WORKING PAPER
CLIMATE ECONOMICS CHAIR

The effect of phasing out energy-inefficient dwellings from the housing market: a housing demand approach

Edouard Civel[1], Anna Creti[2], Gabrielle Fack[3], Daniel Herrera-Araujo[4]

March 2026

1. Chaire Economie du Climat
2. Université Paris-Dauphine, LEDA (CGEMP), and Chaire Economie du Climat, Place du Maréchal de Lattre de Tassigny
3. Université Paris-Dauphine, LEDA (LEGOS), Place du Maréchal de Lattre de Tassigny
4. **Corresponding author**, Mines Paris PSL, 60 Boulevard Saint Michel, 75006, Paris (e-mail: daniel.herrera@minesparis.psl.eu)

Launched in October 2010, the Climate Economics Chair is designed to be a research platform recognized academically and open to the practionners world. Its research program, validated by an international scientific committee, is focused on the economics of climate change.
The Chair has set as a priority the training of young researchers by hosting trainees, doctoral and post-doctoral students from various backgrounds.
The Climate Economics Chair contributes to the dissemination of scientific research through the organization of seminars open to professionals and researchers.

Access all the Climate Economics Chair Working Papers on our website

Chaire Économie du Climat · Palais Brongniart, 4 ème étage · 28 place de la Bourse · 75002 PARIS

The effect of phasing out energy-inefficient dwellings from the housing market: a housing demand approach

Edouard Civel, Anna Creti, Gabrielle Fack, Daniel Herrera-Araujo*

This version: March 2026

Abstract

We evaluate France's phased rental ban on energy-inefficient housing using nationwide housing transactions (2016–2023), linked to Energy Performance Certificates (EPCs) and occupancy records. Focusing on apartments, we use a reduced-form design around the 2021 policy announcement. We find that the annual probability of sale for treated units (EPC G dwellings that were rented prior to sale) increases by 0.72 percentage points. Relative to a baseline annual sale probability of 2.9% for rented EPC G dwellings, this corresponds to a 25% increase. The reform also changes post-purchase use. Within two years of sale, the share of sold EPC G apartments that are rented declines by about 3.8 percentage points, while owner-occupation rises by roughly 2 percentage points. Given that around 36% of transacted apartments are rented in the pre-announcement period, a 3.8-point decline implies about a 10% reduction in the share of newly purchased apartments that enter (or remain in) the rental market. Prices also re-sort across EPC ratings in ways consistent with expected renovation and compliance costs. Finally, we estimate a structural equilibrium model of housing demand that matches these moments and observed market shares. Counterfactual simulations suggest that, under full implementation, owner-occupiers and second-home buyers increase their market shares at the expense of landlords, with price effects concentrated among the least energy-efficient dwellings.

Keywords: Energy performance certificates, hedonic models, sorting demand estimation models.

JEL Codes: Q51, Q58, R21.

Acknowledgments: For helpful discussions, comments and suggestions, we thank Matthieu Glachant. This work benefited from a grant ADEME, within the framework of the "Climate Finance" (HERMES). We are grateful to seminar participants at Chaire Economie du Climat (2024), Mines Paris-PSL (2025), the European Society for Benefit Cost Analysis Conference (Athens, 2025).

*Herrera-Araujo: Corresponding author, Mines Paris PSL, 60 Boulevard Saint Michel, 75006, Paris (e-mail: daniel.herrera@minesparis.psl.eu); Civel: Chaire Economie du Climat; Fack: Université Paris-Dauphine, LEDA (LEGOS), Place du Maréchal de Lattre de Tassigny; Creti: Université Paris-Dauphine, LEDA (CGEMP), and Chaire Economie du Climat, Place du Maréchal de Lattre de Tassigny;

1 Introduction

Residential buildings are a major contributor to energy use and emissions, yet housing markets often underinvest in cost-effective energy-efficiency improvements. Policymakers have tried to close this gap with information disclosure, most notably mandatory energy performance certificates (EPCs), to help buyers and tenants account for expected energy costs and to reward efficient dwellings (Myers, Puller and West, 2022). But information alone is often insufficient. First, the impact of disclosure depends on accuracy and credibility, and evidence in France and elsewhere raises concerns about rating reliability (Fowlie, Greenstone and Wolfram, 2018; Davis, Fuchs and Gertler, 2020).[1] Even when ratings are accurate, housing markets are characterized by agency frictions: in rental housing, landlord–tenant split incentives weaken renovation effort because landlords bear the upfront costs while tenants capture much of the benefits. In response, policymakers have increasingly moved beyond disclosure toward outcome-based regulation that conditions rental eligibility on meeting minimum EPC thresholds. By directly restricting the use of low-rated dwellings in the rental market, these policies are intended to alter investment incentives, asset values, and tenure choices.

The policy debate surrounding such measures therefore centers on their welfare and market consequences. Advocates argue that stricter standards raise the private returns to renovation, reallocate non-compliant units toward owner-occupation, and ultimately reduce energy demand while improving housing quality. Critics counter that binding standards may contract rental supply in the short run, exert upward pressure on rents, reprice low-rated assets with distributional consequences, and strain limited renovation capacity.

This paper provides new evidence on the effects of one of the most ambitious French national efforts to tie rental use to energy performance: France's phased rental-eligibility standard enacted by the *Climate and Resilience Law* (2021). The policy targets dwellings that fail minimum Energy Performance Certificate (EPC) standards. Over a schedule running through 2034, the program freezes rents for F–G units (August 2022), bars new leases above 450 kWh/m^2/year (January 2023), and prohibits rentals of G, F, and E units in 2025, 2028, and 2034, respectively.

Our analysis proceeds in two steps. First, we study anticipation effects around the 2021 announcement using administrative data that cover the national market for metropolitan France. For the broader 2010–2024 period, the raw ADEME EPC database contains about 13 million certificates; within our 2016–2023 analytic window, it contains over ten million records. We match these records to exhaustive property transactions to identify buyer types and link transactions to pre- and post-reform EPC ratings. Second, we use the reduced-form moments as discipline for an equilibrium residential sorting model with heterogeneous buyers. Because of computational constraints, however, the structural model is estimated on a 20-department subsample and calibrated to the corresponding reduced-form moments from that same subsample. Throughout the paper, we therefore distinguish between evidence from the national reduced-form sample and counterfactual results generated from the smaller structural sample.

[1] See 60 millions de consommateurs, "Nouveau DPE: erreurs, pagaille," May 24, 2022.

Our reduced-form results highlight three stylized facts. First, the annual probability of sale increases for treated units—EPC G dwellings that were rented prior to sale—by 0.72 percentage points after the announcement. Relative to a baseline annual sale probability of 2.9% for rented EPC G dwellings, this corresponds to a 25% increase. Second, use after purchase shifts: within two years of sale, the share of sold EPC G apartments that are rented declines by about 3.8 percentage points, while owner-occupation rises by roughly 2 percentage points. Given that around 36% of transacted apartments are rented in the pre-announcement period, this implies roughly a 10% reduction in the share of newly purchased apartments that enter (or remain in) the rental market. Third, prices re-sort across EPC categories consistent with capitalization of expected renovation and compliance costs.

To interpret these patterns and to project longer-run effects, we estimate an equilibrium residential sorting model with heterogeneous buyers. The model lets owner-occupiers, landlords, and second-home buyers value dwelling and neighborhood attributes differently, while prices clear the sales market. Buyer types are observed in the transactions data, so the model takes the composition of buyer types as given and estimates type-specific valuations of dwelling attributes. We estimate preferences using pre-policy transactions with a BLP-style two-step procedure: a maximum likelihood step with a contraction mapping to recover mean indirect utilities, followed by a second-stage decomposition of mean utility into observables and an unobserved quality term. To address price endogeneity in the second stage, we use spatial instruments constructed from dwelling characteristics in "donut" distance bins around each transacted unit (excluding the closest ring). We then combine estimated demand with a supply curve and solve for equilibrium prices that clear the market, which delivers predicted allocations and market shares by EPC and buyer type. Counterfactual policies enter the model by changing the relative attractiveness and effective cost of low-rated dwellings, and we re-solve for the resulting equilibrium prices and sorting outcomes.

Our counterfactual design separates two policy channels. We model compliance as an upgrade of affected dwellings to EPC D that imposes a renovation cost and may also generate non-price utility gains ("comfort") for buyers. We calibrate the renovation cost so that the model matches the reduced-form change in the rental-oriented use after purchases for EPC G units. That is, we find the renovation costs that rationalize the 3 percentage point decline estimated on the 20-department structural subsample (the national estimate is approximately 3.8 percentage points). The implied renovation costs amount to €92.5/m^2. We then extend the exercise beyond EPC G to the broader E–F–G phase-out that is central to the French schedule, using the calibrated renovation costs of €92.5/m^2 for G-to-D, and an educated guess of €70/m^2 for F-to-D, and €40/m^2 for E-to-D.

Our contributions fall into three categories. *Data:* we assemble nationwide transactions linked to EPC records and post-sale occupancy outcomes, allowing us to track not only prices and sales but also how purchased dwellings are used after the transaction. *Reduced form:* we estimate announcement effects on sales, prices, and post-purchase use, documenting that the policy shifts EPC G apartments away from rental use and toward owner-occupation. To our knowledge, this is the first evidence for France using nationwide transactions linked to post-sale occupancy to show that minimum-energy standards affect dwelling use, not only capitalization into prices. *Structural:* we estimate an equilibrium sorting model with heterogeneous buyer types and use it to decompose the mechanisms behind these changes and to simulate counterfactual policy designs, including the prospective E–F-G schedule.

Our study relates to the literature on how energy labels shape residential decisions by making energy efficiency more salient to households. Most of this work estimates how energy efficiency capitalizes into house prices using hedonic methods, typically finding a sales premium for more efficient dwellings.[2] The closest related evidence comes from the UK, where privately rented properties must meet minimum energy efficiency standards. Reduced-form estimates there point to modest price effects (Ferentinos, Gibberd and Guin (2023)) and very small rent increases, with potential unintended consequences for aggregate emissions (Clara et al. (2025)).[3] We contribute to this literature by moving beyond capitalization. We study how a rental-ban policy reshapes who owns and occupies energy-inefficient units, that is, how dwellings reallocate across market segments and uses. We also connect to structural urban models of household location choice, following Bayer, Ferreira and McMillan (2007), which are used to evaluate place-based policies such as public housing demolitions (Almagro, Chyn and Stuart (2023)). To the best of our knowledge, this is among the first studies to use an equilibrium residential sorting model with heterogeneous buyer types to study the consequences of a rental-eligibility policy for energy-inefficient dwellings.

The remainder of the paper is organized as follows. Section 2 presents the French EPC system, the 2021 reform, and the rental-ban timeline. Section 3 describes the datasets and matching procedures. Section 4 documents announcement effects and motivates the structural analysis. Sections 5 and 6 present and estimate the sorting model. Section 7 reports counterfactual simulations, with a focus on the prospective E–F–G schedule, and Section 8 concludes.

2 Background

2.1 The Regulatory Framework for Energy Performance Certificates in France

The French Energy Performance Certificate (EPC), or Diagnostic de Performance Energétique (DPE), is part of a broader EU regulatory framework intended to reduce building energy consumption and greenhouse gas emissions.[4] France has required an EPC for property sales and

[2]Studies include Eichholtz, Kok and Quigley (2010) and Kahn and Kok (2014) in the US, and Brounen and Kok (2011) in the Netherlands, Hyland, Lyons and Lyons (2013) and Fuerst et al. (2015) in Ireland and England. In France, Civel (2019) also finds a significant green premium across several local markets. Several papers specifically measure how prices react to the disclosure of EPC labels (Aydin, Brounen and Kok (2020) and Myers, Puller and West (2022)). Some papers also find evidence of manipulation (Civel et al. (2025), Lu and Spaenjers (2025) or strategic renovations at label thresholds (Comerford, Lange and Moro (2018) and Sejas-Portillo, Moro and Stowasser (2025)).

[3]Reusens et al. (2025) analyses the Flemish housing market, where energy-inefficient dwellings were entirely banned, and also reports limited effects on prices.

[4]Directive 2010/31/EU on the energy performance of buildings mandates that EU countries implement a certification scheme to inform prospective buyers or tenants about a building's energy performance, building

rentals since 2007 and has mandated its inclusion in property advertisements since 2011. A certified and independent expert issues the EPC based on an on-site assessment of dwelling characteristics (e.g., insulation, heating, ventilation), and the certificate reports two A–G labels: an energy label (modeled consumption and expected energy costs) and a greenhouse-gas label (carbon intensity). The EPC methodology was reformed in 2021 under the *Climate and Resilience Law* (Loi n° 2021-1104), which harmonized the approach across the housing stock and aimed to improve comparability and transparency.

2.2 France's policy for phasing out energy-inefficient dwellings

A central element of the *Climate and Resilience Law* is the gradual removal from the rental market of the most energy-intensive dwellings, known as *passoires thermiques*. The policy leverages the mandatory EPC rating to target dwellings classified G, F, and eventually E, and it conditions rental eligibility on meeting minimum energy thresholds rather than imposing a direct renovation mandate at the point of sale.[5] Restrictions are phased in as follows:

Table 1: Timeline of Rental Market Restrictions for Energy-Inefficient Dwellings

Effective Date	**Restriction Imposed on Rental Market**
August 2022	Ban on rent increases for dwellings rated F or G.
January 1, 2023	Ban on new rentals consuming $>$ 450 kWh/m^2/year.
January 1, 2025	Ban on new rentals of all G-rated dwellings.
January 1, 2028	Ban on new rentals of all F-rated dwellings.
January 1, 2034	Ban on new rentals of all E-rated dwellings.

Notes: Restrictions apply to new leases and, in certain cases, to renewals of existing leases. EPC classes refer to the post-2021 DPE methodology. The schedule covers metropolitan France. Sources: Loi n° 2021-1104 (*Climat et Résilience*) and implementing decrees.

Table 1 summarises the phase-in schedule of rental restrictions. These measures apply primarily to new leases (and, in certain cases, renewals) and create incentives for landlords to renovate, reallocate the dwelling to owner-occupation, or sell.

Internationally, similar minimum-standard approaches have been used. The United Kingdom introduced the *Minimum Energy Efficiency Standards* (MEES), which since April 2020 have prohibited landlords in England and Wales from renting properties rated F or G unless an exemption applies.[6] The UK initially planned to tighten the requirement to a minimum EPC C by 2025 for new tenancies and 2028 for all tenancies, but these plans were abandoned in 2023 amid concerns about impacts on the rental market. The legal minimum therefore remains at EPC E, and the policy continues to generate debate about the trade-offs between environmental ambition and housing market outcomes.

upon the foundational framework established by the 2002 directive.

[5] A 2015 proposal by then-Housing Minister Cécile Duflot to mandate renovations at sale encountered constitutional objections and did not proceed, highlighting the legal and political challenges of imposing direct renovation requirements. This episode helped shape the more gradual, rental-eligibility approach embedded in the 2021 *Climate and Resilience Law*.

[6] https://www.gov.uk/guidance/domestic-private-rented-property-minimum-energy-efficiency-standard-landlord-guidance [Last visited 10 August 2025]

3 Data and Sample Construction

This study examines the impact of France's phased rental ban on housing market outcomes. All analyses and results in this paper are restricted to apartments (flats) and exclude houses. To do so, we combine data from multiple sources, including land use data, real estate transactions, parcel maps, and Energy Performance Certificates. This section describes concisely the datasets and the methodology employed to construct the final analytical sample. A detailed description of the construction of the dataset is reported in Appendix A.

The empirical exercises use two related but distinct samples. The reduced-form analysis uses the national apartment sample in metropolitan France over 2016–2023. The structural analysis uses pre-policy transactions from a 20-department subsample over 2016–2020. Whenever we move from reduced-form evidence to structural calibration, we refer to the moment computed on the structural subsample rather than the national estimate.

We start with annual land use data from 2016 to 2024, covering the French metropolitan area, obtained from CEREMA.[7] This database provides detailed information on land usage, premises, and property rights, which includes whether the dwelling is rented or not.

We combine it with the real estate transaction data from 2016 to 2023, also provided by CEREMA. It includes an exhaustive list of property and land sales with detailed characteristics.[8] Our dataset excludes non-expensive transfers, transfers not involving a notary, and complex transfers.[9] The sample is filtered to include only single-unit transactions with specific physical characteristics. Additionally, we restrict the sample based on price per square meter, land value, and estimated rental value.[10]

The CEREMA dataset does not contain information regarding dwelling energy efficiency. To obtain this information, we use the EPC dataset collected by ADEME since 2010. The raw database covers roughly 13 million certificates over 2010–2024; in the 2016–2023 analytic window used in this paper, it contains over 10 million observations, each representing an EPC with detailed property characteristics. Prior to 2021, approximately 20% of EPCs have incorrect geographical coordinates. To address this, we use the SNA API to match addresses with accurate coordinates.[11] A multi-step process, detailed in Appendix A, is employed to identify and correct inaccurate coordinates using postal code boundaries. After correcting for geographical inaccuracies, missing EPCs, and implausible energy consumption or emission values are removed. New or very efficient dwellings (EPC class A) are also excluded.

[7] The Center for Studies and Expertise on Risk, Environment, Mobility and Urban Development (CEREMA) is a public establishment under the supervision of the Ministry of Ecological Transition and Territorial Cohesion. Website: https://www.cerema.fr/fr/cerema [last visited 2 February 2025].

[8] This dataset was initially produced under the name DVF by the Direction Générale des Finances Publiques (DGFiP) from two sources of tax datasets: (1) the Fichier Informatisé des Données Juridiques Immobilières (FIDJI) dataset and the MAJIC (Mise à Jour des Informations Cadastrales) database. Now it is administered by the CEREMA.

[9] Complex transfers correspond to parcels with very large condominiums or with an unresolved dispute.

[10] Rental value is the theoretical monthly rent that could be obtained from a property if it were let without a vacancy period. The appendix provides a detailed description on how rental value is estimated.

[11] The SNA contains all the addresses in France. Website: https://adresse.data.gouv.fr [Last visited February 2025.]

A closest-distance matching algorithm is used to combine the EPC and the combined land-use/real estate transaction datasets, considering geographic coordinates and year of sale.[12,13] Third, coverage expanded and quality control increased, potentially altering the composition of observed dwellings. Finally, post-2021 data no longer differentiate between certificates issued for sales and rentals, complicating direct comparison of transaction-specific EPC distributions across periods. We recognize that these issues complicate the evaluation of the policy announcement, and we therefore address them explicitly in Section 4. The final dataset is restricted to dwellings with energy consumption between 0 and 700 kWh per m2 per year and greenhouse gas emissions up to 200 kg CO2/m2/year.

To increase EPC coverage beyond directly matched transactions, we extrapolate the energy rating of each identified dwelling to all other dwellings sharing the same parcel, construction year, surface-area quartile, and apartment status (apartment versus single-family house). This imputation assumes that dwellings within such narrowly defined groups exhibit similar thermal characteristics, a reasonable approximation given the homogeneity in building materials and design at that level of granularity, especially for apartments, which are the focus of this paper. This step substantially increases the share of dwellings with an assigned EPC, while preserving variation across parcels, vintages, and size categories.

4 Reduced-Form Evidence

This section presents evidence on the relationship between the announcement of the phasing-out policy and dwelling market outcomes. We document that the announcement is associated with an increase in landlords' likelihood of selling their rented dwellings, and that sold dwellings were less likely to be rented. These facts motivate our structural analysis of impacts in Section 5.

4.1 Effect of announcement on sales probability

We begin by analyzing how the policy announcement influenced the likelihood that dwellings were transacted—providing insight into market activity on the extensive margin. We estimate the following linear probability model using a dwelling-year panel, where the dependent variable is an indicator equal to one if the dwelling is sold in a given year:

$$Sale_{jt} = \beta_1 A_t \times R_j + \beta_2 A_t \times G_j + \beta_3 A_t \times G_j \times R_j + \lambda_j + \lambda_t + \varepsilon_{jt}, \tag{1}$$

[12] As for our period of analysis, an EPC is valid for a period of 10 years, we retain only certificates issued in the year of the transaction or earlier, excluding any DPE established after the mutation date. This implies that a dwelling sold in 2022 may be matched to a DPE issued in 2020. For each potential match, we calculate the geographic distance and discard matches located more than 150 meters from the transacted dwelling.

[13] The reform of the DPE in 2021 entailed substantial methodological and operational changes, which introduce potential comparability issues between the two datasets. First, the calculation method shifted from an energy-consumption-based approach (partly reliant on occupant-reported bills) to a model-based estimation using building characteristics, leading to systematic reclassification of certain property types (e.g., small apartments disproportionately moved to lower ratings). Second, the distribution of ratings changed markedly, with more B–D ratings and F–G ratings at the expense of E ratings in the post-2021 data.

where j indexes dwellings and t indexes years; the indicator G_j equals one if dwelling j holds an EPC rating of G. Main effects G_j, R_j, and $G_j \times R_j$ are time-invariant and are absorbed by the parcel fixed effects λ_j. Rental status R_j equals one if the apartment is observed as rented at least once during the period of observation, using both pre-announcement and post-announcement occupancy records. This definition identifies apartments that ever participate in the rental market.[14] Finally, A_t is an indicator for the post-announcement period, equal to one for years 2021–2023 and zero otherwise; parcel fixed effects λ_j absorb all time-invariant unobservables at the dwelling (parcel) level, while λ_t controls for year fixed effects; the error term ε_{jt} captures idiosyncratic shocks to the probability of sale.

The reform of the DPE in 2021 altered the measurement and distribution of EPC ratings, raising comparability concerns when evaluating the policy announcement's impact. Our triple-difference (DDD) design is intended to mitigate these concerns by exploiting variation across three dimensions: (i) time (pre- vs. post-announcement), (ii) treatment intensity (EPC rating category), and (iii) rental status prior to sale. The design is credible only if the measurement and coverage shifts introduced by the 2021 reform affect rented and non-rented dwellings symmetrically within a given EPC category. Under that assumption, differencing across rental status within EPC groups, and then differencing these effects between treated and control EPC categories, nets out spurious changes in measured rating composition that are unrelated to the announcement. We therefore interpret the DDD coefficients as evidence consistent with an announcement effect, rather than as fully isolating the causal effect of the policy in a setting free of measurement concerns.

Figure 1 provides descriptive evidence on pre-policy trends. Panel (a) plots the share of apartments that are rented at least once during the observation window (R_j=1) that are sold each year by EPC rating. The pre-announcement period exhibits parallel trends across EPC categories, with only modest differences in levels. Following the 2021 announcement, sales shares for rented EPC G units rise sharply relative to better-rated units (EPC B–D), consistent with anticipatory divestment by landlords facing future rental restrictions.

Panel (b) repeats the exercise for apartments that were never rented during the period of observation (R_j=0). Here also, pre-announcement trends are broadly parallel, but a noticeable shift appears around 2021 across all EPC categories. This shift is plausibly driven by the reclassification and coverage changes documented above rather than a true behavioral response. The DDD design absorbs this common shift by contrasting the change in rented apartments against the contemporaneous change for never-rented apartments within the same EPC category.

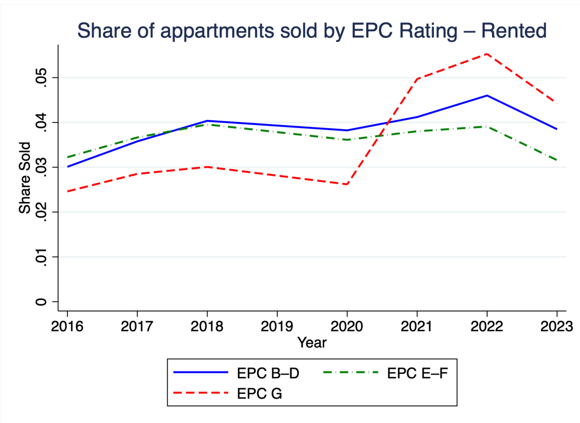
Table 2 reports the estimated coefficients from equation (1). The coefficient on *Rented* $\times$ *Announcement* is negative and significant, reflecting that, absent treatment, rented dwellings experienced a slightly smaller increase in sales probability than never-rented dwellings post-announcement. The triple-interaction term *Rented* $\times$ *Announcement* $\times$ *EPC G* is positive and

[14] We also estimate specifications that define rental status using pre-announcement years only; estimated announcement effects are very similar. We prefer the "ever rented" definition because it better captures the relevant rental market over our observation window.

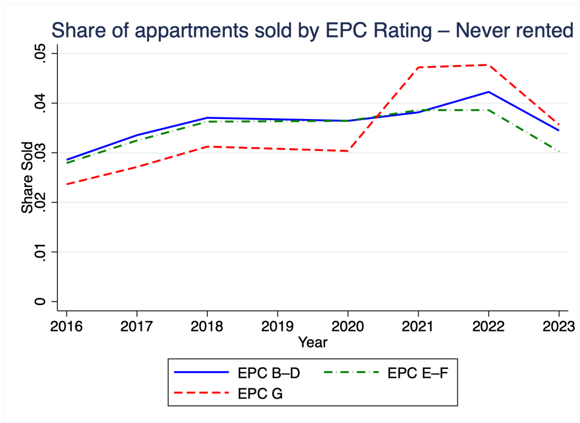
highly significant (0.0072, $p < 0.01$), indicating that rented EPC G dwellings became 0.72 percentage points more likely to be sold post-announcement relative to the within-EPC-group counterfactual of never-rented dwellings. Given that the annual sale probability for EPC G dwellings rented prior to the announcement is around 2.9%, this corresponds to an increase of roughly 25%. The corresponding estimate for EPC E–F in column (2) is smaller (-0.00329) but still statistically significant at the 1% level. Relative to their baseline sale probability of about 3.8%, this effect represents a decrease of approximately 8.6%.

These results are consistent with the descriptive patterns: the announcement is associated with a marked increase in sales probability for low-rated rental properties, with the largest effects concentrated in the most severely rated category (EPC G). By construction, the DDD estimates net out shifts in measured EPC composition induced by the 2021 methodological reform under the maintained symmetry assumption, and the results are therefore consistent with a behavioral response to anticipated regulation. The negative coefficient for EPC E–F should be interpreted more cautiously because those categories were less directly exposed in the near term and may remain more sensitive to residual composition changes.

Figure 1: Annual share of apartments sold, by EPC rating, metropolitan France 2016–2023.



(a) Apartments rented at least once (R_j=1).



(b) Apartments never rented (R_j=0).

Notes: Annual share of apartments sold (sold in year t / total stock) by EPC rating, metropolitan France, 2016–2023. The sample is restricted to apartments satisfying the filters described in Section 3. The vertical dashed line marks 2021, the year of the policy announcement. Panel (a) shows apartments rented at least once over the observation window ; Panel (b) shows apartments never rented. Data source: CEREMA transactions linked to ADEME EPC records.

4.2 Effect of announcement on dwelling use

To assess the impact of the 2021 policy announcement on post-purchase use, we estimate an event-study model on the sample of sold apartments. Our outcome is a dwelling-use indicator for rental status, Use^{rent}_{jt}, equal to one if apartment j is rented in year t and zero otherwise. We compare EPC G apartments to EPC B–D apartments before and after the announcement in the years surrounding the sale. We define $T = 0$ as the year of sale and include a broad set of event-time indicators, while focusing the post-announcement triple interactions on the window from three years before to three years after the transaction. The omitted category is four years

Table 2: Effect of Policy Announcement on Probability of Sale

	(1) G	(2) E+F
Rented x Announcement	-0.00461***	-0.00461***
	(0.00046)	(0.00062)
Rented x Announcement x EPC G	0.0072***	
	(0.00149)	
Rented x Announcement x EPC E|F		-0.00329***
		(0.00063)
Constant	0.0381***	0.0387***
	(0.000177)	(0.000199)
Observations	15,362,105	23,364,740
R-squared	0.115	0.116

Notes: This table reports OLS estimates (LPM) for the probability of dwelling sale. Rented is defined as rented at least once during the period of observation. Sample period is 2016-2023. All models control dwelling fixed effects and time fixed effects. All standard errors are clustered at the commune level. *** p<0.01, ** p<0.05, * p<0.1.

before the sale ($T \leq -4$), ensuring that coefficients capture differential changes in use relative to a pre-sale baseline. Because this design is estimated only on transacted units, it describes reallocation conditional on sale; it does not separately identify how much of the observed change comes from who sells versus who buys. Formally, we estimate:

$$
\begin{aligned}
Use_{jt}^{rent} = & \sum_{k=-7}^{+7} \beta_k^S \times \mathbb{1}\{T_{jt} = k\} \\
& + \sum_{k=-7}^{+7} \beta_k^{SG} \left[G_j \times \mathbb{1}\{T_{jt} = k\}\right] \\
& + \sum_{k=-7}^{+7} \beta_k^{SA} \left[A_t \times \mathbb{1}\{T_{jt} = k\}\right] \\
& + \sum_{k=-3}^{+2} \beta_k^{SAG} \left[A_t \times G_j \times \mathbb{1}\{T_{jt} = k\}\right] + \lambda_j + \lambda_t + \varepsilon_{jt},
\end{aligned} \tag{2}
$$

where j indexes dwellings and t indexes years. The dependent variable Use_{jt}^{rent} equals one if dwelling j is rented in year t and zero otherwise. The indicator G_j equals one if the dwelling holds an EPC rating of G (the omitted group is EPC B–D). $\mathbb{1}\{T_{jt} = k\}$ is an event-time dummy equal to one if year t is exactly k years from the transaction date of dwelling j. A_t is an indicator for post-announcement calendar years, equal to one for 2021–2023 and zero otherwise. The β_k^{SAG} are the event-time-specific interaction coefficients between the announcement and EPC G status at event time k; these are our main coefficients of interest and we focus on $k \geq 0$ to describe post-sale reallocation. Parcel fixed effects λ_j absorb all time-invariant unobservables at the dwelling (parcel) level, while λ_t controls for year fixed effects. The error term ε_{jt} captures idiosyncratic shocks to dwelling use.

Beyond the EPC measurement changes introduced in 2021, a second complication arises from the progressive phase-out of the *taxe d'habitation* between 2018 and 2023.[15] This degradation may bias estimates if changes in reporting quality correlate with EPC category or sale timing. Our DiD strategy mitigates this concern by comparing rental before and after the announcement, and then differencing between treated and control EPC categories. Under the identifying assumption that any reporting deterioration affects dwellings symmetrically within EPC groups, this approach nets out bias from measurement changes unrelated to the policy.

Figure 2 plots event-study coefficients for rental use of EPC G apartments, using EPC B–D apartments as the reference group. In the pre-sale years ($T \leq -1$), coefficients are close to zero, supporting the parallel-trends assumption. Post-sale, rental use falls sharply for EPC G units: at $T = +1$, the probability is about 3.8 percentage points lower (roughly 10% relative to baseline) and remains negative through $T = +2$. Figures 3 and 4 report the corresponding event-study estimates for owner-occupied and secondary-home use. By $T = +2$, the owner-occupied share rises by around 2 percentage points and secondary-home use increases by about 1.5 percentage points, absorbing the remainder of the rental decline. By construction, changes across rental, owner-occupied, and secondary use categories sum to zero.

The DDD results indicate that the policy announcement is associated with a sizable reallocation of sold EPC G apartments away from the rental market post-sale, with most of the reduction absorbed by owner-occupation and a smaller share by secondary use. Taken together with the sales-probability results in Section 4.1, these findings provide a coherent picture of the announcement's impact. The earlier analysis showed that EPC G dwellings, especially those previously rented, were disproportionately sold after the policy announcement. Conditional on sale, the present results are consistent with those units transitioning toward owner-occupation and secondary use: roughly two-thirds of the reallocation appears to move into owner-occupation, and most of the remainder into second-home use. We view this pattern as suggestive of forward-looking reallocation in anticipation of the rental ban, while recognizing that part of the measured effect may also reflect selection in which units are sold after the announcement.

[15] Before the reform, local tax offices updated the dwelling use status annually to calculate the *taxe d'habitation*, which applied differently to primary residences, rentals, and secondary homes. As the tax was progressively abolished, first for lower-income households and then for almost all households, the fiscal incentive for accurate and timely updates declined. This led to a gradual degradation of the rental-status variable in the property tax records, particularly for households no longer subject to the tax, increasing the likelihood of outdated or missing rental information.

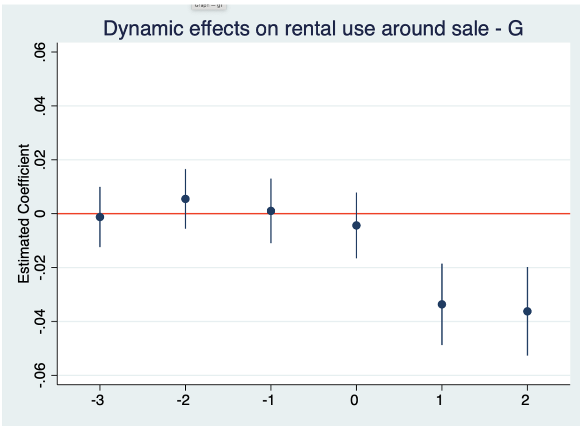


Figure 2: Event-study estimates for rental use of sold EPC G apartments

Notes: Triple-difference (DDD) event-study coefficients $\hat{\beta}_k^{SAG}$ from Equation (2). The dependent variable Use_{jt}^{rent} is an indicator equal to one if dwelling j is rented in year t (0/1). Coefficients measure the differential change in rental use for sold EPC G apartments relative to EPC B–D apartments, before and after the 2021 announcement ($T = 0$ is the year of sale). The y-axis is in percentage points. Bars indicate 95% confidence intervals. Sample: sold apartments in metropolitan France, 2016–2023. The omitted event-time category is $T \leq -4$.

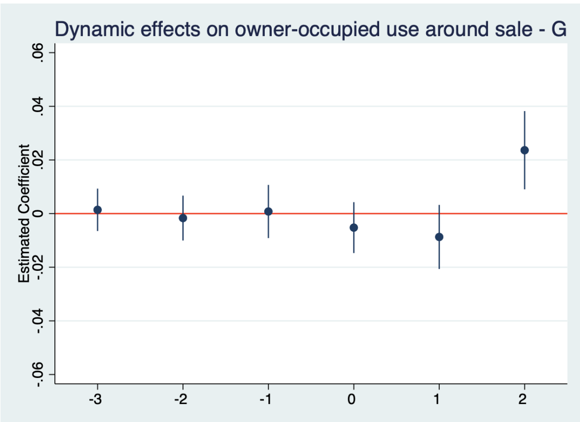


Figure 3: Event-study estimates for owner-occupied use of sold EPC G apartments.

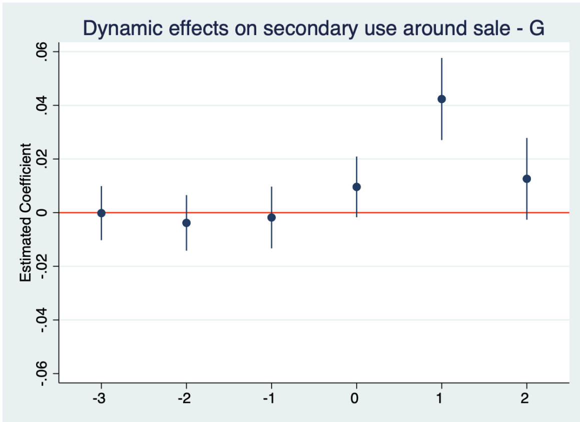


Figure 4: Event-study estimates for secondary-home use of sold EPC G apartments.

Notes to Figures 3 and 4: Notes: Triple-difference (DDD) event-study coefficients from Equation (2). The dependent variables are indicators for owner-occupied use (Figure 3) and secondary-home use (Figure 4), respectively. Coefficients measure the differential change in each tenure outcome for sold EPC G apartments relative to EPC B–D apartments. The y-axis is in percentage points. Bars indicate 95% confidence intervals. Sample: sold apartments in metropolitan France, 2016–2023. The omitted event-time category is $T \leq -4$.

4.3 Effect of announcement on prices

We next turn to the impact of the policy announcement on the pricing margin. This analysis evaluates whether anticipated restrictions on energy-inefficient dwellings were capitalized into transaction values. We estimate a hedonic price model where the dependent variable is the log of the sale price. The main explanatory variables are EPC class dummies and their interactions with the post-announcement indicator:

$$\ln(p_{jt}) = \sum_k \beta_k^E E_{jk} + \sum_k \beta_k^{AE} [A_t \times E_{jk}] + X_j\gamma_1 + A_t \times X_j\gamma_2 + \lambda_{nt} + \xi_{jt}, \tag{3}$$

where j indexes transactions and t indexes years. The dependent variable $\ln(p_{jt})$ is the natural logarithm of the transaction price. E_{jk} is an indicator equal to one if dwelling j is in EPC category k, with one category omitted as the reference group. The coefficients β_k^E capture average price differences by EPC category in the pre-policy period, while β_k^{AE} capture the differential changes in prices for each EPC category after the policy. The vector X_j includes dwelling characteristics such as size, property type, and tenure status. λ_{nt} denotes neighborhood-by-year fixed effects, which absorb all local time-varying shocks, such as seasonality, local economic conditions, or regional housing cycles. The error term ξ_{jt} captures idiosyncratic transaction-level shocks.

The analysis of prices relies on a difference-in-differences (DiD) framework that compares the change in transaction prices for each EPC class before and after the announcement, relative to EPC D dwellings as the reference category. While this approach absorbs time-varying local shocks through neighborhood-by-year fixed effects, it faces an important limitation: prices are set in an integrated local housing market, and expectations of future rental restrictions are likely to spill over to all EPC categories.

In this setting, the coefficients should be interpreted as *relative capitalization effects*, that is, the change in price for each EPC category relative to the reference group, rather than absolute causal effects on market prices. This caution is particularly relevant because two forces coincide: the 2021 methodological change in the EPC and the announcement of the progressive rental phase-out. Both may alter relative prices across EPC categories, and the DiD design cannot fully separate their influence.

Table 3 reports five specifications. Columns (1)–(3) reproduce the baseline analysis for all apartments, those rented at least once during the period of observation, and those never rented. Columns (4)–(5) extend the analysis by splitting municipalities according to their pre-period exposure to low-efficiency dwellings. Specifically, we first compute, for each municipality, the share of transactions classified as EPC E, F, or G prior to 2021. We then rank municipalities by this share across the full sales sample and define two subsamples: municipalities at or below the 25th percentile (low EPC-E+F+G density) and those at or above the 75th percentile (high density). This allows us to assess whether capitalization effects are stronger in markets more exposed to inefficient housing.

Across columns (1)-(3), the pre-announcement EPC coefficients display an efficiency gradient. Relative to EPC D, EPC B apartments sell for about 3.3% more in the full sample, EPC C for about 1.2%, while EPC E and F are not statistically different from D, and EPC G is slightly higher but imprecisely estimated. The gradient is similar across rented and never-rented subsamples, with a somewhat larger premium for EPC B among never-rented units (3.7%). Post-announcement interaction terms reveal a significant repricing of low-efficiency dwellings. In the full sample (column 1), EPC G apartments sell for 2.66% less after the announcement relative to EPC D ($p < 0.01$). The discount is stronger among previously rented units (-3.21%) and similar for never-rented units (-2.58%), suggesting broad market revaluation rather than a purely rental-segment effect. EPC F units lose about 1.54% in the full sample (-1.92% for rented), while

EPC E units decline by 1.46% (all statistically significant at conventional levels).

Conversely, post-announcement, higher-efficiency dwellings appreciate relative to EPC D. EPC B units gain about 1.57% in the full sample (1.86% among rented), and EPC C units increase by 0.68% overall, with stronger effects in the rented subsample (1.01%). These patterns are consistent with a forward-looking capitalization of future rental constraints.

Columns (4) and (5) show that these effects are heterogeneous across local housing markets. In municipalities with a low baseline share of EPC E+F+G dwellings ($\leq P_{25}$), the post-announcement discount for EPC G is small and not statistically significant (-0.6%), whereas EPC E units experience a larger relative decline (-2.25%). In contrast, in high-exposure municipalities ($\geq P_{75}$), EPC G apartments face a discount of 2.61%, and EPC F units decline by about 1.2%. The EPC E penalty remains negative but somewhat smaller in magnitude (-1.31%). Despite the difference between low- and high-exposure municipalities, it is not statistically significant , which may reflect limited statistical power rather than a homogeneous market-wide effect.

These patterns suggest that capitalization of the announcement is stronger for the worst-rated dwellings precisely in markets where inefficient units are more prevalent. One interpretation is that in high-density markets, buyers anticipate tighter local supply adjustments, stronger competition among low-efficiency units amplifying relative price discounts. In contrast, in low-density municipalities, market exposure to the regulation is more limited, attenuating the repricing of the very worst EPC category.

Reduced-form in a nutshell. Across the three outcomes, probability of sale, post-sale use, and transaction price, the evidence paints a consistent picture. Consistent with the DDD design, the announcement appears to have accelerated the sale of low-efficiency rented dwellings, shifted them away from the rental market and toward owner-occupation or secondary use, and lowered their relative market value while increasing the price premium for higher-efficiency units. These patterns are consistent with forward-looking reallocation of housing assets in anticipation of binding future rental restrictions.

Yet, the reduced-form hedonic regression recovers marginal willingness to pay only under restrictive conditions, such as homogeneous preferences or the absence of endogenous sorting on unobserved dwelling characteristics. When households differ in preferences and sort accordingly, the hedonic gradient does not generally identify structural valuation parameters. The coefficients above should therefore be interpreted as equilibrium capitalization effects in transaction prices. Observed price gradients embed both marginal valuation and sorting-induced composition effects. The structural model developed below relaxes these assumptions by explicitly modeling heterogeneous preferences and equilibrium sorting.

5 Sorting Model

Having established the reduced-form stylized facts, we now turn to the structural model. We estimate an equilibrium residential sorting model with three observed buyer types— owner-

Table 3: Impact of EPC Rating and Post-Announcement Period on Apartment Prices

	(1) All Appartments	(2) Rented At least once	(3) Rented Never	(4) EPC E-F-G density $\leq P_{25}$	(5) EPC E-F-G density $\geq P_{75}$
EPC class B	0.0329***	0.0304***	0.0365***	0.0219**	0.0291**
	(0.00723)	(0.00809)	(0.00854)	(0.0109)	(0.0130)
EPC class C	0.0123***	0.0109***	0.0119***	0.00754	0.00583
	(0.00320)	(0.00352)	(0.00457)	(0.00544)	(0.00542)
EPC class E	0.00627	0.00609	0.00640	0.00634	0.00512
	(0.00438)	(0.00496)	(0.00426)	(0.00545)	(0.00456)
EPC class F	0.00243	0.00236	0.000242	-0.00265	-0.00360
	(0.00602)	(0.00746)	(0.00628)	(0.00988)	(0.00458)
EPC class G	0.00854	0.00569	0.0144*	-0.0121	0.00273
	(0.00588)	(0.00643)	(0.00822)	(0.0246)	(0.00679)
EPC_B x After announcement	0.0157**	0.0186**	0.0153	0.0169	0.0148
	(0.00785)	(0.00927)	(0.0111)	(0.0109)	(0.0170)
EPC_C x After announcement	0.00682*	0.0101**	0.00548	0.00847	0.0186**
	(0.00402)	(0.00474)	(0.00517)	(0.00579)	(0.00768)
EPC_E x After announcement	-0.0146***	-0.0160***	-0.0150***	-0.0225***	-0.0131**
	(0.00380)	(0.00436)	(0.00457)	(0.00677)	(0.00593)
EPC_F x After announcement	-0.0154***	-0.0192***	-0.00898	-0.0166	-0.0120*
	(0.00461)	(0.00608)	(0.00628)	(0.0110)	(0.00619)
EPC_G x After announcement	-0.0266***	-0.0321***	-0.0258***	-0.00586	-0.0261***
	(0.00584)	(0.00660)	(0.00890)	(0.0286)	(0.00833)
Constant	7.591***	7.825***	7.272***	7.021***	8.109***
	(0.0608)	(0.0599)	(0.0734)	(0.118)	(0.0749)
Observations	963,281	600,807	329,437	236,820	243,486
R-squared	0.893	0.895	0.893	0.848	0.899

Notes: Dependent variable is the natural logarithm of the transaction price. $\leq P_{25}$ and $\geq P_{75}$ to the sub-sample of municipalities at or below the 25th percentile (low EPC-E+F+G density) and those at or above the 75th percentile (high density), respectively. All regressions include a full set of section-year fixed effects. Standard errors are clustered at the commune level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

occupiers, investors/landlords, and second-home buyers— and recover type-specific valuations of dwelling and neighborhood attributes using pre-policy data. The model is designed to answer a demand-side question: how the announcement changes who buys which dwellings and how prices adjust to clear the market. We therefore take buyer-type composition as exogenous (buyer type is observed in the data) and we do not model entry into the market or switching across types. On the supply side, we do not specify a structural model of construction or seller optimization; instead, we close the model with an exogenously calibrated supply curve that can accommodate an exogenous "supply shifter," which we hold fixed in counterfactual simulations. To reduce computation, we work with a subsample of 20 departments selected for computational tractability.

5.1 Dwelling's Demand

We specify a characteristic-based dwelling demand model, where preferences over dwelling units are parameterized as a function of both observed and unobserved property attributes and household characteristics (Berry et al., 1995). Let $i = 1, \ldots, I$ index individuals participating in the market. Each individual i considers the purchase of a single dwelling unit j from a feasible choice set $\mathcal{J}_i$ of dwellings that are affordable given their wealth y_{it} (buyer

i's budget at time t). We explicitly distinguish three buyer types, captured by a type mapping $\theta(i) \in \{\text{Owner-occupier (O)}, \text{ Investor/Landlord (I)}, \text{ Secondary-home (S)}\}$. We model the choice of *buyers*: each individual i chooses a dwelling j to maximize utility given their type $\theta(i)$. For investors ($\theta(i) = \text{I}$), the valuation of dwelling and neighborhood characteristics reflects beliefs about tenants' willingness to pay for those characteristics (i.e., expected rental cash flows net of costs).

The indirect utility for an individual i of type $\theta(i)$ who chooses dwelling unit j, located in neighborhood n, is:

$$U_{ijnt}^{\theta(i)} \;=\; V_{ijnt}^{\theta(i)}(y_{it} - p_{jnt},\, X_{jnt},\, Z_{it},\, N_{jnt}) \;+\; \xi_{jnt} \;+\; \epsilon_{ijnt}, \tag{4}$$

where X_{jnt} are dwelling characteristics, p_{jnt} is the transaction price, Z_{it} are individual i's observed characteristics (other than wealth), N_{jnt} are neighborhood attributes, ξ_{jnt} are unobserved characteristics such as expected cost of renovations, and ϵ_{ijnt} captures idiosyncratic variation. As a neighborhood's dwelling unit does not change, we abstract from the n subscript below. In the empirical implementation, the observed components of Z_{it} mainly capture buyer-type indicators and other purchaser information observed in the transactions and occupancy data. For investors, $V_{ijt}^{\theta(i)=\text{I}}$ should be interpreted as reduced-form expected profit utility that maps characteristics into anticipated rental demand (tenants' willingness to pay) and costs.

We assume a household chooses a dwelling unit j if the utility from that option is at least as large as the utility from any other dwelling unit $j^{'}$. That is,

$$U_{ijt}^{\theta(i)} \;\geq\; U_{ij't}^{\theta(i)} \qquad \forall j \neq j^{'}.$$

An individual's choice of a dwelling unit j depends on all the available, and affordable, choices and characteristics in the choice set of individual i. We model the purchase decision explicitly: each household either buys one dwelling from its feasible choice set or chooses not to purchase (the outside option). For this, we assume that the outside option has utility $U_{i0t}^{\theta(i)} = \epsilon_{i0t}$.

Next, we assume that the random shocks to utility, ϵ_{ijt}, are distributed i.i.d. Type I Extreme Value. Integrating over ϵ_{ijt} yields the closed-form choice probability of the individual's i purchase of dwelling unit j, as a function of the characteristics of all dwelling types in the market and of the buyer's type:

$$P_{ijt}^{\theta(i)}(\mathbf{X}, \mathbf{N}, Z_i, \mathbf{p}_t, \xi; \boldsymbol{\beta}) \;=\; \frac{\exp\big(V_{ijt}^{\theta(i)} + \xi_{jt}\big)}{1 + \sum_{j' \in \mathcal{J}_i} \exp\big(V_{ij't}^{\theta(i)} + \xi_{j't}\big)}\,, \tag{5}$$

where $\mathbf{X}$ is the matrix of dwelling characteristics, $\mathbf{N}$ is the matrix of neighborhood attributes, Z_i is the vector of household i's demographics, $\mathbf{p}_t$ is the vector of prices of all dwelling units, ξ collects unobserved quality terms, and $\boldsymbol{\beta}$ contains all parameters in the model. Note that individuals are allowed not to purchase a dwelling unit. By construction, type heterogeneity enters through $V_{ijt}^{\theta(i)}$: for owner-occupiers and secondary-home buyers it captures consumption value, while for investors it captures expected rental-market value based on beliefs about renters' willingness to pay for dwelling and neighborhood attributes.

5.2 Market-clearing conditions and the sorting equilibrium

In the dwelling market, choices of individual households aggregate to total dwelling demand, and dwelling prices adjust to equate demand and supply. The expected market demand for a specific dwelling unit j is obtained by aggregating individuals' choice probabilities P_{ijt} whose choice set includes property j:

$$D_{jt}(\mathbf{X}, \mathbf{N}, Z_i, \mathbf{p}, \boldsymbol{\xi}, \boldsymbol{\beta}) = \sum_i P_{ijt}^{\theta(i)}(\mathbf{X}, \mathbf{N}, Z_i, \mathbf{p}_t, \xi; \boldsymbol{\beta}),$$

which depends on the vector of dwelling prices $\mathbf{p}$ at period t.

We close the model by assuming a dwelling supply curve. We consider two scenarios for dwelling supply. In the first case, the dwelling supply is fixed: $S_j(\mathbf{p}) = 1$ (the supply of each property is one). In the second case, we assume that dwelling supply has a constant elasticity and adjusts at the dwelling unit level in response to the price. This assumption mimics owners' consideration to sell or not the dwelling unit at period t. Specifically, we assume that :

$$S_{jt} = \kappa_{jt} p_{jt}^{\varphi},$$

where κ_{jt} is a supply shifter and φ is the supply elasticity which we obtain based on prior studies. We treat this supply curve as a reduced-form representation of listing and selling behavior. We calibrate the elasticity parameter φ externally and interpret κ_{jt} as an exogenous shifter that captures baseline differences in sale propensity and market thickness. We do not use supply shifts for identification and we do not vary κ_{jt} in the policy counterfactuals.

Sorting equilibrium A sorting equilibrium is defined as a vector of dwelling prices, $\mathbf{p}^*$, such that, the dwelling market clears for all j properties:

$$D_{jt}(\mathbf{X}, \mathbf{N}, Z_i, \mathbf{p}^*, \boldsymbol{\xi}, \boldsymbol{\beta}) = S_{jt}(\mathbf{p}^*).$$

Although our model follows the class of equilibrium sorting models with local spillovers studied in Bayer and Timmins (2005) and more closely in Bayer et al. (2007), we do not yet allow for local spillovers to arise.[16]

[16] Following Jia et al. (2024), we address the possibility of multiple equilibria by solving our model with different starting values. Our results always converge to the same equilibrium outcomes, providing empirical evidence for uniqueness in our setting.

6 Estimating the sorting model

Estimation of the sorting model follows a two-step procedure. To simplify the exposition, we first empirically specify the indirect utility function and then provide a detailed description of the first and second steps of the estimation procedure.

6.1 Econometric implementation

We empirically specify the dwelling unit-level utility as a function of observed and unobserved dwelling characteristics, and time fixed effects. We write the indirect utility as:

$$V_{ijt}^{\theta(i)} = \Theta_{jt} + \Gamma_{ijt}^{\theta(i)} + \epsilon_{ijt} \tag{6}$$

where Θ_{jt} defines mean preferences for dwelling units of type j and period t, $\Gamma_{ijt}^{\theta(i)}$ defines household heterogeneity as idiosyncratic deviations from mean preferences, and ϵ_{ijt} rationalizes the remaining idiosyncratic variation. Formally, let the mean utility from purchasing dwelling unit j in period t, Θ_{jt}, be given by:

$$\Theta_{jt} = X_j\beta_X + N_j\beta_N - \alpha\ln(p_{jt}) + \xi_{jt}. \tag{7}$$

where X_j is a $J \times K$ matrix with K the dimension of the observed dwelling characteristics, and N_j is a $J \times L$ matrix with L the dimension of the observed neighborhood characteristics and ξ_{jt} captures the remaining unobserved variation.

We allow household heterogeneity to enter the model through dwelling and neighborhood attribute coefficients, which are functions of observed household characteristics. That is, $\Gamma_{ijt}^{\theta(i)}$ is given by:

$$\Gamma_{ijt}^{\theta(i)} = X_j\Sigma_{XZ}Z_{it} + N_j\Sigma_{NZ}Z_{it}, \tag{8}$$

where Z_{it} is a $Z \times 1$ vector containing observed Z individual characteristics, including preference shifters for buyer type (i.e., owner-occupier, second-home buyer, landlord), Σ_{XZ} is a $K \times Z$ matrix, and Σ_{NZ} is a $L \times Z$ matrix. Assuming that ϵ_{ijt} are independent and identically distributed extreme value, given Equations (7) and (8) the conditional logit probability of household i choosing dwelling j emerges as

$$P_{ijt}^{\theta(i)}(\mathbf{X}, \mathbf{N}, Z_i, \mathbf{p}_t, \xi; \boldsymbol{\beta}) = \frac{\exp(\Theta_{jt} + \Gamma_{ijt}^{\theta(i)})}{1 + \sum_{j' \in \mathcal{J}_i} \exp(\Theta_{j't} + \Gamma_{ij't}^{\theta(i)})}. \tag{9}$$

We assume that households' location decision is optimal given the location decision of all other households and market clearing prices.

We estimate the heterogeneous parameters and mean indirect utilities using maximum likelihood. The maximum likelihood estimator is based on maximizing the probability that the model correctly matches each household with its chosen dwelling choice (Bayer et al., 2007).

The log-likelihood function is given by:

$$LL = \sum_t \sum_i \sum_j Y_{ijt} \ln(P_{ijt}^{\theta(i)}), \tag{10}$$

where Y_{ijt} is an indicator variable that equals one if household i chooses dwelling choice jt in the data and zero otherwise. Standard errors for the non-linear parameters are computed using the sandwich (outer-product-of-gradients) estimator, which corrects for potential misspecification by weighting the inverse Hessian by the outer product of individual score contributions.

6.2 Setting up the first stage

The maximum-likelihood estimation is well defined, but we face several empirical challenges. First, because we consider the universe of dwellings transacted in 20 French departments between 2016 and 2020, we have a large number of choice occasions drawn from an even larger choice set, which creates a dimensionality problem.

Computational and data constraints require limiting the number of alternatives in demand estimation. A natural approach would restrict each buyer's choice set to nearby or affordable dwellings. However, such attribute-based restrictions may introduce bias if the choice set is defined using variables correlated with unobserved preferences (Banzhaf and Smith, 2007). We therefore rely on choice-based sampling, which delivers consistent estimates in multinomial logit and mixed logit models under appropriate weighting (Wasi & Keane, 2012; Guevara & Ben-Akiva, 2013). Specifically, for each purchase we draw a 1% random sample of dwellings sold within a two-month window around the transaction date, yielding an average choice set of approximately 40 properties.[17] In our departments, there were 66,612 transactions over 2016–2020, implying roughly 1,300 properties within a two-month window. Including all such dwellings would generate implausibly large choice sets and substantially increase computational burden.

A second issue is the estimation of all dwelling-type mean taste parameters. We deal with it using a contraction mapping routine. When a full set of alternative-specific constants is included, the first order conditions for maximum likelihood in the conditional logit model imply that the model's share of individuals choosing dwelling unit j must match the sample average probability for dwelling type j. This equality condition allows the recovery of Θ_{jt} at each iteration of the search routine, while the remaining parameters are updated through the usual gradient-based step (Berry 1994; Bayer et al., 2007; Klaiber et al. 2010).[18]

[17] An alternative approach would define discrete dwelling types—e.g., by location, size, and time—rather than individual properties as alternatives.

[18] The contraction mapping and price-instrument equilibrium computation both normalize market size by assuming that for each realized purchase there are 20 potential buyers who opted for the outside option (non-purchase). This implies that the no-purchase probability is approximately 20 times the purchase probability for any given dwelling. The normalization affects the predicted level of market shares (the absolute probability that a given buyer type purchases a given dwelling), but leaves all relative choice probabilities across alternatives and buyer types unchanged—a direct consequence of the independence of irrelevant alternatives (IIA) property of the logit model. Structural parameters and counterfactual ratios are therefore invariant to this normalization.

6.3 Setting up the second stage

Once mean taste parameters are recovered, the second step of the estimation decomposes Θ_{jt} into observable and unobservable components according to Equation (7). There is a direct algebraic relationship between the hedonic model in Equation (3) and the structural model in Equation (7). Re-arranging Equation (7) yields:

$$\ln(p_{jt}) + \frac{\Theta_{jt}}{\alpha} = X_j \frac{\beta_X}{\alpha} + N_j \frac{\beta_N}{\alpha} + \frac{\xi_{jt}}{\alpha}. \tag{11}$$

That is, when preferences are heterogeneous, the estimated mean indirect utility Θ_{jt} provides an adjustment to the hedonic price equation and ensures that the price regression accurately captures mean preferences.

Nevertheless, to consistently estimate Equation (11) we need to deal with the correlation between log price, $\ln(p_{jt})$, and the unobserved component, ξ_{jt}. We are unlikely to capture all variables that affect the household's purchase decision. These unobserved attributes are likely to be correlated with the observed price. As we are dealing with used dwellings, a strong candidate for this correlation is the size of refurbishment a household needs to carry out on the property prior to moving in. Households will consider the full price of the dwelling, which includes the transaction cost and the expected cost of renovations, but our data only include transaction costs.[19] The larger the expected costs of renovations, the smaller the transaction cost. Because of this, it is necessary to find a suitable instrument for prices and estimate the second stage via an instrumental-variables approach.

Following the recent literature on structural demand estimation, we exploit the spatial structure of the dwelling market, along with assumptions on the pricing structure of the market. First, we assume that observed prices are the result of an equilibrating process that depends on dwelling-type attributes from across the extent of the market. Second, we assume that characteristics of dwellings and neighborhoods that are sufficiently far away influence the equilibrium in the dwelling market, hence affecting prices, but have no direct effect on the utility of purchasing a dwelling. Provided these two assumptions hold, exogenous attributes of distant neighborhoods provide potentially good instruments for price.

Accordingly, we construct observed variables with different neighborhood attributes from across the market. Let $\tilde{X}$ and $\tilde{N}$ denote observed dwelling and neighborhood characteristics for neighborhoods located in distance bins relative to dwelling jt: less than 1 km (used as controls only), 1–2 km, 2–3 km, 3–6 km, 6–10 km, and 10–20 km. To make the instrument operational, we implement the following algorithm:[20]

Step 1. Set a starting value for the price coefficient, say α_o

[19]People spend on average between 150 and 450 euros per square metre for a light renovation (plumbing, painting, etc.). For partial renovations (changing the layout of the rooms, renovating the kitchen or bathroom, moving or creating a shower room, etc.), spending is on average 550 to 1000 euros per square meter. For a total renovation of a used dwelling, a budget between 1100 and 1700 euros per square meter is typically required.

[20]This algorithm allows supply for dwellings to vary with price.

Step 2. Estimate $\Theta_{jt} + \alpha_o \ln(p_{jt}) = \tilde{X}_{jt}\beta_X + \tilde{N}_{jt}\beta_N + \tilde{\xi}_{jt}$ using OLS.

Step 3. Obtain $\hat{\Theta}_{jt}^{net}$, the net-of-price predicted mean tastes from step 2 , i.e., the fitted value $(\Theta_{jt} \widehat{+ \alpha_o \ln} p_{jt})$ from the OLS in Step 2.

Step 4. Generate the price instruments solving for the predicted price $\mathbf{p}^{iv}$ that satisfies the market-clearing condition: $D_{jt}(\tilde{\mathbf{X}}, \tilde{\mathbf{N}}, \mathbf{p}^{iv}, \boldsymbol{\xi}{=}\mathbf{0}, \hat{\boldsymbol{\beta}}) = S_{jt}$.

Step 5. Estimate $\Theta_{jt} = X_j\beta_X + N_j\beta_N - \alpha \ln(p_{jt}) + \xi_{jt}$ using $\ln(p_{jt}^{iv})$ as an IV for $\ln(p_{jt})$.

Step 6. Update α_o using the estimated $\hat{\alpha}$ from step 5.

Step 7. Repeat step 2 to step 6 until convergence criterion is satisfied.

According to our algorithm, the instrumental variable $\mathbf{p}^{iv}$ effectively captures the residual variation in Step 4 that comes from outside the 1 km ring. By design, this variation reflects the market characteristics outside the specified control rings, and it is assumed to be exogenous. Consequently, the instrument's power originates from the variability of these neighborhood attributes located outside the rings. Although these external attributes are considered exogenous due to their distance, they are linked with prices through the overall housing market equilibrium.

A potential issue with this method is that creating the price instrument relies on having an initial estimate for the very price coefficient we are trying to determine. To address this circularity and remove dependence on the initial guess, our algorithm uses an iterative procedure. We start by constructing the instrument using an initial guess and run the IV estimation. The resulting price coefficient estimate is then used as a new, updated guess to rebuild the price instrument and re-run the entire IV process. This cycle is repeated multiple times. Repeating the process ensures the price coefficient estimate settles to a stable value, making the final result independent of the initial guess. The algorithm converges when $|\Delta\hat{\alpha}| < 10^{-6}$ across successive iterations; in our application, convergence is achieved within a moderate number of iterations (up to 100 maximum).

Our instruments are constructed as follows: we assign every dwelling to a set of mutually exclusive distance bins relative to the focal dwelling: less than 1 km; 1–2 km; 2–3 km; 3–6 km; 6–10 km; 10–20 km; and beyond 20 km. The <1 km bin is used for direct controls and is excluded from instrument construction; the beyond-20 km bin is defined but likewise excluded from instrument construction. The first set of instruments aggregates the average characteristics of dwellings in the surrounding donut region between 1 km and 20 km (bins 2–6): the theoretical rental value, the average construction year, and the total surface area, each averaged across the five bins; each of these three variables enters the instrument set separately (as $\log(\cdot + 1)$ transforms). The second instrument measures market thickness as the total number of dwellings sold within the same 1–20 km donut (also entered as $\log(\cdot + 1)$). By construction, these variables capture variation in broader neighborhood and market attributes that is plausibly exogenous to unobserved shocks affecting the focal dwelling, while still being correlated with local prices through housing market equilibrium.

6.4 Results from the structural model

In the following section we report the first stage non-linear coefficients that capture heterogeneity in preferences across buyer types using a conditional logit specification. In the second stage, we recover mean utility parameters via a linear regression, estimated either by OLS or IV.

6.4.1 Linear Estimates (Second Stage)

Table 4 reports the second-stage linear estimates from the structural demand model, with Columns (1)–(3) corresponding respectively to OLS, IV using direct instruments, and IV using market-equilibrium instruments.

The OLS estimates in Column (1) exhibit patterns that are both counterintuitive and implausible from a demand perspective, most notably, a positive and large coefficient on price. This reflects the severe endogeneity of prices in our setting: the estimation sample is restricted to dwellings that were sold, and within this set, higher prices are mechanically correlated with unobserved quality attributes that also increase choice probability. In other words, the OLS specification confounds willingness to pay with the selection of higher-quality, more desirable units into the observed transaction set. As a result, OLS price effects are upward biased and uninterpretable.

Columns (2) and (3) address price endogeneity using instrumental variables. Column (2) implements a standard two-stage least squares (2SLS) specification in which log price is instrumented using the spatial instruments defined in the previous section (i.e., "direct instruments"). Column (3) instead uses the market-equilibrium instrument constructed by our algorithm. This instrument exploits variation in neighborhood attributes outside the 1 km control ring, capturing price movements arising from equilibrium adjustments in the broader housing market. Using direct instruments (Column 2) yields a negative and significant price coefficient, while the market-equilibrium instruments in Column (3) produce an even larger magnitude, consistent with economic theory.[21] Using Column (3), where the estimated coefficient on $\ln(p)$ is -6.278, the EPC penalties relative to EPC B translate into the following WTP losses: C: $-$€10.8k, D: $-$€12.0k, E: $-$€12.3k, F: $-$€13.6k, G: $-$€14.0k. These magnitudes are economically coherent, larger (more negative) for worse labels, and align with the reduced-form price results as well as the policy's rental-ban timeline. Turning to other attributes, living space has a large and significant positive coefficient, while older properties generally attract lower valuations, all else equal. Age effects follow a monotonic pattern, with the steepest discount for pre-1949 stock.

The status-transition variables (e.g., Rent $\rightarrow$ Rent) capture tenure changes between 2016 and 2024 and allow for systematic differences in unobserved quality across tenure histories. Units allocated to rental and owner-occupied segments may differ persistently in dimensions not fully captured by observed characteristics, including maintenance intensity, amenity levels, or building quality.[22] The transition indicators flexibly control for such tenure-specific heterogeneity.

[21] Because utility depends on $\ln p$, the implied marginal willingness to pay for attribute k is $\text{WTP}(\bar{p}) = -(\beta_X^k/\alpha)\bar{p}$. We evaluate at $\bar{p} =$ €100K, the average price of apartments in our 20 department sample.

[22] *Rent $\rightarrow$ Rent* equals one if the dwelling was rented in 2016 and remains rented in 2024; *Home $\rightarrow$ Rent* equals

All estimated transition coefficients are negative , implying a discount relative to continuously owner-occupied units. Continuously owner-occupied properties command a premium. Properties that remain in the rental sector (Rent → Rent) display one of the largest discounts. Units transitioning into owner-occupation (Rent → Home) also trade at a discount relative to stable owner-occupied dwellings.

Table 4: Results for linear coefficients

	(1) Model OLS	(2) Model IV Direct	(3) Model IV Market
Log Price (in 2019 euros, in 100,000s)	0.005	-1.961	-6.278
	(0.006)	(0.031)	(0.123)
Rental value (in 2019 euros, in 100s)	-0.128	0.121	0.668
	(0.003)	(0.005)	(0.018)
Living space (in 100s m^2)	1.248	2.292	4.582
	(0.021)	(0.036)	(0.104)
EPC N	0.163	-0.702	-2.601
	(0.025)	(0.042)	(0.114)
EPC C	0.056	-0.136	-0.556
	(0.022)	(0.034)	(0.087)
EPC D	0.123	-0.110	-0.621
	(0.021)	(0.033)	(0.085)
EPC E	0.138	-0.105	-0.638
	(0.021)	(0.033)	(0.086)
EPC F	0.097	-0.153	-0.701
	(0.022)	(0.035)	(0.090)
EPC G	0.170	-0.111	-0.726
	(0.025)	(0.039)	(0.101)
1949–1974	0.193	0.321	0.602
	(0.008)	(0.013)	(0.032)
1975–1988	0.085	0.618	1.789
	(0.010)	(0.018)	(0.051)
1989–2000	0.164	0.690	1.845
	(0.012)	(0.020)	(0.057)
2001–2005	0.242	0.796	2.012
	(0.010)	(0.018)	(0.054)
2006–2025	0.041	0.732	2.249
	(0.034)	(0.054)	(0.142)
Rent → Rent	-0.118	-0.497	-1.328
	(0.008)	(0.014)	(0.040)
Rent → Home	-0.009	-0.184	-0.567
	(0.009)	(0.014)	(0.036)
Rent → Empty	-0.095	-0.546	-1.536
	(0.016)	(0.026)	(0.070)
Home → Rent	-0.014	-0.277	-0.856
	(0.010)	(0.017)	(0.044)
Home → Empty	-0.012	-0.277	-0.859
	(0.018)	(0.028)	(0.072)
Empty → Empty	-0.075	-0.768	-2.290
	(0.023)	(0.038)	(0.101)

Notes: This table reports the second set of estimates from the structural demand estimation. The dependent variable is equal to one when the individual is observed to select the alternative, and zero otherwise. The choice set of each individual consists of 50 alternatives: 1 the choice she selects + 49 random alternatives. The period of analysis is between 2016 and 2020. The EPC of reference is B. The estimates are obtained using a two-step method reported in section 6. All standard errors are clustered at the commune level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Overall, the IV results are internally consistent and align with theory: price sensitivity is negative and significant, energy-inefficient dwellings are penalized in proportion to their EPC

one if it was owner-occupied in 2016 but rented in 2024 (i.e., the year 2024 is out-of-sample); etc. The omitted category is *Owner → Owner*.

rating, larger and newer properties command higher valuations, and stable owner-occupied histories are valued most highly.

6.4.2 Non-Linear Estimates (First Stage)

Table 5 reports the non-linear coefficients from the first-stage logit model, allowing preferences to vary by buyer type. The reference category is owner-occupiers purchasing a primary residence; "Secondary" refers to second-home buyers, while "Investor" denotes purchasers intending to rent out the unit.[23]

The coefficient on *Rental value*, a cadastral measure that incorporates structural quality, location, amenities, and upkeep[24], is positive for all buyer types but notably smaller for investors (-0.104 relative to owner-occupiers) than for secondary buyers (-0.134). This suggests that while higher-quality properties are valued across the board, investors place less incremental value on quality improvements, consistent with a stronger focus on yield maximization rather than consumption amenities.

The *Living space* interaction reveals marked heterogeneity: investors value additional space less than owner-occupiers, with a large and significant negative coefficient (-0.430), while secondary-home buyers value it more ($+0.154$). This aligns with the idea that investors seek smaller, more easily rentable units, whereas secondary buyers, purchasing for leisure or lifestyle, are willing to pay more for additional space.

Energy-performance interactions (EPC $\times$ Buyer Type) are generally small and imprecise, suggesting that relative preferences for EPC classes do not differ strongly across buyer types once other attributes are controlled for. The mild negative loadings for EPC F and G among investors (-0.120 and -0.027) hint at greater sensitivity to low energy efficiency, potentially due to anticipated regulatory constraints on rentals.

Age-of-construction interactions show a consistent pattern: both secondary buyers and investors assign lower utility to newer stock, with the largest penalties for post-2001–2005 dwellings among second-home buyers and for pre-1974 dwellings among investors. The relatively favorable valuation of 2006–2025 dwellings by investors ($+0.147$) is consistent with the impact of the *Loi Pinel*,[25] which provides tax incentives for investment in new or rehabilitated rental properties.

Overall, these results confirm substantial heterogeneity in attribute valuation across buyer types: investors prioritize smaller, potentially higher-yield properties and are less responsive to quality upgrades, while secondary buyers emphasize space and comfort; both groups discount older stock, and investors show some preference for newer properties potentially linked to tax-

[23] We identify the intention of purchase by looking at the usage-status in 2024. Hence, an owner is identified as an investor if the purchased property is rented in 2024.

[24] The theoretical rental value is an administrative construct used as the basis for local taxation in France. It is derived from a weighted surface measure that adjusts the actual area for qualitative factors such as building category, architectural character, state of repair, equipment, and location within the municipality. See Appendix B for full details. It is therefore a broad index of intrinsic dwelling quality rather than a direct market rent measure.

[25] The *Loi Pinel*, introduced in 2014, offers income-tax reductions to investors purchasing new or rehabilitated properties for rental, subject to rent caps and tenant income ceilings. This scheme likely increased demand for newer units in eligible areas.

incentive schemes.

Table 5: Results for non-linear coefficients

	Secondary	Investor
Rental value × Buyer Type	-0.134	-0.104
	(0.009)	(0.005)
Living space × Buyer Type	0.154	-0.432
	(0.068)	(0.034)
EPC N × Buyer Type	0.075	0.009
	(0.065)	(0.031)
EPC C × Buyer Type	-0.045	-0.016
	(0.059)	(0.026)
EPC D × Buyer Type	-0.033	0.019
	(0.057)	(0.026)
EPC E × Buyer Type	-0.029	-0.003
	(0.058)	(0.026)
EPC F × Buyer Type	-0.075	-0.125
	(0.060)	(0.028)
EPC G × Buyer Type	-0.026	-0.033
	(0.068)	(0.032)
1949–1974 × Buyer Type	-0.096	-0.133
	(0.024)	(0.012)
1975–1988 × Buyer Type	-0.171	-0.069
	(0.030)	(0.013)
1989–2000 × Buyer Type	-0.157	-0.124
	(0.034)	(0.015)
2001–2005 × Buyer Type	-0.197	-0.066
	(0.032)	(0.014)
2006–2025 × Buyer Type	-0.236	0.144
	(0.105)	(0.038)

Notes: This table reports the first set of estimates (i.e., non-linear estimates) from the structural demand estimation. The dependent variable is equal to one when the individual is observed to select the alternative, and zero otherwise. The choice set of each individual consists of 49 alternatives: 1 the choice she selects + 49 random alternatives. The period of analysis is between 2016 and 2020. The EPC of reference is B. All standard errors are clustered at the commune level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

7 Counterfactual Simulation Results

Before presenting the simulations, it is useful to recall the main insight from the reduced-form analysis: the 2021 announcement of rental restrictions for EPC G dwellings is associated with a measurable decline in their rental-oriented transactions, consistent with investors exiting rather than with stronger owner-occupier demand. This pattern provides a natural target for calibrating the structural model. The counterfactual exercises below do not provide a second independent estimate of the policy effect. Instead, they translate those reduced-form moments into an equilibrium sorting framework under explicit assumptions about renovation costs, buyer heterogeneity, and supply. Because we estimate the structural model on a sample of 20 departments, we re-estimate the reduced-form effects on this subsample and match the structural model to the corresponding reduced-form moments.

7.1 Linking to Reduced-Form Evidence and Calibration

We first connect the structural model to the reduced-form results presented in Section 4. The reduced-form estimates documented an increase in EPC G sales following the 2021 announcement, which we interpret as consistent with investor divestment when retrofitting is costly or unattractive. Such behavior could be represented as a shift in the supply function, an exogenous change in the composition of units offered for sale. Our current framework does not explicitly incorporate such supply shifts; instead, we use the model to replicate the observed change in the share of units sold for rental purposes.

We compute this share directly from the simulated purchase probabilities by EPC class and buyer type. In the baseline, investors account for approximately 30.2% of EPC G transactions.[26] We then calibrate the renovation cost so that the simulated decline in rental-oriented EPC G acquisitions matches the reduction estimated in the reduced form.

The *Cost* scenario in Table 6 applies only the monetary component of renovation costs to buyers' utility, holding non-monetary attributes fixed. The *Cost + Comfort* scenario additionally incorporates the utility gain associated with upgrading to EPC D, capturing potential improvements in comfort, health, and energy efficiency. The comparison between these specifications is therefore informative about how much of the simulated response comes from the utility value of an improved dwelling rather than from the monetary cost alone. We calibrate renovation costs under the *Cost + Comfort* scenario using a bisection algorithm on the renovation cost parameter so that the simulated change in landlords' share of EPC G acquisitions equals -3 percentage points, matching the reduced-form estimate computed on the 20-department structural subsample. The corresponding national estimate is approximately 3.8 percentage points (Section 4). This implies an average renovation cost of approximately €92.5/m^2, with upgraded units assumed to reach EPC D and to receive the corresponding change in latent utility.[27] For EPC E and EPC F, by contrast, the values of €40/m^2 and €70/m^2 should be read as scenario assumptions rather than estimates disciplined by an external calibration target.

We implement counterfactual equilibria using two solvers that differ in their supply-side treatment. For the *Baseline* sub-scenario, we use a fixed-supply solver ($S_j = 1$) that recovers the price vector clearing demand at pre-policy characteristics. For the *Cost* and *Cost + Comfort* sub-scenarios, we use a variable-supply solver with the constant-elasticity supply curve $S_{jt} = \kappa_{jt} p_{jt}^{\varphi}$ described in Section 5, where the elasticity $\varphi = 0.53$ is taken from Almagro, Chyn and Stuart (2023) and κ_{jt} is normalised so that supply equals one at the observed pre-policy price. In the variable-supply solver, the renovation cost enters via the price.

[26] Computed as Prob(G $\times$ Landlord)/(Prob(G $\times$ Owner) + Prob(G $\times$ Second) + Prob(G $\times$ Landlord)) using baseline values from Table 6.

[27] Formally, the latent utility change u_{change} applied to affected landlord-owned dwellings combines two components: (i) the mean utility gain from upgrading EPC class, given by $\hat{\beta}^{\text{EPC}_D} - \hat{\beta}^{\text{EPC}_G}$ (the linear coefficient difference, common across buyer types); and (ii) the investor-specific non-linear utility gain $\hat{\sigma}^{\text{EPC}_D \times \text{Investor}} - \hat{\sigma}^{\text{EPC}_G \times \text{Investor}}$, using only the Investor (LO) buyer-type interaction terms. The non-linear component is restricted to the investor type because the ban directly affects the utility of the investor who owns the property and must bear renovation costs; the owner-occupier and secondary-buyer response to the EPC upgrade is captured entirely by the common mean utility shift (i).

The resulting reallocation across buyer types is not targeted. The model implies increases in purchases by owner-occupiers and second-home buyers of about 2.5 and 0.5 percentage points, respectively. These shifts are reflected directly in Table 6: under the "Cost + Comfort–G" counterfactual, EPC G purchases by owner-occupiers and second-home buyers rise by about 3.7% and 5.7% (ratios 1.037 and 1.057).[28] While the qualitative ranking across buyer types aligns with the reduced-form evidence, the model overstates the owner-to-landlord and secondary-to-landlord differentials for EPC G units, suggesting that some adjustment in the data may reflect frictions or heterogeneity in regulatory awareness not captured by the framework.

Table 6: Counterfactual Simulation Results: simulated phaseout EPC G, and E, F, G

	Baseline	EPC G changes only		EPC E, F, G change	
		Cost	Cost + Comfort	Cost	Cost + Comfort
Prob. EPC N × Owner	0.0009547	0.99967	1.0018	0.99786	1.0156
Prob. EPC B × Owner	0.00061857	0.99986	1.0014	0.99895	1.014
Prob. EPC C × Owner	0.0032828	0.99986	1.0015	0.99892	1.0141
Prob. EPC D × Owner	0.0097206	0.99984	1.0015	0.99884	1.0143
Prob. EPC E × Owner	0.008968	0.99982	1.0016	1.0368	1.0301
Prob. EPC F × Owner	0.0041157	0.99983	1.0015	1.055	1.0188
Prob. EPC G × Owner	0.0012197	1.0752	1.0366	1.0743	1.0476
Prob. EPC N × Second	0.00022592	0.99963	1.0019	0.99768	1.0158
Prob. EPC B × Second	9.0069e-05	0.99981	1.0015	0.99875	1.0144
Prob. EPC C × Second	0.00043243	0.9998	1.0015	0.99867	1.0149
Prob. EPC D × Second	0.001316	0.99978	1.0016	0.99856	1.0153
Prob. EPC E × Second	0.001194	0.99976	1.0018	1.046	1.038
Prob. EPC F × Second	0.00050636	0.99978	1.0017	1.0714	1.0282
Prob. EPC G × Second	0.00015825	1.1052	1.0567	1.104	1.0695
Prob. EPC N × Landlord	0.00076872	1.0005	0.99672	1.0034	0.97188
Prob. EPC B × Landlord	0.00044806	1.0002	0.99736	1.0017	0.97431
Prob. EPC C × Landlord	0.0021133	1.0003	0.99703	1.002	0.97133
Prob. EPC D × Landlord	0.0061795	1.0003	0.99683	1.0022	0.97043
Prob. EPC E × Landlord	0.00511	1.0004	0.99628	0.91385	0.92935
Prob. EPC F × Landlord	0.0019819	1.0004	0.99591	0.84985	0.94719
Prob. EPC G × Landlord	0.00059544	0.79309	0.89703	0.79519	0.86746
Outside option	0.95	1	1	1.0001	1.0001

Notes: This table reports counterfactual simulations in which the policy-induced renovation requirement is assumed to affect (1) EPC G dwellings and (2) EPC E, F and G. We assume that the required upgrades move the EPC to EPC D, with renovation costs of 92.5/m^2, 70/m^2, and 40/m^2, respectively. The *Baseline* column reports the model-implied probability that a unit is purchased by EPC class and buyer type. Counterfactual columns report values as ratios relative to the baseline. The *Cost* scenario assumes only a cost shock on the demand side, while the *Cost + Comfort* case also allows for a utility change from the lower EPC to EPC D.

7.2 Phasing Out EPC E, F, and G

We next simulate a broader policy tightening in which units rated E, F, and G all require renovation to at least EPC D to remain rentable. We assume renovation costs of €92.5/m^2 for G-to-D, €70/m^2 for F-to-D, and €40/m^2 for E-to-D, with utility changes reflecting both the monetary cost and the upgrade in non-price attributes as above.

The results in Table 6 show that this extension of the ban produces heterogeneous adjustments across buyer types. Landlords experience the sharpest contraction in the targeted

[28]By construction, landlord purchases of EPC G units fall by about 10.3% (ratio 0.897).

segment: their probability of acquiring EPC G units falls by about 13.3% relative to the baseline (ratio 0.867), while EPC F purchases decline by about 5.3% (ratio 0.947) and EPC E by about 7.1% (ratio 0.929). When aggregating E–G purchases, the reduction for landlords reaches about 7.1% (Table 7), with a smaller 2.9% decline in their purchases of B–D dwellings.

In contrast, owner-occupiers and second-home buyers expand their presence in the affected segment. For example, the probability that an owner purchases an EPC G unit increases by about 4.8% (ratio 1.048), while EPC E and EPC F acquisitions rise by about 3.0% and 1.9%, respectively (ratios 1.030 and 1.019). Second-home buyers show a similar but more pronounced shift toward the worst-rated class: EPC G purchases rise by about 7.0% (ratio 1.070), while EPC E rises by about 3.8% and EPC F by about 2.8% (ratios 1.038 and 1.028). As reported in Table 7, aggregating the direct probabilities from the baseline and counterfactual, we observe an increase in E–G purchases by 2.9% and B–D by 1.4% for owner-occupiers, while second-home buyers expand by 3.9% and 1.5%, respectively. This confirms that landlords are the main group exiting the low-efficiency segment, with the resulting slack absorbed by non-investor buyers.

Relative to the first counterfactual—where only EPC G was targeted—the inclusion of E and F amplifies these shifts. The effects on EPC G itself become larger in absolute terms despite the identical assumed cost shock. Within our model, this arises from equilibrium feedback: removing E and F from the effective rental stock further tightens the set of options available to landlords, pushing more of them out of the market altogether. The resulting drop in simulated investor demand for G-rated units implies that owner-occupiers and second-home buyers increase their equilibrium purchase shares for lower-rated dwellings.

Table 7: Aggregated baseline and counterfactual purchase probabilities by buyer type and EPC group: Cost + Comfort scenario (EPC E, F, G ban)

	EPC B–D			EPC E–G		
Buyer type	Baseline	Policy	% Change	Baseline	Policy	% Change
Owner	0.0146	0.0148	+1.43%	0.0143	0.0147	+2.83%
Second-home	0.0021	0.0021	+1.52%	0.0019	0.0019	+3.80%
Landlord	0.0095	0.0092	−2.91%	0.0077	0.0071	−7.08%

Notes: This table reports baseline and counterfactual purchase probabilities and their percentage change under the Cost + Comfort scenario in which EPC E, F, and G dwellings all require renovation to EPC D. "Baseline" refers to the pre-policy sorting equilibrium; "Policy" refers to the counterfactual equilibrium after the ban. "% Change" is defined as (Policy/Baseline − 1) × 100. Purchase probability is the unconditional probability that a buyer of a given type purchases a dwelling in the specified EPC group. Buyer types are identified from post-sale occupancy status in 2024: a purchaser is classified as a Landlord if the dwelling is observed as rented in 2024, as Second-home if vacant/secondary, and as Owner otherwise. "EPC B–D" aggregates classes B through D; "EPC E–G" aggregates E through G. Entries are simulated quantities from the structural model estimated on the 2016–2020 pre-policy sample from a sample of 20 departments (selected for computational tractability; see Section 6); no standard errors are reported. Assumed renovation costs: €92.5/m^2 for G-to-D (calibrated to match a 3 pp decline in landlord purchases in the 10-department subsample; the national estimate is approximately 3.8 pp), €70/m^2 for F-to-D (assumed), €40/m^2 for E-to-D (assumed).

7.3 Impact on equilibrium prices

Table 8 reports hedonic regressions of simulated equilibrium transaction prices on dwelling attributes, focusing on energy performance certificates (EPC). The dependent variable is the logarithm of transaction price, so coefficients on EPC dummies can be interpreted as log price differentials relative to the omitted category (EPC B). In the two counterfactual columns, entries report coefficient multipliers relative to the baseline. For example, the EPC G entry of 1.050 under "Cost + Comfort–G" implies that the EPC G coefficient scales from -0.143 in the baseline to -0.150 in the counterfactual.

Comparison with reduced-form results for EPC G. We first compare the structural counterfactual where only EPC G dwellings face increased renovation costs and utility gains ("Cost + Comfort–G") with the reduced-form capitalization patterns in Table 3. In the reduced form, the post-announcement interaction implies that EPC G apartments become about 2.7% cheaper after the announcement, relative to EPC D (and about 3.2% in the rental-market subsample). In the structural "Cost + Comfort–G" scenario, the EPC G coefficient becomes 5% more negative than in the baseline (multiplier 1.050), while the EPC D coefficient is essentially unchanged (multiplier 1).[29] Put differently, the model matches the direction of capitalization for the targeted class, but generates limited spillovers to higher-rated dwellings in this short-run exercise.

This gap reflects that the reduced-form interaction captures a relative repricing at the time of the announcement in an integrated housing market, where transaction composition, endogenous sorting, and the contemporaneous change in EPC methodology can all affect observed price gradients. Our structural counterfactual instead isolates the incremental equilibrium repricing implied by the calibrated cost-and-comfort shock, holding the housing stock fixed and abstracting from additional supply-side adjustments. The two objects therefore need not coincide, and the reduced-form coefficients should be interpreted as equilibrium capitalization patterns rather than structural willingness-to-pay parameters.

Extending to EFG changes. The third column of Table 8 extends the counterfactual to EPC E, F, and G dwellings ("Cost + Comfort–EFG"), applying renovation costs of €40/m^2 to E, €70/m^2 to F, and €92.5/m^2 to G. In this broader counterfactual, the EPC coefficients become more negative for the newly regulated categories: the EPC G coefficient increases in magnitude by 4.7% (multiplier 1.047), while EPC E and EPC F become 2.6% and 1.1% more negative than in the baseline (multipliers 1.026 and 1.011). Overall, the model implies that equilibrium repricing remains concentrated in the regulated segment, with limited ripple effects on higher-rated categories.

[29] In Table 8, the baseline EPC G coefficient is -0.143; under "Cost + Comfort–G" it becomes $-0.143 \times 1.050 \approx -0.150$.

Table 8: Hedonic Regression Estimates Under Counterfactual Renovation Costs

	Baseline	Cost+Comfort-G	Cost+Comfort-EFG
Rental value (in 2019 euros, in 100s)	0.127	1.001	1.006
	(0.001)		
Living space (in 100s m^2)	0.531	0.998	0.990
	(0.013)		
EPC N	-0.440	1.000	0.999
	(0.015)		
EPC C	-0.097	1.000	1.001
	(0.013)		
EPC D	-0.118	1.000	1.001
	(0.013)		
EPC E	-0.123	1.000	1.026
	(0.013)		
EPC F	-0.127	1.000	1.011
	(0.014)		
EPC G	-0.143	1.050	1.047
	(0.015)		
1949–1974	0.065	1.001	1.020
	(0.005)		
1975–1988	0.271	1.000	1.007
	(0.006)		
1989–2000	0.268	1.000	1.006
	(0.007)		
2001–2005	0.282	1.000	1.004
	(0.006)		
2006–2025	0.352	1.000	1.002
	(0.021)		
Rent → Rent	-0.193	1.001	1.008
	(0.005)		
Rent → Home	-0.089	1.001	1.004
	(0.005)		
Rent → Empty	-0.229	1.001	1.005
	(0.010)		
Home → Rent	-0.134	1.002	1.011
	(0.006)		
Home → Empty	-0.135	1.001	1.004
	(0.011)		
Empty → Empty	-0.353	1.001	1.006
	(0.014)		

Note: This table reports estimates from a hedonic price regression where the dependent variable is the logarithm of transaction price. Each column reflects a scenario: the baseline, a counterfactual with increased renovation costs for EPC E/F/G homes (by 40/70/92.5 per m^2), and a third specification that includes in addition the appropriate utility changes.

8 Conclusion

Housing market regulations that target the energy performance of dwellings have the potential to reshape both the composition of the housing stock and the distribution of ownership. At the same time, the way households and investors respond to such regulations determines their effectiveness and efficiency. This paper develops and applies an equilibrium sorting framework to evaluate the impacts of France's phased ban on renting energy-inefficient dwellings. The framework incorporates preference heterogeneity across buyer types and accounts for general equilibrium feedback between the sales market and the rental sector.

Our analysis delivers three main conclusions. First, the reduced-form evidence documents three distinct market responses to the 2021 announcement. On the extensive margin, the annual probability of sale for rented EPC G dwellings rises by 0.72 percentage points—a 25% increase relative to the pre-announcement baseline—consistent with investor divestment. On the intensive margin, use after purchase shifts: rental-oriented purchases of G-rated dwellings fall by about 3.8 percentage points, which translates into roughly a 10% reduction in the share of newly purchased G-rated apartments entering the rental market. Conditional on sale, owner-occupiers and second-home buyers absorb most of the reallocation away from landlord purchases, with only limited positive spillovers toward more efficient dwellings. Finally, prices re-sort across EPC classes in a pattern consistent with capitalization of expected renovation and compliance costs: EPC G apartments become roughly 2.7% cheaper relative to EPC D after the announcement (3.2% in the rental-market subsample), while EPC B and EPC C dwellings appreciate by approximately 1.6% and 0.7%, respectively.

Second, the structural model provides a way to map those short-run moments into longer-run equilibrium scenarios, but its conclusions remain conditional on strong assumptions. In particular, buyer-type composition is taken as exogenous, renovation costs for EPC E and EPC F are assumed rather than estimated, and the model is calibrated on a 10-department subsample rather than the full national market. The counterfactual results should therefore be interpreted as model-based scenarios rather than reduced-form estimates.

Third, within that framework, extending the rental ban from EPC G to the full E–F–G schedule amplifies market reallocation. Aggregating across EPC E, F, and G, landlord purchase probabilities fall by about 7.1% while those of owner-occupiers and second-home buyers rise by 2.8% and 3.8%, respectively. Even within the cleaner B–D segment, landlord purchases contract by 2.9%, reflecting equilibrium feedback as options for investors narrow. Price effects are concentrated in the regulated segment: EPC G coefficients become 4.7% more negative in absolute terms, with EPC E and EPC F following at 2.6% and 1.1%. At the same time, the present framework does not model rental-market supply shortages, tenant welfare, or realized post-sale renovation outcomes directly.

The calibrated renovation cost of €92.5/m^2 for a G-to-D upgrade warrants specific discussion, as it is small relative to observed renovation expenditures. It is useful to distinguish two types of renovation work. Cosmetic renovation—painting, minor fixtures, kitchen and bathroom refits—typically range from €150/m^2 to €800/m^2 depending on scope. Energy renovations are generally more costly: a deep intervention combining full insulation, window replacement, and a heating system upgrade runs roughly €300/m^2 for an apartment, while the more targeted objective of upgrading a dwelling from G to D requires around €150/m^2 for an apartment (Civel, 2019). Against these benchmarks, the €92.5/m^2 implied by our calibration sits on the lower end of this range. One interpretation consistent with this gap is that investors acquiring G-rated dwellings after the announcement may not view the rental ban as fully credible. If they expect the policy to be only partially enforced, they discount the expected compliance costs accordingly. Continued investor activity in the G-rated segment would therefore not necessarily reflect ignorance of renovation costs, but rather a rational response to policy uncertainty or to expected post-renovation income gains.

Our analysis also leaves out several channels that could matter for the long-run effects of the regulation, including new construction, endogenous migration across regions, and landlord-tenant matching. Quantifying those channels would require additional data and modeling extensions, which we leave for future research Finally, the paper documents reallocation of EPC G dwellings away from the rental market but does not model rental-market supply shortages or tenant welfare directly. The documented increase in the probability of sale for rented EPC G units (0.72 percentage points, or 25% relative to baseline) implies a meaningful reduction in the stock of rentable G-rated units over the period. Whether this supply contraction translates into upward pressure on rents for remaining low-rated units, or is absorbed by substitution toward higher-rated rentals, is an open empirical question that we leave for future work. Readers should therefore interpret the reallocation documented here as a market-structure effect, not as evidence of welfare improvement or deterioration for tenants.

References

Allcott, Hunt, and Michael Greenstone. 2012. “The Energy Efficiency Gap.” *Journal of Economic Perspectives*, 26(1): 3–28.

Almagro, Milena, Eric Chyn, and Bryan Stuart. 2023. “Urban Renewal and Inequality: Evidence from Chicago’s Public Housing Demolitions.” FRB of Philadelphia Working Paper.

Andrulis, John. 1982. “Intra-urban workplace and residential mobility under uncertainty.” *Journal of Urban Economics*, 11(1): 85–97.

Atasoy, Ayse. 2020. “Behavioral responses of green builders to discontinuous certification schemes.” *Resource and Energy Economics*, 60(C): S0928765518300708.

Aydin, Erdal, Dirk Brounen, and Nils Kok. 2020. “The capitalization of energy efficiency: Evidence from the housing market.” *Journal of Urban Economics*, 117: 103243.

Bartel, Ann P. 1979. “The Migration Decision: What Role Does Job Mobility Play?” *The American Economic Review*, 69(5): 775–786.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. “A unified framework for measuring preferences for schools and neighborhoods.” *Journal of political economy*, 115(4): 588–638.

Berry, Steven T. 1994. “Estimating discrete-choice models of product differentiation.” *The RAND Journal of Economics*, 242–262.

Brounen, Dirk, and Nils Kok. 2011. “On the economics of energy labels in the housing market.” *Journal of Environmental Economics and Management*, 62(2): 166–179.

Civel, Edouard. 2019. “The Red and the Green: essays on the economics of information in the sustainable habitat market.” PhD diss. Université de Nanterre-Paris X.

Civel, Edouard, Anna Creti, Gabrielle Fack, and Daniel Hererra. 2025. “Certification, manipulation and competition: evidence from Energy Performance Certificates.” CEPR Working Paper.

Clara, Nuno, João F. Cocco, S Lakshmi Naaraayanan, and Varun Sharma. 2025. “Investments that Make our Homes Greener: The Role of Regulation.” NBB Economic Review.

Collins, Matthew, and John Curtis. 2018. “Bunching of residential building energy performance certificates at threshold values.” *Applied Energy*, 211(C): 662–676.

Comerford, David A., Ian Lange, and Mirko Moro. 2018. “Proof of concept that requiring energy labels for dwellings can induce retrofitting.” *Energy Economics*, 69: 204–212.

Commission, European. 2017. “European Union Energy in figures.” *Statistical pocketbook*.

Crane, Randall. 1996. “The Influence of Uncertain Job Location on Urban Form and the Journey to Work.” *Journal of Urban Economics*, 39(3): 342–356.

Crawford, Gregory S, Rachel Griffith, and Alessandro Iaria. 2021. “A survey of preference estimation with unobserved choice set heterogeneity.” *Journal of Econometrics*, 222(1): 4–43.

Crawley, Jenny, Phillip Biddulph, Paul J. Northrop, Jez Wingfield, Tadj Oreszczyn, and Cliff Elwell. 2019. “Quantifying the Measurement Error on England and Wales EPC Ratings.” *Energies*, 12(18).

Davis, Lucas W., Alan Fuchs, and Paul Gertler. 2020. “How Accurate Are Energy-Efficiency Predictions? Evidence from a Field Test.” *American Economic Review*, 110(7): 2106–2144.

Eichholtz, Piet, Nils Kok, and John M Quigley. 2010. “Doing well by doing good? Green office buildings.” *American Economic Review*, 100(5): 2492–2509.

Ferentinos, Konstantinos, Alex Gibberd, and Benjamin Guin. 2023. “Stranded houses? The price effect of a minimum energy efficiency standard.” *Energy Economics*.

Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram. 2018. “Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program.” *Quarterly Journal of Economics*, 133(3): 1597–1644.

Fuerst, Franz, Patrick McAllister, Anupam Nanda, and Peter Wyatt. 2015. “Does energy efficiency matter to home-buyers? An investigation of EPC ratings and transaction prices in England.” *Energy Economics*, 48: 145–156.

Hardy, A., and D. Glew. 2019. “An analysis of errors in the Energy Performance certificate database.” *Energy Policy*, 129: 1168–1178.

Hyland, Marie, Ronan C Lyons, and Sean Lyons. 2013. “The value of domestic building energy efficiency? Evidence from Ireland.” *Energy Economics*, 40: 943–952.

Kahn, Matthew E, and Nils Kok. 2014. “The capitalization of green labels in the California housing market.” *Regional Science and Urban Economics*, 47: 25–34.

Kan, Kamhon. 2002. “Residential mobility with job location uncertainty.” *Journal of Urban Economics*, 52(3): 501–523.

Klaiber, H Allen, and Daniel J Phaneuf. 2010. “Valuing open space in a residential sorting model of the Twin Cities.” *Journal of environmental economics and management*, 60(2): 57–77.

Le Saout, Ronan, Bérangère Mesqui, and Rathle Jean-Philippe. 2022. “Le Parc de Logement par Classe Energétique au 1er janvier 2022.” Observatoire National de la Transition Energétique Working Paper.

Linneman, Peter, and Philip E. Graves. 1983. "Migration and job change: A multinomial logit approach." *Journal of Urban Economics*, 14(3): 263–279.

Lu, X., and C. Spaenjers. 2025. "Energy labels and heuristic hedonic models in the housing market." *Real Estate Economics.*

Myers, Erica, Steven L. Puller, and Jeremy West. 2022. "Mandatory disclosure and the housing market: Evidence from energy performance certificates." *Journal of Environmental Economics and Management*, 112: 102613.

ONPE. n.d.. "Le tableau de bord 2018." *Observatoire National de la Precarite energetique.*

Reusens, P., F. Vastmans, J. Vandenbergh, T. van Kempen, and S. Damen. 2025. "Impact of the Flemish energy renovation obligation on house prices." NBB Economic Review.

Rosen, Sherwin. 1974. "Hedonic prices and implicit markets: product differentiation in pure competition." *Journal of political economy*, 82(1): 34–55.

Sejas-Portillo, Rodolfo, Mirko Moro, and Till Stowasser. 2025. "The Simpler the Better? Threshold Effects of Energy Labels on Property Prices and Energy-Efficiency Investments." *American Economic Journal: Applied Economics.*

Siegel, Jay. 1975. "Intrametropolitan migration: A simultaneous model of employment and residential location of white and black households." *Journal of Urban Economics*, 2(1): 29–47.

Simpson, Wayne. 1980. "A simultaneous model of workplace and residential location incorporating job search." *Journal of Urban Economics*, 8(3): 330–349.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process. Statement: During the preparation of this work the authors used Claude in order to check the consistency between the code and the manuscript, as well as to improve the writing of the paper. After using this tool/service, the authors reviewed and edited the content as needed and takes full responsibility for the content of the published article.

A Appendix: Dataset construction

As not all relevant information is contained in a single dataset, we need to combine different data sources. Moreover, as information is not always available at the same spatial and temporal scales, some assumptions are required to merge the information from different sources into a single dataset. Our datasets include land-use, real estate transaction, and parcel-map data for France. These data can be readily matched at a highly disaggregated level. In contrast, as there is no unique identifier between the dataset containing household energy-efficiency information and the real estate transaction dataset, we are constrained to use a closest-distance matching algorithm. The following section describes each dataset used, and the steps taken to clean and merge the different sources of information.

A.1 Land use data

We exploit land use data provided by the Center for Studies and Expertise on Risk, Environment, Mobility and Urban Development, CEREMA.[30] The raw database is produced by the tax authorities on the basis of tax returns, which is then restructured and enriched by the CEREMA. Land use data provide in great detail land usage in France, the premises and the property rights linked to them. Updated annually, they constitute snapshots of the French territory on January 1st of each year.[31] We exploit data from 2016 to 2024 for the French metropolitan area. The reduced-form analyses use the national sample, while the structural model later focuses on a 10-department subsample for computational reasons.

A.2 Real estate transaction data

Real estate transaction data contain an exhaustive list of prices and characteristics from all real estate and land transactions in France from 2016 to 2023.[32] Characteristics include the type of sale (i.e., heritage, standard,...), location of the property, surface area of the parcels and premises concerned, the types of premises, and the number of rooms, amongst many others.

[30] The CEREMA is a public establishment under the supervision of the Ministry of Ecological Transition and Territorial Cohesion. It supports the State and local authorities in the development, deployment and evaluation of public planning and transport policies. Website: https://www.cerema.fr/fr/cerema [last visited 04 April 2025].

[31] The version used in this study is the 2024 version.

[32] This dataset was initially produced under the name DVF by the Direction Générale des Finances Publiques (DGFiP) from two sources of tax datasets: (1) the Fichier Informatisé des Données Juridiques Immobilières (FIDJI) dataset and the MAJIC (Mise à Jour des Informations Cadastrales) database. Now it is administered by the CEREMA. The CEREMA improved on the initial DVF by enriching it with data from the land-use database. As the linked database is not exhaustive, we supplement this dataset with information from the land-use dataset.

Our dataset includes all real estate transactions, with three exceptions: non-expensive transfers, transfers that do not go through a notary, and "complex" transfers.[33] The transactions are characterized at the scale of a parcel, and they include information on the evolution of the properties after the sale.

After removing rare properties or specific sales conditions, we implement two types of filters for our sample selection. Our first filter deals with the physical characteristics of the transacted dwellings. That is, only dwellings with the following conditions are kept: (1) transactions involving a single dwelling; (2) dwellings with fewer than three floors; (3) transactions involving fewer than eight parcels; (4) dwellings with at least one main room (bedroom or dining room); (5) dwellings with fewer than three kitchens; (6) dwellings with fewer than three garages.

Finally, we limit our sample to dwellings with a price per square meter between 280 euros and 4200 euros, a land value greater than or equal to 10,000 euros and an estimated rental value greater than 150 euros.[34]

A.3 EPC dataset

Collected by ADEME since 2010, the database includes information on the diagnosis, the expert executing the certification, and the building concerned. For the period 2010–2024, the raw dataset contains roughly 13 million EPCs from the French metropolitan departments (one observation corresponds to one diagnosis) and a set of more than 80 variables detailing property characteristics. Our analysis focuses on the subset relevant for 2016–2023 transactions. The main variables of interest for the study are the location of the property, the estimated energy consumption, the estimated greenhouse gas emissions, the year of construction of the building, its living area, and the method used for the diagnosis.

We estimate that about 20% of the EPCs have incorrect geographical coordinates for the dataset prior to 2021. To correct this problem, we use the "Site National des Adresses" (SNA) API that matches addresses with correct geographic coordinates.[35] We take the following steps:

Step 1: We constitute a single dataset containing addresses of all French departments in our dataset.

Step 2: For each postal code, we determine the minimum and maximum longitude and latitude. We assume that the postal code is correctly included in the dataset. Next, we define wrong coordinates as those observations not within the boundaries of the postal code. That is, a wrong coordinate is one that exceeds the minimum and maximum longitude and latitudes per postal code. Since cities are not squares, it is only an approximation.

Step 3: We keep only the observations with latitude and longitude outside the approximate borders of postal code, and observations with missing latitudes or longitudes. We

[33] Complex transfers correspond to parcels with very large condominiums or with an unresolved dispute.

[34] Rental value is the theoretical monthly rent that could be obtained from a property if it were let without a vacancy period. The appendix provides a detailed description on how rental value is estimated.

[35] The SNA contains all the addresses in France. Website: https://adresse.data.gouv.fr [Last visited February 2025.]

recalculate the longitude and latitude using the SNA's API on this subset of observations.

We reduced the number of incorrect observations to 5%, and we remove them from our sample. Next, all missing EPCs, dwellings with negative energy consumption, and dwellings with negative greenhouse gas emissions were excluded from the sample. As our focus is on used dwellings, we excluded all dwellings with EPC belonging to class A.

A.4 Merging EPC, land use and transaction datasets

As there is not a unique identifier between the EPC dataset and the real estate transaction dataset, we exploit the characteristics available in both data sets to combine them. For a given transaction, we identify the closest EPC using the following characteristics: (1) geographic coordinates of the building concerned; (2) its year of construction. The matching between datasets is done using the following algorithm:

Step 1. Identify all EPC created within [-10, 0] of the year the transaction occurred.[36]

Step 2. Within the set identified by Step 1, match the transaction with the closest EPC as measured by distance.

Using the matched dataset, we restrict the dataset to dwellings where energy consumption is between 0 and 700 kWh per m2 per year. We also restrict the analysis to dwellings with a maximum of 200 kg CO2/m2/year of greenhouse gas emissions.

A.5 Summary statistics

Table 9 presents descriptive statistics for the variables used in the analysis. The table is divided into two panels to distinguish between the broader building stock (Panel A) and the observed transaction sample (Panel B).

Panel A summarises the dwelling stock after assigning an energy performance certificate (EPC) to as many units as possible. Direct EPC matches from administrative records are complemented with imputed ratings for dwellings sharing the same parcel, construction period, surface-area quantile, and apartment/house status. This approach relies on the assumption that units within such narrowly defined groups exhibit similar thermal characteristics, which is especially plausible for apartments built under standardised designs. The resulting coverage allows for stock-level measures of policy exposure. In this imputed dataset, about 38% of dwellings are in the policy-targeted E–G range, with 5% rated EPC G. Average floor area is around $56\,\mathrm{m}^2$, and tenure is split between owner-occupied (46%), rented (44%), and second-home or vacant (10%). The age distribution of the stock is skewed toward older buildings: one-third were built between 1949–1974, and only about 5% after 2012.

Panel B restricts attention to the 364,644 dwellings transacted during the study period, which form the basis for the econometric analysis. The mean transaction price is 177,000 euros

[36]For our period of analysis, an EPC is valid for a period of 10 years.

(in 2019 euros), with an average theoretical rental value of 612 euros per month. Transacted dwellings are slightly larger on average ($59\,\text{m}^2$) than the stock overall, and the distribution of EPC ratings is more favourable: only 4.1% are in class G, compared to 8.9% in class F and 24.8% in class E. Roughly one-third of sales involve rental properties, just over half are owner-occupied, and the remainder are second homes or vacant units. Amenities are relatively common: 28% have a kitchen over $9\,\text{m}^2$, 70% feature a terrace, and 23% include a garage.

By construction, Panel A is suited to evaluating the stock-wide reach of regulatory measures, while Panel B provides the transaction-level variation needed to identify policy effects in the housing market. Overall, the comparison between Panel A and Panel B highlights the broader representativeness of the imputed dataset for stock-level analysis, while Panel B captures the narrower set of properties that enter the transaction market, which may differ in energy performance and other characteristics.

B Appendix: Computation of the Theoretical Rental Value

The *valeur locative cadastrale* (theoretical rental value) is an administrative estimate of the annual rent a property could generate if let continuously at market conditions prevailing on 1 January 1970. It is calculated for each dwelling by the French tax administration and serves as the basis for local property taxation (e.g., *taxe foncière*, formerly *taxe d'habitation*). Although it is not a direct measure of current market rents, it constitutes a standardized, quality-adjusted rent index that remains comparable across time and space.

The computation process combines physical dwelling characteristics, comfort features, and locational factors into a single quality-adjusted measure. It begins with the actual floor area, which is adjusted to obtain a *weighted surface area* by applying coefficients that reflect qualitative attributes such as building category, construction quality, architectural design, internal layout, and equipment. These adjustments apply both to main rooms and to outbuildings such as garages, balconies, or cellars.

Next, *comfort coefficients* are added to capture the presence of amenities. Each feature increases the weighted surface by a fixed number of square metres—for example, gas supply (+2 m^2), running water (+4), electricity (+2), shower (+4), bathtub (+5), washbasins (+3), toilets (+3), and per-room heating (+2). These bonuses also extend to outbuildings.

The resulting surface measure is then modified through *condition and location corrections*. These take into account the dwelling's state of maintenance, its location within the municipality, and the presence of a lift. Such adjustments, defined in Articles 324 Q, 324 R, and 324 S of Annex III of the General Tax Code, are expressed as additive coefficients that further refine the weighted surface. Once this net weighted surface area is obtained, it is multiplied by the rental value per square metre of a reference dwelling in the same municipality. Reference rates are determined by the *Commission Communale des Impôts Directs* for each building category.

All dwellings are classified into one of eight national *building categories*, ranging from (1)

Table 9: Summary statistics for key variables

Panel A: Characteristics on extended panel

Variable	Mean	SD	Min	Max
Rented at least once	0.620	0.485	0	1
EPC class G	0.0519	0.222	0	1
EPC classes E–G	0.409	0.492	0	1
Living space (in m^2)	54.43	26.88	1	646
Rented	0.467	0.499	0	1
Second-home or empty	0.104	0.305	0	1
Owner-occupied	0.429	0.495	0	1
Period of construction: 0–1948	0.224	0.417	0	1
Period of construction: 1949–1974	0.277	0.448	0	1
Period of construction: 1975–1988	0.158	0.365	0	1
Period of construction: 1989–2000	0.131	0.338	0	1
Period of construction: 2001–2005	0.0511	0.220	0	1
Period of construction: 2006–2012	0.104	0.306	0	1
Period of construction: +2012	0.0540	0.226	0	1

Panel B: Variables for transacted dwellings

Variable	Mean	SD	Min	Max
Transaction value (in 2019 euros)	233,077	255,799	9,620	1.026e+07
Rental value (in 2019 euros)	712.2	426.5	144.3	6,629
Rented at least once	0.637	0.481	0	1
EPC class B	0.0273	0.163	0	1
EPC class C	0.163	0.369	0	1
EPC class D	0.346	0.476	0	1
EPC class E	0.259	0.438	0	1
EPC class F	0.0985	0.298	0	1
EPC class G	0.0509	0.220	0	1
EPC class N	0.0556	0.229	0	1
Living space (m^2)	56.65	26.59	9	200
Number of rooms	2.597	1.133	1	8
Number of dining rooms	0.992	0.364	0	8
Presence kitchen larger than 9 m^2	0.221	0.415	0	1
Presence of garages	0.255	0.436	0	1
Floor number	2.287	1.924	0	10
Number of floors of building	4.541	2.771	0	52
Rented	0.360	0.480	0	1
Owner-occupied	0.500	0.500	0	1
Second-home or empty	0.132	0.338	0	1
Rented 2016 – Rented 2023	0.197	0.398	0	1
Rented 2016 – Owner-occupied 2023	0.154	0.361	0	1
Rented 2016 – Second-home 2023	0.0616	0.240	0	1
Owner-occupied 2016 – Owner-occupied 2023	0.251	0.434	0	1
Owner-occupied 2016 – Rented 2023	0.0723	0.259	0	1
Owner-occupied 2016 – Second-home 2023	0.0429	0.203	0	1
Second-home 2016 – Second-home 2023	0.0144	0.119	0	1
Second-home 2016 – Rented 2023	0.0398	0.195	0	1
Second-home 2016 – Owner-occupied 2023	0.0355	0.185	0	1
Period of construction: 0–1948	0.259	0.438	0	1
Period of construction: 1949–1974	0.249	0.432	0	1
Period of construction: 1989–2000	0.125	0.331	0	1
Period of construction: 2001–2005	0.107	0.309	0	1
Period of construction: 2006–2012	0.0493	0.217	0	1
Period of construction: +2012	0.110	0.313	0	1

Notes: The number of observations in Panel A is 26 million, corresponding to around 4 million dwellings per year, while in panel B the number of dwellings is 995,199. Transaction prices are in 2019 euros deflated using the INSEE inflation index. Number of rooms includes kitchens and bedrooms. Living space represents the floor surface measured between walls within an apartment, excluding dependencies. Presence of kitchens larger than 9 m^2, garages, terraces, or annexes are dummy variables equal to one if the feature is present.

Très luxueux ("highly lavish") to (8) *Particulièrement défectueux* ("particularly defective").[37] This category determines the baseline quality multiplier applied to the surface area and thus plays a central role in the valuation.

The calculation produces a 1970-reference rental value (*VLC70*). To obtain the gross rental value reported on current tax rolls, VLC70 is multiplied by a departmental "updating coefficient" (revised annually) and a national revaluation coefficient for the current year. These updates maintain temporal and regional comparability while the underlying structure of the valuation—surface, quality adjustments, and location factors—remains anchored to the 1970 base period.

Because it jointly reflects location, building type, construction quality, equipment, and maintenance condition, the cadastral rental value serves as a comprehensive indicator of *intrinsic dwelling quality* rather than a measure of contemporaneous market rent. In our empirical framework, this variable captures quality dimensions that are otherwise difficult to observe in transaction data, making it a valuable control in both hedonic regressions and structural demand estimations.

[37] Intermediate categories (e.g., 5M) are allowed, representing 20% adjustments relative to the base category. Category boundaries are defined nationally (Article 324-H-I of Annex III), though municipalities may apply their own interpretation when assigning dwellings.