Cursed Numbers and Blessed Streets: Secularisation and the Fading of Superstition in the UK Housing Market

Yangfanqi Liu, Helen X. H. Bao*

26th August 2026

Abstract

Britain's secularisation is usually documented through what people say about their beliefs. This paper asks whether it can also be read from what they pay. We study the price of two symbolic address attributes at opposite poles of the Christian tradition: the house number 13, long held to be cursed, and ecclesiastical street names such as Church Lane, which carry the blessing of sacred association. In roughly eight million matched Rightmove and HM Land Registry transactions across Great Britain between 2006 and 2025, both symbols are priced even under tight conditioning. Number 13 sells at a 0.23 per cent discount to its immediate neighbours, and an ecclesiastical address at a 0.72 per cent premium. Both effects, however, are disappearing. The discount reaches zero by the mid 2010s and the premium by the early 2020s. The early discount and premium alike are concentrated in the districts whose Christian share fell least between the 2001 and 2021 censuses, while the fastest-secularising districts appear to have completed the process before our window opens. The results read the housing market as an instrument of cultural observation. Surveys record what people profess and transactions record what they pay for. In Britain, the two records of religious decline agree.

Keywords: superstition, Christian affiliation, hedonic pricing, housing market

JEL Classification: R21, R30, Z12, D91

1 Introduction

Britain has spent the past half century quietly losing its religion. In the 2001 Census, 71.7 per cent of people in England and Wales identified themselves as Christian; this number fell

*Department of Land Economy, University of Cambridge, 17 Mill Lane, Cambridge CB2 1RX, United Kingdom. Contact: Yangfanqi Liu (ly350@cam.ac.uk), Helen X. H. Bao (hxb20@cam.ac.uk).

to 59.3 per cent in 2011 and to 46.2 per cent in 2021.[1] Secularisation on this scale is usually documented through what people say about their beliefs. This paper asks whether it can also be read from what people pay. We study the price of two symbolic address attributes rooted in Christian tradition: the house number 13, long held to be cursed, and ecclesiastical street names such as Church Lane or Vicarage Road, which carry the blessing of sacred association. Under a typical hedonic model, neither should move prices at all, and any systematic differential therefore measures the market value of a symbol. If that value rests on religious sentiment, its trajectory across decades offers a price-based record of secularisation.

The aversion to 13, or *triskaidekaphobia*, has deep roots in the Christian and wider European imagination. Twelve is a number of completeness, whereas thirteen is regarded as unlucky.[2] Ecclesiastical street names are the positive counterpart within the same tradition. They record the historical centrality of the parish church to British settlement, and connote antiquity, continuity, and for some buyers sanctity itself. Together the two attributes span both poles of Christian symbolism, a curse and a blessing, within a single national market.

Reports show that the British property market does price in these symbols. Rightmove, the country's largest property portal, valued more than ten million homes numbered 1 to 100 and found a £5,000 gap between number 13 and the average.[3] Zoopla, another popular online portal, reported earlier that number 13 properties sell for roughly 3 per cent less than average, and that about a quarter of British streets carry no number 13 at all.[4] These stakes are material whatever one makes of the belief itself. For a household, a discount of this size is a real wealth loss at sale. For local authorities, whether the discount exists directly influences their street numbering policy.[5] Even a sceptical buyer who holds no belief must care about the beliefs of the buyer to whom the property will one day be sold.

Headline gaps of this kind are, however, exactly what a hedonic lens should examine more closely. Properties numbered 13 are not a random sample of the housing stock, since their supply is shaped by developer and council conventions, and streets with ecclesiastical names cluster in old settlement cores whose properties differ systematically in age, size, and location. The broader academic literature on superstition clearly considers such issues, but three limits

[1] Office for National Statistics, *Religion, England and Wales: Census 2021*. The longer view from the British Social Attitudes survey is consistent: 66 per cent of respondents identified as Christian in 1983 against 38 per cent in 2018.

[2] Twelve is related to the number of apostles, months, and hours; thirteen is related to Judas, the thirteenth guest at the Last Supper who betrayed Jesus. Similarly, in Norse myth, it is Loki who arrives thirteenth and brings turmoil to the feast.

[3] Number 13 homes were valued at £354,793 against an average of £360,126, a gap of £5,333 (£5,521 in the 2024 update). See no. 13 valued lowest of 100 door numbers and 5000 lower than average.

[4] See superstitious househunters avoid unlucky number.

[5] In the UK, street numbering policy is a local matter and differs from council to council. As of 2026, more than a quarter of councils still explicitly require the omission of 13.

still stand out. First, the evidence concerns Chinese numerology almost exclusively, whether in East Asian markets (Bourassa and Peng, 1999; Chau et al., 2001; Shum et al., 2014; He et al., 2020) or among buyers of Chinese heritage in Western neighbourhoods (Fortin et al., 2014; Humphreys et al., 2019). The address superstition native to Western culture is nearly untested, with Poupakis (2025) the principal exceptions. Second, method matters. Effects estimated under coarse spatial controls shrink or vanish as the conditioning tightens. Poupakis (2025) shows for England and Wales that the well-publicised number 13 discount disappears entirely once postcode fixed effects absorb local variation. Third, the evidence is almost entirely static. Chau et al. (2001) find lucky floor premia in Hong Kong only during booms, and He et al. (2020) document informed buyers arbitraging such premia, but no study has tracked a symbolic price effect against a measured change in the culture that sustains it. Street names have their own emerging literature, from the pricing of fluency and uniqueness in Sydney (Agarwal et al., 2022) to the discount attached to streets named for Black cultural icons in Harlem (Rennert et al., 2025), yet names as carriers of religious meaning remain unexamined.

Our test aims to fill in these gaps by examining the price effect of number 13 and ecclesiastical names since 2006. The dataset matches Rightmove listings, which record property characteristics rarely observed in administrative data, to the corresponding HM Land Registry Price Paid entries, yielding roughly eight million transactions across Great Britain between 2006 and 2025. The number 13 analysis follows Poupakis (2025) to compare properties numbered 11 to 15 within the same postcode unit, so that the comparison group consists of immediate neighbours on the same street; the street name analysis draws on the full sample with 250 m grid cell fixed effects, and classifies name specifics through a combination of keyword matching and large language model tagging. Three specifications, each building on the last, structure the analysis. The first estimates the pooled price differential over the whole period. The second interacts each attribute with the year of sale to trace its path. The third interacts that path with the change in each Local Authority District's (LAD) Christian share between the 2001 and 2021 censuses, splitting the country into the districts that secularised fastest and slowest and asking whether the price paths differ accordingly.[6]

Pooled over two decades, both symbols are priced. Number 13 carries a discount of 0.23 per cent, equivalent to £600 for the average transaction, and ecclesiastical names a premium of roughly 0.7 per cent. The pooled estimates, however, mask a trend in which both effects fade over time. The discount ran at between 1 and 1.5 per cent through the late 2000s and early 2010s and has been statistically indistinguishable from zero since around 2014. The ecclesiastical premium stood at roughly 2 per cent around 2010, eroded through the following decade, and is gone in

[6] The census data covers only England and Wales. Therefore, the third specification excludes Scottish districts.

the 2020s. It is the curse that lifts first and the blessing that lingers. The third specification ties this dissipation to religious change. The early discount and premium alike are concentrated in districts whose Christian share fell least between 2001 and 2021. Districts with the largest falls show little price effect at any point in our window, implying that the same process had largely run its course there before 2006. In the slower secularising districts, the number 13 discount closes by the mid 2010s, while the ecclesiastical premium erodes by the early 2020s. Read together, the results say that Christian address symbolism was genuinely priced in the British market at the start of our sample, but that these effects disappeared as society became more secular.

The contribution is threefold. Methodologically, this is to our knowledge the largest and most finely conditioned test of address symbolism to date. Listing-matched characteristics permit controls unavailable in registry-only studies, narrow-window neighbour comparisons and highly granular fixed effects absorb the local variation that inflates raw gaps, and large language models extend name classification beyond the reach of keyword lists, with agreement checked both across two models and against the keyword benchmark. Substantively, we provide the first systematic evidence on the pricing of Christian symbolic attributes in a Western housing market, covering a negative and a positive symbol within a single design, where existing work centres on Chinese numerology. Conceptually, we show that a housing market can serve as an instrument for observing societal change. Surveys record what people profess; transactions record what they pay for. That the two move together here, with prices shedding their symbolic content as the population sheds its affiliation, suggests the hedonic record can be read as a piece of social history.

The remainder of the paper proceeds as follows. Section 2 reviews the related literature on superstition, address attributes, and hedonic pricing, and sets out our major hypotheses. Section 3 sets out the empirical framework and the three specifications. Section 4 describes the data, the matching of listings to registered prices, and the construction of the attribute tags. Section 5 reports the results, moving from pooled estimates through year-by-year paths to their interaction with local religious change. Section 6 presents robustness checks covering the underlying data and the tagging procedure. Section 7 discusses interpretation and implications, and Section 8 concludes.

2 Related Literature and Hypotheses

This section reviews the literature in two parts. The first part covers property numbers. It sets out the hedonic baseline, the evidence that numbers carry a price, and the lessons for our design. The second part covers property names. It sets out what street names are worth, and what is known about the price of religious markers more broadly. The section closes with our three hypotheses.

2.1 Property numbers and superstition

The hedonic tradition treats a dwelling as a bundle of characteristics and its price as the sum of their implicit values (Rosen, 1974). In that framework an address is informative only insofar as it indexes location. The symbols it carries, a number on a door or a word on a sign, should command no price of their own. This idea, however, has been challenged by a number of studies. For example, Salzman and Zwinkels (2017) summarises a range of behavioural biases in the housing market. Pope et al. (2015) shows that negotiating parties settle on round focal prices. These studies illustrate that, contrary to the traditional understanding, irrationality or superstition also plays a role in price. Numbers, over which people of different cultures do have preferences, may also show their influence when other conditions are carefully controlled.

The superstition literature in the housing market does provide such evidence. Bourassa and Peng (1999) provide the first hedonic evidence, a premium of around 2 per cent for *Feng Shui* lucky house numbers in Auckland. A substantial literature has followed, almost entirely within Chinese numerology, where 8 is auspicious and 4 inauspicious. Lucky floors command premia in Hong Kong (Chau et al., 2001). In Chengdu, resale flats on floors ending in 8 fetch 235 RMB per m^2 more and sell 6.9 days faster (Shum et al., 2014). In Vancouver, addresses ending in 4 sell at a 2.2 per cent discount and those ending in 8 at a 2.5 per cent premium (Fortin et al., 2014). Singapore shows both a lucky premium and an unlucky discount, along with fewer transactions on inauspicious lunar calendar days (He et al., 2020). The pattern survives migration. Chinese-American buyers, identified through a machine learning name classifier, pay 1 to 2 per cent more for addresses containing an 8 and about 1 per cent less for those containing a 4, with no premium when Chinese sellers face non-Chinese buyers (Humphreys et al., 2019). Sellers participate too, crafting lucky asking prices (Kuo and Hsieh, 2025); Larsen (2026) reviews the number superstition literature in full. Against all this, the superstition native to Western culture is nearly untested with two exceptions. Poupakis (2025) uses data from England and Wales but does not find any significant discount for house number 13 when the spatial fixed effects are tight.

Three lessons from this literature shape our design. The first concerns mechanism. A symbolic price need not reflect the buyer's own belief. Instead, it can rest on beliefs about the beliefs of future buyers (Bourassa and Peng, 1999; Fortin et al., 2014). The second lesson is that symbolic prices move. Hong Kong's lucky floor premia appear only during property booms (Chau et al., 2001); informed buyers in Singapore arbitrage the premium away (He et al., 2020), echoing the decay of published anomalies in financial markets (Marquering et al., 2006); and in theory the gap between lucky and non-lucky assets varies with market liquidity and the persistence of preference heterogeneity. The third lesson is that identification decides magnitude. Estimated effects shrink as spatial conditioning tightens, and in England and Wales the widely reported number 13 discount of around 3 per cent disappears entirely once postcode fixed effects absorb local variation (Poupakis, 2025). Taken together, although several studies of Chinese numerology have found superstition pricing, the rigorous end of the literature produces small effects or none, raising the question of whether Western address superstition is priced at all.

2.2 Street names as cultural signals

Street names carry meaning beyond their function as labels. The critical toponymy literature treats naming as an inscription of collective memory, identity, and heritage into the landscape (Azaryahu, 1996; Rose-Redwood et al., 2010; Alderman, 2008), which is the natural frame for ecclesiastical names as cultural signals, though its methods are qualitative. A small quantitative literature confirms that the name is priced. In Sydney, name fluency matters: longer names command a premium of around 0.6 per cent, unique names of 1.6 per cent, and royal names of about 3 per cent (Agarwal et al., 2022). In Harlem, properties on streets named for Black cultural icons transact at markedly lower prices than comparable properties nearby, revealing that names absorb the valuations a society attaches to what they commemorate (Rennert et al., 2025). Name effects, however, are not automatic. The gender of a street name has no detectable effect on perceived location quality or willingness to pay (Celse and Grolleau, 2025), so any premium found for a thematic category is informative rather than guaranteed. Moreover, most of what an address appears to be worth is location rather than label, as geo-spatial decompositions of Irish prices make clear (Hurley and Sweeney, 2024). Therefore, name effects must be estimated within very fine spatial cells if they are to be believed. Within this literature, religious names, among the oldest layers of British street naming, remain unexamined.

The work closest to our ecclesiastical attribute prices religious sites themselves, although the evidence is divided. Churches lower nearby house prices within about 850 feet in San Diego, by up to 3 per cent (Do et al., 1994), yet living next to godliness carries a premium of about 3.1 per cent in Henderson, Nevada (Carroll et al., 1996). Proximity to a temple raises values (Thompson et al., 2012), and in Hamburg the relationship is non-linear, with premia of up to 4.8 per cent at 100 to 200 metres but nothing within 100 metres (Brandt et al., 2014). The most recent European evidence finds churches and cemeteries capitalising as amenities in Stockholm

(Wilhelmsson, 2025). The divergence of signs across these settings suggests that whether a religious marker registers as an amenity or a nuisance depends on local context. This study differs from the literature in two ways. First, these studies price the building, with its services, sounds, and architecture, whereas a street name prices only the association, since the name persists whether or not a church still stands nearby. Second, our comparisons are drawn within 250 m grid cells, which hold shared surroundings approximately fixed, so that what remains is the symbolic layer.

2.3 Hypotheses

The literature thus establishes three propositions: symbolic attributes can be priced; symbolic prices move with the state of the market and with the population that holds the belief; and only finely conditioned designs yield credible magnitudes. Combined with the secularisation of the British population over our sample period, these motivate three hypotheses.

Hypothesis 1 (*symbolic pricing*). *Properties numbered 13 sell at a discount, and properties on ecclesiastically named streets sell at a premium.*

Under a strictly hedonic model both differentials are zero. The predicted discount derives from *triskaidekaphobia*, through a mechanism similar to the one documented in Chinese settings. The predicted premium is less automatic, since religious markers split sign across contexts and name effects can be null; but in a market whose population still reported a Christian majority at the start of the sample, the dominant association of ecclesiastical names should be favourable. Hypothesis 1 is tested by the pooled specification of Eq. (1).

Hypothesis 2 (*erosion over time*). *Over the sample period from 2006 to 2025, the magnitudes of both the number 13 discount and the ecclesiastical premium decline toward zero.*

The prior literature makes clear that symbolic effects are in any case not fixed parameters. If part of a symbolic price rests on religious sentiment, the sharp fall in Christian affiliation over these two decades should thin the population of buyers who price the symbols, and the differentials should shrink. Hypothesis 2 therefore predicts decline, and it is tested by the year interaction specification of Eq. (2).

Hypothesis 3 (*link to secularisation*). *Districts that experienced large and small declines in their Christian population share between the 2001 and 2021 censuses display different trajectories of the symbolic price effects.*

If the price of a symbol scales with the share of local residents for whom it retains meaning, then districts should display the differentials, and lose them, on a schedule set by their own religious trajectories. Hypothesis 3 is tested by the triple interaction specification of Eq. (3).

3 Empirical Framework

The key to measuring the influence of number 13 and ecclesiastical street names is to compare near-identical properties that differ in these attributes. Because symbolic attributes are not randomly assigned across space, a raw comparison conflates the attribute with the neighbourhoods and property types in which it happens to occur. We therefore absorb this variation with a comprehensive set of physical controls and granular spatial fixed effects. This setting ensures that the attribute's effect is estimated by comparing properties transacted in the same narrow locality and quarter, with controls for housing characteristics. The identifying assumption is that, conditional on these controls, the presence of the attribute is uncorrelated with unobserved determinants of price. The plausibility of this assumption depends on how finely local variation is absorbed, and we absorb spatial variation at either the postcode unit level or the 250 m grid cell level to ensure the accuracy of estimation.

The first specification tests whether the variables of interest have a significant relationship with price. Because the price effect of address elements is small relative to that of other well-explored housing characteristics, a comprehensive set of controls is needed to isolate it from potential confounders. These controls include property characteristics[7], street-name generic groups (see Appendix D), and detailed spatial and temporal fixed effects which absorb local and time-varying variation. The specification is as follows:

$$\ln P_{i,t} = \beta\, D_i + \mathbf{X}_i'\boldsymbol{\gamma} + \lambda_{g(i)} + \alpha_{s(i)} + \tau_t + \varepsilon_{i,t} \tag{1}$$

where $\ln P_{i,t}$ is the natural logarithm of the transaction price of property i sold in quarter t; D_i is the attribute of interest, typically a dummy indicating whether the property has a given characteristic (for example, a house number of 13 or an ecclesiastical street name), with coefficient β; $\mathbf{X}_i$ is a vector of property characteristic controls with coefficients $\boldsymbol{\gamma}$; $\lambda_{g(i)}$ denotes street-name generic-group fixed effects, which enter only when the spatial fixed effects are at the 250 m grid-cell level; $\alpha_{s(i)}$ denotes spatial fixed effects, absorbed at the postcode-unit or 250 m grid-cell level depending on the specification; τ_t denotes quarter fixed effects; and $\varepsilon_{i,t}$ is the error term. Standard errors are clustered at the spatial fixed-effect level.

The first specification yields a single effect averaged across the full sample period. To test

[7] The property characteristics included in this study are property type, floor area, number of bedrooms, number of bathrooms, number of living rooms, EPC rating, year built, and four binary indicators signalling whether the property is under shared ownership, is a retirement property, is sold at auction, or is newly built.

whether the effect evolves, the second specification replaces the single dummy with a full set of attribute–year interactions. The specification is as follows:

$$\ln P_{i,t} = \sum_{y} \theta_y \left(D_i \cdot Y_t^y\right) + \mathbf{X}_i' \boldsymbol{\gamma} + \lambda_{g(i)} + \alpha_{s(i)} + \tau_t + \varepsilon_{i,t} \tag{2}$$

where Y_t^y is a dummy equal to one if the sale occurs in year y and zero otherwise, and θ_y is the attribute's price effect in year y relative to properties without the attribute. We estimate this specification as a fully saturated set of attribute-year interactions with 2006 as the omitted base year, and recover each θ_y using `lincom` in Stata[8]; these coefficients trace the event-study line reported in Figure 3. Standard errors are clustered at the spatial fixed-effect level. All other terms are defined as in Eq. (1).

The second specification documents whether the effect changes over time but not what drives that change. To connect any such trend to a candidate mechanism, the third specification interacts the attribute–year terms with an indicator for areas experiencing a large decline in their Christian population share. This tests whether the time path of an attribute's price effect is systematically related to secularisation. The specification is as follows:

$$\ln P_{i,t} = \sum_{y} \theta_y \left(D_i \cdot Y_t^y\right) + \sum_{y} \phi_y \left(D_i \cdot Y_t^y \cdot C_{\ell(i)}\right) + \mathbf{X}_i' \boldsymbol{\gamma} + \lambda_{g(i)} + \alpha_{s(i)} + \tau_t + \varepsilon_{i,t} \tag{3}$$

where $C_{\ell(i)}$ is a dummy for the high-decline group. Local authority districts are ranked by the fall in their Christian population share between the 2001 and 2021 censuses and divided into terciles. Districts in the top tercile form the high-decline group, those in the bottom tercile the low-decline group, and the middle tercile is excluded from estimation so that the two groups are drawn from opposite ends of the distribution. The coefficient θ_y is the attribute's price effect in year y for the low-decline group, and ϕ_y is the additional effect in the high-decline group, so that the year-y effect for the high-decline group is $\theta_y + \phi_y$. The two sets of coefficients trace the two event-study lines reported in Figure 4 and 5. We estimate Eq. (3) as a fully saturated triple interaction and recover θ_y and $\theta_y + \phi_y$ as linear combinations of the estimated coefficients using `lincom`. As in Eq. (2), 2006 is the omitted base year, and the main effect of $C_{\ell(i)}$ and its interactions with the year effects are absorbed by the spatial and time fixed effects. Standard errors are clustered at the local authority district level. All other terms are defined as in Eq. (1).

[8]All estimation is carried out in Stata/MP 19.5 (StataNow), 4-core edition, under a University of Cambridge network licence. StataCorp. 2025. *Stata Statistical Software: Release 19.* College Station, TX: StataCorp LLC. Stata is proprietary software and replication requires a valid licence. High dimensional fixed effects are absorbed with `reghdfe` (Correia, 2017), and `gtools` is used for grouping operations on large datasets. Figures are produced in Python with `matplotlib`.

4 Data

4.1 Data sources and processing

This study draws on several datasets covering transaction records, listing records, and geographic layers carrying socio-demographic information. Property transaction prices, which serve as the dependent variable throughout the analysis, come from the HM Land Registry (HMLR) Price Paid Data.[9] These are nominal sale prices recorded at the date of transaction. Address information and the property characteristics used as controls come from Rightmove listings.[10] These characteristics include property type, floor area, number of bedrooms, bathrooms, and living rooms, and EPC rating, along with four binary indicators recording whether a property is under shared ownership, is a retirement property, is sold at auction, or is newly built. The remaining data layers serve as controls. These include the ONS Census at LSOA level (2011 and 2021) and at LAD level (2001, 2011, and 2021); the ONS Postcode Directory (February 2026); the HPSSA house price series at LSOA level; and the English Indices of Deprivation.[11] One thing worth noting is that the Rightmove data covers Great Britain, while several of the other variables are available only for England and Wales. The geographic scope of the analysis therefore begins at the level of Great Britain and narrows as specifications draw on variables with more limited coverage.

The listing data requires several cleaning steps before being merged with the HMLR and Census data. First, all rental properties and entries without a valid UPRN[12] are dropped. Among the remaining sales records, duplicate listings for the same property during the same time period but from different agents are merged[13]. Finally, records of changes for each listing are also merged[14] to ensure each listing appears only once in the dataset.

[9] Contains HM Land Registry data © Crown copyright and database right 2026. This data is licensed under the Open Government Licence v3.0. Price Paid Data is available from HM Land Registry. The address elements of the data are processed against Ordnance Survey's AddressBase Premium product, which incorporates Royal Mail's PAF database, and are used here for non-commercial research purposes only.

[10] The listing data were supplied directly by Rightmove for academic use. They are proprietary and cannot be redistributed.

[11] Census, Postcode Directory, and HPSSA data are sourced from the Office for National Statistics and licensed under the Open Government Licence v3.0. Use of the ONS Postcode Directory additionally requires the following statements: contains OS data © Crown copyright and database right 2026; contains Royal Mail data © Royal Mail copyright and database right 2026; source: Office for National Statistics licensed under the Open Government Licence v3.0. The HPSSA series is itself derived from HM Land Registry records. The English Indices of Deprivation 2019 are published by the Ministry of Housing, Communities and Local Government under the same licence.

[12] The Unique Property Reference Number (UPRN) is a persistent numeric identifier assigned to every addressable location in Great Britain. Since July 2020 the identifiers have been available as open data under the Open Government Licence v3.0.

[13] A single property can appear in listings from several different agents, since homeowners are free to switch agents.

[14] Rightmove creates a separate row in its dataset every time a listing is modified, so the same listing can appear several times in the dataset.

We then match the cleaned Rightmove listings to the authoritative HMLR Price Paid Data using postcode, house number, and, where applicable, flat number. The match performs well: every registered price in the Rightmove data is recovered in the matched dataset wherever a match exists. Following the convention in the literature, a property is treated as sold if an HMLR transaction is registered within one year of its first listing date. We also tested a two-year window, but the conventional one-year definition proved the more appropriate choice.

The final dataset spans the twenty years from 2006 to 2025. Listings before 2006 are excluded for two reasons: the Rightmove data begins in 2006, and online property portals were not yet an established part of the UK housing market before that point. Listings in 2026 are included in the regression but not reported in most tables because the data only covers the first quarter.

4.2 Variable construction and sample

The unit of observation is a property sale record. Each row corresponds to a single transaction linking a Rightmove listing to an HMLR sale. A property that sells more than once appears once per sale, in the respective year. A property advertised simultaneously by several agents contributes only one record, following the deduplication described above.

The dependent variable is the natural logarithm of the transaction price. The variables of interest, house numbers and street names, are derived from the address of each property. For a house, the number of the property itself is taken; for a flat, the flat number serves as the number of interest. Where a number is followed by a letter, as in 13A, the letter is dropped, on the reasoning that the effect of interest should be present as long as the number itself appears.

Street names are tagged via either keyword matching or large language models. Where possible, a name is separated into a generic element (the final component such as Road, Lane, or Close) and a specific element (the distinctive remainder). Generics are grouped entirely by keyword matching and enter the regressions as controls, since they proxy for property characteristics not otherwise observed.[15] Specifics are sorted into ten thematic categories[16], among which the ecclesiastical group is the primary focus of this study. To improve on keyword matching, which cannot capture the large share of invented or context-dependent specifics, the ecclesiastical specifics are identified a second time using two different large language models. Full detail on the tagging procedure and the keyword lists is provided in Appendix A and Appendix C.

Housing characteristics are grouped into bands for use as controls. Bedrooms range from 0 to 5

[15] For example, *Wharf* or *Quay* often signals a waterside or converted industrial property, while *Estate* is common among social housing.

[16] The ten categories are: prestige/regal, commemorative, ecclesiastical, nature-pleasant, topographic, water, pastoral/rural, industrial, civic/institutional, and toponymic. Their shares of the sample, and the associated keyword lists are reported in Appendix C.

or more, bathrooms from 1 to 3 or more, and living rooms from 0 to 2 or more. Floor area is split into six bands: under 50 square metres, 50 to 69, 70 to 89, 90 to 109, 110 to 149, and 150 or more. EPC rating enters as a categorical variable spanning the standard A to G bands. Year of construction is grouped according to major shifts in UK housing policy[17]: missing, pre-1900, 1900 to 1929, 1930 to 1949, 1950 to 1966, 1967 to 1975, 1976 to 1990, 1991 to 2002, 2003 to 2011, and 2012 onwards. Records that fail to match are dropped, but missing values on individual characteristics are retained as a separate category rather than excluded.

Spatial variation is absorbed at the most granular level available. Earlier studies have shown that a sufficiently fine spatial fixed effect is essential to isolate the effect of the address attributes examined here (Poupakis, 2025). We adopt the same postcode-unit fixed effects in the number 13 analysis. For the street-name analysis, however, the postcode unit is not a workable choice, because most properties sharing a postcode lie along the same street and would therefore share the same street name. These specifications instead use a 250-metre grid, which groups nearby properties across different streets while still controlling for fine spatial variation.

The estimation sample varies across specifications. The analysis of house number 13 follows Poupakis (2025) to restrict the sample to properties numbered 11 to 15. The ecclesiastical name analysis, by contrast, draws on the full sample of matched transactions.

4.3 Descriptive statistics

Table 1 reports summary statistics for the baseline estimation sample of 7,952,077 transactions. The mean transaction price is £277,930 and the median £219,950, with the gap between the two reflecting the right skew typical of housing prices and motivating the log transformation used throughout the analysis. The property-type composition is consistent with the national housing stock: semi-detached (27.7%) and detached (28.9%) properties are the two largest groups, followed by terraced (19.0%), flats (11.2%), and bungalows (10.0%). Number-13 properties account for 1.3% of the sample and ecclesiastical street names for 3.6%, corresponding to roughly 103,000 and 283,000 transactions respectively. Although these shares are small, the absolute counts remain large. Two control variables have substantial missing values: EPC rating is unrecorded for 25.2% of transactions and the built-year band for 32.6%. Rather than dropping these observations, both variables retain an explicit missing category, which preserves sample

[17] The bands mark broad eras in the UK housing stock, reflecting shifts in construction policy, building standards, and market conditions. Notable anchors include the move to mass slum clearance under the Housing Act 1930, the Parker Morris space standards made mandatory for public housing from 1967, the spread of Right to Buy and owner-occupation through the 1980s, and the market downturn of the early 1990s. Grouping by era rather than by individual year captures systematic differences in build quality, size, and layout across construction periods.

coverage. Finally, the sample is subject to two minor trims that affect only the tails of the distribution. Floor area is capped at 1,000m^2, and transactions priced below £30,000 or above £10,000,000 are dropped.

Table 1: Summary statistics

	N	Mean	SD	Median	Min	Max
Outcome						
Transaction price (in £)	7,952,077	277,930	244,381	219,950	30,000	10,000,000
Variables of interest						
Number-13 dummy	7,952,077	0.013	0.112	–	–	–
Ecclesiastical street dummy	7,952,077	0.036	0.185	–	–	–
Housing characteristics						
Floor area (in m^2)	7,952,077	108.01	47.92	98	15	1,000
Bedrooms	7,952,077	2.89	0.91	3	1	7
Bathrooms	7,952,077	1.35	0.62	1	1	5
Living rooms	7,952,077	1.53	0.71	1	0	5
Property type						
Bungalow	795,125	10.0%				
Flat	890,615	11.2%				
Terraced	1,510,218	19.0%				
Semi-detached	2,204,355	27.7%				
Detached	2,300,732	28.9%				
Unspecified	251,032	3.2%				
EPC rating						
A	4,598	0.1%				
B	274,539	3.5%				
C	1,632,752	20.5%				
D	2,824,026	35.5%				
E	971,465	12.2%				
F	192,911	2.4%				
G	45,312	0.6%				
Missing	2,006,474	25.2%				
Built-year band						
Pre-1900	374,420	4.7%				
1900–1929	946,747	11.9%				
1930–1949	802,119	10.1%				
1950–1966	1,002,135	12.6%				
1967–1975	589,714	7.4%				
1976–1990	661,793	8.3%				
1991–2002	489,490	6.2%				
2003–2011	338,052	4.3%				
2012–now	157,043	2.0%				
Missing	2,590,564	32.6%				
Transaction-type flags						
New build flag	7,952,077	0.002	0.040	–	–	–

Continued on next page

(Continued from previous page)

	N	Mean	SD	Median	Min	Max
Shared ownership flag	7,952,077	0.005	0.067	–	–	–
Retirement flag	7,952,077	0.009	0.094	–	–	–
Auction flag	7,952,077	0.010	0.101	–	–	–

Notes: The sample is the estimation sample used in the baseline specification. Dummy variables report the sample share as the mean; median, minimum, and maximum are not reported for these and for categorical variables, whose category counts (with sample shares) are shown instead. Transaction price is the nominal HMLR price paid. Floor area is capped at 1,000 m^2, and transactions priced below £30k or above £10m are dropped.

The sample spans the twenty years from 2006 to 2025 and covers Great Britain. Transactions are distributed across the period without heavy concentration in any single year, with annual shares ranging from 3.2% to 7.8% for the fully covered years. Geographically, the sample is highly granular: it spans 120 postcode areas, 2,695 postcode districts, 9,127 postcode sectors, 110,353 postcode sub-sectors, and 555,203 postcode units, and maps onto 269,706 distinct 250m grid cells.

Figure 1 plots the total number of transactions and the mean sale price for each house number from 1 to 100, and reveals several patterns. First, transaction volume drops sharply at number 13 relative to its immediate neighbours. Poupakis (2025) attributes this to a smaller supply of properties numbered 13, with some local planning conventions explicitly recommending that developers skip the number. More generally, transaction volume declines gradually as the house number rises. This likely reflects the shrinking housing stock at higher numbers. Every street begins at number 1, but not every street extends as far as 100. Second, mean price falls steeply from house number 1 to around 10, then stabilises across the remaining range. The increasing fluctuation at higher house numbers reflects shrinking sample sizes rather than substantive variation; the median price series is essentially flat across the full range. Third, within the 1 to 10 range, mean price decreases monotonically with house number. Two explanations are possible. Houses numbered 1 and 2 are often corner properties, which may carry a price premium for reasons unrelated to the number itself. Moreover, the pattern may reflect composition rather than price. Low numbers include a greater proportion of the more expensive detached houses, so that the apparent price decline is driven by the changing mix of property types rather than by the numbers themselves. Finally, and most importantly, the mean price line shows no visible drop at number 13. The avoidance of the number is clearly evident in the quantity of transactions, but any effect on price is not discernible to the naked eye.

Ecclesiastical street names are identified for 282,884 transactions, or 3.6 per cent of the estimation sample. These are spread across 1,418 distinct specifics, though the category is dominated by

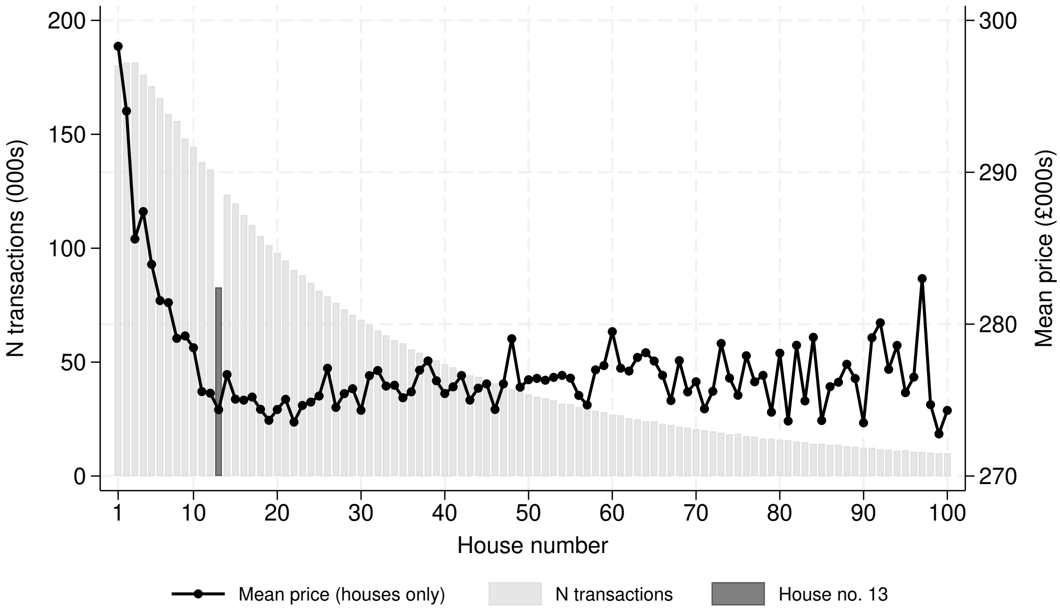

Figure 1: Total number of transactions and average price by house number (1 to 100)

a small number of high-frequency names.[18] The two large language models used to verify the tagging broadly agree with the keyword classification: Claude Sonnet 4.6 recovers 97.7 per cent of the keyword-tagged transactions and MiniMax M3 recovers 91.2 per cent, while each also identifies a further 60,000 or so transactions that keyword matching misses. Full detail on the comparison is provided in Appendix A.

Ecclesiastical streets are not a random subset of the housing stock. They are older, and their properties slightly larger, than those in the rest of the sample.[19] This is consistent with the historical pattern by which such streets cluster around old parish centres. Geographically, however, they are not concentrated in particular regions. The ecclesiastical share of transactions averages 3.6 per cent across the 365 local authority districts in the sample, with an interquartile range of 2.6 to 4.5 per cent and a maximum of 9.3 per cent.

The identifying variation for the street name analysis comes from grid cells that contain both ecclesiastical and non-ecclesiastical transactions. There are 22,616 such cells, and they contain 87.6 per cent of all ecclesiastical transactions, so the majority of the treated observations have a within-cell comparison group. The corresponding figure for the number 13 analysis is higher:

[18]The ten most common ecclesiastical specifics, with their shares of the ecclesiastical subsample, are: CHURCH (18.9%), CHAPEL (5.5%), ST JOHNS (3.8%), PRIORY (3.3%), ST MARYS (3.3%), VICARAGE (3.0%), GLEBE (2.6%), ABBEY (2.6%), RECTORY (2.6%), and ST ANDREWS (2.2%). Together these account for roughly half of all ecclesiastical transactions; the remaining 1,408 specifics form a long tail.

[19]Mean year of construction is 1939 on ecclesiastical streets against 1950 elsewhere; mean floor area is 112.2 m^2 against 107.9 m^2. The property type composition differs only modestly: ecclesiastical streets carry more detached properties (20.4 versus 18.9 per cent) and flats (12.2 versus 11.2 per cent), and fewer terraced properties (27.4 versus 29.0 per cent).

98.2 per cent of number 13 transactions occur in a postcode unit that also contains at least one transaction of a property carrying a different number.

Unconditionally, properties on ecclesiastical streets sell at a premium. The mean sale price is £292,878 against £277,379 elsewhere, and the median £229,950 against £218,500. A regression of log price on the ecclesiastical indicator alone yields a coefficient of 0.052 (s.e. 0.001), implying a raw premium of 5.2 per cent. This figure should not be read as the price of the name. The comparison above makes clear that ecclesiastical streets differ systematically from the rest of the sample, and the raw gap therefore conflates the name with the age, size, and location of the properties that carry it. For number 13, the raw data show nothing and a regression is needed to detect an effect; for ecclesiastical names, the raw data show a good deal and a regression is needed to find out how much of it is real.

5 Results

5.1 Overall effect

The two address attributes explored in this study, number 13 and ecclesiastical street names, carry a measurable price differential once physical characteristics and fine spatial variation are absorbed. A property numbered 13 sells for 0.23 per cent less than its immediate neighbours in the same postcode unit (coefficient -0.0023, s.e. 0.0007, see Figure 2); a property on a street with an ecclesiastical name sells for 0.72 per cent more than a property in the same 250 m grid cell whose street carries no such name (coefficient 0.0072, s.e. 0.0012, see Table 2). At the mean transaction price of £277,930, these correspond to a discount of roughly £600 and a premium of roughly £2,000.

Figure 2 establishes that the number 13 discount is specific to the number rather than an artefact of the estimator, the sample window, or house numbers in general. The figure repeats the estimation for every focal house number from 5 to 25,[20] in each case comparing the focal number with its two neighbours on either side under the same specification. If the design were picking up something mechanical, we would expect to see comparable coefficients scattered across the range. Every number other than 13 is indistinguishable from zero at conventional levels, and the coefficients are tightly centred on zero across the full range. Number 13 alone lies outside the confidence bands of its neighbours. This placebo shows that the effect appears where superstition predicts it and nowhere else.

[20] Houses with a number smaller than 5 are not tested because their neighbours will include corner properties numbered 1 or 2; houses with a number larger than 25 are not tested because the total sample size decreases as the focal number becomes larger.

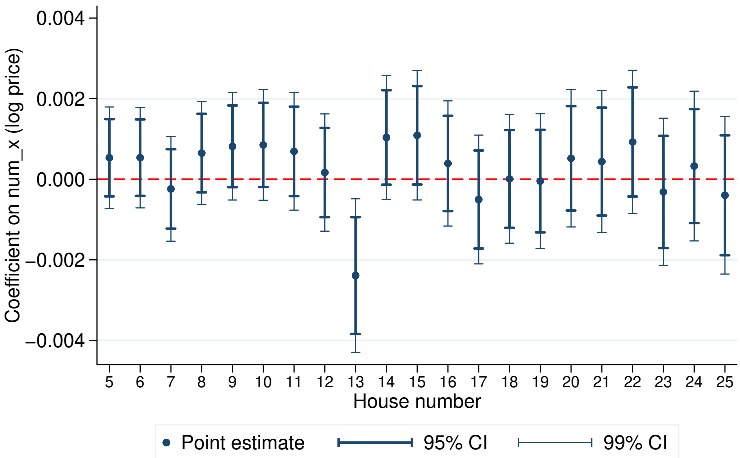

Figure 2: Price effect of each house number versus ±2 neighbours.

Note: Each marker reports the coefficient on the focal-number dummy from a separate regression of Eq. (1). For each focal number n from 5 to 25, the sample is restricted to transactions of properties numbered within two positions of n, so that the comparison group consists of the immediate neighbours on the same street. The dependent variable is the natural log transaction price. Controls comprise property type, number of bedrooms, floor area, number of bathrooms, number of living rooms, EPC rating, year built band, and indicators for shared ownership, retirement, auction, and new build sales. Postcode unit and quarter of transaction fixed effects are absorbed using the Stata command `reghdfe` (Correia, 2017). Standard errors are clustered at the postcode unit level. Thick bars give 95 per cent confidence intervals and thin bars 99 per cent. The estimation sample covers transactions from 2006 to 2025.

For ecclesiastical street names, the regression works in the opposite direction. Section 4.3 reported a raw premium of 5.2 per cent, and Table 2 shows how little of it survives. Conditioning on property characteristics, street name generics, and 250 m grid cell and quarter fixed effects reduces the premium to 0.72 per cent. Roughly seven eighths of the unconditional gap is the age, size, and location of the properties that sit on these streets rather than the name itself. The remaining eighth is not accounted for by anything we observe.

The remaining coefficients in Table 2 are structural and lexical features of street names identified in earlier work (Agarwal et al., 2022), and they are reported here for two reasons. First, they replicate. Street names with longer words and with rarer specifics command higher prices, and the signs and magnitudes are consistent with the fluency literature. Second, they do not disturb the estimates of interest. The ecclesiastical coefficient moves from 0.0088 in column (2) to 0.0072 in the full specification in column (6), and the number 13 coefficient is unchanged to the third decimal place whether or not the street name variables are included (columns 5 and 6). The symbolic attributes we study are not proxying for the lexical properties of the names that carry them.

It is worth noting that both effects are small. With a sample of seven million transactions, statistical significance is common, and a coefficient of -0.0023 invites the objection that we have detected something too minor to matter. Here we are more cautious because these are pooled estimates, averaged over twenty years, and an average is a poor summary of a quantity that is moving. If the superstition attached to number 13 and the premium on an ecclesiastical name have been fading with the congregation that once filled the church at the end of the street, then the pooled coefficient understates the effect at the start of the period and overstates it at the end. The next subsection replaces the single dummy with a full set of year interactions and asks whether this is what the data show.

5.2 The trend over time

Following Section 5.1, we estimate the second specification, Eq. (2), which replaces the single attribute dummy with a full set of attribute-by-year interactions. Each coefficient θ_y measures the attribute's price effect in year y, recovered via `lincom` so that it is reported as a level rather than as a difference from a base year. Figure 3 plots the resulting series for both attributes, with 95 and 99 per cent confidence intervals.

For house number 13, the discount is near -0.015 in 2006 and remains significantly negative through 2013, corresponding to a discount of roughly 1.5 per cent on an otherwise identical neighbour. From around 2014 the coefficient begins to rise, and for most of the years thereafter

Table 2: Influence of Address Elements on Property Prices

	(1) Number 13	(2) Thematic info	(3) Structural info	(4) Lexical info	(5) All – num13	(6) All
Number 13 dummy	-0.0083***					-0.0082***
	(0.0006)					(0.0006)
Ecclesiastical dummy		0.0088***			0.0073***	0.0072***
		(0.0011)			(0.0012)	(0.0012)
Number of words (base: 2)						
1			-0.0093***		-0.0083***	-0.0083***
			(0.0014)		(0.0014)	(0.0014)
3 or more			0.0057***		0.0038***	0.0037***
			(0.0007)		(0.0007)	(0.0007)
Number of syllables (base: 2)						
1			0.0056***		0.0056***	0.0056***
			(0.0006)		(0.0006)	(0.0006)
3 or more			0.0021***		0.0023***	0.0023***
			(0.0004)		(0.0004)	(0.0004)
Name frequency (base: most common third)						
middle				0.0006	0.0006	0.0006
				(0.0004)	(0.0004)	(0.0004)
bottom				0.0031***	0.0027***	0.0027***
				(0.0005)	(0.0005)	(0.0005)
Dictionary test				0.0011**	0.0022***	0.0022***
				(0.0005)	(0.0005)	(0.0005)
House characteristics controls	Yes	Yes	Yes	Yes	Yes	Yes
Street name generic controls	Yes	Yes	Yes	Yes	Yes	Yes
250m grid fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,952,077	7,952,077	7,914,081	7,914,081	7,914,081	7,914,081
R-squared	0.9176	0.9177	0.9178	0.9177	0.9178	0.9178

Notes: The dependent variable is the natural logarithm of the transaction price. All columns estimate Eq. (1) on the full sample of matched transactions for Great Britain, 2006 to 2025. The two attributes of interest are house number 13 and ecclesiastical street names. The number 13 dummy is included in this table only to confirm that it does not interact with the street name attributes, and the estimate reported here is not the paper's preferred estimate of the number 13 effect, because the spatial fixed effects in this table are coarser than those used for the number 13 analysis, and the sample is not restricted to the immediate neighbours of number 13. The remaining street name variables are structural and lexical features tested in earlier literature. Number of words and number of syllables count the words and syllables in the street name. Name frequency records how often the name occurs across the national universe of street names, in terciles from the most to the least common. The dictionary test indicates whether the specific is a real English word rather than an invented one. All specifications control for house characteristics (property type, floor area band, bedrooms, bathrooms, living rooms, EPC rating, built year band, and indicators for shared ownership, retirement, auction, and new build sales) and for street name generic groups (Road, Close, Avenue, and so on; see Appendix D). Spatial variation is absorbed at the 250m grid cell and time variation at the quarter of sale. Robust standard errors clustered at the 250m grid level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

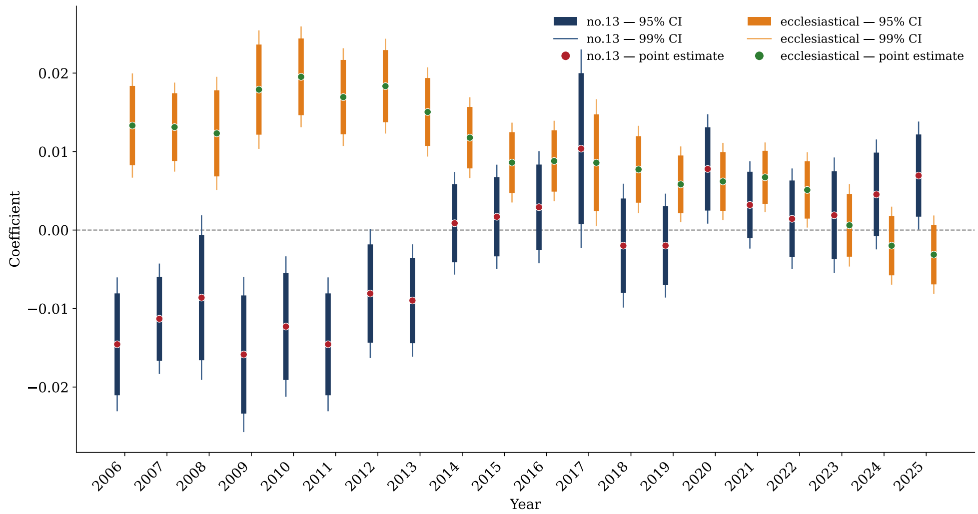

Figure 3: Price effect of house number 13 and ecclesiastical street names by year of sale.

Notes: Each point is the attribute's price effect in the given year, estimated from Eq. (2) and recovered as a level via `lincom` rather than as a difference from the 2006 base year. The two series come from separate regressions and are plotted on common axes for comparison. The number 13 series restricts the sample to properties numbered 11 to 15, so that the comparison group consists of the immediate neighbours on the same street, and absorbs spatial variation at the postcode unit. The ecclesiastical series uses the full sample, absorbs spatial variation at the 250m grid cell, and includes street name generic controls and the remaining thematic categories. Both specifications control for house characteristics (property type, floor area band, bedrooms, bathrooms, living rooms, EPC rating, built year band, and indicators for shared ownership, retirement, auction, and new build sales) and include quarter of sale fixed effects. Standard errors are clustered at the spatial fixed effect level. Thick bars give 95 per cent confidence intervals and thin bars 99 per cent. The figure covers years of sale from 2006 to 2025.

its confidence interval includes zero. By the final years of the sample the point estimate is indistinguishable from zero.

The ecclesiastical premium follows the mirror-image path. At the start of the sample the coefficient is large and positive, near +0.018 in the late 2000s, a premium of close to 1.8 per cent, and it remains firmly positive through 2013. Thereafter it declines, and the decline continues until the premium has closed by the final years and the point estimate rests on or just below zero.

Together, the two series tell a single story. Two symbolic attributes that begin the sample at opposite poles converge over twenty years on the same point of no effect. The cursed number and the blessed street name each lose their pricing power, and they do so along parallel timelines.

5.3 Religious change and the trajectory of the effect

The second specification establishes that both effects move over time but leaves open what drives the movement. The third specification connects the two trends to the broader societal secularisation by asking whether the areas where each effect fades are the same areas where the Christian population has receded most sharply. We measure the decline in religious affiliation by comparing the share of residents identifying as Christian at the LAD level between the 2001 and 2021 Censuses. The change in this share over the two decades is used to split the sample into two groups: a large-change tercile group, where the Christian share fell most, and a small-change tercile group, where it fell least. Equation (3) then traces the year-by-year effect of each attribute separately within the two groups. The initial hypothesis was straightforward: if the two effects are rooted in Christian belief, they should decay fastest where the Christian population has thinned most, so the large-change group should show the steeper downward trajectory.

For number 13, Figure 4 shows that the two groups follow different paths. In the large-change areas, the coefficient is roughly flat near zero across the whole period, with no discount to speak of at any point. In the small-change areas, by contrast, the number 13 penalty begins as a clear discount in 2006 and rises steadily towards zero over the following decade, flattening out from around the mid-2010s. The fading of the number 13 penalty is therefore concentrated in the areas where the Christian population changed least.

For ecclesiastical names, Figure 5 shows the same pattern. In the large-change areas, the ecclesiastical premium is roughly flat and close to zero throughout, with no clear trend. In the small-change areas, by contrast, the premium starts high and declines steadily year by year, following the same trajectory of gradual erosion seen for number 13. The two attributes therefore tell a consistent story: in both cases the effect fades within the areas whose Christian population changed less, while the areas that lost the most Christian residents show little movement across the sample.

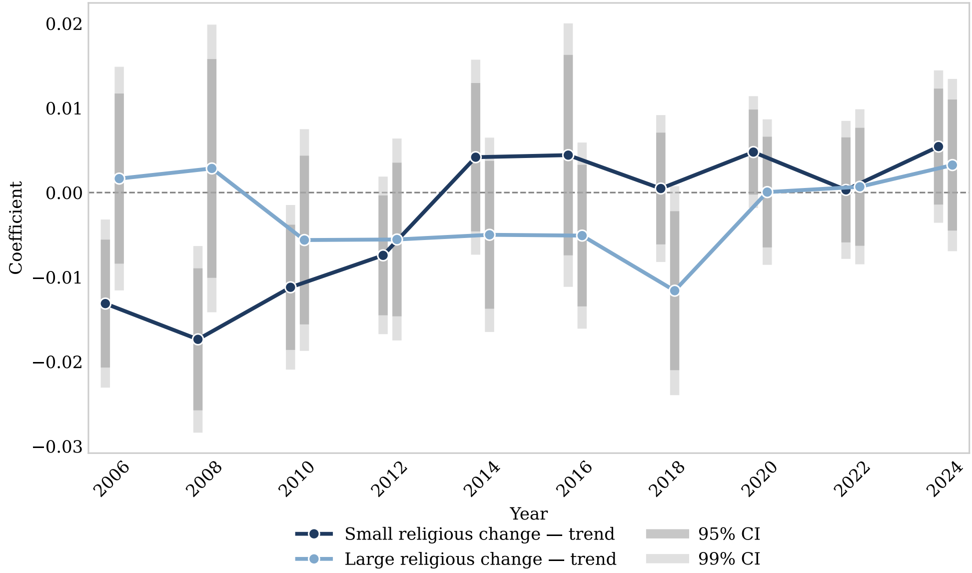

Figure 4: Yearly price effect of house number 13 in different areas

Notes: Each point is the price effect of house number 13 in the given year for the group indicated, estimated from Eq. (3) and recovered as a level via `lincom` Both lines come from a single fully saturated triple interaction. The sample is restricted to properties numbered 11 to 15, so that the comparison group consists of the immediate neighbours on the same street, and spatial variation is absorbed at the postcode unit. Controls comprise house characteristics (property type, floor area band, bedrooms, bathrooms, living rooms, EPC rating, built year band, and indicators for shared ownership, retirement, auction, and new build sales) and quarter of sale fixed effects. Thick bars give 95 per cent confidence intervals and thin bars 99 per cent. The figure covers years of sale from 2006 to 2025.

Both regressions show that effects erode where religious composition is most stable rather than where it changes most. With additional results from Appendix B, our interpretation is that the large-change areas had already completed their secularisation before the sample begins in 2006, whereas the small-change areas entered the period with the penalty still intact and worked through the same decline over the following decade. Our measure of religious change captures the decline in Christian affiliation between 2001 and 2021, but this is only the most recent stretch of a much longer process. The large-change areas are not places that never held the superstition, but places whose secularisation ran ahead of our window and had already dissolved the premium and the penalty before 2006. The small-change areas, by contrast, entered the sample with both effects intact and are observed working through their own decline.

6 Robustness Checks

The estimates reported above depend on many choices ranging from sample period selection to the details of variable construction. This section reports what we have done to replicate the

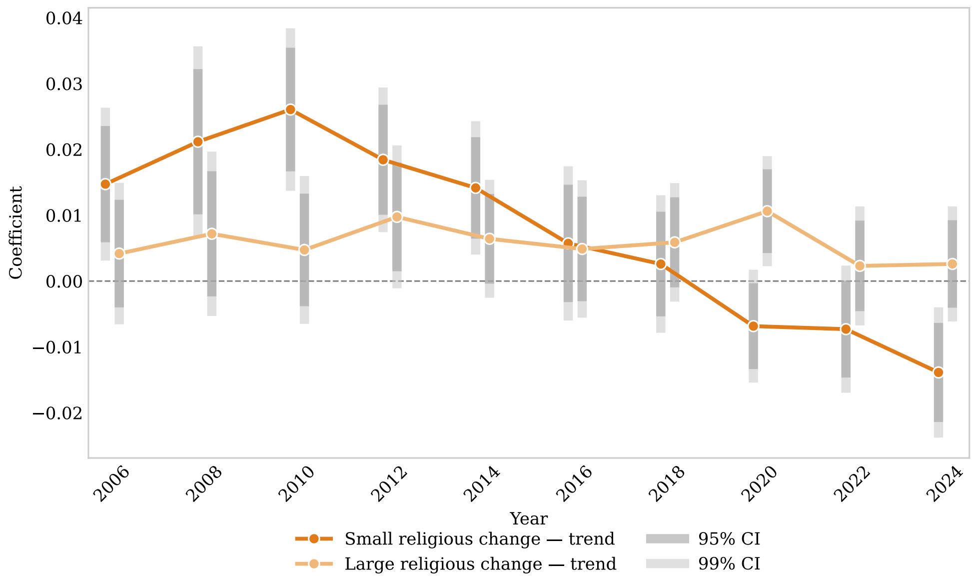

Figure 5: Yearly price effect of ecclesiastical street names in different areas

Notes: Each point is the price effect of an ecclesiastical street name in the given year for the group indicated, estimated from Eq. (3) and recovered as a level via `lincom` All other settings are the same as in Figure 4 except that the spatial fixed effects are at the 250 m grid level.

aforementioned results. Full results are reported in the appendices indicated, and none of them disagrees with the conclusions of Section 5.

One common concern is that the findings are an artefact of the matched Rightmove and HM Land Registry dataset, which covers only the transactions we were able to match and rests on a matching procedure of our own construction. We therefore re-estimate the number 13 specifications on the HM Land Registry Price Paid Data alone, which is openly available and records the universe of registered sales from 1995 onwards. Over the years both sources cover, the registry and the matched sample return coefficients that differ by less than one standard error of either, so the matched transactions are not a selected group with respect to the effect being measured. The year by year path is also recovered on the registry data alone, with no listing characteristics anywhere in the specification. The results (Figure 6) establish that the trend documented in Section 5.2 does not depend on the Rightmove attributes, although adding these attributes makes the estimation more accurate. The same appendix repeats the year by year design for every house number from 5 to 24 and finds that number 13 is the only one that moves. Appendix B reports these results in full.

Another concern is that the ecclesiastical tag is built from a fixed word list and can recognise only the names that list contains. We construct an alternative tag by asking two large language models from different developers whether each distinct street name specific carries a dominant ecclesiastical association for an ordinary English speaker. The two sets of labels agree on 98.9

per cent of cases overall and on 77.4 per cent of the cases either model tags as ecclesiastical, and both recover almost everything the keyword list finds while reaching a further 60,000 or so transactions apiece. Re-estimating Eq. (1) under each of the three tags leaves the premium positive and significant at the 1 per cent level, with the three point estimates lying within a quarter of a percentage point of one another. The choice of tagging procedure therefore does not determine the sign, the significance, or the order of magnitude of the estimate. Appendix A reports the procedure, the agreement statistics, and the three estimates.

7 Discussion

7.1 Interpreting the estimates

The first thing the estimates establish is that much of a symbolic price gap is not symbolic at all. The unconditional ecclesiastical premium is 5.3 per cent, but the conditional premium is 0.72 per cent. Roughly seven eighths of the raw gap is the age, size, and location of the properties that happen to sit on streets named for churches, and only the remaining eighth is attributable to the name. The same pattern holds for number 13, where the coefficient falls as the spatial fixed effects tighten and settles at -0.0023 once the comparison is restricted to immediate neighbours in the same postcode unit (Appendix B). Popular reports of a discount above 3 per cent are measuring something real, but most of what they measure is composition rather than superstition.

The second thing the estimates establish is that the effects of superstition are not fixed parameters. As mentioned in Section 2, most prior literature reports a single pooled coefficient and reads it as a property of the market. Our year-by-year estimates show that for both attributes the pooled figure is an average taken across a substantial decline. The number 13 discount runs near -0.015 in the late 2000s and is indistinguishable from zero from around 2014, while the ecclesiastical premium reaches close to $+0.018$ around 2010 and closes over the following decade. A study of either attribute conducted in 2008 and a study conducted in 2022 would reach opposite conclusions from the same market. More broadly, superstition is culturally specific, and a culturally specific belief has no reason to hold its price when the culture that sustains it is changing.

The third specification links the decline of both effects to the decline of the Christian population. Two symbolic attributes fading together over the same twenty years may have many common causes, but the district-level split adds a cross-sectional prediction. If the effects rest on Christian belief, their trajectories should differ according to local religious change. Our results show that in districts whose Christian share fell least between 2001 and 2021, both effects begin the sample intact and erode across the following decade. In districts whose share fell most, neither effect is detectable at any point. Our explanation is that small-change districts are late secularisers, and the fading of superstition in these districts is still happening between 2006 and 2025. Meanwhile, the large-change group is a group whose secularisation ran ahead of our observation window. Results from Figure 6 also support our interpretation.

7.2 Implications

For participants in the market as it now stands, these symbols no longer carry a price. A developer choosing whether to number a house 13 does not need to factor it in. Nor does an agent advising on a listing, or a buyer wondering whether an address will cost them at resale. The same holds for local authorities that maintain numbering conventions omitting 13. If the discount has closed, the administrative case for the convention has closed with it. Skipping the number also has the perverse consequence of making 13 conspicuous where it does survive.

We would not push this to a general recommendation, because a null result at the aggregate level reveals only the mainstream preferences people hold, and a minority neighbourhood within it may hold the superstition as firmly as ever. A neighbourhood in which a fifth of buyers retain the aversion will show little in a district average and a good deal in its own market. Section 5.3 makes the point, showing that districts differ substantially in when, and whether, the effect disappeared.

Two implications follow for the literature. The first is methodological and has been stated already. Small symbolic effects require spatial conditioning at a granular level, and the estimate is a function of that choice. The second is conceptual. Secularisation is typically measured by asking people what they believe, via survey or interview, and these methods capture profession rather than practice. A buyer's payment, however, reveals their preference at a cost. The price record and the census record moving together across twenty years suggests that the housing market can be read as an instrument for observing cultural change. The method is not confined to religion. The stock of candidate beliefs is large, and any belief that attaches to an observable, non-physical attribute of a property leaves a trace in prices that can in principle be tracked over time.

7.3 Limitations

Several limitations qualify what we have shown. The first is the selection problem. Because streets in the most superstitious places are the most likely to omit number 13 altogether, the properties we observe are drawn disproportionately from places where the number was tolerated, and our estimates of the discount are correspondingly conservative.

The second concerns our measure of religious change. The census religion question is asked at a decadal frequency, so we observe two points and infer a trajectory. Such information is available at local authority district level, which is far coarser than the grid cells and postcode units used elsewhere in the paper, so districts classified as small-change almost certainly contain neighbourhoods that were not.

The third is that the claim in Section 5.3, that the large-change districts had already secularised before 2006, is an inference from the shape of the data rather than an observation. It is the reading that makes the two figures cohere, and it is consistent with results reported in Appendix B, but we do not have accurate estimates for the pre-2006 period and cannot demonstrate it.

8 Conclusion

This paper has asked whether symbolic address attributes carry a price in the British housing market, and whether that price has moved with the country's religious composition. We answer it with roughly eight million matched Rightmove and HM Land Registry transactions covering Great Britain between 2006 and 2025, using two attributes drawn from opposite poles of the same tradition.

Our results show that pooled over the sample period, a property numbered 13 sells for 0.23 per cent less than its immediate neighbours in the same postcode unit, and a property on an ecclesiastically named street for 0.72 per cent more than a property in the same 250 m grid cell whose street carries no such name.

The pooled estimates, however, average across a substantial decline. Replacing the single dummy with a full set of attribute-by-year interactions shows that the number 13 discount runs near 1.5 per cent in the late 2000s before becoming statistically indistinguishable from zero, and that the ecclesiastical premium reaches close to 1.8 per cent around 2010 before closing over the decade that follows. The two attributes that begin the sample at opposite poles converge on the same point of no effect.

We further tie this dissipation to religious change. Splitting local authority districts by the change in their Christian share between the 2001 and 2021 censuses, we find that the early discount and premium alike are concentrated in the districts whose Christian share fell least, while districts with the largest falls show little price effect at any point in our window. This trend is consistent with the broader secularisation process in the UK.

Three contributions follow. Methodologically, this is to our knowledge the largest and most finely conditioned test of address symbolism attempted. The estimates are robust to alternative data sources, to granular spatial and temporal fixed effects, and to the choice of tagging procedure. Substantively, we provide systematic evidence on the pricing of Christian symbolic attributes in a Western housing market where the existing evidence concerns Chinese numerology almost exclusively. Conceptually, we show that a housing market can be read as an instrument for observing cultural change. Surveys record what people profess; transactions record what they pay for. The overlap of the price and census records across time and districts suggests that the hedonic record can be read as a piece of social history.

Two extensions suggest themselves. The first is to look further back. If the argument of Section 5.3 is right, the districts that show no effect in our window were losing it before 2006, and a longer price series matched to the 2001 census could verify that claim. The second is comparative. The method here is not specific to religion or to Britain. Instead, any belief that attaches to an observable, non-physical attribute of a property leaves a trace in prices that can be tracked, and the numerological superstitions documented in East Asian markets sit in cultures whose relevant affiliations are not in comparable retreat. Whether those effects remain stable now that the superstition in the UK has dissolved would say a good deal about whether what we have measured is the decay of a particular belief or the decay of belief as a category of thing that housing markets price. On the evidence here, the cursed number and the blessed street were both worth money in Britain twenty years ago, and neither is worth much now.

References

Sumit Agarwal, Maggie R. Hu, and Adrian D. Lee. Street name fluency and housing prices. *The Journal of Real Estate Finance and Economics*, 65(2):181–229, 2022. doi: 10.1007/s11146-021-09825-5.

Derek H. Alderman. Place, naming and the interpretation of cultural landscapes. In Brian Graham and Peter Howard, editors, *The Ashgate Research Companion to Heritage and Identity*, pages 195–213. Ashgate, Aldershot, 2008.

Maoz Azaryahu. The power of commemorative street names. *Environment and Planning D: Society and Space*, 14(3):311–330, 1996. doi: 10.1068/d140311.

Steven C. Bourassa and Vincent S. Peng. Hedonic prices and house numbers: The influence of Feng Shui. *International Real Estate Review*, 2(1):79–93, 1999.

Sebastian Brandt, Wolfgang Maennig, and Felix Richter. Do houses of worship affect housing prices? evidence from germany. *Growth and Change*, 45(4):549–570, 2014. doi: 10.1111/grow.12066. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/grow.12066.

Thomas M. Carroll, Terrence M. Clauretie, and Jeff Jensen. Living next to godliness: Residential property values and churches. *The Journal of Real Estate Finance and Economics*, 12(3):319–330, 1996. doi: 10.1007/BF00127540.

Jeremy Celse and Gilles Grolleau. The influence of street name gender on perceptions of location quality and rental decisions. *Review of Behavioral Economics*, 12(3):291–310, May 2025. doi: 10.1561/105.00000213. URL https://ideas.repec.org/a/now/jnlrbe/105.00000213.html.

K. Chau, Vincent Ma, and Daniel Ho. The pricing of 'luckiness' in the apartment market. *Journal of Real Estate Literature*, 9(1):29–40, 2001. doi: 10.1080/10835547.2001.12090095. URL https://doi.org/10.1080/10835547.2001.12090095.

Sergio Correia. Linear models with high-dimensional fixed effects: An efficient and feasible estimator. Working paper, 2017. URL http://scorreia.com/research/hdfe.pdf.

A. Quang Do, Robert W. Wilbur, and James L. Short. An empirical examination of the externalities of neighborhood churches on housing values. *The Journal of Real Estate Finance and Economics*, 9(2):127–136, 1994. doi: 10.1007/BF01099971.

Nicole M. Fortin, Andrew J. Hill, and Jeff Huang. Superstition in the housing market. *Economic Inquiry*, 52(3):974–993, 2014. doi: 10.1111/ecin.12066.

Jia He, Haoming Liu, Tien Foo Sing, Changcheng Song, and Wei-Kang Wong. Superstition, conspicuous spending, and housing market: Evidence from Singapore. *Management Science*, 66(2):783–804, 2020. doi: 10.1287/mnsc.2018.3198.

Brad R. Humphreys, Adam Nowak, and Yang Zhou. Superstition and real estate prices: Transaction-level evidence from the US housing market. *Applied Economics*, 51(26):2818–2841, 2019. doi: 10.1080/00036846.2018.1558361.

Aoife K. Hurley and James Sweeney. Irish property price estimation using a flexible geo-spatial smoothing approach: What is the impact of an address? *The Journal of Real Estate Finance and Economics*, 68(3):355–393, 2024. doi: 10.1007/s11146-022-09888-y.

Fang-Chang Kuo and Chia-Ming Hsieh. Crafting lucky prices in the housing market. Working Paper 5390335, August 2025. URL https://ssrn.com/abstract=5390335.

James E. Larsen. An independent literature review: superstitious numbers' influence on real estate markets. *International Journal of Housing Markets and Analysis*, 19(4):1132–1149,

2026. ISSN 1753-8270. doi: 10.1108/IJHMA-02-2025-0034. URL https://doi.org/10.1108/IJHMA-02-2025-0034.

Wessel Marquering, Johan Nisser, and Toni Valla. Disappearing anomalies: a dynamic analysis of the persistence of anomalies. *Applied Financial Economics*, 16(4):291–302, 2006. doi: 10.1080/09603100500400361. URL https://doi.org/10.1080/09603100500400361.

Devin G. Pope, Jaren C. Pope, and Justin R. Sydnor. Focal points and bargaining in housing markets. *Games and Economic Behavior*, 93:89–107, 2015. ISSN 0899-8256. doi: 10.1016/j.geb.2015.07.002. URL https://www.sciencedirect.com/science/article/pii/S0899825615001013.

Stavros Poupakis. The lack of superstition in the housing market in England and Wales. Working paper, SSRN, 2025. URL https://ssrn.com/abstract=5101113.

Lindiwe Rennert, Sonia Torres Rodríguez, and Katharine Elder. Can the names of black cultural icons save the neighborhood: Relating street names to property values and neighborhood racial composition change in Harlem, New York City. *Cities*, 163:106011, 2025. doi: 10.1016/j.cities.2025.106011.

Reuben Rose-Redwood, Derek Alderman, and Maoz Azaryahu. Geographies of toponymic inscription: New directions in critical place-name studies. *Progress in Human Geography*, 34(4):453–470, 2010. doi: 10.1177/0309132509351042.

Sherwin Rosen. Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34–55, 1974. ISSN 00223808, 1537534X. URL http://www.jstor.org/stable/1830899.

Diego Salzman and Remco C.J. Zwinkels. Behavioral real estate. *Journal of Real Estate Literature*, 25(1):77–106, 2017. doi: 10.1080/10835547.2017.12090455. URL https://doi.org/10.1080/10835547.2017.12090455.

Matthew Shum, Wei Sun, and Guangliang Ye. Superstition and 'lucky' apartments: Evidence from transaction-level data. *Journal of Comparative Economics*, 42(1):109–117, 2014. doi: 10.1016/j.jce.2013.10.001.

Eric Thompson, Roger B. Butters, and Benjamin T. Schmitz. The property value premium of a place of worship. *Contemporary Economic Policy*, 30(2):215–222, 2012. doi: 10.1111/j.1465-7287.2011.00255.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1465-7287.2011.00255.x.

Mats Wilhelmsson. The economic geography of churches: Housing market effects in stockholm.

Working Paper Series 25/14, Royal Institute of Technology, Department of Real Estate and Construction Management & Banking and Finance, Dec 2025. URL https://ideas.repec.org/p/hhs/kthrec/2025_014.html.

Appendices

A Alternative identification of ecclesiastical names

The ecclesiastical tag used throughout the main text is built from a fixed word list, reproduced in Appendix C. A list of this kind can only recognise the names it has been told to recognise. Street names, however, are often coined words, family names, or local forms that occur too rarely for any list to enumerate. LLM tagging overcomes this issue because it draws on a much broader body of background knowledge. The two tags are therefore alternative measures of the same construct. Keyword matching is narrow, fully transparent, and reproducible. LLM tagging is broader and reaches names no list can reach, at the cost of a classification rule that cannot be inspected line by line. The main text reports the keyword tag because of its reproducibility, and this appendix serves as an alternative check, showing that the results remain robust under another identification method.

The tagging proceeds as follows. Classification is carried out at the level of the unique specific rather than the transaction. Each distinct specific appearing in the Rightmove dataset is submitted once, receives one label, and that label is then merged onto every transaction whose street carries it. Different measures are taken throughout the process to make sure the results are as rigorous as possible. First, models see the specific in isolation. Second, the generic element is stripped, no location is supplied, and nothing about the property, the transaction, or the price is included in the input. Third, the model temperature is set to 0 to avoid randomness.

The prompt sets out the single question to be answered and gives examples on both sides of the line. It is reproduced in full below.

```
You are an expert in English toponymy and the history of street naming
in the UK. You are given the "specific" element of a street name in
isolation, that is, the distinctive word with its generic (Road, Lane,
Close, Street, etc.) and any location context removed. For example, you
see "Church" (from "Church Lane") or "Abbas" (from "Abbas Road"), not
the full street name and not where the street is.

Because you see the specific in isolation, you CANNOT know the true
etymology of any individual street. Do not try to. Instead, judge a
single question: does this word, used as a street-name specific in the
UK, carry a DOMINANT ecclesiastical association in the mind of a typical
English speaker?
```

```
A specific is ECCLESIASTICAL (1) when its dominant, most-plausible
reading refers to the Christian church, its institutions, buildings,
clergy, saints, religious orders, or closely related concepts. Clear
examples: Church, Chapel, Abbey, Priory, Friars, Bishop, Vicarage,
Cloister, Minster, Trinity, Convent, and saints' names used in a
religious sense.

A specific is NOT ecclesiastical (0) when its dominant reading is
secular, even if a faint religious reading is possible, and even if it
sounds old or grand. Examples: Acacia, High, Victoria, Mill, Station,
Oak, Manor, Queen.
```

Each specific returns three fields. The first is the binary label indicating ecclesiastical or not. The second is a single word giving the reason for the decision, such as SAINT, SURNAME, SETTLEMENT, or PLACENAME. The third is a three-level confidence indicator. Only the label enters the regressions. The other two fields are collected so that we can inspect the results.

The same task was given independently to two models from different developers, Claude Sonnet 4.6 and MiniMax M3, so that agreement between them carries some information about whether the labels reflect a shared reading of the language or are an artefact of one training corpus. Even with the temperature set to zero, identical output cannot be guaranteed on re-running, and we state plainly that this procedure is reproducible in method rather than in exact result.

The two sets of labels can be compared with each other and with the keyword tag. Table 3 gives four specifics and the full output of each model. The first and last are the ordinary cases, where both models agree and give the same reason. The middle two are harder to categorise. PAUL is read by one model as a saint and by the other as a surname, and SAN MARCOS as a saint by one and a placename by the other. Neither disagreement reflects an error about the world.

Table 4 cross-tabulates the two sets of labels. Raw agreement is 98.9 per cent, since the overwhelming majority of specifics are secular and both models say so. The two models agree on 77.4 per cent of the specifics that either of them tags. Set against the keyword tag, both models recover almost all of what the list finds and a good deal besides. Claude Sonnet 4.6 assigns a positive label to 97.7 per cent of the transactions the keyword list tags, and MiniMax M3 to 91.2 per cent. Each also tags a further 60,000 or so transactions that the list does not reach.

Table 5 re-estimates Eq. (1) three times, changing only how the ecclesiastical attribute is measured. The premium is positive and significant at the 1 per cent level under all three tags, and the three point estimates lie within a quarter of a percentage point of each other. Whatever else the choice of tagging method does, it does not determine the sign, the significance, or the order of magnitude of the estimate.

Table 3: Example outputs of the two models

	Claude Sonnet 4.6			MiniMax M3		
Specific	Label	Reason	Confidence	Label	Reason	Confidence
ST GREGORYS	1	saint	high	1	saint	high
PAUL	0	surname	middle	1	saint	high
SAN MARCOS	1	saint	high	0	placename	high
SEFTON	0	settlement	high	0	placename	high

Table 4: Cross-tabulation of the two sets of labels

	Claude Sonnet 4.6		
MiniMax M3	0	1	Total
0	7,912,034	56,617	7,968,651
1	31,007	300,873	331,880
Total	7,943,041	357,490	8,300,531

Notes: Counts are transactions.

Table 5: Ecclesiastical premium under alternative tagging methods

	(1) Keyword matching	(2) Claude Sonnet 4.6	(3) MiniMax M3
Ecclesiastical dummy	0.0073***	0.0067***	0.0075***
	(0.0011)	(0.0011)	(0.0011)
Tagged transactions	282,884	357,490	331,880
House characteristics controls	Yes	Yes	Yes
Street name generic controls	Yes	Yes	Yes
250m grid fixed effects	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes
Observations	7,914,081	7,914,081	7,914,081
R-squared	0.9177	0.9177	0.9177

Notes: The dependent variable is the natural logarithm of the transaction price. The three columns differ only in the construction of the ecclesiastical dummy. Robust standard errors clustered at the 250m grid level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Replication with open registry data

The design used in the main analysis takes its starting point from Poupakis (2025), who shows for England and Wales that the widely reported discount on house number 13 does not survive fine spatial conditioning. That finding is the reason the main specification compares number 13 with its immediate neighbours on the same postcode unit rather than with the wider neighbourhood, and the reason the sample is restricted to properties numbered 11 to 15. Building on that result, we use the HM Land Registry Price Paid Data (referred to as the registry data below) in this appendix and take three further steps: we move the sample window forward to the period covered by the Rightmove listing data, we add the housing characteristics that the registry does not record, and we replace the single pooled dummy with a full set of year interactions. This appendix takes those steps one at a time on openly available data, so that the contribution of each can be seen separately.

The short conclusion is that, when estimated as a single coefficient over the full registry period, the effect of number 13 is indistinguishable from zero. Estimated year by year on the same records, the effect is large and negative in the 1990s, crosses zero in the second half of the 2000s, and drifts upward thereafter. A pooled coefficient reports the average of that path, and the average of a quantity that changes sign within the window is close to zero by construction. The significant positive coefficients from the registry data largely disappear once housing characteristics from the Rightmove dataset are added.

B.1 The pooled estimate and the sample window

The registry data spans 1995 to 2025. Here we use the same method as in Section 5.1. Table 6 moves from the registry to the Rightmove sample one step at a time, holding the spatial fixed effect at the postcode unit throughout. Column (1) estimates the pooled effect on the full registry period. Column (2) restricts the same data to 2006 onwards. Column (3) restricts the sample further to those transactions that also appear in the listing data, retaining only the registry controls. Column (4) adds the housing characteristics that the listing data supplies, on exactly the same transactions as column (3).

Column (1) reaches the same conclusion as Poupakis (2025). Column (2) restricts the sample to transactions from 2006 onwards. The coefficient nearly triples and becomes significant at the 5 per cent level, even though the sample is smaller. The result shows that either the effect is not constant across the thirty-year window, or the composition of the transactions contributing to the comparison shifts over that window in ways the fixed effects do not absorb. Column (3) restricts the sample again to the transactions that are also observed in the Rightmove data,

Table 6: House number 13 and log sale price, by sample window and control set

	(1)	(2)	(3)	(4)
Number 13	-0.0005	-0.0014^{**}	-0.0021^{**}	-0.0023^{***}
	(0.0006)	(0.0007)	(0.0009)	(0.0007)
Data source	Registry	Registry	Rightmove	Rightmove
Sample window	1995–2025	2006–2025	2006–2025	2006–2025
Registry controls	Yes	Yes	Yes	Yes
Housing characteristics	No	No	No	Yes
Postcode unit fixed effects	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes
Observations	2,578,533	1,412,925	588,864	588,864
R-squared	0.932	0.924	0.947	0.962

Notes: The dependent variable is the natural logarithm of the transaction price. The sample is restricted throughout to properties numbered 11 to 15, so that number 13 is compared with its immediate neighbours. Columns (1) and (2) use the HM Land Registry Price Paid Data in full; columns (3) and (4) use the subset of those transactions matched to a Rightmove listing, and are estimated on identical samples so that the control set is the only difference between them. Registry controls are property type, tenure, and new build status. Housing characteristics are the listing attributes used in the main analysis, namely floor area band, bedrooms, bathrooms, living rooms, EPC rating, built year band, and indicators for shared ownership, retirement, auction, and new build sales. All columns absorb spatial variation at the postcode unit and time variation at the quarter of sale. Standard errors clustered at the postcode unit in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

while keeping the controls unchanged. The coefficient moves less than one standard error of either estimate, and the standard error itself rises from 0.0007 to 0.0009 as the sample shrinks. Column (4) adds the housing characteristics as controls. The point estimate moves only from -0.0021 to -0.0023, but the standard error falls from 0.0009 to 0.0007 and the estimate becomes significant at the one per cent level, while the R-squared rises from 0.947 to 0.962.

B.2 Every house number, year by year

Figure 6 estimates the effect separately for each year and for each focal house number from 5 to 24. For each focal number n, the sample is restricted to properties numbered within two positions of n, and the coefficient reported for year y is the price differential between properties numbered n and their immediate neighbours transacting in the same year, on the same postcode unit. The design is the year by year counterpart of the placebo reported in Figure 2, and it is estimated on the registry data alone, with no housing characteristics anywhere in the specification.

Number 13 traces a clear and largely monotone path, while the grey lines remain around zero. Number 13 sits below the grey band until the mid 2000s and above it from the early 2010s, and is outside it for most of the sample. It shows that the movement is specific to number 13 rather than a general property of the registry data, of the estimator, or of house numbers as a class. The trend is visible without any housing characteristics at all, so the results documented in Section 5.2 do not depend on the listing data. The clear trend, however, should not be read as a series of price effects. A discount of 2 per cent in 1995 is plausible; a premium of 2 per cent in 2025 is not. There is no account of superstition, of arbitrage, or of buyer behaviour under which a house numbered 13 is worth two per cent more than the house next door. Two explanations are available, and they are not mutually exclusive. First, the registry omits characteristics that may be related to the number itself. Since number 13 is the only house number in Britain that developers and local authorities routinely skip, the houses that carry it are a selected group in a way that houses numbered 12 or 14 are not. Second, the composition of the sample is not stable across thirty years. Because the skipping convention has been applied unevenly over time and across places, the properties contributing the comparison in 2020 are not drawn from the same population as those contributing it in 1996.

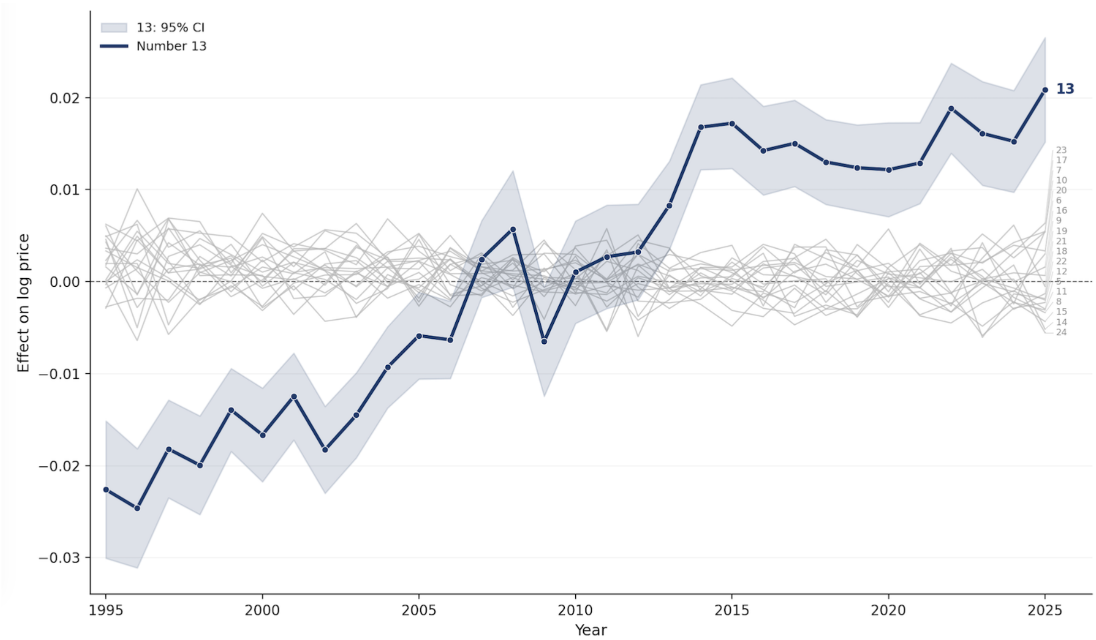

Figure 6: Yearly price effect of every house number from 5 to 24, 1995 to 2025

Notes: Each line traces the price effect of one focal house number by year of sale, estimated from Eq. (2) run separately for each focal number and recovered as a level via `lincom`. For focal number n, the sample is restricted to properties numbered within two positions of n, so that each estimate is the price differential between properties numbered n and their immediate neighbours transacting in the same year on the same postcode unit. All series are estimated on the registry data alone, with no housing characteristics anywhere in the specification. Number 13 is shown in navy with a 95 per cent confidence band; numbers 5 to 24 are shown in grey.

B.3 What the housing characteristics contribute

The years in which both datasets are available allow the contribution of the characteristics to be read directly. Figure 7 estimates the year by year specification twice over 2006 to 2025, once on the registry data with registry controls only and once on the Rightmove data with the full

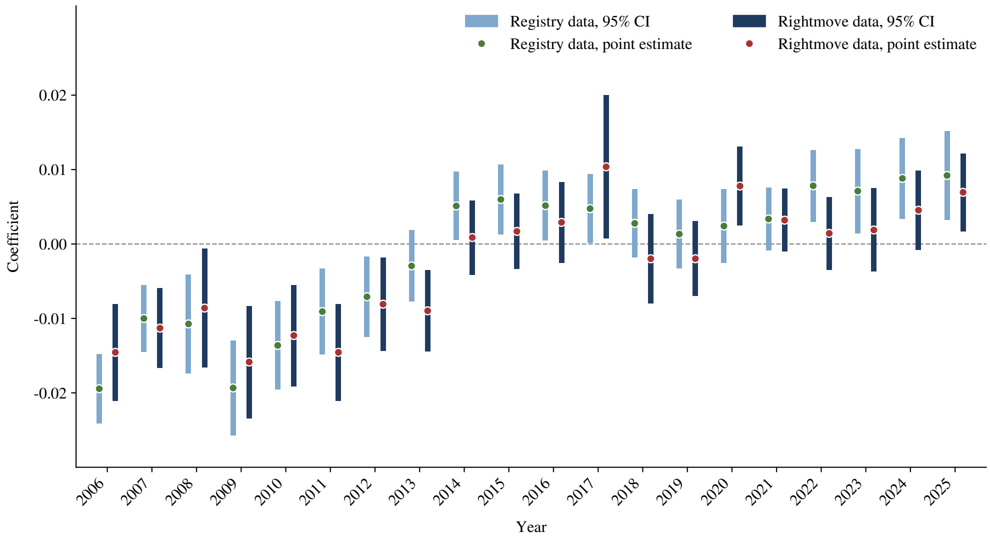

Figure 7: Yearly price effect of house number 13, registry versus Rightmove data

Notes: Each point is the price effect of house number 13 in the given year, estimated from Eq. (2) and recovered as a level via `lincom`. Both series use the same specification, sample restriction, and fixed effects, and differ only in the transactions that enter and the controls available. The registry series uses 1,412,925 transactions with property type, tenure, and new build status as the only controls; the Rightmove series uses 588,864 transactions with the full set of housing characteristics. The confidence intervals on the Rightmove series are wider throughout, which reflects the smaller sample rather than the control set. Two years in the Rightmove series, 2017 and 2020, sit noticeably away from their neighbours; the remaining years track a smooth path, and we do not read anything into the two exceptions.

control set.

Both series follow the same path; where they differ is in how far the effect is said to have travelled. On the registry data the coefficient moves from roughly -1.9 per cent in 2006 to roughly $+0.9$ per cent in 2025. On the Rightmove data with the housing characteristics included, the same movement runs from roughly -1.5 per cent to roughly $+0.7$ per cent. Around a quarter of the apparent movement in the uncontrolled series is therefore absorbed once the physical attributes of the properties are held constant.

C Keyword lists for street name specifics

The ecclesiastical attribute used throughout the main text is measured by keyword matching, so the word lists behind it are reported here in full. Specifics are sorted into ten thematic categories, all of which enter the regressions as controls even though only the ecclesiastical coefficient is reported. Together the ten categories tag 19.7 per cent of transactions in the estimation sample.

The lists were assembled from a frequency ranking of specifics across the sample. The 300 most common specifics were inspected in descending order of frequency and assigned to a category wherever the meaning was clear, with ambiguous cases left untagged. The lists are therefore not exhaustive descriptions of what each category could contain. They are the frequent end of a long tail, chosen so that the great majority of transactions a given category would ever tag are in fact tagged.

A transaction is tagged for a category if the specific of its street name contains any entry from that category's list. Matching is performed on upper case text with apostrophes and punctuation removed, and multi-word entries such as CHURCH FARM are matched as written, including the space. Because the rule is containment rather than equality, examples such as CHURCHFIELDS and OLD CHURCH are both caught by the entry CHURCH. The ecclesiastical category carries one further rule, which tags any specific whose first token is ST.

Containment is qualified by a short list of exclusions, which remove specifics that embed an entry but whose dominant reading lies elsewhere. For example, CHURCHILL contains CHURCH but reads as a commemorative surname rather than as a reference to a church. The exclusions are imposed before the containment rule, so a specific on the exclusion list is left untagged even though it contains an entry from the category's list.

Table 7 reports the ten categories, the number of transactions each tags, and the share of the estimation sample this represents.

Table 7: Keyword lists for the ten thematic categories of street name specifics

Category	**%**	**Keywords**
Nature / pleasant	5.39	OAK, OAKS, OAKWOOD, OAKFIELD, OAKLANDS, OAKLAND, OAK TREE, OAKTREE, OAKDALE, OAKDENE, OAKHILL, ELM, ELMS, ELMWOOD, ELMFIELD, ELM TREE, ASH, ASH TREE, ASHWOOD, ASHGROVE, BEECH, BEECHES, BEECHWOOD, BIRCH, BIRCHWOOD, CEDAR, CEDARS, CEDARWOOD, PINE, PINEWOOD, PINE TREE, FIR, FIRS, FIR TREE, YEW, YEW TREE, HOLLY, HOLLYBUSH, HAWTHORN, CHERRY, CHERRY TREE, CHESTNUT, MAPLE, SYCAMORE, POPLAR, HAZEL, HAZELWOOD, WILLOW, WILLOW TREE, WILLOWS, MULBERRY, ROWAN, ALDER, LINDEN, LIME, LIMES, LIME TREE, ROSE, ROSEWOOD, ROSEBANK, PRIMROSE, BLUEBELL, HEATHER, FERN, BRACKEN, GORSE, BROOM, BRAMBLE, CLOVER, IVY, BLOSSOM, LILY, DAISY, LAVENDER, ORCHARD, ORCHARDS, VINE, WOOD, WOODS, WOODLAND, WOODLANDS, WOODSIDE, WOODFIELD, FOREST, COPSE, COPPICE, SPINNEY, HOLT, COVERT, SHAW, HURST, GROVE, GROVES, FOX, FOXES, SWAN, DEER, BADGER, BADGERS, HERON, HERONS, KINGFISHER, WOODPECKER, HORSE, HORSESHOE

continued on next page

Table 7 *continued from previous page*

Category	**%**	**Keywords**
Ecclesiastical	3.57	CHURCH, CHURCHFIELD, CHURCHFIELDS, CHURCH FARM, CHURCH HILL, CHURCH VIEW, CHRIST CHURCH, CHRISTCHURCH, CHAPEL, CHAPELFIELD, ABBEY, ABBEYDALE, PRIORY, ALL SAINTS, TRINITY, TEMPLE, GLEBE, GLEBELAND, BISHOP, BISHOPS, RECTORY, VICARAGE, PARSONAGE, FRIAR, FRIARS, FRIARY, MONK, MONKS, MONKSWOOD, CLOISTER, CLOISTERS, MINSTER, CATHEDRAL, CHANTRY, CONVENT, DEANERY, NUNNERY, HOLYWELL, HOLYROOD, and any specific whose first token is ST
Prestige / regal	2.52	KING, KINGS, KING'S, KING EDWARD, KING GEORGE, KING CHARLES, KINGS HEAD, QUEEN, QUEENS, QUEEN'S, QUEEN ELIZABETH, QUEEN ANNES, QUEEN MARY, ROYAL, CROWN, PRINCE, PRINCES, PRINCESS, DUKE, DUKES, EARL, EARLS, LORD, LORDS, KNIGHT, KNIGHTS, NOBLE, CASTLE, PALACE, REGENT, REGENTS, MANOR, MANOR HOUSE, TOWER, TOWERS, GROSVENOR, WINDSOR, BALMORAL, SANDRINGHAM, RICHMOND, HAMPTON, MARLBOROUGH, KNIGHTSBRIDGE
Topographic	1.87	HILL, HILLSIDE, HILLCREST, HILL TOP, HILLTOP, HILLVIEW, HILLFIELD, HILL VIEW, MOUNT, MOUNT PLEASANT, MOUNTAIN, CLIFF, CLIFFE, DALE, DALES, VALE, GLEN, GLENDALE, DENE, RIDGE, RIDGEWAY, COOMBE, COMBE, DOWN, DOWNS, KNOLL, FELL, SUMMIT, BROW, BEACON, CRAG, MOOR, MOORLAND, MOORSIDE, MOORFIELD
Pastoral / rural	1.37	FARM, MANOR FARM, HOME FARM, HALL FARM, CHURCH FARM, PARK FARM, FARMHOUSE, MEADOW, MEADOWS, MEAD, MEADS, GRANGE, GRANGE PARK, BARN, TITHE BARN, BARNFIELD, PADDOCK, PADDOCKS, BARTON, PASTURE, FIELD, FIELDS, CROFT, CROFTS, HARVEST, SHEPHERD, DAIRY, FURLONG
Water	1.24	BROOK, BROOKFIELD, BROOKSIDE, SPRING, SPRINGFIELD, SPRINGWELL, RIVER, RIVERSIDE, RIVERBANK, LAKE, LAKESIDE, POND, FORD, BOURNE, AVON, HAVEN, MERE, STREAM, BURN, BURNSIDE, BECK, WELL, WELLS, FOUNTAIN, WATERSIDE, MARSH, MOSS, CARR, FEN
Commemorative	1.18	VICTORIA, VICTORIA PARK, ALBERT, WELLINGTON, NELSON, NELSONS, CHURCHILL, GLADSTONE, PEEL, DISRAELI, MOUNTBATTEN, KITCHENER, HAVELOCK, CLIVE, CECIL, STANLEY, JUBILEE, CORONATION, EMPIRE, MEMORIAL, ARMISTICE
Civic / institutional	1.10	SCHOOL, SCHOOL HOUSE, GRAMMAR SCHOOL, COLLEGE, STATION, MARKET, HALL, HALL PARK, HALL FARM, HALLS, LIBRARY, HOSPITAL, INSTITUTE, ACADEMY
Industrial / working	0.78	MILL, MILLS, MILLFIELD, MILLBROOK, MILLBANK, MILL HILL, MILL HOUSE, CORN MILL, SILK MILL, MILL POND, WINDMILL, FORGE, FOUNDRY, KILN, BRICK KILN, LIME KILN, LIMEKILN, SMITHY, QUARRY, GAS, BRICK, IRON, COLLIERY, TANNERY, WHARF, WHARFE

continued on next page

Table 7 *continued from previous page*

Category	%	Keywords
Toponymic	0.65	KINGSTON, KINGSLEY, KINGSBURY, QUEENSBURY, QUEENSBERRY, EARLSTON, KNIGHTON, CASTLETON, CASTLEFORD, MONKTON, CHAPELTON, BISHOPTON, TEMPLETON, CHURCHTOWN, ASHTON, ASHFORD, ASHBOURNE, ASHBY, ASHBURTON, BURNHAM, BURNS, BURNETT, DOWNHAM, DOWNING, DOWNTON, MOORE, MOORES, MOORHOUSE, WELLINGTON, WELLAND, WELLESLEY, FORDHAM, BECKETT, BECKET, FENTON, FENWICK, MARSHALL, CARRINGTON, CARRICK, FOXTON, SWANSTON, BARNES, BARNETT, BARNARD, BARNTON, CROFTON, HALLAM, HALLIDAY, MEREDITH, CLIFFORD, WOODFORD, CHRISTIE, STANLEY, CECIL, CLIVE

Notes: Categories are ordered by share of the estimation sample. A transaction is tagged if the specific of its street name contains any entry in the corresponding list, matched on upper case text with punctuation removed. Shares are net of the exclusions described above. Apart from those exclusions the categories are not mutually exclusive, and a specific may carry more than one tag. All ten dummies enter the regressions reported in the main text as controls, where only the ecclesiastical coefficient is shown.

D Street name generics

A typical street name divides into two elements. The generic is the final component, the word identifying the thing named as a street of a particular kind, such as Road, Close, or Lane. The specific is the distinctive remainder, the part that names something in the world, such as Church, Victoria, or Oak. The two elements do different work. The generic is descriptive, telling the reader what sort of street this is and, by implication, what sort of property sits on it. The specific is where the meaning of a name resides, and it is the element this paper is concerned with. Generics enter the analysis only as controls.

They are controlled because they are informative about property characteristics that the transaction record does not capture. Some of this information is physical. *Wharf* and *Quay* indicate a waterside setting or a conversion from industrial use; *Mews* indicates converted stabling in an older urban core; *Bungalows* indicates single-storey stock. Some of it concerns layout and provenance. *Close* and *Drive* are characteristic of post-war cul-de-sac development, while *Lane* and *Street* are older and more central; *Estate* is common among social housing. None of this is recorded in the listing fields, and none of it is fully absorbed by the year built band or by a 250 m grid cell. Leaving generics out would require the specific to carry information about built form that properly belongs elsewhere.

The generic is taken as the final element of the street name and matched against a list of 140 common forms. These cover almost 95 per cent of transactions in the estimation sample, and the distribution is heavily concentrated. Table 8 reports the thirty most frequent forms. Road alone accounts for 31.4 per cent of transactions, the six most common forms for 70 per cent, the ten most common for 81 per cent, and the thirty in the table for 91 per cent. A small number of dummies therefore does most of the work, and the remaining 110 forms cover less than 4 per cent of the sample between them.

The 140 forms do not enter as 140 separate dummies. They were first estimated individually, with Road as the omitted baseline on the grounds that it is by some distance the most common form and therefore the most natural point of comparison. Forms whose coefficients could not be distinguished from that baseline were then pooled, since a dummy that does not move the estimate adds parameters without adding information. The retained set is the set plotted in Figure 8 and the set entering Eq. (1) to Eq. (3) as $\lambda_{g(i)}$.

Table 8: The Thirty Most Common Street Name Generics

Generic	Freq.	%	Cum. %	Generic	Freq.	%	Cum. %
ROAD	2,602,728	31.36	31.36	PARK	63,584	0.77	87.27
CLOSE	908,285	10.94	42.30	VIEW	53,054	0.64	87.91
AVENUE	686,743	8.27	50.57	RISE	46,693	0.56	88.47
STREET	627,790	7.56	58.14	GREEN	38,404	0.46	88.94
DRIVE	532,923	6.42	64.56	MEWS	29,817	0.36	89.29
LANE	445,468	5.37	69.92	SQUARE	26,218	0.32	89.61
WAY	366,705	4.42	74.34	CROFT	20,139	0.24	89.85
CRESCENT	216,531	2.61	76.95	MEADOW	17,630	0.21	90.07
GARDENS	197,305	2.38	79.33	MEAD	13,620	0.16	90.23
COURT	135,435	1.63	80.96	GATE	12,357	0.15	90.38
GROVE	135,315	1.63	82.59	END	11,646	0.14	90.52
PLACE	100,483	1.21	83.80	BANK	10,929	0.13	90.65
WALK	79,539	0.96	84.76	ROW	10,876	0.13	90.78
TERRACE	77,358	0.93	85.69	FIELD	9,761	0.12	90.90
HILL	67,586	0.81	86.50	CHASE	9,738	0.12	91.02

Notes: Frequencies are transactions in the matched Rightmove and HM Land Registry sample for Great Britain, 2006 to 2025. The generic is the final element of the street name. Forms are ranked by frequency and read down the left column and then down the right. The 140 forms identified in total cover almost 95 per cent of transactions; the thirty shown here cover 91 per cent.

Figure 8 plots the estimated coefficient on each retained generic relative to the Road baseline, taken from the full specification of Eq. (1). The pattern is legible in terms of the property types the forms indicate. Forms associated with waterside and converted stock sit at the top of the distribution, forms associated with post-war estate development sit near the middle, and forms associated with social and lower value stock sit at the bottom. The ordering is what one would expect if the generic were serving as a summary of the physical and locational characteristics of the street. That said, these coefficients should not be read as the value of a name. They are conditional associations that stand in for characteristics the data do not observe. For example, the appropriate interpretation of a large coefficient on *Quay* is that properties on quays differ from properties on roads in ways the controls do not otherwise capture, not that the word itself is worth money. The main analysis in Section 5 includes the generics shown in Figure 8 as controls, to proxy for characteristics not recorded in the Rightmove dataset.

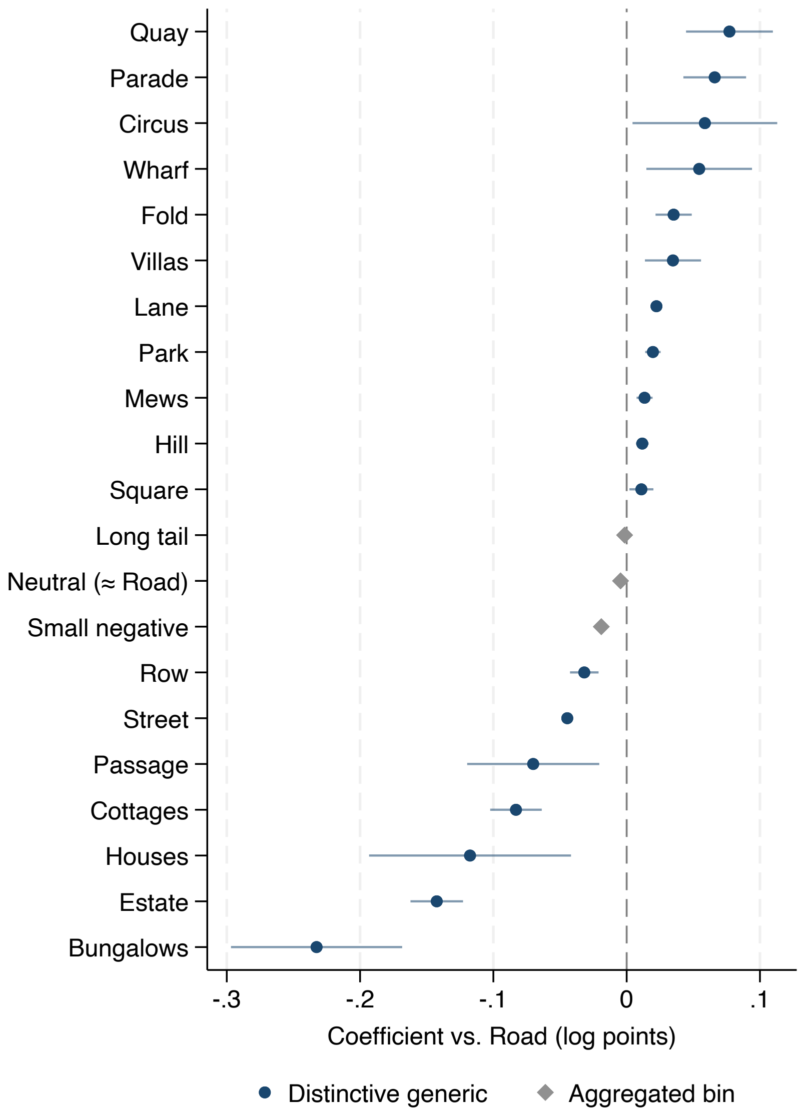

Figure 8: Estimated Price Effect of Street Name Generics

Notes: Coefficients from Eq. (1) estimated on the full sample, expressed relative to the Road baseline. Controls, fixed effects, and clustering are as in Table 2. Forms whose coefficients were not distinguishable from the baseline are pooled and are not shown separately.