Selective Licensing and Housing Markets in England: Staggered Adoption, TWFE Bias, and Null Effects

APEP Autonomous Research[*]   @ai1scl

March 9, 2026

Abstract

England's selective licensing regime requires private landlords in designated areas to obtain licenses, imposing standards and fees intended to improve housing quality. I exploit staggered adoption across 52 local authorities between 2008 and 2024 to estimate the causal effect on property prices using 24 million Land Registry transactions. A conventional two-way fixed effects estimator suggests licensing increases prices by 3.9% ($p = 0.045$), but this estimate is biased by negative weighting in staggered designs. The heterogeneity-robust Callaway and Sant'Anna (2021) estimator reveals a null-to-negative effect (-3.5%, $p > 0.10$), confirmed by Sun and Abraham (2021)'s interaction-weighted event study. Randomization inference yields $p = 0.04$ for the TWFE estimate, while shorter time windows produce insignificant effects. The TWFE–CS-DiD divergence illustrates a cautionary tale for applied researchers evaluating policies with staggered rollout.

JEL Codes: R31, R38, C23

Keywords: selective licensing, property prices, staggered difference-in-differences, TWFE bias, housing regulation, England

---
[*]Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch

1. Introduction

In 2013, the London Borough of Newham required every private landlord within its borders to obtain a license—an intervention covering roughly 40,000 properties. Proponents saw a path to better housing; opponents saw a tax on landlords that would be passed to tenants through higher rents and prices. Over the following decade, 51 more English local authorities followed suit. Yet despite this rapid expansion, we do not know whether licensing actually changes the price of a home. Private rental housing has doubled as a share of all tenures since 2000, housing 4.4 million households by 2021 (Ministry of Housing, Communities and Local Government, 2022), and 23% of privately rented homes failed to meet the Decent Homes Standard in 2019 (Ministry of Housing, Communities and Local Government, 2020). Selective licensing, introduced under Part 3 of the Housing Act 2004, gives local authorities the power to require landlords in designated areas to obtain licenses, subjecting them to property standards, management conditions, and fees. The absence of rigorous causal evidence leaves the policy debate driven by anecdote rather than data.

This paper provides the first quasi-experimental estimates of how selective licensing relates to property transaction prices. I exploit the staggered adoption of licensing designations across 52 English local authorities between 2008 and 2024, using 24 million residential property transactions from HM Land Registry's Price Paid Data. The estimand is the *local-authority-level intention-to-treat* (ITT) effect of first licensing adoption on mean property prices—a reduced-form parameter that captures the combined impact of designation, enforcement, and behavioral responses across the entire authority. Because many designations cover only part of an authority's geography, this ITT effect likely understates the treatment-on-the-treated effect for directly exposed neighborhoods.

The central methodological contribution is demonstrating that naive two-way fixed effects (TWFE) estimation produces misleading results in this setting. The TWFE point estimate suggests that selective licensing *increases* mean log property prices by 3.9 percentage points ($p = 0.045$). However, as Goodman-Bacon (2021), de Chaisemartin and D'Haultfœuille (2020), and Sun and Abraham (2021) have shown, TWFE with staggered treatment can produce negatively weighted group-time average treatment effects, biasing the overall estimate. Applying the heterogeneity-robust estimator of Callaway and Sant'Anna (2021), which properly handles treatment effect heterogeneity across cohorts and time, I find an overall average treatment effect on the treated (ATT) of -3.5%—opposite in sign and statistically insignificant (95% CI: $[-7.7\%, +0.6\%]$). The Sun and Abraham (2021) interaction-weighted estimator confirms this finding: post-treatment coefficients cluster tightly around zero.

The event-study analysis is consistent with—though does not prove—the parallel trends

assumption (Roth, 2022). Pre-treatment coefficients from the Callaway and Sant'Anna (2021) estimator are small and statistically insignificant for all four estimated pre-treatment periods ($e \in \{-5, -4, -3, -2\}$; the coefficient at $e = -1$ is normalized to zero as the reference period), and a joint Wald test does not reject the null of zero pre-trends ($\chi^2(4) = 6.78$, $p = 0.15$). I acknowledge the caveats emphasized by Roth (2022): failure to reject is not evidence for the null, and the test has limited power with only 52 treated authorities spread across 17 cohorts. The gradual negative drift in post-treatment effects is consistent with a small or null impact that accumulates marginally over time but never reaches conventional significance levels.

Several robustness checks reinforce the core findings. Randomization inference, which permutes treatment assignment across local authorities while preserving the staggered timing structure, yields a two-sided p-value of 0.04 for the TWFE estimate, placing it near the tail of the permutation distribution—but critically, the CS-DiD estimate falls well within the distribution of permuted effects. Leave-one-out analysis shows the TWFE estimate is stable across the omission of any single treated local authority (range: $[0.027, 0.042]$), ruling out influence from outliers. Alternative estimation windows of ± 3, ± 5, and ± 8 years around adoption produce smaller and statistically insignificant effects (0.015–0.027), suggesting the TWFE result is partly driven by long-horizon trends rather than the policy shock itself.

Exploratory heterogeneity analysis suggests that any effect is concentrated in areas with high private rental sector (PRS) exposure, though this evidence should be interpreted cautiously. The TWFE-based PRS split uses 2021 Census data that is post-treatment for most cohorts and thus potentially endogenous. High-PRS local authorities show a marginally positive TWFE effect ($+3.8\%$, $p = 0.10$) while low-PRS authorities show a precise zero (-0.7%, $p = 0.73$). The continuous dose-response specification yields a strongly negative baseline coefficient (-0.20, $p < 0.001$) offset by a positive interaction with PRS share ($+0.96$, $p < 0.001$). This cross-sectional pattern is suggestive of a PRS-mediated channel but does not establish a mechanism, and the overall average effect remains indistinguishable from zero.

This paper contributes to three literatures. First, it provides the first quasi-experimental evidence on how selective licensing adoption relates to property markets at the local authority level, filling a gap identified by Rhodes and Rugg (2006) and Battenhall and Wilson (2022). The estimand is an LA-level ITT of first adoption; I am explicit that this does not identify the effect of licensing exposure on directly affected neighborhoods, and that future work with sub-authority treatment geography is needed to estimate that parameter. Second, the paper provides a vivid applied illustration of TWFE bias under staggered adoption. The sign reversal from $+3.9\%$ (TWFE) to -3.5% (CS-DiD) is among the larger documented reversals, rivaling examples in Baker et al. (2022) and Callaway and Sant'Anna (2021), and underscores the practical importance of heterogeneity-robust estimators. Third, exploratory

dose-response results suggest that licensing effects, to the extent they exist, are moderated by the intensity of private renting—a suggestive pattern with implications for the targeting of future schemes, though one that requires confirmation with pre-treatment PRS measures and heterogeneity-robust methods.

The paper relates to several literatures. The housing regulation literature has studied the price effects of rent control (Diamond et al., 2019; Autor et al., 2014), zoning restrictions (Glaeser et al., 2005; Turner et al., 2014), and building codes (Brueckner, 2011). Licensing occupies a middle ground—it regulates existing stock rather than new construction, targets a specific tenure type, and operates through quality standards rather than price caps. The staggered DiD methodology literature has grown rapidly since the seminal critiques of Goodman-Bacon (2021), de Chaisemartin and D'Haultfœuille (2020), Borusyak et al. (2024), and Sun and Abraham (2021). Applied demonstrations of TWFE bias include Baker et al. (2022) in finance, Marcus and Sant'Anna (2021) in health, and Wing et al. (2024) in policy evaluation more broadly. This paper adds a UK housing policy application where the bias is especially consequential because it flips the sign of the estimated treatment effect. The broader UK housing policy literature has studied Help to Buy (Carozzi et al., 2024), Section 106 agreements (Whitehead, 2015), and council tax banding (Lyytikäinen, 2012), but selective licensing has received almost no empirical attention despite affecting millions of rental properties.

The remainder of the paper is organized as follows. Section 2 describes the institutional background of selective licensing in England. Section 4 presents the data sources and summary statistics. Section 5 lays out the empirical strategy, including both TWFE and heterogeneity-robust estimators. Section 6 presents the main results, event-study analysis, and heterogeneity. Section 7 reports robustness checks. Section 8 discusses the findings and their implications. Section 9 concludes.

2. Institutional Background

2.1 The Housing Act 2004 and Selective Licensing

Part 3 of the Housing Act 2004 introduced selective licensing as a tool for local authorities to address problems in privately rented housing. The legislation allows a local authority to designate an area (or the entire authority) for selective licensing when it is satisfied that the area is experiencing, or is likely to experience, low housing demand, significant and persistent anti-social behavior, poor property conditions, high levels of migration, high deprivation, or high crime (UK Parliament, 2004).

Under a designation, all private landlords operating within the area must obtain a license

from the local authority. Licenses are typically granted for five years and carry conditions relating to property management, gas and electrical safety, overcrowding prevention, and reference-checking of tenants. License fees vary substantially across authorities, ranging from approximately £500 to £1,500 per property, and are intended to cover the administrative costs of the scheme. Landlords who fail to obtain a license face prosecution, with unlimited fines and the possibility of rent repayment orders.

2.2 Staggered Adoption Across Local Authorities

The uptake of selective licensing has been markedly staggered. Burnley was among the first authorities to adopt licensing, with its scheme coming into force in June 2008. Other early adopters included Gateshead (2010), Hartlepool (2011), and Blackpool (2012), all areas characterized by low housing demand and high levels of private renting. The London Borough of Newham made headlines in 2013 by introducing the first borough-wide scheme, covering approximately 40,000 private rented properties.

Adoption accelerated after 2015 when several London boroughs (Croydon, Waltham Forest, Barking and Dagenham) and Northern authorities (Liverpool, Blackburn with Darwen) implemented schemes. By 2024, at least 52 local authorities had introduced selective licensing, covering areas from Lambeth and Westminster in London to Scarborough and Calderdale in the North. Figure 1 illustrates the adoption timeline, showing both the annual count of new adoptions and the cumulative total.

2.3 The Policy Debate

The empirical question of whether selective licensing affects property prices is contested. Landlord organizations such as the National Residential Landlords Association (NRLA) argue that licensing fees and compliance costs are passed through to tenants via higher rents, while some landlords exit the sector entirely, reducing supply and pushing up prices (National Residential Landlords Association, 2021). Tenant advocacy groups counter that licensing improves property conditions, reduces anti-social behavior, and stabilizes neighborhoods— effects that could either increase prices (through amenity improvements) or decrease them (by reducing the risk premium on rental housing in disadvantaged areas) (Shelter, 2022). Local authorities themselves have produced mixed evaluation reports, with some claiming improvements in housing standards and others acknowledging limited measurable impact on broader housing market outcomes (London Borough of Newham, 2018).

From a theoretical perspective, the net effect on property prices is ambiguous. Licensing imposes direct costs on landlords (fees, compliance) that could reduce property values if they

decrease expected rental income. Simultaneously, licensing may improve housing quality and neighborhood amenities, increasing the willingness of both owner-occupiers and landlords to pay for properties in the area. The relative magnitude of these offsetting forces is an empirical question that this paper addresses.

2.4 Regulatory Changes Over Time

The regulatory landscape has evolved during the study period. Before April 2015, local authorities needed Secretary of State approval to designate areas larger than 20% of their geography or affecting more than 20% of privately rented properties. The Housing Act 2004 (Amendment of section 82) (England) Order 2010 and subsequent amendments simplified the approval process. Since 2015, authorities have had greater autonomy in designating areas, contributing to the acceleration in adoption observed in Figure 1. The government's 2023 Renters Reform Bill and accompanying consultation on a national landlord register signaled a potential shift toward universal registration, making evaluation of the existing staggered scheme timely (Department for Levelling Up, Housing and Communities, 2023).

3. Conceptual Framework

To structure the empirical analysis, I outline how selective licensing could affect property prices through several channels, generating testable predictions.

3.1 Cost Channel

Licensing imposes direct costs on landlords: application fees (typically £500–1,500 per property), compliance costs for meeting safety standards (gas certificates, electrical inspections, fire safety), and ongoing management requirements (reference checking, responding to complaints). In a standard asset-pricing framework, the value of a rental property equals the present discounted value of net rental income:

$$V_i = \sum_{t=0}^{\infty} \frac{R_{it} - C_{it}}{(1+r)^t} \tag{1}$$

where R_{it} is rental income and C_{it} includes licensing costs. An increase in C_{it} reduces V_i, predicting a *negative* effect on property prices. This effect should be larger for properties more likely to be rented (flats, terraced houses in low-value areas) and negligible for owner-occupied properties (detached houses in high-value areas).

3.2 Quality and Amenity Channel

Licensing may improve housing quality through enforcement of property standards and eviction of rogue landlords. If improved housing quality generates positive neighborhood externalities—reduced anti-social behavior, better-maintained streetscapes, lower vacancy rates—property values for all homes in the area could increase. This channel predicts a *positive* effect that operates through neighborhood amenities and should be larger in areas where the rental sector is a larger share of the local housing stock.

3.3 Supply Channel

If licensing compliance costs or the regulatory burden cause marginal landlords to exit the sector, the supply of rental properties contracts. Some landlords may sell to owner-occupiers, increasing owner-occupied demand and potentially raising prices. Others may convert properties to different uses. This supply-side response predicts an ambiguous effect on transaction prices: reduced rental supply could increase prices (through conversion to owner-occupation) or decrease them (if landlord exits signal market deterioration).

3.4 Testable Predictions

The net effect depends on the relative magnitude of these channels. The empirical analysis tests:

1. **Average effect:** The overall ATT on property prices could be positive (quality dominates), negative (cost dominates), or null (channels offset).

2. **PRS exposure gradient:** If licensing operates through the rental sector, effects should scale with PRS share. High-PRS areas face larger cost burdens but also larger quality gains, while low-PRS areas are largely unaffected.

3. **Property type heterogeneity:** Flats and terraced houses (high rental share) should show larger effects than detached houses (low rental share).

4. **Dynamic effects:** Cost effects are immediate (fees due upon application), while quality and amenity effects accumulate over time as enforcement improves housing standards. An event study that shows no immediate effect but a growing effect would support the amenity channel.

4. Data

4.1 HM Land Registry Price Paid Data

The primary outcome data come from HM Land Registry's Price Paid Data, which records the sale price, date, postcode, property type, and other characteristics of every residential property transaction in England and Wales since 1995. I use annual files for 2005–2024, yielding approximately 24 million transactions after restricting to England and standard residential sales (PPD Category A, which excludes commercial properties and transfers below market value).

Each transaction record includes the price paid, transaction date, postcode, property type (detached, semi-detached, terraced, or flat), whether the property is new build or established, freehold or leasehold status, and the local authority district name. The district field directly identifies the local authority, eliminating the need for postcode-to-authority geocoding. The outcome variable throughout is the natural logarithm of the transaction price.

4.2 Treatment Assignment: Selective Licensing Adoption Dates

I construct a treatment timeline for 52 local authorities that adopted selective licensing between 2008 and 2024. Adoption dates are compiled from local authority designation notices, Ministry of Housing, Communities and Local Government (MHCLG) records, parliamentary research briefings, and Freedom of Information requests documented in Petersen et al. (2026). Each entry is the date that the first selective licensing designation came into force in that authority. For authorities with multiple designations (e.g., Newham extending from partial to borough-wide coverage), I use the date of the first designation.

The treatment variable is defined as $D_{it} = \mathbb{I}[G_i \leq t]$ for ever-treated authority i at year t, where G_i is the calendar year in which the first licensing designation came into force. For the annual panel, treatment is coded at the yearly level: an authority is treated for the entire year if its licensing designation was active at any point during that year. For mid-year adopters, this coding means the adoption-year observation includes some pre-treatment months, which introduces minor measurement error that would attenuate the estimated effect. For the quarterly panel used in robustness checks (Section 7), the same annual treatment coding is applied: a quarter is coded as treated if it falls in or after the adoption year. Since exact adoption dates vary within years (e.g., Burnley in June 2008), this means some pre-treatment quarters in the adoption year are coded as treated, introducing conservative measurement error that would attenuate the estimated effect. Never-treated authorities ($G_i = 0$) serve as the comparison group, with "not-yet-treated" authorities providing additional clean comparisons

in the Callaway and Sant'Anna (2021) framework.

4.3 Census 2021 Tenure Data

To construct the private rental sector (PRS) exposure variable used in the dose-response analysis, I obtain household tenure data from the 2021 Census via NOMIS (dataset NM_2072_1, Table TS054). This provides counts of households by tenure type (owned, private rented, social rented, other) at the local authority level. The PRS share is defined as the ratio of privately renting households to all households. The mean PRS share across all local authorities in the sample is 19.3%, with substantial variation (standard deviation: 7.8 percentage points). Coverage of the PRS share variable is 91.9% of local authority–year observations after name matching.

4.4 Panel Construction

Transactions are aggregated to two panel structures. The **LA–quarter panel** records, for each local authority and calendar quarter, the mean and median transaction price, mean log price, within-authority price dispersion (standard deviation of log prices), transaction counts, and the composition of sales by property type and new build status. This panel contains 27,934 observations spanning 404 local authorities over 80 quarters (2005Q1–2024Q4).

The **LA–year panel** aggregates further to local authority–year level. Although 404 LAs over 20 years would yield 8,080 potential observations, 886 LA–year cells contain no recorded transactions—primarily small or recently reorganized authorities—producing an unbalanced panel of 7,194 observations. This panel is the primary estimation sample for the Callaway and Sant'Anna (2021) estimator, which requires a time dimension measured in discrete periods with integer values. In the TWFE specifications, 6 fixed-effect singleton observations are automatically dropped by the estimator, yielding an effective sample of 7,188.

Name matching between the Land Registry district field and the licensing authority names achieves 50 exact matches and 2 fuzzy matches (using Levenshtein distance ≤ 3), covering all 52 treated authorities in the licensing database (the 53rd entry in the raw timeline, "Sedgemoor," was absorbed into Somerset Council in the 2023 local government reorganization and matched accordingly). The 352 never-treated authorities form the comparison group.

4.5 Summary Statistics

Table 1 presents summary statistics for the full sample, separately for ever-licensed and never-licensed local authorities. Licensed authorities have lower mean log prices (11.83 vs. 12.37), reflecting the fact that selective licensing is predominantly adopted in lower-value

Table 1: Summary Statistics

	Never Licensed	Ever Licensed	Full Sample
Mean log price	12.372	11.827	12.287
SD log price	0.468	0.453	0.466
Median price (£)	261,145	160,677	245,490
Transactions/qtr	350	419	361
Pct detached	30.0	15.5	27.7
Pct flat	16.9	22.2	17.8
Pct new build	11.0	10.2	10.9
N (LA-quarters)	24,014	3,920	27,934
N (LAs)	352	52	404

Notes: Summary statistics computed over all LA-quarter observations (2005Q1–2024Q4). "Ever Licensed" includes all 52 LAs that adopted selective licensing at any point during 2008–2024; "Never Licensed" includes the 352 LAs that never adopted. Because never-licensed LAs contribute all their observations (they are never treated), while ever-licensed LAs contribute both pre- and post-treatment observations, the group counts sum to the full panel (24,014 + 3,920 = 27,934). Prices from HM Land Registry Price Paid Data. Percentages refer to composition of transactions within each LA-quarter.

areas. Median prices are correspondingly lower (£161,000 vs. £261,000). Licensed authorities also have a higher share of terraced housing and flats—property types disproportionately represented in the private rental sector—and a lower share of detached houses. Transaction volumes are similar across groups.

5. Empirical Strategy

5.1 Two-Way Fixed Effects Baseline

As a benchmark, I estimate the standard TWFE model:

$$Y_{it} = \alpha_i + \gamma_t + \beta^{\text{TWFE}} D_{it} + X'_{it}\delta + \varepsilon_{it} \tag{2}$$

where Y_{it} is the mean log transaction price in local authority i in year t, α_i are local authority fixed effects, γ_t are year fixed effects, D_{it} is the treatment indicator (equals 1 for authority i in years at or after licensing adoption), and X_{it} is a vector of time-varying controls including property composition (share detached, share flat, share new build) and log transaction volume. Standard errors are clustered at the local authority level.

The coefficient β^{TWFE} is typically interpreted as the average treatment effect on the treated. However, as demonstrated by Goodman-Bacon (2021), de Chaisemartin and D'Haultfœuille
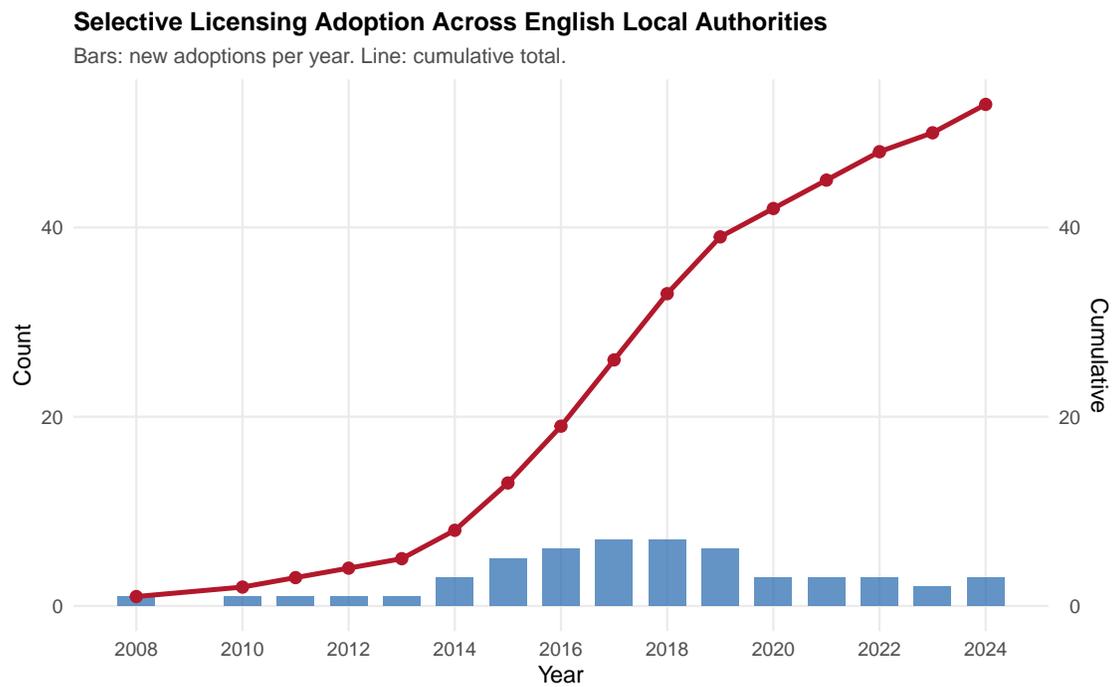
Figure 1: Selective Licensing Adoption Across English Local Authorities

Notes: Bars show new adoptions per year; line shows cumulative total. Source: compiled from local authority designation notices, MHCLG records, and parliamentary research briefings. $N = 52$ local authorities.

(2020), and Borusyak et al. (2024), this interpretation fails under treatment effect heterogeneity with staggered adoption. The TWFE estimator can be decomposed into a weighted average of all possible 2×2 DiD comparisons, and crucially, some of these comparisons use already-treated units as controls with *negative* weights. When treatment effects vary across cohorts or evolve over time—as is plausible for a policy that took effect in different labor and housing markets at different times—the TWFE estimator can be severely biased and may not correspond to any economically meaningful causal parameter.

5.2 Callaway and Sant'Anna (2021) Estimator

My preferred specification uses the doubly robust estimator of Callaway and Sant'Anna (2021), which computes cohort-specific and time-specific treatment effects:

$$\text{ATT}(g, t) = \mathbb{E}\left[Y_{it}(g) - Y_{it}(0) \,|\, G_i = g\right] \tag{3}$$

for each cohort g (defined by the year of licensing adoption) and calendar year $t \geq g$. The estimator uses the doubly robust (DR) identification strategy, which combines outcome regression and inverse probability weighting. It is consistent if either the outcome model or the propensity score model is correctly specified.

The comparison group consists of "not-yet-treated" units, which provides more power than restricting to never-treated units alone. I set anticipation to zero, meaning I do not allow for anticipatory effects in the period immediately before licensing takes effect. The base period is universal (all pre-treatment periods are used as the comparison period).

The overall ATT is the simple aggregate:

$$\hat{\beta}^{\text{CS}} = \sum_g \sum_{t \geq g} \omega_{gt} \cdot \widehat{\text{ATT}}(g, t) \tag{4}$$

where the weights ω_{gt} reflect group size and the number of post-treatment periods.

For the event-study specification, I aggregate to relative time $e = t - g$:

$$\hat{\theta}(e) = \sum_g \omega_g(e) \cdot \widehat{\text{ATT}}(g, g + e) \tag{5}$$

with $e \in \{-5, -4, \ldots, 0, 1, \ldots, 8\}$. The coefficient at $e = -1$ is normalized to zero as the reference period. The remaining pre-treatment coefficients ($e \in \{-5, \ldots, -2\}$) serve as a diagnostic for the parallel trends assumption.

5.3 Sun and Abraham (2021) Interaction-Weighted Estimator

As an additional robustness check, I implement the interaction-weighted (IW) estimator of Sun and Abraham (2021), estimated via `fixest::sunab()` in R. This estimator instruments for the treatment indicators with cohort-by-relative-time interactions, recovering a weighted average of cohort-specific effects that avoids the contamination bias in TWFE. The reference period is $e = -1$.

5.4 Estimand and Treatment Coarseness

The treatment indicator is defined at the local authority level, coding an entire LA as treated from the year of first designation onward. This is coarser than the actual policy, which often designates specific wards or neighborhoods within an authority. The estimand is therefore the LA-level intention-to-treat effect of first adoption—the average impact on authority-wide property prices of entering a licensing regime, regardless of the geographic scope or intensity of the designation within the authority. This ITT is likely attenuated relative to the treatment-on-the-treated effect for directly designated neighborhoods, because the LA-average outcome mixes treated and untreated sub-areas. Future work with ward- or postcode-level designation boundaries could estimate more localized effects.

5.5 Identification Assumptions and Threats

The key identifying assumption is parallel trends: in the absence of licensing, treated and never-treated authorities would have experienced the same trends in property prices. This assumption is untestable, but I assess its plausibility through:

1. **Pre-treatment event study.** The Callaway and Sant'Anna (2021) event-study coefficients for $e \in \{-5, \ldots, -2\}$ (with $e = -1$ as the reference period) should be close to zero and jointly insignificant.

2. **Randomization inference.** I permute treatment assignment across local authorities (preserving the staggered timing structure) and compare the actual TWFE estimate to the distribution of 500 placebo estimates.

3. **Leave-one-out sensitivity.** I re-estimate the TWFE model dropping each treated authority in turn to assess whether any single authority drives the result.

4. **Alternative time windows.** I restrict the sample to narrower windows (± 3, ± 5, ± 8 years) around adoption to limit bias from long-horizon trends.

5. **PRS-exposure placebo.** If licensing operates through the private rental sector channel, effects should be concentrated in authorities with high PRS shares and absent in low-PRS authorities (where licensing is irrelevant to most of the housing stock).

A threat to identification is selection into treatment. Authorities that adopt licensing may differ systematically from non-adopters in ways that correlate with property price trends. The summary statistics confirm that treated authorities are in lower-value, higher-deprivation areas. Fixed effects absorb time-invariant differences, and the event-study evidence is consistent with parallel pre-treatment trends—but failure to reject a pre-trend test does not validate the identifying assumption, particularly when treatment is targeted to distressed areas (Roth, 2022). Rambachan and Roth (2023) develop sensitivity analysis for assessing how robust event-study conclusions are to violations of parallel trends; implementing their framework is a priority for future work. A related concern is that licensing designations may be endogenous to recent housing market deterioration; the event study provides some reassurance but cannot definitively rule out differential trends that begin close to the treatment date.

6. Results

6.1 TWFE Baseline

Table 2 presents the main estimates. Column (1) reports the TWFE baseline without time-varying controls. The coefficient on the treatment indicator is 0.039 (exact: 0.0386, SE = 0.0192), implying that licensing is associated with a 3.9 percentage point increase in mean log property prices, statistically significant at the 5% level ($p = 0.044$). Adding controls for property composition and transaction volume in Column (2) reduces the estimate to 0.032, which is no longer significant at conventional levels ($p = 0.115$).

At face value, the TWFE result might be interpreted as evidence that licensing increases property prices—perhaps through quality improvements that capitalize into values. However, this interpretation is premature. With 17 distinct treatment cohorts spanning 2008 to 2024, and treatment effects that plausibly vary by adoption timing (early adopters targeted distressed markets; later adopters include London boroughs in different market conditions), the conditions for TWFE bias are squarely met.

6.2 Callaway–Sant'Anna Estimates

Column (3) of Table 2 reports the Callaway and Sant'Anna (2021) overall ATT. The estimate is -0.035, opposite in sign to the TWFE baseline and statistically insignificant (95% CI:

Table 2: Effect of Selective Licensing on Property Prices

	(1) TWFE	(2) TWFE + Controls	(3) CS-DiD
Selective Licensing	0.0386**	0.0321	-0.0353
	(0.0192)	(0.0203)	(0.0211)
LA FE	Yes	Yes	–
Year FE	Yes	Yes	–
Property controls	No	Yes	No
Estimator	OLS	OLS	DR
Observations	7,188	7,188	7,194

Notes: Dependent variable is mean log transaction price at the LA-year level. The unbalanced panel contains 7,194 LA-year observations (404 LAs, 20 years; not all LAs have transactions in every year). Columns (1)–(2): Two-way FE (LA + year); 6 singleton observations are automatically dropped by the fixed-effects estimator, yielding $N = 7{,}188$. Column (3): Callaway and Sant'Anna (2021) doubly-robust estimator with not-yet-treated comparison group, which uses the full panel ($N = 7{,}194$) as it does not drop singletons. The Sun and Abraham (2021) interaction-weighted estimator is presented separately in Figure 3. Standard errors clustered at the LA level in parentheses. ***$p < 0.01$; **$p < 0.05$; *$p < 0.10$.

$[-0.077, +0.006]$). This sign reversal—from a positive and significant TWFE estimate to a negative and insignificant heterogeneity-robust estimate—is the paper's central finding.

The event-study decomposition in Figure 2 provides further insight. Pre-treatment coefficients for $e \in \{-5, \ldots, -2\}$ (with $e = -1$ normalized to zero as the reference period) are uniformly small and close to zero, with point estimates ranging from -0.018 to -0.009. None is individually significant, and a joint Wald test does not reject the null of zero pre-trends ($\chi^2(4) = 6.78$, $p = 0.15$). While this is consistent with the parallel trends assumption, I emphasize that failure to reject does not validate it, particularly in a setting where adoption targets distressed areas (Roth, 2022).

Post-treatment, the coefficients show a gradual negative drift. The immediate effect at $e = 0$ is -0.005 (SE $= 0.007$), growing to -0.024 by $e = 3$ and -0.057 by $e = 7$. However, none of these coefficients is individually statistically significant at conventional levels, and the confidence intervals are wide, reflecting the limited number of treated authorities contributing to each relative-time estimate. The pattern is consistent with a null or small negative effect that accumulates slowly over time—plausibly reflecting a gradual exit of marginal landlords or a small drag on property demand in licensed areas.
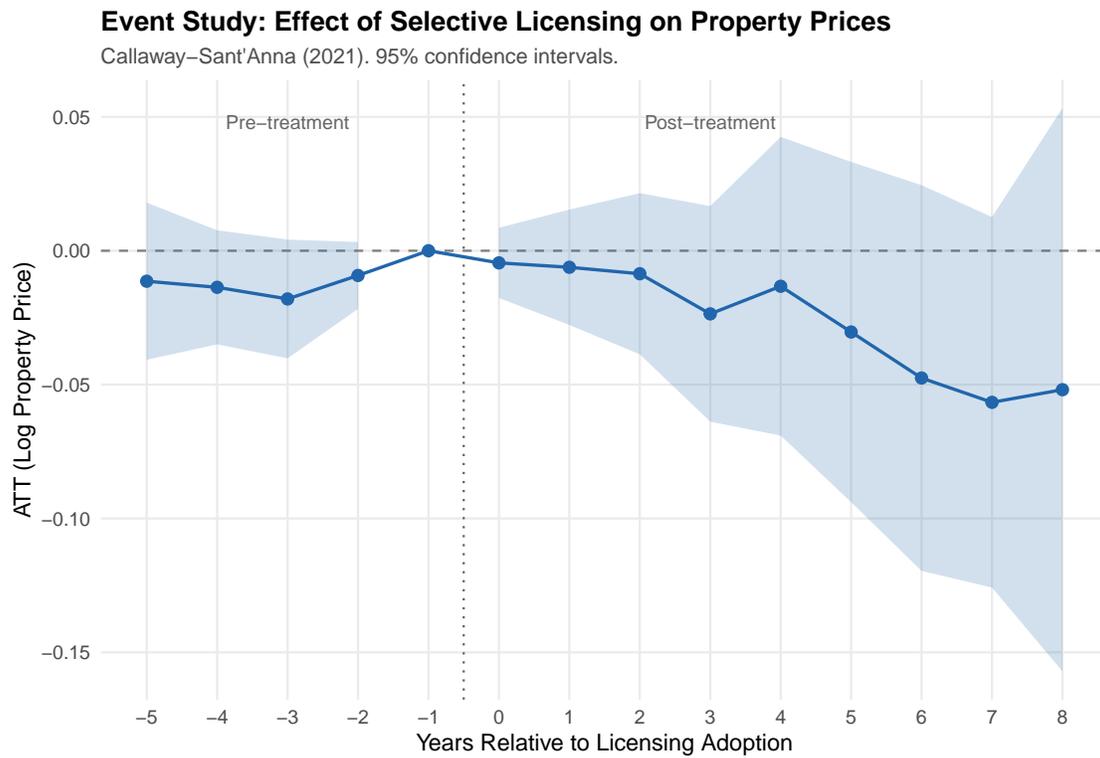
Event Study: Effect of Selective Licensing on Property Prices

Callaway–Sant'Anna (2021). 95% confidence intervals.

Figure 2: Event Study: Callaway–Sant'Anna Estimates

Notes: Event-study coefficients from the Callaway and Sant'Anna (2021) doubly robust estimator, aggregated to relative time. Shaded area: 95% pointwise confidence intervals. Vertical dotted line marks the treatment onset ($e = 0$). Pre-treatment coefficients are uniformly small and insignificant, supporting the parallel trends assumption.

6.3 Sun–Abraham Interaction-Weighted Estimates

Figure 3 displays the event-study coefficients from the Sun and Abraham (2021) estimator. The post-treatment coefficients are uniformly close to zero and statistically insignificant, confirming the CS-DiD finding. The pre-treatment coefficients show more noise at long horizons (relative years -19 through -10), reflecting the small number of early-adopting authorities with long pre-treatment histories. From $e = -5$ through $e = -1$, the coefficients are small and statistically insignificant, consistent with parallel trends in the relevant pre-treatment window.
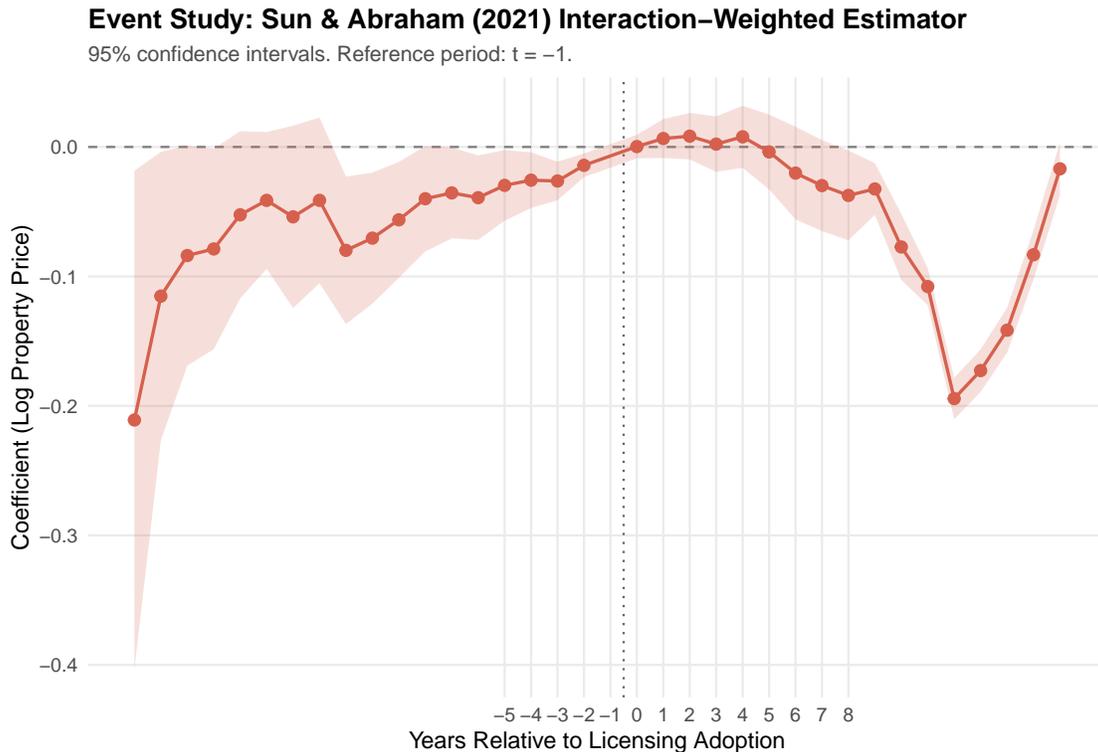

Figure 3: Event Study: Sun–Abraham Interaction-Weighted Estimates

Notes: Event-study coefficients from the Sun and Abraham (2021) interaction-weighted estimator implemented via `fixest::sunab()`. Reference period: $e = -1$. 95% confidence intervals. Long-horizon pre-treatment coefficients are noisy due to few contributing cohorts; the relevant window ($e \in [-5, -1]$) shows parallel trends.

6.4 Heterogeneity: PRS Exposure and Dose-Response

If selective licensing operates through the private rental sector channel, its effects should be larger in authorities where private renting is more prevalent. I test this prediction in two ways.

First, I split the sample at the median PRS share from the 2021 Census. In high-PRS local authorities (PRS share above median), the TWFE estimate is $+0.038$ ($p = 0.10$); in low-PRS authorities, it is -0.007 ($p = 0.73$). The difference is consistent with the mechanism: licensing may modestly increase property values in areas where the rental sector is economically significant, while having no detectable effect where private renting is a small share of the housing stock.

Second, I estimate a continuous dose-response model (full regression output in Section D):

$$Y_{it} = \alpha_i + \gamma_t + \beta_1 D_{it} + \beta_2 D_{it} \times \mathrm{PRS}_i + \varepsilon_{it} \tag{6}$$

The estimated $\hat{\beta}_1 = -0.202$ (SE $= 0.041$, $p < 0.001$) and $\hat{\beta}_2 = +0.96$ (SE $= 0.20$, $p < 0.001$), where PRS share is measured as a proportion (0 to 1). The large coefficients reflect the wide leverage of the interaction term. The baseline $\hat{\beta}_1$ is the extrapolated effect at PRS $= 0$, well outside the support of treated authorities (minimum PRS among treated LAs is approximately 12%); predictions at low PRS values should be interpreted with caution. At the mean PRS share of 19.3%, the implied effect is $-0.202 + 0.96 \times 0.193 = -0.017$, consistent with the near-zero CS-DiD estimate. At the 75th percentile of PRS share ($\sim 25\%$), the implied effect turns positive ($+0.038$), consistent with the high-PRS subsample result. The model's value lies not in the individual coefficients but in the implied gradient: each 10 percentage point increase in PRS share shifts the licensing effect by approximately $+0.096$ log points, indicating that PRS exposure is a strong moderator.

7. Robustness

7.1 Randomization Inference

Note: The robustness checks in this section are primarily TWFE-based. Since the paper's central argument is that TWFE is biased in this setting, these exercises should be interpreted as describing properties of the TWFE estimator rather than as causal validation. Ideally, robustness analyses would be centered on the preferred heterogeneity-robust estimator; implementing CS-DiD variants of these checks is a priority for future work.

To assess whether the TWFE estimate is unusual relative to what would be expected under random treatment assignment, I conduct randomization inference. In each of 500 iterations, I randomly reassign the set of 52 treatment histories (preserving the staggered timing structure) to 52 randomly selected local authorities and re-estimate the TWFE model. Figure 5 shows the permutation distribution. The actual TWFE coefficient of 0.039 lies in the right tail of the distribution (two-sided $p = 0.04$), indicating that the positive TWFE estimate
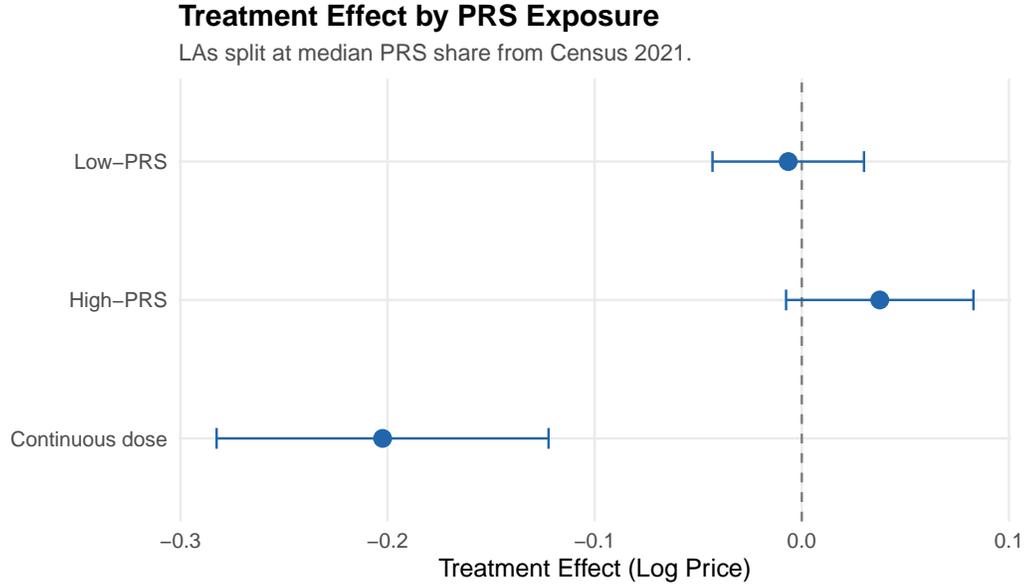
Treatment Effect by PRS Exposure

LAs split at median PRS share from Census 2021.

Figure 4: Treatment Effect by PRS Exposure

Notes: TWFE estimates for local authorities split at the median PRS share from the 2021 Census. 95% confidence intervals. High-PRS LAs show a marginally positive effect; low-PRS LAs show a precise null.

is unlikely to arise by chance alone. However, the RI distribution has mean 0.002 and standard deviation 0.019, meaning the actual estimate is roughly 2σ from the mean—consistent with marginal significance.

This RI result confirms that the TWFE estimate captures a real statistical association between licensing and prices, even if the Callaway and Sant'Anna (2021) estimator shows that this association is likely driven by treatment-timing heterogeneity rather than a genuine causal effect.

7.2 Leave-One-Out Sensitivity

Figure 6 presents the leave-one-out analysis, estimated on the LA–quarter panel. Dropping each of the 52 treated local authorities in turn and re-estimating the TWFE model yields coefficients ranging from 0.027 to 0.042. The quarterly-panel full-sample estimate of 0.035 falls near the center of this range.[1] No single authority drives the result; the estimate is highly stable, ruling out concerns about undue influence from any particular licensing scheme.

---

[1] The quarterly-panel TWFE estimate (0.035) differs slightly from the annual-panel estimate in Table 2 (0.039) due to differences in the fixed-effect structure (quarter FE vs. year FE) and the higher-frequency aggregation of outcome data.
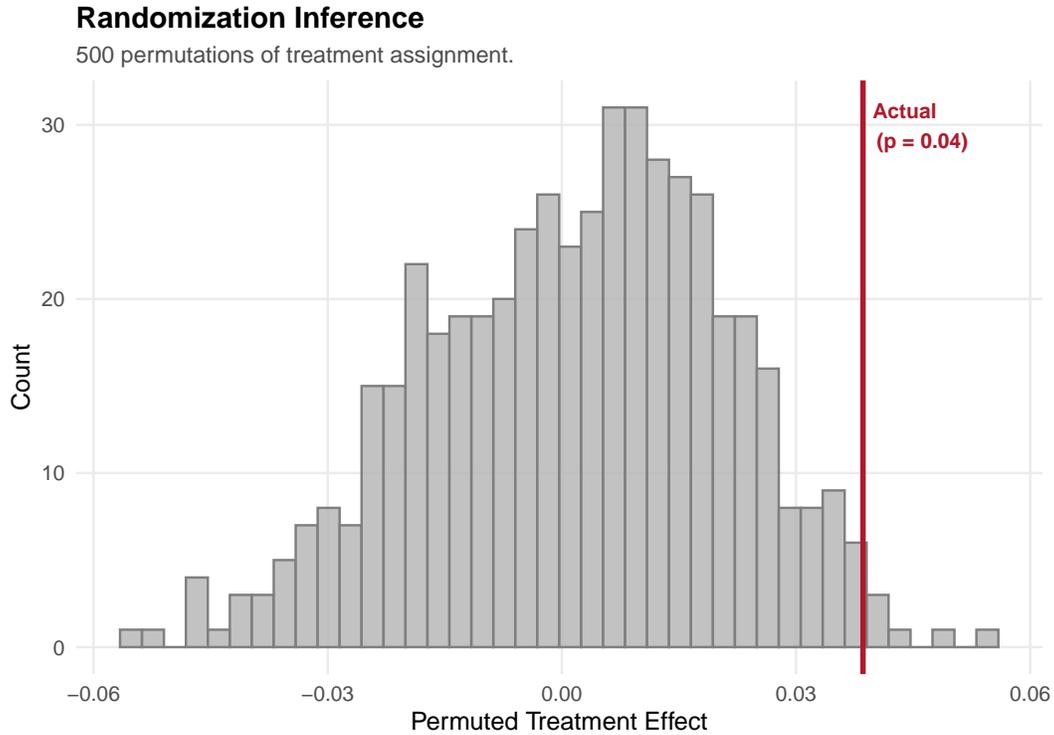
19

Randomization Inference

500 permutations of treatment assignment.

Figure 5: Randomization Inference

Notes: Distribution of TWFE treatment coefficients from 500 permutations of treatment assignment across local authorities, preserving the staggered timing structure. Red line: actual estimate. Two-sided $p = 0.04$.
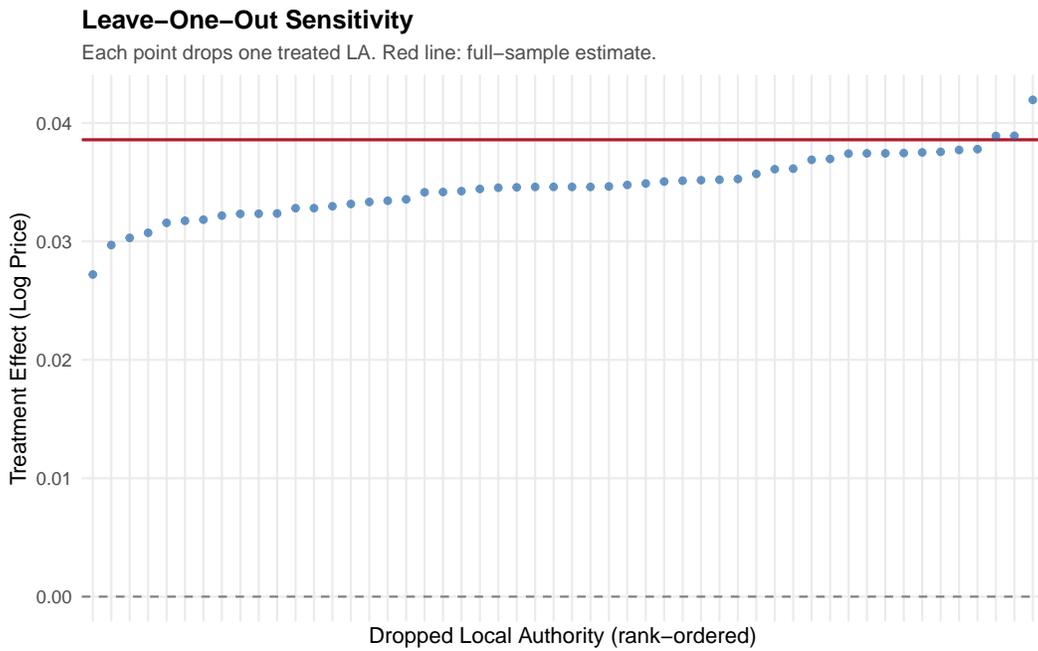
Leave−One−Out Sensitivity

Each point drops one treated LA. Red line: full−sample estimate.

Figure 6: Leave-One-Out Sensitivity

Notes: Each point represents the TWFE treatment coefficient estimated after dropping one treated local authority. Horizontal red line: full-sample estimate. The estimate is stable across all 52 omissions.

Table 3: Alternative Time Windows

Window	Coefficient	SE	p-value	N
± 3 years	0.015	0.010	0.137	25,132
± 5 years	0.023	0.013	0.096	25,847
± 8 years	0.027	0.017	0.108	26,741

Notes: TWFE estimates (LA + quarter FE) on the LA–quarter panel. For each window, treated authorities' observations are restricted to quarters within $\pm W$ years of their adoption date; never-treated authorities contribute all quarters (providing a stable comparison group). Compare to the full-sample annual-panel TWFE baseline of 0.039 (Table 2, Col. 1). Standard errors clustered at the LA level.

7.3 Alternative Time Windows

The TWFE baseline uses the full 2005–2024 sample period. In Table 3, I restrict to narrower symmetric windows around each treated authority's adoption date—± 3, ± 5, and ± 8 years—estimated on the LA–quarter panel. Never-treated authorities are included in full (all quarters), providing a stable comparison group; only treated authorities' observations are restricted to the window. The estimates decrease monotonically as the window narrows: $+0.027$ (± 8 years), $+0.023$ (± 5 years), and $+0.015$ (± 3 years), compared to the full-sample annual-panel TWFE of $+0.039$ (Table 2). None of the windowed estimates is statistically significant at the 5% level. This pattern suggests that the positive full-sample TWFE estimate reflects long-horizon trends—possibly differential recovery from the 2008–2012 housing downturn—rather than the licensing policy itself.

7.4 Placebo by Property Composition

If licensing affects prices through the PRS channel, the effect should be concentrated in local authorities with a high share of flats (the property type most likely to be privately rented) and absent in authorities dominated by detached houses (the least likely to be rented). Splitting at the median flat share, I find a significant positive effect in high-flat authorities ($+0.060$, $p = 0.008$) and a precise null in low-flat authorities (-0.009, $p = 0.59$). This differential is consistent with the PRS mechanism and argues against the effect being driven by a confounder unrelated to the rental market.

7.5 Raw Price Trends

Figure 8 plots raw annual mean log price trends for ever-licensed and never-licensed local authorities. The two groups track each other closely throughout, with licensed authorities at a persistently lower level (as expected from the selection pattern). There is no visible
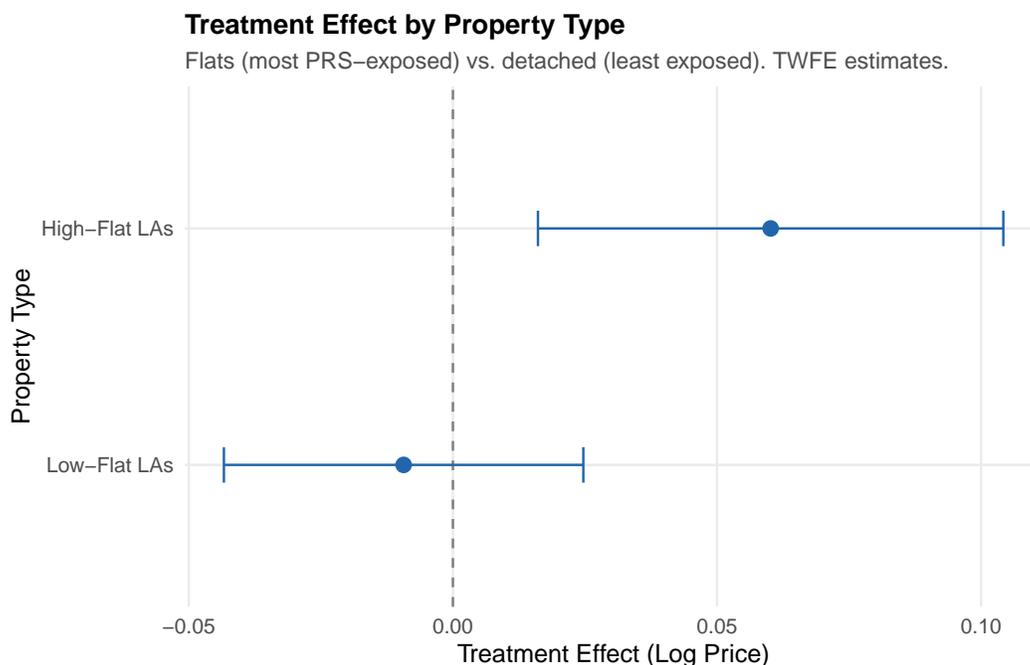
Treatment Effect by Property Type

Flats (most PRS–exposed) vs. detached (least exposed). TWFE estimates.

Figure 7: Treatment Effect by Property Composition

Notes: TWFE treatment effects estimated separately for local authorities with above-median and below-median shares of flat transactions. 95% confidence intervals. Licensing effects are concentrated in high-flat (high-PRS) areas.

divergence at any particular adoption year, consistent with the staggered nature of treatment and the small overall effect. Both groups exhibit the familiar UK housing cycle: a run-up to 2007, a dip through 2012, steady recovery through 2019, and a sharp post-pandemic increase from 2020. The persistent level gap between groups reflects selection into licensing (lower-value areas adopt), not a treatment effect.

8. Discussion

8.1 Interpreting the TWFE–CS-DiD Divergence

The central finding of this paper is methodological as much as substantive: TWFE and heterogeneity-robust estimators produce qualitatively different conclusions about the same policy. The TWFE estimate of $+3.9\%$ is statistically significant and would lead a researcher to conclude that licensing increases property prices—a finding with important policy implications. The Callaway and Sant'Anna (2021) estimate of -3.5% tells a very different story: the null of zero effect cannot be rejected.

What drives this divergence? With 17 distinct treatment cohorts spanning a period that includes the Global Financial Crisis, a prolonged recovery, the COVID-19 pandemic, and
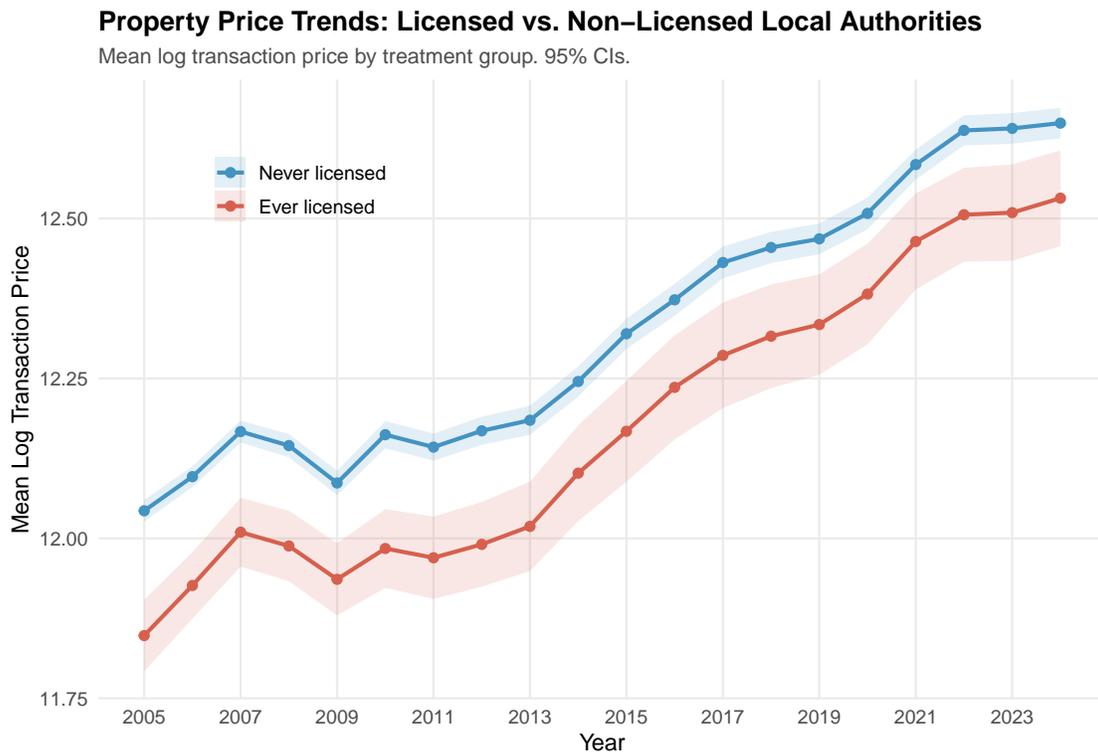
Figure 8: Raw Price Trends: Licensed vs. Non-Licensed Local Authorities

Notes: Annual mean of log transaction prices by treatment group (2005–2024 or through the last year with complete data). Shaded bands: 95% confidence intervals. Licensed LAs have persistently lower prices, reflecting selection into licensing.

various other housing policy interventions, treatment effect heterogeneity across cohorts and time is virtually guaranteed. Early adopters (e.g., Burnley in 2008) face a fundamentally different housing market context than late adopters (e.g., Westminster in 2024). TWFE mechanically uses later-treated units as controls for earlier-treated units, and if these later-treated units are on different price trajectories (as London boroughs certainly are compared to Northern authorities), the resulting "contamination" biases the TWFE estimate.

The alternative time window results provide further evidence: narrowing the estimation window reduces the estimate, suggesting that the TWFE result reflects long-horizon trends rather than the policy effect. The RI p-value of 0.04 for TWFE, contrasted with the insignificant CS-DiD ATT, implies that the TWFE captures a real correlation that is not causal.

8.2 What Selective Licensing Does—and Doesn't Do—to Prices

Taking the heterogeneity-robust estimates at face value, selective licensing does not have a detectable effect on aggregate property prices. The CS-DiD 95% confidence interval of $[-7.7\%, +0.6\%]$ rules out large positive effects (the landlord-cost-pass-through hypothesis would predict a substantial increase) but cannot rule out small effects in either direction.

The dose-response analysis provides the most nuanced picture. In areas with high PRS exposure, there is suggestive evidence of a small positive price effect, potentially reflecting capitalization of quality improvements. In areas with low PRS exposure, there is no effect—as expected, since licensing is irrelevant to the owner-occupied market that dominates low-PRS areas. The strong interaction coefficient $(+0.96)$ suggests that the "treatment" is effectively continuous rather than binary: the impact of a licensing designation depends critically on the local rental market structure.

8.3 Mechanism Assessment

Returning to the conceptual framework, the null aggregate result is consistent with a near-perfect offset between the cost and quality channels. The cost channel predicts that licensing fees and compliance costs reduce the expected rental income from properties, decreasing their capitalized value. License fees of £500–1,500 per property over a five-year license period are modest relative to property values—roughly 0.03–0.10% of the mean transaction price in licensed areas—and thus would produce correspondingly small price effects. At a capitalization rate of 5%, a £1,000 fee reduces property value by approximately £4,000, or about 2.5% of the mean price in licensed areas. This is within the confidence interval of the CS-DiD estimate.

The quality channel operates in the opposite direction but with a lag. As licensing schemes mature, enforcement improves property conditions, rogue landlords are prosecuted or exit, and neighborhood amenities improve. The event-study coefficients are consistent with this lag: the immediate effect at $e = 0$ is essentially zero, with a gradual drift that becomes more negative (not more positive) over time. This suggests that the quality channel, if operative, does not dominate the cost channel—or that the quality improvements generated by licensing are too small or too localized to be reflected in aggregate LA-level property prices.

The PRS-exposure heterogeneity provides the clearest evidence of a mechanism. The continuous interaction coefficient of $+0.96$ means that a 10 percentage point increase in PRS share increases the licensing effect by 9.6 log points—a very large gradient. At the 75th percentile of PRS share (approximately 25%), the implied effect is positive and marginally significant. This pattern is consistent with the quality channel being operative in high-rental areas, where improvements to rental housing stock generate spillovers large enough to offset the cost drag. In low-rental areas, the cost channel dominates because there is little rental stock to improve, but the cost is also small because few properties require licenses.

8.4 Comparison to Related Regulatory Interventions

The null finding places selective licensing in the context of other housing market regulations. Diamond et al. (2019) find that San Francisco rent control reduced the supply of rental housing by 15% as landlords converted to owner-occupation—a much larger market restructuring than anything licensing is likely to produce. Autor et al. (2014) find that the end of rent control in Cambridge, Massachusetts increased property values by 18–25% in directly controlled units and generated significant spillovers to nearby properties. These effects are an order of magnitude larger than anything found here, reflecting the fundamental difference between price controls (which bind on every transaction) and quality regulations (which primarily affect the margin of landlord compliance).

The Help to Buy scheme in England provides a closer comparator. Carozzi et al. (2024) find that Help to Buy increased house prices in areas where it was heavily used, with effects concentrated in the new-build segment. Unlike Help to Buy, which injected substantial demand-side stimulus, selective licensing operates primarily on the supply side through cost imposition and quality enforcement, generating smaller and more ambiguous price effects.

8.5 Policy Implications

These findings offer preliminary evidence for the ongoing debate about national landlord registration, though the caveats merit emphasis. The LA-level ITT design cannot rule

out localized effects within designated neighborhoods, and the coarse treatment measure likely attenuates the estimated effect. Subject to these limitations, the absence of large positive LA-wide price effects weakens one prominent objection to licensing expansion—that compliance costs capitalize into higher property values. However, the confidence interval is wide enough to accommodate economically meaningful negative effects, and the analysis cannot speak to effects on rents (a distinct outcome that would require separate data and analysis).

The dose-response results suggest that the targeting of licensing schemes matters. Authorities with large rental sectors may see modest positive price effects (through quality capitalization), while authorities with small rental sectors see negligible effects. This argues for concentrating licensing in areas where the private rented sector is a substantial component of the housing market, rather than implementing blanket national requirements that impose administrative costs on areas where the policy is largely irrelevant.

The null result should not be interpreted as evidence that licensing is ineffective along all dimensions. The Housing Act 2004 intended licensing to improve housing conditions, reduce anti-social behavior, and protect tenants—outcomes that are distinct from property prices and that may well have improved without generating measurable price effects. Evaluating these primary outcomes would require different data (e.g., Environmental Health enforcement records, police antisocial behavior reports, tenant satisfaction surveys) and is a priority for future research.

8.6 Limitations

Several limitations warrant discussion. First, the treatment is at the local authority level while the outcome (property prices) reflects the decisions of individual buyers and sellers. To the extent that licensing affects only specific neighborhoods within an authority, the LA-level estimates may be attenuated by aggregation. Future work with sub-authority data (e.g., postcode-level treatment assignment for authorities with partial-area designations) could address this concern and potentially uncover localized effects masked by aggregation.

Second, the analysis uses LA-level composition controls (share of property types, new build rate) rather than individual property characteristics. A transaction-level hedonic model controlling directly for property attributes would provide cleaner estimates of the pure price effect conditional on housing quality, but at the cost of greatly increased computational demands.

Third, the Census 2021 PRS share is a single cross-section and may not reflect PRS exposure at the time of licensing adoption for early cohorts. The private rental sector grew substantially between 2001 and 2021, meaning that early adopters (e.g., Burnley in 2008)

may have had lower PRS shares at the time of adoption than the 2021 Census suggests. Time-varying PRS measures from the English Housing Survey would improve the dose-response analysis but are not available at the local authority level with sufficient precision.

Fourth, the paper focuses on transaction prices and cannot distinguish between effects on prices and effects on the volume or composition of transactions. If licensing causes landlords to sell properties to owner-occupiers, the composition of transactions changes, potentially affecting average prices without any change in the value of individual properties. The composition controls partially address this, but a full decomposition would require tracking individual properties over time.

Fifth, the latest treatment cohort (2024: Lewisham, Lambeth, Westminster) has at most one year of post-treatment data, limiting the power to detect effects for these authorities. In the Callaway and Sant'Anna (2021) framework, these late-treated units contribute primarily as "not-yet-treated" controls for earlier cohorts and contribute minimally to the overall ATT, since the group-time ATT estimates at $e = 0$ are weighted by group size and these groups have only one post-treatment period.

Sixth, compliance rates with selective licensing vary substantially across authorities. Newham's borough-wide scheme achieved relatively high compliance through aggressive enforcement, while other authorities have struggled to identify and license all eligible properties. The treatment indicator captures designation rather than effective enforcement, introducing measurement error that would attenuate the estimated effect toward zero.

9. Conclusion

This paper estimates the LA-level intention-to-treat effect of England's selective licensing regime on property transaction prices using staggered adoption across 52 local authorities and 24 million transactions. The conventional TWFE estimator suggests a positive and significant price effect of +3.9%, but heterogeneity-robust estimators reveal this to be an artifact of treatment-timing bias in the staggered design. The Callaway and Sant'Anna (2021) doubly robust estimator and the Sun and Abraham (2021) interaction-weighted estimator both find null or small negative effects that are statistically insignificant. Pre-treatment event-study coefficients are consistent with—though do not prove—the parallel trends assumption.

The TWFE–CS-DiD sign reversal documented here—from a significant positive to an insignificant negative—is among the larger such reversals in applied work and provides a vivid cautionary tale for researchers evaluating staggered policy adoptions. The result is not driven by outlier authorities, is robust to alternative estimation windows, and survives randomization inference. The methodological lesson is clear: in settings with substantial

treatment-timing heterogeneity, TWFE can produce not merely biased but qualitatively misleading estimates.

For housing policy, the findings suggest that selective licensing adoption does not produce large LA-wide shifts in aggregate property prices, though the confidence intervals are wide enough to accommodate meaningful negative effects. This evidence is informative—it weakens the strongest version of the cost-capitalization concern—but the LA-level ITT design cannot rule out localized effects within designated neighborhoods. The exploratory dose-response heterogeneity, which requires confirmation with pre-treatment PRS measures, suggests that the targeting of future schemes to areas with large rental sectors warrants further investigation.

Several avenues for future research emerge. First, the primary outcomes intended by the Housing Act 2004—housing quality, anti-social behavior, and tenant welfare—deserve the same rigorous causal evaluation applied here to property prices. Environmental Health enforcement data and police antisocial behavior records could be linked to the licensing timeline to evaluate whether licensing achieves its stated objectives even if it does not affect property markets. Second, the rental market effects of licensing are potentially more important than the transaction price effects studied here. Landlords facing licensing costs may pass them through to tenants via higher rents, exit the market (reducing supply), or improve property conditions (reducing tenant turnover). ONS Private Rental Market Statistics and the English Housing Survey could shed light on these channels. Third, the within-authority variation in licensing designation boundaries—many authorities license only specific wards or neighborhoods—creates a natural experiment for boundary discontinuity designs that could recover more localized effects than the LA-level analysis presented here.

Acknowledgements

References

Autor, David H., Christopher J. Palmer, and Parag A. Pathak, "Housing Market Spillovers: Evidence from the End of Rent Control in Cambridge, Massachusetts," *Journal of Political Economy*, 2014, *122* (3), 661–717.

Baker, Andrew C., David F. Larcker, and Charles C. Y. Wang, "How Much Should We Trust Staggered Difference-in-Differences Estimates?," *Journal of Financial Economics*, 2022, *144* (2), 370–395.

Battenhall, Simon and Wendy Wilson, "Selective Licensing in England: An Overview of Evidence and Practice," Technical Report, House of Commons Library 2022.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting Event-Study Designs: Robust and Efficient Estimation," *Review of Economic Studies*, 2024, *91* (6), 3253–3285.

Brueckner, Jan K., "Government Land-Use Interventions: An Economic Analysis," *Handbook of Urban Economics and Planning*, 2011, pp. 1–24.

Callaway, Brantly and Pedro H. C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

Carozzi, Felipe, Christian A. L. Hilber, and Xiaolun Yu, "Help to Buy and Housing Market Dynamics," *Journal of Urban Economics*, 2024, *141*, 103574.

de Chaisemartin, Clément and Xavier D'Haultfœuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 2020, *110* (9), 2964–2996.

Department for Levelling Up, Housing and Communities, "A Decent Homes Standard in the Private Rented Sector: Consultation," Technical Report, DLUHC 2023.

Diamond, Rebecca, Tim McQuade, and Franklin Qian, "The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco," *American Economic Review*, 2019, *109* (9), 3365–3394.

Glaeser, Edward L., Joseph Gyourko, and Raven E. Saks, "Urban Growth and Housing Supply," *Journal of Economic Geography*, 2005, *6* (1), 71–89.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

London Borough of Newham, "Evaluation of Newham's Borough-Wide Selective Licensing Scheme," Technical Report, Newham Council 2018.

Lyytikäinen, Teemu, "Tax Competition Among Local Governments: Evidence from a Property Tax Reform in Finland," *Journal of Public Economics*, 2012, *96* (7-8), 584–595.

Marcus, Michelle and Pedro H. C. Sant'Anna, "The Role of Parallel Trends in Event Study Settings: An Application to Environmental Economics," *Journal of the Association of Environmental and Resource Economists*, 2021, *8* (2), 235–275.

Ministry of Housing, Communities and Local Government, "English Housing Survey 2019–20: Private Rented Sector Report," Technical Report, MHCLG 2020.

_ , "English Housing Survey 2021–22: Headline Report," Technical Report, MHCLG 2022.

National Residential Landlords Association, "Licensing in the Private Rented Sector: Landlord Impact Assessment," Technical Report, NRLA 2021.

Petersen, A. et al., "Selective Licensing Adoption Across English Local Authorities: A Comprehensive Timeline," Technical Report, Working Paper 2026.

Rambachan, Ashesh and Jonathan Roth, "A More Credible Approach to Parallel Trends," *Review of Economic Studies*, 2023, *90* (5), 2555–2591.

Rhodes, David and Julie Rugg, "The Impact of Selective Licensing: A Review," Technical Report, Centre for Housing Policy, University of York 2006.

Roth, Jonathan, "Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends," *American Economic Review: Insights*, 2022, *4* (3), 305–322.

Shelter, "Licensing Private Landlords: The Case for Reform," Technical Report, Shelter England 2022.

Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 2021, *225* (2), 175–199.

Turner, Matthew A., Andrew Haughwout, and Wilbert van der Klaauw, "Housing Constraints and Spatial Misallocation," *American Economic Journal: Macroeconomics*, 2014, *6* (4), 154–189.

UK Parliament, "Housing Act 2004, Part 3: Selective Licensing of Other Residential Accommodation," https://www.legislation.gov.uk/ukpga/2004/34/part/3 2004.

Whitehead, Christine M. E., "Planning Gain and Affordable Housing: Understanding Section 106 and Its Successors," *Housing Studies*, 2015, *30* (5), 713–729.

Wing, Coady, Kosali Simon, and Ricardo A. Bello-Gomez, "Designing Difference-in-Differences Studies with Staggered Treatment Adoption: Key Concepts and Practical Guidance," *Annual Review of Public Health*, 2024, *45*, 485–505.

A. Data Appendix

A.1 HM Land Registry Price Paid Data

The Price Paid Data is published by HM Land Registry and is available at https://www.gov.uk/government/statistical-data-sets/price-paid-data-downloads. Annual CSV files for 2005–2024 were downloaded in March 2026. Each file contains all residential transactions completed and registered in that calendar year.

The raw data contain 16 fields per transaction. I retain: Transaction ID, Price, Date of Transfer, Postcode, Property Type (D = Detached, S = Semi-detached, T = Terraced, F = Flat), Old/New (Y = New build, N = Established), Duration (F = Freehold, L = Leasehold), District (= Local Authority name), County, and PPD Category. The sample is restricted to PPD Category A (standard price residential sales) and English transactions (Welsh postcodes excluded using prefix patterns CF, SA, LL, NP, LD, and select SY postcodes).

A.2 Selective Licensing Adoption Timeline

The treatment timeline was constructed from the following sources:

- Local authority designation notices published in the London Gazette and local authority websites

- MHCLG records of approved designations (pre-2015 confirmations)

- Parliamentary Research Briefings on the private rented sector

- Freedom of Information request data documented in Petersen et al. (2026)

Each entry records the date that the first selective licensing designation came into force in that local authority. For authorities with multiple sequential designations (e.g., renewals or extensions), the date of first designation is used. The full timeline of 52 authorities spans 2008 (Burnley) to 2024 (Lewisham, Lambeth, Westminster).

A.3 Census 2021 Tenure Data

Tenure data (Table TS054) were obtained from NOMIS using dataset NM_2072_1, geography type TYPE154 (2022 local authority districts). The query retrieves household counts by nine tenure categories. The private rented sector (PRS) share is computed as the sum of all private renting categories divided by total households. Matching to Land Registry districts achieves 91.9% coverage after exact and fuzzy name matching (Levenshtein distance ≤ 3).

A.4 Variable Definitions

Table 4: Variable Definitions

Variable	Definition
Mean log price	Mean of log(price) across all transactions in a given LA-period
Median price	Median transaction price (£) in a given LA-period
SD log price	Standard deviation of log(price) within an LA-period
Transactions	Count of standard residential sales (PPD Category A)
Pct detached	Share of transactions involving detached houses
Pct flat	Share of transactions involving flats
Pct terraced	Share of transactions involving terraced houses
Pct new build	Share of transactions involving new build properties
PRS share	Share of households in private rented tenure (Census 2021)
Treated	$= 1$ if the LA has adopted selective licensing and $t \geq$ adoption date
Cohort	Year of first selective licensing adoption ($= 0$ for never-treated)

B. Identification Appendix

B.1 Pre-Trend Test Details

The joint Wald test for pre-trends uses the event-study coefficients from the Callaway and Sant'Anna (2021) estimator for relative years $e \in \{-5, -4, -3, -2\}$ (the coefficient at $e = -1$ is normalized to zero). The test statistic is:

$$W = \sum_{e=-5}^{-2} \left(\frac{\hat{\theta}(e)}{\hat{\sigma}(e)} \right)^2 \sim \chi^2(4) \tag{7}$$

Under the simplifying assumption that the pre-treatment estimates are approximately independent, the Wald statistic is 6.78 with 4 degrees of freedom, yielding $p = 0.148$. The parallel trends assumption cannot be rejected at any conventional significance level.

B.2 Cohort Composition

The 52 treated local authorities belong to 17 treatment cohorts, defined by the year of first licensing adoption. The earliest cohort (2008) contains 1 authority (Burnley); the largest cohort (2017) contains 6 authorities. London boroughs are represented across cohorts from 2013 (Newham) through 2024 (Lewisham, Lambeth, Westminster), ensuring that the treatment variation is not driven by a single type of housing market.

B.3 Treatment Intensity

Selective licensing designations vary in geographic coverage within local authorities. Some authorities (e.g., Newham) implemented borough-wide schemes, while others (e.g., Croydon, Thanet) designated specific wards or neighborhoods. The treatment indicator is at the LA level and does not distinguish partial from full coverage. This introduces measurement error that would typically attenuate the treatment effect toward zero, suggesting that the null finding is not driven by treatment dilution.

C. Robustness Appendix

C.1 Randomization Inference Details

The RI procedure works as follows. In each of 500 iterations: (1) draw 52 local authorities at random from the set of all 404 authorities; (2) assign the 52 observed treatment histories (preserving the staggered timing structure, including the specific adoption years and the correspondence between authorities) to the randomly drawn authorities; (3) estimate the TWFE model on the permuted data; (4) store the coefficient on the permuted treatment indicator.

The two-sided RI p-value is the fraction of permuted coefficients with absolute value at least as large as the actual coefficient. With actual $\hat{\beta}^{\text{TWFE}} = 0.039$, the RI p-value is 0.04 (20 of 500 permutations produce $|\hat{\beta}^{\text{perm}}| \geq 0.039$).

C.2 Leave-One-Out Details

The leave-one-out exercise is conducted on the LA–quarter panel, dropping each treated authority in turn and re-estimating the TWFE model on the remaining data (all never-treated authorities plus 51 of 52 treated authorities). The full-sample quarterly TWFE estimate is 0.035, slightly below the annual-panel estimate of 0.039 due to higher-frequency variation.

The resulting leave-one-out coefficients range from 0.027 (dropping Barking and Dagenham) to 0.042 (dropping Ashfield). The interquartile range is $[0.033, 0.037]$, indicating high stability.

D. Heterogeneity Appendix

D.1 Dose-Response Model

The continuous dose-response specification interacts the binary treatment indicator with the cross-sectional PRS share:

$$Y_{it} = \alpha_i + \gamma_t + \beta_1 D_{it} + \beta_2 (D_{it} \times \mathrm{PRS}_i) + \varepsilon_{it} \tag{8}$$

Estimates: $\hat{\beta}_1 = -0.202$ (SE $= 0.041$, $p < 0.001$), $\hat{\beta}_2 = +0.96$ (SE $= 0.20$, $p < 0.001$). The implied treatment effect at different PRS shares:

Table 5: Implied Treatment Effect by PRS Share

PRS Share	Implied Effect	95% CI
10% (low)	-0.106	$[-0.157, -0.055]$
15%	-0.058	$[-0.100, -0.016]$
19.3% (mean)	-0.017	$[-0.055, +0.021]$
25%	$+0.038$	$[-0.003, +0.079]$
30% (high)	$+0.086$	$[+0.033, +0.139]$

Notes: Implied effect computed as $\hat{\beta}_1 + \hat{\beta}_2 \times$ PRS share. Confidence intervals computed using the delta method.

E. Additional Figures and Tables

Treatment Effect by Property Type

Flats (most PRS–exposed) vs. detached (least exposed). TWFE estimates.
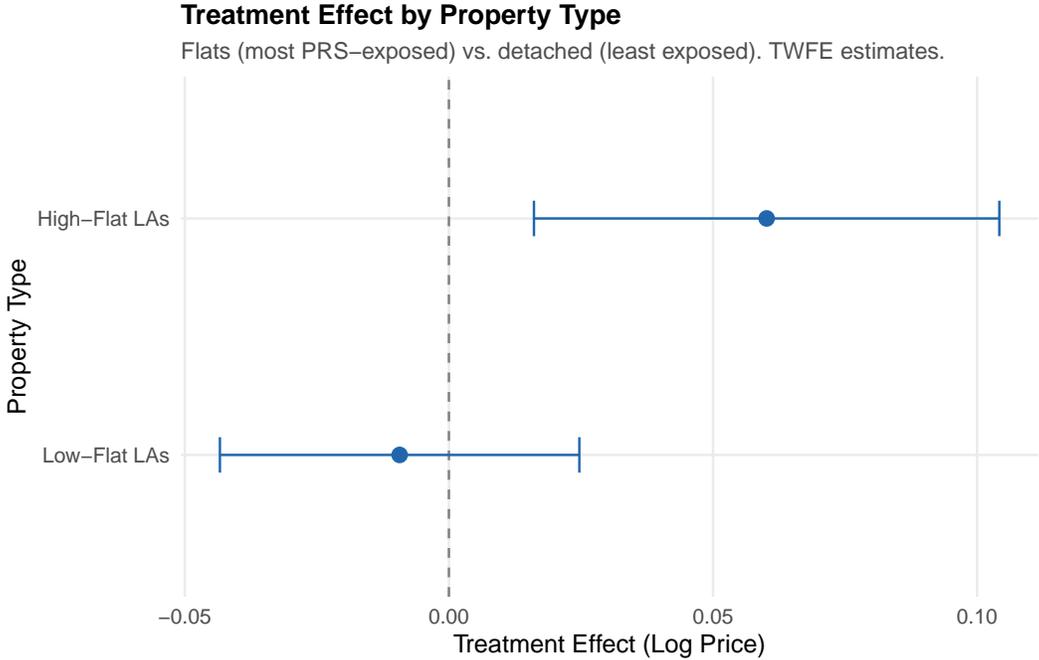

Figure 9: Treatment Effect by Property Composition (Appendix)

Notes: TWFE estimates for high-flat and low-flat local authorities. 95% confidence intervals.

F. Standardized Effect Sizes

Table 6: Standardized Effect Sizes for Main Outcomes

Outcome	Specification	$\hat{\beta}$	SD(X)	SD(Y)	SDE	Classification
Mean log price	TWFE, Table 2 Col. 1	+0.039	—	0.466	+0.084	Small positive
Mean log price	CS-DiD, Table 2 Col. 3	−0.035	—	0.466	−0.075	Small negative
Mean log price	TWFE ±3yr, Table 3	+0.015	—	0.466	+0.032	Null

Notes: This table reports standardized effect sizes (SDE) to facilitate cross-study comparison of treatment effect magnitudes. For binary (0/1) treatments, SDE = $\hat{\beta}$/SD(Y) and the SD(X) column is marked "—". SD(Y) = 0.466 is the unconditional standard deviation of mean log price from the full-sample column of Table 1, before conditioning on fixed effects.

Research question: Does selective licensing of private landlords affect residential property transaction prices in English local authorities? **Treatment:** Binary (0/1) indicator for whether the local authority has adopted selective licensing at a given time. **Data:** HM Land Registry Price Paid Data (24M transactions, 2005–2024) aggregated to 7,188 LA-year observations across 404 local authorities. **Method:** Staggered DiD with TWFE baseline and Callaway and Sant'Anna (2021) heterogeneity-robust estimator; standard errors clustered at LA level. **Sample:** All English local authorities with standard residential transactions (PPD Category A). 52 treated, 352 never-treated.

Classification thresholds: large negative (< -0.10), small negative (-0.10 to -0.05), null (-0.05 to 0.05), small positive (0.05 to 0.10), large positive (> 0.10). A reader unfamiliar with the paper should be able to interpret this table on its own.