Supporting Homeownership Cost-Effectively: Evidence from 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

November 21, 2023

Abstract

This article estimates the causal effect of the French policy of subsidised loans both on the number of new homeowners (the extensive margin) and on housing choice and price distortions (the intensive margin). The identification relies on the spatial and temporal variations of policy schemes over the last decade, controlling for endogenous treatment assignment by semi-parametric multi-level propensity scores. Our doubly robust estimations cannot rule out that the policy has no causal effect at the extensive margin, whereas it has significant effects at the intensive margin. Considering that policy makers value positively the extensive margin and negatively the intensive margin, we compute the returns on investment of government spending for counter-factual policy schemes and credit market conditions. Our simulations show that the current French policy is not cost-efficient for a wide range of monetary values of the two margins and cannot be improved by increasing overall budget cost.

JEL classification: H81, R21, R38

Keywords: Housing policy ; return on investment ; unconfoundedness ; generalized additive model ; spatial smoothing.

[*]UMR CESAER, INRAE, Institut Agro, Univ. Bourgogne Franche-Comté, F-21000, Dijon, France.
[†]CEREMA Hauts-de-France, F-59000 Lille, France.

We are grateful to Florence Goffette-Nagot, Carl Gaigné, Elisa Dienesch, Jean-Claude Driant, Jean Dubé, Sophie Legras, Benoît Schmutz and Olivier Dupré for precious comments. We are grateful to the 70^{th} AFSE seminar participants for insightful remarks. We also thank the CEREMA and the SGFGAS for database access. We thank the French Ministry of Ecological Transition for financial support through a 3-year PhD grant.

1 Introduction

In most developed countries, supporting homeownership is one of the largest housing policy (Andrews and Caldera Sánchez, 2011). These public interventions are motivated by the positive externalities expected to be associated with homeownership, usually thought to be larger than the negative ones (Coulson and Li, 2013). Homeownership increases children school performance (Green and White, 1997; Harkness and Newman, 2003), housing maintenance (Harding, Miceli, and Sirmans, 2000), and improves self-employment resulting from capital gains (Harding and Rosenthal, 2017) despite a risk of negative equity (Cunningham and Reed, 2013) and lower mobility (Green and Hendershott, 2001). Furthermore, this public support is a response to the affordability crisis (Fisher and Gervais, 2011; Carozzi, Hilber, and Yu, 2020), by possibly mitigating the growing importance of parental assistance (Boehm and Schlottmann, 1999; Lee et al., 2020) and providing a more equal access to housing throughout the population (Green and Vandell, 1999).

Despite these positive externalities expected from the extensive margin, these policies may have unintended consequences at the intensive margin, by distorting recipients' choices or reinforcing possibly undesirable high housing prices. Previous research indeed found that such policies favour opportunistic behaviour as subsidized households purchase more expensive (Davis et al., 2020) and larger houses (Hanson, 2012), and that their location choices change (Daminger and Dascher, 2022). The policy-induced increase in demand has also been shown to lead to higher housing prices (Martin and Hanson, 2016; Kunovac and Zilic, 2022) and to lower supply of affordable housing (Sommer and Sullivan, 2018). This aligns with the evidence that decreased interest rates result in an increase in housing prices (Barone et al., 2021). In regards to either small or insignificant effects at the extensive margin (Glaeser and Shapiro, 2003; Hanson, 2012; Hilber and Turner, 2014; Bäckman and Lutz, 2020; Kunovac and Zilic, 2022), the unintended effects seem to dominate the expected ones in the literature. However, to the best of our knowledge, the overall cost-effectiveness of supporting homeownership across both margins is not well established.

Whereas homeownership support is based on mortgage contracts in almost all countries, determining the balance between extensive and intensive margin effects is challenging due to the variations in policy designs. We can distinguish two main approaches in existing policies. In the first, the subsidy is channelled through income tax as households deduct their payments of interests from their taxable income. The main US housing policy (the Mortgage Interest Deduction, Glaeser and Shapiro, 2003; Hilber and Turner, 2014) follows this pattern, such as policies of other countries like Belgium (Hoebeeck and Inghelbrecht, 2017) and Denmark (Gruber, Jensen, and Kleven, 2021). In these cases, the subsidy amount depends on household income through differentiated marginal tax cuts. In the second approach, subsidized loans cover a share of the total amount of the mortgage independently of the income of the recipients. Treatment intensity is capped by an eligible share of the overall mortgage value and a specific population is targeted by income thresholds. Countries such as England (Carozzi, Hilber, and Yu, 2020), Croatia (Kunovac and Zilic, 2022), Germany (Daminger and Dascher, 2022), and the US (in a complementary policy to support homeownership after the 2008 crisis, Hembre, 2018) adopt this design. The French IFL (interest-free loan) policy follows this second approach with, given household eligibility, a public support that depends exclusively on the characteristics of the housing loan (overall value, loan maturity, and interest rate).

In this paper, we estimate the cost-effectiveness of supporting homeownership from the French policy. We account for both targeted effects (the number of additional homeowners) and the unintended effects (price capitalisation, housing choice distortion). Our cost-effectiveness definition depends on a wide range of monetary valuation for each margin according to policy objectives. We disentangle the cost-effectiveness according to source of subsidy variation (the *primitives* variables in the remaining of the paper) that channel the subsidy variation. We identify two variables that are under the control of policy makers (loan ceiling value and coverage share, which are related to the policy design) and two variables that result from credit market conditions (loan maturity and interest rate). Despite the differences in policy designs across countries, we consider these primitive policy variables are common to all subsidy schemes and would allow to compare different policies supporting homeownership between them.

To identify and estimate the needed statistics to infer cost-effectiveness, we employ an intensity-varying treatment framework as it allows to disentangle the source of variation in treatment intensity, depending on whether it is driven by policymaker decisions or macroeconomic conditions. Our intensity-varying framework relies on spatial variations of subsidy. The French policy adopts a four-level classification to define tenseness which in turn determine the policy primitives. Subsidy variations are consequently endogenous to municipality characteristics which need to be addressed to perform causal inference. Yet, the ABC classification still recovers an arbitrary part, resulting from subjectivity in the assignation process as outlined by the French Court of Auditors. We leverage a selection-on-observables approach to isolate the part due to the arbitrariness of the ABC zoning. The initial stage estimates the probability of each municipality to receive each level of treatment according to observable characteristics using a generalised propensity score (Imbens, 2000). To consider the ordered structure of the treatment intensity design, we employ an ordinal logit. In the second stage, we regress outcomes of the extensive (new homeowners) and intensive (house values, recipient housing choices) margins using an inverse probability weighting mechanism derived from GPS. We estimate treatment effects that are relevant to policy by examining the impact of a particular treatment level on a municipality group when compared to their receiving another level of treatment.

Inverse probability weighting give more weight to observations with higher counterfactual power. Finally, we relate estimated effects for bilateral combinations in treatment level to primitive variations. It provides dose-responses functions that allows us to compute the cost-effectiveness of the policy, using exogenous implicit valuation of extensive and intensive margin effects by policymakers.

Our approach does not rely on a specific natural experiment exploited by strategies such as difference-in-difference models (e.g Davis et al., 2020) or difference-in-discontinuity designs (e.g. Carozzi, Hilber, and Yu, 2020). Indeed, such strategies typically hinder the applicability of the results to other policy contexts given common LATE issues (for a discussion, see Deaton, 2010; Imbens, 2010). Consequently, on the one hand, it limits the possibility to distinguish the sources of subsidy variation. Natural experiments usually arise from policy reforms that affect either specific primitives or all primitives simultaneously.[1] Hence, it limits the possibility to distinguish the effect of each primitive. On the other hand, it does not allow estimating the effect of subsidy variations on specific populations, as these strategies exploiting natural experiments primarily provide average treatment effects on treated individuals (Roth et al., 2023). Therefore, findings may have limited external validity as they depend on the policy reform (which differs considerably across countries) and the macroeconomic conditions in which it takes place.

Our approach to ensure the internal validity of the selection on observable approach is threefold. First, we collect a significant number of pre-treatment variables that are likely to determine the degree of housing market tenseness. These variables include both demand-side (e.g. population density, income, socio-economic status) and supply-side (e.g. housing construction, past prices, past neighbourhood prices) pre-treatment variables. Second, as unobserved variables may violate the assumption of unconfoundedness (Lewbel and Yang, 2016), we account for non-observable spatial variables (Gilbert et al., 2023) by introducing smooth functions of the spatial coordinates of homeowners' locations in a semiparametric generalised additive model (GAM, Wood, 2017). For instance, we expect housing market tenseness to depend on geographical constraints (Saiz, 2010), housing supply elasticity (Accetturo et al., 2021) restrictions on land development (Turner, Haughwout, and Van der Klauw, 2014), or demand for particular amenities (Bayer et al., 2016). Third, these variables are also included in the outcome regression, thus providing a doubly robust estimation. The consistency of policy-relevant treatment effects can be ensured if the first or second step of the identification strategy is well-defined (Robins and Rotnitzky, 1995; Słoczyński and Wooldridge, 2018). Whereas our results depends on a selection on observables approach, our results are also consistent with a difference-in-difference approach,which reinforces the credibility of our identification strategy. Finally, a placebo analysis does not allow to reject the unconfoundedness assumption.

[1] For instance, most French policy reforms about the IFL policy affect the covering share.

Our contribution to the literature on policy support for homeownership is threefold. First, our identification strategy allows us to disentangle the effect of each primitive. Consequently, we can distinguish their contribution to both extensive and intensive margin effects. Using this approach, we derive policy recommendations that can be readily generalised across countries considering the external validity of our identification strategy. Second, we consider the cost-effectiveness of the policy by taking into account both the targeted and unintended effects of the policy. The balance between these two margins results from policymakers' implicit valuation of the effects that stem from the policy objectives. Third, we find that subsidy variations produce most of their effects on the intensive margin, while we cannot rule out that they affect the number of homeowners. Therefore, increasing the subsidy through policy-controlled primitives is likely to be inefficient. The cost-effectiveness of increasing public spending depends mainly on a positive valuation of price capitalisation by policymakers, which is unlikely.

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 objectives that define policy relevance, taking into account the externalities 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 following Section 5 provides the empirical results from the two-step approach and the 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 subsidized loan with no interest to pay for a given share of a maximum ceiling value of their overall housing credit. The cost for public finance equals the sum of interests not at charge for recipients, as the government consents a tax reduction to private lenders (banks) supplying such contracts. As shown in Appendix A.1 of Online Appendix (OA), the number of recipients since the beginning of the policy is about 3.2 millions, for an overall budget cost of about 26.1 billions euros (about 8,000 euros by recipient). The policy excludes a small share of high-income households (about 10% of richer tenants according to Sotura, 2020) and targets specifically the newly built housings (although existing housings are eligible conditionally on renovation 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.[2]

The mechanism that defines IFL intensity has remained stable over time and is mainly based on four primitive parameters. 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 primitives under control for the government to implement the IFL policy.[3] The main difference between both policy-controlled primitives is that a variation does not affect the same population. Covering share affect the entire recipients population, while the ceiling value variations affect the most expensive operations.[4] The budget cost for the policy also depends on two other primitives from the credit market, which 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. As shown in Appendix B.1 of OA, the budget cost c of a IFL contract equals the subsidy-equivalent for the recipient, which is:

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

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

2.2 The ABC Zoning

Both ceiling values and eligibility shares depend on the location of IFL contracts, from an exhaustive and mutually exclusive classification of French municipalities based on the tenseness of the housing markets.[5] This ABC zoning introduces four ordered degrees of tenseness, from C the lowest level, to B_2 and B_1 the intermediate levels, and A the

[2]For the French market, an usual condition to access to a classical loan is that reimbursement payments cannot be higher than 33% of income. Consequently, this also restricts access to the IFL policy.

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

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

[5]According to official documents, tenseness is defined "from the imbalance between the housing supply and the housing demand" (French Ministry of Ecological Transition).

highest level. This official zoning was updated four times since its introduction in 2003, the latest update of October 2014 is stable for the 2015–2019 period under study. The choice of municipality as 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. It represents the smallest jurisdiction in France, with 34,970 units in 2019.

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 the classification assigns quite similar municipalities to different levels: the French administration in charge of monitoring public expense noted in 2012 the lack of transparency of the zoning (Cour des Comptes, 2012). It concluded that the zoning does not depend exclusively on objective characteristics, suggesting potential subjectivity in the assignment. Most existing quasi-experimental approaches dealing with the endogeneity of IFL treatment assignment rely on the arbitrariness of this zoning (Labonne and Welter-Nicol, 2015; Beaubrun-Diant and Maury, 2021; Chareyron, Ly, and Trouvé-Sargison, 2021).

Beyond variations introduced by the design of the IFL policy, credit market conditions (interest rate and loan maturity) also vary over the considered period (Panel C of Table 1). These variations are not clearly correlated with the ABC zoning, as they mainly depend on the exogenous economic context. From the variations of the four primitives, the budget cost of the policy experienced sizeable changes across multiple cross-sectional and temporal dimensions, while it is a national policy. Hence, the causal evaluation of the IFL policy has to be considered with a varying treatment intensity framework rather than a more classical counter-factual experiment of policy removal.

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

Variable	Period	Country	*ABC Zoning Areas* A	B_1	B_2	C
			A - Pre-treatment variables			
Number of Municipalities	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 clas sified 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.

2.3 Data

We aggregate three exhaustive individual data sources at the municipal level ($N = 34{,}970$) and match them with demographic data. We remove the municipalities of the Corsican island due to geographic constraints related to the spatial smoothing terms we introduce in the econometric specification (360 observations) and municipalities of the *Alsace-Moselle* region (1,605 observations) as transaction data are missing for this region due to administrative and historical 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 variables.[6] Consequently, potential selection bias on our estimation is at worse marginal as these observations would have small impact due to low counterfactual power.

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

[6]Considering a selection model within municipalities C that explains whether latent variable is estimated, 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.

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

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.

3 Cost-Effectiveness Measures

We consider the cost-effectiveness of subsidy variations from a policymaker perspective according to credible policy objectives. We assume that policymaker objectives can be proxied through the marginal valuation of respectively one additional homeowners, price capitalisation on the housing market and distortion of recipients' housing choices. We first detail our extensive and intensive margin definitions related to the definition of cost-effectiveness. We then present first-order response to primitive variation and highlight needed statistics to compute cost-effectiveness measures in regard with policy objectives.

3.1 Policy Objectives

We develop a framework for evaluating IFL policy that directly derives from the potential objectives set by policymakers. We consider that policymakers' objectives relate to the definitions of the extensive and intensive margins. 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. On the other hand, the intensive margin consists of two different effects. First, 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 $\widetilde{P}$). Second, policy-induced demand could affect the average purchase price of the whole transaction, especially in the context of low supply elasticity. We consider that the impact on the dynamics of the housing market is of interest considering the context of affordability issues. We then consider the effect of subsidy variation on average housing price (noted V).

In addition to the objectives set by policymakers, we expect policy costs to be relevant for efficiency. The total cost of the policy depends on the number of recipients (denoted $\widetilde{N}$) and the average cost per recipient (denoted $\widetilde{C}$). The latter quantity is directly related to the primitive that defines the amount of subsidy at the household level. We assume that policymakers try to achieve their objectives while minimising the total cost of the policy.

3.2 Empirical Definition

Assuming that policy objectives are reflected through the monetary valuation of both intended and unintended effects, we define policy efficiency from the perspective of the policymaker, denoted E, as

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

with ϕ, ψ and θ representing respectively the policymaker's valuation of an additional homeowner, price capitalisation 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 policy efficiency by varying the subsidy. Following the definition of efficiency, we derive first order Equation 2 according to the primitive source of subsidy variation a, such as

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

Then, the sign of e_a is the main criterion to determine whether increasing public spending by primitive a is efficient according to the policy objectives. Indeed, if e_a is positive, the policymaker would consider that the benefits are greater than the drawbacks and public expenditure.

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 papers address the valuation of externalities associated with homeownership (Coulson and Li, 2013), it is not clear how policymakers value unintended effects relative to intended ones. We then discuss the cost-effectiveness of increasing public spending through IFL subsidies, following scenarios that differ according to how these effects are valued from the perspective of policymakers. In credible scenarios, additional homeowner is positively valued as it is the policy stated objective, while price capitalisation on the housing market is likely to be negatively valued in regards with affordability concerns. Finally, the valuation of distortion 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 primitive 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 cost-effectiveness of raising IFL subsidy through each primitive. Moreover, we must estimate similar quantity for average cost per recipient and number of recipients are needed to assess the impact on policy cost.

3.3 Dose-response Functions

We recover the marginal effect 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 primitives of interest. Variations of the IFL policy across the four ABC zones and the three periods define a multi-valued treatment taking $G = 12$ levels. Let g denote a level of treatment and T_g a dummy variable 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 primitive. We exploit this structure of the IFL policy to map policy-relevant treatment effects to primitive 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 primitives 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 impact on policy objectives.

4 Identification strategy

4.1 Identifying Assumptions

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

Assumption 1 *Weak Unconfoundedness.*

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

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 multi-valued treatments as in the IFL policy:

Assumption 2 *Overlap*

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

Under the two previous assumptions, Słoczyński and Wooldridge (2018, Lemma 3.2) demonstrate that counter-factual treatment effects can be identified from usual data. 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 concept of tenseness of the housing market used to establish ABC zoning, we define a unobserved latent variable η_i^* crossing thresholds to determine the classification of municipalities. The propensity for a municipality i to be high in the hierarchy depends on the J pre-treatment variables x_{ji} used to proxy the political decision, a bivariate smooth function of the geographical coordinates of its centroid $\mathbf{z}_i$ (longitude and latitude, Gilbert et al., 2023), and a random term ε_i representing the arbitrary part of the zoning. This latter term is assumed to be logistically distributed to produce an ordered logit model. The latent variable describing 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.

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. The main difference is that each outcome Y is modeled 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 shrunk during the estimation procedure. 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 cost-effectiveness 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 $\hat{y}_\ell(g') \equiv \hat{\alpha}^Y_{g'} + \sum_{j=1}^{J} \hat{f}^Y_{g'j}(x_{j\ell}) + \hat{h}^Y_{g'}(z_\ell)$ comes from the estimation of the outcome Y for the subset of municipalities that receive treatment g'. It is simply the predicted outcome values for the whole population of municipalities with $\ell = 1, \ldots, N$.

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 our results section with the measures of cost-effectiveness resulting from IFL subsidy variations according to each primitive.

5.1 First-stage Models from ABC Zoning

We estimate a semi-parametric ordered logit GAM on the ABC classification of municipalities, given a set of pre-treatment variables on housing supply and demand. To reach our identification restriction of unconfoundedness, we include a maximum of variables that can be used by the French administration to construct this classification, including the pre-treatment unitary house prices of the municipality and the neighbouring ones. We include the geographical coordinates of the centroids of each municipality through bivariate smoothing splines to control for spatial confounders (Gilbert et al., 2023). As allowed by the GAM framework, all variables enter semi-parametrically with a degree of smoothing that is endogenously shrunk by the penalised estimation procedure. Table 2 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.

The results present both high pseudo-R^2 and share of good predictions in the bottom panel

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

in rows. 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 modeling 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, most remaining errors (77.2% of overall errors) concern municipalities that are commonly used as the basis for natural experiments to assess housing policy, exploiting spatial discontinuity designs (Chapelle, Vignolles, and Wolf, 2018; Bono and Trannoy, 2019). The underlying assumption is that the variations in the treatment derived from the ABC classification are as good as random for municipalities close to the area boundary. Furthermore, 29.2% of the prediction errors concern municipalities that experience a change in treatment level following the 2014 reform. Our prediction is consistent with the previous classification for 72.7% of these errors. Therefore, our prediction errors mainly concern observations that are considered to be quasi-randomly assigned in the ABC classification or that have experienced a recent change in classification, which strengthens the credibility of the GPS estimate.

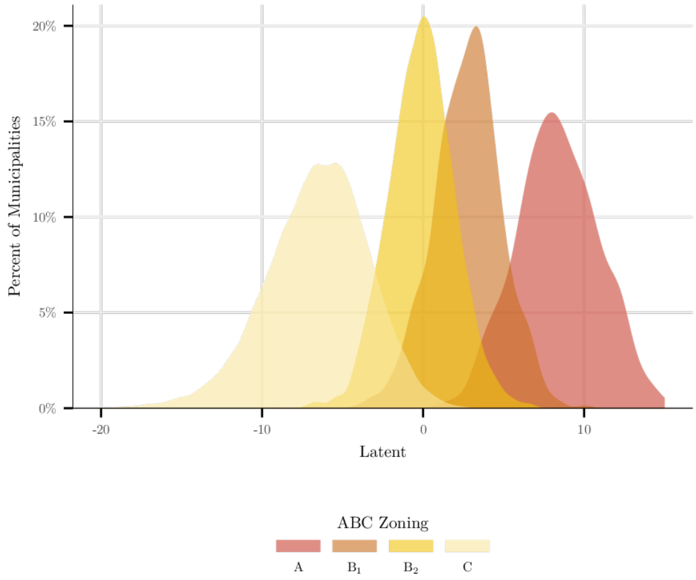
Since overlap is crucial to recover consistent effects and likely to be reduced for high-dimension model variables (D'Amour et al., 2021), we compare the distribution of the latent variable underlying the classification process (Fig. 1). Latent distributions follow the ordered structure of the ABC classification, 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 Appendix C.6 to assess the statistical power of pre-treatment variables. We report spatial smoothing splines functional forms in Appendix C.5, as it represents our main contribution to control unobserved heterogeneity.

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

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.

pretreatment 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 primitive 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 primitive 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 primitive values between treatment level g and g'. Since we consider small primitive 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

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

(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 primitive used for policy reforms, it has no significant effect on the number of homeowners (column 1, Table 4). Thus, according to our results, increasing the covering share is unlikely to achieve policy objectives, as the number of homeowners does not increase significantly with the amount of subsidy. Meanwhile, increasing the ceiling value (the second primitive that policymakers can control) has no significant effect on the number of homeowners, while it increases the number of policy recipients (+8.1%, i.e. *almost 16,000 recipients at the country level*). Given the 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. Our results show that it causes a shift in demand from unsubsidised to subsidised housing without affecting tenure decisions at the aggregate level. 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

Table 4. 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	0.003	0.029***	0.004	0.005	-0.034**	0.002	0.003	-0.003	0.041***
	(0.027)	(0.007)	(0.004)	(0.008)	(0.013)	(0.002)	(0.003)	(0.002)	(0.002)
Ceiling Value	-0.014	-0.004	-0.007	-0.005	0.081***	-0.013***	-0.040***	0.029***	-0.006**
	(0.026)	(0.006)	(0.004)	(0.007)	(0.009)	(0.002)	(0.003)	(0.002)	(0.002)
Interest Rate	-0.001	0.000	-0.002	-0.004**	0.018***	-0.005***	-0.008***	0.004***	-0.001
	(0.007)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
Loan Maturity	0.007	-0.000	0.000	-0.002	-0.021***	0.005***	0.022***	-0.018***	0.022***
	(0.016)	(0.005)	(0.003)	(0.005)	(0.005)	(0.001)	(0.002)	(0.001)	(0.001)
Constant	0.000	-0.000	0.000	0.000	-0.174***	0.010***	-0.015**	0.042***	-0.013**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.021)	(0.004)	(0.006)	(0.005)	(0.006)
R^2	0.169	0.529	0.542	0.640	0.461	0.285	0.411	0.479	0.888
Adj. R^2	0.062	0.468	0.483	0.594	0.391	0.193	0.335	0.412	0.874
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 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.
*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

market, while opportunistic behaviour is favoured in a context where credit conditions are less favourable for households (higher interest rates, shorter loan duration).

Whereas the covering share has at best a weak effect on the number of homeowners, it has a significant effect on the intensive margin. A one-unit increase in the covering share raises the price of all transactions by 2.9%, leading to a significant unintended effect on the housing market. Thus, the variation in subsidy resulting from the covering share specifically affects the intensive margin, with no significant effect on the distortion of recipients' housing choices. Meanwhile, raising the ceiling value has no significant effect on the housing market, while the effect on recipients' housing choices is more mixed. Indeed, while it has a negative impact on both purchase price and area (columns 6 and 7, Table 4), the price per area unit covaries positively with the ceiling value (column 8). We interpret these results as a potential change in the location choice of recipients, as increasing the ceiling value has a significant effect on the financial support for the most expensive dwellings, while leaving the financial support for the cheapest ones unaffected (for more details see Appendix B.2). Thus, raising the cap could increase the attractiveness of operations in the most expensive areas. Most of the effects for the primitive variations controlled by policymakers concern the intensive margins, with differentiated effects on recipients' housing choices. Finally, policy recipients are also sensitive to the characteristics of the loan, as an increase in the interest rate (respectively, the duration of the loan) causes recipients to reduce (respectively, increase) their purchase price.

These results are consistent with alternative results from natural experiments. Using a difference-in-difference technique (refer to Appendix E for the alternative results), we investigate the 2018 reform, which changes 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, with no significant effect on the number of homeowners. However, our study extends these results by accounting for the primitive source of variation and overcoming the LATE issue.

5.4 Simulation According to Policy Objectives

We finally provide counterfactual simulations that produce cost-effectiveness measures for increasing public spending according to each primitive. For comparison purpose, we choose to simulate primitive variations that have a similar impact on the total cost of the policy. We choose to 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 primitive 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 negative 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 primitive, although it is not statistically significant for some primitives.

Policymakers' assessment of individual effects at both ends is inherently unobservable. We only consider cases where increasing the number of homeowners is positively valued, as this is the main objective. In line with the results of Coulson and Li (2013), we set the value of an additional homeowner at 10k euro.[7] We also consider a range of price capitalisation valuations from -2k and $+1$k per 1,000 euro of price capitalisation, with -1k and 0k as intermediate values, to discuss cost-effectiveness according to the population policymakers are trying to benefit. Finally, for housing distortion cases, we consider a range from -1k to 1k per 1k euro of increase in the recipient's purchase price, including the indifference situation ($\theta = 0$). We report the results in Table 5.

[7]This corresponds to the average subsidy amount at the country level. However, our results are robust to the specification of the value of an additional homeowner, although it affects statistical significance.

Table 5. Cost-Effectiveness Measures According to Different Valuation of Extensive and Intensive Margin Effects from a Policy Maker Perspective

	Efficiency Measure with $\phi = 10k$											
	ψ = -2k			ψ = -1k			ψ = 0k			ψ = 1k		
θ	-1k	0k	1k	-1k	0k	1k	-1k	0k	1k	-1k	0k	1k
Covering Share	-22.20	-21.45	-20.69	-10.45	-9.69	-8.94	1.30	2.06	2.81	13.05	13.81	14.57
	(24.9)	(25.2)	(25.4)	(24.1)	(24.4)	(24.6)	(23.6)	(23.8)	(24.1)	(23.4)	(23.6)	(23.9)
Ceiling Value	-1.30	-1.71	-2.12	-1.42	-1.83	-2.24	-1.55	-1.96	-2.36	-1.67	-2.08	-2.49
	(1.8)	(1.8)	(1.8)	(1.7)	(1.8)	(1.8)	(1.7)	(1.7)	(1.8)	(1.7)	(1.8)	(1.8)
Interest Rate	-0.80	-1.39	-1.99	-0.76	-1.36	-1.96	-0.73	-1.33	-1.93	-0.70	-1.30	-1.90
	(1.9)	(2.0)	(2.0)	(1.9)	(1.9)	(2.0)	(1.9)	(1.9)	(1.9)	(1.9)	(1.9)	(1.9)
Maturity	50.94	76.60	102.25	49.28	74.94	100.59	47.62	73.27	98.93	45.96	71.61	97.27
	(178.4)	(179.1)	(180.0)	(172.2)	(172.9)	(173.8)	(169.5)	(170.4)	(171.3)	(170.7)	(171.6)	(172.5)
	Efficiency Measure with $\phi = 1k$											
	ψ = -2k			ψ = -1k			ψ = 0k			ψ = 1k		
θ	-1k	0k	1k	-1k	0k	1k	-1k	0k	1k	-1k	0k	1k
Covering Share	-24.95***	-24.20***	-23.44***	-13.20***	-12.45***	-11.69**	-1.45	-0.69	0.06	10.30***	11.06***	11.81***
	(6.2)	(6.3)	(6.6)	(3.8)	(3.9)	(4.2)	(2.3)	(2.4)	(2.8)	(3.4)	(3.4)	(3.6)
Ceiling Value	-0.44	-0.85	-1.26	-0.56*	-0.97	-1.38	-0.69**	-1.10	-1.50**	-0.81	-1.22	-1.63**
	(0.4)	(0.4)	(0.4)	(0.3)	(0.3)	(0.3)	(0.2)	(0.2)	(0.2)	(0.2)	(0.3)	(0.3)
Interest Rate	-0.50	-1.10	-1.70	-0.47**	-1.07	-1.66**	-0.43***	-1.03	-1.63***	-0.40***	-1.00	-1.60**
	(0.4)	(0.4)	(0.4)	(0.3)	(0.3)	(0.3)	(0.2)	(0.2)	(0.2)	(0.2)	(0.3)	(0.3)
Maturity	-15.90	9.75	35.41	-17.57	8.09	33.75	-19.23	6.43	32.08*	-20.89	4.77	30.42
	(54.2)	(54.1)	(54.4)	(31.0)	(31.0)	(31.6)	(16.7)	(17.0)	(18.4)	(29.9)	(30.2)	(31.2)

Notes: Exploiting coefficients derived from the dose-response function (Table 4), we calculate cost-effectiveness for a cost-normalised increase of the overall IFL budget using Equation 3 for the four sources of primitives. 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 cost-effectiveness 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.
*** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$

Our cost-effectiveness measures are robust to the valuation of an additional homeowner from a policy maker perspective. It does, however, substantially affect the statistical significance of our results. Indeed, running our simulations with $\phi = 1k$ (bottom panel, Table 5) reduces the standard errors and leads to more precise results. We explain this behaviour by the lack of a significant effect for the number of homeowners, regardless of the primitive (column 1, Table 4). We therefore discuss the cost-effectiveness results for $\phi = 1$k.

From a policymaker's perspective, increasing the subsidy amount is likely to be inefficient according to our simulations. Indeed, regardless of the valuation of the distortion of recipients' housing choices (denoted by θ) and price capitalisation (denoted by ψ), increasing the ceiling value is inefficient as $e_a < 0$. This means that the ceiling value mostly affects the policy cost. The only situation with a positive and significant cost-effectiveness measure corresponds to the situation where price capitalisation is favourable. Assuming that these effects are targeted, increasing the covering share appears to be highly efficient ($e_a > 10$), regardless of how policymakers assess the distortion of recipients' housing choices. However, the positive valuation of price capitalisation favours current homeowners rather than homeownership candidates. If price capitalisation is undesirable, increasing the covering share is inefficient. Thus, the decision to increase the covering share depends to a large extent on the population that policymakers wish to favour. In addition, raising the interest rate reduces the efficiency of the policy, regardless of the policy objectives, as the impact on the cost of the policy outweighs the potential benefits.

In conclusion, the cost-effectiveness of raising the subsidy through the covering share mainly depends on the population that policymakers are trying to target. If policymakers are targeting potential first-time homeowners, increasing public spending through a policy-driven primitive seems inefficient given the lack of an extensive margin effect combined with pronounced price capitalisation and impact on policy cost. The valuation of the distortion of recipients' housing choices has a weak effect on efficiency. Conversely, if policymakers are targeting current homeowners, an increase in public spending trough the covering share is likely to be efficient as price capitalisation becomes desirable.

6 Conclusion

The French IFL policy aims to increase the number of homeowners through interest cuts for first-time owners. We leverage spatial variation of treatment using selection-on-observables restriction 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 variables. From the linear dose-response functions, we discuss the policy results for different scenarios that differ according to the policymaker valuation of extensive and intensive margins effects. We then derive conclusions about the cost-effectiveness of raising the IFL subsidy.

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 primitives 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, which are directly related to inflationary effects and housing choice distortion. Given the growing concerns about affordability and potential indifference for housing choices distortion, we expect these effects not to be targeted by policymakers, leading to the inefficient result of increasing the IFL subsidy.

However, if policymakers aim to favour current homeowners, increasing the subsidy is likely to be efficient. Hence, the cost-effectiveness of additional public spendings into the IFL depends on the population that policymakers are trying to benefit. For example, if the targeted population consists of current homeowners (leading to positive valuation of price capitalisation effect), increasing the coverage share will generate benefits rather than damage. The efficiency therefore crucially depends on the population of interest (the entire society, the recipients or the current homeowners).

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. Since French administration aims to reduce land consumption through higher constraints (the net-zero-artificialisation implementation), policy contribution is of concern.

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.

Andrews, Dan and Aida Caldera Sánchez (2011). "The Evolution of Homeownership Rates in Selected OECD Countries: Demographic and Public Policy Influences". In: *OECD Journal: Economic Studies* 2011.1, pp. 1–37.

Bäckman, Claes and Chandler Lutz (2020). "The Impact of Interest-Only Loans on Affordability". In: *Regional Science and Urban Economics* 80.103376.

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.

Boehm, Thomas P. and Alan M. Schlottmann (1999). "Does Home Ownership by Parents Have an Economic Impact on Their Children?" In: *Journal of Housing Economics* 8, pp. 217–232.

Bono, Pierre-Henri and Alain Trannoy (2019). "The Impact of the 'Scellier' Income Tax Relief on Building Land Prices in France". In: *Economie et Statistique / Economics and Statistics* 507d, pp. 91–114.

Carozzi, Felipe, Christian A. L. Hilber, and Xiaolun Yu (2020). "On the Economic Impacts of Mortgage Credit Expansion Policies: Evidence from Help to Buy". In: *CEPR Discussion Paper* DP14620.

Chapelle, Guillaume, Benjamin Vignolles, and Clara Wolf (2018). "Impact of a Housing Tax Credit on Local Housing Markets: Evidence from France". In: *Annals of Economics and Statistics* 130, pp. 101–131.

Chareyron, Sylvain, Tidiane Ly, and Yohann Trouvé-Sargison (2021). "Ownership Incentives and Housing Affordability: Evidence from France". 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.

Daminger, Alexander and Kristof Dascher (2022). "Homeowner Subsidy Repeal and Housing Recentralization". In: *Land Economics*, 081821–0095R.

Davis, Morris A., Stephen D. Oliner, Tobias J. Peter, and Edward J. Pinto (2020). "The Impact of Federal Housing Policy on Housing Demand and Homeownership: Evidence from a Quasi-Experiment". In: *Journal of Housing Economics* 48.101670.

Deaton, Angus (2010). "Instruments, Randomization, and Learning about Development". In: *Journal of Economic Literature* 48.2, pp. 424–455.

Fisher, Jonas D. M. and Martin Gervais (2011). "Why Has Home Ownership Fallen Among the Young?" In: *International Economic Review* 52.3, pp. 883–912.

Gilbert, Brian, Abhirup Datta, Joan A. Casey, and Elizabeth L. Ogburn (2023). "A Causal Inference Framework for Spatial Confounding". In: *arXiv:2112.14946*.

Glaeser, Edward L. and Jesse M. Shapiro (2003). "The Benefits of the Home Mortgage Interest Deduction". In: *Tax Policy and the Economy*. MIT Press. Vol. 17. NBER Chapters. National Bureau of Economic Research, Inc, pp. 37–82.

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

Green, Richard K. and Kerry D. Vandell (1999). "Giving Households Credit: How Changes in the U.S. Tax Code Could Promote Homeownership". In: *Regional Science and Urban Economics* 29.4, pp. 419–444.

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

Gruber, Jonathan, Amalie Jensen, and Henrik Kleven (2021). "Do People Respond to the Mortgage Interest Deduction? Quasi-experimental Evidence from Denmark". In: *American Economic Journal: Economic Policy* 13.2, pp. 273–303.

Hahn, Jinyong (1998). "On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects". 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.

Hilber, Christian A. L. and Tracy M. Turner (2014). “The Mortgage Interest Deduction and Its Impact on Homeownership Decisions”. In: *Review of Economics and Statistics* 96.4, pp. 618–637.

Hoebeeck, Annelies and Koen Inghelbrecht (2017). “The Impact of the Mortgage Interest and Capital Deduction Scheme on the Belgian Mortgage Market”. In: *Working Paper Research: Bank of Belgium* 327.

Imbens, Guido W (2010). “Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)”. In: *Journal of Economic Literature* 48.2, pp. 399–423.

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

Kunovac, Davor and Ivan Zilic (2022). “The Effect of Housing Loan Subsidies on Affordability: Evidence from Croatia”. In: *Journal of Housing Economics* 55.101808.

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

Lee, Hyojung, Dowell Myers, Gary Painter, Johanna Thunell, and Julie Zissimopoulos (2020). “The Role of Parental Financial Assistance in the Transition to Homeownership by Young Adults”. In: *Journal of Housing Economics* 47, p. 101597.

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.

Roth, Jonathan, Pedro H.C. Sant’Anna, Alyssa Bilinski, and John Poe (2023). “What’s Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature”. In: *Journal of Econometrics* 235.2, pp. 2218–2244.

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.

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.

Turner, Matthew, Andrew Haughwout, and Wilbert Van der Klauw (2014). “Land Use Regulation and Welfare”. In: *Econometrica* 82.4, pp. 1341–1403.

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

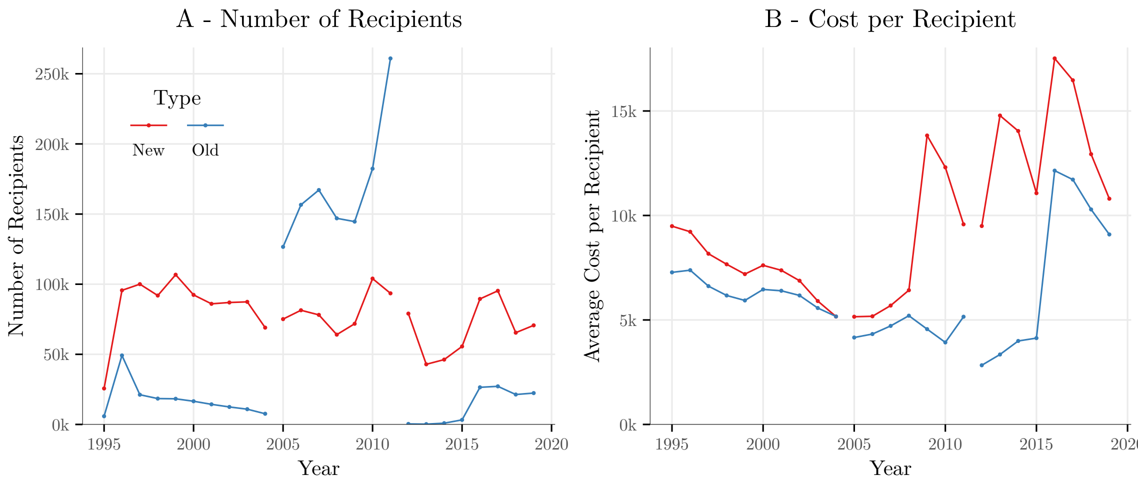


Figure A.1. Number of IFL Recipients and Average Subsidy since the Policy Introduction

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

Table A.2. Descriptive statistics for the Pre-Treatment Variables for Municipality included in the GPS Specification

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

Notes: The average density of the 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

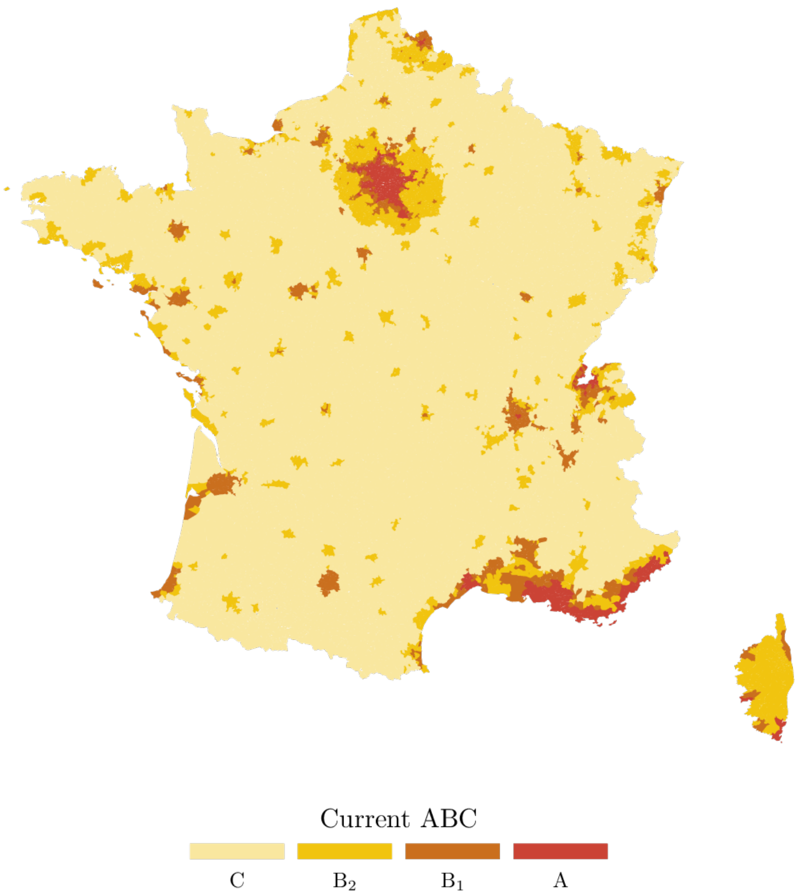


Figure A.4. Current ABC Zoning

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:

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

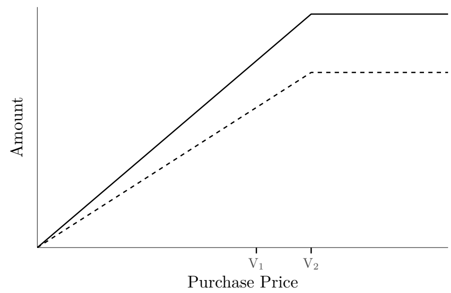
yielding an overall cost for the household to

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

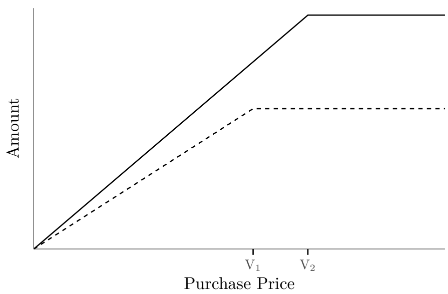
B.2 Difference in Treatment According to Policy-Controlled Primitives

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

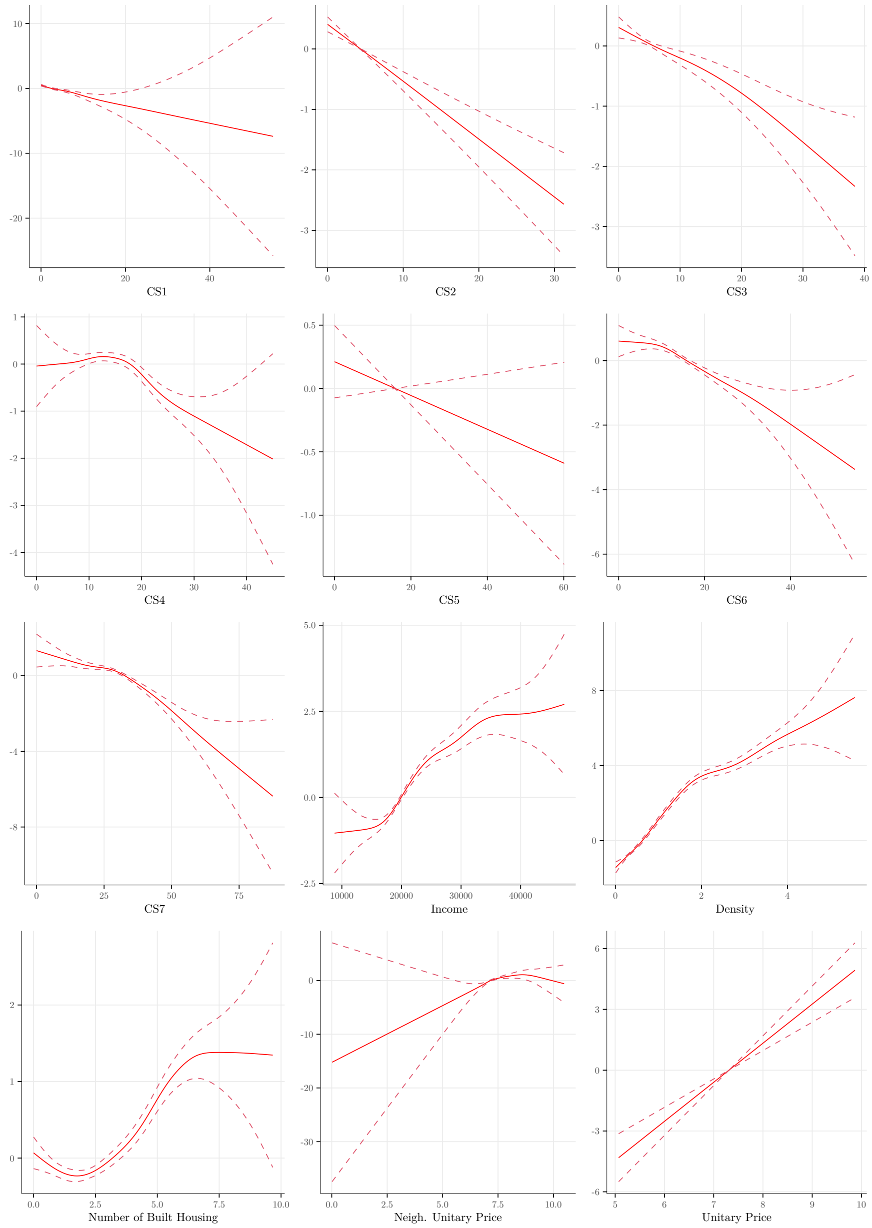


Figure C.1. Contribution for 1D-variable in the GPS Estimation

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

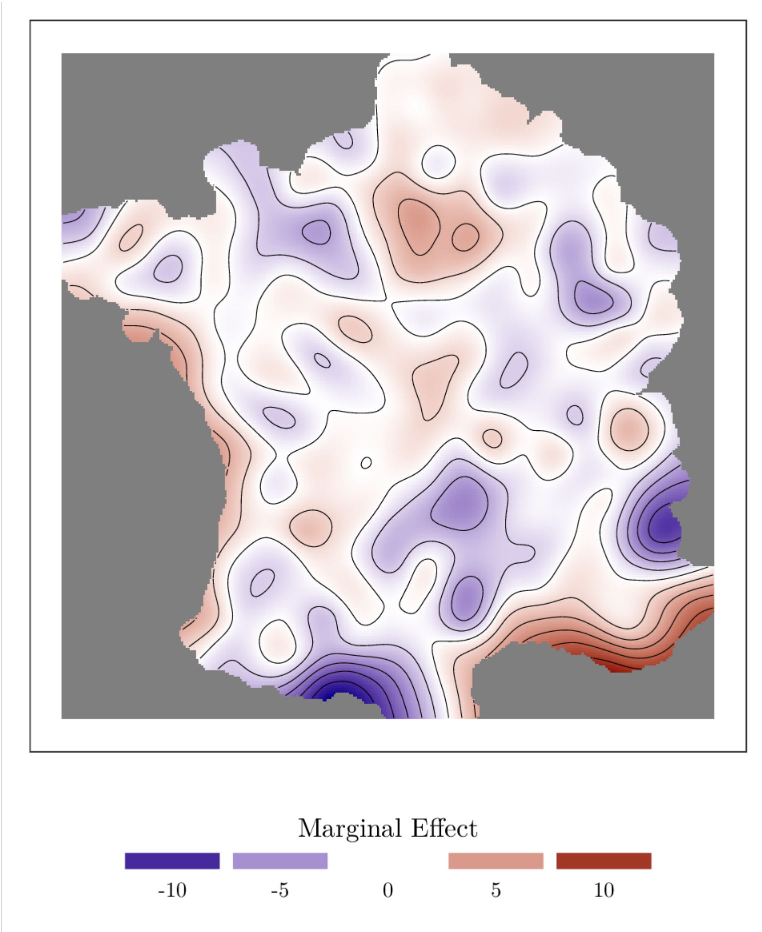


Figure C.2. Spatial Smoothing Function for the GPS Estimation Based On Municipality Coordinates

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

(a) Predicted Zoning (b) Prediction Results

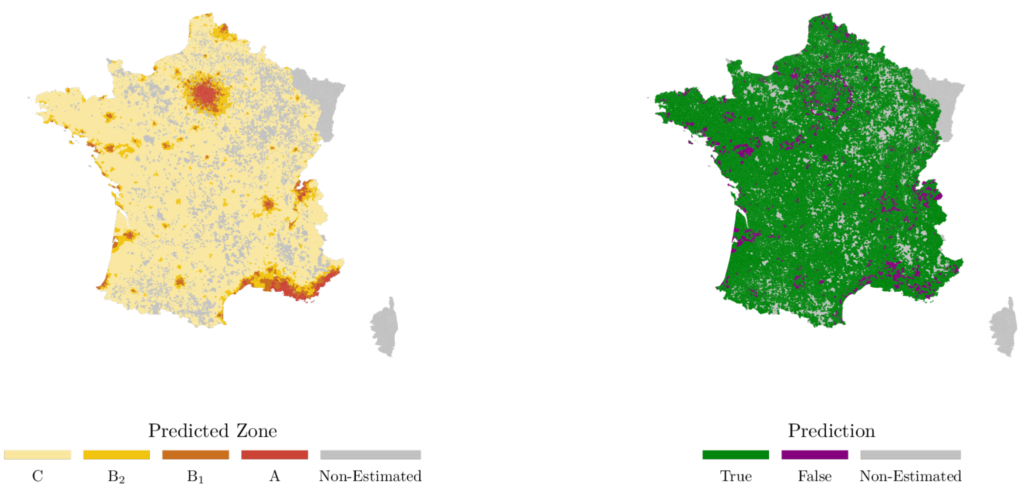


Figure C.3. Estimated Classification of the Municipalities

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)

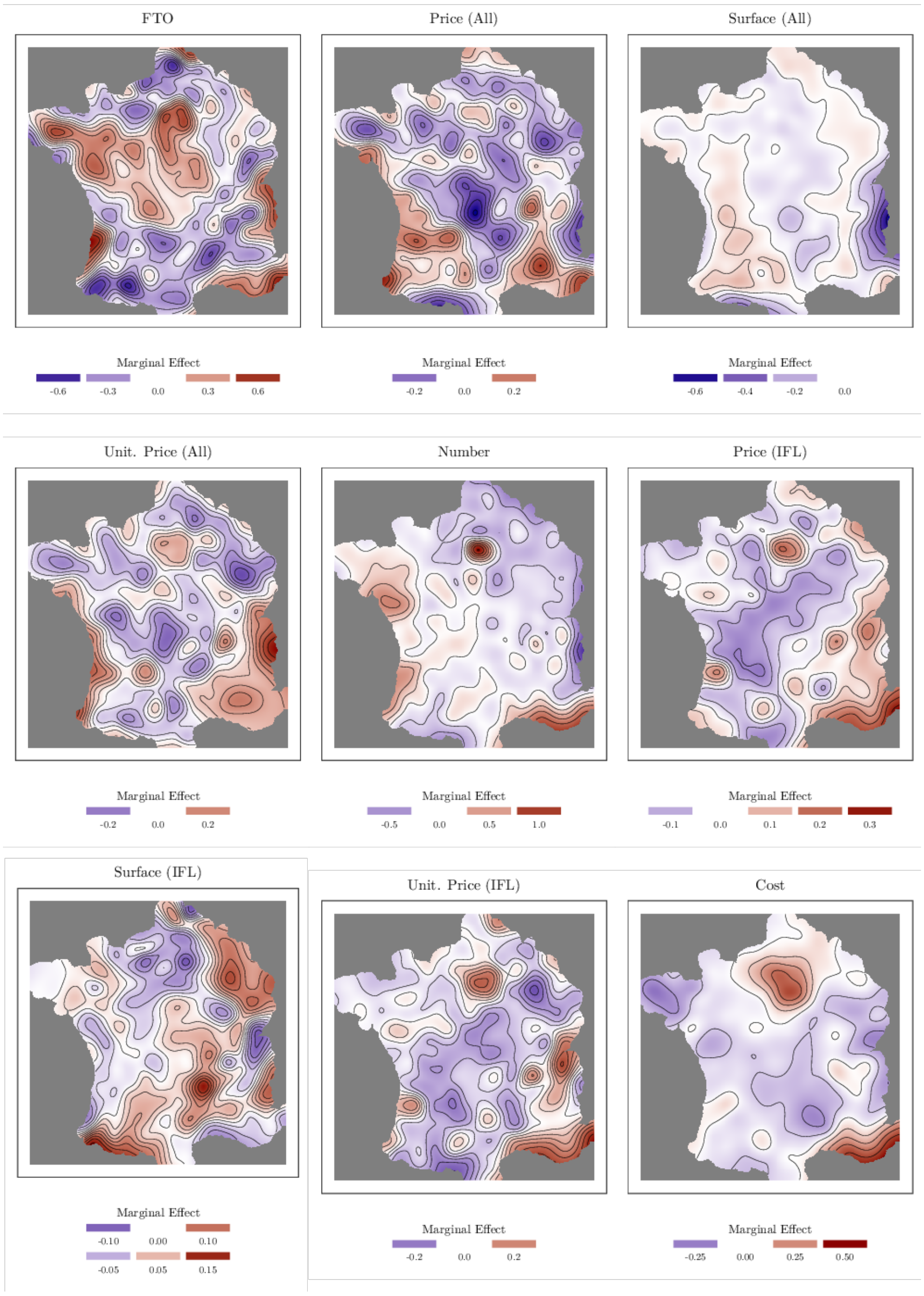


Figure C.6. Marginal Contribution for Spatial Coordinates (Second Step)

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
	-	(0.980)	(0.965)	(0.965)	-	(1.261)	(1.245)	(1.245)	-	(0.882)	(0.872)	(0.870)
B_1	-0.242	-	0.209	0.227	1.180	-	0.821***	0.672***	-1.168	-	-0.142	-0.189
	(0.980)	-	(0.175)	(0.176)	(1.261)	-	(0.184)	(0.182)	(0.882)	-	(0.163)	(0.159)
B_2	-0.451	-0.209	-	0.018	0.359	-0.821***	-	-0.149***	-1.026	0.142	-	-0.048*
	(0.965)	(0.175)	-	(0.027)	(1.245)	(0.184)	-	(0.029)	(0.872)	(0.163)	-	(0.025)
C	-0.469	-0.227	-0.018	-	0.508	-0.672***	0.149***	-	-0.979	0.189	0.048*	-
	(0.965)	(0.176)	(0.027)	-	(1.245)	(0.182)	(0.029)	-	(0.870)	(0.159)	(0.025)	-
					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
	-	(0.190)	(0.150)	(0.150)	-	(0.321)	(0.319)	(0.319)	-	(0.286)	(0.280)	(0.280)
B_1	0.431**	-	0.301***	0.232**	-0.446	-	0.043	0.010	-0.294	-	-0.496***	-0.542***
	(0.190)	-	(0.112)	(0.111)	(0.321)	-	(0.044)	(0.043)	(0.286)	-	(0.068)	(0.067)
B_2	0.129	-0.301***	-	-0.070***	-0.488	-0.043	-	-0.033***	0.202	0.496***	-	-0.046***
	(0.150)	(0.112)	-	(0.011)	(0.319)	(0.044)	-	(0.011)	(0.280)	(0.068)	-	(0.008)
C	0.199	-0.232**	0.070***	-	-0.456	-0.010	0.033***	-	0.248	0.542***	0.046***	-
	(0.150)	(0.111)	(0.011)	-	(0.319)	(0.043)	(0.011)	-	(0.280)	(0.067)	(0.008)	-
					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
	-	(0.161)	(0.156)	(0.156)	-	(0.160)	(0.157)	(0.157)	-	(0.157)	(0.157)	(0.157)
B_1	0.081	-	0.213***	0.194***	-0.177	-	-0.030	-0.010	-0.270*	-	-0.142***	-0.112***
	(0.161)	-	(0.031)	(0.031)	(0.160)	-	(0.035)	(0.035)	(0.157)	-	(0.024)	(0.023)
B_2	-0.132	-0.213***	-	-0.019***	-0.147	0.030	-	0.021***	-0.128	0.142***	-	0.030***
	(0.156)	(0.031)	-	(0.006)	(0.157)	(0.035)	-	(0.005)	(0.157)	(0.024)	-	(0.004)
C	-0.113	-0.194***	0.019***	-	-0.168	0.010	-0.021***	-	-0.158	0.112***	-0.030***	-
	(0.156)	(0.031)	(0.006)	-	(0.157)	(0.035)	(0.005)	-	(0.157)	(0.023)	(0.004)	-
					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
	-	(0.198)	(0.181)	(0.181)	-	(0.363)	(0.361)	(0.359)	-	(0.230)	(0.202)	(0.202)
B_1	0.116	-	0.209**	0.177**	-0.209	-	-0.229***	-0.312***	-0.259	-	-0.404***	-0.485***
	(0.198)	-	(0.088)	(0.087)	(0.363)	-	(0.051)	(0.050)	(0.230)	-	(0.116)	(0.116)
B_2	-0.092	-0.209**	-	-0.031***	0.021	0.229***	-	-0.082***	0.146	0.404***	-	-0.081***
	(0.181)	(0.088)	-	(0.011)	(0.361)	(0.051)	-	(0.011)	(0.202)	(0.116)	-	(0.010)
C	-0.061	-0.177**	0.031***	-	0.103	0.312***	0.082***	-	0.227	0.485***	0.081***	-
	(0.181)	(0.087)	(0.011)	-	(0.359)	(0.050)	(0.011)	-	(0.202)	(0.116)	(0.010)	-

Notes: We report the bilateral combinations effect for ATE type estimator. The four panels correspond to the four outcome concerned by ATE estimation and derived from fiscal data. Then, we have three main columns that represent the stable period for the IFL scheme, with four subcolumns related to the ABC classification. In rows, we have again the levels contained in the ABC classification. Hence, the bilateral combinations are reported for each intersection, and must be understand as "if (rows) have received (cols), difference in outcome would be (results)". We also report in 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***
	-	(0.041)	(0.150)	(0.049)	-	(0.049)	(0.115)	(0.105)	-	(0.053)	(0.105)	(0.109)
B_1	3.446***	-	-0.196***	-0.290***	0.275***	-	-0.205***	-0.547***	-1.715***	-	-0.210***	-0.576***
	(0.047)	-	(0.040)	(0.013)	(0.080)	-	(0.043)	(0.018)	(0.041)	-	(0.039)	(0.016)
B_2	3.809***	-0.145***	-	-0.163***	0.185	-0.204***	-	-0.233***	-0.279**	-0.124***	-	-0.261***
	(0.325)	(0.043)	-	(0.018)	(0.286)	(0.055)	-	(0.016)	(0.132)	(0.045)	-	(0.019)
C	5.385***	0.275	-0.009	-	2.255**	-0.095	0.002	-	0.670	-1.472***	0.066*	-
	(1.064)	(0.226)	(0.032)	-	(0.974)	(0.283)	(0.038)	-	(0.456)	(0.210)	(0.036)	-
					$\widetilde{V}_g$: Average Housing Price (Subsidized Housing)							
A	-	-0.028***	-0.191***	-0.076***	-	-0.022***	-0.011	-0.060***	-	-0.069***	-0.161***	-0.046***
	-	(0.008)	(0.023)	(0.013)	-	(0.008)	(0.015)	(0.015)	-	(0.008)	(0.019)	(0.010)
B_1	-0.801***	-	-0.043***	-0.034***	-0.118***	-	-0.050***	-0.036***	-0.293***	-	-0.048***	-0.048***
	(0.011)	-	(0.009)	(0.002)	(0.008)	-	(0.006)	(0.003)	(0.008)	-	(0.006)	(0.002)
B_2	-1.000***	0.096***	-	-0.030***	-0.163***	0.054***	-	-0.019***	-0.090***	0.103***	-	-0.032***
	(0.082)	(0.008)	-	(0.004)	(0.048)	(0.010)	-	(0.002)	(0.026)	(0.009)	-	(0.003)
C	-1.508***	0.160***	0.075***	-	0.222	0.100	0.033***	-	0.207***	0.307***	0.037***	-
	(0.272)	(0.038)	(0.010)	-	(0.147)	(0.069)	(0.004)	-	(0.061)	(0.033)	(0.004)	-
					$\widetilde{Q}_g$: Average Housing Size (Subsidized Housing)							
A	-	-0.109***	0.167***	0.654***	-	-0.073***	0.077***	0.242***	-	-0.083***	0.051**	0.350***
	-	(0.008)	(0.031)	(0.016)	-	(0.007)	(0.018)	(0.014)	-	(0.010)	(0.021)	(0.006)
B_1	-1.027***	-	0.022**	0.125***	-0.409***	-	0.011*	0.057***	-0.448***	-	0.023***	0.117***
	(0.014)	-	(0.009)	(0.002)	(0.009)	-	(0.006)	(0.002)	(0.014)	-	(0.007)	(0.001)
B_2	-1.062***	0.026***	-	0.014***	-0.996***	0.089***	-	0.010***	-0.734***	0.019**	-	-0.001
	(0.170)	(0.009)	-	(0.004)	(0.095)	(0.010)	-	(0.002)	(0.027)	(0.009)	-	(0.002)
C	-1.997***	0.216***	0.054***	-	-2.287***	0.473***	0.012**	-	-0.941***	0.262***	-0.046***	-
	(0.494)	(0.038)	(0.008)	-	(0.380)	(0.059)	(0.006)	-	(0.125)	(0.026)	(0.006)	-
					$\widetilde{P}_g$: 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***
	-	(0.009)	(0.024)	(0.014)	-	(0.007)	(0.015)	(0.013)	-	(0.007)	(0.021)	(0.010)
B_1	0.231***	-	-0.071***	-0.128***	0.382***	-	-0.063***	-0.093***	0.212***	-	-0.069***	-0.165***
	(0.013)	-	(0.008)	(0.002)	(0.007)	-	(0.006)	(0.002)	(0.009)	-	(0.007)	(0.001)
B_2	0.116	0.037***	-	-0.044***	0.978***	-0.042***	-	-0.026***	0.624***	0.077***	-	-0.032***
	(0.178)	(0.009)	-	(0.004)	(0.045)	(0.013)	-	(0.002)	(0.026)	(0.008)	-	(0.004)
C	0.484	-0.046	0.018*	-	1.822***	-0.366***	0.026***	-	1.219***	0.079***	0.084***	-
	(0.500)	(0.049)	(0.010)	-	(0.151)	(0.115)	(0.006)	-	(0.084)	(0.030)	(0.008)	-
					$\widetilde{C}_g$: Average Cost per IFL							
A	-	-0.125***	-0.516***	-1.514***	-	-0.056***	-0.226***	-0.173***	-	-0.110***	-1.096***	-0.745***
	-	(0.010)	(0.064)	(0.020)	-	(0.010)	(0.021)	(0.022)	-	(0.008)	(0.028)	(0.014)
B_1	-0.096***	-	-0.527***	-1.148***	0.155***	-	-0.200***	-0.266***	-0.229***	-	-0.874***	-0.922***
	(0.023)	-	(0.017)	(0.003)	(0.013)	-	(0.009)	(0.004)	(0.015)	-	(0.009)	(0.002)
B_2	0.245***	0.509***	-	-0.490***	-0.151***	0.144***	-	-0.110***	0.664***	0.876***	-	-0.126***
	(0.080)	(0.015)	-	(0.008)	(0.041)	(0.009)	-	(0.003)	(0.027)	(0.019)	-	(0.006)
C	0.532**	0.905***	0.555***	-	-0.432***	0.508***	0.139***	-	1.197***	0.976***	0.140***	-
	(0.255)	(0.101)	(0.016)	-	(0.128)	(0.056)	(0.011)	-	(0.133)	(0.075)	(0.010)	-

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)

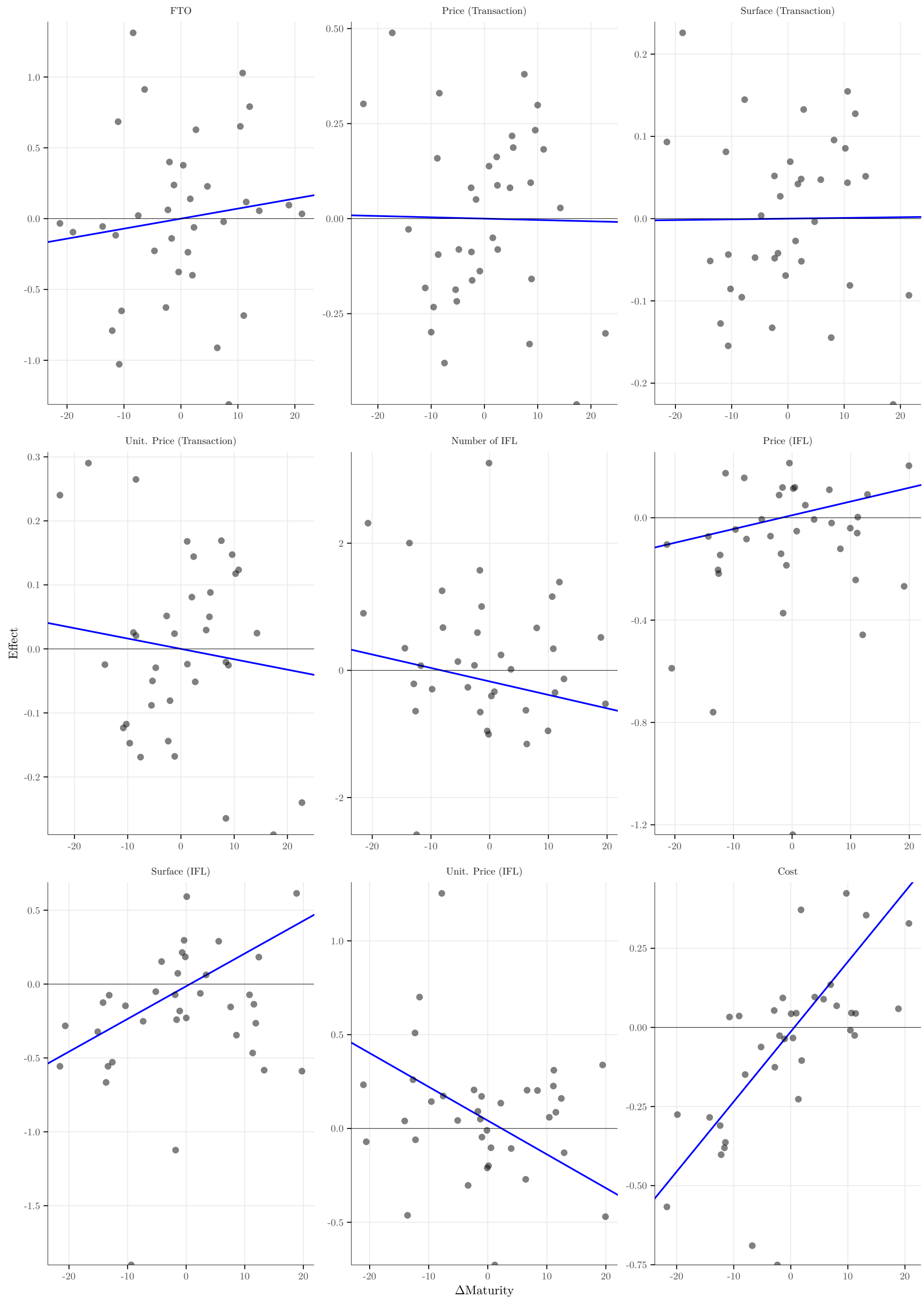


Figure C.9. Dose-Response Plots

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)

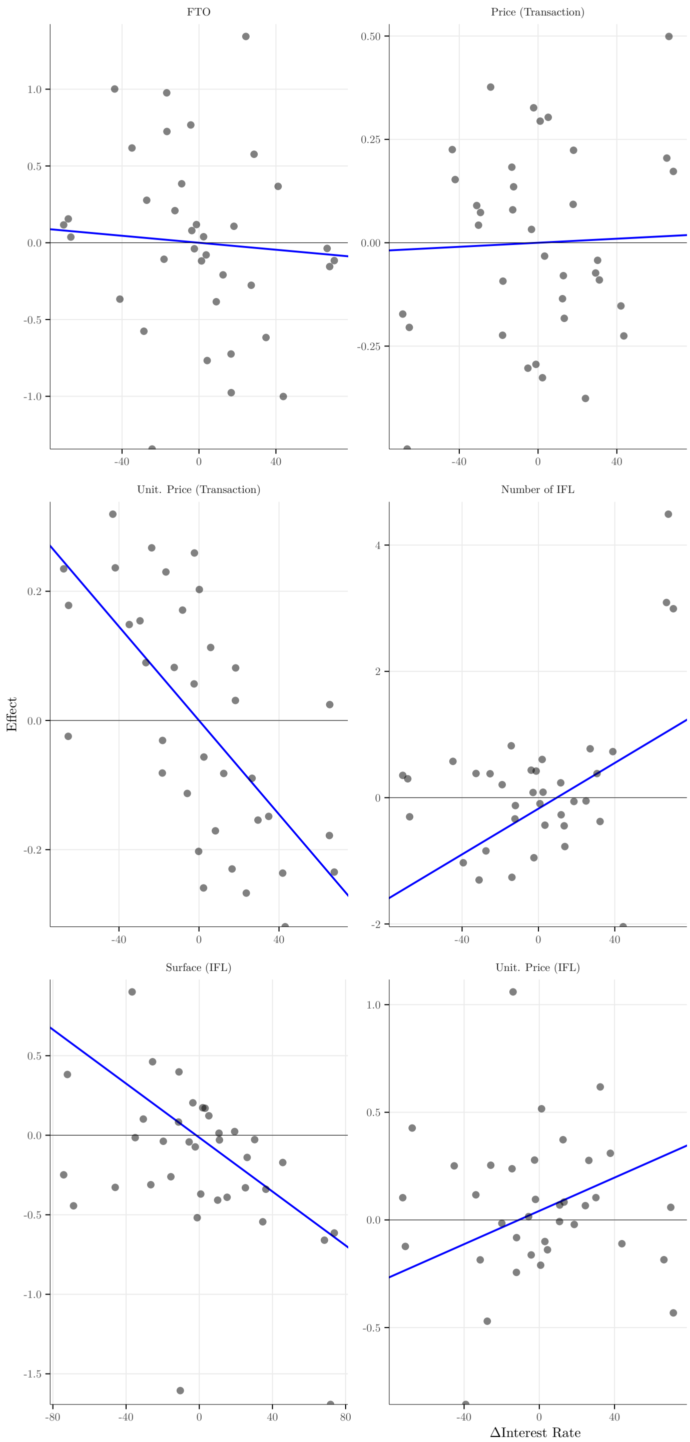


Figure C.10. Dose-Response Plots

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)

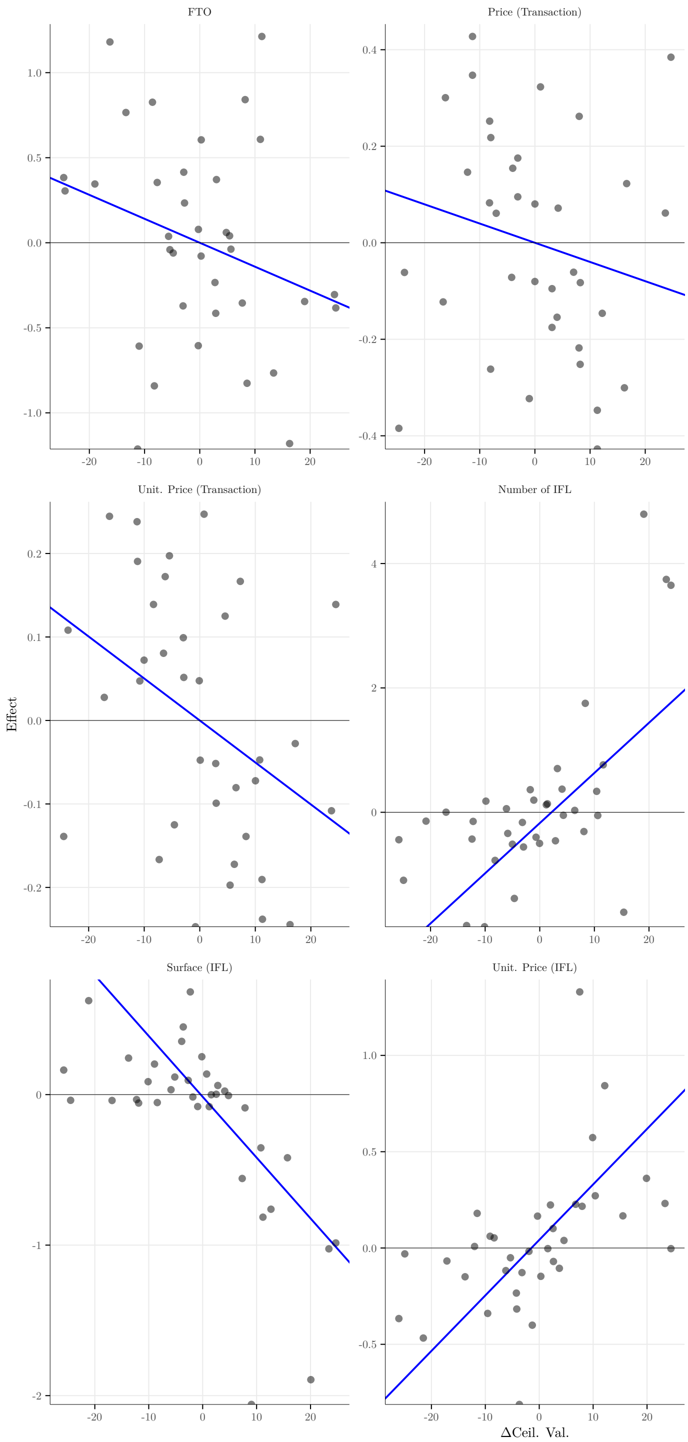


Figure C.11. Dose-Response Plots

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)

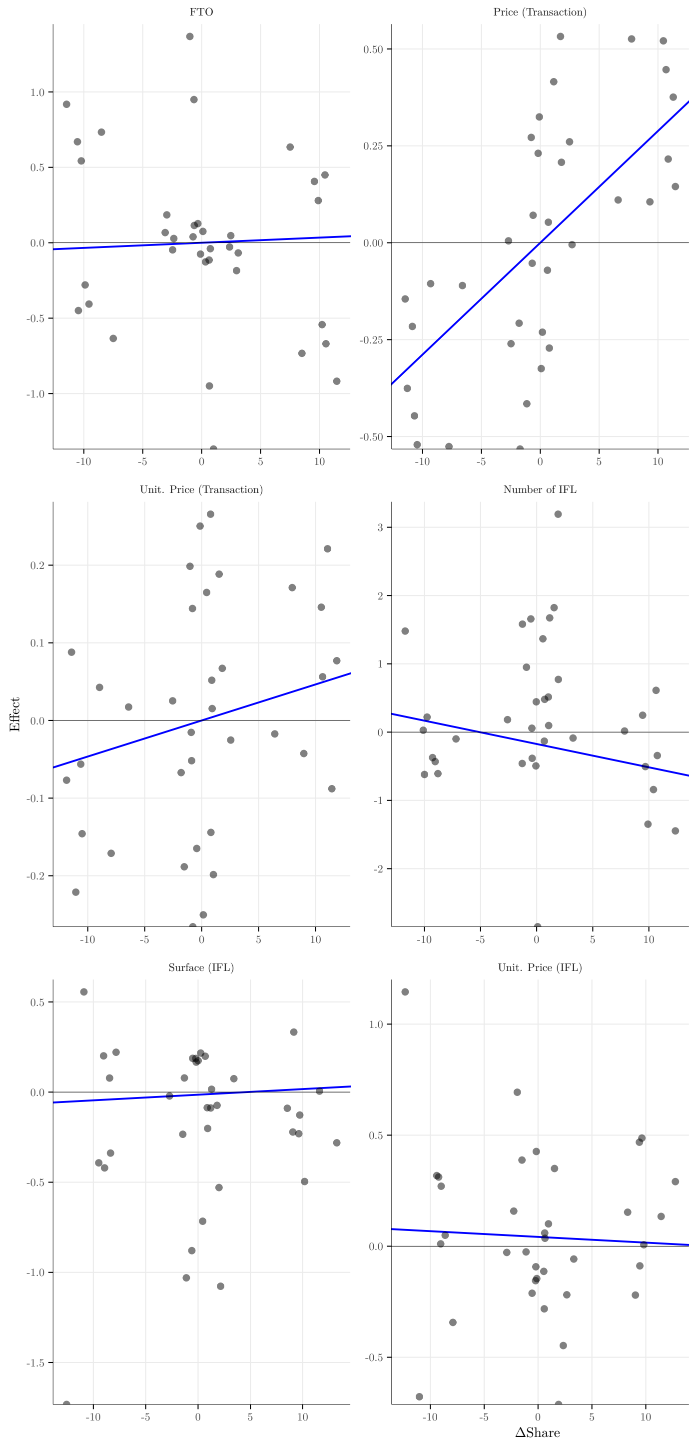


Figure C.12. Dose-Response Plots

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 home-owners, 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	477.2	-3,025.2***	3,502.5***
	(1,166)	(1,104)	(202)
Ceiling Value	6,302.0***	6,821.3***	-519.3***
	(846)	(776)	(192)
Interest Rates	1,475.5***	1,539.0***	-63.5
	(244)	(224)	(55)
Loan Maturity	40.8	-1,840.8***	1,881.6***
	(487)	(447)	(118)

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

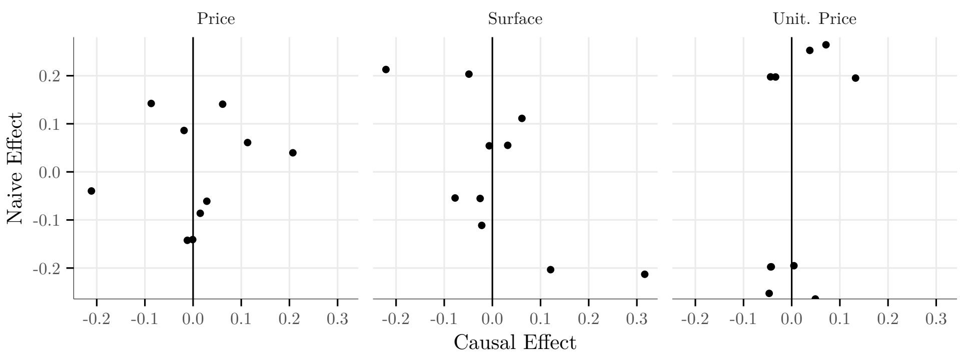
D Placebo Analysis

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).

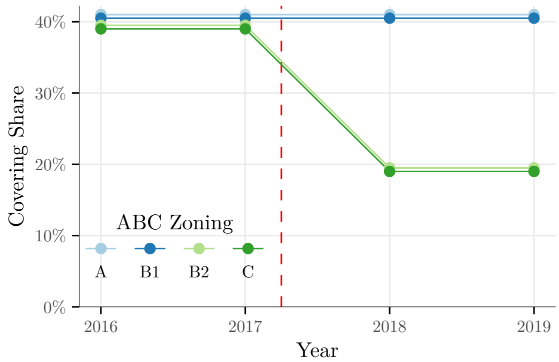


Figure E.1. Covering Share per ABC Areas According to Time

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

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

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

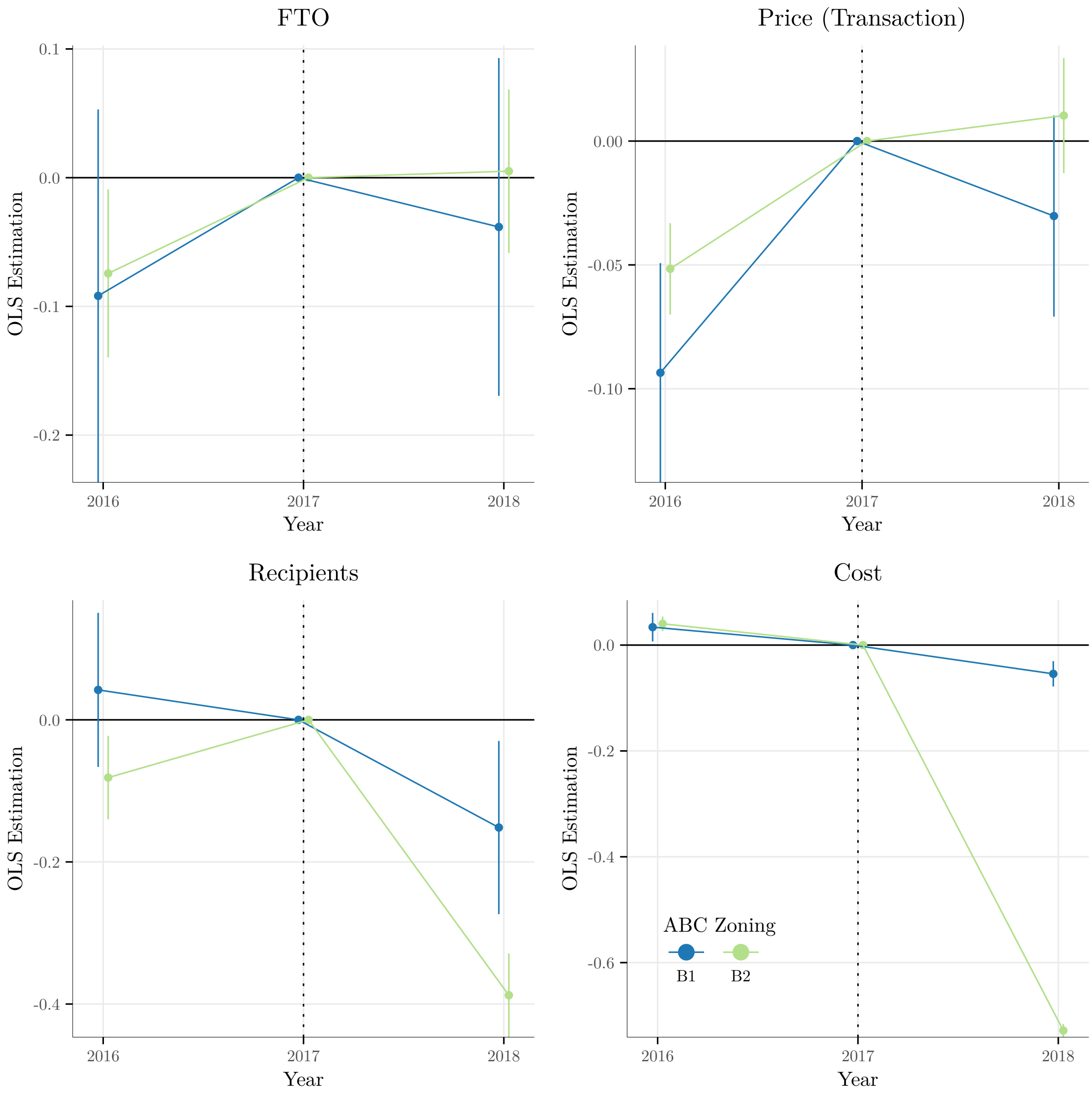


Figure E.2. OLS Results from DiD Identification Strategy restricted to B_1 and B_2 Areas

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

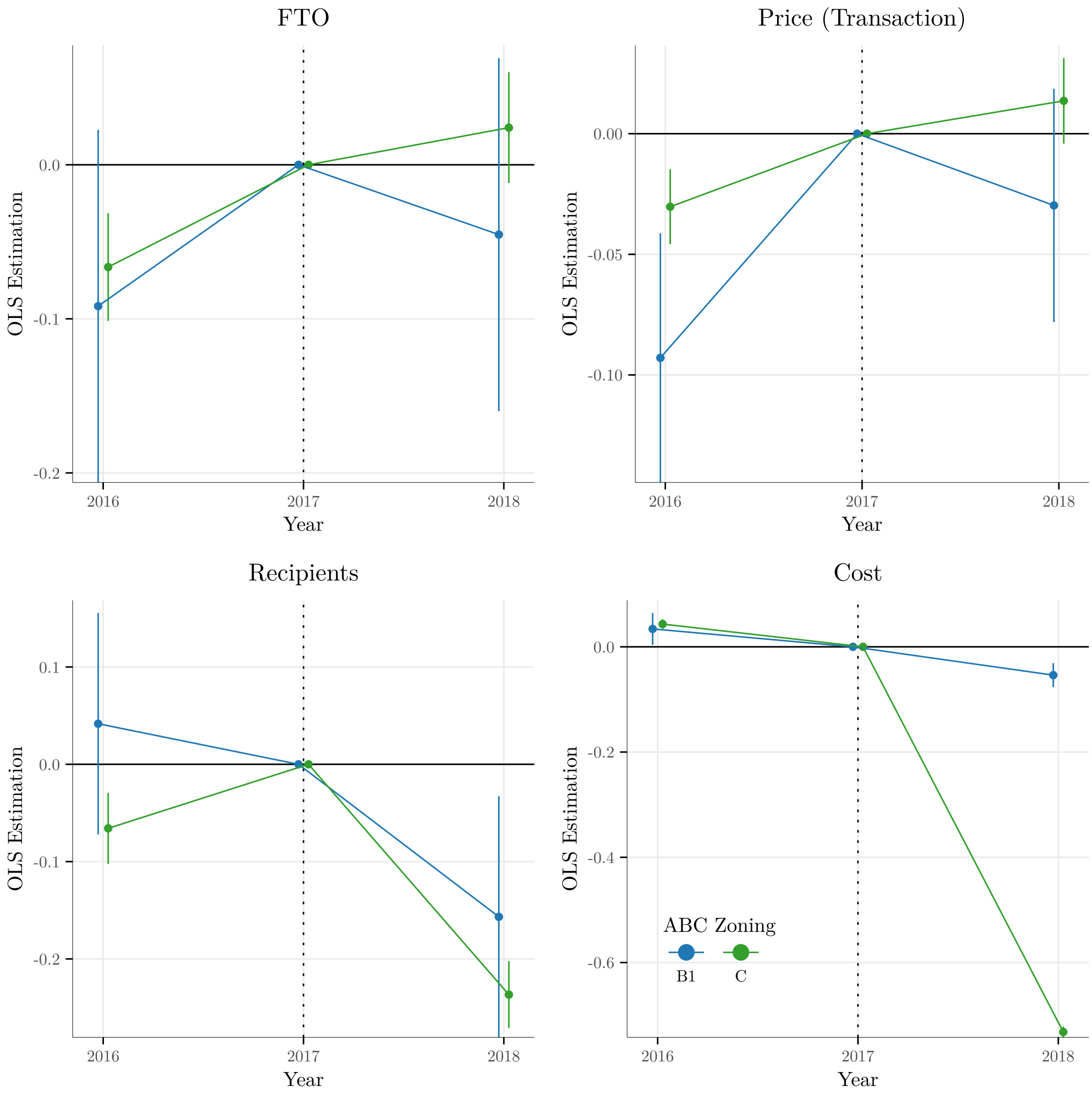


Figure E.3. OLS Results from DiD Identification Strategy restricted to B_1 and C Areas

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

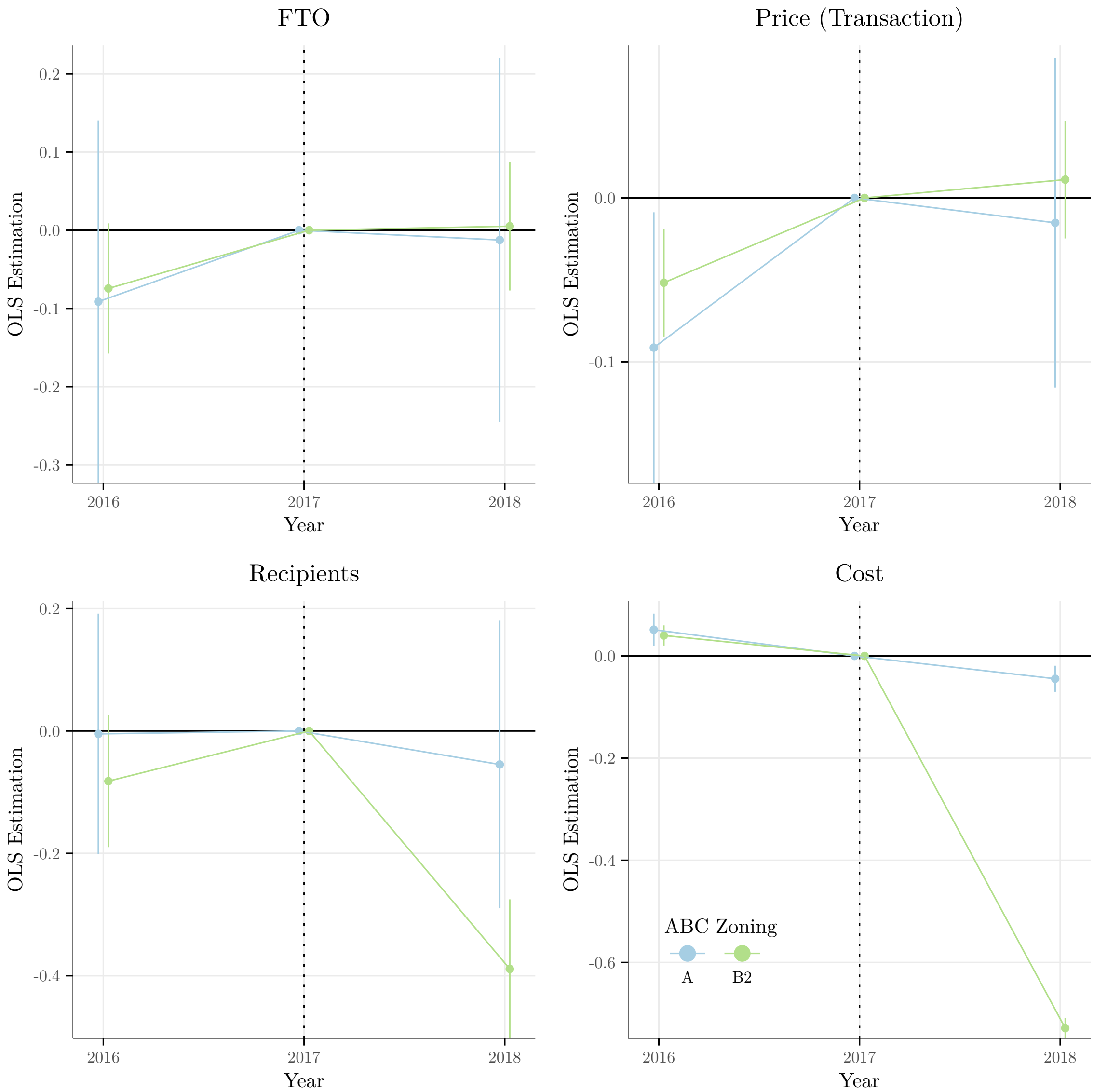


Figure E.4. OLS Results from DiD Identification Strategy restricted to A and B_2 Areas

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

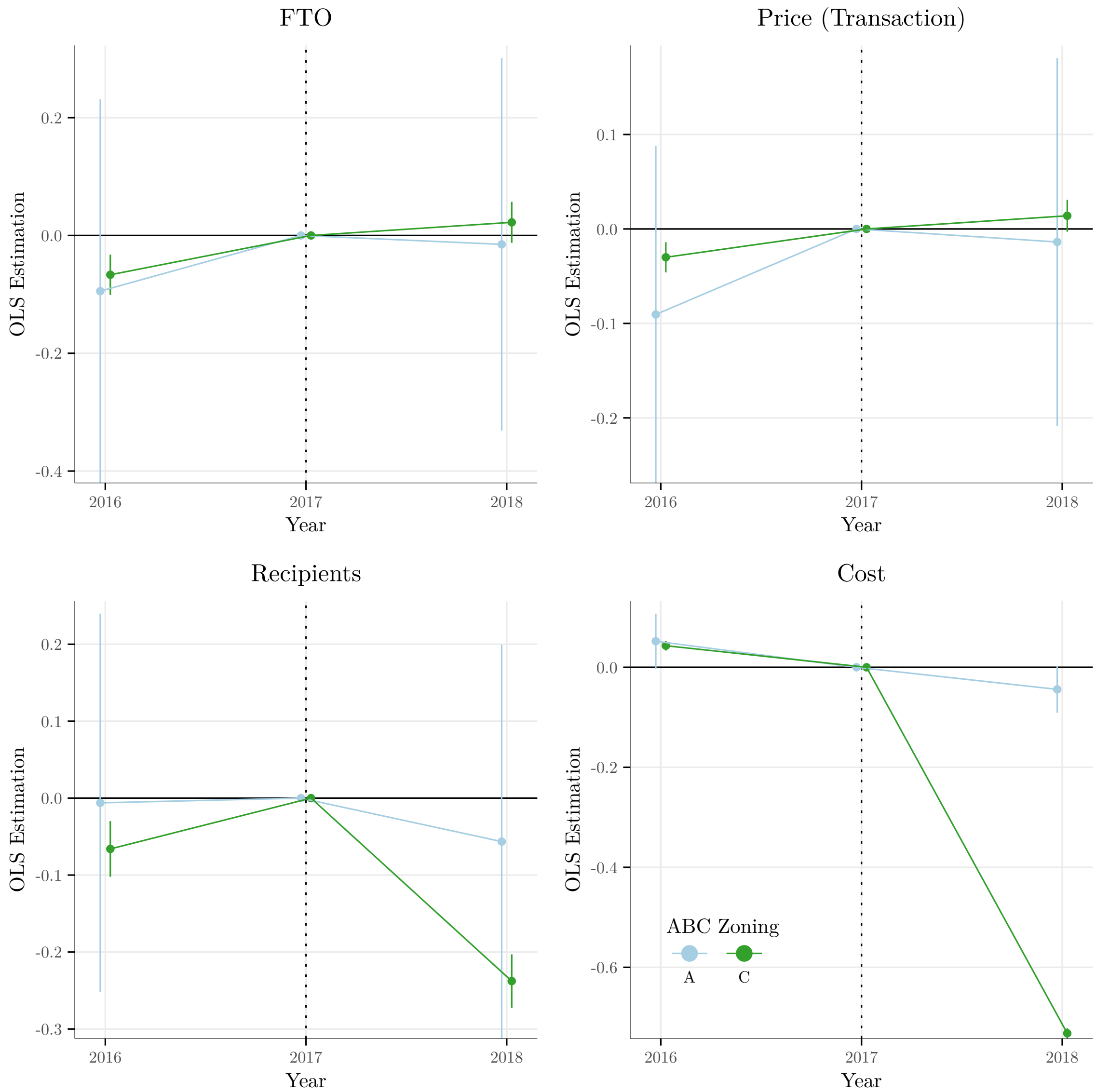


Figure E.5. OLS Results from DiD Identification Strategy restricted to A and C Areas

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

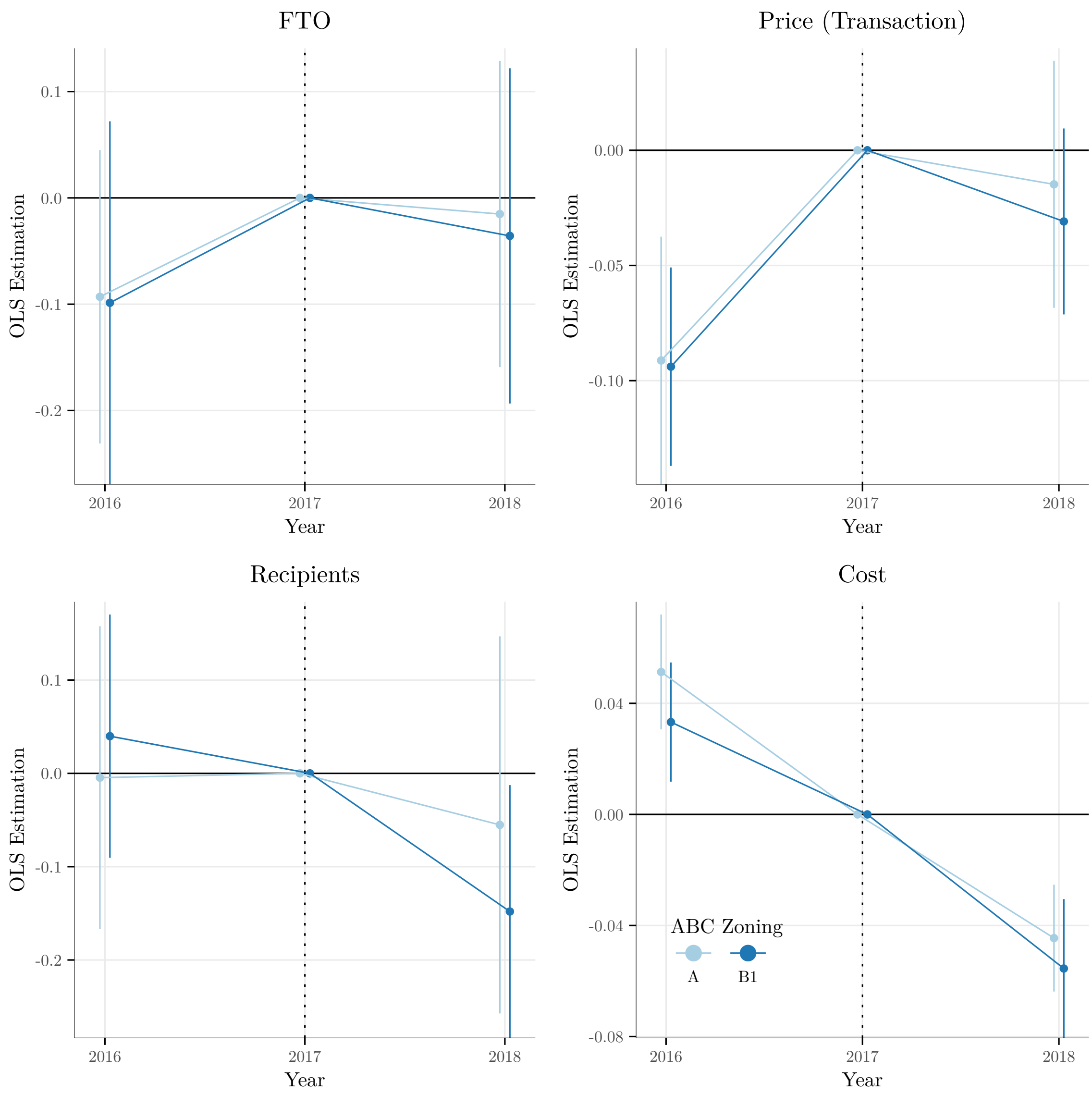


Figure E.6. OLS Results from DiD Identification Strategy restricted to A and B_1 Areas

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

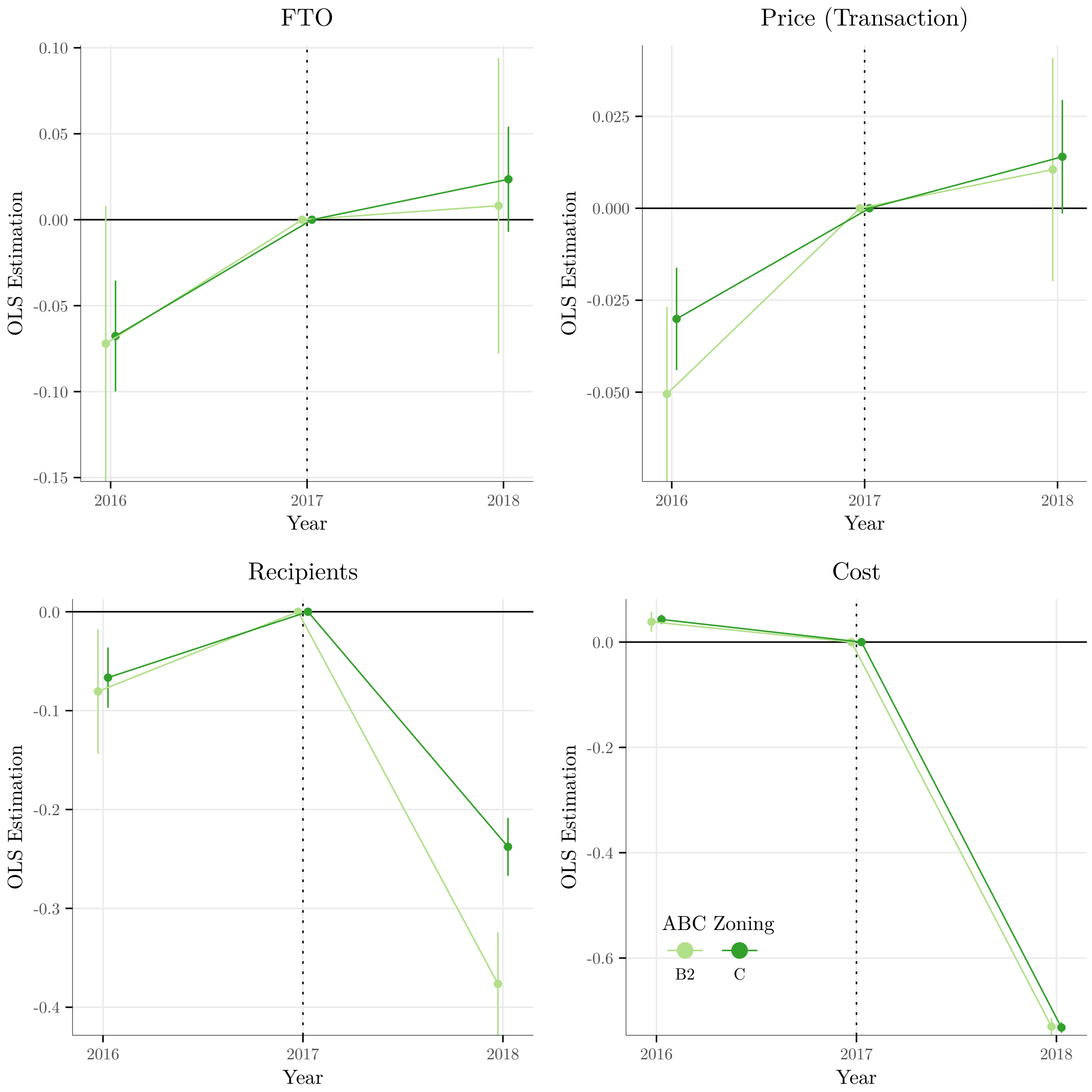


Figure E.7. OLS Results from DiD Identification Strategy restricted to B_2 and C Areas

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.