American Economic Review 2026, 116(3): 1014–1051
https://doi.org/10.1257/aer.20240513

Dynamics of the Long-Term Housing Yield: Evidence from Natural Experiments†

By Verónica Bäcker-Peral, Jonathon Hazell, and Atif Mian*

Each month, a fraction of UK property leases are extended by 90 years or more. We construct a new dataset using thousands of these natural experiments since 2000 and estimate the expected long-term housing yield, y^*. After remaining steady at around 5 percent, y^* starts to decline when the Great Recession hits and reaches a low of 2.7 percent in 2024. The decline is steeper in inelastic markets, while y^* remains higher in regions more exposed to long-run climate risk. Our estimate of y^* is updated in real time using public data. (*JEL* E32, G12, Q54, R31, R38)

Expected long-term yields are of natural interest in macroeconomics and finance as they reveal market's expectation about long-run equilibrium after medium-term shocks have dissipated. For instance, there is great interest in how much long-term yields have declined in recent decades, why this has happened, and whether the recent rise in yields is temporary (Blanchard 2023). However, expected long-term yields are notoriously difficult to measure—one needs data that can separate the long-term versus the medium-term behavior of market prices.

This paper exploits a natural experiment from the United Kingdom in order to estimate the expected long-term yield for housing. In the United Kingdom, most apartments are bought and sold as long duration leases, or "leaseholds," with initial lease terms that typically exceed 100 years. Periodically, a fraction of leases have their duration extended by 90 years or more.[1] We assemble a new administrative dataset on all such lease extensions between 2000–2024, with over 130,000 extensions in total. The data contains market prices of leaseholds both before and after

* Bäcker-Peral: Princeton (email: veronicabp@princeton.edu); Hazell: London School of Economics (email: j.hazell@lse.ac.uk); Mian: Princeton and NBER (email: atif@princeton.edu). Yuriy Gorodnichenko was the coeditor for this article. We thank the Julis-sRabinowitz Center for Public Policy and Finance at Princeton for financial support and Cristian Badarinza, Saleem Bahaj, Ambrogio Cesa-Bianchi, Matthias Doepke, Angus Foulis, Juan Herreño, Christian Hilber, Karin Hobelsberger, Ethan Ilzetzki, Andreas Mense, Ben Moll, Dmitry Mukhin, Christina Patterson, Tarun Ramadorai, Ricardo Reis, Karthik Sastry, Johannes Stroebel, Alwyn Young and seminar participants at NYU, LSE, Princeton University, Bank of England, and NBER SI Monetary Economics meetings for comments. We thank Chris Buckle, Alison Draper, Eileen Morrison, Oliver Saxby, and Leigh Shapiro for sharing their expertise on leaseholds and lease extensions. We thank Sam Barnett and Ruxandra Iosif for excellent research assistance. This paper was previously titled "Measuring the Natural Rate Using Natural Experiments". A website (https://sites.google.com/view/ystar-dynamics/home) accompanying this paper displays real-time updates to our estimates, data, and replication instructions.

† Go to https://doi.org/10.1257/aer.20240513 to visit the article page for additional materials and author disclosure statement(s).

[1] The typical extension is for an additional 90 years, but a large share of leases extend by 700 years or more, offering variation at various parts of the long-term yield curve.

extension. We develop an empirical methodology to estimate the real-time dynamics of the expected long-term housing yield, using the gain in leasehold market value due to duration extension. We have made our data and code public (https://sites.google.com/view/ystar-dynamics/home) and provide real-time updates so that other researchers can track and estimate y^* in real time.

Formally, y_t^* is the market's expectation of the long-term housing yield at a point in time, that is, $y_t^* = \lim_{u\to\infty} E_t\left[\frac{R_{t+u}}{P_{t+u}}\right]$, where $\frac{R}{P}$ is the rent-price ratio. y_t^* can equivalently be written as, $y_t^* \equiv r_t^* + \zeta_t^* - g_t^*$, where r_t^* is the time t expected return on safe assets in the long-term, ζ_t^* is the expected long-term risk premium for housing, and g_t^* is the expected long-term dividend growth rate. y^* contains information about important macroeconomic fundamentals, such as r_t^*. By measuring the expected rent-price ratio, y^* also tracks the price at which markets expect the housing market to clear in the long term.

Our empirical methodology for estimating y^* starts by estimating the increase in market value of a leasehold due to duration extension. We estimate the increase in value using a difference-in-differences estimator that compares the price growth of a leasehold, before versus after extension, with the price growth of an otherwise similar group of control leaseholds that do not extend at the same time. We combine the difference-in-differences (DiD) estimate of the gain in market value due to extension with a simple discounted cash-flow pricing equation, to estimate y^* via nonlinear least squares. y^* is measured in real terms and is unaffected by inflation expectations.

Figure 1 plots the estimated average y^* from 2000 to 2024, and the shaded area represents the 95 percent confidence interval. There are two key findings. First, though y^* stays stable around 5.2 percent between 2000 and 2006, there is a trend fall at the onset of the Great Recession, culminating in a low of 2.7 percent in 2024. The magnitude of this decline is large, equivalent to a near-doubling of the long-term price-rent ratio. Second, y^* remained relatively stable through the pandemic and afterward. The sample size expands over time, resulting in more precise estimates in recent years.[2]

An important advantage of our dataset is that it is not only public, but also updated in real time on monthly basis. Since y^* contains information about important macroeconomic parameters such as r^*, our methodology and real-time estimates should be useful for policy-relevant questions. There are close to a thousand lease extension natural experiments reported every month, allowing for precise real-time estimates. Figure 1 illustrates real-time estimates, plotting separately the monthly updates for 2024.[3]

Our natural experiment and microdata based approach for estimating the expected long-term yield has three main advantages. First, the approach "differences out" the shorter end of the yield curve. As such, we can identify the long-term yield y^* without making hard-to-verify assumptions about the medium

[2] Our methodology allows for real-time updates to our y^* estimates. In Figure 1 and Figure 9 we provide the most up-to-date series. However, for the remainder of the paper, we utilize the vintage of our data downloaded on February 1, 2024, for replicability.

[3] See the paper's website (https://sites.google.com/view/ystar-dynamics/home) for data, code, real-time estimates and replication instructions.

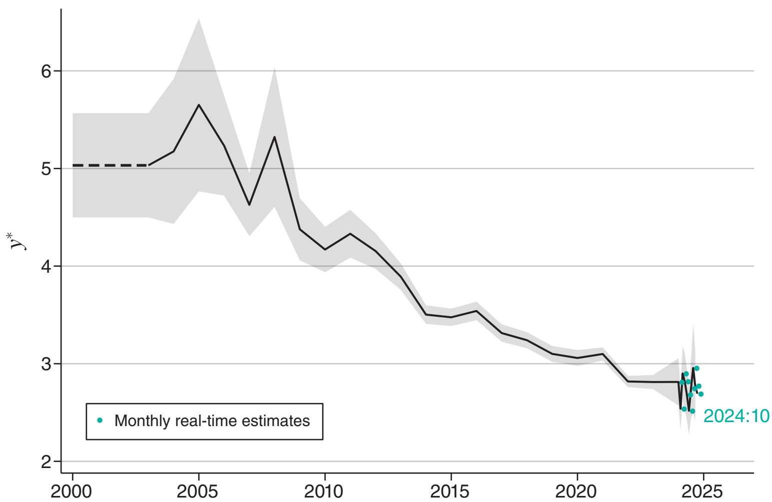

Figure 1. Time Series of the Expected Long-Term Housing Yield

Notes: The figure shows estimates of y^* over time for the full sample of lease extensions. Estimates for 2000–2003 are pooled and are reported in the dashed black line. Yearly estimates of y^* for 2004–2023 are plotted in the solid black line. Monthly estimates for January 2024 through October 2024 are also reported. The shaded area shows 95 percent confidence intervals for the estimates. Standard errors are heteroskedasticity robust.

run shocks affecting the economy. Intuitively, the price before extension measures short-term value, whereas the price after extension measures both the short- and long-term value of the same property. Hence the difference in prices depends almost exclusively on the long-term yield. Short- and medium-term shocks, such as monetary tightening, do not contaminate our estimate of y^*. Our estimator uses very long-term information about prices, almost a century in the future—meaning even persistent medium-term shocks will not matter. By contrast, forward yields from financial assets are typically available at most 10–30 years in the future.

Second, and relatedly, our method does not rely on structural assumptions about the macroeconomy in order to estimate expected long-term yields. A large literature attempts to estimate expected long-term yields for safe assets, making structural assumptions in order to filter out short-term disturbances (e.g., Laubach and Williams 2003; Holston, Laubach, and Williams 2017). This work is justly celebrated for tackling a difficult and important problem. However structural methods face model mis-specification concerns that can turn out to be important. For instance, the Holston, Laubach, and Williams (2017) estimates were discontinued during the pandemic due to mis-specification issues (Williams 2023).

Third, lease-extension natural experiments generate variation in lease duration for the *same property*. Thus our estimator is immune to spurious correlations between leasehold duration and unobserved heterogeneity in the service flow of housing. Similarly, our estimate is unaffected by short-term shocks to the service flow of housing. For instance, a shock to demand for a particular segment of London property, which raises its service flow, is also "differenced out" by our estimator.

Our methodology relies on the "parallel trends" assumption, meaning the service flow of housing must grow similarly for extended and control properties.

This assumption could fail if extending properties improve relative to controls, e.g., through renovations. However, using textual analysis of listings, we find no evidence that extended properties are more likely renovate. Additional results further validate the parallel trends assumption. In particular, there is no difference in pre-trends of control and extended property prices. These groups are also balanced on hedonic characteristics, and our estimator is robust to saturating variation coming from observed hedonics (Altonji, Elder, and Taber 2005; Oster 2019). Most importantly, rents measure housing service flow and show no difference in growth between extended and control properties.

In the final part of the paper, we examine the macroeconomic implications of our estimates and draw three key conclusions. First, although our estimate of y^* is based on UK housing, long-term yields exhibit similar behavior across other markets. Long-term forward bond yields closely track y^* throughout our sample, as do long-run trends in the yields of global housing and other assets. This suggests that the decline in expected long-term yields is broad based rather than housing specific. Given this widespread decline, a plausible driver is the fall in expected long-term safe asset yields, r^*, rather than housing-specific shifts in dividend growth, g^*, or risk premia, ζ^*. Consistent with this view, we find suggestive VAR-based evidence linking the decline in y^* to falling r^*, in line with a broader literature documenting similar trends (Holston, Laubach, and Williams 2017).

Second, the gap between current and long-term expected yields provides a real-time indicator of whether current housing valuations are transitory or persistent under market's expectations. We assess this by comparing current rent-price ratios to y^*, the expected long-term rent-price ratio. While these ratios aligned in 2000 and from 2010 onward, they diverged in the early 2000s, suggesting a temporary house price boom—consistent with other evidence. However, recent comparison between rent-price ratio and y^* suggests that the market now anticipates housing valuations to be more persistent.

Third, cross-sectional variation in y^* highlights the effects of inelastic land supply and long-run risk on yields. y^* declines more in supply-constrained areas, where rising demand boosts valuations. The level of y^* is higher in regions with greater long-run climate risk, measured by flood and subsidence exposure, suggesting that y^* embeds a significant housing risk premium.

Related Literature.—Our paper builds on the seminal work of Giglio, Maggiori, and Stroebel (2015), who observed that variation in UK apartment lease durations can be used to estimate long-run yields.[4] Giglio, Maggiori, and Stroebel (2015) estimate the average level of y^* by comparing properties with different lease durations while controlling for hedonic characteristics. In contrast, we focus on the dynamics of y^* and provide a new quasi-experimental design based on lease extensions that exploit within-property variation in leasehold duration and price. The natural experiment approach accounts for unobserved heterogeneity in housing service flows across different properties and provides real-time monthly estimates of y^*. Our

[4] See also Badarinza and Ramadorai (2015); Giglio, Maggiori, and Stroebel (2016); and Bracke, Pinchbeck, and Wyatt (2018).

methodology also explicitly incorporates the option value created by lease extensions and introduces a new bunching estimator to measure this effect.

Our paper also relates to a literature inferring the yield of capital using data from national accounts (e.g., Gomme, Ravikumar, and Rupert 2015; Farhi and Gourio 2018; Reis 2022; Vissing-Jorgensen 2022). This literature notes that the ratio of profits to the capital stock—which proxies for the yield of capital when there is perfect competition and constant returns to scale—has been stable, whereas we find that the expected long-run yield of housing and potentially other forms of capital has fallen. As Farhi and Gourio (2018) show, rising monopoly power in goods markets can reconcile these two phenomena.

Our paper relates to existing work measuring the evolution of long-term yields. Jordà et al. (2019) study long-run trends in the time series of various asset prices across decades. Their approach is well suited to tracking gradual evolution in expected long-term asset prices, by "averaging out" shorter-term factors without making strong assumptions. However one cannot measure higher frequency or real-time dynamics of long-term expected forward yield this way. A second approach, taken by Laubach and Williams (2003) and Holston, Laubach, and Williams (2017), specifies a structural macroeconomic model to estimate expected long-term yield on safe assets r^*.[5] As we have discussed, the structural approach may be more vulnerable to mis-specification than our method.

Outline.—The rest of the paper is structured as follows. Section I defines the expected long-run yield of housing. Section II describes the data. Section III presents our empirical methodology. Section V discusses the implications of our estimates for the macroeconomy. Section VI concludes with a discussion of possible future work based on our new methodology.

I. The Expected Long-Term Housing Yield

Our main object of interest is the expected long-term housing yield, which is the expected yield on housing once short- and medium-run forces have subsided. Consider the price of a unit of housing at time t, given by the expected present discounted value of its rent, R_t, as

$$P_t = R_t E_t\left[\int_0^\infty e^{-\int_0^s (r_t(u)+\zeta_t(u)-g_t(u))\,du} ds\right], \tag{1}$$

where $r_t(u)$ is the risk free rate u periods forward at time t, $\zeta_t(u)$ is the risk premium, and $g_t(u)$ is rental growth.[6] The housing yield is defined as $y_t(u) \equiv r_t(u) + \zeta_t(u) - g_t(u)$. The long-term expected housing yield, y_t^*, can be defined as

$$y_t^* \equiv \lim_{u\to\infty} E_t\big[y_t(u)\big] = r_t^* + \zeta_t^* - g_t^*.$$

[5] Related papers include Kiley (2015); Lubik and Matthes (2015); Johannsen and Mertens (2016); Crump, Eusepi, and Moench (2016); Hamilton et al. (2016); Rachel and Smith (2017); Christensen and Rudebusch (2017); Del Negro et al. (2017); Rachel and Summers (2019); and Del Negro et al. (2019).

[6] For simplicity this derivation omits a "rational bubble" term.

This is also equivalent to the long-run rental-price ratio, that is,

$$y_t^* = \lim_{u \to \infty} E_t\left[\frac{R_{t+u}}{P_{t+u}}\right] = R_t^{\infty}/P_t^{\infty}.$$

The expected long-term yield of housing contains important information about long-term safe asset yields r^*, long-term risk premia for housing ζ^*, and long-term dividend growth for housing g^*. These are obviously very important objects for both macroeconomics and finance. y^* also reflects expectations about long-run equilibrium structure of the economy, that is, the housing yield that is expected to bring the forces of supply and demand in balance in the long run.

The challenge in empirical identification of y^* is that we need to strip out any effects of short-run, transient shocks on housing yields, such as monetary policy, credit booms, short-run bubbles and adjustment costs. It is also difficult to observe the true rental value or "service flow" for housing. Section III explains how our empirical approach addresses these challenges in detail.

II. Data and Lease Extension Details

This paper builds a new administrative dataset on all lease extensions in the United Kingdom from 2000 onward. We are also making this dataset public, along with its real-time updates.[7] This section describes the construction of the dataset.

A. *Data*

Properties in the United Kingdom are divided into two categories: freeholds and leaseholds. Freeholds are a perpetual claim to the ownership of a property. Leaseholds are long duration leases to the property that can be bought and sold, and which typically last for many decades at origination. Leaseholds are typically apartments, and, importantly, are periodically extended, after negotiating a price for extension with the freeholder. As of 2022, there are 20 million freehold and 5 million leasehold dwellings in England. This paper builds a new dataset on extensions in the duration of leaseholds.

Constructing Public Leasehold Extension and Transaction Price Data.—We build a new administrative dataset of leasehold extensions based on historical archives of leasehold titles from His Majesty's Land Registry (HMLR), (HM Land Registry 2024b). The key challenge in identifying leasehold extensions is that public records only report information on current *open* leases. For example, a lease record will tell us about how many years are left in the current lease of a particular apartment. But publicly available records do not provide information on *closed* leases, that is, the prior lease of the same property that was extended at some point in time to create the current open lease.

Since we want to build a dataset on all historical lease extensions, we need (i) data on closed leases that were extended, and (ii) to match each closed lease with the

[7] See Bäcker-Peralh, Hazell, and Mian (2026)and https://sites.google.com/view/ystar-dynamics/home.

corresponding new extended lease. In order to do this, we purchased the universe of closed lease titles associated with transactions of flats in England and Wales before May 2023 from HMLR. These are lease titles which have been removed from the Lease Register because they have been overwritten or cancelled. We then identify which closed leases were subsequently extended by matching them with publicly available data on open leases. We obtained legal permission from the land registry office to make this newly compiled data on lease extensions publicly available on our website. This procedure gives us information on all lease extensions conducted prior to May 2023.

We also want to build our public lease extension dataset in a way that it is automatically updated in real time every month. Doing so enables us to estimate y_t^* in real time going forward—something that we hope will be of use to academics, central bankers and other policy professionals. In order to track every new lease extension going forward, we start with the full Lease Register of May 2023, and update it every month with the latest lease register vintage to identify new lease extensions. As such our public data website keeps getting updated with new real-time monthly estimates of y^*. As of the time of writing, we have identified approximately 319,000 leasehold extensions since 2000. The distribution of extension times and lengths is presented in Table A.1.

HMLR also publishes data on all property transactions in England and Wales starting in 1995 and updated monthly (HM Land Registry 2024a). The dataset includes the exact date, price and address for each transaction. We were provided with exact transaction IDs to merge transactions to closed lease titles. The open lease register does not include transaction IDs, so we conduct a fuzzy merge based on provided addresses to match open leases to transactions, as we detail in Supplemental Appendix A.3.[8]

Private Hedonics Data.—We augment our public data with two private datasets on housing characteristics and rental prices. While we cannot make this data public, as we later show, this restricted data is not necessary for any of our main results. We only use the private data in this paper to show that identification based on public data is quite robust to concerns regarding unobserved hedonics.

The first proprietary dataset is from Rightmove, Inc. and spans 2006 to the present (Rightmove Inc. 2006–2022). The data includes information on the number of bedrooms, number of bathrooms, number of living rooms, floor area, property age, parking type, heating type, and property condition (rated as Good, Average, or Poor) of listed properties. It also includes rents for rented properties. These data must be purchased from Rightmove, however, our main analysis can be carried out without these data.

The second dataset of housing characteristics and rental prices is from Zoopla Inc., and is provided for free to researchers by the Urban Big Data Centre (2010–2021). This dataset also provides number of bedrooms, bathrooms, and rents. Additionally, it includes the number of floors and receptions of the property. We are able to match

[8] We exclude 0.02 percent of our transactions, which have implied negative lease terms at the time of transaction. We also exclude 0.6 percent of properties which are sold both as a leasehold and a freehold within our sample.

approximately 80 percent of transactions to the Rightmove and Zoopla listing data. Rental data is available for about 40 percent of properties.

B. *Lease Extension Details*

This section provides institutional details about the process through which leases can be extended. The distinction between freeholds and leaseholds dates to medieval England, during which permanent ownership of land and property, known as "freehold" ownership, was available only to feudal nobility.[9] During this time, leasehold estates were granted to serfs who would work the land for a set period of time and in exchange would pay a portion of the harvest to the freehold landowner. During the twentieth century, cash-poor landowners began to issue long leaseholds of 99 and 125 years, providing immediate liquidity without giving up ownership of the underlying land.

Today, leaseholds are very common in England and Wales, comprising 97 percent of all 2.5 million transacted apartments and 5 percent of houses. The freeholds underlying UK flats are typically owned by landed estates (e.g., the Cadogan Estates) which are privately managed, and other private landlords, developers, and investment companies. A very small proportion of these freeholds are owned by the Crown or the Church of England.

The lease length at the time of its issuance is referred to as "initial lease term," and the lease length at any future time is referred to as "remaining lease term." The distribution of leases can be divided into two groups; about 70 percent of leasehold flats in our sample are *short leaseholds* with remaining terms of 250 years or less and the other 30 percent are *long leaseholds* with remaining terms of 700 years or more. There are practically no properties with remaining terms between 250 and 700.[10] The most common initial terms for short leaseholds are for 99 and 125 years, which account for 77 percent of short leaseholds. The most common initial term for long leaseholds is 999 years, which account for 96 percent of all long leaseholds. Lease registration became mandatory in October 2003, so after this date, we capture all extensions of leasehold properties.

In principle, the freeholder and leaseholder can mutually agree to extend the existing length of a lease at any point in time for an agreed upon price. A 1993 Leasehold Reform, Housing and Urban Development Act (1993 Act) granted flat leasehold owners the right "to acquire a new lease" 90 years longer than the original lease, conditional on a one-off negotiated payment to the freeholder that reflects the market value of the gain in leasehold value from extension.[11] The legal recourse for lease extension is seldom used as it is costly for both parties to hire qualified surveyors, and the legal process can be lengthy. However the law does affect the

[9] The first known use of the term "freeholder" is in the Domesday Book published in 1086 under the reign of William the Conqueror.

[10] This distribution is illustrated in Supplemental Appendix Figure A.1.

[11] The 1993 Act also gives leaseholders the right to, upon extension, buy out the payment of future ground rents, which are annual payments to the freeholder. Ground rents are very small, however, with a median ground rent of £10 annually according to English Housing Survey data for 2009–2017 (UK Data Service 2022)

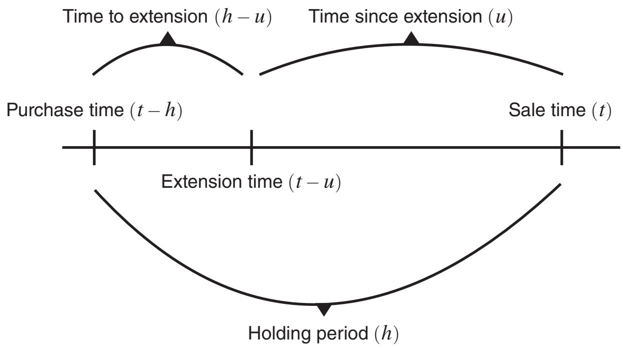

Figure 2. Diagram of Extension Time

Notes: The figure is a diagrammatic representation of the notation we will use in the paper. We say a property is purchased at time $t - h$, sold at time t and held for an amount h of years. We say that a property extends at time $t - u$, where $0 < u < h$.

outside option of negotiating parties in the event that negotiation breaks down. As such, tribunals can affect the price of extensions, even though they are rarely used.[12]

We introduce some notation in Figure 2 that will be useful in describing how we use the lease extension natural experiment to estimate y^*. The figure gives the example of a leasehold apartment that is first purchased at time $t - h$, sold at time t, and has its lease is extended at time $t - u$ between the two transactions, where $u < h$. These lease extensions, which are transacted twice—once before, and once after extension—form the "treatment" group in our estimation. We assume for illustrative purposes that the lease is extended for 90 years.

We denote the lease term to maturity, or lease duration, at purchase time as $T + h$ and its duration at sale time as $T + 90$ (notice that its duration would have been T at sale had the lease failed to be extended). We denote the transaction price of a property i of duration T at time t by P_{it}^{T}. The transacted prices before and after lease extension, and the lease duration before extension, will be the key inputs into our estimation. We do not observe the extension payment paid by leaseholders to freeholders.

As we see in Supplemental Appendix Figure A.2, properties with a very short holding period, h, are very likely to be "flippers" who purchase, renovate, add bedrooms, and sell properties with the explicit intention of making a quick profit. Since such properties may violate our exclusion restriction, as we will discuss, we exclude all lease extensions with a holding period $h \leq 2$ years from our analysis.

We restrict our sample to properties that were extended by more than 30 years, since leases that are extended by very short amounts may be unique. We also exclude properties which were extended with very short remaining durations, since we are focused on the long end of the yield curve. Finally, we remove outliers, defined as the 1 percent of data that experienced highest/lowest price growth over

[12] We download data on the universe of lease extension cases determined by tribunals since 1993, which are published by the Leasehold Advisory Service (https://decisions.lease-advice.org/). There are 4,420 cases since January 1, 2000, which is about 1 percent of extensions.

TABLE 1—NUMBER OF EXTENSIONS

Extension amount	90	700+	Other	Total
2000–2005	819	2,793	2,806	6,418
2006–2010	3,416	7,107	5,364	15,887
2011–2015	11,753	15,806	10,755	38,314
2016–2020	16,996	19,434	12,199	48,629
2021–2023	13,281	14,263	6,822	34,366
Total	46,265	59,403	37,946	143,614

Notes: The table reports the number of extended leases that have transaction data for each time period. The first column includes 90 year extensions, the next column includes 700+ year extensions, and the third column includes others, which are almost all non-90 under 200 year extensions. The first columns denotes the time of sale after extension.

the extension window relative to their controls. However, our main results are robust to including these properties.

Table 1 shows the distribution of lease extension length over time for our treatment pool of lease extensions. These are extended leases in our sample that have a transacted market price before and after lease extension. There are 143,614 such lease extensions. 90 years is the most common lease extension length, accounting for about 30 percent of all extensions, and there is also a high fraction of extensions that are for 700+ years. These extensions effectively convert short leaseholds into long leaseholds.

Supplemental Appendix A.4 describes how we can use the lease extension data to estimate the hazard rate of extension for a leasehold conditional of the number of years left in its duration. Figure A.3 plots this hazard rate and shows that almost no property gets extended with more than 90 years remaining in maturity. After a property hits 80 years remaining, its extension probability jumps to a probability of extension of about 5 percent, and then slowly falls back to 2 percent or so. Figure A.4 plots the implied cumulative hazard rate and shows that the cumulative hazard rate approaches one as duration of a leasehold approaches zero. In other words, a very tiny fraction of leaseholds expire without getting an extension.

We report a number of other descriptive statistics in the Supplemental Appendix for lease extension as well. Supplemental Appendix Figure A.5 shows that the median duration before extension for leaseholds getting extended by 90 years is around 70 years. Figure A.6 shows the median time between transactions is 10 years. Supplemental Table A.2 and Figure A.7 show that leasehold property owners are broadly similar to freehold property owners in terms of demographic and mortgage characteristics, as well as price-to-rent ratio over the business cycle. Supplemental Appendix Figure A.8 shows the geographical heatmap of number of extensions by geographical region in England and Wales.

C. *Transaction Lags*

Since we also provide real-time monthly updates of y^* on our project website, it is important to point out that there is a natural lag between when a house price is agreed upon by transacting parties versus when the transaction is actually recorded in our data. The date in the land registry relates to when the form to transfer the property is signed by both parties. However, the parties typically agree on a price several months earlier, and then undergo a process of finalizing mortgage and contract details.

We can estimate the amount of time between the date in which the buyer and seller agree on a price and the date in which the price is recorded by the land registry by using property listing data from Rightmove. If we assume that sellers stop posting house listings before agreeing on a price with a buyer, then the last date that a listing is posted must precede the date in which the buyer and seller agree to a price. The median amount of time elapsed between the last property listing on Rightmove and the date recorded by the land registry for transactions which have an associated Rightmove listing is 3.7 months, with a mean of about 5 months. The full distribution of the time elapsed between the listing date and the land registry date recorded is presented in Figure A.9. Therefore our real-time estimate is based on information with a lag of around five months.

III. Empirical Methodology

This section explains how to use the price gain from lease extension to estimate the expected long-term housing yield. We use the example of a lease that extends by 90 years as our illustrative example. The price of a leasehold P_t^T, with T years until expiration, and an option to extend by 90 years on expiration, is

$$(2)\quad P_t^T = R_t \int_0^T e^{-\int_0^s y(u)\,du} ds + \max\left[0, (1-\alpha) R_t \int_T^{T+90} e^{-\int_0^s y(u) du} ds + \ldots\right].$$

This equation starts with equation (1) that represents the present value of service flows from housing over the first T periods before the lease expires. The second-term represents the option value of additional extensions. $(1-\alpha)$ is the share of the price gain from extension going to the leaseholder, after deducting the negotiated payment to the freeholder and various costs that this negotiation entails. These terms multiply the present value of service flows from the lease, over the 90 year period after the extension. The ellipsis refers to the value of future extensions after $T + 90$, which have a similar structure. The max operator acknowledges that option value is nonnegative—instead of extending the lease, the leaseholder can choose not to extend and receives zero payoff.

Equation (2) clarifies that the option value of lease extension raises the value of a leasehold. Consider two cases. If $\alpha = 0$, the leaseholder receives the entire value of a lease extension, and $P_t^T = R_t \int_0^\infty e^{-\int_0^s y(u)\,du} ds$. With $\alpha = 0$, the price of a finite duration leasehold is the same as the price of its equivalent freehold (we abstract from uncertainty in the rate of return for simplicity).

On the other hand, if $\alpha = 1$, the leaseholder receives none of the value from extension, and the price of a leasehold is $P_t^T = R_t \int_0^T e^{-\int_0^s y(u)\,du} ds$. The service flows after T have no value to the leaseholder, since they go to the freeholder. With intermediate values $\alpha \in (0,1)$, the price of a T duration leasehold is between the duration T price and the infinite duration price.

What is the correct value of α for empirical analysis? UK law requires that leaseholders pay freeholders the entire value of lease extensions, that is, $\alpha = 1$ from a legal standpoint. While the leaseholder has the right to get a 90 year extension, they must pay the freeholder the full market value of the extension. In case the two parties cannot agree on that, the matter is settled in a tribunal. Supplemental Appendix A.5 provides full details on the process.

Supplemental Appendix A.5 also presents empirical estimate of α that exploits a discontinuity in the lease extension pricing function used by courts at $T = 80$. The analysis reveals that α is very close to one, especially in the early half of our sample when y^* is high. We also incorporate our empirical estimate of α in the estimation of y^* and show that it does make any material difference. We refer the reader to Supplemental Appendix A.5 for full details. Based on these results, the rest of the paper assumes $\alpha = 1$.

A. *A DiD-NLLS Estimator of y^**

We combine difference-in-difference and nonlinear least squares (DiD-NLLS) estimators to estimate y^*. First, we embed equation (2), the formula for lease extension price into a difference-in-differences estimator in order to identify y^*. We start with a difference-in-differences estimate, $\underline{\Delta}_{it}$, for the change in market value of leasehold i as a result of the lease extension:

$$\underline{\Delta}_{it} \equiv \left[\log P_{it}^{T+90} - \log P_{i,t-h}^{T+h}\right] - \left[\log P_{jt}^{T} - \log P_{j,t-h}^{T+h}\right]. \tag{3}$$

$\log P_{it}^{T+90} - \log P_{i,t-h}^{T+h}$ is the price change for property i bought h periods ago, which extends by 90 years from a duration of T to $T + 90$ years. We write $\underline{\Delta}_{it}$ with a double line to denote that it is a difference-in-difference. Property j is a suitably chosen control property, bought and sold in the same periods, with the same duration as property i prior to extension. Substituting the formula (2) for the price of a leasehold, with $\alpha = 1$, into the difference-in-differences estimator gives us

$$\underline{\Delta}_{it} = \log\left(\int_0^{T+90} e^{-\int_0^s y(u)\,du} ds\right) - \log\left(\int_0^{T} e^{-\int_0^s y(u)\,du} ds\right) + \Delta_{t,t-h}\left(\log R_{it} - \log R_{jt}\right). \tag{4}$$

In this equation, the first two terms represent discounting of the extended lease versus its control property. The final term is the difference between the growth rate of the service flow of housing, for the treatment versus the control group, over the length of the holding period.

Using equation (4) we can state our identification assumption. Estimating y from the difference-in-differences estimate $\underline{\Delta}_{it}$ requires a "parallel trends" assumption: The growth in service flows for the treatment and control properties, before versus after extension, must be the same; meaning prices for treated and control properties would have evolved similarly absent extension. If so, then the final term from equation (4) vanishes. The parallel trend assumption will fail if, for instance, properties that extend their lease are more likely to renovate than control properties. More generally, extending properties cannot have had different price growth from their controls, had they not extended. Section IIID presents placebo tests that probe the identification assumption.

In order to use equation (4) to estimate y^*, we need to parameterize the shape of the yield curve $y(s)$. We make the simple assumption that the yield curve is horizontal with $y(s) = y^*$. This assumption might appear extreme at first. However, as we show below, a key advantage of our difference-in-differences estimator is that it "differences out" differences in the yield curve at shorter maturities. The reason is that when T is large, the difference-in-differences estimator is primarily identified from long duration flows between T and $T + 90$. Thus the $y(s) = y^*$ assumption practically translates into the assumption that $y(s)$ asymptotes to y^* eventually. With this parameterization and the parallel trends assumption, our estimating equation becomes

$$\underline{\Delta}_{it} = \log\left(1 - e^{-y^*(T+90)}\right) - \log\left(1 - e^{-y^*T}\right). \tag{5}$$

Equation (5) can be estimated using nonlinear least squares (NLLS) to estimate y^*.

B. *Properties of the DiD-NLLS Estimator*

There are two important advantages of the DiD-NLLS estimator. First, it differences out the service flow of housing under the parallel trends assumption. As such none of the terms related to service flow appear in the estimation equation (5). This is a significant advantage because in practice it is extremely difficult to observe the true service flow of housing especially for owner occupied housing. Service flow of housing includes various hard to observe terms like taxes, depreciation and utility from consuming housing. The service flow may also vary across time and space for unobserved reasons. For instance, consider a temporary increase in demand for a narrow segment of London property. This shock does not affect the long-run return but does affect service flows and prices for certain properties in the short run—our estimator eliminates this variation.[13]

The second advantage of our estimator is that it differences out the effect of short-term yields on asset prices, and hence, isolates the long end forward of interest. We illustrate this feature of our estimator by numerically simulating a wide range of yield curves that all have the same long-term y^*, but differ greatly in the shape of yield curve over short to medium run horizon. The black solid line in Figure 3 presents one possible parameterization of the forward yield $y(s)$, where the forward yield curve $y(s)$ flattens out to y^* for $s \geq 40$ years, with y^* equal to around 4.0 percentage points.[14]

Given our parameterization of $y(s)$, we can solve for $\log P_{it}^{T+90} - \log P_{it}^{T}$ for all T. Then, for each T, we can solve for our estimator of y^* as a function of T numerically. The resulting values of our estimator, which we term $\hat{y}^*(T)$, are plotted in blue in Figure 3. Our estimator $\hat{y}^*(T)$ closely approximates the true long-run rate y^* for durations T after which the forward curve has flattened. We also plot the point

[13] As we show in Supplemental Appendix Section A.6, the estimator also differences out multiplicative taxes on property value, including the Stamp Duty Land Tax.

[14] We choose a flexible functional form $y(s) = \beta_1 - \beta_2 \cdot \beta_3^{-\beta_4(s-\beta_5)}$ and estimate the β parameters such that $y(0)$ is equal to spot rate on 1-year bonds, $y(10)$ is equal to the 10-year gilt yield, and the average of $y(s)$ for $10 \leq s \leq 25$ is equal to the 10 Year 15 Year gilt forward yield, averaged for the 2010–2020 period. For all the bond yields, we use the mean yield for our sample period. We present other possible parameterizations of $y(s)$ in Supplemental Appendix A.7.

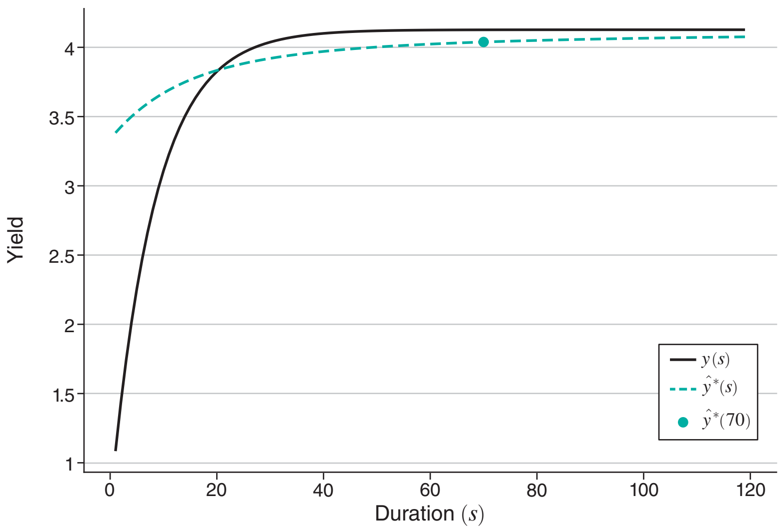

FIGURE 3. A PARAMETERIZATION OF THE HOUSING YIELD CURVE

Notes: The black line presents one parameterization of the forward yield curve, $y(s)$, which is chosen so that its shape matches the forward curve implied by the 1 year gilts, the 10 year gilts, and the 25 year gilts, averaged over the 2010–2020 period. $\hat{y}^*(s)$ is our estimator of y^* for each T, described below.

estimate of $\hat{y}^*$ at $T = 70$, which is approximately the median duration of leaseholds at extension. However there remains a small gap between $\hat{y}^*(T)$ and y^* even at long horizons. The reason is that in our difference-in-differences estimate (4), the effect of yields on price is nonlinear, and shorter-term yields do not entirely "cancel out" of the difference-in-difference estimator. However the effect of these shorter-term yields is small.

In fact, our estimator produces tight estimates for a wide range of yield curves. In Figure 4, panel A we present a time-varying yield curve for which the short end fluctuates tremendously over time but the long end (y^*) is constant. For instance, the yield curve labeled d is very downward sloping, whereas yield curve g is very upward sloping. Then, the solid line in Figure 4, panel B shows how our estimator $\hat{y}^*$ reacts to changes in the short end of the yield curve, where $\hat{y}^*$ is estimated for a lease with 70 years remaining. The points corresponding to each instance of the yield curve in Figure 4, panel A are labeled accordingly. For instance, point d in Figure 4, panel B corresponds to estimates of y^* for the downward sloping yield curve d of Figure 4, panel A. Our estimator is relatively stable despite the fluctuations in the short end of the yield curve and remains within 0.1 percent of y^*. This logic suggests that our estimator can successfully estimate the dynamics of y^*, even in the presence of volatile shocks to short-term rates.[15]

The reason why our estimator is able to provide a close approximation of y^* is because T is large, which effectively differences out most of the yield curve $y(s)$ for $s < T$. As T becomes smaller, the effect of the short end on $\hat{y}^*(T)$ increases, meaning estimates of y^* become increasingly biased. To see this, consider an alternative

[15] In Supplemental Appendix A.7 we show that we can use $\hat{y}^*(s)$ to bound the approximation error of our estimator, and therefore, get a lower and upper bound for y^*.

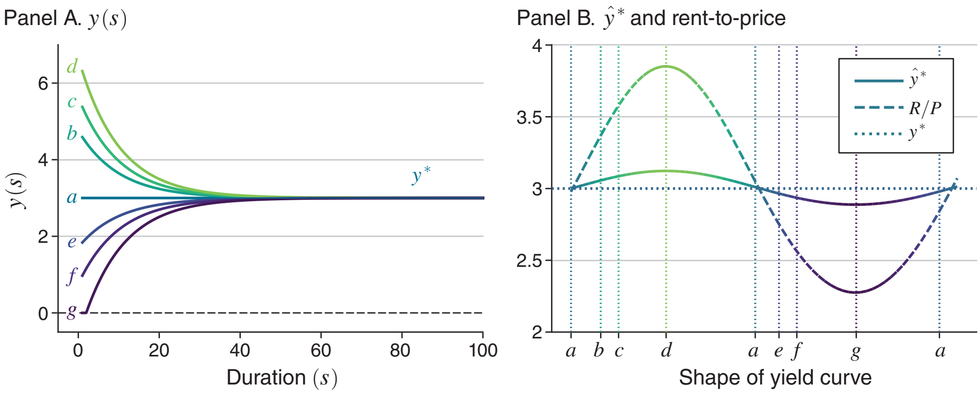

Figure 4. "Differencing Out" the Short End

Notes: The figure illustrates the effect of fluctuations at the low end of the yield curve on $\hat{y}^*$ and the rent-to-price rate, the rent-to-price ratio. Panel A presents several instances of the yield curve, $y(s)$ over time. Panel B indicates the estimates $\hat{y}^*$ and the rent-to-price ratio at each of these instances. For instance, point d on the right panel corresponds to the true value of y^*, the estimated value of y^*, and the rent-to-price ratio; given a yield curve d on the left panel. The rent-to-price ratio is estimated such that $P_{it}^{\infty}/R_{it} = \int_t^{\infty} e^{-\int_t^s y(u)du} ds \equiv \frac{1}{R/P}$.

estimator: the rent-to-price ratio of a freehold property, $\frac{R_{it}}{P_{it}^{\infty}}$. Like our estimator, the price-to-rent estimator cancels out the flow value of housing; it does not, however, difference out the short end of the yield curve and is therefore far more susceptible to changes in short-term forward rates. The dashed line in Figure 4 panel B indicates the value of the rent-to-price ratio, as the short end of the yield curve shifts. Changes in the short end of the yield curve affect the rent-to-price ratio by almost an order of magnitude more than they affect $\hat{y}^*(70)$. These results demonstrate that to effectively capture y^*, we must take the difference between *two* long duration assets; one does not suffice.

One important assumption in our simulations is that the yield curve is flat after 40 years. Therefore there is a unique expected housing yield at all sufficiently long horizons. Alternatively, the yield curve could be upward or downward sloping, even at long horizons in excess of 70 years. In the coming section, we present several pieces of evidence consistent with a flat long-term yield curve for housing.

Our approach does not require a particular structural model of why short run yields vary. Short run yields may fluctuate due to cyclical movements in housing risk premia, safe interest rates or liquidity conditions. Bubbles of the form studied by Harrison and Kreps (1978) also manifest in short run yields, provided that these bubbles disproportionately affect short duration valuations. Regardless, our approach differences out this short run volatility in order to estimate the expected long-term yield. Therefore our estimator does not require us to commit to a structural model of the economy, which might raise concerns about misspecification. Our estimator also differences out any variation due to "rational bubbles."[16]

[16] Our approach is related to how forward yields are calculated on financial assets such as zero coupon bonds, but is available at far longer horizons than are normally available for financial assets.

C. *Implementing the Estimator and Selecting a Control Group*

We now describe how to implement our estimator via nonlinear least squares and select controls. According to equation (5), for each individual property i our difference-in-differences estimator is

$$\underline{\Delta}_{it} = \log\left(1 - e^{-y_t^*(T_{it}+90)}\right) - \log\left(1 - e^{-y_t^* T_{it}}\right). \tag{6}$$

Here, we have generalized the expression of the estimator to allow a time varying expected long-term yield, and to acknowledge that the duration of the property before extension, T_{it}, can vary. Equation (6) shows that we can estimate y_t^* by nonlinear least squares. The estimator is valid at any point in time, hence we can estimate the dynamics of y_t^*. Two statistics inform y_t^* in the estimator. First, the difference-in-difference $\underline{\Delta}_{it}$ can be calculated for every property i, as the difference in price growth between the extending property and its control. Second, the covariance between $\underline{\Delta}_{it}$ and the duration before extension T_{it} also helps to identify y^*. Regarding inference, we cluster standard errors at the level at which treatment is assigned, following standard practice (Abadie et al. 2023). Since treatment is assigned at the level of each extending property i, heteroskedasticity-robust standard errors without clustering suffice.

We select a control group separately for each extending property, from neighboring properties of a similar duration that did not extend during the purchase and sale transaction window of the treated property. Selecting controls presents a challenge. Ideally, one selects a control that is bought and sold at the same time as the treated property. However this procedure reduces the effective sample size, because many extending properties do not have neighboring controls that are bought and sold simultaneously. We therefore use repeat sales methods to expand our set of controls. For each extending property, we measure its counterfactual price growth using a repeat sales index of similar duration neighboring properties that are bought and sold at similar times to the extender. The repeat sales index is calculated separately for every extending property, using the individual property's sample of control observations. We only select control properties that do not extend, avoiding the "forbidden comparisons" problem (Borusyak, Jaravel, and Spiess 2021).

To construct the repeat sales control, we consider a treated property i, which was purchased at time $t - h$, sold at time t, and extended for 90 years at some time $t - h < t - u \leq t$.[17] Suppose this property has duration $T + h$ at purchase and duration $T + 90$ at sale. The set of control properties, specific to property i, is those properties that (i) do not extend between $t - h$ and t; (ii) have a duration within five years of property i; and (iii) are within d km of property i. d is the smallest possible Haversine distance such that it is feasible to construct a repeat sales index for the counterfactual price of property i between $t - h$ and t. We discard property i if there are not enough controls within a 20 km radius.[18] This procedure

[17] In our primary analysis, we measure t at a quarterly frequency. We show robustness to using other frequencies in Figure A.23.

[18] In Supplemental Appendix Figure A.10 we show the distribution of radii. Nearly 80 percent of experiments have a radius under 5km. In Supplemental Appendix Figure A.11 we show that the radius is typically smallest in dense urban areas.

automatically adjusts for the different density of housing in urban and rural areas. In Supplemental Appendix Figure A.11 we verify this by showing that the control radius is strongly correlated with local population density, and is always small for very dense areas (ONS 2021b). Finally, we produce a repeat sales index for property i's control group, using a three-part procedure similar to Case and Shiller (1989) (see Supplemental Appendix A.8 for details). In Supplemental Appendix A.9, we show that these controls are similar to their extended counterparts on all of the main hedonic measures we observe in the Rightmove and Zoopla data. We sometimes denote the set including property i and its control properties as experiment i.

There are two advantages from using repeat sales methods. First, the identification assumption underlying our difference-in-differences estimator is essentially unchanged. Between $t-h$ and t, the growth of service flows for the treated property must be the same as for the control properties that constitute the treated property's repeat sales index. Second, the repeat sales index method lets us construct controls for many more extending properties. With this method, we are able to construct controls for 130,971 of our 143,614 lease extension experiments.[19, 20]

D. *Validating the Identification Assumption*

Our identification assumption is parallel trends: Growth in the service flow of housing should not differ for extending properties and controls on average. This assumption fails if those who extend the lease are also more likely to improve other characteristics of the property, for instance through renovations. We provide some indirect tests of the parallel trends assumption in this section.

Why are leaseholds extended in practice? It should be kept in mind that as we have shown, the cumulative hazard rate calculation tells us that ultimately all leaseholds are extended before expiring. So our D-in-D estimator is taking advantage of the fact that some happen to be extended earlier than others. The main reason why people extend leases is that it is a way to accumulate savings in the property, and resell the property at a higher value. For example, according to Leasehold Advisory Service,[21] a government-funded source of legal advice, the main reason to extend is that "as the lease gets shorter and the number of years goes lower, the value of the lease decreases and it becomes more expensive when you extend the lease. This is why it is often a good idea to increase the term of the lease especially if you want to sell the property." This motive for extending leases is consistent with parallel trends.

In any event, we take the endogenous renovation concern seriously and provide two different tests for checking the parallel trends assumption.[22] As a first test of parallel trends, which directly focuses on renovations, we show that extending properties are no more likely to make home improvements than controls. First, we identify

[19] To provide intuition for why a repeat sales index expands the available controls, consider an extending property bought at $t-10$ and sold at t, with two neighboring properties: the first bought at $t-10$ and sold at $t-5$; the second bought at $t-5$ and sold at t. Using both properties one can create counterfactual price growth for the extending property.

[20] In robustness, we will also estimate y^* using only the subset of extending properties with controls that are bought and sold at the same time.

[21] https://www.lease-advice.org/faq/why-should-i-extend-my-lease/

[22] Additionally, in Supplemental Appendix A.9 we provide evidence that extended and control properties are similar on observable characteristics.

properties that were recently renovated based on textual analysis of Rightmove listing descriptions. In particular, we create a renovation dummy that is equal to one if a property's listing contains certain key words associated with renovation such as "renovated," "refurbished", or "improved." About 9 percent of all listings make reference to renovation. To verify that this is a valid measure of home improvement, we regress it against change in bedroom count for the renovated flat from the listing before renovation to the listing after renovation. The regression coefficient is positive and statistically significant; the probability that a property's bedroom count increases is more than twice as high if the property was marked as renovated in that year. Moreover, properties that were purchased and immediately resold ("flippers"), have much higher renovation rates and are far more likely to increase their bedroom count, as shown in Supplemental Appendix Figure A.2. This is suggestive that both our text-based measure and the change in bedroom count are good indicators of renovations.[23] Our other main hedonic measures—bathroom count, living room count, and floor area—experience nearly zero growth across all flats, which is consistent with the United Kingdom's strict regulation of construction.

In Table 2, we present the difference in the renovation rate of extended properties relative to the average renovation rate for their controls. The outcome variable in column 1 is our text-based measure of renovation. The outcome variables in columns 2–5 are the change in hedonic characteristics for properties that have two distinct listings before and after extension time. We exclude cases where the implied change in housing characteristics is negative, which are most likely data entry errors. In all specifications, we find no evidence that extended properties renovate at a different rate than other properties around extension episodes.

As a second test of parallel trends, we study the behavior of market rents before and after extension. Rents are only available for a subset of properties, however they are an observable proxy for the service flow, which allows us to assess parallel trends relatively directly. Consistent with our identification assumption and with the absence of excess renovations for treated properties, we find that rents evolve similarly for the control and treatment group.

For each experiment i, we consider treated and control properties for which there are at least two rental transactions recorded by Rightmove or Zoopla between the purchase date $t-h$ and the sale date t of the property. We study a regression

$$\Delta \log R_{i,j,t',t''} = \alpha_{i,t',t''} + \beta \mathbf{1} \times \text{Extension}_j + \epsilon_{i,j,t',t''}, \tag{7}$$

where $\Delta \log R_{i,j,t',t''}$ is growth in rents for property j between periods t' and t'' for which there are observed rents for both the treatment and control, $\alpha_{i,t',t''}$ are experiment $\times$ rental transaction years fixed effects. We control for whether the property was transacted between rental listings, since properties are more likely to experience greater rental price growth when they change owners. The regression measures whether rental growth was higher for extended properties than for control properties

[23] Both of these measures are also positively correlated with change in both sale and rental price. Therefore, if extended properties were more likely to renovate, our identification assumption would not hold. Fortunately, as we will see below, extended properties are no more likely to renovate by either measure, and so our identification assumption holds.

Table 2—Renovation in Extended versus Nonextended Flats

	Renovation rate (1)	Δ Bedrooms (2)	Δ Bathrooms (3)	Δ living rooms (4)	Δ floor area (5)
Extension	−0.001	−0.002	0.000	−0.001	−0.002
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
Experiment fixed effects	✓	✓	✓	✓	✓
Control mean	0.091	0.04	0.002	0.001	0.022
Observations	153,400	43,612	33,078	31,934	34,564
N. experiment	76,700	21,806	16,539	15,967	17,282

Notes: The table reports renovation rates for extended properties relative to their nonextended counterparts for all experiments. For each experiment we have two observations indexed with j: the renovation rate for the extended property, and the mean renovation rate across all control properties. We run the following regression, $X_{ij} = \alpha_i + \beta \mathbf{1}(\text{Extended}_j) + \epsilon_{ij}$, for renovation measures X_j where α_i are experiment fixed effects. The renovation rate for treated properties in column 1 is based on the post-extension sale listing. The renovation rate for columns 2–5 is based on change in hedonic characteristics before and after extension time. Standard errors are heteroskedasticity robust. Standard errors in parentheses.

that have rental transactions in the same years. Table 3 presents the results and finds small and insignificant differences in the behavior of rents between the treatment and control group. The first column is the baseline regression.

Column 2 of Table 3 collapses the rental growth comparison between extended and nonextended properties at the experiment level, as in Table 2. We relax the restriction that control and extended properties must have listings in the same years and compare mean annualized rent growth for extended properties and their nonextended controls for each experiment. Thus each experiment has two observations and receives equal weight.

Column 3 reestimates regression equation (7), where growth in control rents is now given by a repeat sales index of rent growth in each extending property's control group. Each observation is a rental cycle for which we have two extended property rental transactions, and we are able to construct a control repeat sales rent growth index. On average, there are approximately two rental cycles within each purchase-sale window. The resulting estimates are small and statistically insignificant.[24]

The analysis above confirms that rental growth is no different between extended properties and nonextended properties over the experiment window. However, the identification assumption in equation (4) also requires that rent growth for extended and control properties is expected to be the same beyond the experiment window. We therefore further test if the long-run growth rate of rents, g^*, is the same for treatment and control properties.

We assess the assumption regarding long-run rental growth by expanding the window over which we study rent growth. Figure 5 plots annual rent growth for extension and control properties against time. We plot annual rent growth starting from the midpoint between the purchase and sale time of the extending property. We construct rent growth for the control properties associated with each extending property, using our baseline repeat sales method. We interpolate across years with

[24] There may be concern that the rental data is stale. However, we find that over 80 percent of our rental pairs experienced a change in the rental price.

TABLE 3—WITHIN-EXPERIMENT WINDOW RENT GROWTH

	Δ log(Rent)		
	(1)	(2)	(3)
Extension	0.0007	−0.0006	−0.0003
	(0.0006)	(0.0006)	(0.0007)
Experiment × rent years fixed effects	✓		✓
Experiment fixed effects		✓	
Annualized		✓	
RSI			✓
Observations	3,943,474	36,410	74,266
N. Experiment	16,569	18,205	18,887

Notes: The table reports rent growth during the experiment window of $t - h$ to t for extended properties relative to control properties. The first column is the estimated coefficient from equation (7). We use mean annualized rent growth as the LHS variable, and do not control for time fixed effects. The third column uses a rent repeat sales index. Standard errors are heteroskedasticity robust. Standard errors in parentheses.

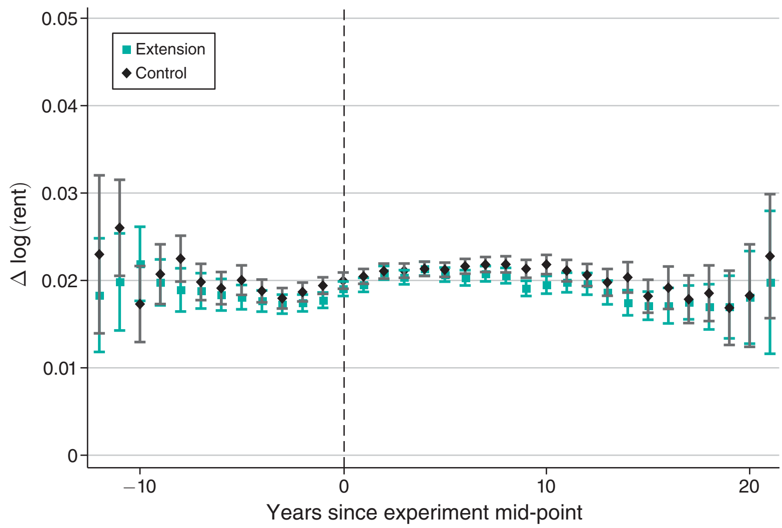

FIGURE 5. RENT GROWTH AT LONG HORIZONS

Notes: The figure presents annual rent growth at all horizons for treated properties relative to a repeat sales index of rental prices for controls. We interpolate across years when rental listings are not consecutive and plot rent growth against time since the experiment window. The x-axis variable is the mid-point between the purchase and sale of the treated property, $(t + (t + h))/2$. Error bars represent 95 percent confidence intervals.

missing rent data.[25] The results show that rent growth is similar for extending and control properties at all horizons including the long term, and is never statistically different. The results in this section should be reassuring regarding the identification assumption: Extended properties do not have a higher propensity to get renovated, and their rental growth is no different from control properties.

[25] For instance, suppose we observe rents R_{t-5} and R_t for a property in years $t - 5$ and t, without data in between. We assume that rent growth equals $(1/5) \times (R_{t-5}/R_t)$ in all intervening years.

IV. Empirical Results

A. *Event Study Analysis of the Effect of Lease Extension*

We develop intuition for our estimator of y^* through an event study representation of the methodology. Consider the modal leasehold extension in our sample that extends its duration from T to $T + 90$ years. These leaseholds have a median duration at sale of 157 years, meaning that had they not extended, they would have had a duration of 67 years at sale.[26] The difference-in-differences estimate $\underline{\Delta}_{it}$, from equation (3), represents the gain in log market value at time t, of a leasehold that extends u periods earlier.

In Figure 6, we plot in blue the estimate $\underline{\Delta}_{it}$ against time between extension and sale, u. For instance, the first entry on the positive x-axis is the mean DiD estimate of the increase in market value for leaseholds that sell one year after extension. One year after extension, the difference in log price between extended properties and their controls jumps by about 0.1 log points. Afterward, the difference continues to grow exactly as predicted by asset pricing equation (2), so that leaseholds that sell 10 years after extension increase their market value by almost 0.2 log points. The increase in market value rises with the time since extension because (i) the value of the control leasehold that did not get extended falls more for each passing year due to its shorter duration, which varies nonlinearly with log price, and (ii) y^* falls over our sample period. The dashed gray presents the post-extension price gain at each value of u predicted by our simple asset pricing model, given the mean duration and estimated y^* level for properties transacted in that period. The predicted trend aligns closely with the observed estimates of $\underline{\Delta}_{it}$.

We also use the event study to investigate pre-trends. On the negative x-axis, we apply the difference-in-differences estimator to the pre period—that is, to the price growth of treated versus control properties over the two transactions prior to the extension. In doing so, we can estimate whether there are pre-trends in price growth before extension. We plot the difference-in-differences estimate in the pre period against the number of years between the final transaction before the extension, and the extension itself. There are small and statistically insignificant pre-trends—extended properties experience similar growth to controls in the pre-period before extension, regardless of the number of years between transaction and extension.[27]

B. *Estimating Average y^**

The difference-in-differences estimate, $\underline{\Delta}_{it}$, of the gain in market value as a result of duration extension is used in the nonlinear equation (6) to estimate y^*. Figure 7 graphically illustrates how the estimation procedure works, using 90-year extensions as example. The figure binscatters $\underline{\Delta}_{it}$ estimates for various durations T_i. The dashed black line is the best-fit prediction from our NLLS estimation of equation (6). Figure 7 shows that the DiD estimate of the increase in market value after extension aligns

[26] Supplemental Appendix Figure A.13 displays the histogram of remaining lease term at sale for the extended and control properties.

[27] Figure A.17 presents the event study for only 90-year extensions.

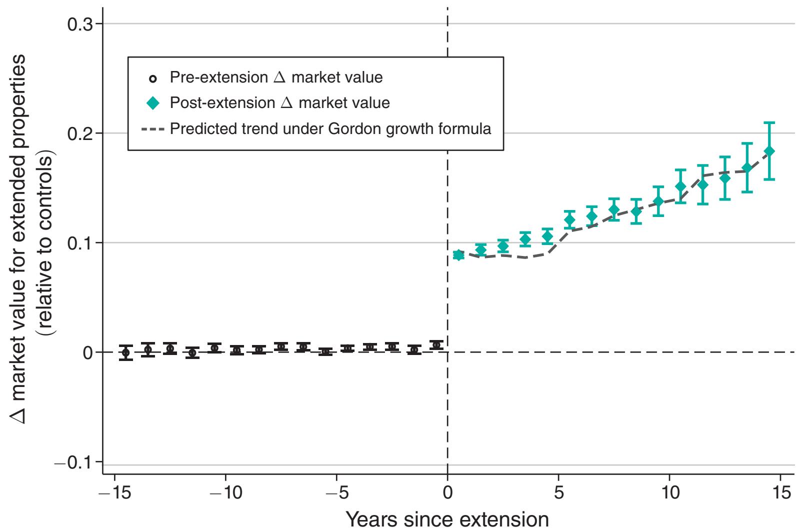

FIGURE 6. EVENT STUDY REPRESENTATION OF LEASE EXTENSION

Notes: The figure presents an event study representation of a lease extension. On the positive x-axis we plot in blue diamonds the mean difference in price change between time $t - h$ and t, $\underline{\Delta}_{it}$, against time since extension, u. On the negative x-axis we plot in black circles the mean difference in price change between time $t - h - h'$ and $t - h$ for extended versus control properties. $t - h - h'$ is the time of the nearest transaction before $t - h$. The dotted line plots the predicted price gain based on mean duration and y^* at each time. The sample includes all extended properties. 95 percent confidence intervals are shown.

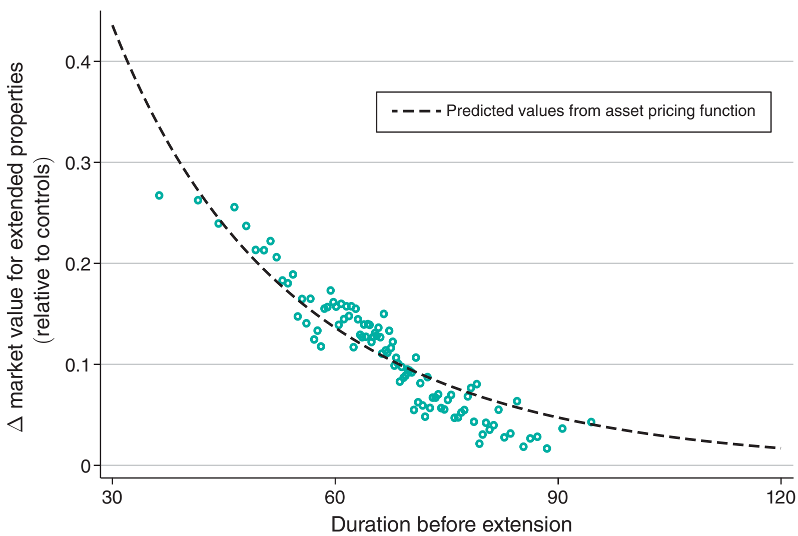

FIGURE 7. DURATION BEFORE EXTENSION VERSUS PRICE GAIN AFTER EXTENSION

Notes: The figure is a binscatter of our difference-in-difference estimator against duration before extension, T, with 100 bins. The sample includes leases that were extended for 90 years. The black line shows fitted estimates of equation (6).

remarkably well with the predicted increase according to the simple asset pricing formula of equation (6). Our simple asset pricing formula predicts that the percent valuation gain from a 90 year extension should be larger for extensions with shorter

Table 4—Estimated y^*

	Constant y^*	Flexible y^*				Constant y^*
	$k = 90$ (1)	$T = 50$ (2)	$T = 60$ (3)	$T = 70$ (4)	$T = 80$ (5)	$k \geq 700$ (6)
y^*	3.51	3.55	3.51	3.48	3.44	3.50
	(0.023)					(0.019)
Observations	42,429					54,581
t-stat (700+ versus 90)						−0.32

Notes: Column 1 presents estimates of y^* from equation (6) for all extensions. Columns 2–6 present estimates of $y^*(T+k)$ for the range of durations in our sample, where we parameterize y^* linearly as a function of $T+k$, $y^*(T+k) = \alpha + \beta \cdot (T+k)$. We estimate y^* using a version of $\underline{\Delta}_{it}$ that has been residualized on year × local authority fixed effects. Standard errors are heteroskedasticity robust. Standard errors in parentheses

duration at the time of extension. Consistent with the formula, the percent gain in property value as a result of extension is decreasing in the duration before extension—ranging from only 7 percent for properties with 90 years remaining at extension, to more than 30 percent for properties with duration of 40 years at extension.

More generally, there are three sources of variation used to identify y^* according to our NLLS DiD estimation procedure. First, for a given remaining duration T at extension, and a given extended duration $T+k$, the valuation gain $\underline{\Delta}_{it}$ is larger when y^* is lower. Second, as shown in Figure 7, the gain from extension increases as the duration before extension T falls. Moreover this increase is stronger when y^* is lower.[28] Third, the valuation gain is larger when the size of the lease extension amount k is bigger, with a larger gain when y^* is lower.

Table 4 presents the y^* estimates implied by Figure 7. Column 1 shows that y^* estimate for properties extended by 90 years is 3.5 percent over the full sample period. As discussed in Section IIIB, our estimator assumes that the long-term forward yield curve for housing is flat, meaning a unique value of y^*. We now test if the assumption of a flat long-term forward yield curve is accurate, using two sources of variation: in the duration T at which a leasehold is extended, and the size k of the increase in lease duration at extension.

We estimate the degree to which estimates of the long-term yield vary with duration, by parameterizing the estimate of y^* as $y(T) = \alpha + \beta \cdot T$, where T is the duration of the leasehold before extension. If the long-term yield curve is sloped, then β will be different from zero. We exploit variation in duration after extension to estimate both α and β. We control for time of experiment by residualizing $\underline{\Delta}_{it}$ on year fixed effects before estimating y^*, to make sure spurious variation in when certain extensions are popular does not drive our result. Columns 2 through 5 of Table 4 present estimates of $y(T)$ under this more flexible form. The results show that $y(T)$ varies little over the range we observe it, even at very long durations, suggesting a flat long-term yield curve.[29]

[28] We can explicitly reject a linear relationship between $\underline{\Delta}_{it}$ and T_{it} by running the following quadratic regression: $\underline{\Delta}_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 T_{it}^2 + \epsilon_{it}$. The estimated β_i's are all statistically significantly different from zero.

[29] In Supplemental Appendix Figure A.18 we estimate the event study of Figure 6 separately for above and below median duration properties, which also documents the lack of pre-trends for both.

As another test of the slope of forward yield curve, we estimate y^* using the sample of leaseholds that extend by more than 700 years. These leaseholds are affected by much longer horizon yields than the baseline sample, which extend by 90 years. Column 6 of Table 4 presents estimates of y^* using only extensions of more than 700 years. Estimates of y^* are remarkably similar for both groups, and statistically indistinguishable according to a *t*-test, the *t*-stat of which is presented at the bottom of column 6. Therefore y^* seems to vary little at very long horizons, consistent with a flat yield curve.[30,31]

One concern about our estimate may be that short duration leaseholds have an additional liquidity premium, because banks might be less willing to issue a mortgage against shorter duration leaseholds. Supplemental Appendix A.10 investigates this concern formally and finds little evidence of liquidity effects in our sample. For example, data from the English Housing Survey shows that shorter duration leaseholds are as likely to have a mortgage, with similar mortgage characteristics in terms of maturity, loan-to-value ratio, and interest rate. Shorter duration leaseholds in our sample are also no different in terms of trading liquidity (e.g., time on the market). Finally, incorporating discontinuities around the leasehold durations where mortgage lending might change does not affect our estimates of y^*.[32,33]

C. *Estimating Dynamics of* y^*

A key advantage of our data and methodology is that we have new leasehold extension natural experiments every month—in recent years close to one thousand per month. We can thus estimate the dynamics of y^* over a period of more than two decades. Long-run dynamics of y^* should be of particular interest to scholars studying questions such as secular stagnation, movements in r^*, and similar.

There is a clear decline in y^* over our sample period. Figure 8 illustrates this by plotting the average gain in market value, $\underline{\Delta}_{it}$, for 90-year extensions against duration at extension, T, for three different time periods. The opacity of bars in the bar graph reflects the number of observations behind each bar. The gain in market value due to 90-year extension of a leasehold has increased by about 10 percentage points between the beginning and end of our sample on average. This is a large increase, which reflects a significant fall in y^* over our sample period.[34]

We now move to the full data, with all extensions, so we can estimate the full dynamics of y_t^* with precision. In total, we have 130,971 leasehold extension natural experiments between 2000 and December 2023. Since there are fewer natural

[30] In Supplemental Appendix Figure A.16 we present a binscatter using properties that were extended for more than 700 years.

[31] The estimate of y^* is not affected by "buy to let" properties. Supplemental Appendix Figure A.14 shows that y^* is similar for buy to let and owner occupied properties; Supplemental Appendix Figure A.15 shows that at the local authority level, y^* is uncorrelated with the share of second homes. Data on second homes is provided by the ONS (2021a).

[32] In Supplemental Appendix A.11, we find that our estimate of y^* is not seasonal, consistent with its interpretation as a long-run object.

[33] Since our estimator constructs a repeat sales index control for every treated observation, there is variance due to "generated regressors" that our baseline standard errors do not account for. In Supplemental Appendix Section A.12 we correct the standard errors using a bootstrap procedure and find that the effect is small.

[34] In Supplemental Appendix Figure A.19 we plot the event study of Figure 6 separately for each period, which also documents the lack of pre-trends at all times.

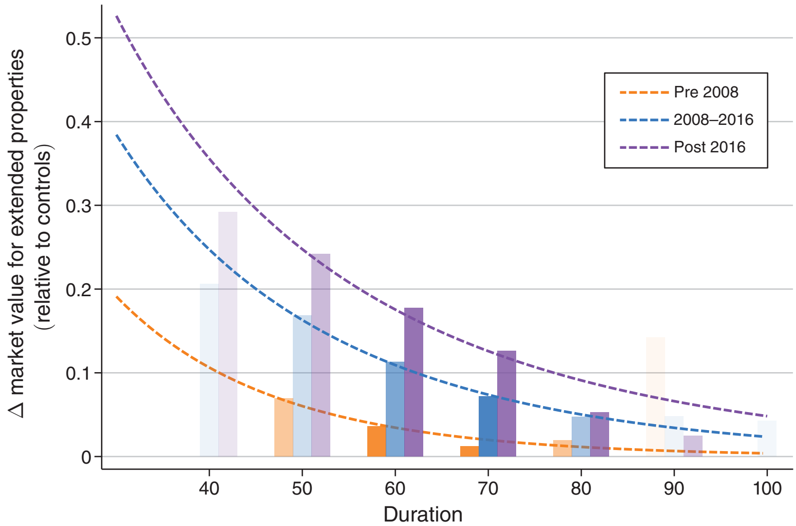

Figure 8. Price Change from Extension over Time

Notes: The figure shows the mean difference-in-difference estimate for various durations in the pre-2008, 2008–2016 and post-2016 periods. The bars show the level of $\underline{\Delta}_{it}$ for each bin and the dashed lines shows fitted estimates. The sample includes only 90-year extensions. Bars are shaded proportionally to the number of observations that make up the bar.

experiments in the beginning of our sample period, we pool the experiments together for the first four years and estimate a single y_t^* for 2000–2003. From then on, we estimate y_t^* separately for each year. Since lease registration became mandatory after 2003, we capture all lease extensions after this time.

Figure 1 in the introduction presents the estimates of y_t^* at annual frequency. The shaded region reflects 95 percent confidence interval. The estimates get more precise over time as the sample size gets larger.[35] y^* is relatively stable around 5.2 percent from 2000 until 2006, when it starts to fall persistently. In total y^* falls from around 5.2 percent to 2.7 percent in 2024, an almost 50 percent decline. The magnitude of this decline is large, corresponding to a doubling of the long-run expected price-rent ratio. Notably, y^* estimate remains stable during the 2020 pandemic, despite considerable volatility in shorter-term asset prices during this period.[36]

D. *Real-Time Dynamics of* y^*

As mentioned earlier, one useful feature of our methodology is that it can be implemented in real-time, with publicly available data. Figure 9 plots our estimate of the expected long-term yield of housing, at monthly frequency until December 2023, the last month for which UK housing data are currently available. The shaded

[35] Supplemental Appendix Figure A.20 plots the time varying estimate y_t^*, separately for short and long extensions.

[36] After 2017, various governments have proposed reducing leasehold extension costs. How would reform affect our estimate of y^*? In Supplemental Appendix Section A.5.7 we show that expectations of a leasehold reform would lead our estimate of y^* to rise, meaning the true decline in y^* may have been greater than our estimate.

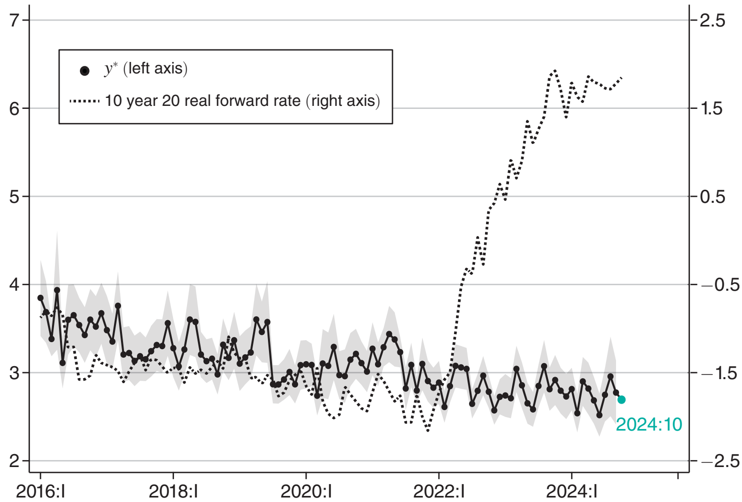

Figure 9. Monthly Estimates of y^* and Rate of Return on Government Bonds

Notes: The solid line presents y_t^* estimated at a monthly frequency using all extensions. The line is plotted from January 2016 to October 2024. 95 percent confidence interval is shaded.

area represents the 95 percent confidence interval of our estimate of long-run rates and shows how precise our standard errors are, even when estimating y_t^* at a monthly frequency. The tight standard errors result from having a reasonably large sample size of around 900 lease extensions per month. As discussed in Section IIC, y_t^* estimates include new information with a lag of about five months.

Two observations from the monthly time-series are worth mentioning. First, there is a barely noticeable change in y^* during the COVID-19 recession, despite considerable volatility in housing markets over this period. This behavior is consistent with the COVID pandemic representing a large but short-term shock to housing markets—which, reassuringly, our measure of long-term expectations is able to "look through." In the future, our estimates should help policymakers to understand in real time whether other shocks have also affected long-run expectations.

Second, the recent post-pandemic tightening cycle of monetary policy starting in January 2022 has had a large effect on real yields on safe assets—even at relatively long durations—but has not affected y^*. The dashed line in Figure 9 indicates that the 10-year-20 real forward rate on government bonds has risen by about 350 basis points since the beginning of 2022. In contrast, y_t^* has remained stable, though it is important to remember that these estimates include a lag of about five months.

E. *Robustness Checks*

We have made our data public, and presented the estimate of y^* in as transparent a manner as possible in the graphical analysis above. The market is remarkably precise in estimating the very long forward that y^* represents. For example, the market increasingly values the gain from extension when lease extension happens with fewer years until maturity. There is also a clear rise in market value gain due to extension in the second half of our sample, reflecting a sharp fall in y^*. This section provides additional robustness checks for our estimates.

The DiD estimator uses a weighted repeat sales index as in Case and Shiller (1989) to estimate price growth for control properties. An alternative would be to use the simple repeat sales index introduced by Bailey, Muth, and Nourse (1963) (BMN). Supplemental Appendix Figure A.23 presents in the black line our baseline estimates and in the dashed gray line the estimates produced with a BMN repeat sales index. The dotted blue line presents estimates when we include control properties with holding periods under two years ("flippers"). The main repeat sales index is produced at a quarterly frequency. The dotted green line presents estimates when we calculate the repeat sales index at a yearly frequency. All four lines are very similar and a suggest an overall decline in y^* of about 2.4 percentage points.

Instead of estimating price growth of control properties between time $t - h$ and t via a repeat sales method, another option is to use controls that are also purchased and sold at times $t - h$ and t. We denote this the "exact control" method. The main disadvantage of the exact control method is that it imposes a much stricter criteria, so fewer extensions are matched, and the controls which are matched tend to be further away from the treated properties. Supplemental Appendix Figure A.24 presents estimates of y^* using the exact control method in the dashed gray line. The black line presents estimates using the baseline methodology on the sample of experiments for which we identify exact controls. The two lines are again very similar.[37]

Finally, we may worry that the observed time-trend is driven by changes in regional composition of extensions over the sample period. To show that this is not the case, we recalculate y^* using weights that fix the composition of the local authorities in the data at (i) the average over the full period and (ii) the end of period distribution. The results are presented in Supplemental Appendix Figure A.26 and are similar to the baseline estimates.[38]

F. *The Advantage of Within-Property Variation in Duration*

Our microdata-based approach to estimating y^* builds on Giglio, Maggiori, and Stroebel (2015), who first noted that the varying durations of UK properties make them ideal for estimating long-term housing yields. They use cross-sectional comparisons of freeholds and leaseholds to estimate the level of y^*. We extend this insight by leveraging the quasi-experiment nature of lease extensions to estimate the dynamics of y^*. The quasi-experimental approach has some statistical advantages over the approach that uses cross-sectional variation.

A key distinction between the cross-sectional and quasi-experimental approaches is the source of variation in duration. The former relies on differences across properties, while the latter exploits within-property variation. Long-duration properties

[37] Further, in Supplemental Appendix Figure A.25 we repeat this analysis but restrict to treated properties for which a control can be found within 1km. We do this because under the exact control method, the average control radius increases, since it is more difficult to find neighboring controls. Again, the results are similar when we restrict to controls within 1km.

[38] One question is whether the tax treatment of different kinds of properties interacts with our result. In Supplemental Appendix Figure A.14, we show that estimates of y^* are similar for "buy-to-let" and owner-occupied properties; and in Supplemental Appendix Figure A.15 we show that estimates of y^* are uncorrelated with the ownership of second homes.

may have different service flows—for instance, as freehold flats might be of higher construction quality. The cross-sectional method depends on hedonic controls to account for price differences due to potential differences in service flow. The quasi-experimental approach uses variation in duration within the same property due to lease extensions, and relies on the "parallel trends" assumption highlighted earlier.

To gauge how unobserved heterogeneity affects the quasi-experimental and the cross sectional estimates, we study a test in the spirit of Altonji et al. (2005) and Oster (2019). According to these papers, one sign of bias from unobserved heterogeneity is that different ways of controlling for *observed* heterogeneity matters. We therefore estimate y^* using both the quasi-experimental and the cross-sectional approach, and control for observed heterogeneity in different ways. We therefore estimate y^* using both the quasi-experimental and the cross-sectional approach, controlling for over 100 different variations of hedonic characteristics. In one variation we do not include any controls. In another, we allow price to vary linearly with number of bedrooms and floor area, and in another we allow price to vary quadratically with these same controls.[39] In the most extreme case, we control for fixed effects of the following seven characteristics: number of bedrooms, number of bathrooms, floor area, year built, heating type, property condition rating, and availability of parking.[40] The other variations include all possible subsets of these seven characteristics. Note that certain specifications, such as the one without any controls, are unlikely to yield plausible estimates of y^* when using a cross-sectional methodology. Nevertheless, we include them deliberately to transparently document the full sensitivity of the estimators to variations in observed heterogeneity. Importantly, substantial variation is also observed among specifications considered plausible—such as comparing linear controls with fixed effects specifications—highlighting the estimator's sensitivity even within realistic modeling choices.

We then plot our estimates of y^* for each specification in Figure 10, for both the quasi-experimental and cross-sectional methodologies. Under the cross-sectional approach, the estimates vary tremendously from 1.78 percent (in the case of the Giglio, Maggiori, and Stroebel (2015) specification) to more than 10 percent in the case of controlling for fixed effects. In contrast, our quasi-experimental estimates of y^* are stable around the baseline estimate. Therefore the form of observed heterogeneity is not very important for the quasi-experimental approach—suggesting that estimates are somewhat robust to unobserved heterogeneity only with this approach.[41,42]

[39] Supplemental Appendix Figure A.21 indicates that most hedonic controls vary relatively linearly with log price.

[40] The fixed effects controls are the same as the main specification Giglio, Maggiori, and Stroebel (2015). Giglio et. al. add an indicator variable for properties with missing hedonics and includes them in the main sample—whereas the current exercise restricts the sample with controls to properties that have hedonic characteristics. We remove properties without hedonics because these properties will not be affected by different ways of adding controls. Moreover, if controls are important, then including properties with missing control information may lead to omitted variable bias.

[41] The standard errors for the cross-sectional estimates range from 0.002 to 75.84, and the standard errors for the quasi-experimental estimates range from 0.013 to 0.028. Therefore in a statistical sense, the cross-sectional estimates are significantly less stable.

[42] The Giglio, Maggiori, and Stroebel (2015) specification and the fixed effects specification differ because including observations with missing hedonics is similar to leaving out hedonic controls entirely.

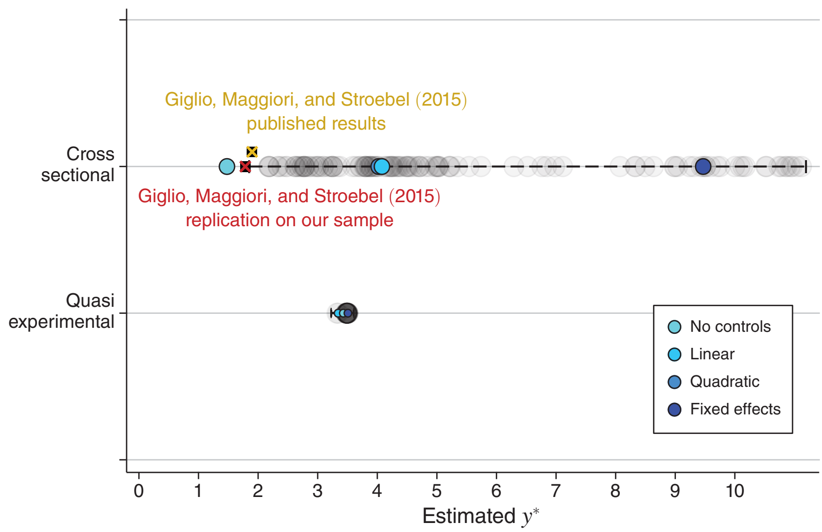

Figure 10. Stability of y^*: Controlling for Observables

Notes: The figure presents estimates of y^* under various choices of hedonic controls. For the cross-sectional estimates, we estimate $\log P_{it}^T - \log P_{it}^\infty = \log\left(1 - e^{-y^*T}\right)$ by NLLS, where P_{it}^∞ is the price of a freehold transacted in the same quarter and local authority as a T duration leasehold. For the quasi-experimental estimates, we follow the methodology described in Section III. For each methodology, we perform over 100 estimations, controlling for different combinations of hedonic characteristics. We indicate in various shades of blue four important sets of controls: no controls, linear controls, quadratic controls, and the full set of hedonic fixed effects. The gold cross presents the y^* estimate from Giglio, Maggiori, and Stroebel (2015), and the red cross presents our replication of their estimate, using the full data from 2000–2023.

Since the quasi-experimental methodology utilizes within-property price change, the hedonics controls from Figure 10 affect our results only to the degree that either (i) they are not time-invariant, for instance if the property has renovated, or (ii) there is a change in the sample of properties for which there is data. A stronger test is to control for hedonic characteristics interacted with time. We present these results in Supplemental Appendix Figure A.22, which again indicate that the effect of hedonic controls on the quasi-experimental design is minimal. Nevertheless, there are limits to this exercise for testing the quasi-experimental method. For instance, the test may not detect certain failures of the parallel trend assumption, due to changing characteristics within the same property that we do not observe in our dataset.

V. Macroeconomic Implications of y^*

The primary object of our paper is to introduce the new real-time public dataset on leasehold extensions, and develop a methodology that uses this data to estimate the expected long-run housing yield, y^*. In this section we illustrate how y^* can be useful in informing us about movement in long-term yields in broader class of assets, as well as the natural rate of interest on safe assets. We also illustrate how our data can be used to estimate y^* at a granular geographic level, and show how dynamics of y^* depend on housing risk premia and supply-side elasticity.

A. *The Secular Decline in Expected Long-Term Yields*

Our estimate of expected long-term yield comes from the UK housing market. Are these estimates useful for other asset classes? Obviously one cannot estimate expected long-term forward yields on other assets since other assets do not have the unique natural experiment that we use. However, we can approximate the dynamics of expected long-term yields for other assets at low frequency, using the long run trend of yields. The long-run trend "averages out" short-term shocks to yields, and therefore tracks low frequency movements in the long-term expected yield.[43] If long-run trends for other forms of capital match y^*, then the dynamics of long-run expected yields are similar for UK housing and other assets, at low frequency.

Figure 11 shows that the long-run trend of various other yields has been similar to y^*, particularly prior to 2020. In panel A we plot long-run safe asset yields for the United Kingdom and worldwide. In panel B we plot rent-price ratios for the United Kingdom and a subset of OECD countries. The series is produced from rent-to-price index ratios published by the OECD, and normalized in 2023 to have the same level as y^*. The long-run trend decline for these other asset yields is quantitatively similar to y^*, particularly before 2020. For instance, in panel A, UK and global real forward rates experience a decline of roughly 4 percentage points between 2000 and 2022, compared to a roughly 3 percentage points fall in y^* for UK housing over the same period. In panel B, peak-to-trough fall in rent to price ratios, either in the United Kingdom or globally, has been virtually identical to y^*. We tentatively conclude that the dynamics of long-run expected yields are similar for housing and other assets. Many previous papers study long-run trends in asset yields, and find declines before 2020 (e.g., Farhi and Gourio 2018; Reis 2022). However our measure of expected long-term yields is also available at high frequency and in real time.

Therefore the decline in $y^* \equiv r^* + \zeta^* - g^*$ seems to be common across various assets, and not specific to housing. This finding suggests in turn that the fall in y^* was caused by a fall in r^*, the expected long-term safe asset yield. The reason is that r^* is common to all assets, whereas movements in risk premia ζ^* and dividend growth g^* are specific to housing. Consistent with this logic, in Supplemental Appendix A.13, we present suggestive evidence that the decline in y^* is due to a decline in r^*. In particular, we estimate long-run housing risk premia and expected dividend growth using a standard vector autoregression (VAR) approach, and show that neither long-run UK housing risk premia, nor dividend growth for housing, can account for the trend decline of y^* for housing. Instead, the cause seems to be a fall in long-run expected safe asset yields, r^*. Our finding is consistent with the large literature that has identified declines in r^* through different methods (Holston, Laubach, and Williams 2017). We stress that this exercise is tentative due to the uncertainty of the VAR based procedure.

[43] Formally, consider the yield of an asset z_t, which follows a driftless autoregressive integrated moving average process (ARIMA). We are interested in the expected long-term yield $z_t^* = \lim_{j\to\infty} E_t z_{t+j}$, which is also the Beveridge and Nelson (1981) trend of z_t. Beveridge and Nelson show that $z_t = z_t^* + \tau_t$, where τ_t is a stationary and mean zero "transitory shock." Therefore, time series averages of actual yields, $\sum_{j=0}^{J} z_t/J$, and expected long-run yields, $\sum_{j=0}^{J} z_t^*/J$, are similar over long horizons J, because the long-horizon average of τ_t converges to zero.

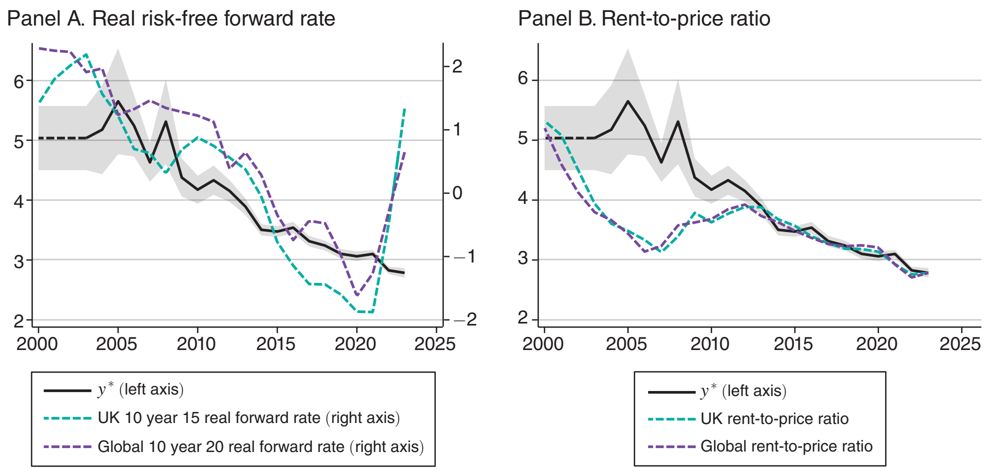

Figure 11. Comparing Other Asset Yields to y^*

Notes: The figures present yields for other assets over the 2003–2023 period. Panel A plots the 10-year-15 real forward rate for the United Kingdom, using data on real forward curves from the Bank of England (Bank of England 1970). It also plots the weighted average 10-year-20 forward rate for nine OECD countries for which data is available from Global Financial Data (2024). The average is weighted by each country's GDP. Since data on inflation-linked bonds is not available for many countries, we define the real forward rate as the nominal forward rate minus 3 percent expected inflation. Panel B plots the UK rent-to-price ratio and the weighted average rent-to-price ratio for 22 OECD countries for which it is available (OECD 2023). The rent-to-price ratio is calculated from a rent-to-price index published by OECD, and normalized to the 2023 y^* level.

In Supplemental Appendix Section A.15, we review available evidence on the behavior of very long-term safe interest rates, with durations of 100 years or more. The data come from the single sovereign debt issuer with 100-year bonds (Austria), as well as very long-term corporate bonds issued by various high credit quality corporates (Trace 2022; Datastream 2024). The evidence is somewhat uncertain, due to liquidity issues with the bonds and a short time series. However on the whole the evidence suggests low and stable r^*, consistent with our findings.

B. *Comparing Current Yields to* y^*

One can usefully compare expected long-term yields to current asset yields. This exercise reveals whether shorter or longer-term factors have been the determinant of current asset price fluctuations, as we illustrate in two applications.

Dynamics of r^* *after 2020.*—One important question is whether r^*, the expected long-term yield of safe assets, has risen since 2020 or stayed at its prior low level. Meanwhile, as Figure 11, panel A shows, medium-term asset yields have risen sharply after 2020. Our estimates help to answer whether these medium-term yields have risen due to longer-term factors, indicating rising r^*; or, due to shorter-term factors. The answer has several implications. For instance, the behavior of r^* affects the conduct of monetary policy. Moreover the behavior of r^* is a key indicator of whether the economy will return to the low interest rates of the "secular stagnation" era (Blanchard 2023).

The stability of y^* after 2020 suggests that r^* has also been stable, despite the rise in medium-term yields. We stress that $y^* \equiv r^* + \zeta^* - g^*$ and r^* are different objects. However, y^* contains a great degree of information about r^*, since a large rise in r^* would likely lead to a large rise in y^*.[44]

The different behavior of medium- and long--term yields suggests that, since 2020, the yield curve has flattened, with short- and medium-run forward yields rising relative to long-run yields. Our estimate of y^* suggests that medium-run interest rates will fall because the current rise in the medium-term forward rate is a temporary deviation from r^*, which has remained low. Alternatively, y^* must ultimately rise to meet the short end of the yield curve. Our real-time estimates of y^* will allow us to track this convergence process on a monthly basis.

However the convergence between medium-term yields in Figure 11 and y^* may be slow, for at least two reasons. First, the duration of the medium- and long-term yields is different. The medium-term yields in the figure are 10 year interest rates starting in 15 years' time. y^* is pinned down by 90 year yields start in 70 years' time. As such, the gap between the two series could last for quite some time. Second, y^* as estimated in the paper could be a slower moving variable than the correct measure of r^*. In housing markets, trading is relatively infrequent, and factors such as norms and advice by real estate agents could be important. If so, then y^* may evolve slowly. On the other hand, measures of r^* inferred from financial market prices might evolve more quickly.

Dynamics of House Prices during the Early 2000s.—The gap between current rent-price ratios and y^*, the expected long-term rent-price ratio, can be a useful diagnostic for understanding whether current housing valuations are likely to persist, or may be more transitory. As an example, consider Figure 11 panel B. Current rent-price ratio tracks y^* roughly over the long-run. However, there can be sustained periods where the two diverge, as in the period from 2000 to 2006.

The divergence during 2000–2006 suggests that the housing boom during this period was driven by a decline in the short end of the housing forward curve, while the long-end yield remained high. Since risk-free yields remained relatively largely stable during this period as well, once could conclude that the divergence between y^* and R/P is driven by a medium-term shift in expectations about the housing risk premium or housing dividend growth, which might prove to be transient. In contrast, the decline in rent-to-price ratios after 2012 tracks the decline in y^* very closely, suggesting that the recent increase in house prices has been driven by a shift in the long end, and hence might be more persistent.

Through this perspective, the high housing valuations of 2000–2006 and 2015 onward are quite different. In 2000–2006, valuations were temporarily higher than their long-run values, and likely to fall, whereas current housing valuations seem to be more persistent and less likely to revert. Put differently, the rent-to-price ratio were at the same level in 2006 and at the end of the 2010s. However long-term yields were much lower in the later part of the sample—suggesting that rent-to-price ratios at the end of the 2010s were more likely to last. Going forward, our estimates

[44] An alternative, and in our view less likely possibility, is that r^* has risen but a second shock has caused ζ^* or g^* to change in an offsetting direction at exactly the same time.

Table 5—Determinants of Long-Run Risk

	y^*	
	(1)	(2)
Flood risk	0.19	
	(0.07)	
Subsidence risk		0.33
		(0.11)
Observations	224	224
R^2	0.05	0.10

Note: Standard errors in parentheses.

can be used to understand whether movements in house prices are driven by transitory or more persistent factors.

C. *Cross-Sectional Variation in y^**

The long-run housing yield $y^* \equiv r^* + \zeta^* - g^*$ contains housing-specific components: The risk-premium ζ^* and the rate of rent growth, g^*. We now investigate which factors affect ζ^* and g^* by exploiting cross-sectional variation in y^*.

y^* and Climate Risk.—One source of variation in regional long-run housing risk may be exposure to climate change related hazards. In the United Kingdom, two major climate risks that will worsen as global temperatures rise are floods and subsidence. There is substantial variation in expected exposure to these risks depending on proximity to bodies of water, temperature, and soil type. We obtain data on flood risk from the UK Environment Agency and data on predicted long-run damage from subsidence from the British Geological Survey GeoClimate project (UK Environment Agency 2019; British Geological Survey 2020). We aggregate these measures at the local authority (LA) level, creating measures of the share of each LA that is exposed to each risk. We then correlate these shares with our local estimates of y^* in the two columns of Table 5. Areas with larger climate risk exposure have higher levels of y^*, consistent with them having higher long-run risk premia.

Δy^* and Supply Elasticity.—The expected long-term housing yield also encodes whether the supply of housing is elastic. Suppose that demand for housing rises, perhaps because of a fall in long-term interest rates r^*. If supply is inelastic, then valuations of housing will rise, increasing the price-rent ratio and lowering y^*. However, elastic land supply will accommodate the growth in demand, and mitigate rising valuations and falling y^*.

Formally, recall the definition $y^* \equiv r^* + \zeta^* - g^*$, and suppose that housing demand rises due to falling r^*. If housing supply is elastic, then new construction leads to slower rent growth, meaning g^* falls and y^* does not change. If housing supply is inelastic, then g^* does not change and falling r^* passes through to falling y^*.

Consistent with this logic, we find that y^* falls by more in areas with more inelastic housing supply. We define areas as local authorities (LAs) and use two measures of

Table 6—Cross-Sectional Heterogeneity in Δy^*

	Δy^*				Refusal rate	Δy^*	
	OLS	OLS	OLS	OLS	1st stage	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Refusal rate	−4.29	−6.35	−6.87			−17.94	−13.74
	(1.25)	(1.27)	(1.36)			(11.19)	(5.44)
Share developed		−1.63	−1.82				−2.38
		(0.32)	(0.54)				(0.72)
Change in delay rate				2.52	−0.17		
				(0.64)	(0.04)		
Region fixed effects			✓				
Observations	224	224	224	224	224	224	224
R^2	0.11	0.30	0.34	0.11	0.08	−1.05	0.00

Notes: The table presents cross-sectional regressions of the long-run change in y^* against local authority level measures of supply-side constraints in the United Kingdom. The long-run change in y^* is defined as the difference between post-2022 y^* and pre-2009 y^* (inclusive). Columns 1 through 4 present OLS regressions of Δy^* against three measures of supply elasticity: the average refusal rate of major construction applications, the share of developable land developed, and the change in delay rate following the 2002 policy reform, all of which are discussed in the main text. In column 3 we also include region fixed effects. Column 6 and 7 present IV regressions, where the refusal rate is instrumented using the change in delay rate after 2002. Standard errors in parentheses.

supply constraints from Hilber and Vermeulen (2016). The first is the share of major construction applications refused by that local authority, averaged over a nearly 30-year period. The second is the share of developable land developed in 1990.

One problem is that the refusal rate of construction applications is endogenous; developers are less likely to submit construction applications in more restrictive LAs that are unlikely to accept them, which in turn reduces the refusal rate of more restrictive LAs. For example, as we discuss in Supplemental Appendix A.14, the application refusal rate in places like Central London is surprisingly low, most likely because construction firms know that applications will be refused and do not submit them. This implies that the measured application refusal rate is an underestimate of what the refusal rate would be if the application submission rate was randomly assigned.

To address this endogeneity concern, Hilber and Vermeulen (2016) develop an instrument based on a 2002 policy reform, which incentivized LAs to make decisions on major construction project applications in a timely fashion. This policy should have especially lowered the application delay rate for inelastic LAs, which were more likely to delay construction applications before the reform. Column 5 of Table 6 presents the first stage for this instrument, indicating that LAs that lowered their application delay rate after 2002, on average had higher refusal rates.[45] In Supplemental Appendix A.14 we elaborate further on the validity of this instrument.

Table 6 shows that local authorities with more inelastic housing supply have greater declines in y^*. We measure the decline, Δy^*, as the long difference between the pre-2009 period to the post-2022 period, although our results are robust to

[45] Supplemental Appendix Figure A.27 presents a binscatter of the first stage.

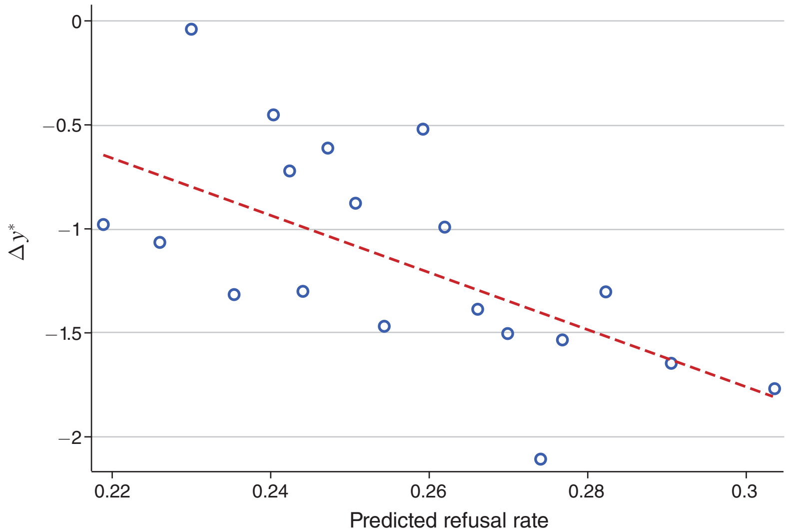

FIGURE 12. THE DECLINE IN y^* DEPENDS ON HOUSING SUPPLY ELASTICITY

Note: The figure presents a binscatter with 20 bins of the long-run change in y^*, Δy^*, against the predicted values from the first stage of the IV in column 6 of Table 6.

choosing other cutoff years. Regressions are weighted by the inverse of the sum of the variances of the pre and post y^* estimates.[46] Column 1 of Table 6 presents an OLS regression of Δy^* on the local authority major construction application refusal rate. In column 2, we also include the share of developable land developed. In column 3, we control for fixed effects for the nine regions of England. In column 4, we present a reduced-form regression of Δy^* against the delay rate instrument from Hilber and Vermeulen (2016). In column 5, we instrument for the refusal rate using the delay rate instrument, and in column 6 we repeat this regression, but also include the share developed as a covariate. After instrumenting for the refusal rate, the share developed control has a smaller effect on the refusal rate coefficient, consistent with the refusal rate measure being endogenous. In all cases, the results indicate that more regulated and land-constrained local authorities experience greater declines in y^*. Based on column 2, and holding the share of land developed constant, the decline in y^* is more than 1 percentage point greater in magnitude for local authorities at the ninetieth percentile of refusal rate, relative to those at the tenth percentile. Figure 12 presents a binscatter of the predicted values from the first stage of column 6 against Δy^*, controlling for the share of land developed, again indicating that more restrictive LAs experienced greater declines in y^*.

Our finding is not surprising: As demand for housing rises, valuations increase more in areas with inelastic supply. However our estimates have implications for the

[46] In all the regressions, we utilize the variation of $\underline{\Delta}_{it}$ built using an annual repeat sales index (RSI), as opposed to a quarterly repeat sales index. This is because we can match the annual RSI to more experiments, which therefore yields a larger sample size. Supplemental Appendix Figure A.23 verifies that estimates using the annual RSI are very similar to the quarterly RSI.

overall economy. The large decline in y^* indicates that housing supply is inelastic in the United Kingdom in aggregate. Previous work has established similar results at higher frequency (e.g., Miles and Monro 2019). We show that inelastic housing supply matters even at long-run frequencies for the UK economy.

VI. Concluding Remarks

This paper estimates the expected long-term yield of housing and its dynamics from 2000 to present for the UK property market. We exploit a natural experiment: extensions of long-duration property leases in the British property market. Our findings show that y^* fell from 5.2 percent before the Great Recession to 2.7 percent in 2024. An important goal of this paper was to assemble an administrative dataset on lease extensions and make it publicly available for free. Our data will be updated in real time going forward. We hope that both policy makers and academics will find the data and empirical methodology useful for estimating the market's expectation of long-term yield in real time.

Long-run yields are valuable because they "look through" the short-term factors affecting asset prices in real time. Our results show that there has been a growing gap between long-run yields and real forward yields from 2022–2024. This gap will narrow either by a fall in spot real yields, as the "secular stagnation view" suggests, or a rise in long-run yields if the post-COVID era reflects a structural regime shift. The consequences of these two scenarios for asset prices and the real economy could of course be very different. In recent months, y^* has remained stable. The real-time estimates of y^* should be helpful to determine the trajectory of long-run yields going forward.

The focus of this paper has been the measurement of the expected long-run housing yield, which, as we have discussed, is difficult to do with precision and minimal assumptions. However, our paper introduces several questions for future research: Why has y^* fallen? And what does this imply about the state of the economy? Although these questions are out of the scope of this paper, we believe that there is much that can be learned about the economy from studying y^* in the cross section. We have argued that the dynamics of y^* encode useful information about the elasticity of housing—future work may be able to generalize these results to housing in other countries, and to other forms of capital.

REFERENCES

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge. 2023. "When Should You Adjust Standard Errors for Clustering?" *Quarterly Journal of Economics* 138 (1): 1–35.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1): 151–84.

Bäcker-Peral, Verónica, Jonathon Hazell, and Atif Mian. 2026. *Data and Code for: "Dynamics of the Long-Term Housing Yield: Evidence from Natural Experiments.*"American Economic Association; distributed by Inter-university Consortium for Political and Social Research. https://doi.org/10.3886/E235801V.

Badarinza, Cristian, and Tarun Ramadorai. 2015. "Long-Run Discounting: Evidence from the UK Leasehold Valuation Tribunal." Unpublished.

Bailey, Martin J., Richard F. Muth, and Hugh O. Nourse. 1963. "A Regression Method for Real Estate Price Index Construction." *Journal of the American Statistical Association* 58 (304): 933–42.

Bank of England. 1970-2022. *Daily Government Liability Curve*. https://www.bankofengland.co.uk/statistics/yield-curves (accessed December 2, 2024).
Beveridge, Stephen, and Charles R. Nelson. 1981. "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'." *Journal of Monetary Economics* 7 (2): 151–74.
Blanchard, Olivier. 2023. "Secular Stagnation Is Not Over." Unpublished.
Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2021. "Revisiting Event Study Designs: Robust and Efficient Estimation." Unpublished.
Bracke, Philippe, Edward W. Pinchbeck, and James Wyatt. 2018. "The Time Value of Housing: Historical Evidence on Discount Rates." *Economic Journal* 128 (613): 1820–43.
British Geological Survey. 2020. *GeoClimate UKCP18 Open.* https://www.bgs.ac.uk/datasets/geoclimateukcp18-open/#:~:text=GeoClimate%20UKCP18%20Open%20is%20provided,case)%20within%20the%20grid%20cell (accessed September 30, 2024).
Case, Karl E., and Robert J. Shiller. 1989. "The Efficiency of the Market for Single-Family Homes." *American Economic Review* 79 (1): 125–37.
Christensen, Jens H. E., and Glenn D. Rudebusch. 2017. "New Evidence for a Lower New Normal in Interest Rates." FRBSF Economic Letter 2017-17.
Crump, Richard K., Stefano Eusepi, and Emanuel Moench. 2016. "The Term Structure of Expectations and Bond Yields." Federal Reserve Bank of New York Staff Report 775.
Datastream. 2024. *Austrian Yield Curves.* URL (accessed [data]).
Farhi, Emmanuel, and François Gourio. 2018. "Accounting for Macro-Finance Trends: Market Power, Intangibles, and Risk Premia." Unpublished.
Giglio, Stefano, Matteo Maggiori, and Johannes Stroebel. 2015. "Very Long-Run Discount Rates." *Quarterly Journal of Economics* 130 (1): 1–53.
Giglio, Stefano, Matteo Maggiori, and Johannes Stroebel. 2016. "No-Bubble Condition: Model-Free Tests in Housing Markets." *Econometrica* 84 (3): 1047–91.
Global Financial Data. 2024. *Global Yield Curves.* URL (accessed [data]).
Gomme, Paul, B. Ravikumar, and Peter Rupert. 2015. "Secular Stagnation and Returns on Capital." *Economic Synopses* 19: 1–3.
Hamilton, James D., Ethan S. Harris, Jan Hatzius, and Kenneth D. West. 2016. "The Equilibrium Real Funds Rate: Past, Present, and Future." *IMF Economic Review* 64 (4): 660–707.
Harrison, J. Michael and David M. Kreps. 1978. "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations." *Quarterly Journal of Economics* 92 (2): 323–36.
Hilber, Christian A. L., and Wouter Vermeulen. 2016. "The Impact of Supply Constraints on House Prices in England." *Economic Journal* 126 (591): 358–405.
HM Land Registry. 2024. *Price Paid Data*. https://www.gov.uk/government/collections/price-paid-data (accessed February 28, 2024).
HM Land Registry. 2024. *Registered Leases*. https://use-land-property-data.service.gov.uk/datasets/leases (accessed February 1, 2024).
Holston, Kathryn, Thomas Laubach, and John C. Williams. 2017. "Measuring the Natural Rate of Interest: International Trends and Determinants." *Journal of International Economics* 108: S59–S75.
Johannsen, Benjamin K., and Elmar Mertens. 2016. "The Expected Real Interest Rate in the Long Run: Time Series Evidence with the Effective Lower Bound." Board of Governors of the Federal Reserve System Notes 1703.
Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M. Taylor. 2019. "The Rate of Return on Everything, 1870–2015." *Quarterly Journal of Economics* 134 (3): 1225–98.
Kiley, Michael T. 2015. "What Can the Data Tell Us About the Equilibrium Real Interest Rate?" Unpublished.
Laubach, Thomas, and John C. Williams. 2003. "Measuring the Natural Rate of Interest." *Review of Economics and Statistics* 85 (4): 1063–70.
Lubik, Thomas A., and Christian Matthes. 2015. "Calculating the Natural Rate of Interest: A Comparison of Two Alternative Approaches." Federal Reserve Bank of Richmond Economic Brief 15-10.
Miles, David, and Victoria Monro. 2019. "UK House Prices and Three Decades of Decline in the Risk-Free Real Interest Rate." Unpublished.
Negro, Marco Del, Domenico Giannone, Marc P. Giannoni, and Andrea Tambalotti. 2017. "Safety, Liquidity, and the Natural Rate of Interest." Unpublished.
Negro, Marco Del, Domenico Giannone, Marc P. Giannoni, and Andrea Tambalotti. 2019. "Global Trends in Interest Rates." Journal of International Economics 118: 248–62.
OECD. 2023. *Housing Prices*. https://www.oecd.org/en/data/indicators/housing-prices.html (accessed February 6, 2024.).

Office of National Statistics (ONS). 2021a. *Number of Vacant and Second Homes, England and Wales: Census 2021*. https://www.ons.gov.uk/peoplepopulationandcommunity/housing/bulletins/numberofvacantandsecondhomesenglandandwales/census2021 (accessed September 17, 2024).
Office of National Statistics (ONS). 2021b. *Population Density*. https://www.ons.gov.uk/datasets/TS006/editions/2021/versions/4 (accessed September 9, 2024).
Office of National Statistics (ONS). 2024. *Annual Survey of Hours and Earnings.* https://www.ons.gov.uk/releases/employeeearningsintheuk2024 (accessed February 20, 2024)
Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business and Economic Statistics* 37 (2): 187–204.
Rachel, Lukasz, and Lawrence H. Summers. 2019. "On Secular Stagnation in the Industrialized World." Unpublished.
Rachel, Lukasz, and Thomas D. Smith. 2017. "Are Low Real Interest Rates Here to Stay?" *International Journal of Central Banking* 13 (3): 1–42.
Reis, Ricardo. 2022. "Which R-Star, Public Bonds or Private Investment? Measurement and Policy Implications." Unpublished.
Rightmove Inc. 2006/2022. *Rightmove Listings*.
Trace. 2022. *Corporate Bond Yields.*
UK Data Service. 2022. *English Housing Survey*. https://datacatalogue.ukdataservice.ac.uk/series/series/200010#abstract (accessed December 25, 2024).
UK Environment Agency. 2019. *Flood Risk Areas*. https://www.data.gov.uk/dataset/42c31542-228d-439b-8dbe-e72135dae71c/flood-risk-areas (accessed September 30, 2024).
Urban Big Data Centre. 2010/2021. *Zoopla Listings.* https://www.ubdc.ac.uk/data-services (accessed August 1, 2023).
Vissing-Jorgensen, Annette. 2022. " 'Has Monetary Policy Cared Too Much About a Poor Measure of r* ?' Discussion." Unpublished.
Williams, John C. 2023. "Measuring the Natural Rate of Interest: Past, Present, and Future." Unpublished.