D Pellegrino (2021). 'Housing Supply Elasticity and Growth: Evidence from Italian Cities'. In: *Journal of Economic Geography* 21.3, pp. 367–396.

Barone, Gugliemo, Francesco David, Guido De Blasio and Sauro Mocetti (2021). 'How Do House Prices Respond to Mortgage Supply?' In: *Journal of Economic Geography* 21.1, pp. 127–140.

Bayer, Patrick, Robert McMillan, Alvin Murphy and Christopher Timmins (2016). 'A Dynamic Model of Demand for Houses and Neighborhoods'. In: *Econometrica* 84.3, pp. 893–942.

Beaubrun-Diant, Kevin and Tristan-Pierre Maury (2021). 'Implications of Homeownership Policies on Land Prices: The Case of a French Experiment'. In: *Economics Bulletin* 41.3, pp. 1256–1265.

Brei, Michael, Leonardo Gambacorta, Marcella Lucchetta and Bruno Maria Parigi (2023). 'How Effective Are Bad Bank Resolutions? New Evidence from Europe'. In: *Journal of Financial Stability* 67, p. 101153.

Carozzi, Felipe, Christian A. L. Hilber and Xiaolun Yu (2024). 'On the Economic Impacts of Mortgage Credit Expansion Policies: Evidence from Help to Buy'. In: *Journal of Urban Economics* 139, p. 103611.

Chareyron, Sylvain, Tidiane Ly and Yohann Trouvé-Sargison (2021). 'Ownership Incentives and Housing Affordability: Evidence from France'. In: *SSRN Electronic Journal*.

Coulson, N. Edward and Herman Li (2013). 'Measuring the External Benefits of Homeownership'. In: *Journal of Urban Economics* 77, pp. 57–67.

Cour des Comptes (2012). *Rapport Public Annuel 2012, Tome I*. Tech. rep. Cour des Comptes, p. 1033.

Crump, Richard K., V. Joseph Hotz, Guido W. Imbens and Oscar A. Mitnik (2009). 'Dealing with Limited Overlap in Estimation of Average Treatment Effects'. In: *Biometrika* 96.1, pp. 187–199.

Cunningham, Chris and Robert R. Reed (2013). 'Negative Equity and Wages'. In: *Regional Science and Urban Economics* 43.6, pp. 841–849.

D'Amour, Alexander, Peng Ding, Avi Feller, Lihua Lei and Jasjeet Sekhon (2021). 'Overlap in Observational Studies with High-Dimensional Covariates'. In: *Journal of Econometrics* 221.2, pp. 644–654.

Deng, Xiangou, Zhaohui Li, Michael J. Seiler and Hua Sun (2023). 'Market Strength and Brokerage Choice in Residential Housing'. In: *The Journal of Real Estate Finance and Economics*.

Gilbert, Brian, Abhirup Datta, Joan A. Casey and Elizabeth L. Ogburn (2024). *A Causal Inference Framework for Spatial Confounding*.

Gobillon, Laurent and David Le Blanc (2008). 'Economic Effects of Upfront Subsidies to Ownership: The Case of the Prêt à Taux Zéro in France'. In: *Journal of Housing Economics* 17.1, pp. 1–33.

Green, Richard K. and Patric H. Hendershott (2001). 'Home-Ownership and Unemployment in the US'. In: *Urban Studies* 38.9, pp. 1509–1520.

Green, Richard K. and Michelle J. White (1997). 'Measuring the Benefits of Homeowning: Effects on Children'. In: *Journal of Urban Economics* 41.3, pp. 441–461.

Hahn, Jinyong (1998). 'On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects'. In: *Econometrica* 66.2, pp. 315–331.

Hanson, Andrew (2012). 'Size of Home, Homeownership, and the Mortgage Interest Deduction'. In: *Journal of Housing Economics* 21.3, pp. 195–210.

Harding, John, Thomas J. Miceli and C.F. Sirmans (2000). 'Do Owners Take Better Care of Their Housing Than Renters?' In: *Real Estate Economics* 28.4, pp. 663–681.

Harding, John P. and Stuart S. Rosenthal (2017). 'Homeownership, Housing Capital Gains and Self-Employment'. In: *Journal of Urban Economics* 99, pp. 120–135.

Harkness, Joseph M and Sandra J Newman (2003). 'Effects of Homeownership on Children: The Role of Neighborhood Characteristics and Family Income'. In: *Economic Policy Review* 9.2, pp. 87–107.

Hembre, Erik (2018). 'An Examination of the First-Time Homebuyer Tax Credit'. In: *Regional Science and Urban Economics* 73, pp. 196–216.

Hoang-Duc, Chinh, Hang Nguyen-Thu, Tuan Nguyen-Anh, Hiep Tran-Duc, Linh Nguyen-Thi-Thuy, Phuong Do-Hoang, Nguyen To-The, Vuong Vu-Tien and Huong Nguyen-Thi-Lan (2024). 'Governmental Support and Multidimensional Poverty Alleviation: Efficiency Assessment in Rural Areas of Vietnam'. In: *The Journal of Economic Inequality*.

Imbens, Guido W. (2000). 'The Role of the Propensity Score in Estimating Dose-Response Functions'. In: *Biometrika* 87.3, pp. 706–710.

Labonne, Claire and Cécile Welter-Nicol (2015). 'Cheap Credit, Unaffordable Houses?' In: *Working Paper*.

Lewbel, Arthur and Thomas Tao Yang (2016). 'Identifying the Average Treatment Effect in Ordered Treatment Models without Unconfoundedness'. In: *Journal of Econometrics* 195.1, pp. 1–22.

Martin, Hal and Andrew Hanson (2016). 'Metropolitan Area Home Prices and the Mortgage Interest Deduction: Estimates and Simulations from Policy Change'. In: *Regional Science and Urban Economics* 59, pp. 12–23.

Robins, James M. and Andrea Rotnitzky (1995). 'Semiparametric Efficiency in Multivariate Regression Models with Missing Data'. In: *Journal of the American Statistical Association* 90.429, pp. 122–129.

Rubin, Donald B. (1974). 'Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.' In: *Journal of Educational Psychology* 66.5, pp. 688–701.

Saiz, Albert (2010). 'The Geographic Determinants of Housing Supply'. In: *Quarterly Journal of Economics* 125.3, pp. 1253–1296.

Słoczyński, Tymon and Jeffrey M. Wooldridge (2018). 'A General Double Robustness Result for Estimating Average Treatment Effects'. In: *Econometric Theory* 34.1, pp. 112–133.

Sodini, Paolo, Stijn Van Nieuwerburgh, Roine Vestman and Ulf von Lilienfeld-Toal (2023). 'Identifying the Benefits from Homeownership: A Swedish Experiment'. In: *American Economic Review* 113.12, pp. 3173–3212.

Sommer, Kamila and Paul Sullivan (2018). 'Implications of US Tax Policy for House Prices, Rents, and Homeownership'. In: *American Economic Review* 108.2, pp. 241–274.

Sotura, Aurélie (2020). 'Housing Policies, Spatial Inequalities and Housing Markets'. PhD thesis. Paris: EHESS.

Valentin, Maxence (2023). 'Subsidizing Housing with Deductions'. In: *Journal of Economic Surveys* 38.4, pp. 1490–1515.

Waxman, Andrew, Yuanning Liang, Shanjun Li, Panle Jia Barwick and Meng Zhao (2020). 'Tightening Belts to Buy a Home: Consumption Responses to Rising Housing Prices in Urban China'. In: *Journal of Urban Economics* 115.103190.

Wood, Simon N. (2017). *Generalized Additive Models: An Introduction with R*. 2nd ed. Chapman and Hall/CRC.

Wood, Simon N., Natalya Pya and Benjamin Säfken (2016). 'Smoothing Parameter and Model Selection for General Smooth Models'. In: *Journal of the American Statistical Association* 111.516, pp. 1548–1563.

Online Appendix (not for publication)

A Additional Descriptive Statistics

A.1 IFL Summary

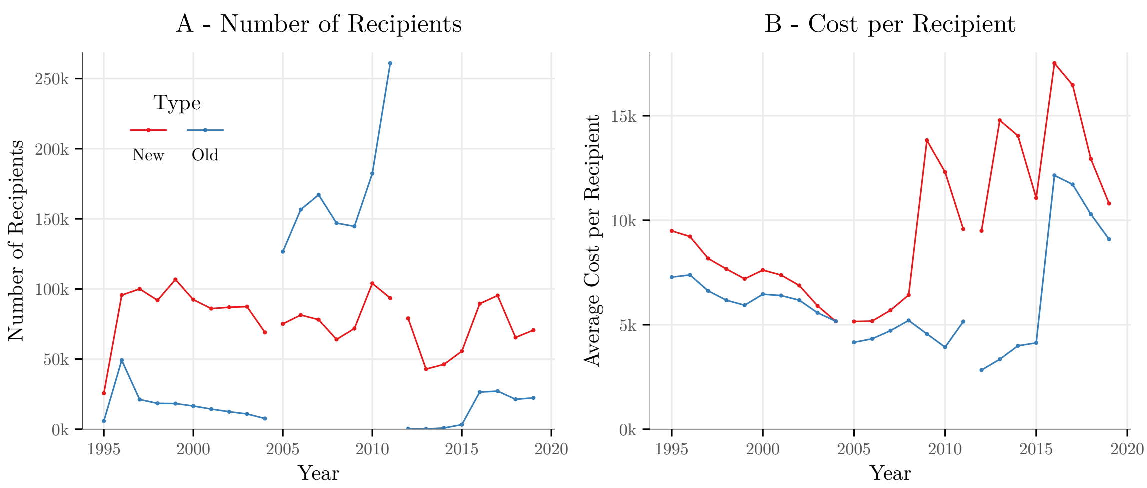


Figure A.1. Number of IFL Recipients and Average Subsidy since the Policy Introduction

Notes: We report for each year the number of households who benefit from the IFL (Panel A) and the average cost of per recipient (Panel B). We distinguish both variables according to whether it concerns existing or newly built housing.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

A.2 Municipalities Pre-Treatment Characteristics

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

Notes: The average density of the municipalities used to estimate the GPS is 193 inhabitants per kilometre square. Our sample is composed of 26,819 observations. CS1 corresponds to share of socio-professional categories within the municipality. 1 corresponds to farmers, 2 to artisans and merchants, 3 to managers, 4 to intermediate professions, 5 to employees, 6 to labour works, 7 to retired.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

A.3 Descriptive Statistics for the Outcomes

Table A.3. Descriptive Statistics for the Outcomes

	N	Mean	Std Dev	Median	Q1	Q3	Min	Max
A								
FTO	1,874	258.7	614.6	118.0	48.0	48.0	1.0	12,760.0
Price (Transaction)	1,874	443,572	426,877	320,763	261,133	261,133	107,760	5,999,507
Surface (Transaction)	1,874	80	17	78	68	68	39	177
Unit. Price (Transaction)	1,874	7,279	11,407	4,416	3,470	3,470	1,197	238,899
Recipients	1,874	46.8	102.6	17.0	5.0	5.0	1.0	2,182.0
Price (IFL)	1,874	281,426	67,118	273,749	232,633	232,633	107,000	660,000
Surface (IFL)	1,874	81	25	75	62	62	22	271
Unit. Price (IFL)	1,874	3,662	1,096	3,407	3,003	3,003	641	10,897
Cost	1,874	19,068	5,295	18,568	15,694	15,694	2,264	40,806
B_1								
FTO	3,295	122.4	348.7	48.0	24.0	24.0	1.0	8,268.0
Price (Transaction)	3,295	309,657	263,593	254,787	208,695	208,695	111,964	7,386,864
Surface (Transaction)	3,295	90	17	89	79	79	39	175
Unit. Price (Transaction)	3,295	4,174	5,056	3,095	2,536	2,536	1,442	119,469
Recipients	3,295	18.7	38.4	8.0	3.0	3.0	1.0	848.0
Price (IFL)	3,295	238,583	53,601	232,025	202,959	202,959	68,441	620,610
Surface (IFL)	3,295	95	23	95	80	80	30	280
Unit. Price (IFL)	3,295	2,595	613	2,499	2,197	2,197	615	10,753
Cost	3,295	17,445	4,734	17,189	14,408	14,408	2,422	41,496
B_2								
FTO	6,200	55.2	129.4	25.0	12.0	12.0	1.0	2,572.0
Price (Transaction)	6,200	213,863	201,699	186,010	154,125	154,125	20,000	6,297,033
Surface (Transaction)	6,200	96	16	95	86	86	38	191
Unit. Price (Transaction)	6,200	2,630	3,528	2,117	1,713	1,713	345	125,116
Recipients	6,200	8.2	11.5	4.0	2.0	2.0	1.0	127.0
Price (IFL)	6,200	207,007	41,762	200,911	179,365	179,365	60,691	479,954
Surface (IFL)	6,200	103	21	101	93	93	1	500
Unit. Price (IFL)	6,200	2,117	3,142	1,978	1,766	1,766	372	172,197
Cost	6,200	11,906	5,642	10,576	7,370	7,370	995	47,635
C								
FTO	29,463	17.5	23.1	11.0	6.0	6.0	1.0	424.0
Price (Transaction)	29,463	161,697	137,386	144,000	114,000	114,000	12,000	6,258,743
Surface (Transaction)	29,463	100	18	99	90	90	20	400
Unit. Price (Transaction)	29,463	1,849	2,728	1,534	1,229	1,229	138	169,770
Recipients	29,463	3.7	5.7	2.0	1.0	1.0	1.0	169.0
Price (IFL)	29,463	181,554	37,411	178,500	157,928	157,928	40,000	492,888
Surface (IFL)	29,463	108	22	104	95	95	1	700
Unit. Price (IFL)	29,463	1,769	2,769	1,706	1,504	1,504	165	243,577
Cost	29,463	9,851	5,117	8,489	5,729	5,729	322	41,061

A.4 ABC Perimeter

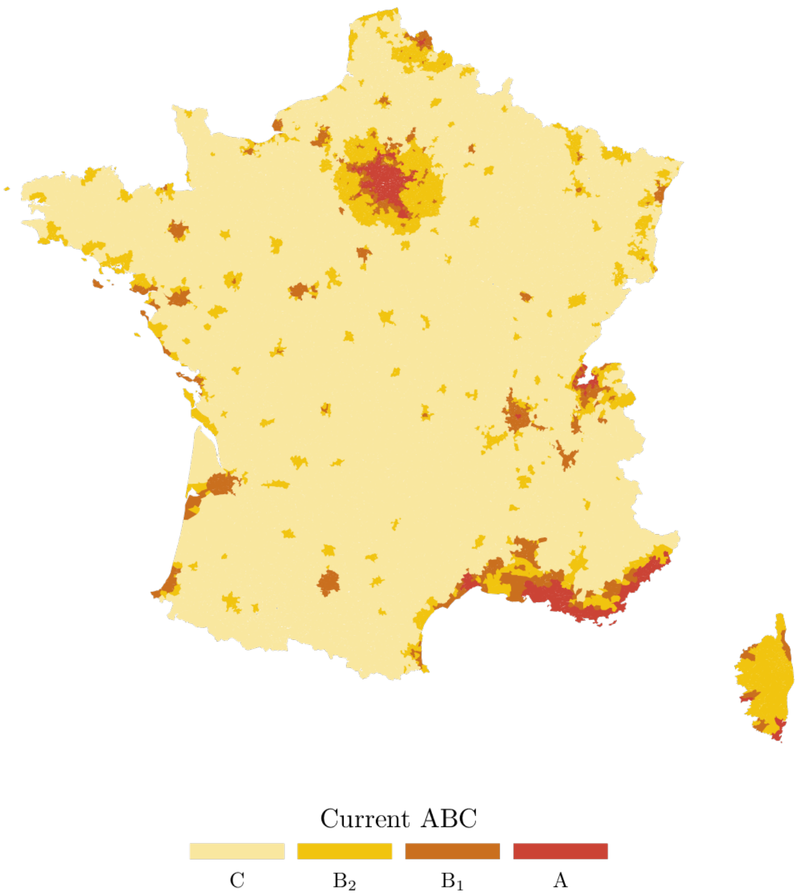


Figure A.4. Current ABC Zoning

B IFL Design

B.1 Cost Calculation

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

$$X_t = X_{t-1} - m + rX_{t-1} = (1+r)\ X_{t-1} - m \tag{11}$$

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

$$X_t = (1+r)^t \left[V_b - \frac{m}{r} \right] + \frac{m}{r} \tag{12}$$

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

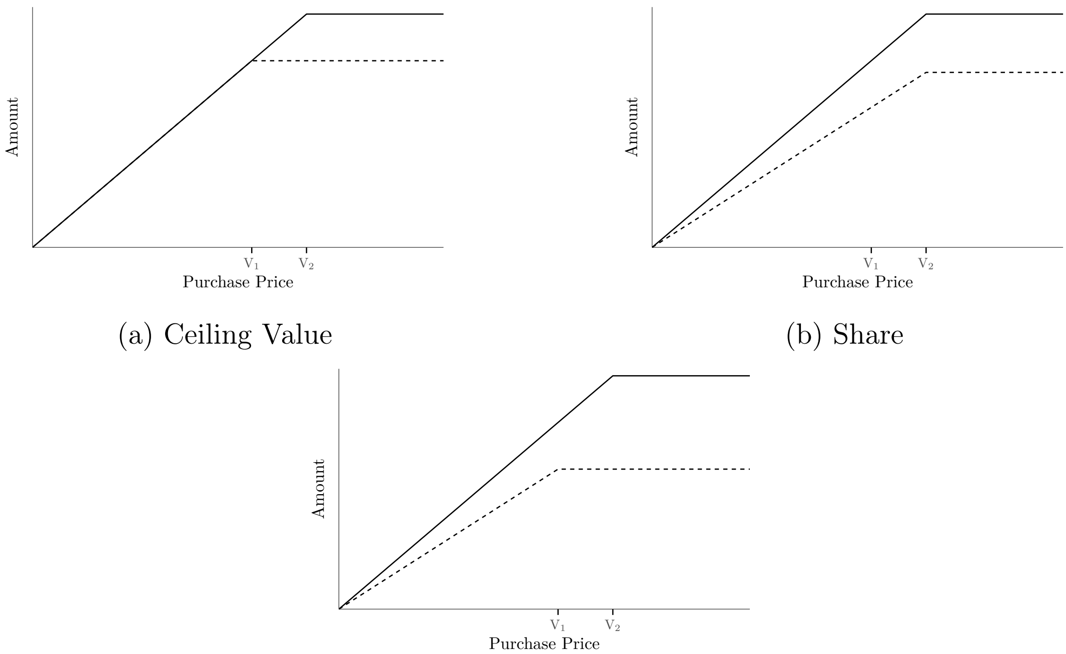
$$m = \frac{rX_0}{1-(1+r)^{-D}} \tag{13}$$

yielding an overall cost for the household to

$$C = \sum_{k=1}^{D} m - V_b = \left[\frac{D\,r}{1-(1+r)^{-D}} - 1 \right] V_b \tag{14}$$

B.2 Difference in Treatment According to Policy-Controlled Primitives

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



(a) Ceiling Value

(b) Share

(c) Ceiling Value and Share

Figure B.2. Difference on Homeownership Cost Induced by the IFL

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

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

C Additional Results from Models Estimations

C.1 Estimated Spline Functions for the GPS specification

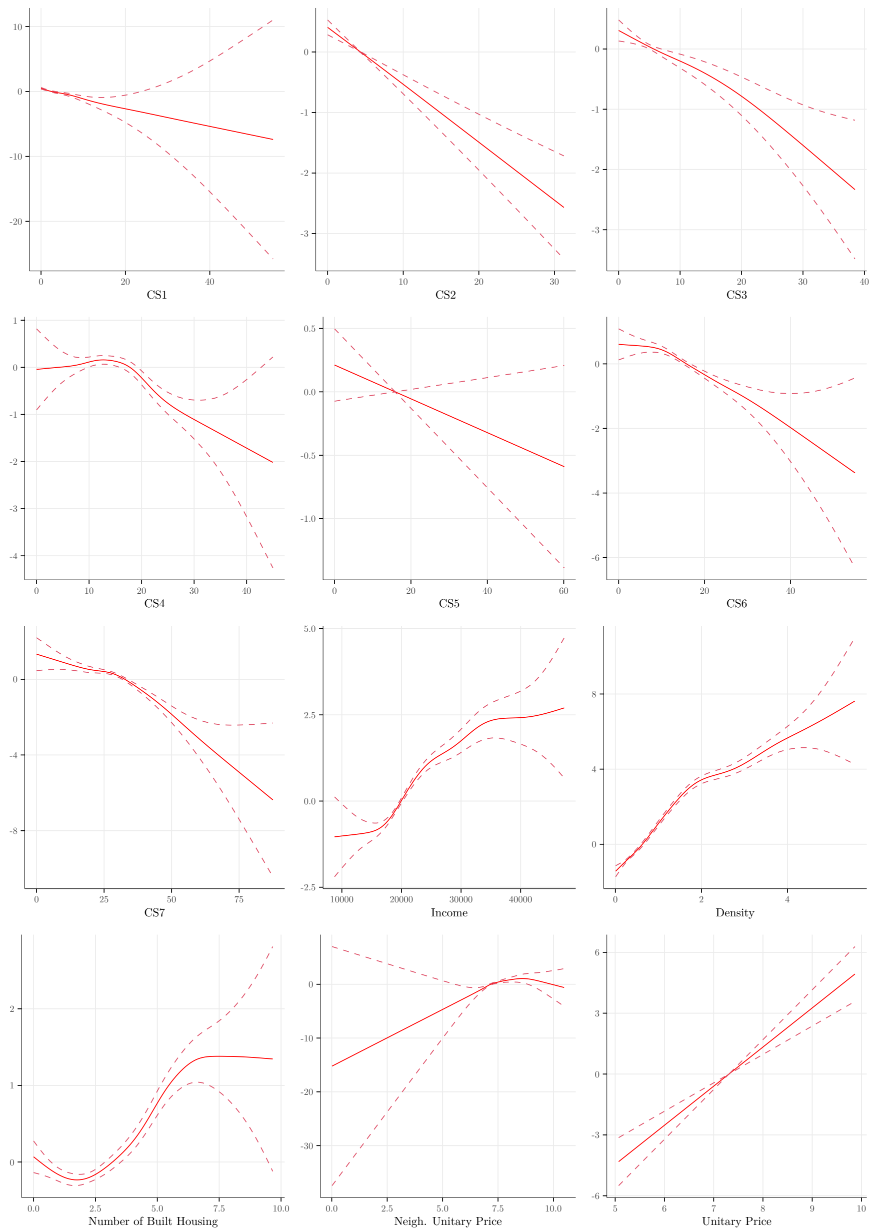


Figure C.1. Contribution for 1D-variable in the GPS Estimation

Notes: For each continuous covariate, we report the functional form in the GPS estimation following the endogenous shrinkage procedure to set the effective degree of freedom. In addition, we report the confidence interval for a 95% level. We exploit the `gam` function from the `mgcv` package.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

C.2 Spatial Smoothing Splines for GPS Estimation

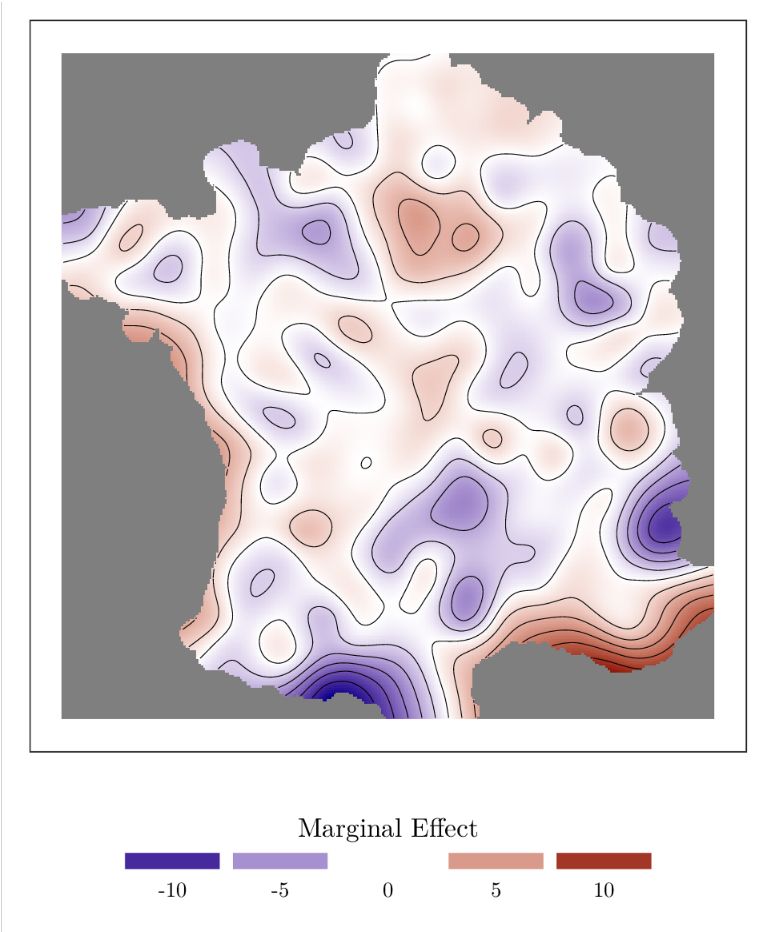


Figure C.2. Spatial Smoothing Function for the GPS Estimation Based On Municipality Coordinates

Notes: We report the spatial smoothing function for the GPS estimation, using bi-variate additive splines. Spline parameters are endogenously shrunk using restricted maximum likelihood approach. The maximum degree of freedom is set to 200. Red (respectively blue) values indicate that the outcome is locally higher than the average. We exploit the `gam` function from the `mgcv` package.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

C.3 Predicted Zoning from GPS Estimation

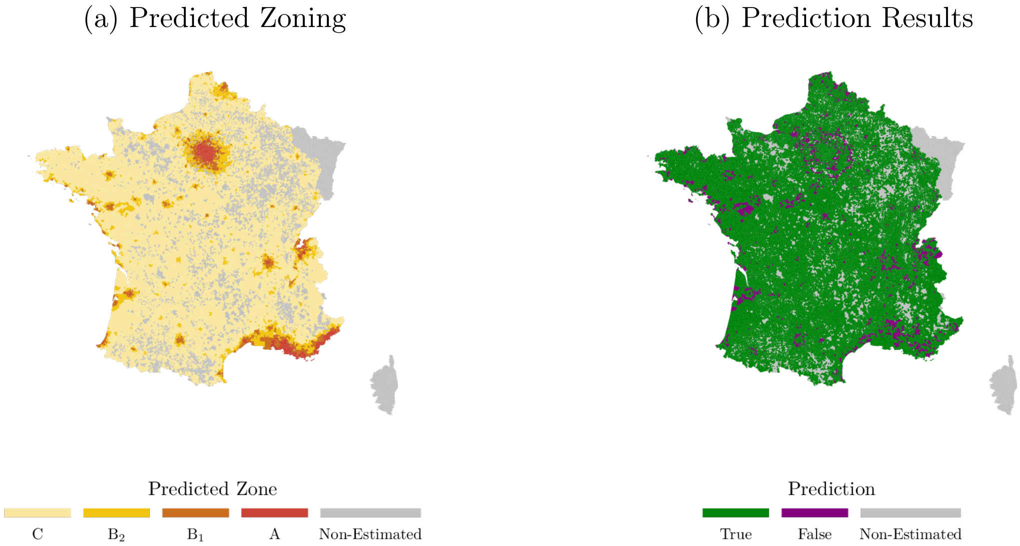


Figure C.3. Estimated Classification of the Municipalities

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

C.4 Descriptive Statistics about Overlap Resulting from GPS Estimation

Table C.4. Overlap

	Share				Number of Obs.			
	A	B_1	B_2	C	A	B_1	B_2	C
A	95.0%	26.6%	3.9%	0.0%	679	190	28	0
B_1	30.3%	95.0%	73.6%	4.7%	398	1,247	966	62
B_2	3.0%	75.5%	95.0%	34.6%	89	2,266	2,850	1,038
C	0.0%	4.2%	20.9%	95.0%	0	909	4,565	20,701

Notes: We report for each pair of treatment level the overlap measured by the share of observations in treatment level g belonging to the 95% range of the latent distribution of the treatment level k. For instance, using the second row of the table, 30.3% of observations classified as B_1 belong to the 95% range of distribution restricted to A observations, according to the latent variable.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

C.5 Joint Significance for Outcomes Specification (Pooled Models)

Table C.5. Joint Significance for Control Variables in the Outcomes Specification

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

Notes: For the nine outcomes Y in columns, we report the average effective degree of freedom, the average χ^2 test and the number of joint significance for control variable in the second step estimation. In addition, we report average regression statistics (bottom rows). These statistics are derived from the estimation of bilateral combinations effects required to obtain our dose-response functions. The unit of observations is municipality. We exploit the `gam` function from the `mgcv` package.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

C.6 Spatial Smoothing Splines Results for Outcomes Specification (Pooled Models)

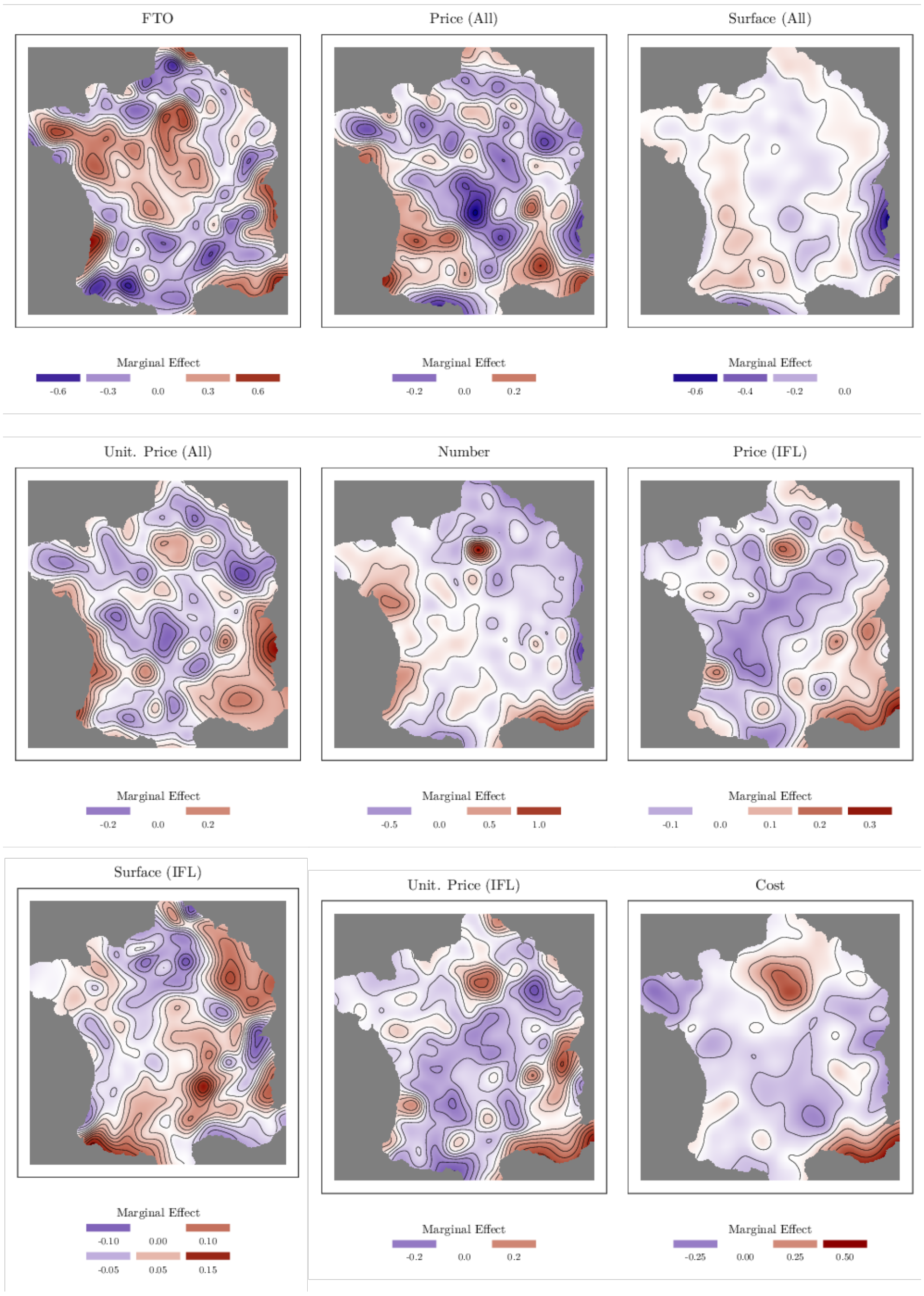


Figure C.6. Marginal Contribution for Spatial Coordinates (Second Step)

Notes: For the nine outcomes Y in columns, we report the spatial smoothing functions for pooled regressions. Our outcome respectively comes from fiscal data or recipients' files. The effective degree of freedom for each function is endogenously shrank. Red (respectively blue) values indicate that the outcome is locally higher than the average. We exploit the `gam` function from the `mgcv` package.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

C.7 Policy Relevant Treatment Effects (ATE)

	Treatement (ATE)											
	2015				2016–2017				2018–2019			
Area	A	B_1	B_2	C	A	B_1	B_2	C	A	B_1	B_2	C
					N_g: Number of FTO							
A	-	0.242	0.451	0.469	-	-1.180	-0.359	-0.508	-	1.168	1.026	0.979
	-	(1.231)	(1.213)	(1.219)	-	(1.032)	(0.953)	(0.950)	-	(1.202)	(1.167)	(1.164)
B_1	-0.242	-	0.209	0.227	1.180	-	0.821**	0.672*	-1.168	-	-0.142	-0.189
	(1.231)	-	(0.185)	(0.174)	(1.032)	-	(0.393)	(0.389)	(1.202)	-	(0.297)	(0.295)
B_2	-0.451	-0.209	-	0.018	0.359	-0.821**	-	-0.149**	-1.026	0.142	-	-0.048
	(1.213)	(0.185)	-	(0.067)	(0.953)	(0.393)	-	(0.061)	(1.167)	(0.297)	-	(0.053)
C	-0.469	-0.227	-0.018	-	0.508	-0.672*	0.149**	-	-0.979	0.189	0.048	-
	(1.219)	(0.174)	(0.067)	-	(0.950)	(0.389)	(0.061)	-	(1.164)	(0.295)	(0.053)	-
					V_g: Average Housing Price (Overall Transaction)							
A	-	-0.431	-0.129	-0.199	-	0.446	0.488	0.456	-	0.294	-0.202	-0.248
	-	(0.784)	(0.730)	(0.729)	-	(0.583)	(0.564)	(0.562)	-	(1.195)	(1.184)	(1.187)
B_1	0.431	-	0.301	0.232	-0.446	-	0.043	0.010	-0.294	-	-0.496***	-0.542***
	(0.784)	-	(0.255)	(0.254)	(0.583)	-	(0.167)	(0.169)	(1.195)	-	(0.142)	(0.135)
B_2	0.129	-0.301	-	-0.070***	-0.488	-0.043	-	-0.033	0.202	0.496***	-	-0.046
	(0.730)	(0.255)	-	(0.021)	(0.564)	(0.167)	-	(0.021)	(1.184)	(0.142)	-	(0.036)
C	0.199	-0.232	0.070***	-	-0.456	-0.010	0.033	-	0.248	0.542***	0.046	-
	(0.729)	(0.254)	(0.021)	-	(0.562)	(0.169)	(0.021)	-	(1.187)	(0.135)	(0.036)	-
					Q_g: Average Housing Size (Overall Transaction)							
A	-	-0.081	0.132	0.113	-	0.177	0.147	0.168	-	0.270	0.128	0.158
	-	(0.294)	(0.286)	(0.284)	-	(0.324)	(0.321)	(0.321)	-	(0.293)	(0.289)	(0.290)
B_1	0.081	-	0.213***	0.194***	-0.177	-	-0.030	-0.010	-0.270	-	-0.142**	-0.112*
	(0.294)	-	(0.074)	(0.072)	(0.324)	-	(0.054)	(0.053)	(0.293)	-	(0.065)	(0.065)
B_2	-0.132	-0.213***	-	-0.019	-0.147	0.030	-	0.021	-0.128	0.142**	-	0.030**
	(0.286)	(0.074)	-	(0.019)	(0.321)	(0.054)	-	(0.015)	(0.289)	(0.065)	-	(0.014)
C	-0.113	-0.194***	0.019	-	-0.168	0.010	-0.021	-	-0.158	0.112*	-0.030**	-
	(0.284)	(0.072)	(0.019)	-	(0.321)	(0.053)	(0.015)	-	(0.290)	(0.065)	(0.014)	-
					P_g: Average Housing Price per m^2 (Overall Transaction)							
A	-	-0.116	0.092	0.061	-	0.209	-0.021	-0.103	-	0.259	-0.146	-0.227
	-	(0.867)	(0.852)	(0.850)	-	(0.588)	(0.562)	(0.561)	-	(1.087)	(1.080)	(1.081)
B_1	0.116	-	0.209	0.177	-0.209	-	-0.229	-0.312**	-0.259	-	-0.404***	-0.485***
	(0.867)	-	(0.208)	(0.203)	(0.588)	-	(0.154)	(0.156)	(1.087)	-	(0.137)	(0.133)
B_2	-0.092	-0.209	-	-0.031	0.021	0.229	-	-0.082***	0.146	0.404***	-	-0.081**
	(0.852)	(0.208)	-	(0.037)	(0.562)	(0.154)	-	(0.022)	(1.080)	(0.137)	-	(0.032)
C	-0.061	-0.177	0.031	-	0.103	0.312**	0.082***	-	0.227	0.485***	0.081**	-
	(0.850)	(0.203)	(0.037)	-	(0.561)	(0.156)	(0.022)	-	(1.081)	(0.133)	(0.032)	-

Notes: We report the bilateral combinations effect for ATE type estimator. The four panels correspond to the four outcome concerned by ATE estimation and derived from fiscal data. Then, we have three main columns that represent the stable period for the IFL scheme, with four subcolumns related to the ABC classification. In rows, we have again the levels contained in the ABC classification. Hence, the bilateral combinations are reported for each intersection, and must be understand as "if (rows) have received (cols), difference in outcome would be (results)". We also report in brackets the standard errors obtained with a bootstrap procedure with 500 iterations.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

C.8 Policy Relevant Treatment Effects (ATT)

	Treatment (ATT)											
	2015				2016–2017				2018–2019			
Area	A	B_1	B_2	C	A	B_1	B_2	C	A	B_1	B_2	C
					$\widetilde{N}_g$: Number of IFL							
A	-	-0.999***	-0.804*	-0.938	-	-0.289	-0.871**	-2.548***	-	-0.095	-0.668	-1.933*
	-	(0.266)	(0.434)	(0.616)	-	(0.180)	(0.408)	(0.600)	-	(0.230)	(0.540)	(1.024)
B_1	3.446**	-	-0.196**	-0.290	0.275	-	-0.205**	-0.547***	-1.715	-	-0.210**	-0.576*
	(1.587)	-	(0.079)	(0.179)	(1.450)	-	(0.090)	(0.172)	(1.209)	-	(0.089)	(0.297)
B_2	3.809*	-0.145	-	-0.163***	0.185	-0.204	-	-0.233***	-0.279	-0.124	-	-0.261***
	(2.036)	(0.158)	-	(0.052)	(2.103)	(0.186)	-	(0.051)	(1.940)	(0.179)	-	(0.059)
C	5.385*	0.275	-0.009	-	2.255	-0.095	0.002	-	0.670	-1.472*	0.066	-
	(3.138)	(0.695)	(0.100)	-	(3.344)	(1.090)	(0.138)	-	(3.067)	(0.814)	(0.131)	-
					$\widetilde{V}_g$: Average Housing Price (Subsidized Housing)							
A	-	-0.028	-0.191**	-0.076	-	-0.022	-0.011	-0.060	-	-0.069*	-0.161***	-0.046
	-	(0.054)	(0.082)	(0.103)	-	(0.027)	(0.054)	(0.096)	-	(0.036)	(0.060)	(0.085)
B_1	-0.801***	-	-0.043**	-0.034	-0.118	-	-0.050***	-0.036	-0.293	-	-0.048***	-0.048*
	(0.297)	-	(0.017)	(0.032)	(0.229)	-	(0.013)	(0.025)	(0.273)	-	(0.014)	(0.028)
B_2	-1.000**	0.096***	-	-0.030***	-0.163	0.054	-	-0.019**	-0.090	0.103***	-	-0.032***
	(0.391)	(0.030)	-	(0.011)	(0.313)	(0.035)	-	(0.008)	(0.378)	(0.028)	-	(0.009)
C	-1.508**	0.160	0.075***	-	0.222	0.100	0.033	-	0.207	0.307**	0.037	-
	(0.613)	(0.123)	(0.022)	-	(0.514)	(0.154)	(0.044)	-	(0.595)	(0.124)	(0.038)	-
					$\widetilde{Q}_g$: Average Housing Size (Subsidized Housing)							
A	-	-0.109	0.167*	0.654***	-	-0.073*	0.077	0.242**	-	-0.083*	0.051	0.350**
	-	(0.087)	(0.090)	(0.160)	-	(0.041)	(0.066)	(0.098)	-	(0.049)	(0.069)	(0.144)
B_1	-1.027**	-	0.022	0.125***	-0.409	-	0.011	0.057**	-0.448	-	0.023	0.117***
	(0.457)	-	(0.027)	(0.045)	(0.290)	-	(0.018)	(0.027)	(0.319)	-	(0.017)	(0.045)
B_2	-1.062*	0.026	-	0.014	-0.996**	0.089*	-	0.010	-0.734	0.019	-	-0.001
	(0.582)	(0.039)	-	(0.013)	(0.408)	(0.053)	-	(0.008)	(0.469)	(0.036)	-	(0.014)
C	-1.997**	0.216	0.054**	-	-2.287***	0.473	0.012	-	-0.941	0.262*	-0.046	-
	(0.894)	(0.155)	(0.027)	-	(0.695)	(0.288)	(0.058)	-	(0.778)	(0.158)	(0.058)	-
					$\widetilde{P}_g$: Average Housing Price per m^2 (Subsidized Housing)							
A	-	0.100	-0.355***	-0.460***	-	0.027	-0.094*	-0.302***	-	0.017	-0.214***	-0.396***
	-	(0.087)	(0.101)	(0.135)	-	(0.037)	(0.052)	(0.093)	-	(0.042)	(0.080)	(0.124)
B_1	0.231	-	-0.071***	-0.128***	0.382	-	-0.063***	-0.093***	0.212	-	-0.069***	-0.165***
	(0.370)	-	(0.022)	(0.039)	(0.263)	-	(0.014)	(0.023)	(0.233)	-	(0.018)	(0.039)
B_2	0.116	0.037	-	-0.044***	0.978**	-0.042	-	-0.026***	0.624*	0.077***	-	-0.032***
	(0.467)	(0.039)	-	(0.014)	(0.482)	(0.070)	-	(0.007)	(0.332)	(0.025)	-	(0.012)
C	0.484	-0.046	0.018	-	1.822**	-0.366	0.026	-	1.219**	0.079	0.084	-
	(0.705)	(0.149)	(0.030)	-	(0.836)	(0.340)	(0.026)	-	(0.548)	(0.119)	(0.067)	-
					$\widetilde{C}_g$: Average Cost per IFL							
A	-	-0.125	-0.516***	-1.514***	-	-0.056	-0.226**	-0.173	-	-0.110**	-1.096***	-0.745***
	-	(0.079)	(0.162)	(0.393)	-	(0.043)	(0.096)	(0.179)	-	(0.049)	(0.115)	(0.148)
B_1	-0.096	-	-0.527***	-1.148***	0.155	-	-0.200***	-0.266***	-0.229	-	-0.874***	-0.922***
	(0.430)	-	(0.042)	(0.106)	(0.226)	-	(0.025)	(0.045)	(0.347)	-	(0.019)	(0.043)
B_2	0.245	0.509***	-	-0.490***	-0.151	0.144***	-	-0.110***	0.664	0.876***	-	-0.126***
	(0.560)	(0.056)	-	(0.030)	(0.364)	(0.037)	-	(0.014)	(0.580)	(0.059)	-	(0.017)
C	0.532	0.905***	0.555***	-	-0.432	0.508**	0.139***	-	1.197	0.976***	0.140***	-
	(0.908)	(0.224)	(0.086)	-	(0.628)	(0.222)	(0.033)	-	(0.979)	(0.232)	(0.039)	-

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

C.9 Partial Plots (1)

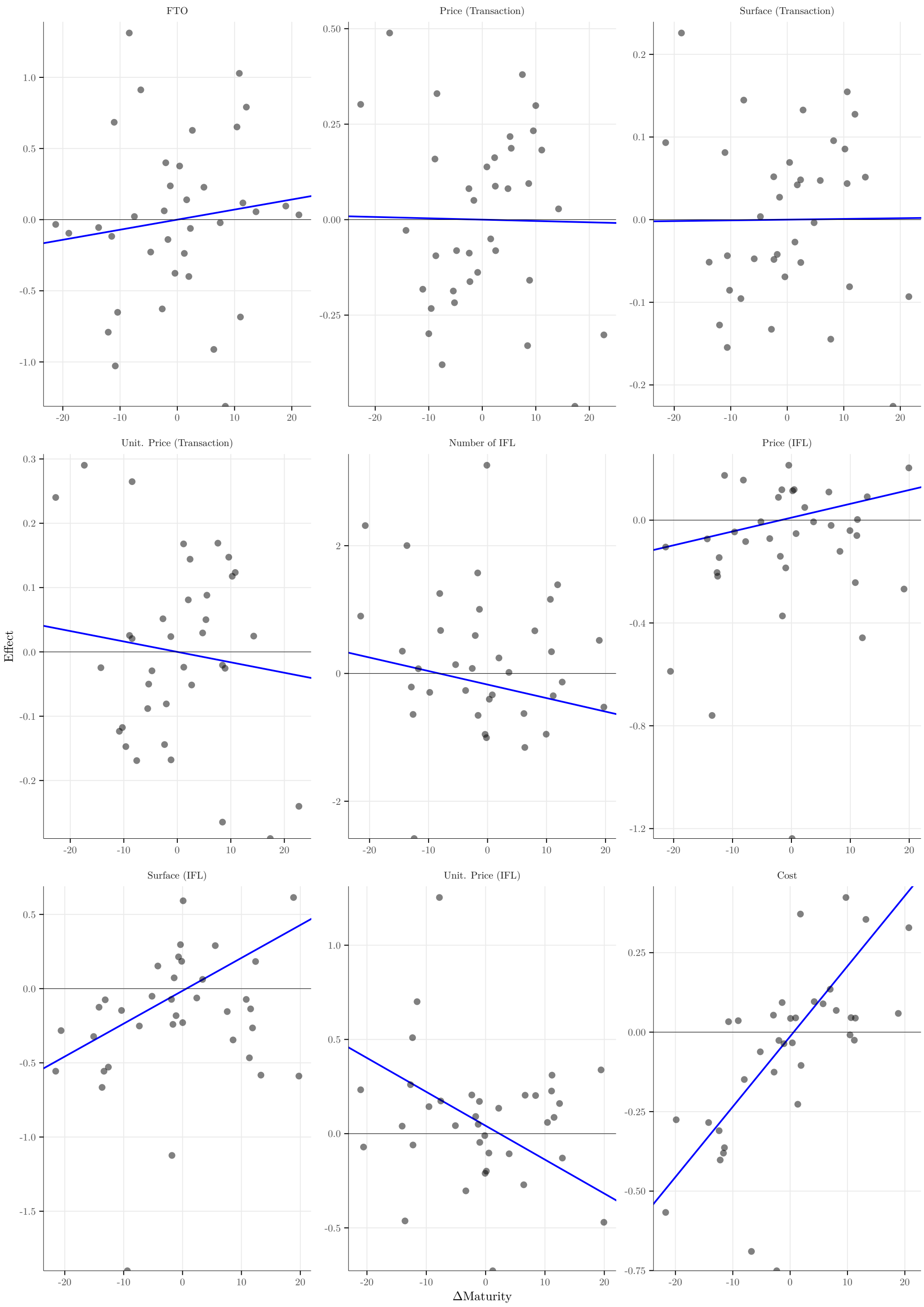


Figure C.9. Dose-Response Plots

Notes: We report the partial plot for dose-response function. The nine partial plot corresponds to our nine selected outcomes, while the x-axis represents a variation in primitive data sources. The observation unit is the bilateral combinations of treatment (either ATE or ATT). The regression is performed using OLS.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

C.10 Partial Plots (2)

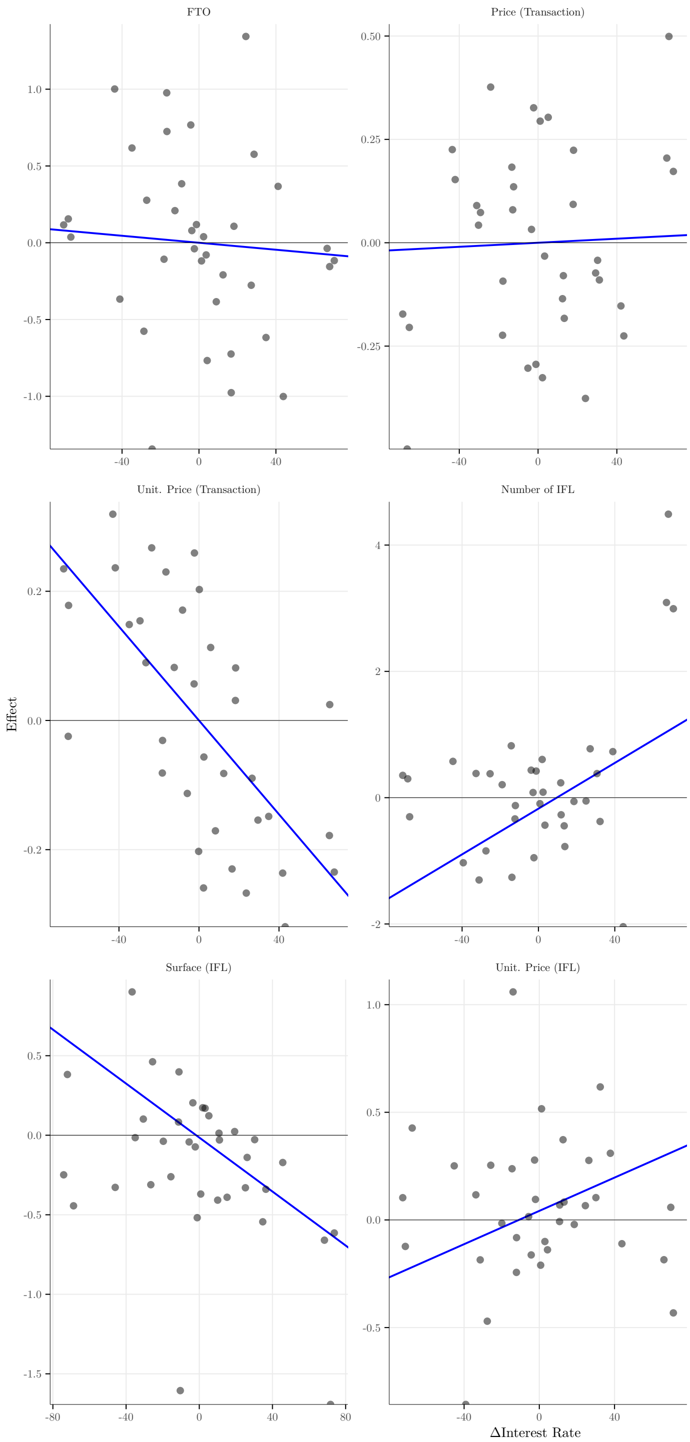


Figure C.10. Dose-Response Plots

Notes: We report the partial plot for dose-response function. The nine partial plot corresponds to our nine selected outcomes, while the x-axis represents a variation in primitive data sources. The observation unit is the bilateral combinations of treatment (either ATE or ATT). The regression is performed using OLS.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

C.11 Partial Plots (3)

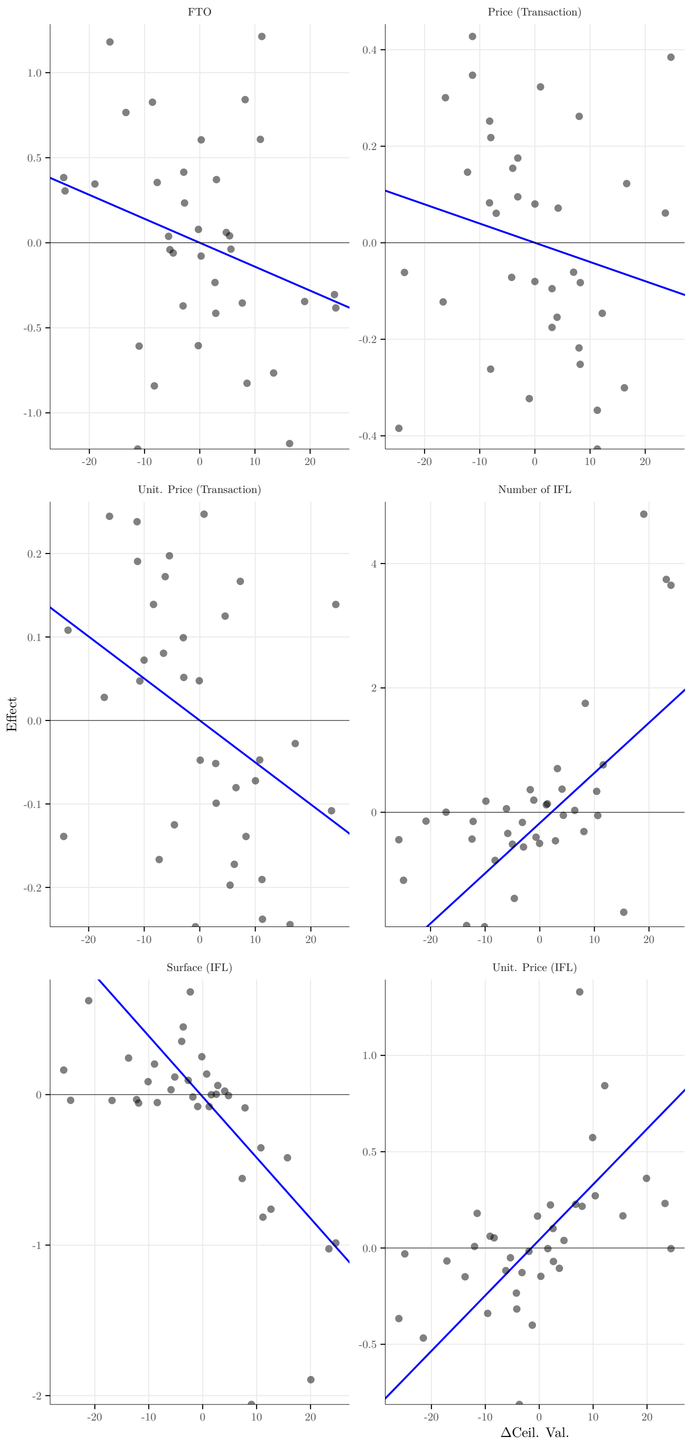


Figure C.11. Dose-Response Plots

Notes: We report the partial plot for dose-response function. The nine partial plot corresponds to our nine selected outcomes, while the x-axis represents a variation in primitive data sources. The observation unit is the bilateral combinations of treatment (either ATE or ATT). The regression is performed using OLS.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

C.12 Partial Plots (4)

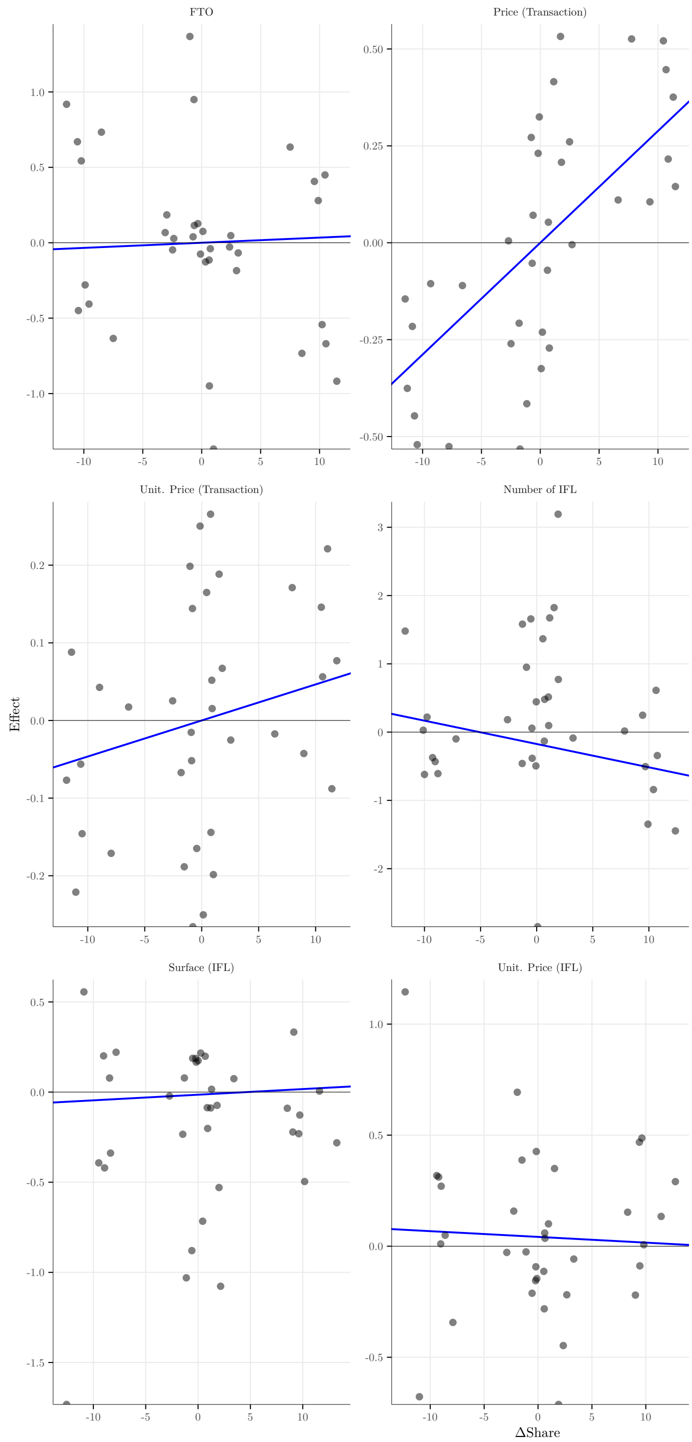


Figure C.12. Dose-Response Plots

Notes: We report the partial plot for dose-response function. The nine partial plot corresponds to our nine selected outcomes, while the x-axis represents a variation in primitive data sources. The observation unit is the bilateral combinations of treatment (either ATE or ATT). The regression is performed using OLS.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

C.13 Main Results With Maximum Degree of Freedom set to 100 for Spatial Smoothing

Table C.13. OLS coefficients for policy primitives from dose-response functions

	Outcome variables from ...								
	Tax	Transaction Data			IFL files				
	N	V	S	P	$\widetilde{N}$	$\widetilde{V}$	$\widetilde{S}$	$\widetilde{P}$	$\widetilde{C}$
Covering Share	0.009	0.010	-0.056	-0.017	0.031	0.020	-0.008	0.013	0.047***
	(0.120)	(0.021)	(0.049)	(0.051)	(0.075)	(0.017)	(0.017)	(0.016)	(0.014)
Ceiling Value	-0.054	-0.007	0.035	0.039	0.040	-0.010	0.015	0.008	0.016
	(0.127)	(0.019)	(0.039)	(0.040)	(0.089)	(0.025)	(0.023)	(0.014)	(0.012)
Interest Rate	-0.013	0.001	-0.004	0.002	0.024	-0.000	0.000	0.002	0.006
	(0.034)	(0.005)	(0.008)	(0.008)	(0.026)	(0.007)	(0.007)	(0.004)	(0.004)
Loan Maturity	0.019	0.007	-0.034	-0.036	0.016	0.008	-0.012	-0.004	0.017**
	(0.074)	(0.017)	(0.027)	(0.031)	(0.039)	(0.009)	(0.009)	(0.007)	(0.009)
Constant	0.000	-0.000	0.000	-0.000	-0.261**	0.025	0.001	0.025	-0.006
	(0.000)	(0.000)	(0.000)	(0.000)	(0.125)	(0.038)	(0.039)	(0.029)	(0.030)
R^2	0.809	0.201	0.318	0.415	0.204	0.129	0.138	0.355	0.834
Adj. R^2	0.784	0.097	0.231	0.339	0.101	0.016	0.027	0.271	0.812
N	36	36	36	36	36	36	36	36	36

Notes: For the nine outcomes Y in columns, the table reports the β_a^Y coefficients associated to each primitive in rows. They are estimated from dose-response functions of Equation 5. N accounts for the number of new homeowners, V for housing value, S for surface, and P for unitary housing price. The variables with a $\sim$ are the same variables computed for IFL recipients, $\widetilde{C}$ is the IFL cost. The unit of observation is the bilateral combination of four ABC zones for the three periods of interest. The maximum degree of freedom for the spatial smoothing in the specifications of both propensity score and outcomes is set to 100. Standards errors in parenthesis are estimated by bootstrap with 500 iterations. ATEs for tax and transaction data are weighted according to the inverse of their bootstrapped standard errors, ATT for IFL variables are additionally weighted according to the number of municipalities receiving the considered treatment levels.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.
*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

C.14 Main Results With Maximum Degree of Freedom set to 50 for Spatial Smoothing

Table C.14. OLS coefficients for policy primitives from dose-response functions

	Outcome variables from ...								
	Tax	Transaction Data			IFL files				
	N	V	S	P	$\widetilde{N}$	$\widetilde{V}$	$\widetilde{S}$	$\widetilde{P}$	$\widetilde{C}$
Covering Share	0.007	0.007	-0.045	-0.030	0.014	0.025**	0.001	0.002	0.041***
	(0.103)	(0.016)	(0.049)	(0.039)	(0.051)	(0.010)	(0.016)	(0.012)	(0.013)
Ceiling Value	-0.116	-0.007	0.022	0.012	-0.096	0.006	0.013	0.013	0.024**
	(0.123)	(0.016)	(0.032)	(0.029)	(0.075)	(0.014)	(0.016)	(0.008)	(0.009)
Interest Rate	-0.026	0.001	-0.005	-0.003	-0.002	0.003	0.002	0.001	0.005*
	(0.033)	(0.004)	(0.007)	(0.006)	(0.023)	(0.004)	(0.005)	(0.003)	(0.003)
Loan Maturity	0.057	0.008	-0.018	-0.015	0.093***	-0.002	-0.010	-0.010**	0.007
	(0.073)	(0.012)	(0.022)	(0.023)	(0.028)	(0.005)	(0.007)	(0.004)	(0.005)
Constant	-0.000	-0.000	0.000	0.000	-0.297***	0.055**	-0.008	0.040**	0.020
	(0.000)	(0.000)	(0.000)	(0.000)	(0.074)	(0.026)	(0.032)	(0.017)	(0.022)
R^2	0.910	0.146	0.372	0.140	0.282	0.167	0.169	0.319	0.937
Adj. R^2	0.899	0.036	0.290	0.029	0.190	0.060	0.062	0.231	0.928
N	36	36	36	36	36	36	36	36	36

Notes: For the nine outcomes Y in columns, the table reports the β_a^Y coefficients associated to each primitive in rows. They are estimated from dose-response functions of Equation 5. N accounts for the number of new homeowners, V for housing value, S for surface, and P for unitary housing price. The variables with a $\sim$ are the same variables computed for IFL recipients, $\widetilde{C}$ is the IFL cost. The unit of observation is the bilateral combination of four ABC zones for the three periods of interest. The maximum degree of freedom for the spatial smoothing in the specifications of both propensity score and outcomes is set to 50. Standards errors in parenthesis are estimated by bootstrap with 500 iterations. ATEs for tax and transaction data are weighted according to the inverse of their bootstrapped standard errors, ATT for IFL variables are additionally weighted according to the number of municipalities receiving the considered treatment levels.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.
*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

C.15 Main Results With No Spatial Smoothing

Table C.15. OLS coefficients for policy primitives from dose-response functions

	Outcome variables from ...								
	Tax	Transaction Data			IFL files				
	N	V	S	P	$\widetilde{N}$	$\widetilde{V}$	$\widetilde{S}$	$\widetilde{P}$	$\widetilde{C}$
Covering Share	0.028	-0.007	-0.066***	-0.053**	0.105*	-0.023***	0.011	0.002	0.036***
	(0.082)	(0.011)	(0.018)	(0.024)	(0.059)	(0.009)	(0.012)	(0.010)	(0.009)
Ceiling Value	0.035	-0.029***	0.029	0.020	-0.009	-0.019**	0.015	-0.003	0.024***
	(0.070)	(0.008)	(0.021)	(0.018)	(0.039)	(0.008)	(0.010)	(0.008)	(0.006)
Interest Rate	0.005	-0.004	-0.007	-0.008	0.012	-0.007	0.002	-0.002	0.004
	(0.204)	(0.025)	(0.062)	(0.053)	(0.128)	(0.024)	(0.032)	(0.026)	(0.021)
Loan Maturity	-0.035	0.024***	-0.032***	-0.029**	0.002	0.006*	-0.011**	-0.001	0.004
	(0.041)	(0.007)	(0.011)	(0.012)	(0.021)	(0.004)	(0.004)	(0.003)	(0.004)
Constant	0.000	0.000**	0.000**	0.000**	-0.226***	0.025	0.000	0.020	-0.003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.076)	(0.015)	(0.016)	(0.016)	(0.021)
R^2	0.323	0.493	0.363	0.430	0.255	0.440	0.368	0.124	0.894
Adj. R^2	0.235	0.428	0.281	0.356	0.159	0.367	0.286	0.011	0.880
N	36	36	36	36	36	36	36	36	36

Notes: For the nine outcomes Y in columns, the table reports the β_a^Y coefficients associated to each primitive in rows. They are estimated from dose-response functions of Equation 5. N accounts for the number of new homeowners, V for housing value, S for surface, and P for unitary housing price. The variables with a $\sim$ are the same variables computed for IFL recipients, $\widetilde{C}$ is the IFL cost. The unit of observation is the bilateral combination of four ABC zones for the three periods of interest. The specifications of both propensity score and outcomes do not include spatial smoothing. Standards errors in parenthesis are estimated by bootstrap with 500 iterations. ATEs for tax and transaction data are weighted according to the inverse of their bootstrapped standard errors, ATT for IFL variables are additionally weighted according to the number of municipalities receiving the considered treatment levels.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.
*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

C.16 Cost Variation for One Unit Increase According to Each Primitive Source

Table C.16. Cost Variation for One Unit Increase According to Each Primitive Source

Primitive	Overall Cost	$(\partial\widetilde{N}/\partial a)\widetilde{C}$	$(\partial\widetilde{C}/\partial a)\widetilde{N}$
Covering Share	-255.4	-3,715.7	3,460.3***
	(4,168)	(4,041)	(976)
Ceiling Value	2,011.1	2,530.4	-519.2
	(3,067)	(2,961)	(810)
Interest Rates	1,046.9	1,132.7	-85.8
	(900)	(864)	(254)
Loan Maturity	1,976.3	224.9	1,751.4***
	(1,904)	(1,826)	(485)

Notes: We report for each primitive source, the overall impact in euros in the government budget at the municipality level. We distinguish the overall impact on policy cost according to the impact resulting from increase of the number of recipients (3^{rd} column) and the average cost per recipient (4^{th} column). The variation corresponds to one unit increase for the primitive. Interest rates is expressed in hundredth of unit. We report standard errors obtained with a bootstrap procedure with 500 iterations.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

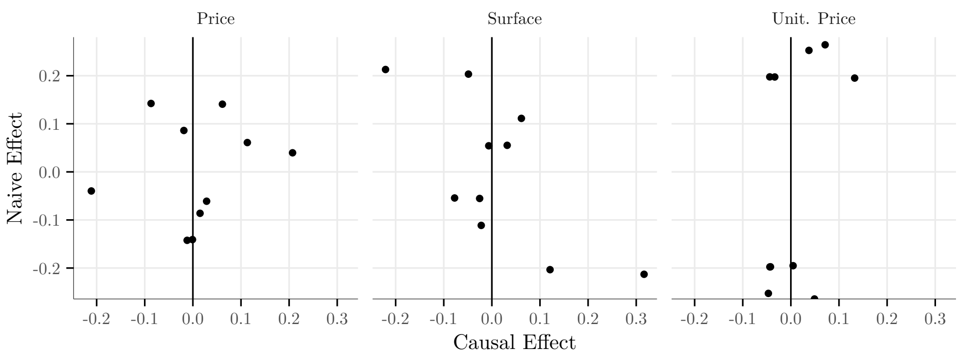
D Placebo Analysis

D.1 Results from the Placebo Analysis

Our placebo analysis relies on specific feature of the IFL design. As shown in Equation 1, the IFL amount is characterised by the two policy primitives, the coverage share and the ceiling value being spatially heterogeneous in line with the ABC perimeter. We take advantage of the fact that differences in IFL subsidy between two ABC areas that have similar covering share concerns the most expensive operations (for more information about IFL subsidy variation, see Appendix B.2). Hence, observations with purchase price under the lowest ceiling value for areas with similar covering share, benefit from the same IFL amount.

Our procedure is as follows. For IFL aggregated outcomes restricted to observations with no difference of IFL amount, we first estimate for treatment level naive regressions corresponding to unconditional average difference, without correcting for ABC perimeter endogeneity. Then, we estimate treatment effect using our doubly robust estimator. We present bilateral effects according to whether it is the naive estimator or the doubly robust one. As we can select observations not subject to differences in treatment for IFL outcomes, we restrict our placebo analysis at the intensive margin related to IFL recipients housing choices. We report in Fig. D.1 bi-variate graphs for policy relevant treatment effects to compare magnitude between the naive and the doubly robust estimations.

Figure D.1. Naive and policy-relevant treatment effects used in placebo analysis



Notes: We report the 36 bilateral combinations of the IFL effects on outcomes restricted to recipients for observations with no difference in treatment intensity. In Y-axis, we report the naive effect, *i.e.* without weighting scheme according to treatment intensity and regression adjustment. In the X-axis, we report the doubly robust estimator using the GPS specification and the regression adjustment. Our choice to restrict placebo analysis to the IFL outcomes is driven by the possibility to select precisely observations with no difference in treatment (see Appendix B.2).
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

The placebo analysis supports the validity of our two-steps procedure. Indeed, while naive estimated effects are sizeable and significant (and confirms the endogeneity issues of the ABC perimeter), our policy relevant treatment effects estimated on population with similar treatment intensity are not significant for unitary housing price outcomes. However, there are still some significant differences for surface and overall purchase price. Finally, the placebo analysis cannot allow to reject the selection-on-observables restriction.

E Difference-in-Difference Results

E.1 Design of the Natural Experiments

We estimate causal effects of IFL policy based on alternative identification strategy by leveraging a natural experiment that occurred in January 2018. Indeed, a major reform affected significantly the covering share of two areas from the ABC zoning (namely the B_2 and C tier) whereas it remains unaffected for two other areas (namely the A and the B_1).

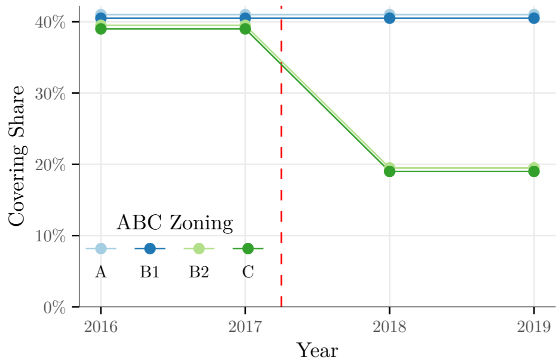


Figure E.1. Covering Share per ABC Areas According to Time

Notes: We represent the coverage share per ABC tier from 2016 to 2019. Green lines corresponds respectively to B_2 and C areas that experience a cut in the covering share that has been introduced in January 2018. The covering share falls from 40% to 20%. In the meantime, areas represented in blue lines does not experience any change of covering share in January 2018.

As the ABC zoning i s endogenous, we adopt a difference-in-difference approach and consider each bilateral combinations of treatment level (for instance the A-B_2), including bilateral combinations of areas with no difference in treatment as placebo test (for instance the A-B_1 combinations). We then estimate the following equation

$$Y_{jt} = \alpha \cdot \mathbb{1}_{t>=2018} + \beta \cdot \mathbb{1}_{j\in 40^-} + \gamma \cdot \mathbb{1}_{t>=2018} \times \mathbb{1}_{j\in 40^-} + \varepsilon_{jt} \tag{15}$$

with Y_{jt} outcome of interest for group j at time t, $\mathbb{1}_{j\in 40^-}$ indicating whether the group j is concerned by the downgrade of the covering share and ε_{jt} the error term. The parameter of interest γcaptures the effect of reducing IFL intensity through the covering share. We estimate this equation by OLS with clustered standard errors at the level of commuting zones. We report results in the following subsections, considering the six possibilities under investigations.

These results allow us to assess the effect of variation in the coverage share. However, given the heterogeneity resulting from LATE issues, these results are only a robustness check of our main results using the selection-on-observables approach. The coverage share has a strong effect on the cost of the policy, while it has no effect on the number of first-time owners. Meanwhile, a reduction in the coverage rate has a significant effect on the number of recipients, strengthening the credibility of the distortion in housing choice. Thus, these results from natural experiment designs are consistent with our main findings.

E.2 DiD Estimation for B_1 and B_2 Areas

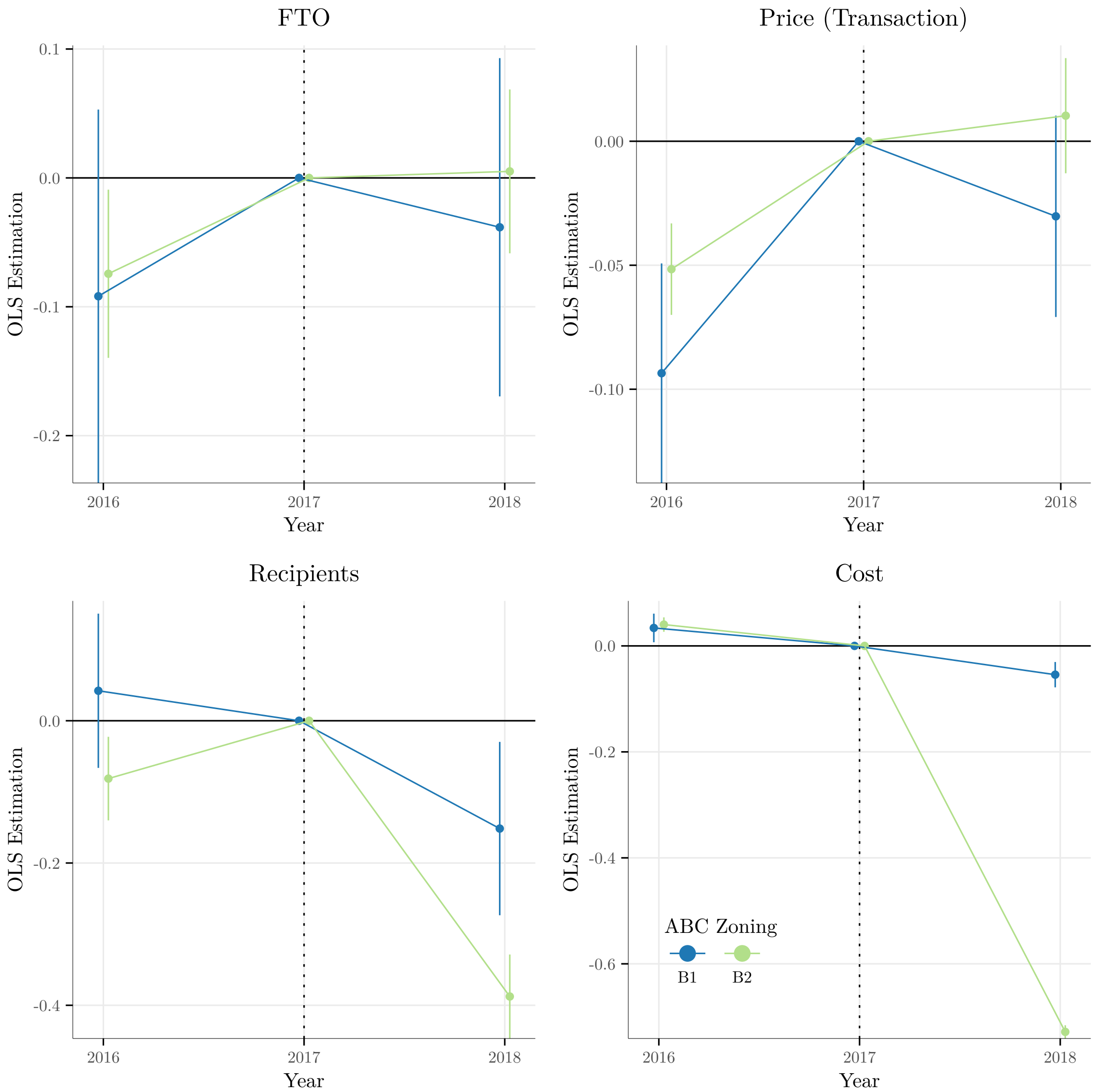


Figure E.2. OLS Results from DiD Identification Strategy restricted to B_1 and B_2 Areas

Notes: We represent results from event study for four outcomes of interest (number of first-time owners, average transaction price, number of recipients and policy cost). We restrict samples to B_1 and B_2 municipalities. B_2 municipalities experience a cut in the coverage share that occurred in January 2018. Conversely, B_1 municipalities do not experience any cut in the covering share. Our unit of observation is the municipality. We report standard error in vertical line with a 95% confidence interval. Standard errors are clustered to the commuting zone.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

E.3 DiD Estimation for B_1 and C Areas

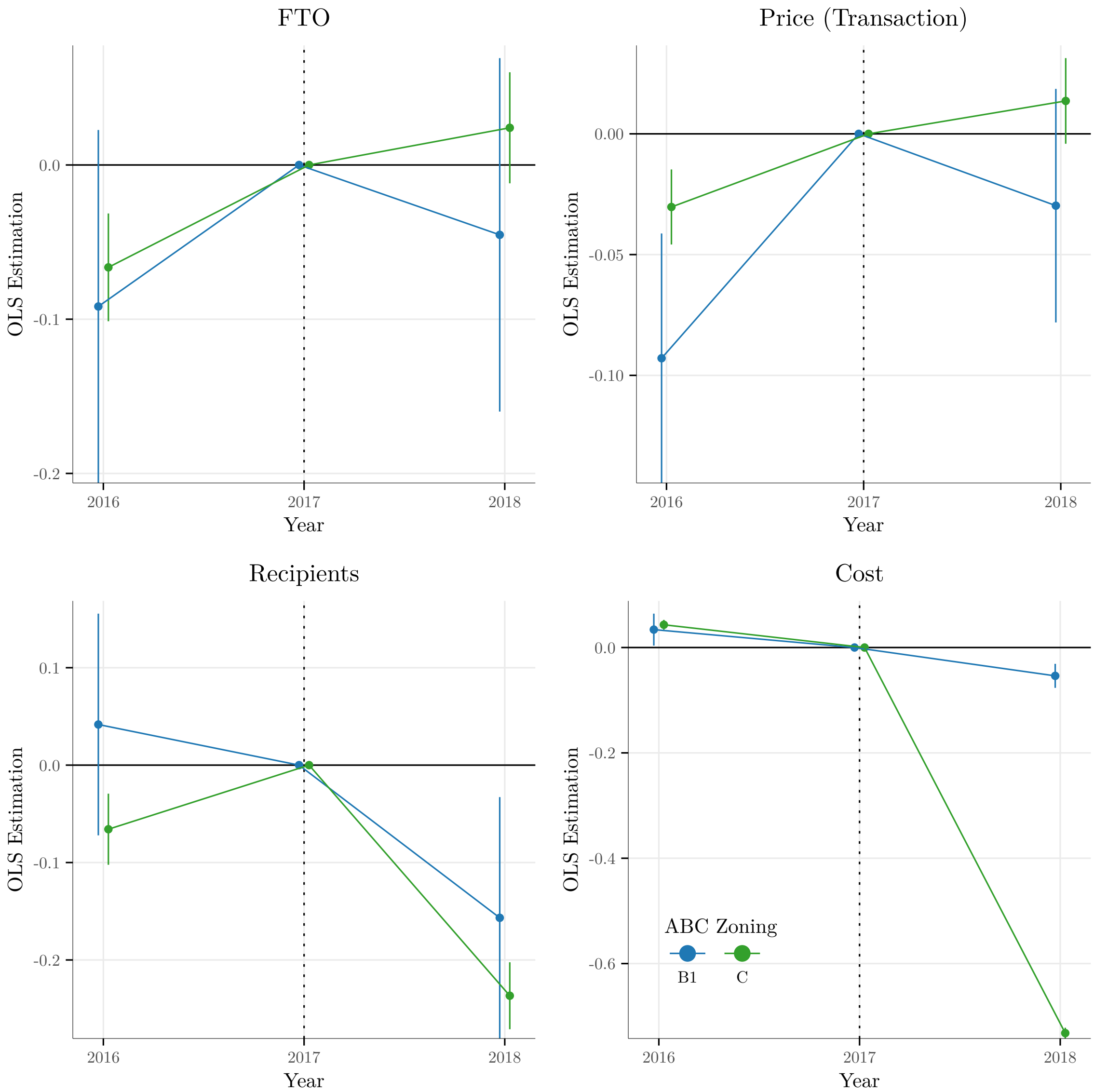


Figure E.3. OLS Results from DiD Identification Strategy restricted to B_1 and C Areas

Notes: We represent results from event study for four outcomes of interest (number of first-time owners, average transaction price, number of recipients and policy cost). We restrict samples to B_1 and C municipalities. C municipalities experience a cut in the coverage share that occurred in January 2018. Conversely, B_1 municipalities do not experience any cut in the covering share. Our unit of observation is the municipality. We report standard error in vertical line with a 95% confidence interval. Standard errors are clustered to the commuting zone.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

E.4 DiD Estimation for A and B_2 Areas

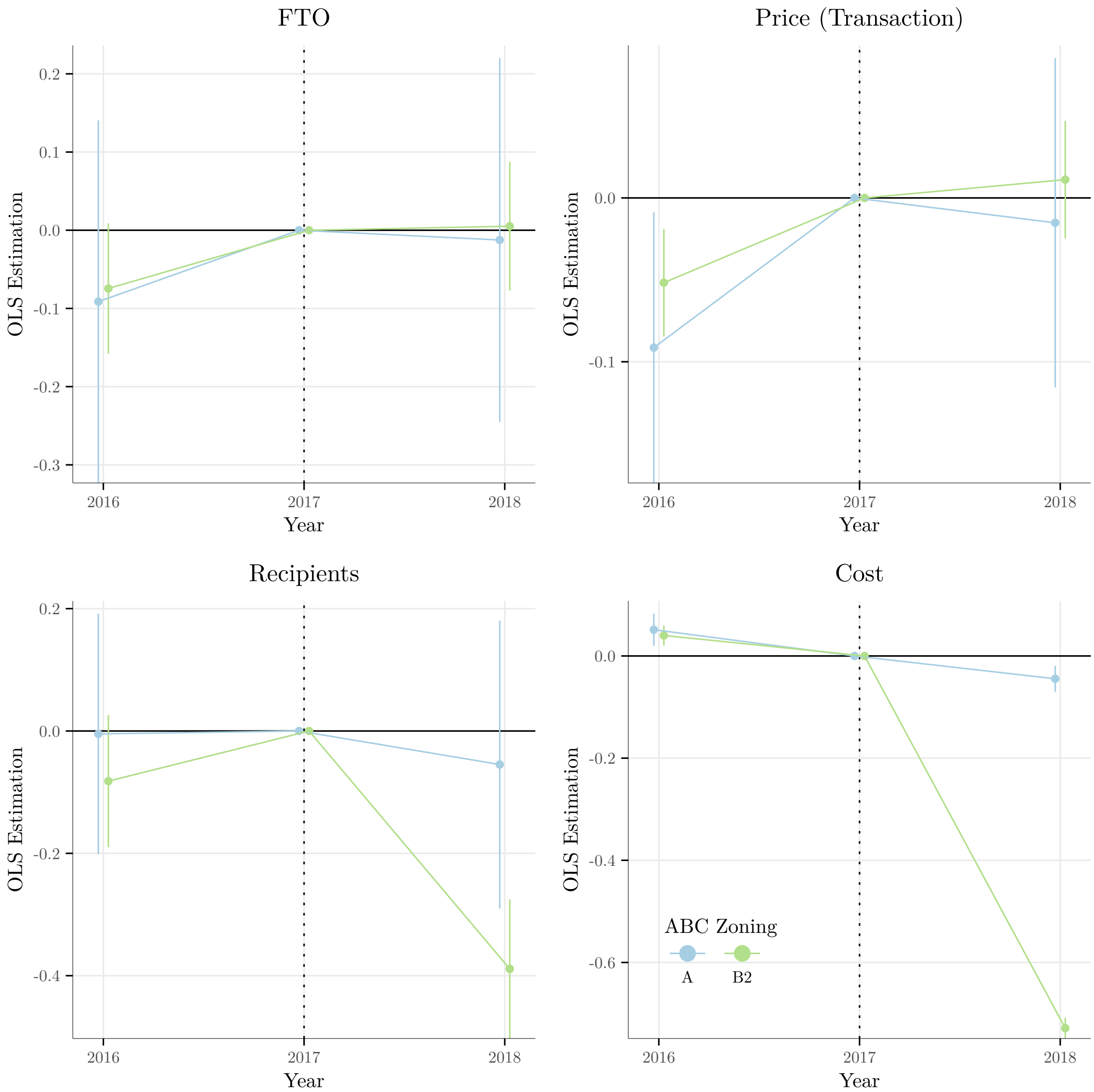


Figure E.4. OLS Results from DiD Identification Strategy restricted to A and B_2 Areas

Notes: We represent results from event study for four outcomes of interest (number of first-time owners, average transaction price, number of recipients and policy cost). We restrict samples to A and B_2 municipalities. B_2 municipalities experience a cut in the coverage share that occurred in January 2018. Conversely, A municipalities do not experience any cut in the covering share. Our unit of observation is the municipality. We report standard error in vertical line with a 95% confidence interval. Standard errors are clustered to the commuting zone.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

E.5 DiD Estimation for A and C Areas

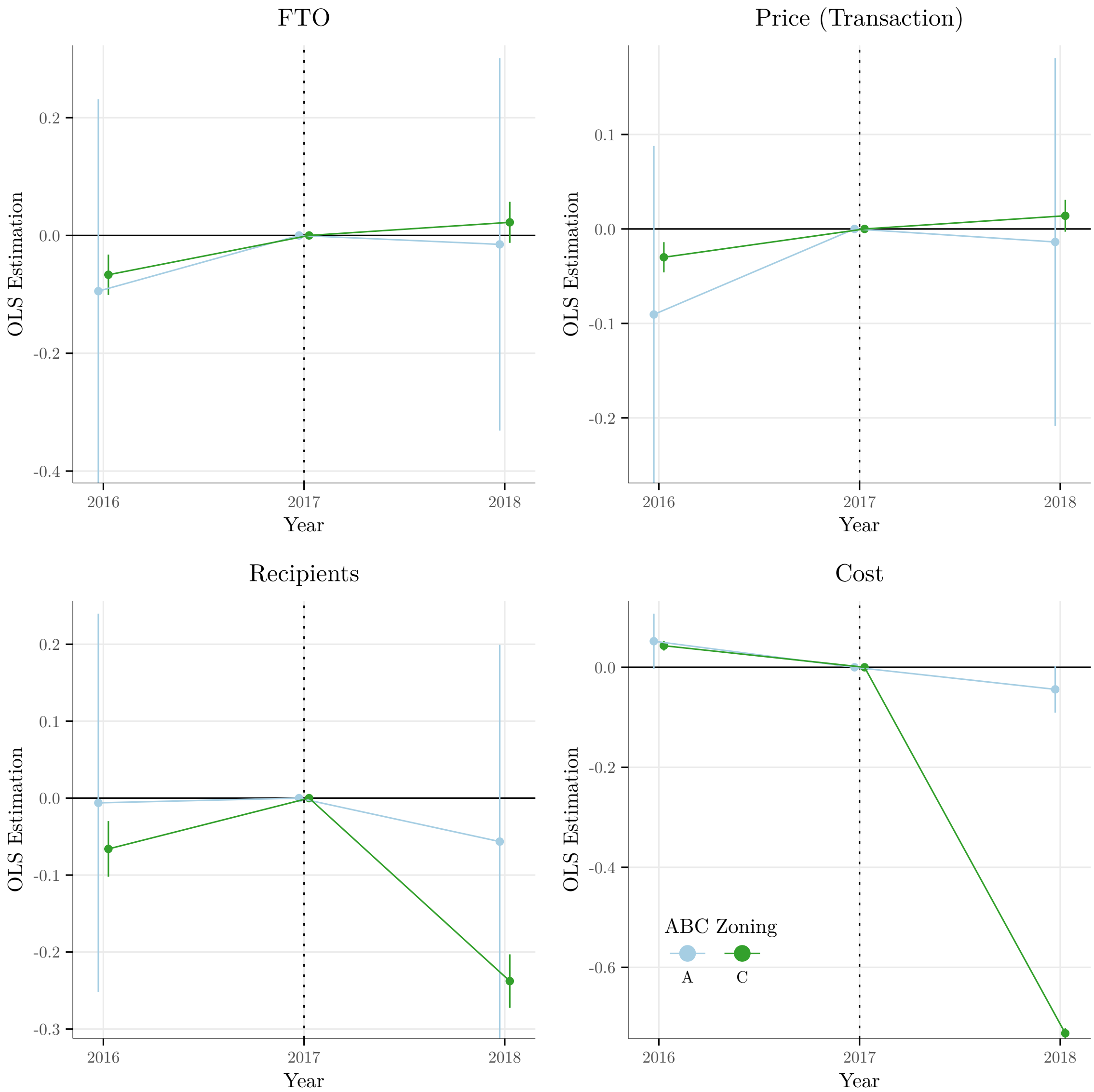


Figure E.5. OLS Results from DiD Identification Strategy restricted to A and C Areas

Notes: We represent results from event study for four outcomes of interest (number of first-time owners, average transaction price, number of recipients and policy cost). We restrict samples to A and C municipalities. C municipalities experience a cut in the coverage share that occurred in January 2018. Conversely, A municipalities do not experience any cut in the covering share. Our unit of observation is the municipality. We report standard error in vertical line with a 95% confidence interval. Standard errors are clustered to the commuting zone.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

E.6 DiD Estimation for A and B_1 Areas

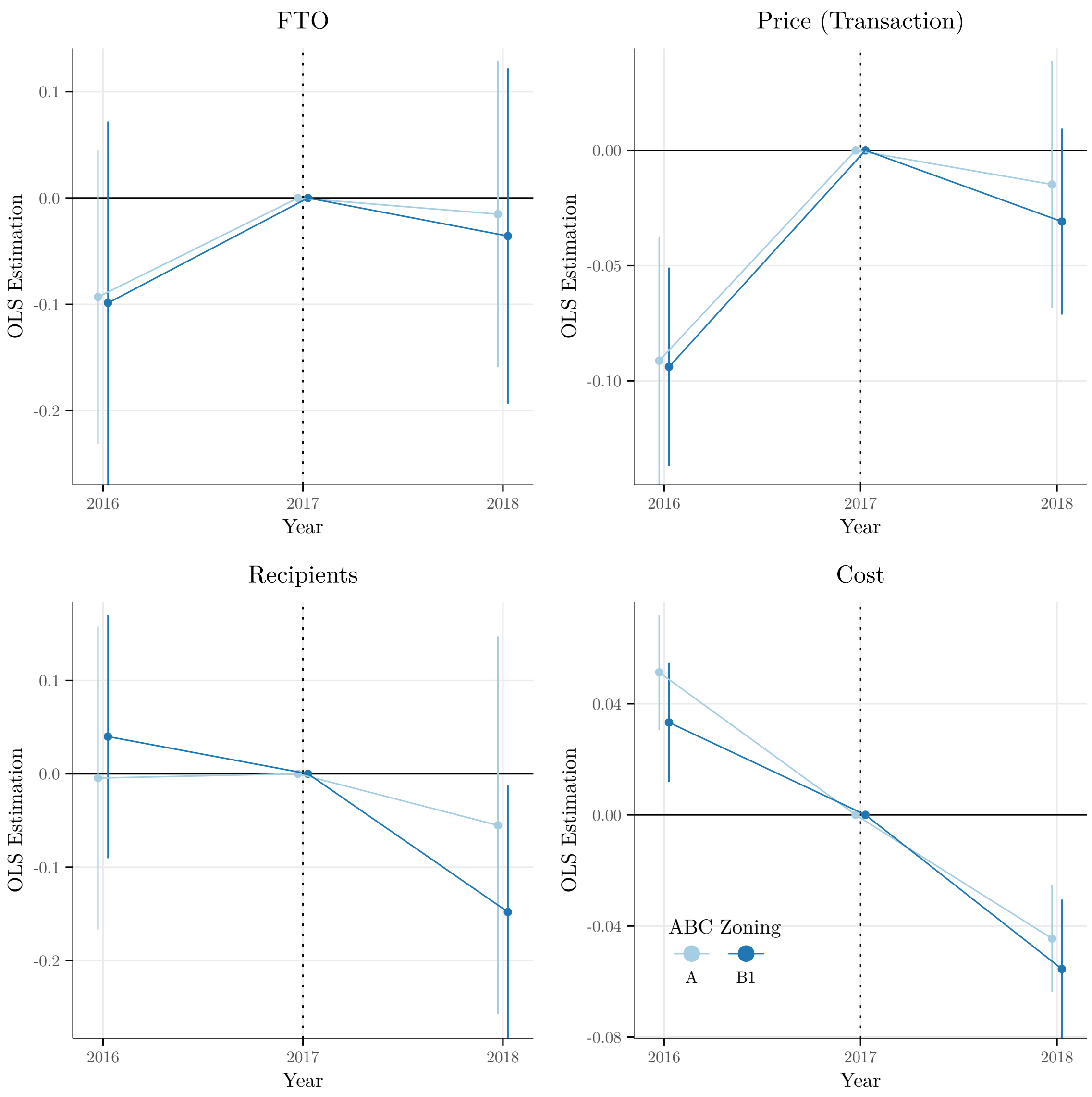


Figure E.6. OLS Results from DiD Identification Strategy restricted to A and B_1 Areas

Notes: We represent results from event study for four outcomes of interest (number of first-time owners, average transaction price, number of recipients and policy cost). We restrict samples to A and B_1 municipalities. Both groups of municipality experience no cut in the covering share. Consequently, it corresponds to a placebo test. We report standard error in vertical line with a 95% confidence interval. Standard errors are clustered to the commuting zone.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.

E.7 DiD Estimation for B_2 and C Areas

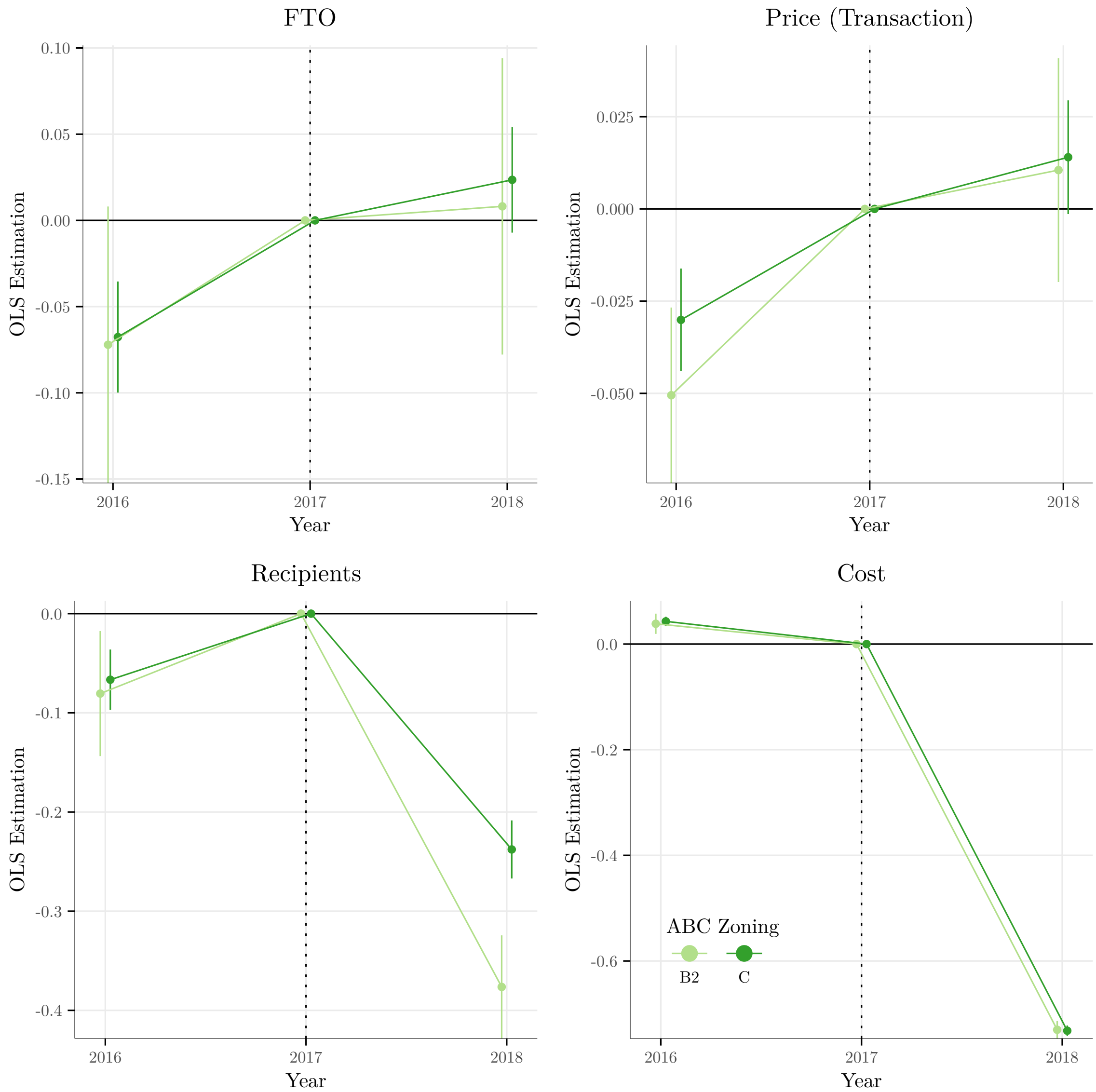


Figure E.7. OLS Results from DiD Identification Strategy restricted to B_2 and C Areas

Notes: Notes: We represent results from event study for four outcomes of interest (number of first-time owners, average transaction price, number of recipients and policy cost). We restrict samples to B_2 and C municipalities. Both groups of municipality experience a cut in the covering share. Considering potential heterogeneity in treatment effect, it cannot be considered as a placebo test. We report standard error in vertical line with a 95% confidence interval. Standard errors are clustered to the commuting zone.
Sources: Authors' Calculation based on SGFGAS, DV3F, *Fichiers Fonciers* and INSEE data.