CAGE

Policy Report

November 2025

Solving Britain's Housing Crisis: Where and how to build 1.5 million homes

Nikhil Datta

Amrita Kulka

With research assistance from Jawad Wehbe, Edoardo Badii, Harshita Anandan and Matthew Stubbs

UNIVERSITY OF WARWICK

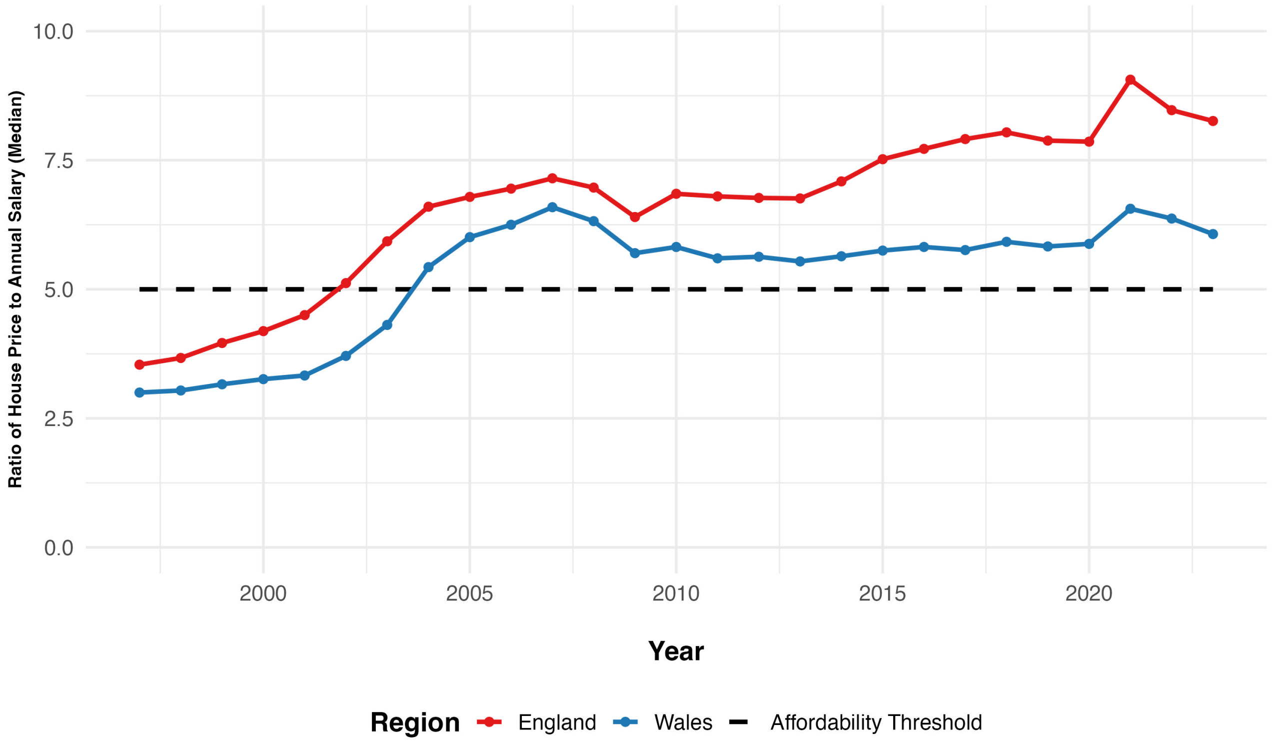
UKRI Economic and Social Research Council

Contents

1 Introduction

Over the past three decades the UK has seen housing affordability plummet. In 1997, the median house price was approximately three times the median annual salary. Today the median house price is more than eight times the median annual salary in England, and six times in Wales (see figure 1), and these changes are no longer being partially offset by low borrowing costs – instead these have been at a 16 year high over the past two years. A lack of new housing is the main culprit in this crisis (Watling and Breach, 2023; Hilber and Vermeulen, 2016) and housing supply has been consistently under the government's own target by a large margin.

Figure 1: Ratio of median house price to median wage



Note: The figure shows the ratio of the median house price to the median wage for England and Wales for the period 1997 to 2023. Source: Office for National Statistics.

The new Labour government has pledged to build "1.5 million" new homes during their five-year term while also making adjustments to housing targets by Local Authority. This report seeks to answer one key question: Where should those homes be built?

We use data on over 20 billion carefully cleaned housing searches from Rightmove—the UK's most popular property portal—together with the full history of rental and sales listings from 2019–2024. Searches, rather than realised prices, provide the closest measure of unconstrained location preferences, as they are not tied to whether a property is actually available in an area nor to how much properties cost.

From these data, we construct two new statistics. The **housing gap** is the difference between the number of searchers in an area and the number of available homes, capturing excess demand: the larger the gap, the larger this excess demand and the more additional homes are needed to meet it. **Tightness** is the ratio of searchers to homes, reflecting competition for a single property and linking directly to how hard or quick it may be to secure one. The higher the tightness, the more competition for a single property. Put another way, the gap measures the number of people affected by the lack of housing, while the tightness reflects the intensity of the effect.

We calculate these two measures for over 200,000 Output Areas in Great Britain, where the median area is one twentieth of a square kilometre, run several validation checks on them to confirm that they are robustly measuring where people want to live, and then use them to demonstrate where housing is most demanded at the neighbourhood level.

Using these measures, in particular the housing gap, we compare where there is excess demand to the government's new targets, which are set at the Local Authority (LA) level. We then disaggregate the demand measures down to within LA, to make detailed comments on the types of areas which are in most need of additional residential development.

In particular, we consider what type of housing projects are needed: densification, involving the development of large towns' and cities' inner cores; urban extensions, which would expand large towns and cities outward; small town extensions, meaning development within and around large villages and small towns; and new rural developments, which create new towns and villages or expand small rural settlements. Demand in the form of urban and town extensions as well as new rural developments often intersects existing planning designations, such as greenbelts, which we present extensive evidence on.

All geographically detailed results on the housing gap and tightness, along with planning designations and other important statistics can be found on our mapping tool linked to this paper: wheretobuild.warwick.ac.uk. We have additionally produced 350 Local Authority-specific reports, all of which can be found on the CAGE website. We hope the results presented here will be of use to policy makers at both the national and local level.

The report is laid out as follows. Section 2 documents how housing has been built over the past two decades. It first presents details on the process of how targets have been set and how Local Planning Authorities have attempted to meet those targets. We then present novel statistics on frictions in the planning system and where housing has been built, including its access to amenities and employment, as well as how fast it has been built.

Section 3 introduces our housing measures in detail and discusses how areas are ranked according to these measures at a highly granular level.

Section 4 is designed to inform both national and local stakeholders. Section 4.1 and 4.2 focus on micro-geographies. Section 4.2 begins by comparing local patterns of housing delivery since 2010 with the areas of highest demand and then moves on to highlighting where housing should be delivered within local planning authorities. Section 4.3 identifies key frictions in the current planning system and shows how they connect to the housing gap at a hyper-local level. Section 4.4 then aggregates these statistics up to the local authority level and compares them with current government targets.

2 How housing has been built in the past 25 years

2.1 From Targets to Planning Permission

The UK has seen repeated reforms to its housing and planning framework, beginning with the Town and Country Planning Act 1947, which established the discretionary planning system. The 1950s brought the first modern greenbelts[1], followed in 1968 by the creation of county-level Structure Plans. From the 1990s onwards, the focus shifted upward to the regional scale, first to Regional Planning Guidance (1991) and later Regional Spatial Strategies (2004). The pendulum then swung back towards localism under the Coalition Government after 2010, marking another major shift in how planning was conceived and delivered.

A consistent feature throughout this period, however, has been the presence of housebuilding targets, variously framed as political pledges, indicative requirements, or legally binding annualised supply figures. The responsibility for setting, calculating, and delivering these targets has shifted repeatedly, and they have often been a source of political contention. Table 1 below summarises how this process has changed in England since 2004, with additional details on the method of calculation for the National Planning and Policy Framework (NPPF).

Each successive approach to housing targets has attracted criticism. Under the Regional Spatial Strategies, allocations were frequently viewed as overly top-down, inflexible, and vulnerable to political adjustment, fueling resentment among local authorities. The Objectively Assessed Need (OAN) framework under the 2012 NPPF proved costly and slow, with councils often reliant on consultants to produce Strategic Housing Market Assessments, which generated significant inconsistencies across areas and became a frequent source of dispute at examination which delayed plan-making. More recently, the standard method introduced in 2018 has been criticised both for its volatility when based on household projections and for its blunt affordability adjustment. The most recent 2024 stock-based revision has been seen as more stable and transparent, but arbitrary and insufficiently sensitive to local demand pressures. In particular, it is not obvious that excess demand will be well correlated with existing housing stock, especially given how the spatial economy in the UK has shifted over the past 50 years.

A key aspect of housing targets is that the level of geography they are targeted at are typically LPAs, the planning arms of Local Authorities. As such, they are quite geographically coarse with only around 400 LPAs across Great Britain. LPAs are guided by their Local Plans, which are statutory documents intended to coordinate development in the local area and which cover several factors beyond residential housing, including job creation, the environment, transport, and local economic growth. Local Plans must look at least 15 years ahead, and are assessed by a Planning Inspector to ensure they are "sound" — that is, positively prepared, justified, effective, and consistent with national policy. In practice, this means that LPAs must show they are planning to meet their assessed housing target while also weighing competing pressures such as land availability, environmental designations, and infrastructure capacity.

The process of deciding where new housing goes is shaped by both evidence and consultation. LPAs typically issue a "Call for Sites," inviting landowners, developers, and residents to propose land for development. These sites are then evaluated for their suitability, availability, and deliverability, often using detailed technical assessments on transport links, flood risk, or environmental impact. LPAs also consult statutory bodies such as the Environment Agency and Natural England, alongside local residents and businesses. The outcome is a mapped set of allocations within the Local Plan that indicate where new homes will be built and in what quantity. While this process is designed to be evidence-led, politics inevitably plays a role: choices between greenfield development and urban intensification, for example, are politically sensitive. Nonetheless, once a Local Plan is adopted, it provides the framework against which individual planning applications are judged, meaning that it is the central mechanism through which LPAs decide both how many homes to build and where they should go.

Importantly, England's discretionary planning system means that the adoption of a Local Plan does not, by itself, guarantee that housing will be built on allocated sites. Developers must still submit planning applications, which are decided on a case-by-case basis by the LPA in light of the Local Plan, material considerations, and national policy. This creates flexibility for site-specific circumstances to be taken into account, but also introduces uncertainty: allocations set the strategic direction, while delivery ultimately depends on discretionary decisions about individual proposals.

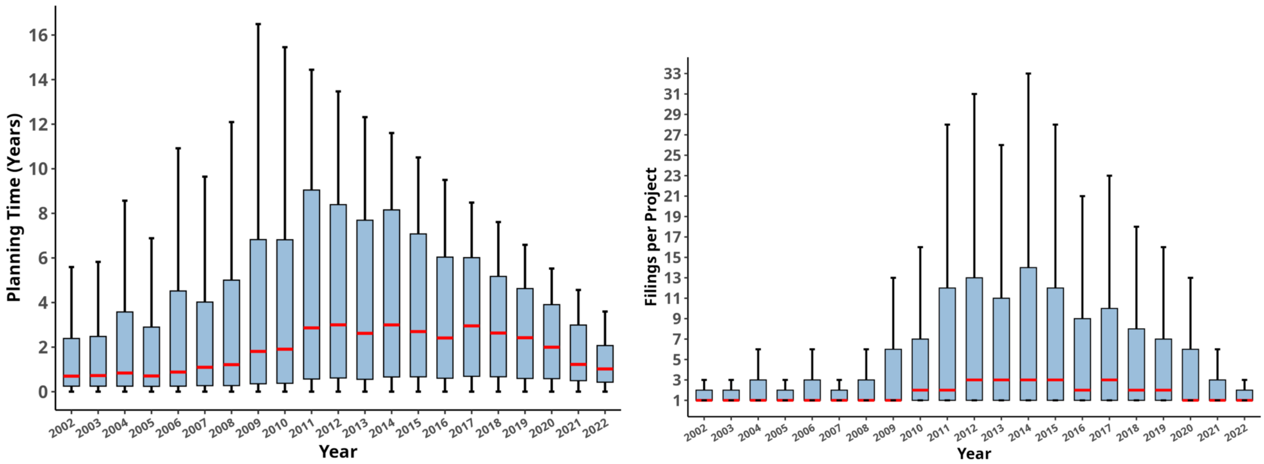
[1] Officially the UK's first greenbelt was in 1938 around London, however this differed somewhat from the modern idea of greenbelts as growth boundaries.

Table 1: UK Housing Target Policy History[2]

Period	Policy	Who Sets Targets?	How Are Targets Calculated?	Legally Binding?
2004–2010	Regional Spatial Strategies (RSS)	Regional Planning Bodies, approved by the Secretary of State	Based on demographic analysis, migration trends, and labour market forecasts.	Yes.
2010–2012	Localism Act (Transition Period)	Local Planning Authorities	Responsibility shifted to Local Planning Authorities, who set targets based on Strategic Housing Market Assessments.	No.
2012–2018	National Planning Policy Framework (NPPF) – Objectively Assessed Need (OAN)	Local Planning Authorities	Determined through Strategic Housing Market Assessments. Calculations used household projections, adjusted for past under-supply, market signals, employment forecasts, and affordable housing requirements.	Expected to meet OAN unless environmental constraints justified otherwise.
2018–2020	National Planning Policy Framework (NPPF) - Standard Method	Local Planning Authorities using formula given by DLUHC	**Step 1: Baseline** Projected average annual household growth over 10 years. **Step 2: Affordability Adjustment** If the workplace-based affordability ratio exceeds 4, the baseline is increased by 0.25% for each 1% that the ratio is above 4. **Step 3: Cap** Housing need is capped at 40% above the most recent requirement—Local Plan if <5 years old, otherwise Spatial Development Plan if <5 years old, otherwise the higher of the Step 1 baseline or the latest plan.	No. Provides a minimum starting point; deviations allowed with evidence.
2020–2024	National Planning Policy Framework (NPPF) - Standard Method (2020 Revision)	Local Planning Authorities using formula given by DLUHC	**+ Step 4: Cities and Urban Centres Uplift** Apply a 35% uplift for urban local authorities. A 35% uplift is then applied for those urban local authorities in the top 20 cities and urban centres list. Where a cap is applied in Step 3, the uplift is applied after the cap.	No. Provides a minimum starting point; deviations allowed with evidence.
2024–present	National Planning Policy Framework (NPPF) - Standard Method (2024 Revision)	Local Planning Authorities using formula given by MHCLG	**Step 1: Baseline** Calculate 0.8% of existing housing stock. **Step 2: Affordability Adjustment** If the five-year average workplace-based affordability ratio exceeds 5, the baseline is increased by 0.95% for each 1% above 5.	No. Provides a minimum starting point; deviations allowed with evidence.

[2]This table shows the history of housing target policies only for England. Wales and Scotland have devolved housing delivery policies with Welsh Local Planning Authorities setting housing numbers in their Local Development Plans informed by the Welsh government's Local Housing Market Assessment Tool. Scottish Local Planning Authorities prepare a Local Housing Strategy based

Figure 2: Distribution of Weighted Project Planning Time and Number of Filings



Note: This figure shows the distribution of weighted (by number of units in a project) planning duration and number of filings per project from 2002 to 2022. The boxes show the Inter-Quartile Range (IQR), which is the range between the 25th percentile (Q1) and the 75th percentile (Q3); the red line is the weighted median. Whiskers extend to the smallest and largest observations within $[Q_1 - 1.5\times\text{IQR}, Q_3 + 1.5\times\text{IQR}]$; outliers are omitted for clarity. The planning data set we use was constructed as described in Appendix B.3.

This discretionary structure has important implications for how planning unfolds in practice. While Local Plans establish a framework, actual delivery depends on multiple rounds of applications, negotiations, and sometimes appeals. Over the past two decades, this has coincided with a notable shift in how long projects spend within the planning system and how much paperwork they require.

As figure 2 illustrates, the average duration of projects has lengthened considerably since the early 2000s, with median timelines rising from around one to two years to several years by the 2010s. At the same time, the number of filings per project has increased, reflecting more iterative and protracted engagement with the planning process.[3] Together, these trends point to a system where bringing forward housing has become progressively more complex and time-consuming, even once land has been allocated in Local Plans.

2.2 Realised Planning & Build Patterns

To examine where residential builds have been applied for and actually built over the past 25 years we use detailed data on 1.4 million planning applications, which we have collected from Local Authority planning portals and 3.2 million residential builds from NHBC. For more details please see Appendix B. We classify builds into four types, based on the location of the residential development and the population of the village, town or city, known more formally as Built up Areas (BuAs) and Built up Area Subdivisions (BuASDs).[4]

We describe these four categories, Densification, Urban Extension, Small Town Extension and Rural Development below in table 2. The categories are selected with regards to the literature on housing supply which has long considered the importance of densifying already built-up areas (e.g. Anagol et al. (2021), Kulka et al. (2024), for Cities (2025), Schuetz (2020)) while town extensions and new towns are categories explicitly considered by the New Towns Taskforce set up by the government in September 2024 to support the delivery of the next generation of new towns (New Towns Taskforce, 2025).

Figure 3 below gives examples of actual residential development sites over the past 20 years in Cambridge, Birmingham, Leamington Spa, and Northstowe, and demonstrates the classification procedure for urban densification, extensions and new rural developments. Further details on the procedure are given in Appendix C.1.

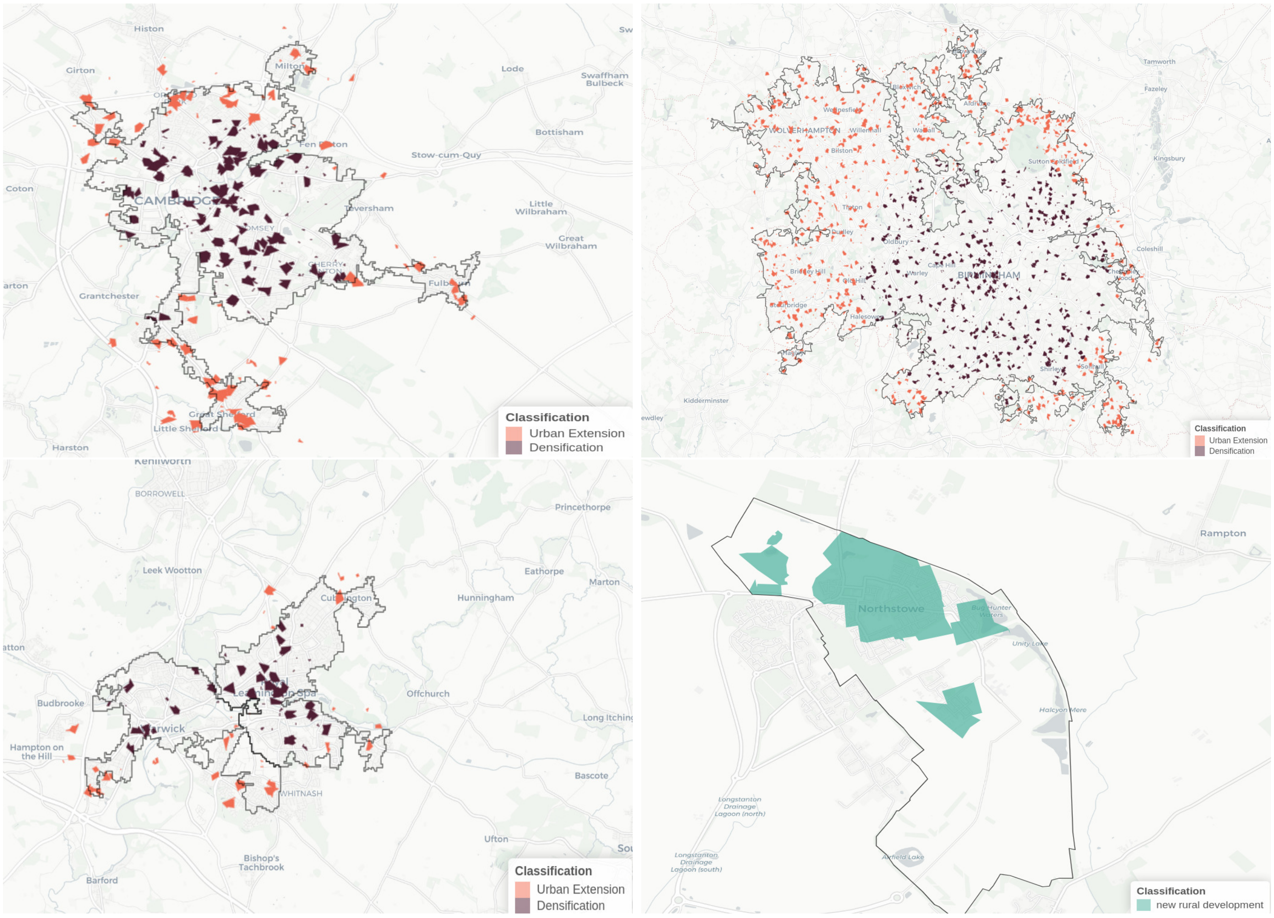
Over the past two decades there has been a fundamental shift in where residential developments have taken

on local Housing Need and Demand Assessments. We take these different systems into account when we discuss housing targets in section 4.4.

[3]The planning data runs up to September 2025. As such the number of filings and planning time is censored for the later years - by construction, a planning application submitted in January 2022 is truncated to have taken less than 4 years, even if it is still within the planning system. Therefore, the reduction in times from 2019 onwards should be treated with due caution.

[4]For details on how we use these definitions, please see appendix C.1. For a map of all BuASDs, see figure C.1 in the appendix.

Figure 3: Maps of Sites by Classification



Note: The figure shows the location of sites in Cambridge, Birmingham, Leamington Spa, and Northstowe according to our classification. The data on sites and builds is collected from NHBC. For Birmingham, Leamington Spa, and Cambridge, the boundary is that of the BuASD. For Northstowe, we use the town boundary from OpenStreetMaps. Due to data confidentiality restrictions, we distort the location and shape of all sites, randomly remove some sites, as well as simulate additional sites. For details regarding the classification, please refer to Appendix section C.1.

Table 2: Classification of Residential Builds

Category	Definition
Densification	a) Development in the inner 70% of any BuASD with a population larger than 20,000. Inner 70% measured around the Lower Layer Super Output Areas (LSOA) with the highest employment density.
	b) Largest 5 BUAs in Great Britain (London, Birmingham, Manchester, Liverpool, Glasgow): Development in the inner 50% of any BuASD with more than 20,000. Inner 50% measured around the LSOA with the highest employment density.
Urban Extension	a) Development in the outer 30% extending to a 1km buffer of any BuASD with more than 20,000.
	b) Largest 5 BUAs in Great Britain (London, Birmingham, Manchester, Liverpool, Glasgow): Development in the outer 50% extending to a 1km buffer of any BuASD with more than 20,000.
Small Town Extension	All development within a 1km buffer of BuASD with population between 5,000-20,000.
Rural Development	Development at least 1km away from any existing BuASD and all new builds within a BuASD with less than 5,000 population.

place. The right-hand side panel in figure 4 shows that two decades ago, approximately 50% of new builds were classified as densification, while that number has dropped to around 15%. Similarly, the proportion of urban extensions roughly doubled, from about 25% up to approximately 50%. From 2010, the combined proportion of rural developments and small town extensions has increased by around 5-10 percentage points. The left-hand side shows similar patterns, though with some deviations for planning applications for residential builds. In particular, it shows a slightly less pronounced drop in densification and increase in urban extensions, while also showing an increase in new rural developments.

It is important to note that planning applications represent intentions or possibilities, while completions represent realised supply. These two figures could therefore differ for several reasons. These include differences over time in acceptance rates, decisions to actually undertake the build by developers (which may be driven by cost considerations and associated conditions attached to acceptances), as well as differences in build-out rates (how long development actually takes).[5]

Amenity and Employment Access

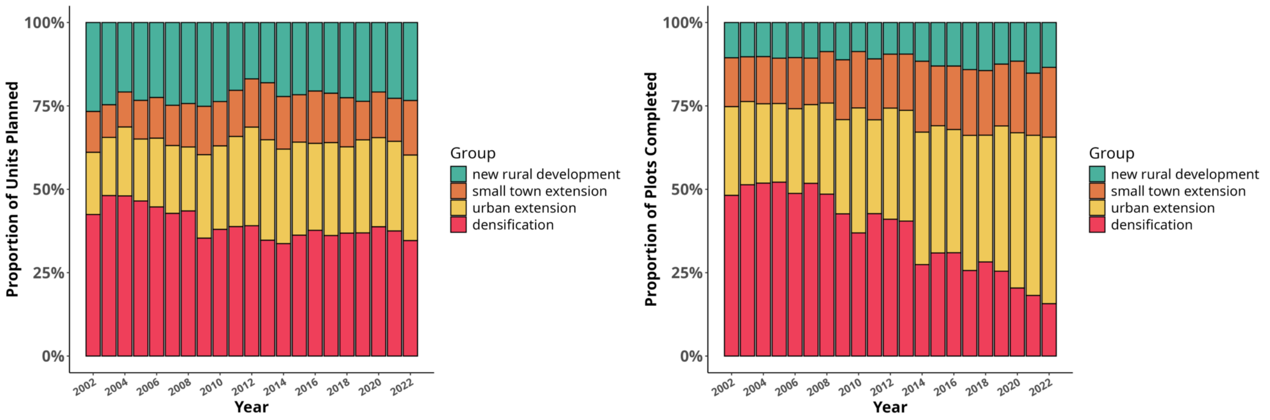
As the locations of residential builds have changed, so has the degree of local amenities. As shown in figure 5, in the early 2000s only around 11% of new builds were located in BuASDs without a GP or secondary school. This figure has risen to and peaked at approximately 18% in recent years. The availability of amenities such as schools, surgeries, shops, and community centers are important considerations for households deciding where to live.[6]

More generally, new builds tend to be located in areas with poorer employment access than the existing housing stock. Employment access is also a major factor determining a location's attractiveness for households (e.g. Ahlfeldt et al. (2015)). To measure employment access, we combined data on the precise locations of new builds, existing population counts by OA/LSOA, and the spatial distribution of jobs by OA/LSOA from the Business Register and Employment Survey. We then used detailed travel time calculations — during regular commuting hours and by both car and public transport — for all OA/LSOA pairwise combinations in Great Britain, obtained from the Travel Time API. This approach allows us to calculate, for both existing housing

[5] There is also a possibility that a small portion of the difference may be driven by selection of the NHBC sample, which has shrunk from representing about 90% of new builds 10 years ago, to about 70% today. However, given a large part of the difference emerged during a stable period of representativeness for the NHBC data, we do not believe this is a major contributing factor.

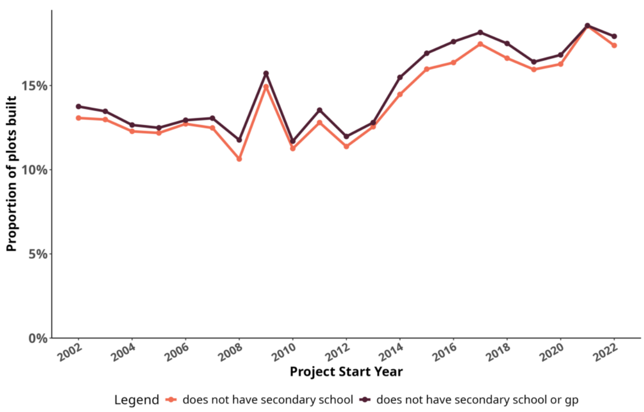
[6] See for example https://www.bbc.co.uk/news/uk-england-cambridgeshire-66156561

Figure 4: Proportions of Units Planned vs Actual Builds by Group Over Time



Note: The left panel shows the proportion of planned units by group from 2002 till 2022. In contrast, the right panel shows the proportion of actual builds by group over the same period. Both figures cover Great Britain. Data on builds is collected from NHBC. For details on construction of planned units and classification, please refer to Appendix sections B.3 and C.1 respectively.

Figure 5: Proportions of Builds in BuASDs in England without Secondary School and/or GP Over Time

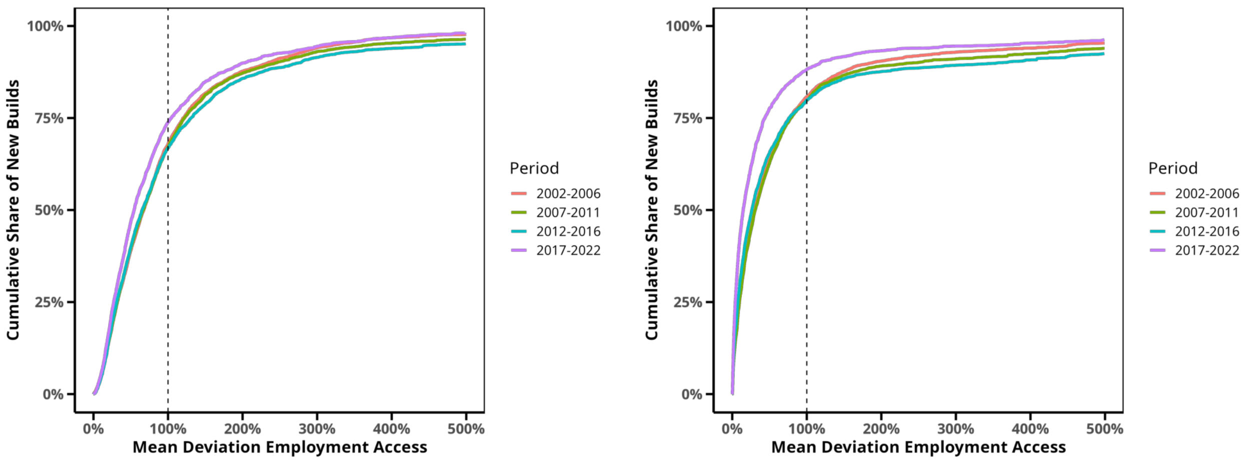


Note: The figure shows the proportion of builds between 2002 and 2022 in BuASDs across England that has no secondary school and/or General Practitioner (GP). Schools and GP coordinates were matched to their corresponding BuASD and an indicator was created for each variable. Data on location of schools and GP are collected from gov.uk and House of Commons Library respectively.

stock and new builds, the number of jobs that can be reached within a given travel time.

Figure 6 below shows employment access for new builds over the past two decades relative to the average access across the entire population, for the fastest transport mode between an origin and destination (left panel), and public transport only (right panel). The horizontal axis measures job access relative to the national average: 100% means equal to the national average, values above 100% mean better access (more jobs within reach), and values below 100% mean worse access. The vertical axis shows the cumulative share of new builds, so the curve traces out how new homes are distributed across areas with different levels of job access. The left hand panel implies that around 70% of all new builds have worse than average employment access by all transport modes, while the right hand panel suggests this is even further exacerbated when considering commuting only via public transport.[7]

Figure 6: Share of Builds Relative to Mean Deviation of Employment Access



Note: Both figures show the share of new builds in Great Britain for different periods relative to the mean deviation of employment access across LSOAs. The left panel uses employment access within 25 minutes of an LSOA, determined by the fastest available mode of transportation (private or public). The right panel uses employment access within 25 minutes of public transport. Data on builds and employment are collected from NHBC and Office of National Statistics respectively. For details on the calculation of employment access and mean deviation employment access, please refer to Appendix section C.2.

Speed of Development

The government's pledge to deliver 1.5 million homes is ambitious not only in scale but also in the speed at which development is expected. Because planning and construction often happen at the project level, where the number of potential units varies, the size of a project may influence how quickly homes can be delivered. Larger projects might benefit from economies of scale, enabling faster build-out. This suggests a possible trade-off: while developments in less dense areas may offer poorer access to amenities, they could support quicker housing delivery. This is partially born out in figure 7, which shows that urban extensions have on average 30 residential units per project, while densification has 24 and small town extension has 22. Surprisingly, new rural developments have even fewer.

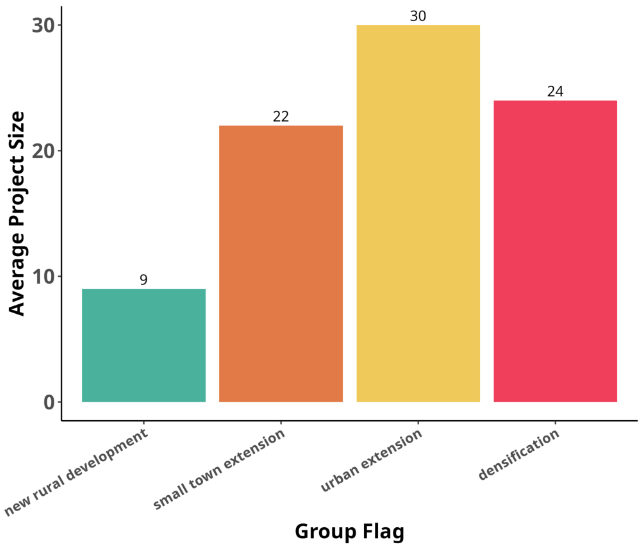
However, despite their larger project sizes, we find no evidence that urban extensions deliver housing more quickly. Figure 8 illustrates this by showing two measures: the *time development projects spend in the planning system*[8] (left-hand panel, measured as the period between the first and last filing), and the *average unit-level build-out time* (right-hand panel, measured as the time from development project start to plot completion), across our four development types.

On average, densification projects spend 37 months in the planning system, compared to 39 months for small town extensions and 49 months for urban extensions. These gaps are even larger at the 75th percentile. Similarly, unit-level build-out averages are fastest for small town extensions (28 months) and densification (29 months) and slowest for urban extensions (33 months).

[7]For more details on the calculations please see section C.2 in the appendix.

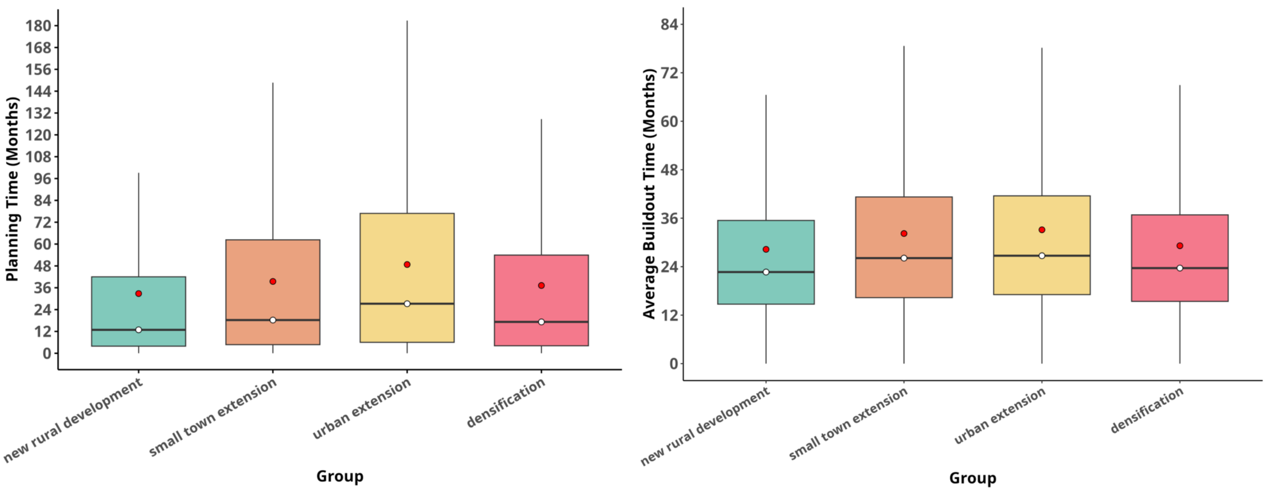
[8]The time taken may not be solely due to the planning system, as some filings, such as discharge of conditions, are also determined by developers.

Figure 7: Average Project Size by Classification



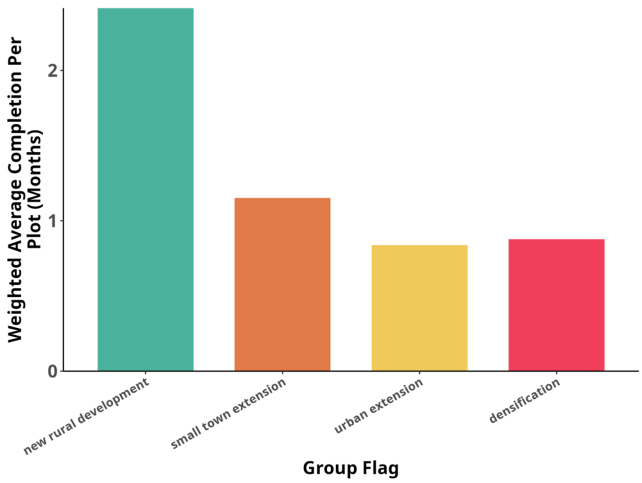
Note: The figure plots the average project size according to the development project's classification. Project size is calculated by summing the number of distinct builds within a project. Data on builds is collected from NHBC. For details on how projects are classified refer to Appendix section C.1.

Figure 8: Average Planning and Build out Time by Classification



Note: The left and right figures respectively show box plots of planning time and average build out time in Great Britain. The project start date is used to calculate the buildout rate when available; otherwise, the first plot start date is used. For the left figure, we expanded each project's corresponding row by the number of units in the project and then plotted the box plot of the planning duration. For the right figure, build out time is at the unit level and defined as the difference between a build's completion time and its corresponding project's start date. The boxes show the Inter-Quartile Range (IQR), which is the range between the 25th percentile (Q1) and the 75th percentile (Q3). The white and red dots are the median and mean respectively. Whiskers extend to the smallest and largest observations in our data set that are within $[Q_1 - 1.5\times\text{IQR}, Q_3 + 1.5\times\text{IQR}]$; we also omit outliers for clarity. Data on builds is from NHBC while planning data is constructed as described in Appendix section B.3. For further details on the construction of a build's completion date, please refer to Appendix section B.2.

Figure 9: Weighted Average Completion Time Plot by Classification



Note: The figure shows the average completion time per plot weighted by a project's size in Great Britain. Completion time is defined as the difference between the last build's completion date and its corresponding project's start date. For each project, the completion time was divided by the project size and then the weighted mean (by size) was calculated across projects within the same classification. Build data is from NHBC. For details on project classification and build's completion date construction, please refer to Appendix section C.1 and Appendix section B.2 respectively.

Finally, Figure 9 also reports a third measure: the *average completion time per unit*, calculated as total project duration divided by the number of units. This again shows little evidence of efficiency gains from larger urban extensions relative to densification. Rural developments additionally stand out as being particularly slow across all measures.

3 How to measure housing demand

As described above, sites for new home delivery are determined first through the central government's targets by local authority and then set out by local plans within LA. Our aim is to provide information on hyper-local housing demand that can help place new homes in locations that households prefer to reside in.

3.1 Two new measures of housing demand

We leverage novel search data from Rightmove on 20 billion housing searches in both the buying and letting market, combined with the full history of rental and sales listings, to introduce two new measures of housing demand relative to housing availability at the OA level: the housing gap and housing tightness.

The search data includes anonymised individual user identifiers, which in turn allows us to allocate each user across all the areas they search in ensuring their weight sums to one and therefore every user is counted only once. We perform several cleaning steps including ensuring that our search data do not contain bots. For more details on this and the calculation please see Appendix B.1. We combine this data with several other sources including planning designations, historical new build supply, planning statistics and details on pre-existing villages, towns and cities in the UK. For details on our data and sample construction, see Appendix B.

We define the following two measures:

1. The **housing gap** is constructed as the difference between demand (i.e. the number of searchers) in an area and available listings in 2024. It essentially represents the excess demand in an area - the larger the housing gap is, the greater the excess of searches relative to listings and the more housing would have to be added to meet demand.
2. **Housing tightness** is constructed as the ratio of demand (i.e. number of searches) over the available listings in an area in 2024 - it therefore captures searches per listing. Intuitively, if searches per listing are high, as many people are viewing a listing, the harder or slower it is to secure a home.

Another way to understand the distinction between these two measures is that the housing gap captures how many people are affected by excess demand, while housing tightness captures how hard it is for each person to find a property.

A few examples help illustrate the distinction: In London there might be 1 million searchers and 900 thousand listings — a gap of 100 thousand homes, meaning many people are affected, but each searcher still faces relatively modest competition (around 1.1 searchers per listing), so it may not take long to find a home. In contrast, in Leamington Spa there might be 10 searchers and only 1 listing — a small gap of 9 homes, so few people are affected overall, but each searcher faces very intense competition (10 searchers per listing), meaning it would take much longer to secure a property.

In short, the gap reflects the scale of the shortage, while tightness reflects its intensity.

3.1.1 Advantages of using search data as a measure of demand

Search data has several advantages when it comes to measuring housing demand. First, it allows us to measure demand at the search location level. These can be as spatially fine as single postcodes, but also include neighbourhoods of towns and cities, and can also cover wider searches based on radius based bufferzones of areas. Second, searchers can look for a location even if there are no listings in that area, or no listings meet there other set of search filters, such as number of bedrooms and price. We observe such searches which are unconstrained by the availability of housing. Third, this is a direct measure of housing demand - we do not need to extract it from a different source.

There may also be concerns around noise in the search data. The demand measure necessarily relies on the accuracy of the search data representing legitimate location interest. Because of the way it is constructed, random noise in individual searches will average out and is therefore unlikely to bias our results. However, non-random noise could pose a concern. In particular, two types of search behaviour may limit the use of housing search data as a measure of genuine location interest.

- *Aspirational* searchers—individuals looking for housing in areas that would remain unaffordable even if prices were to fall somewhat due to new supply.

- *Self-censoring* searchers—individuals restricting their searches to areas at the upper end of their current affordability, and who would therefore be underrepresented in areas that would become attainable if prices were to fall modestly.

We now discuss each of these cases in turn.

Aspirational searchers: An example of such behaviour is given by users frequently searching for high-end properties in areas such as Mayfair without serious intent to buy there. To address this, we conduct a series of robustness checks restricting the sample to users who have actively engaged with the platform—such as those who have saved properties or contacted estate agents during their searches—and find that our results remain unchanged. For more details please see section C.4.1 in the appendix.

Self-censoring searchers: An example of such behaviour is someone not searching in an area that they are only marginally unable to afford. While self-censoring in search behaviour can introduce bias, it is not necessarily problematic. In some locations, marginal construction costs are so high that prices would remain above what many households could ever afford, even under large supply expansions. In such places, non-search simply reflects structural affordability limits rather than a behavioural distortion. Put another way, self-censoring to an extent mitigates aspirational searchers.

If individuals fully censor their searches based on prevailing prices, then the observed "excess demand" at each location already reflects affordability constraints. In that case, the implied demand curve at the going price is perfectly flat: the market could absorb additional units at the same price, and supply increases would therefore not lead to lower prices. Put differently, under complete price-based self-censoring, our measure describes how much extra housing could be taken up at current prices. In that world, because prices would not respond to additional supply, there is no bias in our measure (Büchler and Lutz, 2024).

The more concerning bias arises for households near the margin of affordability — those who might have searched in a location if prices were only slightly lower. If individuals never search above their effective budget constraint, we understate latent demand for these areas and the potential reallocation that could occur if prices fell. However, under the more realistic assumption that people search somewhat above and below their budget thresholds, the observed pattern of searches still provides a reasonable approximation to underlying location preferences — and hence to how search behaviour would shift in a counterfactual world with lower housing costs.

We support this assumption with evidence from th search data. In the data, we observe that approximately 50% of searches do not specify a maximum price filter, while there is around a 20% probability of adjusting the maximum price filter in any subsequent search within a session. We also observe that searchers use both broad spatial parameters (e.g., radii of 10 miles, where location-specific prices vary widely) and much narrower geographical parameters (e.g., postcodes). This descriptive evidence suggests that searchers typically begin within a bandwidth around their affordability but refine their search as they proceed. Thus, price-based self-censorship in the search data does not appear so strict as to prevent some degree of spatial substitution if prices were to adjust following new supply.

3.1.2 Comparison to prices as a measure of demand

Measuring demand at very local levels is often constrained by the availability of data at fine spatial scales, or by data that are unbiased by other factors. For example, household location decisions observed through housing transactions capture people's preferences only conditional on the housing stock that is available. If we observe no one moving into a certain area, this could either indicate a lack of demand or instead reflect that there is demand but insufficient supply. Transaction prices can help disentangle these two cases: we would expect prices to be higher in the latter than in the former.

However, prices can also be misleading indicators of excess demand, as they are shaped by property-specific characteristics, local amenities, and geographic or policy-induced constraints.

At a fundamental level, house prices reflect both local amenities and wage levels (Roback, 1982; Rosen, 1979). There is ample evidence that positive income shocks to residents capitalise into land and housing prices——even in lower-cost segments of the market. For example, increases in the minimum wage have been shown to raise rental prices (Yamagishi, 2021), suggesting that improvements in ability-to-pay can directly translate into price adjustments.

Consequently, cities or towns with fewer amenities—and therefore lower average house prices—may nonetheless experience high excess demand relative to their available housing stock. Moreover, house prices are only observable for completed transactions. Low turnover can thus result in small sample sizes and, at best, noisy local estimates.

Finally, local prices also reflect local development costs. These vary substantially across areas due to differences in key inputs such as land prices, labour costs (i.e., local wages), regulatory barriers, and geography (Glaeser et al., 2005; Saiz, 2008). Land costs, in particular, are sensitive to ownership concentration and the potential exercise of market power.

3.2 National distribution of the housing gap and tightness

Figures 10 and 11 show the national rank in housing gap per km^2 and housing tightness respectively for all OAs in Great Britain. Darker colors indicate a higher gap or tightness. Both maps demonstrate that the highest gaps and tightness are within the major cities, though their ranking differs to some extent.

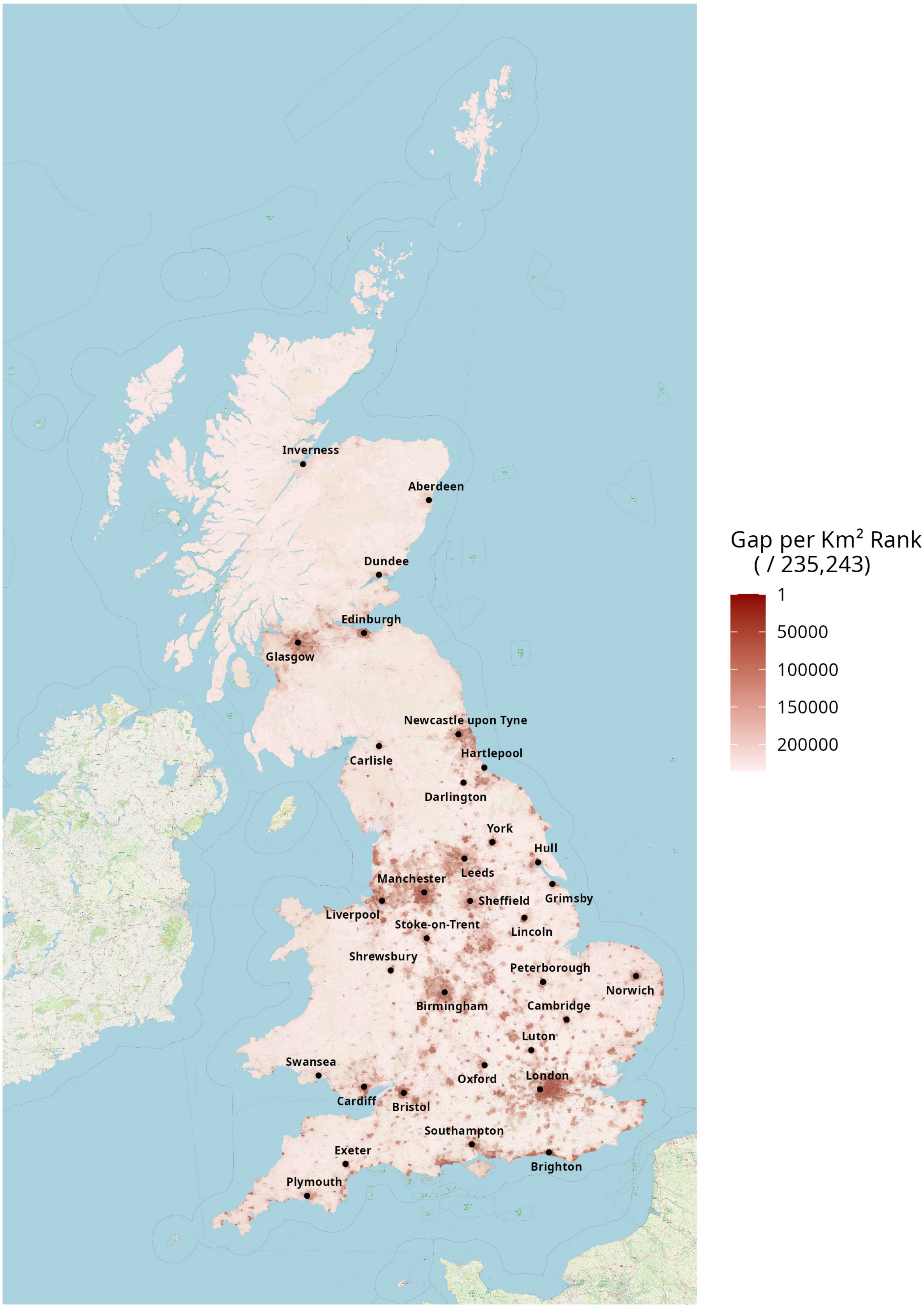
The London boroughs that contain neighborhoods with the highest gaps in Great Britain are Wandsworth, City of London, Kensington & Chelsea, Westminster, Camden, Tower Hamlets, Hammersmith and Fulham, Islington, Hackney and Newham. Additionally, areas of Bristol, Manchester, Salford, Edinburgh, Portsmouth, and Sheffield also rank in the top half percentile of OAs with the highest gap. While the London boroughs are well-known to lack housing and are ranked highly in terms of their housing gap according to our measure, the area around Manchester as well as Bristol are also noteworthy.

While big cities stand out in Figure 10, the map also shows a large degree of heterogeneity in where people want to live. This is reflective of segmented housing markets across Great Britain as well as strong preferences for home location and high moving costs (Venator, 2020; Costa et al., 2024). Most people tend to stay close to their place of birth, evidenced for example by a median moving distance of only 3.2 miles in the UK (Lomax, 2021), and 65% of moves in the US taking place within a county (Frost, 2020).

On the other hand, areas with the highest housing tightness are less concentrated. Cities in Scotland stand out as having areas with the highest housing tightness, among them Edinburgh, Glasgow and Dundee. Additionally, areas of Bristol, Cambridge, Lancaster, Liverpool, Manchester, Hull, Portsmouth, Sheffield, Swansea, Cardiff and York are among the cities with neighborhoods that experience high housing tightness. All of these cities contain output areas in the top 2% of housing tightness. Remarkably, none of the London boroughs show up in the top 2% of housing tightness, highlighting the much higher turnover of new units in London. It is important to reiterate here that the housing gap and housing tightness capture different dimensions of housing need. Therefore, housing tightness can be high in locations which have a low number of listings (not always large towns and cities) even while the housing gap, i.e. the excess demand for housing relative to available supply, can be low in these areas. This is particularly visible in rural areas around the Scottish Highlands, and some coastal areas.

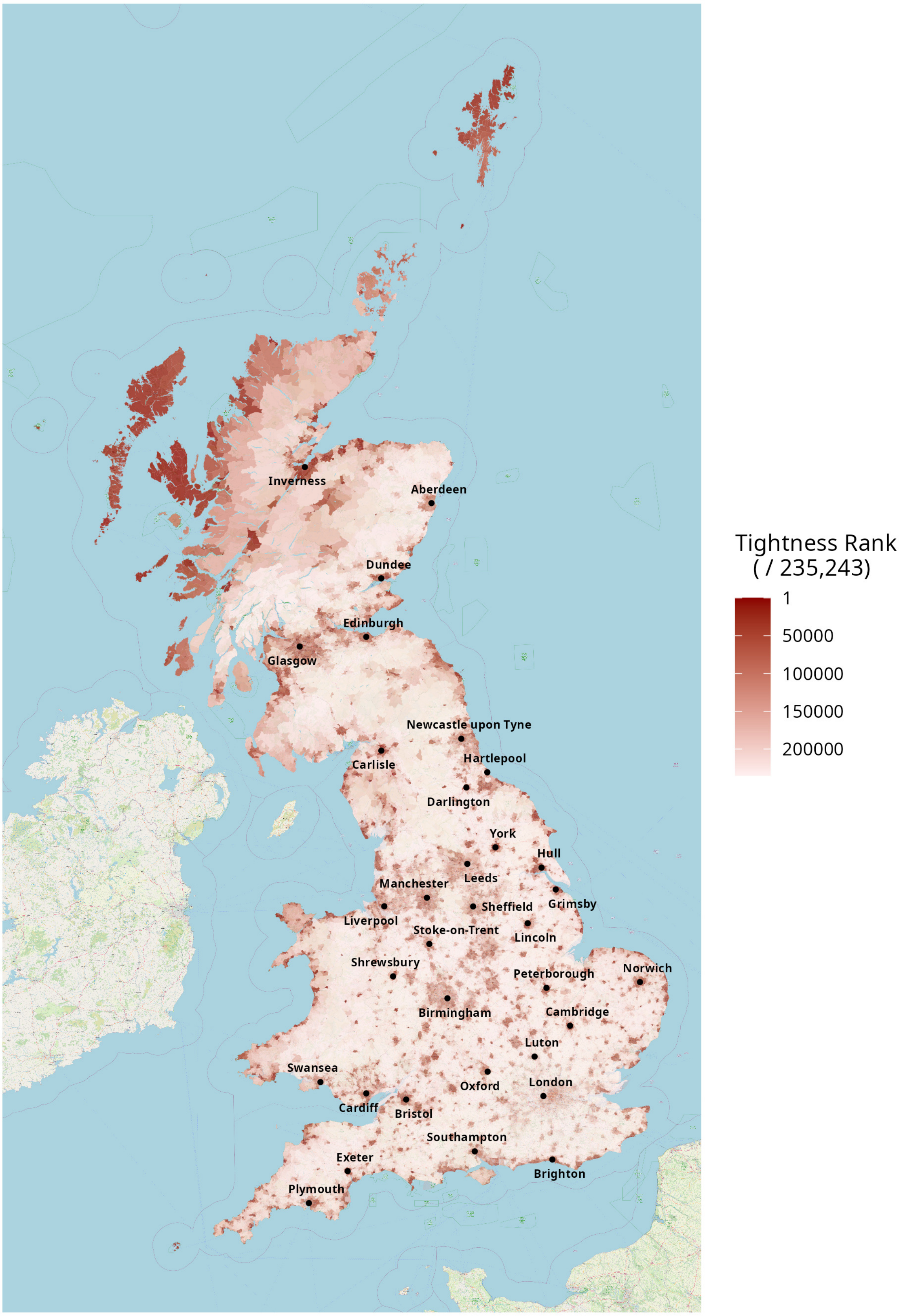
The comparison between these two sets of areas highlights that there is separate variation in housing tightness and the housing gap and that – while some cities have both a very large housing gap and a tight housing market, suggesting long wait times for searchers until they find a match, there are also areas that have a high gap and low tightness and vice versa. Areas with a high gap and low tightness are likely to be housing markets that see a rapid turnover: Since new listings appear regularly, searchers are able to find units faster even though there may be significant excess demand relative to the available listings. On the other hand areas with high tightness imply long search times even in areas where the gap may not be large. Essentially, areas with a high gap require many new units to meet demand, but areas with a high housing tightness may need them urgently to improve people's welfare by lowering wait times.

Figure 10: Map of Gap per km^2 Rank



Note: The above map shows the gap per km^2 rank at the OA level for Great Britain. The gap per km^2 variable is constructed from Rightmove data. For details on the construction of the variable, please refer to Appendix C.3.

Figure 11: Map of Tightness Rank



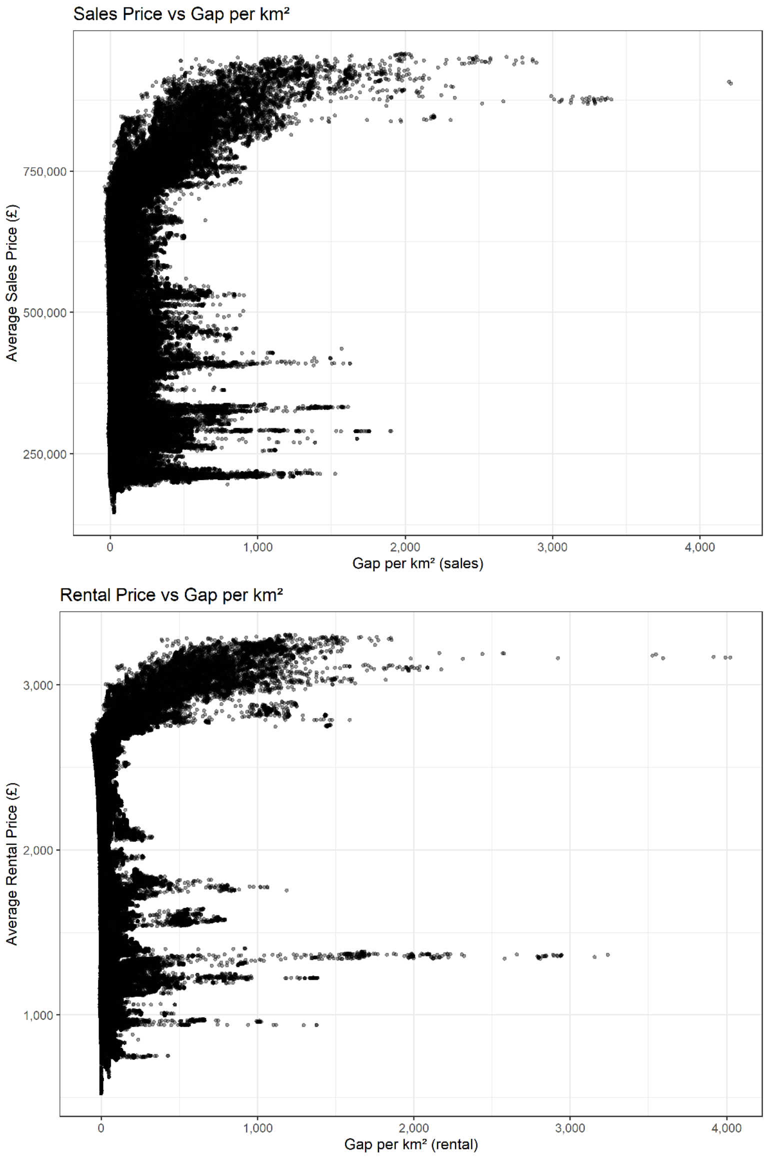
Note: The above map shows the tightness rank at the OA level for Great Britain. The variable is constructed from Rightmove data. For details on the construction of the variable, please refer to Appendix C.3.

3.3 Validating and understanding demand measures

As these are novel measures of housing demand, we now provide further context for what aspects of demand these measures capture through further analysis.

First, we show in Figure 12 below the correlation between the housing gap at the OA level for the sales market and average sales prices above as well as the housing gap at the OA level for the rental market against average rental prices below. The patterns of correlation are similar across both and reveal that – while there certainly is a positive correlation between higher housing gaps and prices, this correlation is most pronounced for very high house prices (roughly sales prices above GBP 725,000 and rental prices above GBP 2,500). For prices below these thresholds, the correlation between the housing gap and prices is much lower, indicating as we discuss above, that excess demand relative to the existing housing stock can be high even in areas with low average housing costs. These aspects of excess housing demand are missed if demand for housing is measured solely based on price. Appendix Figure A.2 shows that the housing gap is much more strongly correlated with sales and rental prices *within* local authorities whereas Figure 12 compares the housing gap and prices *across* LAs.

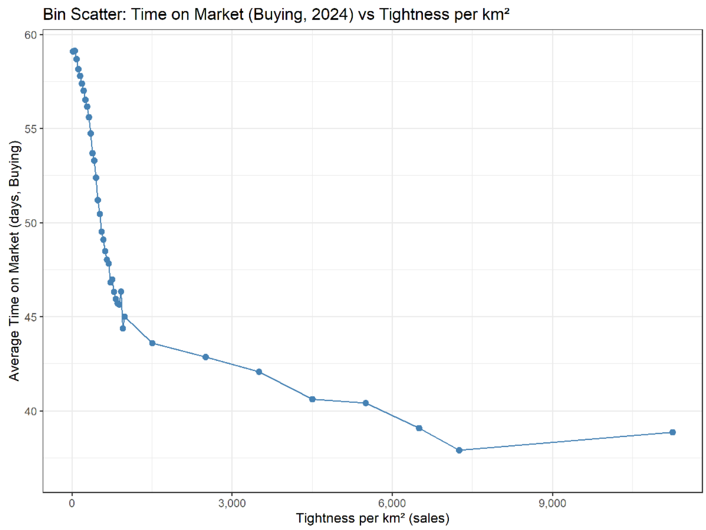
Figure 12: Relationship between average sales/rental prices and the housing gap per square kilometer



Note: The above figures show the scatter plot at the OA level of the sales and rental prices against the gap per km^2. Price data are collected from Rightmove, which is also used to construct the gap per km^2. For details regarding the construction of the gap, please refer to Appendix section C.3.

The differences in the rental and sales markets are important to explore but beyond the scope of this report. For the purposes of this report, we calculate an overall housing gap that adds the gap in sales and rental markets together to determine overall housing demand. However, we note that Figure 12 shows that there is a high degree of correlation between between sales (top panel) and rental gaps (bottom panel) and their respective prices implying that conclusions for the separate markets are unlikely to differ widely on average.

Figure 13: Housing Tightness and Time on the Market



Note: The figure above shows the binned scatter plot of the average time on the market against the tightness per km^2 in the buying market. Both variables are constructed using Rightmove data. For details regarding the construction of the tightness variable, please refer to Appendix section C.3.

Figure 13 supports the interpretation of areas with high housing tightness laid out above as areas with high average wait times for searchers. The flip side of such markets with high demand per unit is that sellers or landlords are able to find matches relatively quickly and units are not on the market for long. The figure shows exactly this pattern for the sales market, indicating that the average time on the market for units more than halves from nearly 60 days in OAs with the lowest housing tightness to between 20 and 30 days for OAs with the highest housing tightness.

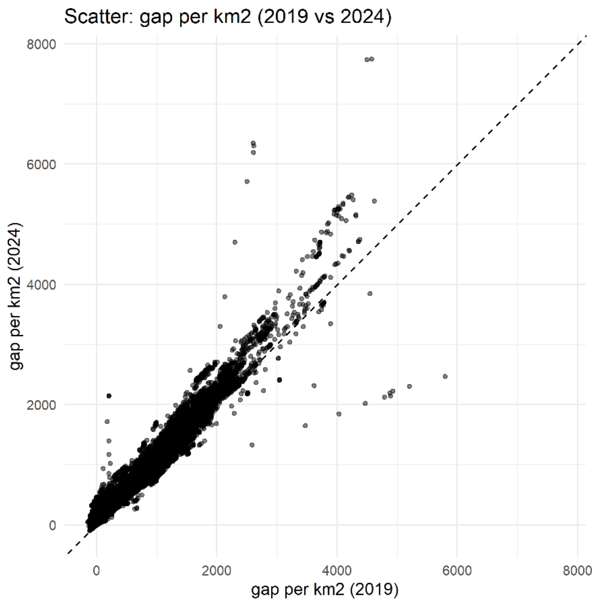
Second, one might worry that the housing gap is not a reliable measure of demand and could be driven by noise such as – potentially temporary – shocks in the housing market by area at a certain point in time. To assuage such valid concerns we also calculate the housing gap with 2019 Rightmove search and listings data. We then plot the 2024 and 2019 housing gaps against each other in Figure 14. The dashed line represents the 45 degree line – if the housing gap remained identical all points would lie along this line. The figure shows that the 2019 and 2024 gap are remarkably similar, despite the intermittent years including the COVID pandemic during which housing demand was severely affected. We are therefore confident that the housing gap is a stable measure of demand for the present decade and potentially beyond.

A third point we explore that is key for the interpretation of our demand measures is how housing search responds to the construction of new homes. In order to use housing searches to make policy recommendations of where to build, it must be the case that searches measure *unconstrained* preferences. This assumption would fail if searches responded to new housing supply. As an example, if the gap is shown to be very low in locations being considered for rural development and small town extensions, but this is as a result of people knowing there is not housing available there, then it may be perfectly reasonable to develop housing in these areas.

To test whether this is the case, we use detailed information on the precise location of new build developments of at least 20 units from 2019 onwards, and examine how searches respond to the presence of these new developments while using areas that did not experience developments as controls. We consider homes that appear in the sales market, and examine how searches evolve in the weeks before, during and after beginning of the construction of the first plot on a site. See Appendix C.4.2 for further details on the calculation. Figure 15 plots the results of this analysis.

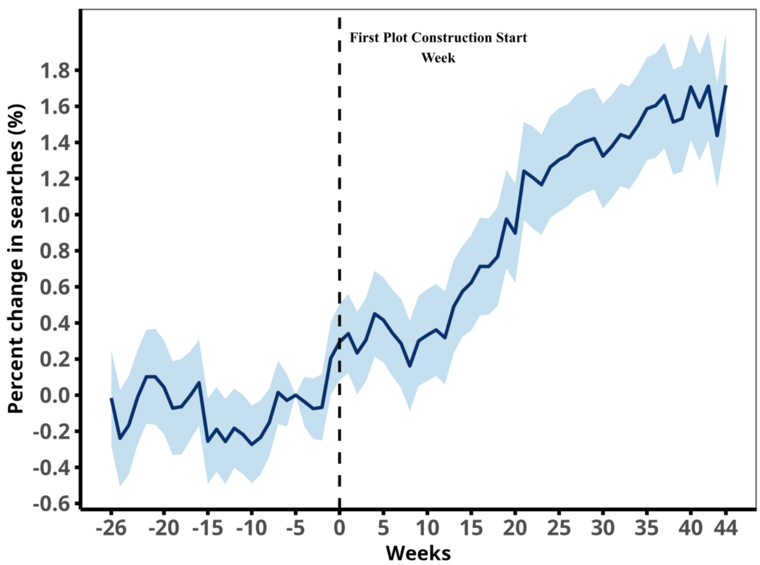
The figure shows that while the number of searches does increase, the change is very small and precisely estimated. Areas with new housing experience searches increasing and leveling off at 1.8% a year after the build

Figure 14: Housing gap per square kilometer in 2019 vs 2024



Note: The above figure shows the scatter plot at the OA level of the gap per km^2 in 2019 vs 2024. For details regarding the construction of the variable, please refer to Appendix section C.3.

Figure 15: Event Study



Note: The figure shows the response of searches in percentage changes across a window of -26 to 44 weeks relative to the shock. The effects are normalised to week -5 relative to the shock. Search and builds data used are collected from Righmove and NHBC respectively. For details on the data set construction and regression specification refer to Appendix section C.4.2.

starts. As the average number of searches is approximately 80 across all areas that experience new builds, this suggests the raw total number of searches only increased by around 1.5 searches. As the the average site in our sample has 106 built new units, the increase in searchers relative to new units is negligible. As such, this demonstrates that our measure of demand is remarkably stable in response to even large housing supply shocks. Appendix Figure A.4 shows that this negligible change holds across new sites of all four classification types (e.g., new rural development, etc.). It is also important to note that prices and rents might fall in these locations as a result of the large supply shock which could induce further searches. Figure 15 shows that additional searches are small despite this potential channel. We view this as further evidence that self-censorship is not a major issue in our data and that the searches are reflecting actual location preferences.

There are two main takeaways from this fact. First, we view these results as a validation of our demand measures. Housing searches capture people's stable desire to live in certain areas which cannot be influenced by the available supply of units. We are therefore comfortable using our measures to discuss where new housing should be built based on people's demand. Second, the stability of demand suggests that new homes built in areas that do not correspond to households' preferences may take longer for developers to sell, and plots may remain undeveloped for longer periods as a result. We discuss this further below in Section 4.3.

Overall, this section has highlighted that the housing gap and housing tightness are relatively stable, capture people's unchanging preferences for location that go beyond the affordability of an area and result in a high degree of variation across Great Britain in terms of where there are large excesses in demand relative to supply (gap) and where housing is demanded urgently (tightness). We next explore what we can learn about demand from variation in the housing gap within local authorities and how the gap be used to inform housing targets and building decisions.

4 Using the new demand measures to inform housebuilding policy

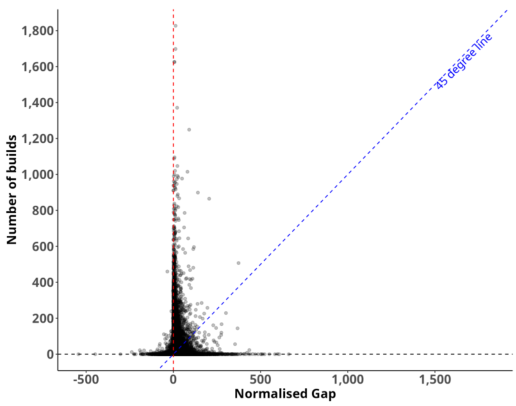
In this section, we discuss how we can use the housing gap and tightness to inform housebuilding decisions. First, we focus on hyper-local, within-LA variation in the housing-gap. Below, in Sections 4.2 and 4.3, we discuss takeaways within-LA, where housebuilding goals are set out in local plans. These local results make the best use of our granular data to discuss local housing gaps. Second, we aggregate up to LA wide targets (Section 4.4). We use our demand measures to calculate housing targets at the LA level that we compare to the government's targets discussed.

4.1 The importance of within-LA decisions

Much of the political debate around housing delivery focuses on nationally-set housing targets at the LA level. However, housing delivery is ultimately a local responsibility and must cater to local needs and conditions. We will show below in section 4.4 that, at LA level, the government's targets align more or less with the targets we can calculate based on our housing gap. However, within LA there may be large degrees of variation in the demand for housing, suggesting that after setting housing targets, determining where exactly to build within an area covered by a local plan is arguably the more important aspect of housing delivery decision making.

To assess how much variation in the housing gap occurs across LAs versus within them, we decompose the housing gap variance using a one-way ANOVA. Only around 3.9% of the total variance in gap is explained by differences in mean values between LAs, while the remaining 96.1% of the variance arises from differences within LAs and across OAs. This indicates that the vast majority of variation in gap reflects local differences inside LAs rather than systematic differences between LAs and highlights the important role of allocating locations for new development sites within Local Plans.

Figure 16: Scatter Plot of Number of Builds vs Normalised Gap



Note: The figure shows the scatter plot of the number of builds from 2010 to 2020 against the normalised 2014 gap; the total gap is normalised to the total number of builds starting between 2010 to 2020. Both variables are the OA level, and sites are matched to OAs based on highest intersection proportion of a site. Data on builds is collected from NHBC and the gap is constructed using Rightmove data. For details regarding the gap construction, refer to Appendix section C.3.

To underscore the importance of local variation, we again consider in which OAs building has taken place in the years between 2010 and 2020. In Figure 16, we correlate these numbers against the housing gap in 2014 (for calculation details please see Appendix C.3) normalised to the total number of builds in this time period. If building took place based on the excess demand measure or predicted future demand, we would expect a one for one correlation, i.e. all points should lie along the 45 degree line in the Figure. This is far from the case, in fact, much of the new builds have taken place in areas with a very low excess demand, some of them (albeit a small fraction) have even been built in areas with an oversupply of housing relative to demand, where the normalised gap is negative.

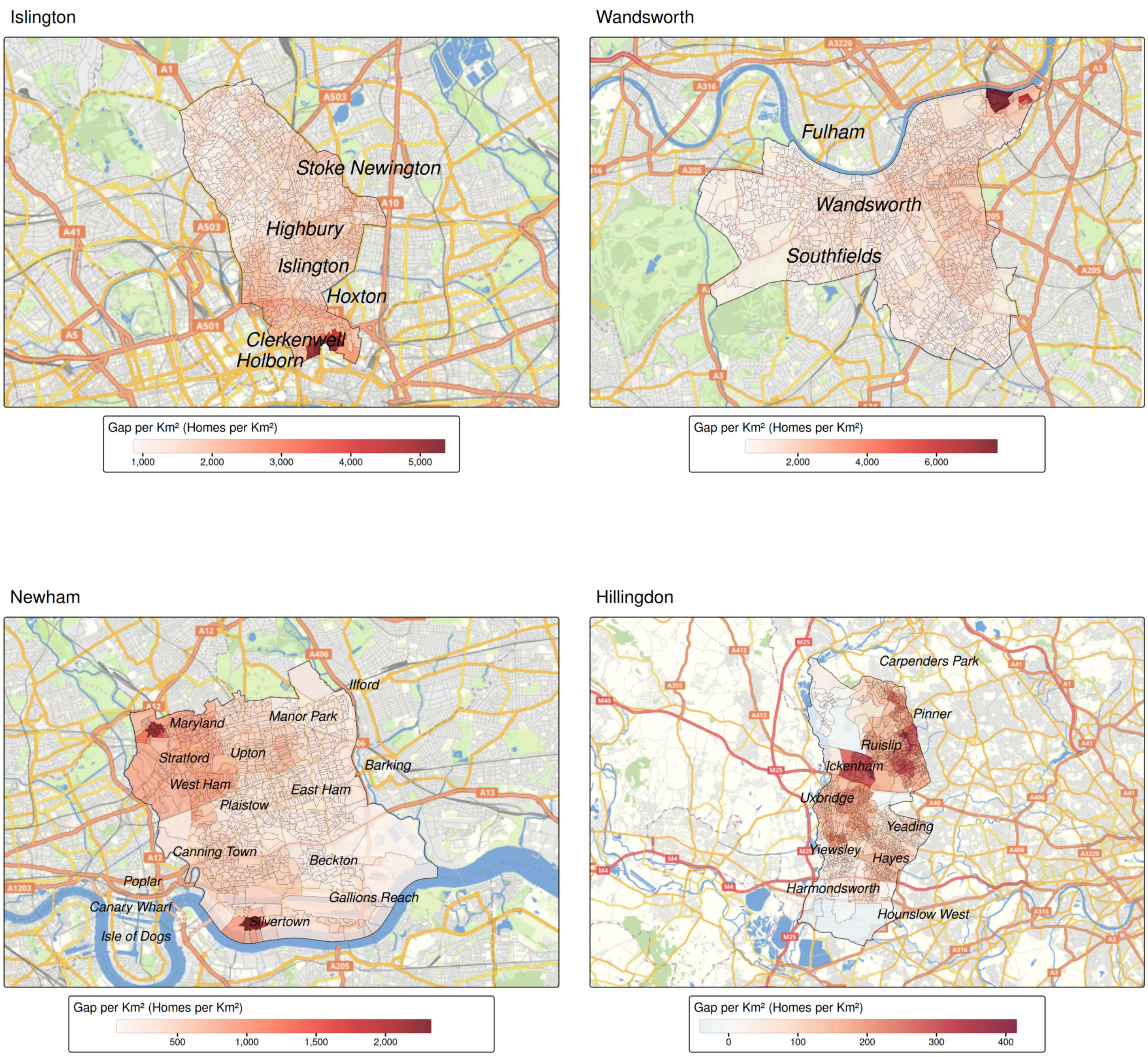
Therefore, localised building within local authorities is far from the neighbourhoods in which excess demand is highest. Next, we look under the hood of this local variation, providing examples of how the mapping tool can inform within-LA building as well as drawing general conclusions about how new house building could take place to lower the housing gap.

4.2 Where to build at the neighbourhood level

We present examples of within-LA housing gaps that shed light on the importance of building in particular locations, and relate results back to our discussion of different residential build classifications, which are based on policy debates on how to deliver 1.5 million new homes (Kulka and Datta, 2024).

Beginning with London, Figure 17 maps the housing gap for four boroughs, each with different scaling to reflect the local Output Area gaps. Above we discussed that certain London boroughs have the highest housing gaps in the country, among them Islington, Wandsworth and Newham. Particularly, Wandsworth and Islington have high gaps on average across the borough, but also display high degrees of variation within the borough. In Islington for example, neighbourhoods closer to the City of London display the highest housing gaps within the borough. In Wandsworth, it's areas close to the Thames some of which are commercial areas, highlighting an increased demand for mixed-use housing[9] while in Newham two areas in the East Village and Docklands neighbourhoods stand out.

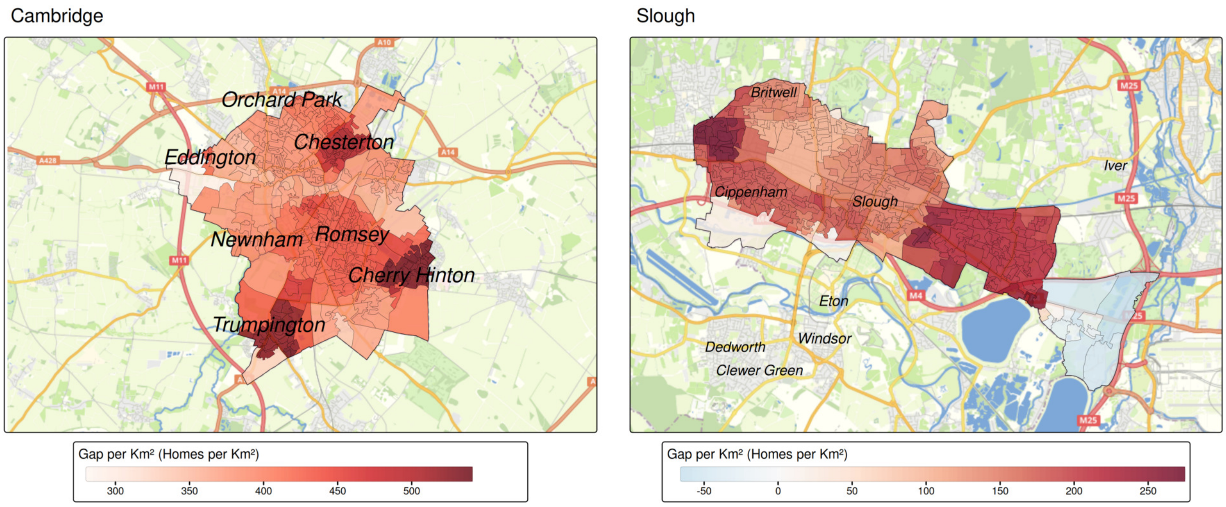
Figure 17: Variation in the housing gap in London boroughs



Note: The above maps show the gap per km^2 at the OA level in Islington, Wandsworth, Newham, and Hillingdon. The gap per km^2 variable is constructed from Rightmove data. For details on the construction of the variable, please refer to Appendix C.3.

[9] Appendix Figure A.1 shows that some areas with high housing gap are likely commercial areas with a lower residential density.

Figure 18: Variation in the housing gap: Cambridge and Slough



Note: The above maps show the gap per km^2 at the OA level in Cambridge and Slough. The gap per km^2 variable is constructed from Rightmove data. For details on the construction of the variable, please refer to Appendix C.3.

At the same time, neighbourhoods within other London boroughs such as Hillingdon (Figure 17), Hounslow, and Croydon lie in the bottom percent of the housing gap, e.g. have some of the lowest demand for housing relative to listings in the country. Parts of these boroughs even see an oversupply of housing relative to listings, i.e. a negative housing gap. Hillingdon is also an example of a LA which displays much less variation in the housing gap within the borough, suggesting there are no clear priority locations for newbuilds. Cambridge (Figure 18) similarly has little variation in its housing gap, but the gap is overall at a higher level than Hillingdon's. Slough (Figure 18) is similar but has an area of notable negative housing gaps – i.e. an excess of available housing relative to demand – in the south-east which is next to Heathrow Airport and therefore impacted by air traffic. In general, OAs with the lowest gaps are often next to unattractive infrastructure such as sewage plants, airports, large road interchanges, or in areas without good access to amenities.

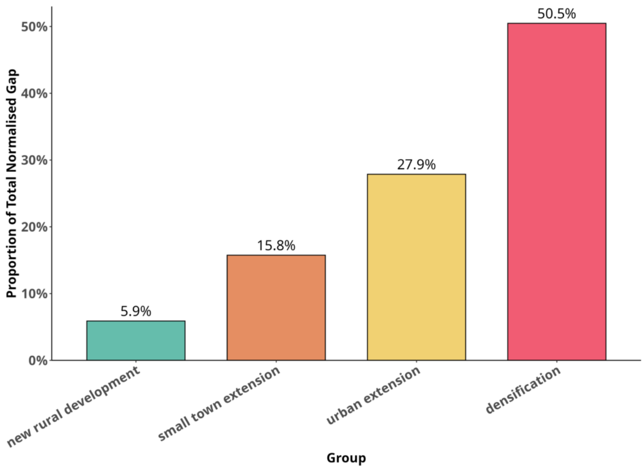
Overall, these maps highlight the tremendous variation that exists within LAs in terms of their housing demand relative to supply. Given the validity of the measure demonstrated in section 3, the figures that can be generated from the mapping tool suggest areas within LAs that could be prioritised for development when local plans are updated. At the same time, they suggest areas where new housing is unlikely to be filled quickly or even at all.

In keeping with our prior discussions, we break down the total housing gap across Great Britain into our four residential building classifications by assigning OAs a classification based on the categories defined in Table C.1. Figure 19 shows that half of the housing gap, 50.5%, is in areas that we would classify as densification were new construction to occur there. In other words, satisfying half of the excess housing demand would require densifying the core of cities, as well as towns with more than 20,000 residents. The inner London boroughs Wandsworth (Figure 17), Islington (Figure 17), and Camden have a particularly high housing gap along with the urban core of Bristol, Manchester, Salford and Edinburgh (Figure 20). These gaps could be satisfied by adding densifying housing units. In practice, this can mean several things such as adding additional floors or units to existing buildings, identifying and building on greyfield sites, or promoting more mixed-use housing.

The other half of the housing gap stems mostly from areas classified as urban extensions (28%) were new housing to be built there, with smaller fractions of excess demand in areas classified as small town extension (16%) and new rural development (6%). Areas with high demand for urban extensions include London's outer-lying boroughs such as Kingston upon Thames, North-West Bristol and South Nottingham (see Figure 20).

London, Bristol, and Nottingham are all encircled by a greenbelt — an urban growth boundary originally intended to permit new residential development only in exceptional cases. Greenbelts have been shown to strongly limit housing supply even while providing some positive amenity effects (Koster, 2024). In the context of new-build classifications, greenbelts are most likely to limit urban and small town extensions. Figure 21 plots

Figure 19: Proportion of Normalised Gap by Classification



Note: The figure shows the proportion of total gap normalised to 1.5 million in Great Britain by classification. The gap is calculated at the OA level, and each OA was classified using the same process applied to build projects from NHBC (see Appendix section B.2). The gap was constructed using Rightmove data. For details regarding gap construction refer to Appendix section C.3.

the housing gap in greenbelt areas around Bristol, Figure 22 does the same for Greater Manchester and Liverpool, and Figure 23 shows the London greenbelt. Strikingly, almost the entire greenbelt area around Greater Manchester and Bristol has a positive housing gap, with demand heavily exceeding available supply. This can be seen clearly for Bristol's southern neighbourhoods as well as southern Salford which lies in the greenbelt around Manchester.

On the other hand, London's greenbelt (Figure 23) shows a somewhat more mixed picture in terms of the housing gap. The areas immediately outside of London's boundaries do have positive excess demand for the most part (for example greenbelt areas bordering the borough of Redbridge), however areas deeper in London's greenbelt often display a low or even negative housing gap.

Finally, in areas without a greenbelt, the demand for urban extensions can also be seen by examining the housing gap in local authorities bordering larger cities and towns. We demonstrate this using the example of the local authority of South Hams (Figure 24) which neighbors the city of Plymouth. The border with Plymouth is highly visible in Western South Hams where the housing gap is visibly higher than in many other areas in the center of the local authority. Once again, our calculation suggests that new housing added on Plymouth's border would likely be bought/rented out fast while new housing added in central South Hams may be slower to match or even remain vacant.

Figure 25 shows two examples of areas with a high demand for small town extensions, i.e. around Dorchester in Dorset and around Saltburn in Redcar & Cleveland. It is important to note that, while Saltburn and Dorchester are small towns, the housing gap is as high or higher as the excess demand in some of the major cities and large towns discussed above. This underlines the fact that our housing demand measure can be used at a highly granular spatial level and – by measuring people's actual demand for housing – paints a more detailed picture of where people want to live than other, more summarised, measures.

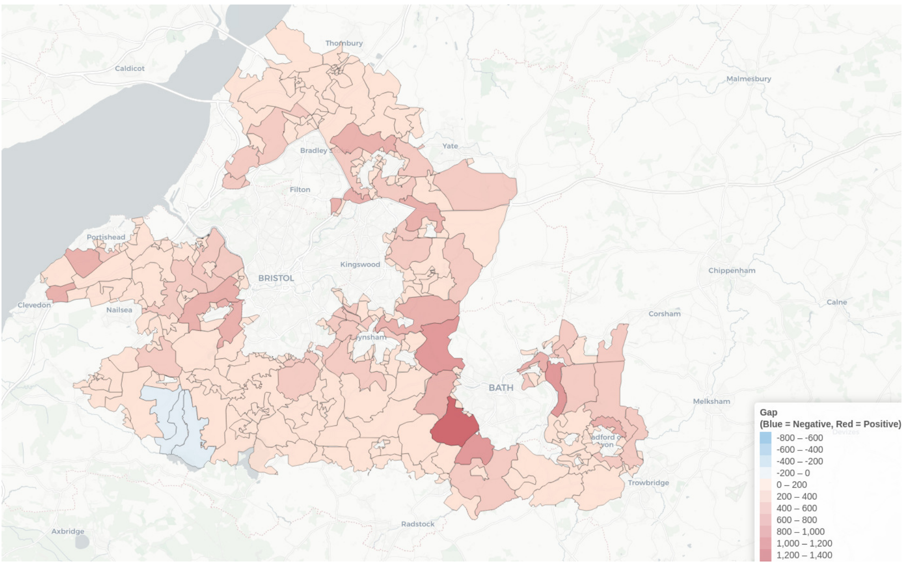
Finally, Figure 26 shows Cornwall and Tendring – two local authorities with the highest demand in the form of new rural developments. In Cornwall, this demand stems from the areas around the village of Mousehole and near Port Isaac and Polperro. In Tendring, the housing gap is high in areas surrounding Clacton & Jaywick sands. We note that Clacton-on-Sea and Mousehole are both popular holiday destinations and therefore, it is possible that the high demand in these areas might be driven by a demand for second or holiday homes.

These maps show how our tool can help local policymakers and other stakeholders identify where housing demand exceeds supply, and where markets are slow to match searchers with homes. Next, we discuss more general policy implications for housing delivery that we can draw from the within-LA data.

Figure 20: Densification

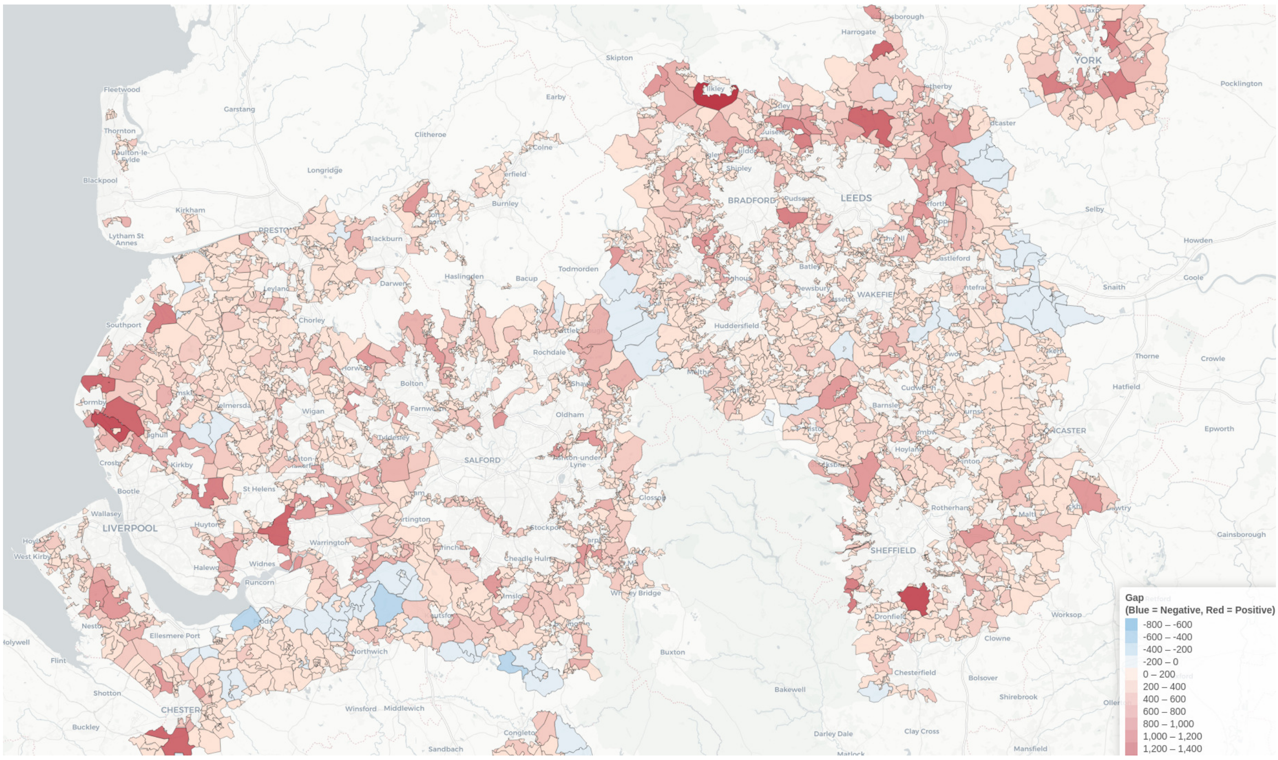
Note: The above maps show the gap per km^2 at the OA level across several LAs. The gap per km^2 variable is constructed from Rightmove data. For details on the construction of the variable, please refer to Appendix C.3.

Figure 21: Bristol Greenbelt



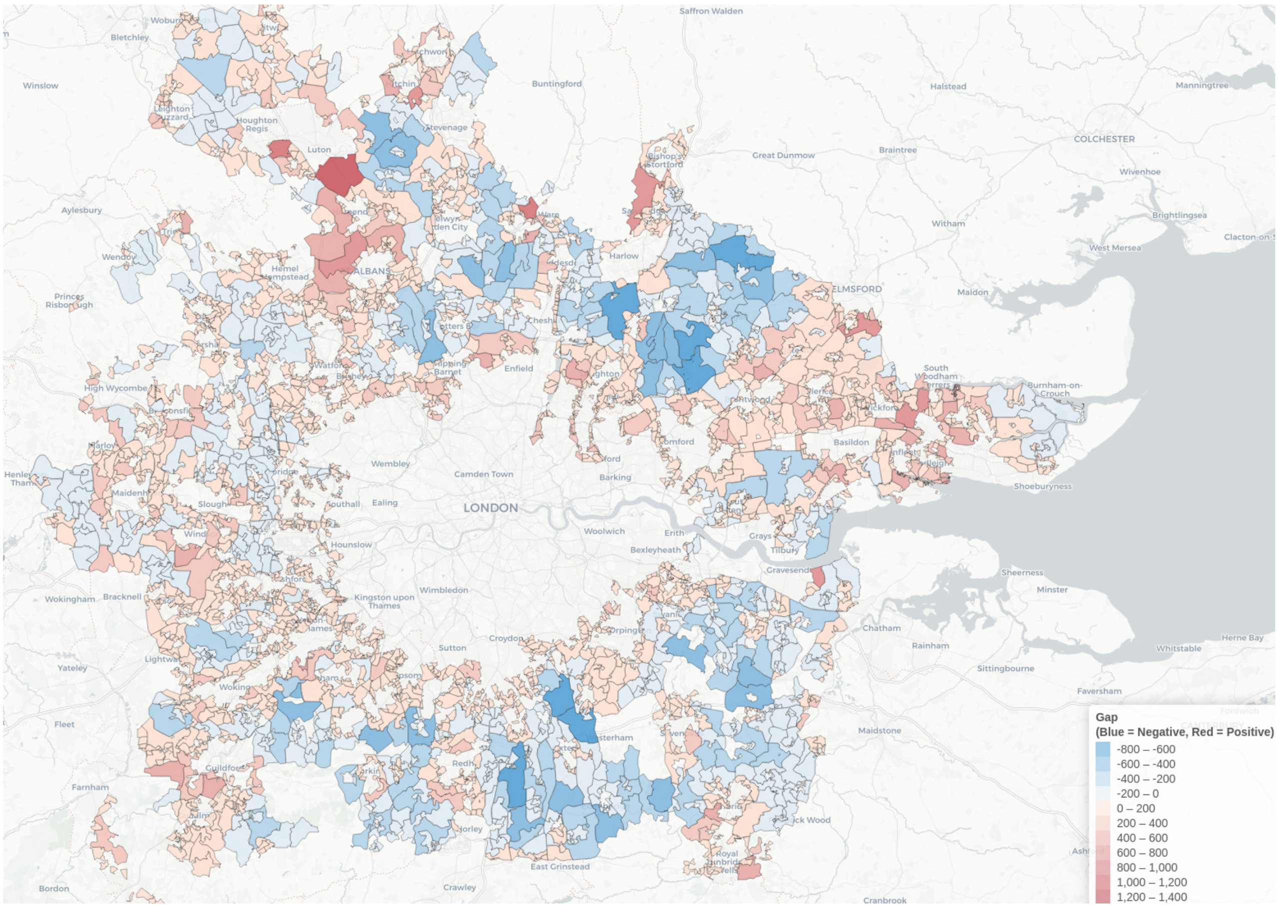
Note: The above map shows the gap at the OA level in Bristol greenbelt. Each OA's gap is mapped to its corresponding (if any) greenbelt region based on highest intersection proportion of the OA. Data on greenbelts' locations is collected from gov.uk and the gap was constructed using Rightmove data. For details on the construction of the gap variable, please refer to Appendix C.3.

Figure 22: North-West Greenbelt



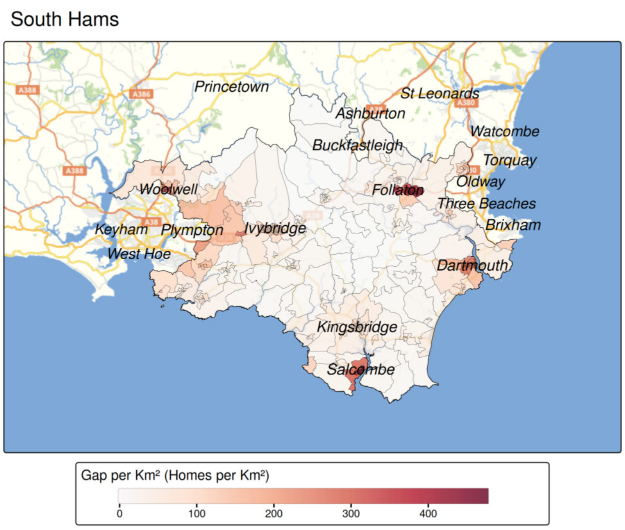
Note: The above map shows the gap at the OA level in North-West greenbelt. Each OA's gap is mapped to its corresponding (if any) greenbelt region based on highest intersection proportion of the OA. Data on greenbelts' locations is collected from gov.uk and the gap was constructed using Rightmove data. For details on the construction of the gap variable, please refer to Appendix C.3.

Figure 23: London Greenbelt



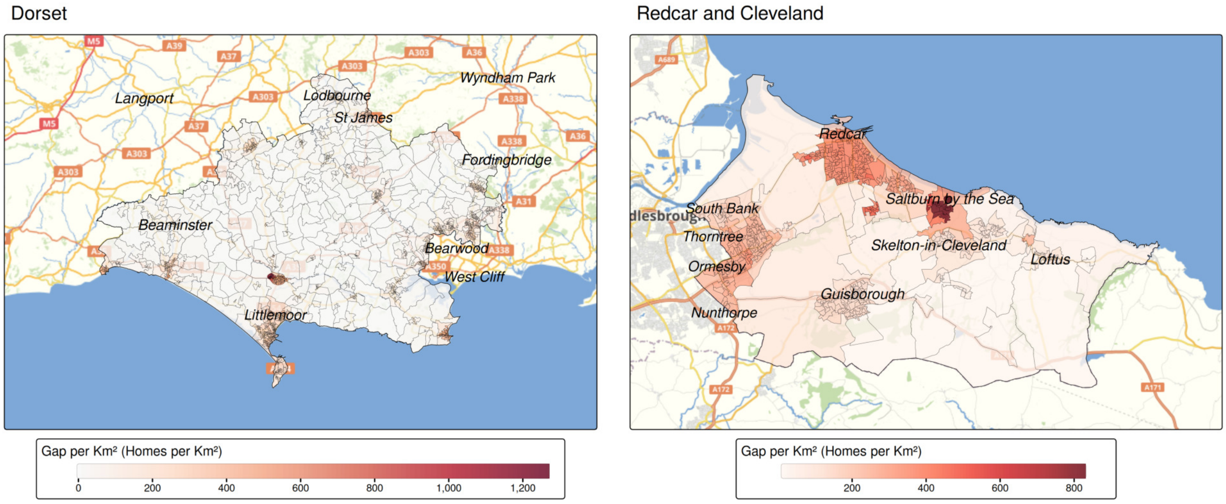
Note: The above map shows the gap at the OA level in London greenbelt. Each OA's gap is mapped to its corresponding (if any) greenbelt region based on highest intersection proportion of the OA. Data on greenbelts' locations is collected from gov.uk and the gap was constructed using Rightmove data. For details on the construction of the gap variable, please refer to Appendix C.3.

Figure 24: Urban Extension: South Hams



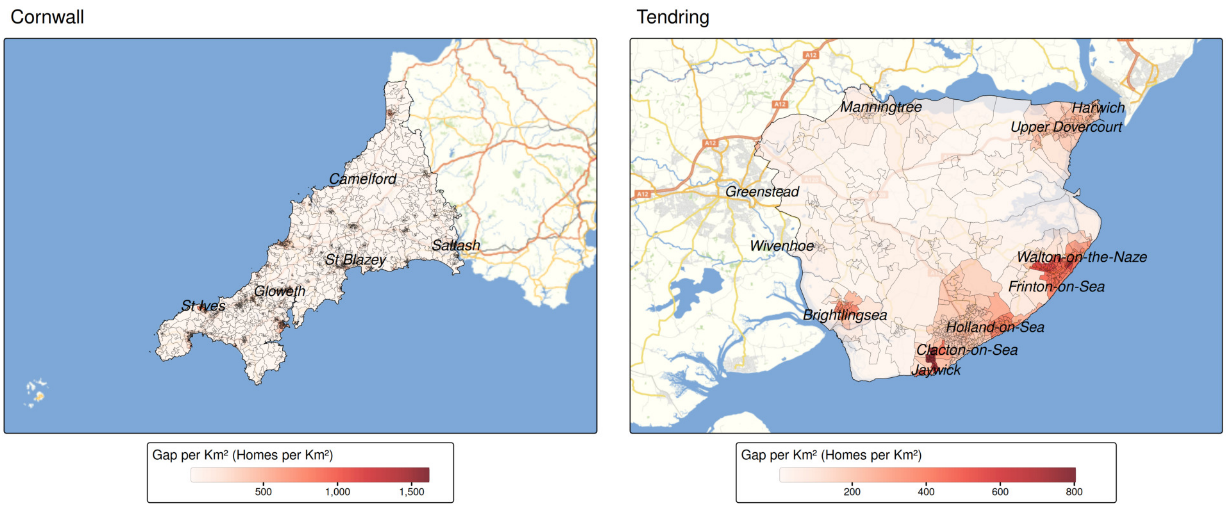
Note: The above map shows the gap per km^2 at the OA level in South Hams. The gap per km^2 variable is constructed from Rightmove data. For details on the construction of the variable, please refer to Appendix C.3.

Figure 25: Small Town Extensions



Note: The above map shows the gap per km^2 at the OA level in Dorset as well as Redcar and Cleveland. The gap per km^2 variable is constructed from Rightmove data. For details on the construction of the variable, please refer to Appendix C.3.

Figure 26: New Rural Developments



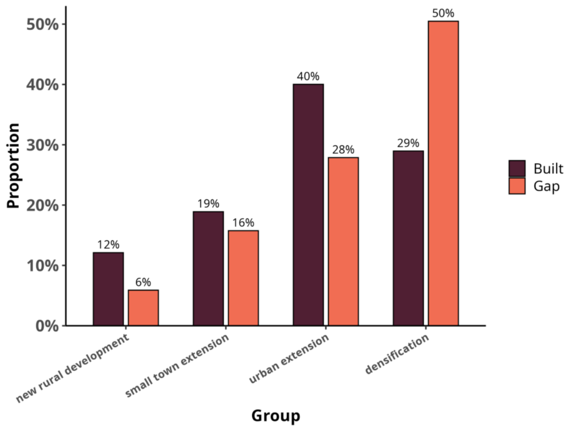
Note: The above map shows the gap per km^2 at the OA level in Cornwall and Tendring. The gap per km^2 variable is constructed from Rightmove data. For details on the construction of the variable, please refer to Appendix C.3.

4.3 Policy conclusions within local authorities

How can our results inform current policy debates around where to build housing? We have already shown (Figure 19) that half of the gap in housing is in areas that would require new housing delivery to be in the form of densification. Our mapping tool allows interested groups and individuals to view where exactly these areas lie. This insight is important as a lot of the current debates centre around the building of new towns and larger town extensions.

Figure 27 makes clear that this debate is also reflected in past building patterns. We show the proportion of new builds from 2010 to 2020 in differently classified areas alongside the proportion of the housing gap in those areas, as it was in 2014.[10] While half of the housing gap lies in areas that required densification, only 30% of new builds since 2010 were built in such areas while 30% of new builds were built in the form of new rural developments and small town extensions and 40% of new builds were built in the form of urban extensions. However, these are not places where the excess demand is highest. In particular, while the fraction of small town extensions built was roughly in line with housing demand in such areas, the fraction of builds as new rural developments was twice that demanded, and even builds in the form of urban extensions were about 40% higher than necessary. Our results emphasise that a debate around identifying and streamlining building on sites in city and town centers is crucial if new homes are to be in areas that households want to live in.

Figure 27: Proportion of Builds vs Gap Across Classifications



Note: The figure shows the proportion of builds from projects that started after 2010 against the proportion of gap in Great Britain by classification. Data on builds is collected from NHBC; the gap is constructed using Rightmove data. For details on the construction of the gap measure refer to Appendix section C.3.

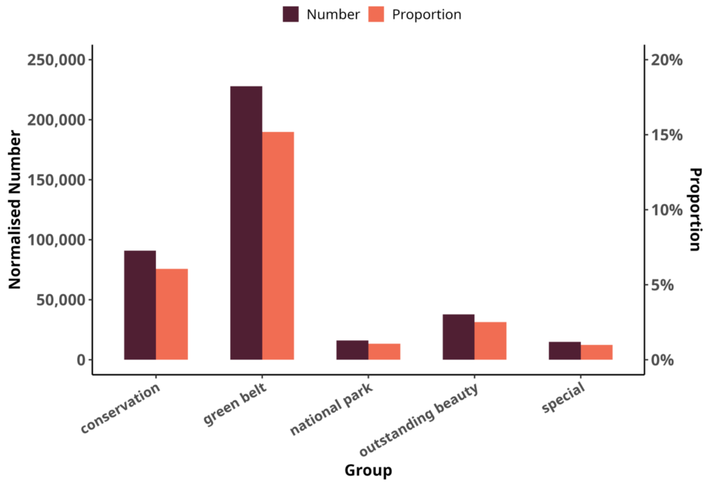
Figure 28 shows the proportion of the total housing gap and the corresponding number of units (where we've normalised the total housing gap to 1.5 million) that lie in various designated land areas on which building is restricted. The largest group is by far greenbelt land with 15% of the gap, and therefore over 200,000 of 1.5 million demanded housing units. Conservation areas – likely reflecting the preferences of local residents (Ahlfeldt et al., 2017) but limiting individual rights to alter or build homes in a specific way – are next, making up 6% of the gap with less than 4% of the total gap on areas that lie in national parks or areas of outstanding beauty.

Desired development on greenbelt areas often requires urban extensions. In addition to the Figures 23, 22 and 21 above that show the housing gap on greenbelt land around London, Bristol, and Manchester, Figure 29 shows the same information for the greenbelts around Oxford and Birmingham. Two things stand out: First, the housing gaps are typically highest in the areas closest to the city which is not surprising as these areas will, at present, have the best employment and amenity access. Second, not all greenbelt areas are in high demand. For example, the eastern part of Oxford's greenbelt has a negative housing gap.

Again, our tool can help fine tune housing policy by pinpointing specific areas in need of housing. Demand on greenbelt areas is quite concentrated in many cases. Figure 30 shows that when aggregating all of England's greenbelt land, allowing construction on only a fifth of this land would satisfy about 70 percent of the excess demand on greenbelts. Allowing construction on only a tenth of greenbelt land with the highest housing gap would already satisfy around 55 percent percent of the excess demand on greenbelts, amounting to more than

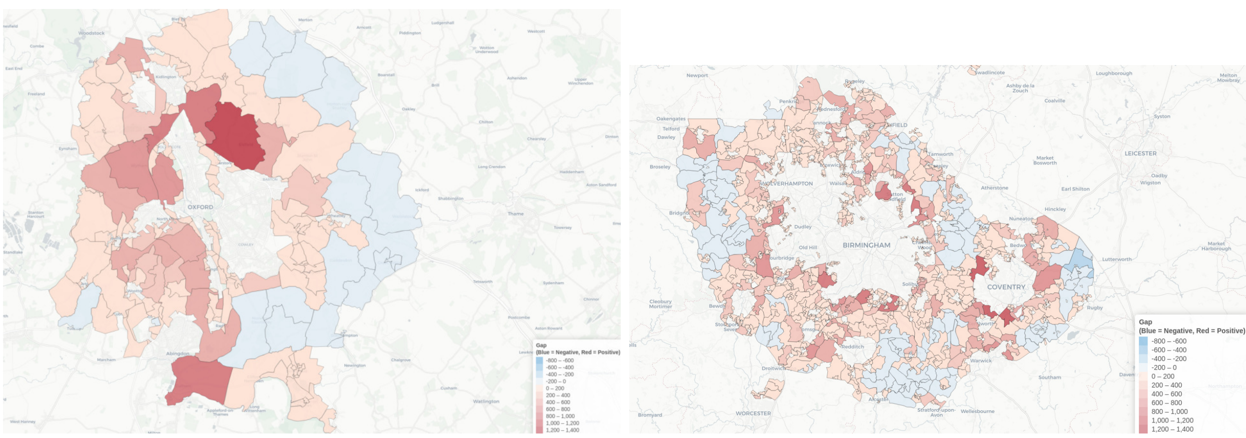
[10]See Appendix C.3.4 for this calculation.

Figure 28: Proportion and Number of Gap by Planning Designation



Note: The figure shows the number and proportion of the normalised gap by designation across England. The total gap across England is normalised to 1.5 million, and the total used when calculating proportions is also 1.5 million. Each OA is matched to a planning designation using the highest intersection proportion of that OA. If an OA is fully covered by more than 1 designation, it was assigned the designation with the largest area. Planning designations data is collected from gov.uk, and the gap is constructed using Rightmove data. For details on the construction of the gap measure, refer to Appendix section C.3.

Figure 29: Maps of Gap across Greenbelts Regions in Oxford and Birmingham

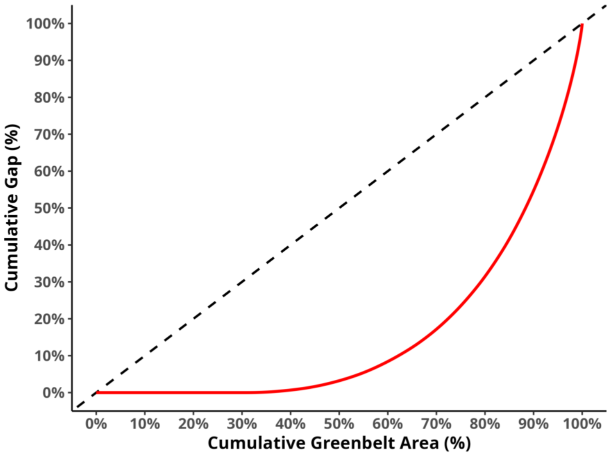


Note: The left and right figures show the gap at the OA level across greenbelt regions in Oxford and Birmingham respectively. Each OA's gap is mapped to its corresponding (if any) greenbelt region based on highest intersection proportion of the OA. Data on greenbelts' locations is collected from gov.uk and the gap was constructed using Rightmove data. For details on the gap's construction, refer to Appendix section C.3.

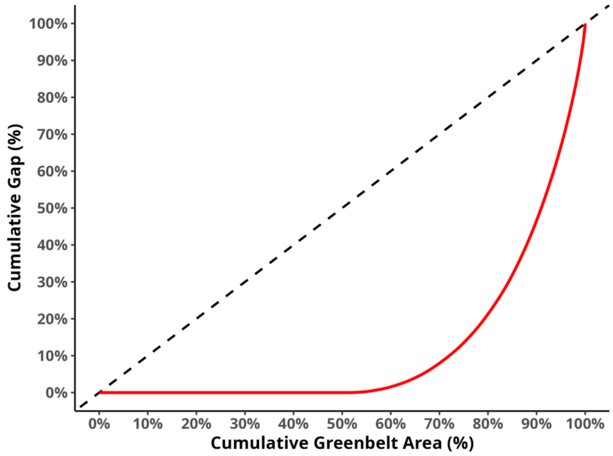
8% of the total gap. Much of the greenbelt land with high excess demand lies in the North-West greenbelt around Manchester and Liverpool and the South-West Yorkshire greenbelt around Leeds and Sheffield. Figure 30 shows that 52% of London's greenbelt has no excess demand at all, while the same is true for only 19% of the South and West Yorkshire greenbelt and 15% of the North-West greenbelt.

Figure 30: Excess Housing Demand on Greenbelts

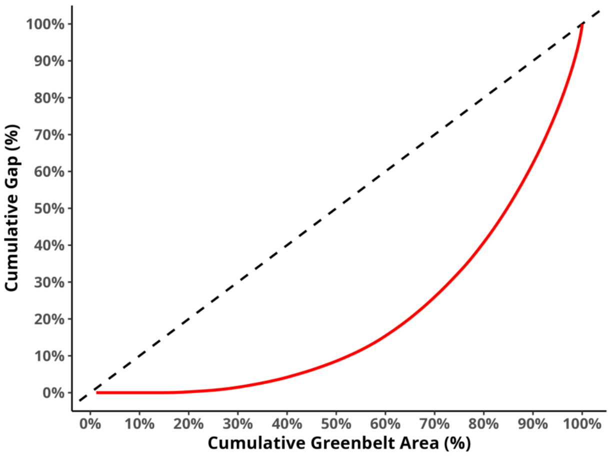
(a) All Greenbelts



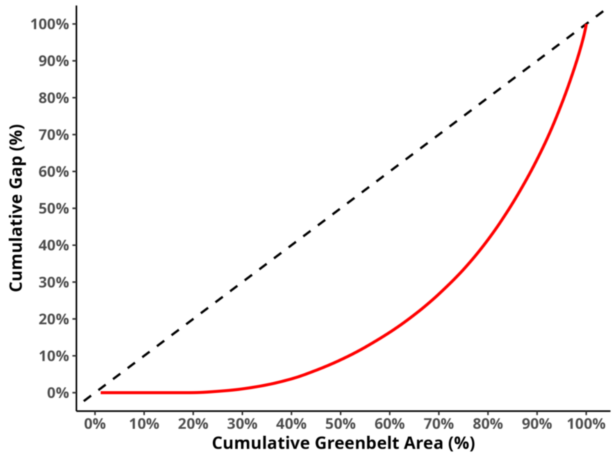
(b) London Greenbelt



(c) North-West Greenbelt



(d) South and West Yorkshire Greenbelt

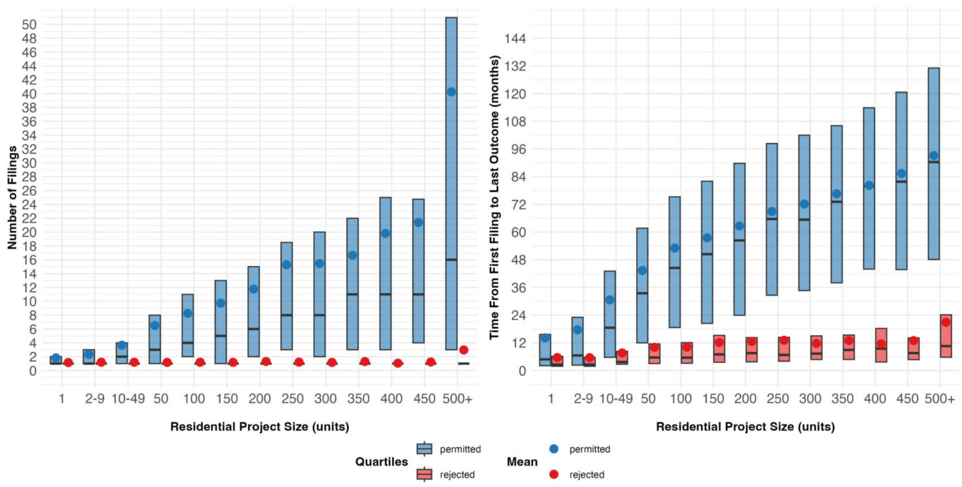


Note: The above figure shows the cumulative percentage of housing gap against the cumulative percentage of greenbelt area across all greenbelts, London, North West, and South and West Yorkshire. To create the North West greenbelt, we combine the following greenbelts: Carnforth, Lancaster, Morecambe, Merseyside, Greater Manchester, and Blackpool (CPRE, 2018). To construct this figure, we intersect OAs with greenbelts, and use the intersecting OAs' areas as the greenbelt area. Negative gaps at the OA level are set to zero. The data is then ordered by gap per km^2, and cumulative sums and percentages of both gap and OA area are calculated. The data on gap is constructed using Rightmove data, and data on greenbelts is collected from planning.gov.uk. For further details on the construction of the gap, please refer to section C.3.

Instead of, for example, allowing building on the entirety of all greenbelts, the housing gap clearly shows where building would help satisfy demand and where it would lead to residential homes on green space that would likely remain empty. Areas with high excess demand are typically located close to the main cities on the inner side of their respective greenbelts. These are areas that should be of priority to policymakers considering which greenbelt land to deregulate. Examples of such high-impact areas include areas north of the A12 near Chadwell Heath and Romford in the East London greenbelt, south and west of Stretford in Manchester in the North-West greenbelt, and areas just west of the river Avon in the Bristol greenbelt.

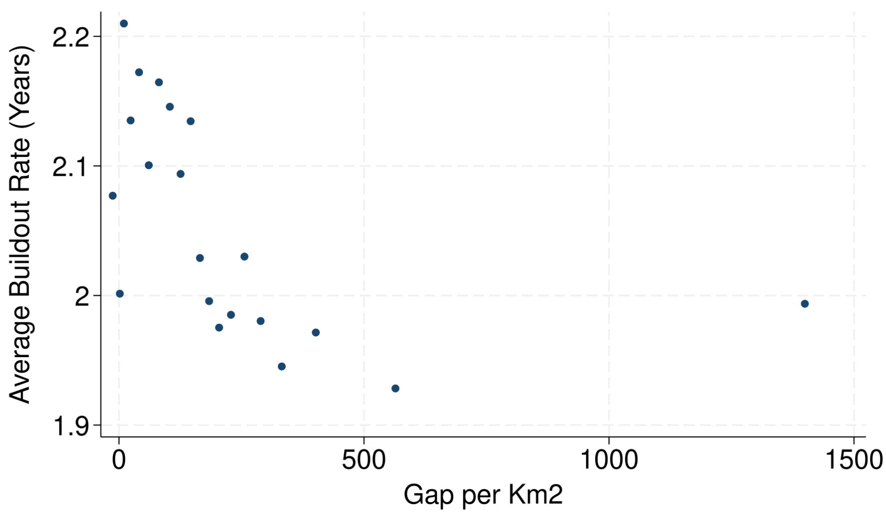
While the above documents clear demand side patterns, supply side frictions must additionally be considered. As the housing crisis is acute, building speeds and building ease ("how to build") need to be considered along with the location ("where to build"). Earlier (Figure 8), we discussed that larger projects – particularly urban extensions – do not seem to be able to deliver housing more efficiently despite their size. Figure 31 shows an even finer breakdown of project sizes and plots on the left hand side the number of filings that projects of different sizes are required to make within the planning system and on the right hand side the time from first filing to last planning outcome. It clearly demonstrates that larger projects, starting at about 50 units or more, face much higher hurdles in terms of their paperwork through the number of filings, and can often spend years in the planning system. On the other hand, smaller projects of one to 50 units, have a much lower filing burden (not more than four filings per accepted project on average) and spend, on average, less than three years in the planning system, less than two years on average for projects with 10 units or less.

Figure 31: Planning barriers by project size



Note: The figure shows boxplots of the number of filings (left panel) and time from first filing to last outcome (right panel) for residential planning applications for new units filed in Great Britain from 2000 through 2025. Applications are binned by number of the applied for units. Red bars are for rejected projects, blue for permitted ones. The blue and red dots show the respective averages and the black line is the median within a group. The box shows the 25th percentile at the lower end and the 75th percentile at the upper end. For further details on the planning data please see section B.3.

Figure 32: Average buildout rate and housing gap



Note: The figure shows the binned scatterplot of the average buildout rate at the OA level against the gap per km^2. The buildout is restricted to builds completed between 2010 and 2020. The project start date is used to calculate the buildout rate when available; otherwise, the first plot start date is used. Data on builds is from NHBC, and the gap was constructed using Rightmove data. For details on the construction of a build's completion date and the construction of the gap variable, please refer to Appendix sections B.2 and C.3 respectively.

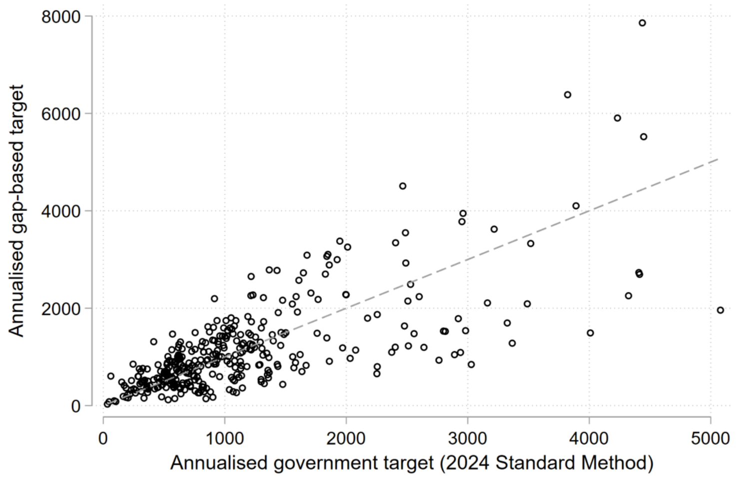
Figure 32 further shows that the buildout rate – a developer decision – is several months faster on average when building takes place in areas with a high housing gap than in areas with a low housing gap. This is intuitive, as builders will wait longer to build in areas with low demand to put new units on the market than in areas with high demand where these units will be filled fast and area-wide prices are likely less affected. This result presents further empirical evidence on the fact documented in Competition and Markets Authority (CMA) (2023) that housebuilders will sell properties at a rate consistent with the local absorption rate, in order to avoid reducing prices.

Taken together, the two figures suggest that the fastest delivery of housing under the current planning system and developer composition would therefore be in the form of smaller developments in areas with a high demand for housing. These tend to be areas in specific parts of the core of medium or larger towns which represent most of the excess demand, followed by housing in areas that would require urban extensions. Much of the urban extensions in highly demanded areas are on the urban fringes of cities surrounded by greenbelt land while some of the new rural town developments are likely to satisfy the demand for second rather than primary homes.

4.4 Aggregating to Local-Authority-level housing targets

As we discussed in Section 2, a key national policy around housing delivery are the LA specific housing targets. The formula for calculating these targets was updated in 2024 to be primarily based on the existing housing stock. We use the aggregated housing gap to allocate the annual equivalent of 1.5 million homes over five years to LAs based on their proportion.[11]

Figure 33: Gap-based LA target vs target under the 2024 Standard Method



Note: The figure shows a scatter plot of the LA-level target calculated based on the housing gap and the LA-level government targets (calculated based on the 2024 Standard Method). Our targets allocate the same number of new builds across local authorities as England, Scotland, and Wales aim to allocate overall. For details on the calculation refer to Appendix section C.5.

Figure 33 shows the correlation between the government's target and the equivalent calculated using the housing gap for all of Great Britain. Overall, it seems that the government's new standard method has set housing targets that are roughly in line with excess demand across LAs in most of Great Britain: There is a strong positive correlation between the government's LA targets and the targets based on the housing gap. For many local authorities the gap-based target aligns closely with the government's target with differences of 100 units or less. However, there are some exceptions in which the gap-based target suggests allocating more or less housing than the government's targets. The biggest positive discrepancies exist in the North of England. Particularly, the gap-based target would allocate over 1600 units more each to Leeds, Manchester, and North Yorkshire LAs. At the same time, the LAs with the largest negative discrepancies are the London boroughs of Barnet, Brent and Ealing. Appendix Figure A.5 plots these differences on a map.[12]

We note that this is not meant to say that less housing should be built in some LAs, for example those in London, in general. In fact, the actual total LA gap we have calculated is always considerably larger and always exceeds the current government LA housing targets, reflecting that all current LA targets are inadequate in comparison to long-run demand. However, *given the government's target of (only) 1.5 million homes*, our measures suggest that a higher fraction of these homes should be allocated to Northern LAs than to some of the London boroughs. While we note that there are large differences between the target and gap in Cornwall and South Hams, due to the higher than average proportion of second homes in these areas, we have reservations

[11] In practice we use separate gaps for England, Scotland and Wales, aggregated from LA targets which amount to England - 371994, Scotland - 20045 and Wales - 9879 homes per year.

[12] Our LA level gap from which we subtract the government's target to give the differences plotted in figure A.5 can be seen in figure A.6 in the appendix.

the excess demand there is entirely driven by a reasonable definition of housing need.

We supplement the housing targets with information on the planning system on average number of filings, average wait time in months from the filing of the first to the outcome of the last filed application for projects, and average refusal rates by local authority. We can then compare these two pieces of data. Table 3 shows local authorities in the top 10% of planning barriers (highest wait times, filings and refusal rates) that are also in the top 25% of the government's housing targets. With the exception of North Northamptonshire and North Yorkshire, these LAs are all London boroughs. While the gap-based target is usually lower in these local authorities, these are nevertheless areas with a high housing demand that are top candidates for planning reform or devoting additional resources in the local planning system.

Table 3: Local authorities with the highest housing targets and planning barriers

Local authority	Target	Gap	Avg wait time	Avg filings	Refusal rate (%)
Bexley	1801	909	17.99	2.37	32.07
Croydon	2763	1045	16.01	2.29	37.70
Hillingdon	2260	798	18.07	2.36	42.48
Lambeth	3041	1695	18.04	3.56	33.34
Lewisham	2470	1194	15.59	2.63	30.11
Merton	1936	1138	15.79	2.06	28.41
N Northamptonshire	2064	2280	15.77	2.02	28.33
North Yorkshire	4232	5898	18.38	2.03	28.16
Southwark	2710	1784	22.07	4.22	31.33
Wandsworth	3880	2731	16.08	2.90	21.29

Table 4 on the other hand shows local authorities in the bottom 10% of planning barriers and top 25% of government housing targets. These local authorities are some of the largest cities in England, namely Birmingham, Manchester, and Leeds. The gap-based target is even higher than the government's targets for almost all of these local authorities, indicating a strong excess demand for housing. At the same time, planning barriers are comparatively low, particularly average wait times and the number of filings required by the planning system. In other words, these are ideal candidates for delivering housing that is both highly demanded as well as in areas that currently have an efficient planning system that imposes few hurdles on new construction.

Table 4: Local authorities with the highest housing targets and lowest planning barriers

Local authority	Target	Gap	Avg wait time	Avg filings	Refusal rate (%)
Birmingham	4974	5515	7.46	1.25	29.79
Dudley	1594	1500	6.88	1.24	23.06
Leeds	4159	6377	7.63	1.23	32.24
Manchester	2686	4504	6.82	1.20	25.62
Somerset	3891	4098	5.37	1.16	17.28

Overall, housing targets are on average allocated according to housing demand under the current new standard method. There are, however, a few exceptions, particularly large cities in the midlands and north of England, that have substantially higher demand than housing targets allocated to them and additionally have planning systems that should facilitate the delivery of large amounts of housing rapidly.

5 Takeaways & Further Work

Over the past two decades, Britain's pattern of housebuilding has not fully aligned with where people most want to live — with an undersupply of new homes in high-demand urban centres and greater building in lower-demand areas.

Using data on over 20 billion housing searches coupled with detailed information on available supply, we construct a new, excess-demand–driven measure — the housing gap — to identify precisely where homes are needed most.

Our analysis shows that around half of Britain's housing gap lies in areas suited to densification—particularly within parts of London such as Wandsworth, Islington, and Camden, and cities including Manchester, Bristol, Salford, Edinburgh, and Portsmouth. Yet, only 29% of new homes since 2010 have been built in these high-demand densification areas. Instead, a growing share of new builds has been concentrated in lower-demand areas, signified by worse-than-average job access and a lack of amenities such as GPs or secondary schools.

We recommend that national and local policymakers reallocate housing targets to reflect actual demand, focusing less on cross–local authority comparisons and more on the neighbourhood-level. Central and local government should increase the share of homes delivered through densification substantially, prioritising high-demand areas. In addition, planning reform should target the 26% of the housing gap that lies within restrictive designations such as greenbelts — particularly around cities like Manchester, Leeds, and Liverpool — to enable much-needed growth of England's northern cities.

Policymakers can use our interactive map and Local Authority reports to identify priority neighbourhoods, align housing targets with demand, and locate areas where planning designations most limit growth. Our next phase of work will deepen this analysis by incorporating micro-level planning barriers—linking our demand measures to the specific frictions that constrain supply on the ground.

References

Ahlfeldt, G. M., K. Moeller, S. Waights, and N. Wendland (2017). Game of zones: The political economy of conservation areas.

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015). The economics of density: Evidence from the berlin wall. *Econometrica 83*(6), 2127–2189.

Anagol, S., F. V. Ferreira, and J. M. Rexer (2021). Estimating the economic value of zoning reform. Technical report, National Bureau of Economic Research.

Büchler, S. and E. Lutz (2024). Making housing affordable? the local effects of relaxing land-use regulation. *Journal of Urban Economics 143*, 103689.

Competition and Markets Authority (CMA) (2023). Housebuilding market study. https://www.gov.uk/cma-cases/housebuilding-market-study. Published 28 February 2023, updated 22 October 2024.

Costa, F., L. Nunes, and F. M. Sanches (2024). How to attract physicians to underserved areas? policy recommendations from a structural model. *Review of Economics and Statistics 106*(1), 36–52.

CPRE (2018). England green belts: Green belt fact sheets — england. Fact sheet, Campaign to Protect Rural England (CPRE). Accessed: 20 October 2025.

for Cities, C. (2025, October). Designing density: How to increase the supply of housing in existing residential areas. https://www.centreforcities.org/blog/designing-density-how-to-increase-the-supply-of-housing-in-existing-residential-areas/. Accessed: 2025-10-12.

Frost, R. (2020). Who is moving and why: Seven questions about residential mobility.

Glaeser, E. L., J. Gyourko, and R. E. Saks (2005). Why have housing prices gone up? *American Economic Review 95*(2), 329–333.

Hilber, C. A. and W. Vermeulen (2016). The impact of supply constraints on house prices in england. *The Economic Journal 126*(591), 358–405.

Koster, H. R. (2024). The welfare effects of greenbelt policy: Evidence from england. *The Economic Journal 134*(657), 363–401.

Kulka, A. and N. Datta (2024, September). Why building new towns isn't the answer to the uk's housing crisis. Accessed: 2025-10-06.

Kulka, A., A. Sood, and N. Chiumenti (2024). Under the (neighbor) hood: Understanding interactions among zoning regulations. *Available at SSRN 4082457*.

Lomax, N. (2021). Estimating household mobility using novel big data. In *Big data applications in geography and planning*, pp. 25–42. Edward Elgar Publishing.

Moreno-Maldonaldo, A. and C. Santamaria (2024). Delayed childbearing and urban revival: A structural approach.

New Towns Taskforce (2025, September). New towns taskforce: Report to government. https://www.gov.uk/government/publications/new-towns-taskforce-report-to-government/new-towns-taskforce-report-to-government. Accessed: 2025-10-12.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of political Economy 90*(6), 1257–1278.

Rosen, S. (1979). Wages-Based Indexes of Urban Quality of Life. In P. Mieszkowski and M. Straszheim (Eds.), *Current Issues in Urban Economics*, pp. 74–104. Baltimore: Johns Hopkins University Press.

Saiz, A. (2008). On local housing supply elasticity. *Available at SSRN 1193422*.

Schuetz, J. (2020, January). To improve housing affordability, we need better alignment of zoning, taxes, and subsidies. Policy 2020 / big ideas, The Brookings Institution. Accessed: 2025-10-12.

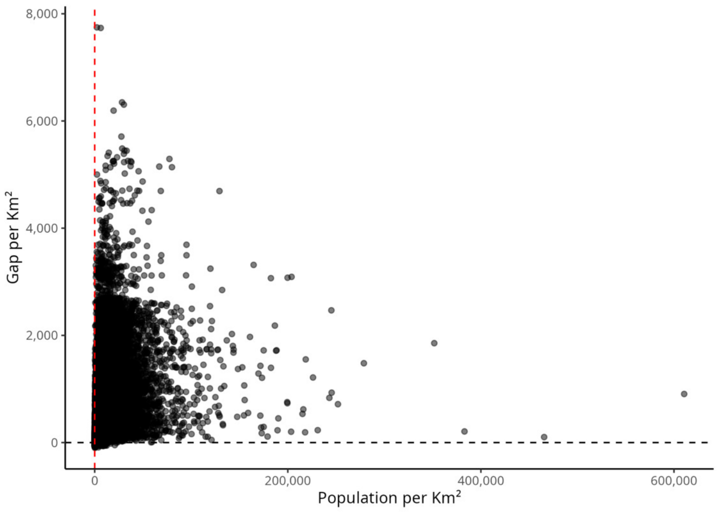
Venator, J. (2020). Dual earner migration, earnings and unemployment insurance. *Manuscript. https://joannavenator. files. wordpress. com/2020/11/venator_jmp_uifortrailingspouses_111920 20.*

Watling, S. and A. Breach (2023, February). The housebuilding crisis: The uk's 4 million missing homes. Accessed: 2025-10-12.

Yamagishi, A. (2021). Minimum wages and housing rents: Theory and evidence. *Regional Science and Urban Economics 87*, 103649.

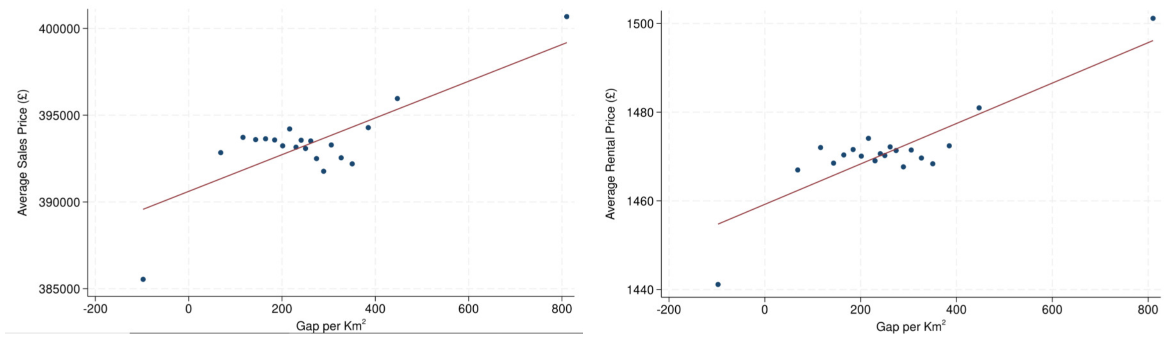
A Additional Figures and Tables

Figure A.1: Gap Per Km Squared vs Population



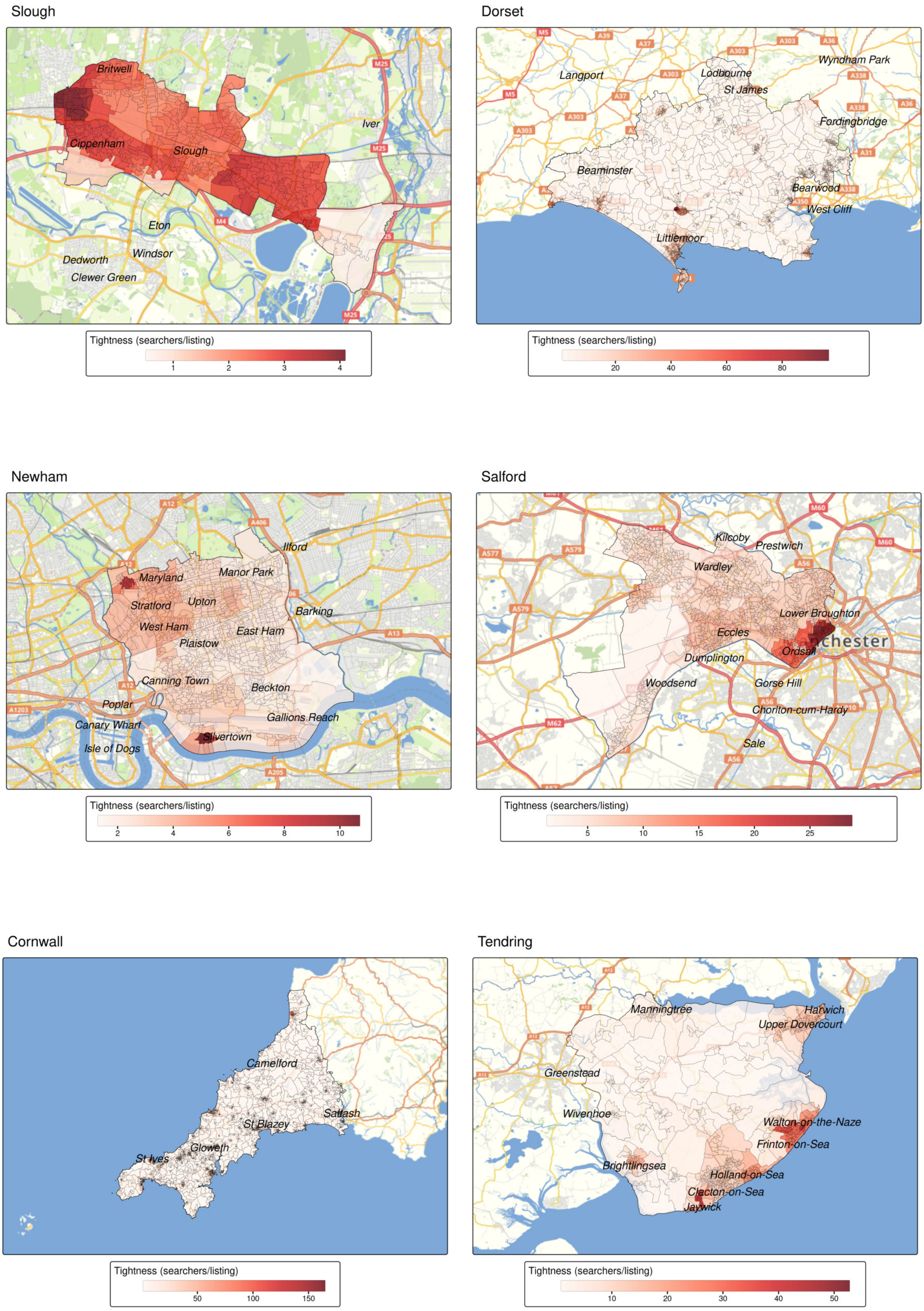
Note: The figure shows the scatter plot of the 2014 gap per km^2 against population per km^2. Both variables are at the OA level. Population estimates are collected from the Office of National Statistics, and the gap is constructed using Rightmove data. For details regarding the construction of the gap variable, please refer to Appendix section C.3.

Figure A.2: Relationship between average sales/rental prices and the housing gap per square kilometer within Local Authorities

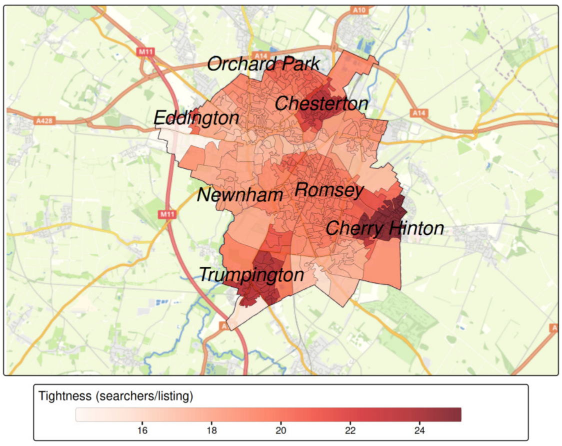


Note: The above figures show the binned scatter plot at the OA level of the sales and rental prices against the gap per km^2 residualised by local authority fixed effects. Price data are collected from Rightmove, which is also used to construct the gap per km^2. For details regarding the construction of the gap, please refer to Appendix section C.3.

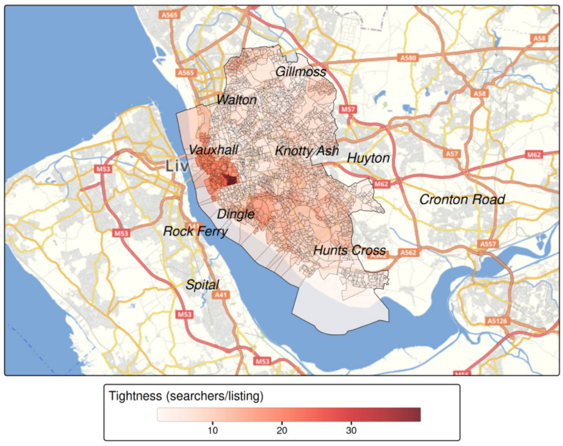
Figure A.3: Tightness



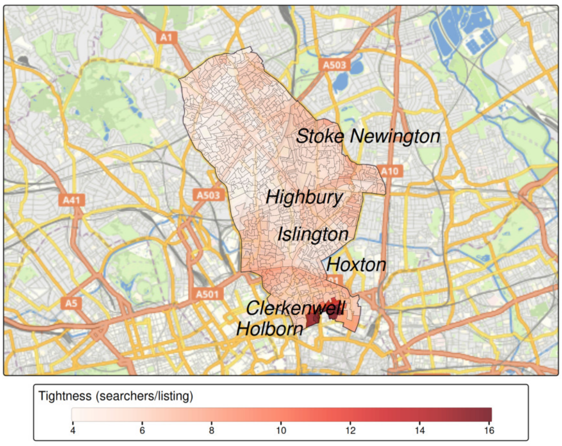
Cambridge



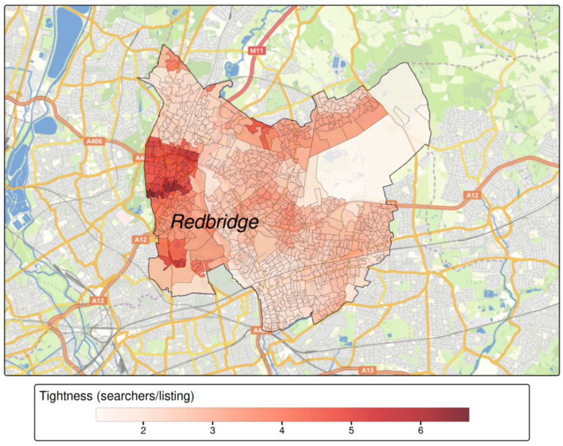
Liverpool



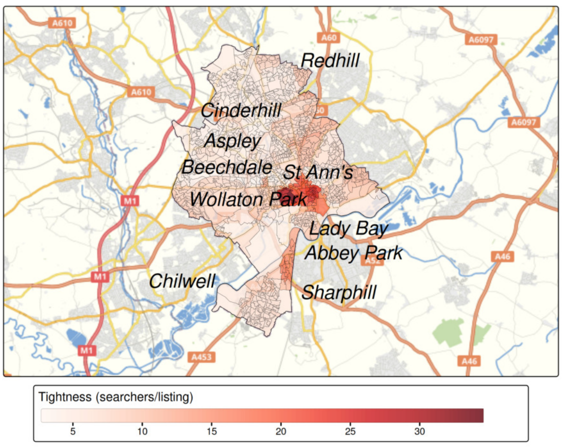
Islington



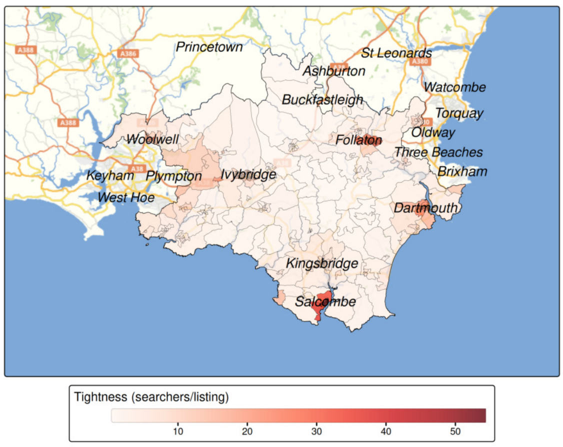
Redbridge

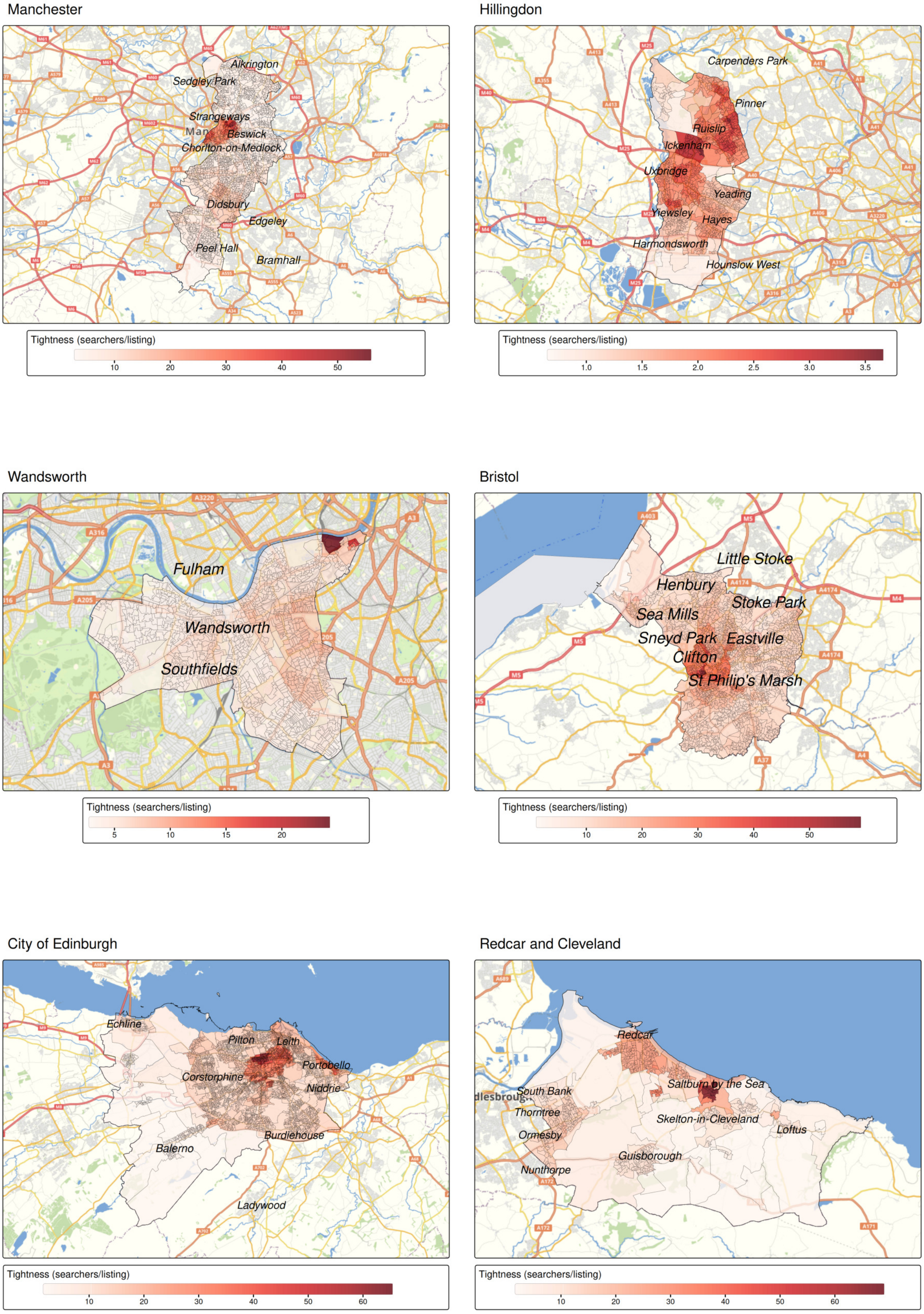


Nottingham



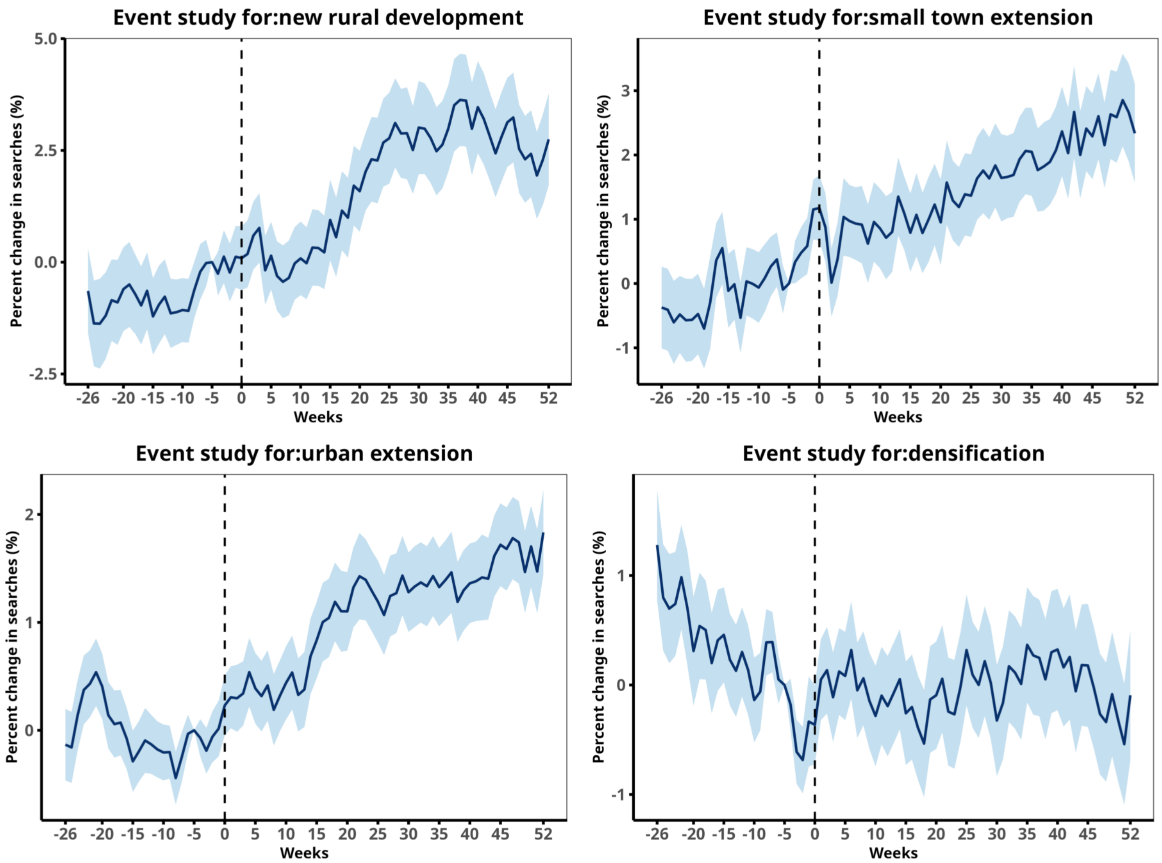
South Hams





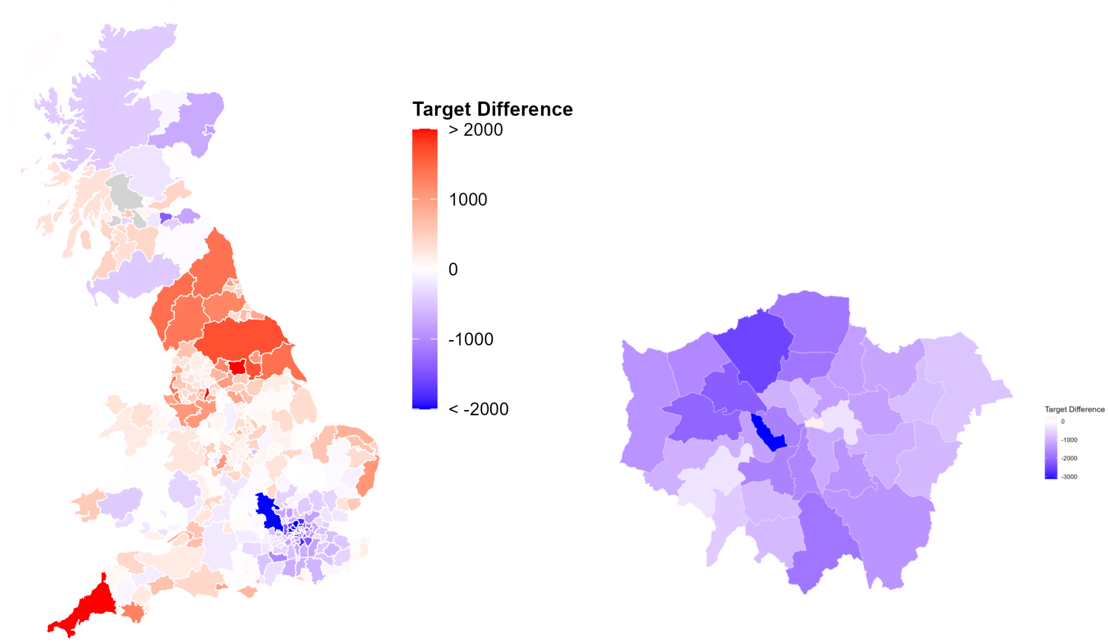
Note: The above maps show the tightness at the OA level across different LAs. The tightness variable was constructed using Rightmove data. For further details on the construction of the variable, please refer to Appendix C.3.

Figure A.4: Event Study by Classification



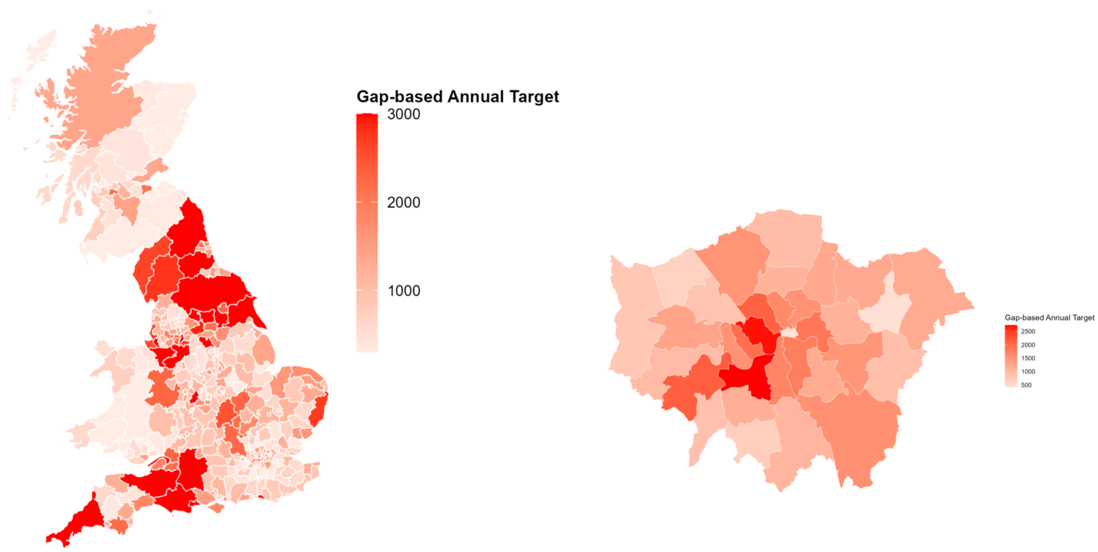
Note: The figure shows the response of searches in percentage change by classification across a window of -26 to 52 weeks relative to a the shock. The effects are normalised to week -5 relative to a shock. Search and builds data used are collected from Rightmove and NHBC respectively. For details on the data set construction and regression specification refer to section C.4.2.

Figure A.5: Difference to government target



Note: The figure plots the difference between the LA-level target calculated based on the housing gap and the LA-level government targets. Our targets allocate the same number of new builds across local authorities as England, Scotland, and Wales aim to allocate overall. Positive numbers (in red) indicate that the housing gap suggests increasing targets, negative numbers (in blue) indicate that the housing gap suggests a decrease in target. For details on the calculation refer to Appendix section C.5.

Figure A.6: 1.5 million homes allocated according to search gap



Note: The figure plots the LA-level target calculated based on the housing gap. London borough's are enhanced and shown separately on the right. Gap-based targets are allocated according to the fraction of the overall gap that lies in a local authority. These fractions are then applied to the same number of total new builds across local authorities as England, Scotland, and Wales aim to allocate overall. Therefore, our targets sum up to the same overall number of units as England's, Scotland's and Wales' targets. For details on the calculation refer to section C.5.

B Data appendix

The three main datasets used are the builds, property searches, and planning data sets described further below. We collected data on builds across Great Britain from NHBC, and data on searches from Rightmove.

B.1 Rightmove data

The Rightmove data contains the universe of all searches made between 2019-2024 for both sales and rentals. The data includes unique, 32-digit hexadecimal hash-keys for users such that they can be tracked over sessions. The search data includes all aspects of the search, including the properties they clicked on, saved, and contacts made to estate and letting agents. Importantly, the data contains the precise search location terms, which could be place names, stations, postal areas, districts and full postcodes, as well as the radius included in the search if used. We additionally have the spatial polygons for all of these searches, which are used to allocate searches to Output Areas. Our baseline restriction of the search data is such that users should have:

1. More than one session
2. A session longer than 60 seconds
3. At least one search
4. Less than 10,000 events

These restrictions ensure that we eliminate search bots. This is an important cleaning step to make sure that we are capturing preferences for residential locations accurately. We describe robustness of our demand measures to further sample restrictions on the types of searchers below in section C.4.1.

On the supply side, the data also contain all sale and property listings for the same time period.

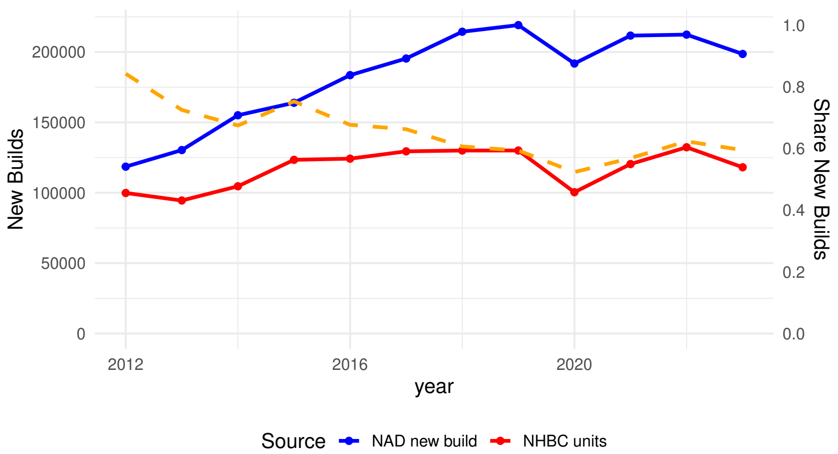
B.2 NHBC

Coverage. The builds data set initially covers the United Kingdom; however, using sites' postcodes, we identified that all construction sites in Northern Ireland have missing boundaries. Consequently, we restrict our analysis to Great Britain by excluding those sites. The NHBC covers a substantial chunk of all new builds. For England, Figure B.1 shows that coverage was as high as 90% of all newbuilds in 2012. While coverage nowadays has dropped closer to 60% this is still a large fraction. Figure B.2 further shows that coverage is high for most areas in England.

The data are organized into construction projects and development sites within these projects. We correct any invalid polygon geometries in the construction project polygons and aggregate all polygons associated with the same site into a single geometry. Lastly, we keep only sites that are active or have been completed. Our final data set consists of approximately 141,000 sites and 3.2 million builds.

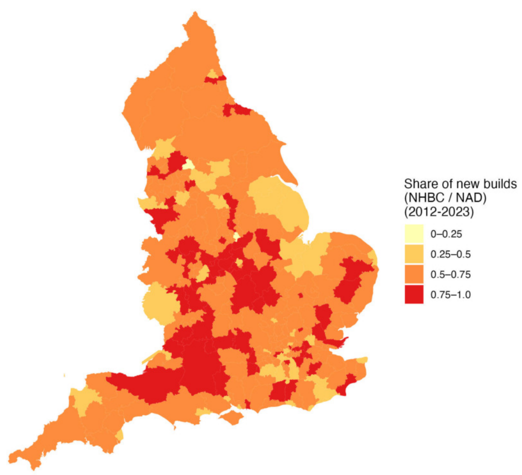
Build completion date. We use the build completion date when calculating the buildout rate and a project's completion time. To identify the completion date for each build, we first picked the earliest date between when the plot was finalized for warranty and the building control final date. If neither of those dates is available, we chose the date when the property was legally completed. For any plots still missing a date, we checked for the last recorded handover inspection, and if that's not available, we used the first pre-handover date (i.e. date at which first pre-handover inspection stage took place), provided the inspection was not marked as failed. If none of these dates were available, we considered the plot incomplete.

Figure B.1: Coverage of NHBC data in England



Note: The figure compares the number of new builds by year in the NHBC sample (red line) to the government's Net Additional Dwelling numbers (blue line). The dashed yellow line plots the ratio $\frac{NHBC}{NAD}$ against the right-hand side y-axis which shows the share.

Figure B.2: Spatial coverage of NHBC data in England



Note: The figure plots the average share of new builds recorded in NHBC data relative to the NAD data ($\frac{NHBC}{NAD}$) by local authority in England for the time period of 2012-2023.

B.3 Planning Dataset Construction

We construct a project-level dataset of residential applications in the UK for the period 2000-2025, with the aim of identifying and grouping applications related to new residential housing projects. These projects range from single homes to developments comprising thousands of units.

The initial dataset consists of approximately 18 million planning applications filed across the UK between 2000 and 2023. It includes applications for new residential dwellings—the focus of our analysis—alongside other kinds of applications, such as residential extensions, changes of use, commercial floor space, advertising permissions, and tree-cutting activities. Importantly, each submission is logged as a separate application, regardless of its nature, with all procedural steps for a single project—outline, approval of reserved matters, discharge of conditions, and amendments—recorded as distinct entries.

For each application, we collect the text describing the project, the address, key dates (filing, consultation, decision), the final decision outcome, the type of application (e.g., outline, full, discharge of condition), and other procedural metadata.

We proceed in four main steps to prepare our dataset. First, we identify new residential development applications by filtering and classifying unstructured text in application descriptions using keyword searches and machine learning models. We then extract the proposed number of housing units from each application by combining regular expressions with language models to interpret and count relevant information in the descriptions. Next, we group applications referring to the same project using explicit cross-references, shared addresses, and clustering algorithms to reconstruct each development's planning process and avoid double-counting. Finally, we aggregate data at the project level by assigning each group a unified decision outcome and number of units using a hierarchy of rules based on the content and outcomes of individual applications.

B.4 Additional Data Sources

To supplement these main datasets, we make use of several publicly available data sources which we list here. In creating our classification, we collected Built-Up Area sub-divisions population data from BuASD Population, Nomis. To create the aggregated London BuASDs, we used the data on the boundaries of the London Boroughs from London.gov.uk. We also used data on Secondary Schools and GP locations from Schools.gov.uk and House of Commons Library respectively. For market access calculations and BuASDs centroid calculation, we used employment data at the LSOA level from the Business Register and Employment Survey (BRES). For Scotland, we used employment data at the data zone level from Data Zones Employment, Nomis. Additionally, we collected data on the travel distance between LSOAs from Travel Time API. We also collected LSOA population data from LSOA Population estimates, ONS. Furthermore, we gathered OA Population Data from OA Population Estimates, ONS. For greenbelts and other planning designations shapefiles, we used the data from Planning.gov.uk. We also collected the shapefiles on data zones boundaries from 2011 Data Zones, ONS. For Scottish settlements and localities boundaries and population, we gathered the data from Localities and Settlements, NRS. Lastly, for OA, LSOA, MSOA, BuASD and BuA shapefiles, we collected the data from Geoportal, ONS.

C Methodology appendix

C.1 Project Classification

Following the classification rule in Table 2, both densifications and extensions are defined for BuASDs that have a population size of more than 20,000. We classify as densification developments that are in the inner 70% of the majority of these BuASDs, spanning out from the employment centre. Urban Extensions are sites that lie in the outer 30% or within a 1 km buffer of the original BuASD. In contrast, small town extensions are sites that lie within a BuASD that has a population between 5,000 and 20,000 or within a 1 km buffer of such a BuASD. As for new rural developments, those are sites that lie inside or within a 1 km buffer of a minor BuASD (population less than 5,000) or sites that are not matched to any BuASD. Importantly, since we only have sites' boundaries, we assign all projects within a site to the site's classification. We now describe the classification procedure in detail.

We began cleaning the data by consolidating the geometries for each site from the NHBC dataset by merging all polygons associated with a given site together. Next, using official BuASD boundaries for England and Wales and Locality boundaries for Scotland, we merged in the population data at the BuASD/Locality level and classified each by size: "minor" ($<$ 5,000 people), "small" ($<$ 20,000), "medium" ($<$ 75,000), "large" ($<$ 200,000), and "major" ($>$ 200,000), as used by the ONS.

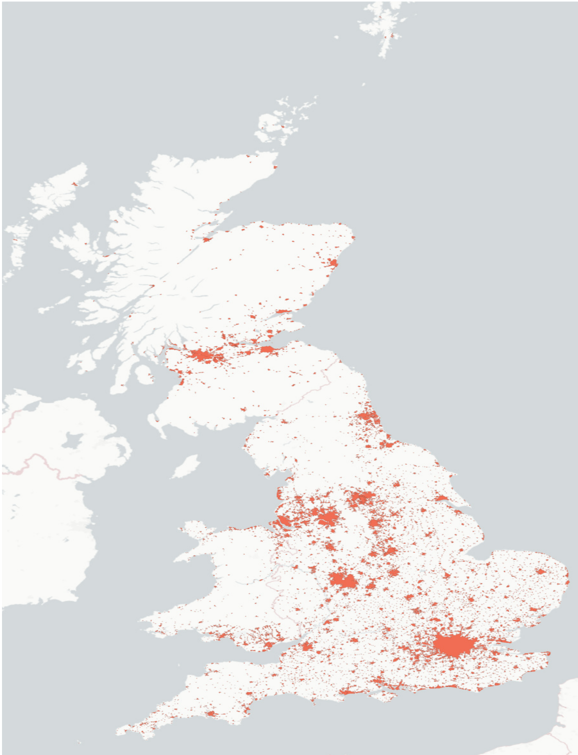
Prior to classifying BuASDs/Localities according to our definitions, we addressed the risk of misclassifying sites in the center of the five largest urban areas in Great Britain (London, Greater Manchester, Birmingham, Liverpool, and Glasgow) as the inner parts of these areas are split into several BuASDs/Localities. For this purpose, we aggregated all BuASDs contained within the BuA of each of these areas with the exception of London and Glasgow into a single unit by spatially intersecting and merging them. For London, we combined all BuASDs that spatially intersect the London Boroughs. We further enclosed the Thames River within the London boundary polygon and merged it accordingly so that demand near the Thames does not get classified as an urban extension. We also closed the gap in the London polygons between approximately Edmonton and Fairlop for the same reason. The population of each of the aggregated four areas was defined as the sum of the populations of their corresponding BuASDs. For Glasgow, we aggregated all localities intersecting the Glasgow settlement and assigned to the resulting geography a population equal to the sum of the populations of its constituent Localities.

Our final dataset includes BuASDs in England and Wales and localities in Scotland, with the exception of the BuASDs for Greater Manchester, Birmingham, Liverpool, and London, as well as the localities for Glasgow, which were replaced by their corresponding aggregated units. For simplicity, we refer to both BuASDs (in England and Wales) and localities (in Scotland) collectively as BuASDs. Figure C.1 below shows a map of the final BuASDs used in Great Britain.

To classify densifications and urban extensions, we first restricted the dataset to BuASDs with populations greater than 20,000. For those BuASDs, we defined the center of each BuASD as the centroid of the LSOA, conditional on the centroid lying within the BuASD, with the highest employment per km squared. Since LSOAs do not exist for Scotland, we follow the same process there but use Scottish Data Zones. Next, we defined the "core" as the area remaining after iteratively expanding circles from the center and intersecting these circles with the original BuASD polygons, stopping when 70% of the polygon is covered. For London, Greater Manchester, Birmingham, Liverpool, and Glasgow, due to their "metro-area" nature, we applied the same process but defined the core as the area where 50% of the original area is preserved (following in spirit the definition of suburbs in Moreno-Maldonaldo and Santamaria (2024)). We then intersected each site with its corresponding (if any) BuASD core; a site was matched to a BuASD if at least 25% of its area overlapped with the BuASD core. We classified sites that matched as "densification".

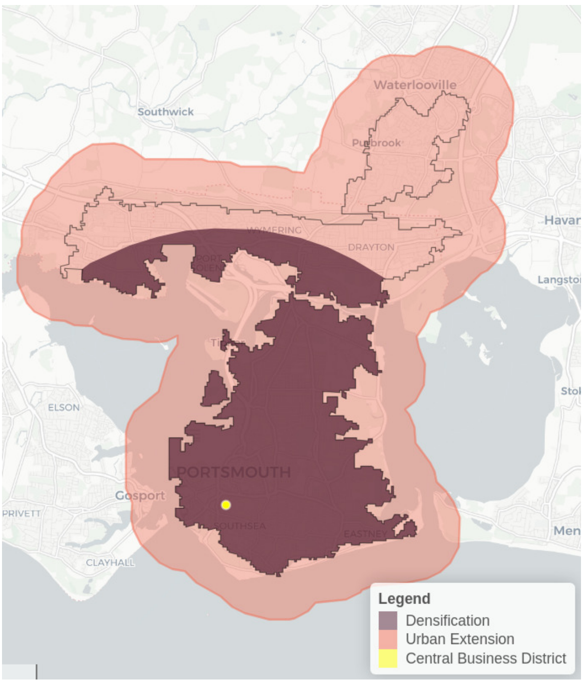
For sites that did not meet this threshold, we applied a 1 km outward buffer to the original filtered BuASD polygons, and checked which sites intersected the buffered polygons. Any intersecting site was then assigned to the buffered BuASD with which they had the largest area of overlap. If a site was fully covered by two separate buffered BuASDs, it was assigned to the larger BuASD. Sites that were matched to the buffered BuASDs were classified as "urban extensions". Figure C.2 below illustrates an example of the resulting polygons for Portsmouth.

Figure C.1: Map of BuASDs in Great Britain



Note: The above figure shows a map of the BuASDs in Great Britain after our merging process for BuASDs in London, Birmingham, Manchester and Liverpool as well as the merging of Localities in Glasgow.

Figure C.2: Densification and Urban Extension Definition Example: Portsmouth

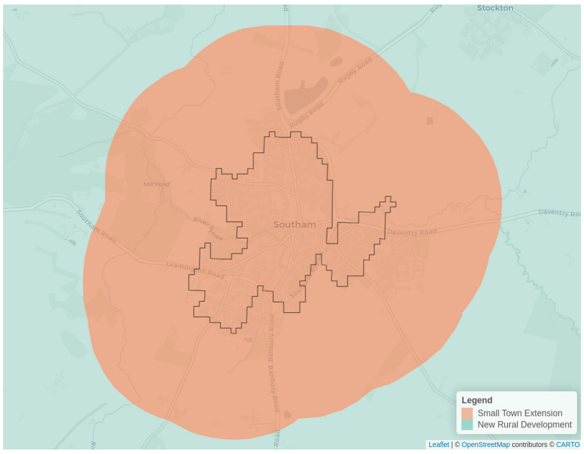


Note: The above figure shows the produced densification and urban extension polygons for Portsmouth. The yellow dot is the inner centeroid of the LSOA with the highest employment per km squared. The black border represents Portsmouth BuASD boundaries.

For small BuASDs, we applied a 1 km buffer on the original polygons and assigned each site based on the highest intersetion area with a BuASD. Similar to extensions, if a site was full covered by 2 separate buffered BuASDs, it was assigned to the larger BuASD. Matching sites were classified as “small town extension”.

Lastly, sites that intersected (using the same intersection rule as urban extensions and small town extensions) a 1 km buffered minor BuASDs or sites that intersected no BuASDs in our process were classified as “new rural development”. Figure C.3 below illustrates an example of small town extension and new rural development around Southam.

Figure C.3: Small Town Extension and New Rural Development Definition Example: Southam



Note: The above figure shows the produced small town extension and new rural development in Southam. The black border represents Southam BuASD boundaries.

C.2 Employment Access and Mean Deviation

To create the 25-minute public transport employment access dataset, we first collected employment numbers at the LSOA level (see Appendix B). For each LSOA, we then identified all other LSOAs within 25 minutes using a time travel matrix (see Appendix B). We defined "employment access" as the sum of the employment level across all LSOAs within the specified time. More formally, the employment access for LSOA i, where time between LSOA i and j is T_{ij} in minutes, and employment in j is E_j, and the population of i is P_i.

$$Employment\ Access_i \ = \ \sum_j E_j \, \mathbf{1}\{\, T_{ij} \leq 25 \,\} \tag{B.1}$$

We produced two versions of this measure: one based on 25-minute public transport travel times, and another using the fastest across both public and private transport.

We then merged in the population data at the LSOA level, and assigned each construction site to the LSOA covering the largest share of the site's area. Next, we calculated the weighted (by population) national mean of employment access. National average employment access is given by:

$$\overline{Employment\ Access} \ = \ \frac{\sum_i P_i \, Employment\ Access_i}{\sum_i P_i} \tag{B.2}$$

Next, we calculate the mean deviation employment access for each LSOA defined in equation B.3 as:

$$\widetilde{Employment\ Access_i} = \frac{Employment\ Access_i}{\overline{Employment\ Access}} \tag{B.3}$$

Finally, we split builds into different periods based on the build's start date (see Figure 6), and calculate the cumulative shares of new builds at the LSOA level within each period after sorting the data by mean deviation.

C.3 Gap and Tightness Calculation

C.3.1 Searcher Calculations

Let g represent a grid, p a polygon, i a user, n a search number, and m a listing. The key ideas in the calculation are:

- Each user should only count once, therefore for any user, summing over all the OAs in the UK should give a value of 1, for that user.
- Each listing should only count once, therefore for any listing, summing over all the OAs in the UK should give a value of 1, for that listing.

Note in this context g is an OA, while p is a search location on the property platform. As such p may be a location ("Bethnal Green"), a postal area/district ("SE1"), a train station ("Manchester Picadilly"), and may also have a radius attached.

Users and listings therefore need to be allocated to OAs. This needs to take into account two factors. Users can search across multiple polygons (and listings can appear in multiple polygons), and polygons contain multiple OAs. For searcher calculations, a user i search n is equal to 1 for polygon p if they searched that polygon, and 0 for all others

$$S_{ipn} = \mathbb{1}[i \ searched \ p] \tag{B.4}$$

Thus, summing over all searches gives a count of searches for each polygon for user i

$$S_{ip} = \sum_n S_{ipn} \tag{B.5}$$

In the first step, we need to allocate the searcher across all the polygons they searched in such that it sums to 1. Hence, let ω_p be some set of weights.

$$\hat{S}_{ip} = \frac{S_{ip}\omega_p}{\sum_p S_{ip}\omega_p} \tag{B.6}$$

We use a weight of $\omega_p = 1$ such that each polygon gets an equal weight.
Once users are allocated across polygons, users still need to be allocated to OAs within those polygons. All OAs receive equal weighting within the polygon. Therefore large polygons give each grid within it a smaller share and vice versa.

$$S_{igp} = \hat{S}_{ip} \ * \ \frac{g \in p}{\sum g \in p} \tag{B.7}$$

To then get the user allocation to a grid, we sum over all polygons:

$$S_{ig} = \sum_p S_{igp} \tag{B.8}$$

To get total grid level searches we sum over users i

$$\mathcal{S}_g = \sum_i S_{ig} \tag{B.9}$$

C.3.2 Listing Calculations

We count each unique property listing only once, in a similar vein to the approach with searchers. We use a similar allocation procedure as we do for searchers to ensure consistency.

A listing m is equal to 1 for polygon p if that listing appears in that polygon p :

$$L_{mp} = \mathbb{1}[m \ \in \ p] \tag{B.10}$$

First, we need to allocate the listing across all the polygons it falls in such that it sums to 1. Let ω_p be some set of weights:

$$\hat{L}_{mp} = \frac{L_{mp}\omega_p}{\sum_p L_{mp}\omega_p} \tag{B.11}$$

Again we use equal weighting ($\omega_p = 1$ for all p). Now, to allocate listings within polygons to OAs, we give each grid equal weighting within the polygon:

$$L_{mgp} = \hat{L}_{mp} \; * \; \frac{g \in p}{\sum g \in p} \tag{B.12}$$

To then get the listing allocation to a grid, we sum over all polygons:

$$L_{mg} = \sum_{p} L_{mgp} \tag{B.13}$$

To get total grid level listings we sum over listings m:

$$\mathcal{L}_g = \sum_{m} L_{mg} \tag{B.14}$$

C.3.3 Gap and Tightness

The gap is defined as:

$$\mathcal{G}_g = \mathcal{S}_g - \mathcal{L}_g \tag{B.15}$$

Tightness is given by:

$$\mathcal{T}_g = \frac{\mathcal{S}_g}{\mathcal{L}_g} \tag{B.16}$$

We multiply searchers in both equations by a scaling factor α for two main reasons. One, it is highly likely that different members within a household are searching simultaneously for their next property, and will use multiple devices (e.g. laptop, phone, tablet). Additionally, a given user may clear their cookies, which, in the absence of a user being logged-in, is how we track individual users. As such this gives us:

$$\mathcal{G}_g = \alpha \mathcal{S}_g - \mathcal{L}_g \tag{B.17}$$

$$\mathcal{T}_g = \alpha \frac{\mathcal{S}_g}{\mathcal{L}_g} \tag{B.18}$$

Mathematically, the scaling factor does not change the ranking of the tightness measure. We use an $\alpha = 0.125$, which in turn allows for two users in a single household, using two devices, who refresh their cookies once during the year. Importantly, tests reveal that the rank-rank correlation of $\mathcal{G}$ is still very high at this level, while reducing the gap to what is seen as a more realistic level.

C.3.4 Simulated gap

While we have data on listings going back to 2010, to have an estimate of the 2014 gap, to compare against builds over the 2010-2020 decade, we require an estimate of the 2014 searchers by OA. We simulate this using data on the 2019-2024 LA level growth rates, and extrapolate back, while adjusting for population growth.

Step 1: Estimating Local Authority (LA) Search Growth. For each local authority ℓ, we estimate an OLS regression through the origin using OA-level search counts in 2019 and 2024:

$$\mathcal{S}_{gl}^{2024} = \beta_\ell \mathcal{S}_{gl}^{2019} + \varepsilon_{gl}. \tag{B.19}$$

The estimated slope is given by:

$$\beta_\ell = \frac{\sum_g \mathcal{S}_g^{2019} \mathcal{S}_g^{2024}}{\sum_g (\mathcal{S}_g^{2019})^2}. \tag{B.20}$$

Note $\mathcal{S}_{gl}^{2024}$ and $\mathcal{S}_{gl}^{2019}$ are already adjusted for α.

If an LA has insufficient data (i.e., $\sum_g (\mathcal{S}_g^{2019})^2 = 0$ or $\beta_\ell \leq 0$), we replace β_ℓ with a global fallback rate:

$$\beta_{\text{global}} = \frac{\sum_\ell \sum_g \mathcal{S}_g^{2019} \mathcal{S}_g^{2024}}{\sum_\ell \sum_g (\mathcal{S}_g^{2019})^2}. \tag{B.21}$$

Step 2: Back-casting OA Searchers to 2014. Given β_ℓ, we back-cast each OA's 2014 searchers as:

$$\widehat{\mathcal{S}}_g^{2014} = \frac{\mathcal{S}_g^{2019}}{\beta_\ell}. \tag{B.22}$$

Any negative simulated values are set to zero.

Step 3: Rescaling to Match Population Growth. We then rescale the simulated 2014 searchers so that the mean ratio of 2024 to 2014 searchers equals 1.07, corresponding to the national population growth rate between 2014 and 2024:

$$\text{scale}_{2014} = \frac{\text{mean}_g \left(\frac{\mathcal{S}_g^{2024}}{\widehat{\mathcal{S}}_g^{2014}} \right)}{1.07}, \tag{B.23}$$

$$\widetilde{\mathcal{S}}_g^{2014} = \widehat{\mathcal{S}}_g^{2014} \times \text{scale}_{2014}. \tag{B.24}$$

Step 4: Computing the 2014 Gap. Finally, we compute the simulated 2014 OA-level gap as the difference between simulated searchers and actual listings:

$$\mathcal{G}_g^{2014} = \widetilde{\mathcal{S}}_g^{2014} - \mathcal{L}_g^{2014}. \tag{B.25}$$

C.4 Robustness and validation checks

C.4.1 Robustness Checks

Given a concern around potential noise issues in the search data we carry out several cleaning restrictions and confirm the results still hold. Noise is most likely among the searchers rather than listings as searches could be generated from bots or from non-serious/hobby searchers. We therefore calculate the number of searchers, $\mathcal{S}_g$ as per equation B.9 for the following user restrictions for our period in 2023-2024:

1. All non-bot users
2. Users who saved a property
3. Users who contacted an estate agent
4. Users who saved a property and contacted an estate agent

We believe that these restrictions ensure that we are capturing various levels of serious searchers who are displaying their true residential location preferences. We then calculate the OA level rank-rank searcher number correlation based on these restrictions. The results of this exercise are presented in the table below. The spearman correlation is extremely high for all restrictions, suggesting the results are not affected by noise on the search platform.

Table A1: Rank–rank correlations between user restriction groups.

Restriction 1	**Restriction 2**	**Rank–Rank Correlation (ρ)**
All non-bot users	Saved a property	0.956
All non-bot users	Contacted estate agent	0.929
All non-bot users	Saved and contacted estate agent	0.954
Contacted estate agent	Saved and contacted estate agent	0.993
Contacted estate agent	Saved a property	0.970
Saved and contacted estate agent	Saved a property	0.991

C.4.2 Validation: Event Study Details

Sample Construction

Defining our shock week as the first week of construction of the first build within a site, we use the following to create the data set used for our analysis. Starting with our data set of 141,000 sites, we restricted the sample to sites that had the first build starting post-2019 corresponding with the start of the search data; this reduced our sample size to around 13,300 sites. We then selected only sites that have a size of 20 builds or more, further reducing our sample to around 5,900 sites. To align our dataset with the search data, we narrowed the sample to shocks that occurred before May 2024, with 5,110 sites remaining. To maintain a full window for our analysis, we also kept only shocks that occurred 26 weeks after January 2019 and 52 weeks before May 2024, making our final shock sample consisting of 3,200 sites. We then found the intersecting OA for each site, and created the corresponding controls for a site through the following:

1. Across all OAs in the UK, we only kept OAs that are in Middle Lower Super Output Area (MSOA) that never intersected a treated OA.
2. From the surviving OAs, we then assigned for each site its control OAs such that the latter lay within the same Travel To Work Area (TTWA) as the intersecting treated OAs.

Lastly, we stacked all treated and control OAs and added a unique identifier for each site's corresponding OA in both groups, creating the final data set with around 265 million observations.

Regression Specification

Using the constructed dataset, Figure 15 is produced by running the following regression specification and clustering the standard errors at the OA-Site level:

$$ln(Searches_{s,o,yw}) = \sum_{\substack{k=-26 \\ k \neq -5}}^{44} \beta_k \; Treated_{s,o,k} + \gamma_{s,o} + \lambda_{s,yw} + \epsilon_{s,o,yw} \tag{B.26}$$

where $k = yw - t$ indexes the relative event time such that t refers to the calendar year–week of the shock. ln(Searches$_{s,o,yw}$) is the log of the number of buying searches in site stack s, OA o, and year-week yw. Treated$_{o,s,\,k}$ is an indicator variable equal to 1 if the treated OA o from stack s is in the k^{th} week relative to the shock (with $k = 0$ being the week of the shock), and 0 otherwise. β_k is the coefficient on the treatment indicator for event week k; the effects are normalised to $k = -5$. Lastly, $\gamma_{s,o}$ and $\lambda_{s,yw}$ denote $output\ area - stack$ and $yearweek - stack$ fixed effects respectively.

Figure A.4 is produced by running the same specification as above but for subgroups based on our classification of sites; we extend the analysis there to the full 52 weeks.

C.5 Target calculations

C.5.1 Local authority targets

To calculate optimal local-authority level targets and compare them to the official government targets, we proceed in the following steps:

First, as our baseline measure of the housing gap is at the output area level, we aggregate the output-area gaps within each local authority to generate a local authority level housing gap. Two notes are important here. a) We allocate output areas to the local authority their maximum area lies in. b) We set negative output area gaps to 0 as housing is fixed and durable and cannot be moved from one output area to another. Therefore, an excess supply of housing in one output area does not mitigate excess demand in a different output area.

Second, we sum across all local authority gaps and then take each LA's fraction of this overall gap. This is the optimal share of any target that should be allocated to each authority based on our housing gap measure. We calculate shares separately by country, as England, Wales and Scotland set housing targets in different ways. For England, the targets are calculated using the 2024 revised standard method discussed in section 2.1, for Wales and Scotland, targets are taken from local development plans.

Third, we aggregate the official local-authority level government targets by country. We then multiply our optimal shares with these country totals to arrive at our optimal local-authority target based on the housing gap (plotted in Figure A.6). The difference between our target and the governments' targets are plotted in Figure A.5.

C.5.2 Output area targets

The optimal OA-level targets follow a similar approach as our optimal LA-level targets. Instead of calculating shares at the LA-level, we calculate each OA's share of the total gap in the LA that it lies in (again, setting negative gaps to 0). We then multiply this share with the LA-target set by the government to give us our optimal OA-targets.

As governments do not set output-area level housing targets, there is no comparison possible in this case.