NBER WORKING PAPER SERIES

DYNAMICS OF THE LONG TERM HOUSING YIELD:
EVIDENCE FROM NATURAL EXPERIMENTS

Verónica Bäcker-Peral
Jonathon Hazell
Atif R. Mian

Working Paper 31760
http://www.nber.org/papers/w31760

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2023, Revised February 2024

We thank the Julis Rabinowitz 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 Ruxandra Iosif for research assistance. This paper was previously titled "Measuring the Natural Rate Using Natural Experiments". A website accompanying this paper displays real-time updates to our estimates, data, and replication instructions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Dynamics of the Long Term Housing Yield: Evidence from Natural Experiments
Verónica Bäcker-Peral, Jonathon Hazell, and Atif R. Mian
NBER Working Paper No. 31760
October 2023, Revised February 2024
JEL No. E4,E40,E5,G5

ABSTRACT

Every month, a fraction of UK property leases are extended for another 90 years or more. We use new data on thousands of these natural experiments from 2003 onwards to estimate y*, the expected long term housing yield, or rent-price ratio. Starting from a level of 5.2%, y* starts to fall during the Great Recession, reaching a low of 2.8% in 2023. Real-time data shows that y* has not risen since 2021. Our approach is precise, avoids misspecification concerns, allows real-time estimates with public data, and provides information about changes in r*, the expected long term yield of safe assets.

Verónica Bäcker-Peral
The Julis-Rabinowitz Center
for Public Policy and Finance
216-1 Julis Romo Rabinowitz Building
Princeton University
Princeton, NJ 08544
United States
veronicabp@princeton.edu

Jonathon Hazell
Department of Economics
London School of Economics
32 Lincoln's Inn Fields
United Kingdom
jadhazell@gmail.com

Atif R. Mian
Princeton University
Bendheim Center For Finance
26 Prospect Avenue
Princeton, NJ 08540
and NBER
atif@princeton.edu

A website accompanying this paper which displays real-time updates to our estimates, in addition to data and replication instructions. is available at
https://sites.google.com/view/naturalrate/home

1 Introduction

This paper provides a new public data set and methodology to measure the expected long term yield of housing, both historically and in real time. The expected long term yield, or rent-price ratio, is important because it contains information about the expected long term equilibrium of the economy, after transitory shocks have passed. For instance, yields on many assets were low and falling during the early decades of the 21st century, but rose after the Pandemic Recession. At the time of writing, a hotly debated question has been whether yields will remain high or revert to former levels (e.g. Blanchard, 2023). The answer sheds light on whether the structure of the economy has changed since 2020. As such, central bankers and financial market participants pay a great deal of attention to expected long term yields.

Formally we measure y^*, the yield that at a point in time, markets expect will prevail in the long term. This yield satisfies $y^* \equiv r^* + \zeta^* - g^*$, where r^* is the expected return on safe assets in the long term, ζ^* is the long term risk premium, and g^* is expected dividend growth in the long term. As such, y^* contains information about some key macroeconomic aggregates. Additionally, since it is the expected rent-price ratio, y^* clears the market for housing in the long term.

Expected long term yields are difficult to measure. Conceptually, the challenge is to "look through" transitory factors such as business cycles and adjustment costs, which affect asset prices only in the short and medium term. There are two approaches to overcoming this challenge with time series data. One approach studies long run trends in the time series of various asset prices across decades (e.g. Farhi and Gourio, 2018; Jordà, Knoll, Kuvshinov, Schularick and Taylor, 2019; Reis, 2022). This approach is well suited to tracking gradual evolutions 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. A second approach specifies a structural macroeconomic model in order to filter out short term shocks and measure expected long term yields in real time. For instance, Laubach and Williams (2003) and Holston, Laubach and Williams (2017) use this approach to measure the expected long term yield on safe assets, r^*.[1] This work is justly celebrated for tackling a difficult and important problem. However, estimates tend to be sensitive to the specification of the structural model. For instance, the Holston et al. (2017) estimates were discontinued during the pandemic due to misspecification (Williams, 2023).

This paper takes a different approach to measuring the dynamics of expected long term

[1] Related papers include Kiley (2015), Lubik and Matthes (2015), Johannsen and Mertens (2016), Crump et al. (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).

yields, by using natural experiments and micro data. In doing so, we are able to calculate expected long term yields in real time, without making hard-to-verify assumptions about the short term behavior of the economy.

We estimate y^* for UK housing from 2003 to present. In the UK, most apartments and 22% of all dwellings, are sold as "leaseholds"—long duration leases starting at ninety years or more, issued by the ultimate owner of the property, or "freeholder". The leaseholder can buy or sell the lease, giving each lease a series of market prices. Moreover leaseholders have the right to extend their lease conditional on paying freeholders the value of the lease extension. Lease extensions typically happen when the current lease has somewhere between 60 to 90 years left. The typical extension is for an additional 90 years, but a large share of leases extend by 700 years or more, letting us estimate y^* at various parts of the long term yield curve.

This paper assembles a new data set on leasehold extensions and transactions from 2003 to the present. We estimate the increase in the market value of a leasehold due to duration extension, by comparing its price before and after lease extension with an otherwise similar control group of leaseholds that do not get extended at the same time. We embed this difference-in-differences estimate of the market value of lease duration extension into a simple discounted cash-flow pricing equation. The equation allows us to estimate *the market's expected* y^* via non-linear least squares using over 120,000 lease extension experiments from 2003 onward. Our measure of y^* is automatically in real terms (i.e. unaffected by inflation expectations).

An important empirical advantage of our natural experiment approach to estimating y^* is that extensions generate variation in lease duration for the *same property*. Therefore our approach can "difference out" shorter term factors without making strong assumptions that are vulnerable to misspecification. 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 measures only the long term value, and short to medium term shocks to rates of return, such as monetary tightening, do not affect our estimate of y^*. Similarly, our estimate is unaffected by short term shocks to the service flow of housing or other hard-to-measure characteristics of any given property. For instance, a shock to demand for a particular segment of London property, which raises its service flow, does not affect our estimate.

The key assumption for our natural experiment to identify y^* is "parallel trends": the service flow of housing must grow similarly for extended properties and their control group. We support this identification assumption in four ways. First, there are no pre-trends, meaning prices of extending properties evolve similarly to the control before extension. Second, the treatment and control group are balanced on a rich set of hedonic characteristics. Third, market rents and hedonic characteristics evolve similarly for extenders and the control group. Fourth, the estimator is not sensitive to controlling for observed heterogeneity, suggesting bias from unobserved characteristics is small (Altonji, Elder and Taber, 2005; Oster, 2019).

Figure 1: Time Series of the Expected Long Term Housing Yield

The figure shows estimates of y^* for each year of data and for the full sample of lease extensions. Monthly estimates for January 2023 through October 2023 are also reported. The shaded area shows 95% confidence intervals for the estimates. Standard errors are heteroskedasticity robust.

Figure 1 plots the average y^* for each year from 2003 to 2023, and the shaded area represents the 95% confidence interval. There are two key findings. First, though y^* stays around 5.2% between 2003 and 2006, there is a trend fall at the onset of the Great Recession, culminating in a low of 2.8% in 2023. 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 afterwards.

Our measure of y^* is precise and available in real time, making it potentially useful for policymakers. The lease extension data base put together in this paper is based on publicly available data that is updated every month, meaning our methodology can be easily replicated and extended in real time. Our estimator exploits cross sectional variation at a point in time, and there are over a thousand leasehold extensions every month—allowing precise estimates of y^* even at a monthly frequency. We illustrate the usefulness of y^* in Figure 1, by presenting estimates separately for each month for 2023, with the latest data point in October 2023. Real-time monthly data shows modest movements in y^* in 2023, with tight confidence intervals.[2]

Our estimates of y^* are specific to housing, but they have several macroeconomic implications. First, our estimates tentatively suggest a decline in expected long term yields on capital throughout the economy, beyond the housing sector. We study long run trends in the yields of other assets, such as equities and bonds, and find they display similar declines to y^*. Long run trends "average out" short run shocks, and therefore measure the same long term component that our estimate of y^* captures. We conclude that the decline in expected long term yields is broad-based, not specific to housing. Importantly, long run trends in other asset prices are only available at low frequency—whereas our measure of y^* is available at high frequency, so that policymakers can it in a timely fashion. Given that the decline in $y^* \equiv r^* + \zeta^* - g^*$ for housing is common to other assets, a plausible cause is falling expected long term safe asset yields r^*, as opposed to housing specific movements in dividend growth g^* or risk premia ζ^*. We find suggestive VAR-based evidence that r^* accounts for the fall in y^*, consistent with a large literature finding declines in r^* through other methods (Holston et al., 2017).

A second implication of our estimate concerns whether r^* has risen since 2020. This question has implications for the conduct of monetary policy, and whether the economy will revert to its pre-2020 equilibrium after the shocks of recent years have passed (Blanchard, 2023). Though $y^* \equiv r^* + \zeta^* - g^*$ and r^* are not equal, the absence of a sharp rise in y^* after 2020 suggests that r^* has not risen either. Over the same time, medium term safe asset yields have risen sharply—for instance, real 30 year forward yields for UK government debt rose by roughly 350 basis points during 2020-2023. Our estimates suggest that this increase was due to short or medium term shocks, such as monetary tightening. Relative to prevailing estimates of r^* using structural time series methods, our method provides two advantages. First, our micro-data approach leads to precise estimates of y^*, which weigh against meaningful increases in r^*. Estimates of r^* with time series methods are typically imprecise, and rule out neither unchanged nor large rises in r^*. Second, our natural experiment method does not risk model mis-specification, whereas structural estimates of whether r^* has risen are sensitive to the model (Baker et al., 2023).

A third implication of our estimates comes from cross sectional heterogeneity in the dynamics of y^*, which reveals information about land supply. Empirically, we show that y^* falls by more in areas with inelastic land supply. This finding is unsurprising: areas with constrained supply should experience rising valuations as demand grows, meaning a falling rent-price ratio y^*. Given this finding, the large fall in aggregate y^* indicates that overall,

[2] See our website for data, real-time estimates and replication instructions.

land supply is inelastic in the UK, consistent with a wealth of other evidence (e.g. Miles and Monro, 2019).

Our paper is closely related to the seminal work of Giglio, Maggiori and Stroebel (2015), who were the first to observe that because UK apartments vary in duration, they are particularly well suited to estimating expected long run yields.[3] Giglio et. al. make a cross-sectional comparison of properties with different duration, and control for a rich set of hedonic characteristics, in order to estimate the level of y^*. Our paper studies the dynamics of y^*, and in doing so builds on their work in two ways.[4] First, we use a quasi-experimental design based on lease extensions, which allows us to use variation in leasehold duration and price *within the same property*. We show that within-property variation is critical to accurately estimating y^*, due to unobserved heterogeneity in the service flow of housing. Our method can then estimate the dynamics of y^* reliably in real time, at monthly frequency. Second, the possibility of lease extension creates option value that might affect the price of UK leaseholds. We explicitly take this option value into account, and in Appendix A.13 we develop a new bunching estimator to measure this option value.

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.[5]

Outline. The rest of the paper is structured as follows. Section 2 defines the expected long run yield of housing. Section 3 describes the data. Section 4 presents our empirical methodology. Section 6 discusses the implications of our estimates for the macroeconomy. Section 7 concludes with a discussion of possible future work based on our new methodology.

2 The Expected Long Term Yield of Housing

This paper studies the yield of housing that is expected by markets in the long term, after medium-run forces have subsided. Formally, consider the price of a unit of housing. The

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

[4]Our estimate of the level of y^* is not strictly comparable to Giglio et al. because our sample of lease extensions have different durations from the cross section of properties that they study.

[5]Farhi and Gourio (2018) show that given the long run expected yield of capital, rising monopoly power raises the profit-to-capital ratio.

price at time t, P_t is given by the present value of its dividend, R_t, as

$$P_t = R_t \int_0^{\infty} e^{-\int_0^s (r(u)+\zeta(u)-g(u))du} ds. \tag{1}$$

In this equation, $r(u) + \zeta(u)$ is the discount rate of housing u periods in the future, which discounts the present value of dividends at a rate that sums the safe rate of return and the risk premium, all in real terms, and $g(u)$ is dividend growth.[6]

We define the yield of housing as $y(u) \equiv r(u) + \zeta(u) - g(u)$. We are interested in the yield of housing that is expected in the long term, y^*. We assume that y^* exists, in which case it is given by $\lim_{u\to\infty} y(u) = y^* = r^* + \zeta^* - g^*$. In the long run, the dividend-price ratio converges to y^*, since by Equation (1) we have $R_\infty/P_\infty = \lim_{u\to\infty} y(u) = y^*$.

The expected long term yield of housing matters for at least two reasons. First, y^* contains information about long term safe asset yields r^*, long term risk premia for housing ζ^*, and long term dividend growth for housing g^*. For instance, a large fall in y^* implies that at least one out of risk premia ζ^*, capital gains g^*, or long term safe asset yields r^*, has fallen. Moreover r^*, ζ^*, and g^* all contain important information about the equilibrium structure of the economy.

Second, the expected long-term yield of housing is an important variable in its own right, beyond information about its various components. In particular, since y^* is the long-term dividend-price ratio, it is the key variable that clears the market for housing in the long run. Since housing is an important asset, y^* is relevant for the long run supply and demand for assets in the economy at large.

Crucially, y^* is independent of short run, transient shocks to rates of return or dividends. As such, the key challenge to estimating expected long term yields is to filter out these shorter run shocks. Current asset prices—even for very long duration assets—depend not only on y^* but also on shorter run rates of return, which include transitory factors such as monetary policy, credit booms, short run bubbles and adjustment costs. The reason is that, as Equation (1) shows, the price of a long duration asset depends on the integral of short and longer term discount rates. Therefore the current dividend-price ratio, R_t/P_t, is affected by these transitory shocks and may greatly differ from the expected long-term ratio of dividends to prices—which is also the expected long term yield. Moreover, the dividend of capital R_t is often difficult to observe. For instance, the service flow of owner-occupied housing cannot be observed from market transactions, and must be imputed. Therefore short run shocks to the dividend of capital—for instance, a temporary increase in demand for a certain segment of London housing—are an additional source of variation that confounds estimates of y^*.

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

3 Data & Lease Extension Details

This section introduces the setting—lease extensions for long duration leasehold properties in the United Kingdom—and explains the data sets we will use for our analysis.

3.1 Data

We use five data sets for the analysis of this paper. Most of the analysis uses data that are publicly available almost in real time. This is an important feature of our method, since it allows for real-time updates and replication. A website accompanying this paper contains data, replication instructions, and real time analysis. We now describe our five data sets.

(1) **Land Registry Transaction Data**: We obtain publicly available data on all property transactions registered in England and Wales between 1995 and October 2023 from the Land Registry. The data set includes the exact date, price and address for each transaction. Properties are also subdivided 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. As we will discuss in detail below, leasehold flat durations are periodically extended, after negotiation with the freeholder. 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.[7] 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 20th 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. To this day, leaseholds are very common in England and Wales, comprising approximately 5% of houses and 97% of apartments. 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.

(2) **Land Registry Lease Data**: Data on the length of lease terms, which vary significantly across leaseholds, are provided in a separate Land Registry data set, also publicly available. The Land Registry lease term data set includes the property address associated to the lease, the term length and origination date of the lease, as well as the date in which the lease was registered with the Land Registry. A majority of leases were originated and

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

registered after the 1950s. Starting in late 2003, lease registration became mandatory, so we capture all registered leases after this period. We will start our analysis in 2003, after mandatory registration.

Each entry in the lease term data set provides the length of the lease term, its registration date, and the start date of the lease. The Land Registry does not provide match keys to merge the transaction and lease term data sets, so we conduct a fuzzy merge based on provided addresses, as we detail in Appendix A.4.[8]

The Land Registry classifies properties into five types: flats, detached houses, semi-detached houses, terraced houses, and others. The three categories of houses are for the most part freehold properties. Even when they are leaseholds, houses tend to have very long lease terms relative to flats, with median remaining lease term over 800. Since leaseholds are mostly flats, we will drop non-flats for the bulk of our analysis.

(3) **Land Registry Extensions Data**: We obtain data on lease extensions from two sources, neither of which have been used in prior academic research to our knowledge. First, beginning in September 2021, the Land Registry began publicly releasing information on lease extensions. Second, we have obtained from the Land Registry all lease extensions before September 2021, which we have made available on the website accompanying this paper.

(4) **Rightmove Hedonics Data**: We obtain data on housing characteristics and rental prices from Rightmove, Inc. spanning 2006 to the present, which include 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.

(5) **Zoopla Hedonics Data**: We supplement this with data from Zoopla, Inc. which is provided for free to researchers by the Urban Big Data Centre. This data set also provides number of bedrooms and bathrooms and rents. Additionally, it includes the number of floors and receptions of the property. We are able to match approximately 80% of transactions to the Rightmove and Zoopla listing data. Rental data is available for about 40% of properties.

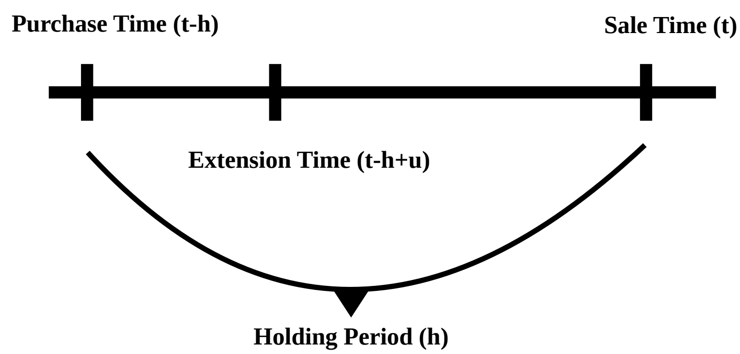
3.2 Lease Extension Details

This section provides additional details about leaseholds in England and Wales and the circumstances under which the duration of leases can be extended. As we described in the previous section, leaseholds are properties with very long but finite leases, moreover these

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

leases can be bought and sold. The lease length at the time it was issued is denoted its "initial lease term" and the lease length it has at any future point in time is denoted its "remaining lease term." The distribution of leases can be divided into two groups; about 70% of leasehold flats in our sample are *short leaseholds* with remaining terms of 250 years or less and the other 30% are *long leaseholds* with remaining terms of 700 years or more. There are practically no properties with remaining terms between 250 and 700.[9]. The most common initial terms for short leaseholds are for 99 and 125 years, which account for 77% of short leaseholds. The most common initial term for long leaseholds is 999 years, which account for 96% of all long leaseholds.

Figure 2: Diagram of Extension Time



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 $t-h+u$, where $0 < u < h$.

Beginning in 1993, the 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.[10] This negotiation is potentially costly, since both the freeholder and the leaseholder may hire qualified surveyors in order to assess the value of the property, and the negotiation can be lengthy. We denote these cases as *lease extensions*. This option is particularly relevant for leasehold owners of short leaseholds for whom lease expiration may be more of a concern than for owners of long leaseholds.

Although the 1993 Act legally provides the option to extend, almost all extensions are

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

[10]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.

negotiated out of court by leaseholders and their freehold landlords. Despite this, the court-recommended amount of 90 years is the most common extension length (accounting for about 30% of all extensions). A fraction of extensions, however, are for approximately 900 years, which effectively convert short leaseholds into long leaseholds.

We will now introduce some notation that we will use throughout the paper to refer to extended properties in our sample, which we show in a diagram in Figure 2. Consider a lease that transacts twice within the Land Registry Transaction Data Set. We say that a property was purchased at time $t-h$ and sold at time t, where h is the amount of time between purchase and sale. We are interested in properties which were extended at some time $t-h+u$ where $u<h$.

We denote the lease term to maturity, henceforth referred to interchangeably as 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 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 Appendix Figure A.2, properties with a very short holding period, h, are very likely to be "flippers" who purchase, renovate and sell properties with the explicit intention of making a quick profit and therefore behave differently from other owners. Unless otherwise specified, we exclude properties with $h \leq 2$ year from our analysis.

Table 1: Number of Extensions

Extension Amount	90	700+	Other	Total
2003-2005	635	1,867	1,910	4,412
2006-2010	3,448	6,832	5,057	15,337
2011-2015	11,707	15,399	10,357	37,463
2016-2020	16,720	18,785	11,633	47,138
2021-2023	11,823	12,351	5,677	29,851
Total	44,333	55,234	34,634	134,201

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.

In Table 1 we present the distribution of extensions for different amounts over time. In total, there are 134,201 lease extensions in our main sample, which is 5% of flats. How common are these lease extensions? In Appendix Figure A.3 we show the hazard rate of extension, defined as the conditional probability a property will extend given how long its remaining term to maturity is. We can see that almost no properties extended with more than 90 years remaining. After a property hits 80 years remaining, its extension probability jumps to a probability of extension of about 5%, and then slowly falls back to 2% or so.[11]

For much of our analysis, we will focus on leases that extend by the typical amount of 90 years. For these leases, the median duration before extension is large, at around 70 years, as shown in Appendix Figure A.5. The median time between transactions is 10 years, with 7 years between purchase and lease extension (see Appendix Figure A.6 and Appendix Figure A.7). In some specifications, we also study extensions of more than 90 years, in order to learn about the shape of the long term yield curve.[12]

3.3 Transaction Lags

There is a lag between when a house price is agreed versus when the transaction occurs. The date in the Land Registry is 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 bound 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 months, with a mean of about 4.5 months. The full distribution of the time elapsed between the listing date and the Land Registry date recorded is presented in Figure A.10. Therefore new information will affect our estimates with a lag of around 4.5 months.

4 Empirical Methodology

This section explains how to use the price gain from lease extension to estimate the expected long term housing yield. We start by deriving the price of a leasehold. The price of a leasehold

[11] We explain how we calculate the hazard rate in Appendix A.11.

[12] In Appendix A.1, we provide additional summary statistics about lease extensions, including a heatmap of the extension rate by region; and the distribution of transaction dates and extension amounts. We also show that leasehold owners are broadly representative of the overall UK housing market in terms of owner and mortgage characteristics and price-to-rent ratio business cycles in Appendix Table A.1 and Appendix Figure A.8, as well as the geographic distribution of extensions in Appendix Figure A.9.

P_t^T, with T years until expiration, and an option to extend by 90 years on expiration, is

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

This equation starts with the asset pricing identity of Equation (1), since the first term is 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, 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 non-negative—instead of extending the lease, the leaseholder can choose not to extend and receives zero payoff.

This equation clarifies that the option value of lease extension raises the value of a leasehold. Consider two cases. First, suppose $\alpha = 0$. Then, the leaseholder receives the entire value of a lease extension. Provided that the value of an extended lease is positive, the leaseholder will always choose to extend—the option of lease extension is always "in the money" (we abstract from uncertainty in the rate of return). Then $P_t^T = R_t \int_0^\infty e^{-\int_0^s y(u)du} ds$, meaning the price of a finite duration leasehold is the same as the price of an equivalent, but infinite duration asset. Since the property always costlessly exercises its option to extend, then the effective duration of the asset is infinite. A second instructive case is $\alpha = 1$. In this case, the leaseholder receives none of the value from extension, and the price of a leasehold is $P_t^T = R_t \int_0^T e^{-\int_0^s y(u)du} ds$. This price is the same as an asset with a duration of exactly T periods. 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.

In the main analysis, we will assume that $\alpha = 1$, so that the value of extension goes entirely to freeholders and not leaseholders. Therefore there is no option value from lease extension. This assumption is appropriate given the institutional features of the UK property market. As we discuss at length in Appendix A.13, the law recommends that leaseholders pay freeholders the entire value of lease extensions. In Appendix A.13 we will also consider a bunching estimator, which identifies α using discontinuities in legally mandated lease extension payments when $T = 80$. Our estimates of expected long term yields change by little when we estimate α directly.

4.1 A Difference-in-Differences Estimator of y^*

Next, we embed the formula for the lease extension price into a difference-in-differences estimator, in order to identify y^* for UK housing. We will study a difference-in-differences estimator Δ_{it} for the percent price change after lease extension:

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

In this equation, $\log P_{it}^{T+90} - \log P_{i,t-h}^{T+h}$ is the price change for a property i bought h periods previously, which extends by 90 years from a duration of T years remaining, at time t. 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 implies

$$\Delta_{it} = \log\left(\int_0^{T+90} e^{-\int_0^s y(u)du} ds\right) - \log\left(\int_0^{T} e^{-\int_0^s y(u)du} ds\right) + \Delta_{t,t-h}\left(\log R_{it} - \log R_{jt}\right). \quad (3)$$

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.

The identification assumption of our estimator is a form of "parallel trends". The growth in service flows for the treatment and control properties, before versus after extension, must be the same. If so, then the final term from Equation (3) vanishes.[13,14]

Finally, in order to implement the estimator we parameterize the shape of the yield curve $y(s)$. We make the simple assumption that $y(s) = y^*$, so that the yield curve is horizontal and equal to its terminal value, which is the long term expected yield. Shortly, we will see that this assumption accurately prices leaseholds in our quasi-experiment, even when there is significant variation in the slope of the yield curve at short horizons. The reason is that when T is large, the difference-in-differences estimator is primarily identified from long duration flows between T and $T + 90$. With this parameterization and the parallel trends assumption, we arrive at the final form of our estimator

$$\Delta_{it} = \log\left(1 - e^{-y^*(T+90)}\right) - \log\left(1 - e^{-y^* T}\right). \quad (4)$$

[13] The difference-in-differences estimator does not depend on the holding period h. Appendix Figure A.11 shows that estimates of Δ_{it} are uncorrelated with h.

[14] Equation (3) also imposes that the growth of long term service flows, after the holding period, does not differ for the treatment versus control property. We provide evidence for this assumption in Section 4.4.

4.2 Advantages of the Estimator

The key advantage of our estimator is that it "differences out" short run variation in property prices, from either short term yields or in the service flow of housing. This differencing out applies regardless of the precise specification of the short run shocks affecting property prices. Therefore our estimator is able to estimate expected long term yields without relying on a potentially mis-specified structural model.

In more detail, a first advantage of the estimator is that it differences out the service flow of housing, provided that that parallel trends assumption holds. Terms related to service flow do not appear in the value of the difference-in-differences estimator (4). The service flow includes taxes, depreciation, as well as the utility from consuming housing. This property of the estimator is appealing because the service flow is difficult to measure directly, especially for owner occupied housing. Moreover there may well be significant unobserved heterogeneity in this service flow across properties, which may vary across time. 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.

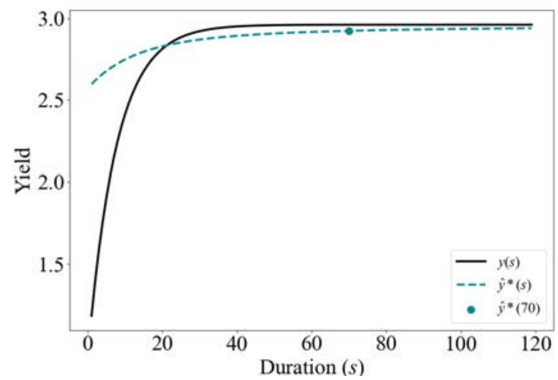
The second advantage of our estimator is that it differences out the effect of short term yields on asset prices, in order to estimate the long term. Therefore a transitory shock to yields, such as a monetary tightening, does not affect our estimate. Intuitively, suppose that the duration before extension T is large. The price before extension capitalizes flows over the first T periods. The price after extension capitalizes flows over $T+90$ periods. Therefore the price change at extension identifies the price of cashflows *after* T, only in the far future. However in practice, this argument is hard to demonstrate analytically.[15]

We illustrate the success of our estimator in differencing out short term yields numerically. In general, the success of our estimator depend on the shape of the yield curve before it converges to y^*, and on the duration of the property before it extends. Numerically, our estimator successfully estimates y^* under an empirically reasonable yield curve, and when it is applied to long duration properties. 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 3.0 pp.[16] 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

[15] In effect, we are studying the price of a positive coupon bond, which is not analytically tractable.

[16] 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 the 3-month London Interbank Offered Rate (LIBOR), $y(10)$ is equal to the 10-year gilt yield, and the average of $y(s)$ for $10 \leq s \leq 30$ is equal to the 10 Year 20 Year gilt forward yield. For all the bond yields, we use the mean yield for our sample period. We present other possible parameterizations of $y(s)$ in Appendix A.5.

Figure 3: A Parameterization of the Housing Yield Curve



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 UK 3-month LIBOR, the 10 year gilts and the 30 year gilts. $\hat{y}^*(s)$ is our estimator of y^* for each T, described below.

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 estimate of $\hat{y}^*$ at $T = 70$, which is approximately the median duration of leaseholds at extension.

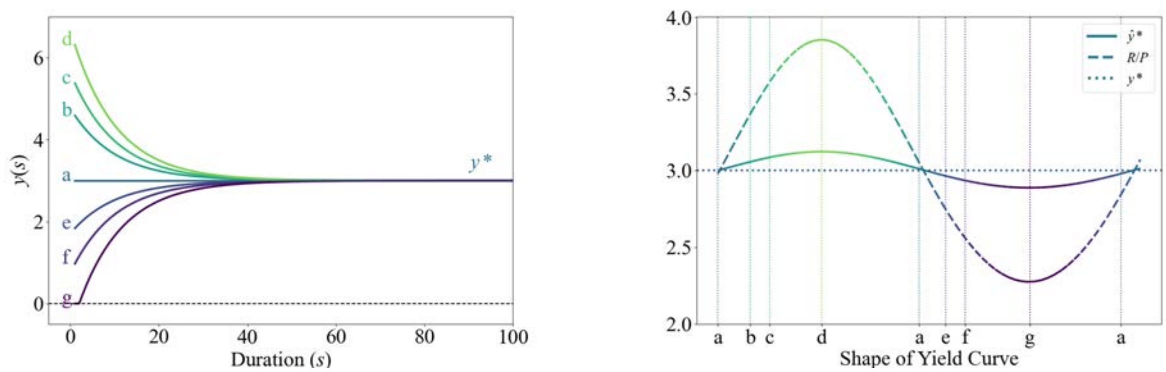
Our estimator produces tight estimates for a wide range of yield curves. In Figure 4a we present a time-varying yield curve for which the short-end fluctuates tremendously over time but the long end (y^*) is constant. For instance, the yield curve labelled d is very downward sloping, whereas yield curve g is very upward sloping. Then, the solid line in Figure 4b 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 4a are labelled accordingly. For instance, point d in Figure 4b corresponds to estimates of y^* for the downward sloping yield curve d of Figure 4a. Our estimator is relatively stable despite the fluctuations in the short end of the yield curve, and remains within 0.1% 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.[17]

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 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

[17] In Appendix A.5 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^*.

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 4b 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.

Figure 4: "Differencing Out" the Short End



(a) $y(s)$ **(b)** $\hat{y}^*$ and Rent-to-Price

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 the rent-to-price ratio is estimated such that $P_{it}^{\infty}/R_{it} = \int_t^{\infty} e^{-\int_t^s y(u)du} ds \equiv \frac{1}{R/P}$.

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".[18]

4.3 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 (4), for each individual property i our difference-in-differences estimator is

$$\Delta_{it} = \log\left(1 - e^{-y_t^*(T_{it}+90)}\right) - \log\left(1 - e^{-y_t^* T_{it}}\right). \tag{5}$$

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 (5) 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 Δ_{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 Δ_{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 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.[19] 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. To expand the available controls, we use repeat sales methods. For each extending property, we measure its counterfactual price growth using a repeat sales index of 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.

To provide intuition, consider an extending property bought at $t-10$ and sold at t. Suppose there is no neighboring property that is bought at $t-10$ and is sold at t. However, there are two other neighboring properties: the first bought at $t-10$ and sold at $t-5$; the second bought at $t-5$ and sold at t. By chaining the price growth of the two neighboring

[18] 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.

[19] Control properties are not treated, avoiding the "forbidden comparisons" problem (Borusyak et al., 2021).

properties, one creates a counterfactual price growth between $t-10$ and t for the extending property. More generally, consider a property that is bought and sold at $t-h$ and t. One can construct its counterfactual price growth using a repeat sales index, where the sample is all non-extending and neighboring properties that are either bought or sold in a window around $t-h$ and t.

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 122,224 of our 134,201 lease extension experiments.[20]

To construct the repeat sales control, we consider a treated property p, which was purchased at time $t-h$, sold at time t, and extended for 90 years at some time $t-h < t-h+u \leq t$.[21] Suppose this property has duration $T+h$ at purchase and duration $T+90$ at sale. The set of control properties, specific to property p, is those properties that (i) do not extend between $t-h$ and t; (ii) have a duration within 5 years of property p; and (iii) are within d km of property p. d is the smallest possible Haversine distance such that it is feasible to construct a repeat sales index for the counterfactual price of property p between $t-h$ and t. We discard property p if there are not enough controls within a 20 km radius. This procedure automatically adjusts for the different density of housing in urban and rural areas. Finally, we produce a repeat sales index for property p's control group, using a three-part procedure similar to Case and Shiller (1989) (see Appendix Section A.6 for details).

4.4 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. One way parallel trends could be violated is a renovation. Suppose that extending properties also renovate during their holding period. If so, then the price change after extension would include not only the value of extensions, but also the value of renovations.

We provide three tests in support of parallel trends. First, our treatment and control groups are close to being balanced on observable characteristics. We use the property charac-

[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] In our primary analysis, we measure t at a quarterly frequency. We show robustness to using other frequencies in Figure A.24.

teristics data provided by Rightmove and Zoopla. The results of regressing the same hedonic characteristics on an extension dummy are presented in Panel A of Table 2. Flats that have been extended at some point tend to have slightly better amenities than neighbouring flats of similar lease length. The differences in magnitude are small, however. For instance, the mean number of bedrooms, after controlling for geographical region, lease length, and listing year, in an extended property is 1.75, as opposed to 1.73 in a property that has not been extended. Average floor area is 67.5 vs 65 square meters for extended and non-extended properties, respectively. As a whole, rental prices, which ought to measure the quality of the property, are about 2% higher in extended properties.[22]

Table 2: Hedonic Characteristics in Extended vs Non-Extended Flats

Panel A: Levels

	(1) Num Bedrooms	(2) Num Bathrooms	(3) Num Living Rooms	(4) Floor Area	(5) Age	(6) Log Rental Price
Extension	0.022***	0.047***	0.005***	2.531***	7.848***	0.021***
	(0.002)	(0.001)	(0.001)	(0.099)	(0.179)	(0.001)
Fixed Effects	✓	✓	✓	✓	✓	✓
Control Mean	1.73	1.18	1.03	65	46.47	9.23
N	2,486,525	2,004,725	1,772,460	2,065,992	1,497,046	1,440,473

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel B: Differences

	(1) Δ Num Bedrooms	(2) Δ Num Bathrooms	(3) Δ Num Living Rooms	(4) Δ Floor Area	(5) Δ Log(Rent)
Extension	0.009	0.000	-0.001	0.102	0.007
	(0.005)	(0.001)	(0.001)	(0.086)	(0.009)
Fixed Effects	✓	✓	✓	✓	✓
Control Mean	.001	0	0	-.002	.012
N	1,155,049	939,406	832,303	963,454	644,156

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel A reports level of hedonic characteristics for extended properties relative to their non-extended counterparts. It presents estimates of β for the specification, $X_{it} = \alpha + \beta \mathbb{1}(\text{Extended}_{it}) + \Gamma_{it} + \epsilon_{it}$, for several hedonic characteristics X_{it} where Γ_{it} are 10-year duration bin × Local Authority × Year fixed effects. The sample is leasehold flats which appear at least once in the Land Registry Transaction Data Set. Panel B reports renovation rates in extended properties relative to their non-extended counterparts. It presents estimates from regression equation (6). For Columns 1-4, the sample includes properties with two separate Rightmove or Zoopla listings corresponding to two separate transactions before and after extension time. For Column 5, the sample includes properties that transact twice and have distinct rental price data before and after extension time.

As a second test of parallel trends, which directly focusses on renovations, we show that

[22] In Appendix Figure A.13 we show how our main hedonic variables vary with log price, controlling for Local Authority fixed effects. Notice that property size, bedroom number, and the log of rental prices all vary approximately linearly with log price. We also regress our hedonic characteristics on an experiment fixed effect and plot the density distribution of the residuals in Appendix Figure A.14. These residual plots visually confirm how similar the amenities distribution is between extensions and controls.

extending properties are no more likely to make home improvements than controls. We study Rightmove and Zoopla data for properties that have two distinct listings before and after extension. For these properties we estimate the following specification,

$$\Delta_h \text{Hedonic Characteristic}_{it} = \alpha + \beta \times \mathbb{1}\text{Extension}_{it} + \Gamma_{it} + \epsilon_{it} \tag{6}$$

where Γ_{it} are Local Authority $\times$ duration $\times$ year of listing fixed effects. The regression results are presented in Panel B of Table 2. We find no significant evidence that extended properties renovate at a different rate than other properties.

As a third 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.

We consider every treated property p and every property in p's control set 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

$$\Delta \log R_{i,t',t''} = \alpha + \beta \mathbb{1} \times \text{Extension}_i + \Gamma_{i,t',t''} + \epsilon_{i,t',t''} \tag{7}$$

where $\Delta \log R_{i,t',t''}$ is growth in rents between periods t' and t'' for which there are observed rents for both the treatment and control, $\Gamma_{i,t',t''}$ are experiment $\times$ rental transaction years fixed effects. The regression measures whether rental growth was higher for extended properties than for control properties 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. In the second column, we exclude rental transaction pairs for which the rental price does not change, which comprise about 15% of the data. In the third column, we only consider rental transactions that are strictly within the experiment window, and exclude those which are listed in the same years as $t-h$ or t, since properties which are rented out the same year that they are sold may be unique. In the fourth column, we do both of these checks. In the fifth column, we expand the sample size by comparing annualized rent growth in treated and control properties, without controlling for the years of listing. For each treated property we calculate its mean annual rent growth and the mean annual rent growth for its controls, and then take the difference. In the sixth column, we produce a repeat sales index for rent growth in each extending property's control group, and then re-estimate regression equation (7) using the repeat sales index control. The resulting estimates are small and negative—the opposite sign of what would happen if the extending property renovates.

Table 3: Within-Experiment Window Rent Growth

	100×Δ log(Rent)					
	(1)	(2)	(3)	(4)	(5)	(6)
Extension	0.12	0.15	-0.04	-0.04	0.03	-0.24*
	(0.09)	(0.12)	(0.10)	(0.13)	(0.06)	(0.11)
Experiment	✓	✓	✓	✓	✓	✓
× Rent Years FE	✓	✓	✓	✓		✓
Incl. Transaction Years	✓	✓				
Incl. Δ log(Rent)=0	✓		✓			
Annualized					✓	
RSI						✓
N	3,744,957	3,086,231	3,209,535	2,636,413	4,062,366	79,348

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports rent growth during the experiment window of $t-h$ to t for extended properties relative to control properties. In the first column, we estimate $\Delta \log R_{it} = \alpha + \beta \mathbb{1}(\text{Extended}_{it}) + \Gamma_{it} + \epsilon_{it}$ where Γ_{it} are experiment x rent listing years fixed effects. In the second column we exclude properties for which $\Delta \log R_{it} = 0$ and in the third column we exclude rental listings at times $t-h$ and t. In the fourth column we conduct both robustness checks. In the fifth column, we only control for experiment fixed effects, and have annualized rent growth as the LHS variable. In the sixth column, we build a repeat sales index of rent growth and compare rent growth in treated properties to the implied control rent growth from the index.

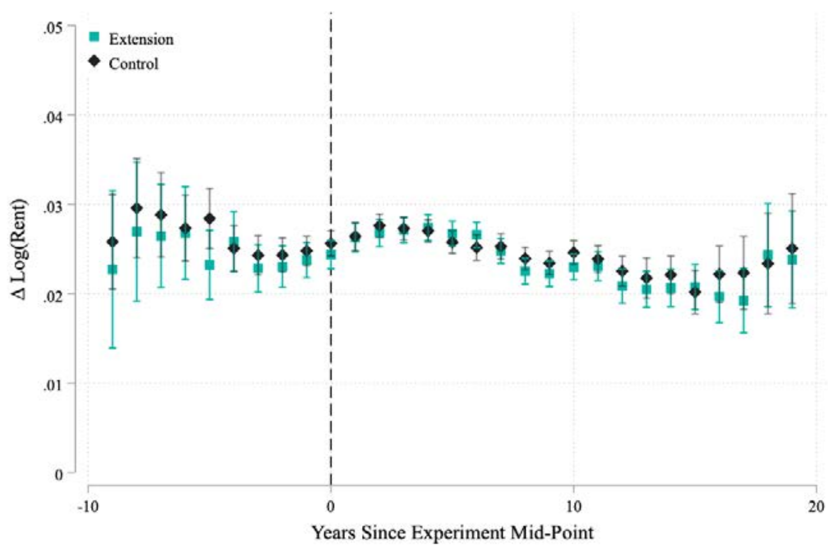
In addition to our tests of the parallel trends assumption, we also construct a second control measure, that explicitly accounts for hedonic differences between the treatment and the control group. In particular, we regress prices on hedonic controls, namely the number of bedrooms and floor area. Then, we create a residualized price variable, $P_{it}^{(res)} = \overline{P} + \epsilon_{it}$ where $\overline{P}$ is the mean price level across all flats. $P_{it}^{(res)}$ represents the price of property i at time t, taking out the average effect of its hedonic characteristics at time t. We can then adjust for hedonic characteristics, by residualizing prices for both extensions and their controls in this way. However we only observe hedonic characteristics for a fraction of our experiments (e.g. we have data on bedroom count for 110 thousand and data on floor area for 88 thousand of our 134 thousand experiments). Therefore the majority of our analysis is conducted using the first control measure, with the hedonic control measure as robustness.[23]

Finally, we revisit an assumption that we made to derive the difference-in-differences estimator (3), namely that the growth of long term service flows after the holding period, is

[23] In much of our analysis, we focus on extensions of 90 years. Appendix Table A.2 repeats the analysis of Table 2 on that subsample only.

the same for the treatment and control property. We can assess this assumption by measuring long term rent growth. In Figure 5 we plot 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 missing rent data.[24] Reassuringly, rent growth is similar for extending and control properties at all horizons including the long term, and is never statistically different.[25]

Figure 5: Rent Growth At Long Horizons



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% confidence intervals.

5 Empirical Results

This section presents our main empirical results. We estimate the level of y^* for UK property to be 3.4%, with a decline from about 5.2% in 2006 to 2.8% in 2023. Moreover the fall takes place before 2020 while y^* is relatively stable afterwards.

[24]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.

[25]For robustness, we repeat Figure 5 excluding observations with zero rent growth, and exclude observations with more than 5 years between listings (Figure A.15).

5.1 Event Study Representation of Lease Extension

The duration of properties when they extend is long. Leaseholds that extend by 90 years have a median duration at sale of 157 years—meaning had they not extended, they would have had a duration of 67 years at sale.[26]

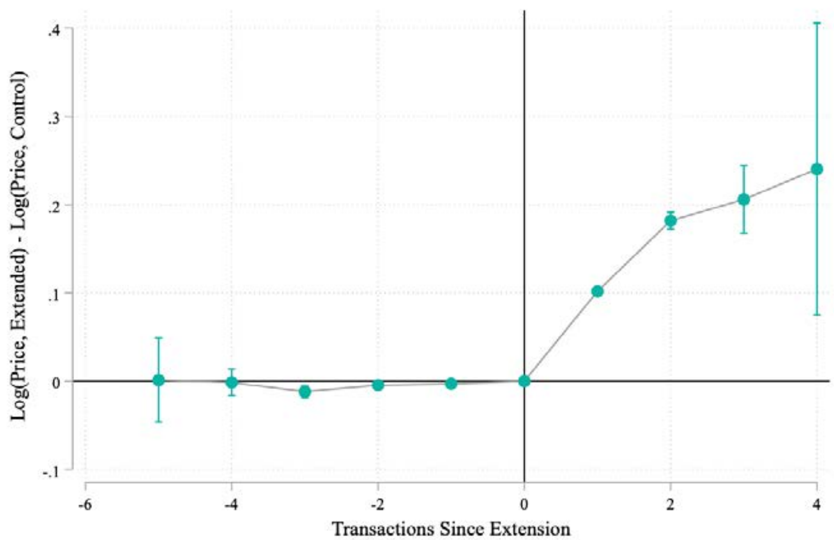


Figure 6: Event Study Representation of Lease Extension

The figure presents an event study representation of a lease extension. The line represents the difference in price growth for extended properties relative to their controls relative to the nearest transaction before extension, which is normalized to zero. The sample includes all 90-year extensions. 95% confidence intervals are shown, though they are so small that they overlap with the scatter plot market for several points.

To visualize the effect of lease extensions on prices, we create an event study for the leases that extend for 90 years. For each transaction of extended properties before and after the extension event, we calculate the mean difference in price change relative to that property's repeat sales index control. We plot these price differences for all of the extension experiments. We use the number of transactions before the extension as our x axis variable, where the transaction at time 0 is the last transaction before the extension. We normalize the difference between the treatment and control to be zero at time 0. The result is in Figure 6. After extension, the difference in log price between extended properties and their controls jumps by 0.1 log points, or about 10.5%. The difference continues to grow after this point because as time from extension increases, the lease term falls more, which results in a greater predicted difference in price between extension and control according to our simple

[26] Appendix Figure A.16 displays this information in a histogram, also including the duration of the control observations that did not extend.

asset pricing model.[27] Moreover, consistent with our identification assumption, there are small and statistically insignificant "pre-trends" before lease extension.[28]

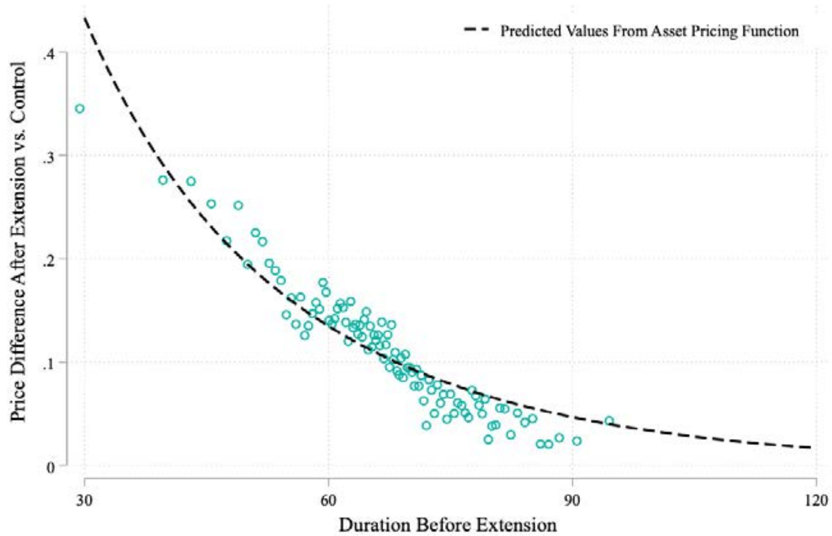
5.2 Estimating the Level of y^*

We present estimates of the level of y^*. We estimate Equation (5),

$$\Delta_{it} = \log\left(1 - e^{-y^*(T_{it}+90)}\right) - \log\left(1 - e^{-y^* T_{it}}\right)$$

using all lease extension experiments i. There are two sources of variation identifying y^*. First, the estimating equation predicts that the price gain is larger when y^* is lower, because a lower value of y^* raises the value of duration. Second, the gain from extension varies with duration, T_i. Leaseholds with shorter remaining terms will receive the benefit of extension sooner, which leads to greater price growth upon extension.

Figure 7: Duration Before Extension vs. Price Gain After Extension



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 (5).

Now we plot the variation from the data that identifies y^* according to our estimator. Figure 7 binscatters the difference-in-difference estimates Δ_{it} for various durations T_i, for

[27] Appendix Figure A.19 investigates this point in detail. There, we show that the price gain of the treated property is increasing versus its control as the years since the extension increase. Intuitively, the gap between the extender and control grows over time because falling duration affects properties with an initially short duration (the controls) more than properties with an initially long duration (the treated).

[28] Figure A.18 presents the event study including extensions of sizes other than 90 years.

the sample of leaseholds that extend by 90 years. As predicted by the estimating equation, the percent increase in property value as a result of extension increases in the duration before extension—ranging from only 7% for 90-year duration properties to more than 30% for properties with duration of 40 years at extension. The black curve in Figure 7 presents fitted estimates from our nonlinear estimator. Inspection of Figure 7 indicates that our estimate of y^* is a good fit across both the average price gain after an extension, and variation in the price gain across T—the two statistics that identify y^*.[29]

Table 4: Estimated y^*

	Constant y^*	Flexible y^*				N
	(1)	(2) $T = 50$	(3) $T = 60$	(4) $T = 70$	(5) $T = 80$	(6)
No Hedonics	3.37***	3.36***	3.37***	3.39***	3.40***	40,633
	(0.02)	(0.03)	(0.02)	(0.02)	(0.04)	
No Hedonics (Hedonics Sample)	3.29***	3.19***	3.27***	3.36***	3.45***	25,343
	(0.02)	(0.04)	(0.03)	(0.03)	(0.05)	
Hedonics	3.32***	3.16***	3.29***	3.43***	3.57***	25,343
	(0.03)	(0.05)	(0.03)	(0.04)	(0.06)	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 01$

Column 1 presents estimates of y^* from Equation (5) for 90 year extensions. Column 2-5 present estimates of $y^*(T)$ for $T \in \{50, 60, 70, 80\}$, where we parameterize y^* linearly as a function of T, $y^*(T) = \gamma + \beta \cdot T$. The first row presents estimates using the baseline procedure to find controls described in Section 4.3, which utilize the raw transaction price variable as the main outcome variable. The second row presents estimates using the raw transaction price variable, but restricted to the sample of observations which are not missing hedonic characteristics. The third row presents estimates of y^* and $y^*(T)$ produced using transaction price residualized on hedonic characteristics as the main outcome variable, as described in Section 4.3. Standard errors are heteroskedasticity robust.

Next, we present estimates of the y^* implied by Figure 7. We estimate $y^* = 3.4\%$, indicated in Column 1 of Table 4. The second and third rows of Table 4 show that these estimates are largely unchanged by inclusion of hedonic controls.

As we discussed in Section 4.2, our estimator assumes that the long term yield curve for housing is flat, meaning a unique value of y^*. We now provide two pieces of evidence for this assumption. First, we estimate the degree to which estimates of the long term yield vary with duration, by parameterizing our estimate of y^* as $y(T) = \gamma + \beta \cdot T$, where T is the duration of the leasehold at extension. If the long term yield curve is sloped, then β will be different from zero. We exploit variation in duration at the time of extension to estimate both γ and β. In Columns 2-5 of Table 4, we present estimates of $y(T)$ under this more flexible form for our sample range of T. We can see that $y(T)$ varies little over the range we

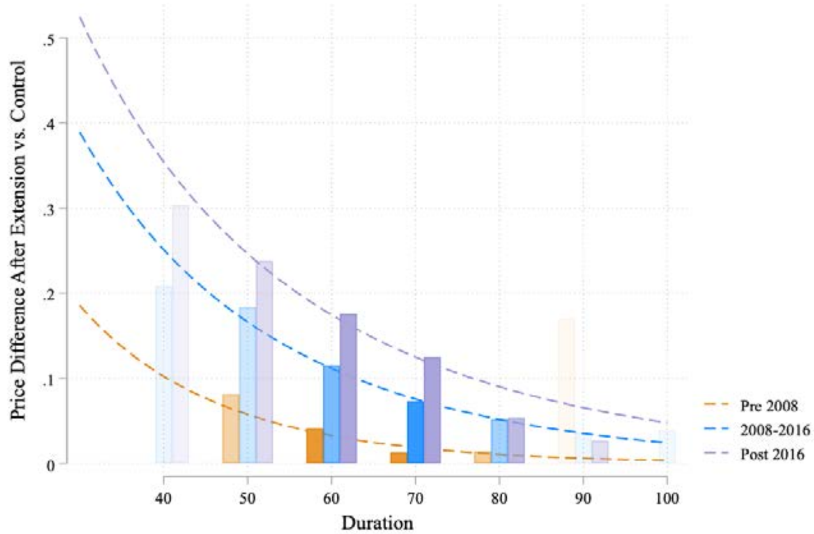
[29]We can explicitly reject a linear relationship between Δ_{it} and T_{it} by running the following quadratic regression: $\Delta_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 T_{it}^2 + \epsilon_{it}$. The estimated β_i's are all statistically significantly different from zero.

observe it, suggesting a flat long term yield curve.[30] As a second test, 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. Appendix Table A.3 reports estimates of y^* for both the longer and shorter extensions. Estimates of y^* are remarkably similar for both groups. Therefore y^* seems to vary little at very long horizons, consistent with a flat yield curve.[31]

One concern about our estimate may be that short duration leaseholds have an additional liquidity premium because banks are less willing to issue a mortgage against shorter duration leaseholds. We investigate this concern in Appendix A.7 and find that liquidity premia do not seem to affect our results: from survey data, short duration leaseholds seem to have similar characteristics to other leaseholders and seem to be similarly liquid. Explicitly incorporating liquidity frictions into our estimator affects the value of y^* little. [32]

5.3 Trend Dynamics of y^*

Figure 8: Price Change From Extension, Over Time



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 Δ_{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.

[30] In Figure A.20 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.

[31] In Figure A.17 we present a binscatter using properties that were extended for more than 700 years.

[32] In Appendix A.8, we investigate the seasonality of our estimate of y^*. Spot UK house prices are highly seasonal (Ngai and Tenreyro, 2014). However our estimate of y^* is not seasonal, consistent with its interpretation as a long run object.

Now we present our first key result: estimates of the trend dynamics of y^*. Interest rates have been falling in the UK, as in much of the rest of the world, for at least four decades. Can the same be said for y^*? We find a sizable fall in y^* across the entire yield curve between the start and end of our sample period. This decline in y^* is depicted in Figure 8. In this figure we show that at all durations, the price growth associated with lease extension Δ_{it}, has increased by approximately 10pp between the beginning and end of our sample. The opacity of the bar graph is shaded by the number of observations in the bar, and contains only 90 year extensions. An increase in the price growth from lease extension implies that the value of long duration cashflows has increased—that is, y^* must have fallen.[33]

Given our large data set, we are able to estimate the dynamics of y^*_t with precision. We start in 2003, at the beginning of our sample. We have already presented the estimates in Figure 1 in the Introduction, which contains estimates of y^*_t at annual frequency. Our time varying estimate includes all lease extensions, instead of only 90 year extensions, in order to raise the precision of our estimates; the shaded region is 95% confidence intervals for the full sample of leases.[34]

The estimates show that the fall in y^* has been relatively consistent throughout this entire sample period. Between 2003 and 2023, we estimate y^* has fallen from around 5.4% to 2.8% — an approximately 50% decline. The magnitude of this decline is large, corresponding to a doubling of the long run expected price-rent ratio. Notably, our estimates are stable during the 2020 pandemic, despite considerable volatility in shorter term asset prices during this period.

Our estimate of a declining *housing* yield before 2022 is related to the well known fact that *government bond* yields fell around the world before 2022 (e.g. Holston et al., 2017). However bond yields could have fallen due to factors that are specific to safe assets, such as quantitative easing or rising demand for the safety and liquidity of government bonds (e.g. Krishnamurthy and Vissing-Jorgensen, 2012; Caballero and Farhi, 2018). These safe asset-specific factors may be unrelated to the yield on capital in the real economy. To the contrary, our estimates show that other forms of capital also experienced large trend declines in yields before 2022, similar to bonds.

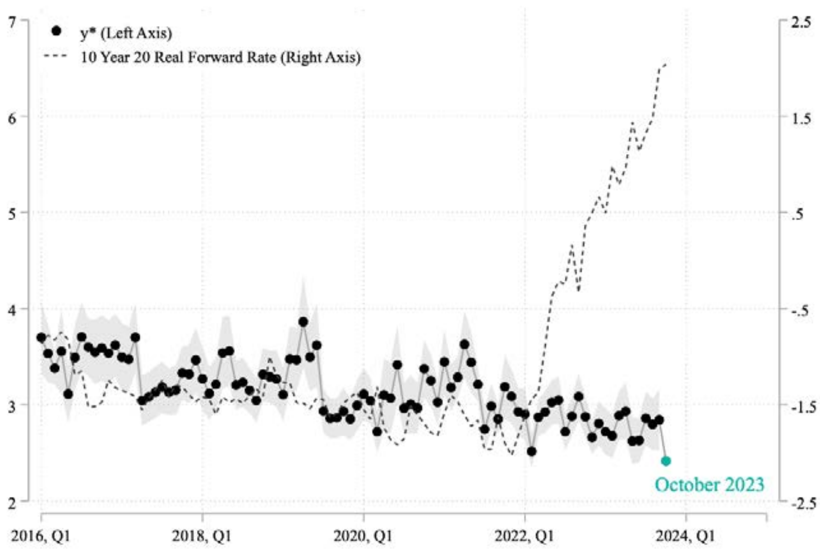
[33] In Figure A.21 we plot the event study of Figure 6 separately for each period, which also documents the lack of pre-trends at all times.

[34] Appendix Figure A.22 plots the time varying estimate y^*_t using various extension amounts shows similar results, again consistent with a flat long term yield curve. We are unable to estimate y^* for 90-year extensions in 2008, because the mean estimated difference-in-difference is negative due to low sample size and volatility from the housing crisis.

5.4 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 October 2023, the last month for which UK housing data are available. The shaded area represents the 95% 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.[35]

Figure 9: Monthly estimates of y^* and rate of return on government bonds



The solid line presents y_t^* estimated at a monthly frequency using all extensions. The line is plotted from January 2016 to October 2023. 95% confidence interval is shaded.

Two observations from the monthly time-series are worth mentioning: the first is that there is a barely noticeable change in y^* during the Pandemic Recession, despite considerable volatility in housing markets over this period. This behavior is consistent with the 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

[35] As discussed in Section 3.3, our y_t^* estimates include new information with a lag of about 5 months.

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 5 months.

5.5 Robustness Checks

There are several variations of our estimator that we can run to verify that the estimate is robust. Our difference-in-difference 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 et al. (1963) (BMN). Appendix Figure A.24 presents in the black line our baseline estimates and in the dashed grey 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 2.4pp.

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. Appendix Figure A.25 presents estimates of y^* using the exact control method in the dashed grey 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.

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 (1) the average over the full period, and (2) the end of period distribution. The results are presented in Appendix Figure A.26 and are similar to the baseline estimates.

5.6 The Advantage of Natural Experiments

Our microdata-based approach to estimating y^* builds on the key insight of Giglio et al. (2015), which was the first paper to observe that UK properties are uniquely well suited to estimating long term housing yields, because of their varying duration. Giglio et al use a cross-sectional comparison of freeholds and leaseholds with different duration to estimate the level of y^*. Building on their insight, our approach uses the quasi-experiment of lease extensions to estimate the dynamics of y^*. We now elaborate on some advantages of quasi-experiments compared with cross-sectional variation. However it is important to bear in mind that estimates of y^* with each method are not strictly comparable, since the sample of properties and their duration is different in each case.

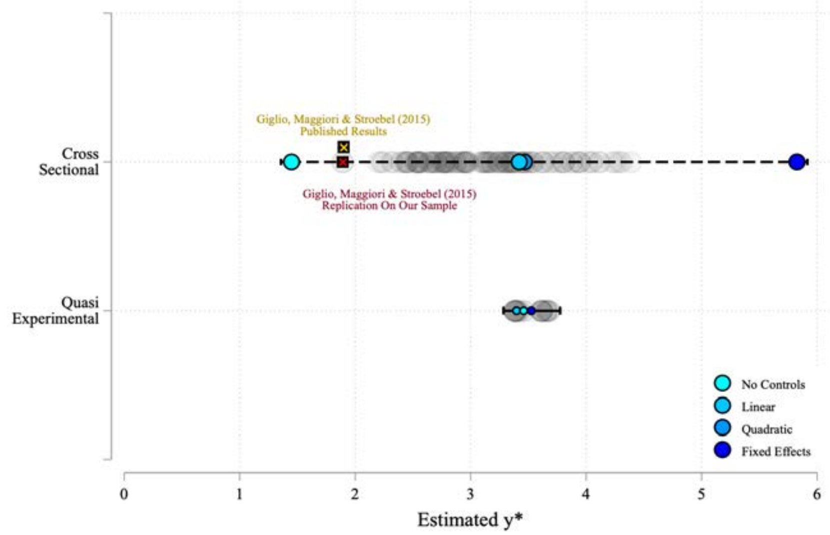
An important difference between the cross-sectional approach and the quasi-experimental approach is that in the former, the primary source of variation for duration is across properties rather than within properties. Long duration properties might have differences in the service flow of housing. For example, freehold flats might have higher quality construction. As such, the cross-sectional approach relies on detailed hedonic characteristics to control sources of variation in property price associated with the service flow. However with either the cross sectional or the quasi-experimental approach, unobserved heterogeneity may bias the estimates.

To gauge the effect of unobserved heterogeneity on the quasi-experimental and the cross sectional estimates, we study the sensitivity of the estimates to *observed* heterogeneity, in the spirit of Altonji et al. (2005) and Oster (2019). 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 yet another we allow price to vary quadratically with these same controls. 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.[36] The other variations include all possible subsets of these seven characteristics.

We then plot our estimates of y^* under each variation for both the quasi-experimental and cross-sectional methodologies in Figure 10. Under the cross-sectional approach, the estimates vary tremendously from 1.3% (in the case of no hedonic controls) to more than 5%. In contrast, our quasi-experimental estimates of y^* are highly stable around 3.5%. These results provide tentative evidence that our quasi-experimental methodology offers estimates of the expected long-term yield housing that are relatively robust to unobserved heterogeneity.

[36] The fixed effects controls are the same as the main specification Giglio et al. (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.

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



The figure presents estimates of y^* under various choices of hedonic controls. For the cross-sectional estimates, we estimate $\log P_{it}^T - \log P_{it}^\infty = \log(1 - e^{-y^*T})$ 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 4. 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 et al. (2015), and the red cross presents our replication of their estimate, using the full data from 2003-2023.

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 (1) they are not time-invariant, for instance if the property has renovated, or (2) 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 Figure A.23, which again indicate that the effect of hedonic controls on the quasi-experimental design is minimal.

6 Macroeconomic Implications of y^*

This section discusses three implications of our estimates of y^* for the macroeconomy. First, we find suggestive evidence that our estimate of y^* captures the behavior of expected long term yields for assets other than UK housing. Second, our estimates suggest that r^*, the expected long term yield of safe assets, has not risen after 2020. Third, cross-sectional heterogeneity in the dynamics of y^* suggests that constrained housing supply is important for its evolution.

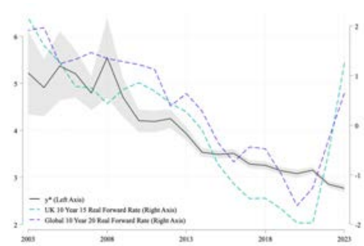
6.1 Expected Long-Term Yields of Other Assets

Our estimate of expected long-term yields comes from UK housing. We now present suggestive evidence that the dynamics of expected long-term yields are likely to be similar for other assets. We show that at *low* frequencies, the dynamics of y^* match the dynamics of long-term expected yields on other assets. This finding suggests that at *high* frequencies, the dynamics of y^* are similar to other assets, though expected long-term yields cannot be calculated for other assets at high frequency.

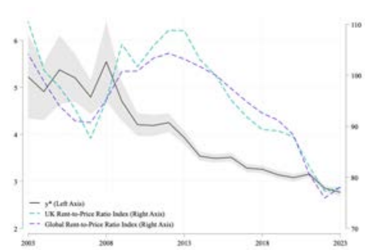
Information about expected long term yields is not directly available for other assets, which lack our natural experiment. However, we can approximate the dynamics of expected long term yields for other assets at low frequency, using the long run trend of yields. The long run trend "averages out" short term shocks to yields, and therefore tracks low frequency movements in the long-term expected yield.[37] If long run trends for other forms of capital match y^*, then the dynamics long run expected yields are similar for UK housing and other assets, albeit at low frequency.

However our measure of expected long term yields has a major advantage, being available at high frequency. Long run trends can only update gradually, meaning policymakers cannot use this information in a timely manner.

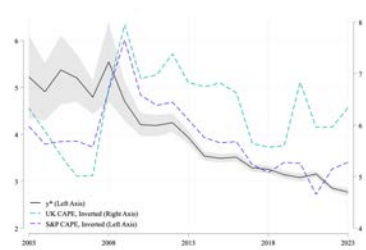
Figure 11: Comparing Other Asset Yields to y^*



(a) Real Risk-Free Forward Rate



(b) Rent-to-Price Ratio



(c) Inverted CAPE

The figures present yields for other assets over the 2003-2023 period. Panel (a) plots the 10 year 15 real forward rate for the UK, using data on real forward curves from the Bank of England. It also plots the weighted average 10 year 20 forward rate for 9 OECD countries for which data from Global Financial Data is available. 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% expected inflation. Panel (b) plots the UK rent-to-price ratio index from OECD Data and the weighted average rent-to-price ratio index for 22 OECD countries for which it is available from OECD data. Panel (c) plots the inverse of the Cyclically Adjust Price-Earnings ratio (CAPE) for the UK, based on data from Barclays and for the S&P based on updated data from Shiller (2006).

[37]Formally, consider the yield of an asset z_t, which follows a driftless ARIMA process. 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 shows that the long run trend of various other yields has been similar to y^*, particularly prior to 2020. In panels (a)-(f), we plot long run safe asset yields for the US and worldwide, rent-price ratios for the US and worldwide, and earnings yields for the UK and US stock markets. In most cases, the long run trend for these other asset yields is similar to y^*, particularly before 2020. 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). Our contribution is a measure that captures similar information but at higher frequency.

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^* that is common to all assets, whereas movements in risk premia ζ^* and dividend growth g^* are specific to housing. Consistent with this logic, in Appendix A.9, 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 capital gains using a standard Vector Autoregression approach (e.g. Campbell and Shiller, 1988), and show that neither long run UK housing risk premia, nor capital gains 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 et al., 2017). We stress that this exercise is tentative due to the uncertainty of the VAR based procedure.

6.2 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. 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).

Our estimates of y^* can help to answer this question. We stress that $y^* \equiv r^* + \zeta^* - g^*$ and r^* are different objects. However, y^* contains a great degree of information about r^*. In particular, the stability of y^* after 2020 suggests that r^* has also been stable.[38] This finding is striking because medium term bond rates have risen sharply since 2020, as Figure 9 shows. 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. Our real-time estimates of y^* will allow us to track this convergence process on a monthly

[38] 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.

basis.

Our estimates have two advantages relative to prevailing information about r^*, which typically relies on structural time series estimates. First, our microdata-based estimate of y^* is precise, and therefore weighs against meaningful increases in r^*. Estimates of long-term yields that do not use microdata, and rely on time series variation, are significantly less precise. For instance, the Holston et al. (2017) estimate of r^* for the UK had a mean standard error of r^* between 2000 and 2019 of 4.3pp — an order of magnitude greater than our real-time quarterly standard errors. The difference in precision is an intrinsic advantage of using microdata. Time series data is relatively uninformative about the long run, meaning inference about long run parameters is uncertain (Farmer et al., 2021). The micro-data approach instead uses informative cross-sectional variation to make inferences about long run parameters.

A second advantage of our natural experiment approach is that it is arguably less vulnerable to model misspecification. Structural time series methods, on the other hand, do potentially suffer from this concern. Some estimates of r^* find that it has risen significantly, whereas other estimates find r^* unchanged (Baker et al., 2023). These estimates are sensitive to the exact structural model used to identify r^*. Our approach, however, relies on a difference-in-differences strategy that is arguably less vulnerable to misspecification.

6.3 Cross-Sectional Heterogeneity in y^*

The expected long term housing yield also encodes other useful information about the economy, such as 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 the rise in valuations and fall in 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 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. Because a Local Authority's land supply elasticity is potentially endogenous, Hilber and Vermeulen (2016) also 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 particularly lowered the application delay rate for inelastic LAs, which were more likely to delay construction applications before the reform.[39]

Table 5: Cross-Sectional Heterogeneity in Δy^*

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) IV	(6) IV
Refusal Rate	-3.54*	-5.51***	-6.28***		-10.06*	-11.15**
	(1.37)	(1.56)	(1.83)		(4.31)	(3.53)
Share Developed		-1.43***	-1.50*			-2.02***
		(0.42)	(0.57)			(0.55)
Change in Delay Rate				2.27**		
				(0.83)		
Region FE			✓			
N	122.00	122.00	122.00	122.00	122.00	122.00
R2	0.08	0.22	0.34	0.09	.	0.05

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

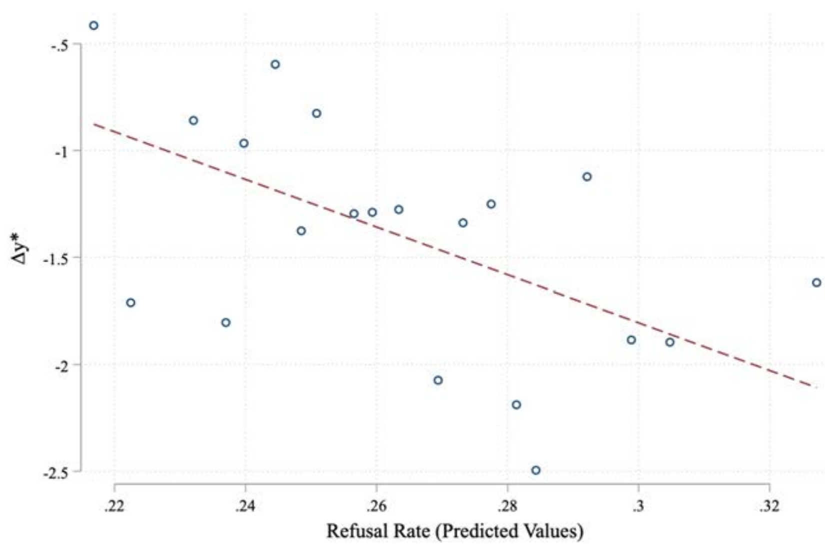
The table presents cross-sectional regressions of the long-run change in y^* against Local Authority level measures of supply-side constraints in the UK. The long-run change in y^* is defined as the difference between post-2022 y^* and pre-2009 y^* (inclusive). We exclude data from 2021 because of the Pandemic Recession's disproportionate effect on London. 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 (5) and (6) present IV regressions, where the refusal rate is instrumented using the change in delay rate after 2002. In all the regressions, we utilize the variation of Δ_{it} built using an annual repeat sales index, 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. Figure A.24 verifies that estimates using the annual RSI are very similar to the quarterly RSI. All the presented estimates are similar if using a quarterly RSI.

Table 5 shows the relationship between housing supply elasticity and the observed decline in y^* for each Local Authority in our sample. 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 choosing other cutoff years. Regressions are weighted by the inverse of the sum of the variances of the pre and post y^* estimates. Column 1 of Table 5 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 region fixed effects. 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,

[39] Appendix Figure A.27 verifies this first-stage, showing that LAs with higher mean refusal rates experienced greater declines in the delay rate after 2002.

but also include the share developed as a co-variate. In all cases, the results indicate that more regulated and land-constrained Local Authorities experience greater declines in y^*. Based on the Column 2 and holding the share of land developed constant, the decline in y^* is more than 1pp greater in magnitude for Local Authorities at the 90th percentile of refusal rate, relative to those at the 10th 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^*.

Figure 12: The Decline in y^* Depends on Housing Supply Elasticity



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 5.

Our finding is not surprising: as demand for housing rises, valuations increase by more in areas with inelastic supply. However our estimates have implications for the overall economy. The large decline in y^* indicate that housing supply is inelastic in the UK 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.

7 Concluding Remarks

This paper estimates the expected long-term yield of housing and its dynamics from 2003 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% in 2006 to 2.8% in 2023.

Long-run yields are valuable because they "look through" the short term factors affecting asset prices in real time. There has been a growing gap between long-run yields and real forward yields in 2022-23. This gap will narrow either by a fall in spot real yields, as the "secular stagnation view" suggest, or a rise in long-run yields if 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, "When Should You Adjust Standard Errors for Clustering?," *The Quarterly Journal of Economics*, 2023, *138* (1), 1–35.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 2005, *113* (1), 151–184.

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

Bailey, Martin J., Richard F. Muth, and Hugh O. Nourse, "A Regression Method for Real Estate Price Index Construction," *Journal of the American Statistical Association*, December 1963, *58* (304), 933–942.

Baker, Katie, Logan Casey, Marco Del Negro, Aidan Gleich, and Ramya Nallamotu, "The Post-Pandemic r*," August 2023.

Beveridge, Stephen and Charles R Nelson, "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*, 1981, *7* (2), 151–174.

Blanchard, Olivier, "Secular Stagnation Is Not Over," January 2023.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting Event Study Designs: Robust and Efficient Estimation," August 2021.

Bracke, Philippe, Edward W. Pinchbeck, and James Wyatt, "The Time Value of Housing: Historical Evidence on Discount Rates," *The Economic Journal*, August 2018, *128* (613), 1820–1843.

Caballero, Ricardo J. and Emmanuel Farhi, "The Safety Trap," *The Review of Economic Studies*, 2018, *85* (1 (302)), 223–274.

Campbell, John Y and Robert J Shiller, "Stock Prices, Earnings, and Expected Dividends," *the Journal of Finance*, 1988, *43* (3), 661–676.

Case, Karl E. and Robert J. Shiller, "The Efficiency of the Market for Single-Family Homes," *The American Economic Review*, 1989, *79* (1), 125–137.

Christensen, Jens H. E. and Glenn D. Rudebusch, "New Evidence for a Lower New Normal in Interest Rates," *FRBSF Economic Letter*, 2017.

Crump, Richard K., Stefano Eusepi, and Emanuel Moench, "The Term Structure of Expectations and Bond Yields," *Staff Reports*, May 2016, (775).

Farhi, Emmanuel and François Gourio, "Accounting for Macro-Finance Trends: Market Power, Intangibles, and Risk Premia," September 2018.

Farmer, Leland, Emi Nakamura, and Jón Steinsson, "Learning about the Long Run," Technical Report, National Bureau of Economic Research 2021.

Genesove, David and Lu Han, "Search and Matching in the Housing Market," *Journal of Urban Economics*, 2012, *72* (1), 31–45.

Giglio, Stefano, Matteo Maggiori, and Johannes Stroebel, "Very Long-Run Discount Rates," *The Quarterly Journal of Economics*, February 2015, *130* (1), 1–53.

_ , _ , and _ , "No-Bubble Condition: Model-Free Tests in Housing Markets," *Econometrica*, May 2016, *84* (3), 1047–1091.

Gomme, Paul, B. Ravikumar, and Peter Rupert, "Secular Stagnation and Returns on Capital," *Economic Synopses*, 2015, (19).

Hamilton, James D., Ethan S. Harris, Jan Hatzius, and Kenneth D. West, "The Equilibrium Real Funds Rate: Past, Present, and Future," *IMF Economic Review*, November 2016, *64* (4), 660–707.

Harrison, J. Michael and David M. Kreps, "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations," *The Quarterly Journal of Economics*, 1978, *92* (2), 323–336.

Hilber, Christian A. L. and Wouter Vermeulen, "The Impact of Supply Constraints on House Prices in England," *The Economic Journal*, March 2016, *126* (591), 358–405.

Holston, Kathryn, Thomas Laubach, and John C. Williams, "Measuring the Natural Rate of Interest: International Trends and Determinants," *Journal of International Economics*, May 2017, *108*, S59–S75.

Johannsen, Benjamin K. and Elmar Mertens, "The Expected Real Interest Rate in the Long Run: Time Series Evidence with the Effective Lower Bound," *FEDS Notes*, February 2016, *2016* (1703).

Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M Taylor, "The Rate of Return on Everything, 1870–2015," *The Quarterly Journal of Economics*, 2019, *134* (3), 1225–1298.

Kiley, Michael T., "What Can the Data Tell Us About the Equilibrium Real Interest Rate?," September 2015.

Krishnamurthy, Arvind and Annette Vissing-Jorgensen, "The Aggregate Demand for Treasury Debt," *Journal of Political Economy*, 2012, *120* (2), 233–267.

Laubach, Thomas and John C Williams, "Measuring the Natural Rate of Interest," *The Review of Economics and Statistics*, November 2003, *85* (4), 1063–1070.

Lin, Zhenguo and Kerry D. Vandell, "Illiquidity and Pricing Biases in the Real Estate Market," *Real Estate Economics*, 2007, *35* (3), 291–330.

Lippman, Steven A. and John J. McCall, "An Operational Measure of Liquidity," *American Economic Review*, 1986, *76* (1), 43–55.

Lubik, Thomas A. and Christian Matthes, "Calculating the Natural Rate of Interest: A Comparison of Two Alternative Approaches," *Richmond Fed Economic Brief*, 2015, (Oct).

Miles, David and Victoria Monro, "UK House Prices and Three Decades of Decline in the Risk-Free Real Interest Rate," December 2019.

Negro, Marco Del, Domenico Giannone, Marc P. Giannoni, and Andrea Tambalotti, "Safety, Liquidity, and the Natural Rate of Interest," March 2017.

_ , _ , _ , and _ , "Global Trends in Interest Rates," *Journal of International Economics*, May 2019, *118*, 248–262.

Ngai, L. Rachel and Silvana Tenreyro, "Hot and Cold Seasons in the Housing Market," *American Economic Review*, December 2014, *104* (12), 3991–4026.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, April 2019, *37* (2), 187–204.

Rachel, Lukasz and Lawrence H. Summers, "On Secular Stagnation in the Industrialized World," August 2019.

_ and Thomas D. Smith, "Are Low Real Interest Rates Here to Stay?," *International Journal of Central Banking*, 2017, *13* (3), 1–42.

Reis, Ricardo, "Which R-Star, Public Bonds or Private Investment? Measurement and Policy Implications.," September 2022.

Shiller, Robert J., *Irrational Exuberance*, 2nd edition ed., New York, NY: Crown Business, May 2006.

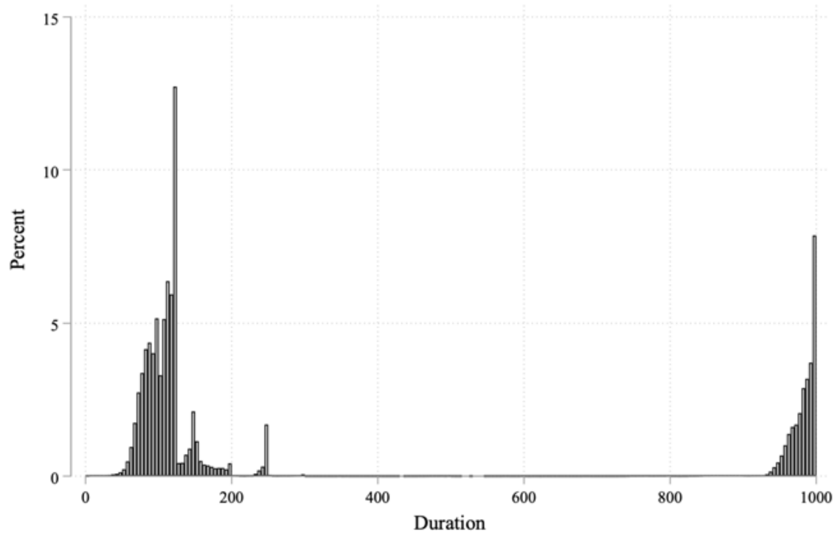
Vissing-Jorgensen, Annette, ""Has Monetary Policy Cared Too Much About a Poor Measure of r* ?" Discussion," May 2022.

Williams, John C, "Measuring the Natural Rate of Interest: Past, Present, and Future," May 2023.

A Appendix

A.1 Additional Figures

Figure A.1: Distribution of Lease Term for Leasehold Flats



The figure is a histogram of the remaining lease term at the time of transaction. The sample is all leasehold flats that transact at least once in the Land Registry Transaction Data Set.

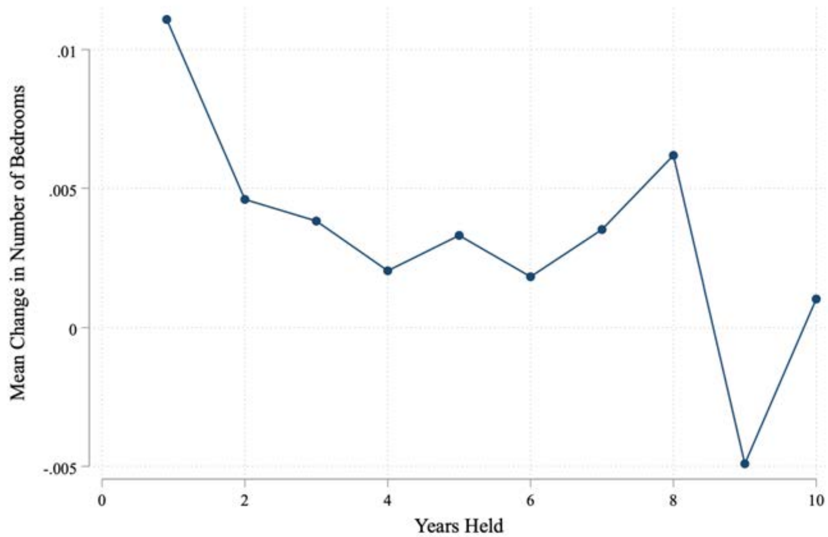
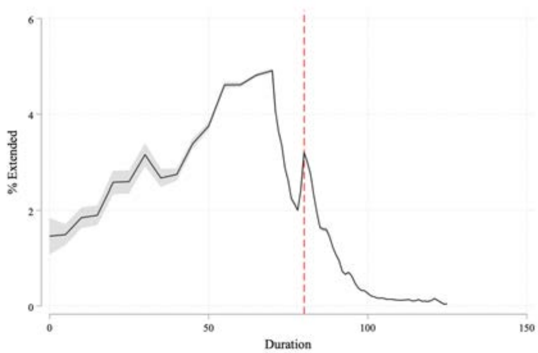


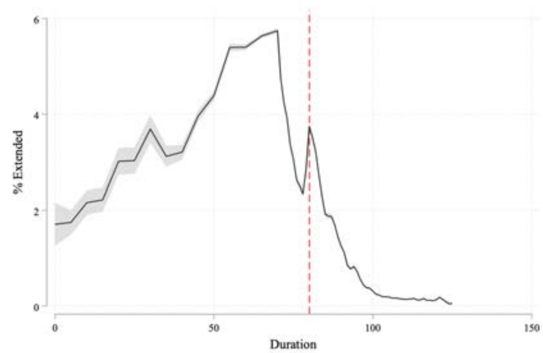
Figure A.2: Renovations By Holding Period

The figure shows the change in number of bedrooms reported relative to property holding period, where very short holds have a disproportionately high renovation rate. We take this as evidence that many of these are "flippers" who buy properties to re-sell them. The sample is all flats for which we observe two different Rightmove or Zoopla listings associated with two different property transactions.

Figure A.3: Hazard Rate of Lease Extension



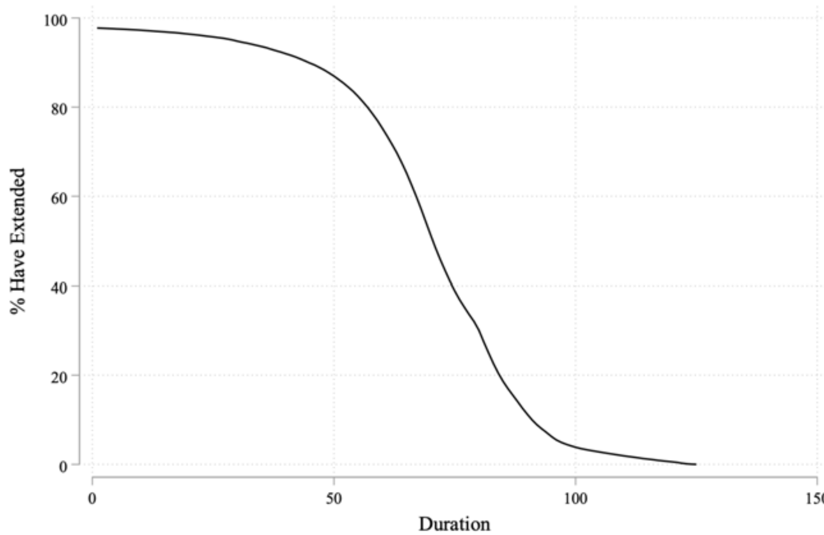
(a) Raw



(b) Corrected

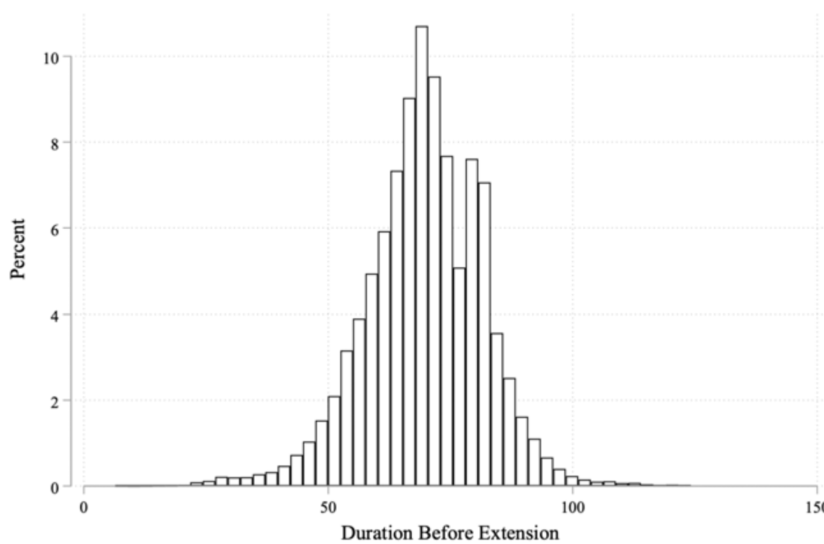
The figure shows the conditional probability of extension, $\theta(T)$ given that a property has duration T. In the first panel, the conditional probability of extension is given by $\theta_1(T) = \frac{N_T^{Ext}}{N_T}$ where N_T^{Ext} is the number of properties which extended with duration T and N_T is the number of properties that reached duration T. In the second panel, the conditional probability of extension is $\theta_2(T) = \gamma \frac{N_T^{Ext}}{N_T}$ where $\gamma = 1.17$ adjusts for the fact that our primary method does not identify properties which never transact before extension. The shaded area shows the 95% confidence interval.

Figure A.4: Cumulative Hazard Rate



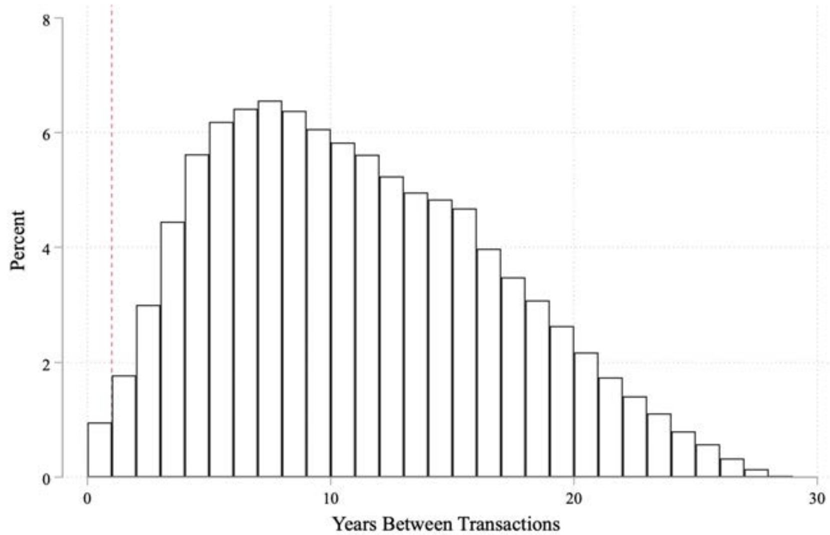
The figure shows the cumulative probability of extending over a property's lifetime. The sample includes all leases with at least one transaction and covers the 2003-2020 period. We exclude the pandemic period due to abnormally low extension rates.

Figure A.5: Histogram of Duration Before Extension



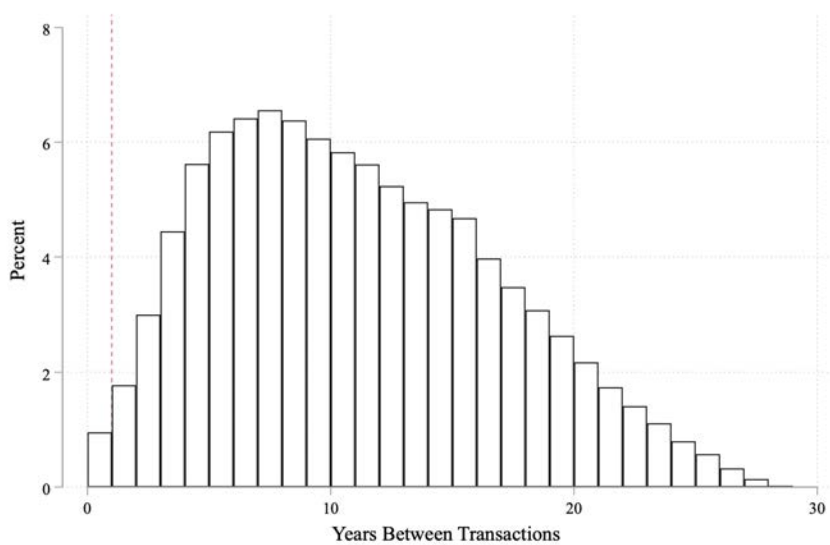
The figure presents a histogram of lease duration immediately before extension. The sample is all extended flats.

Figure A.6: Histogram of Years Between Transaction



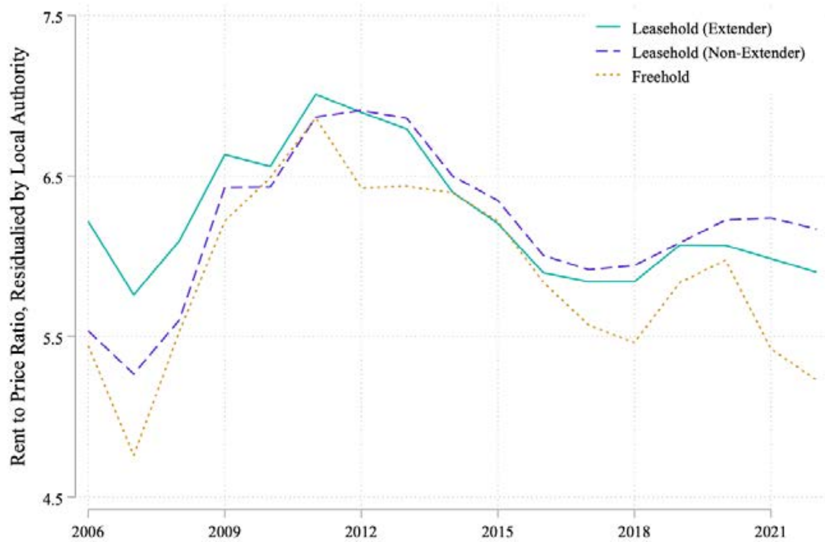
The figure shows a histogram of the holding period, h, for lease extensions which have a recorded transaction before and after extension. The dotted line shows the $h = 1$ cutoff; properties below the cutoff are not included in our primary sample. The sample is all extended flats.

Figure A.7: Histogram of Years Between Transaction and Extension



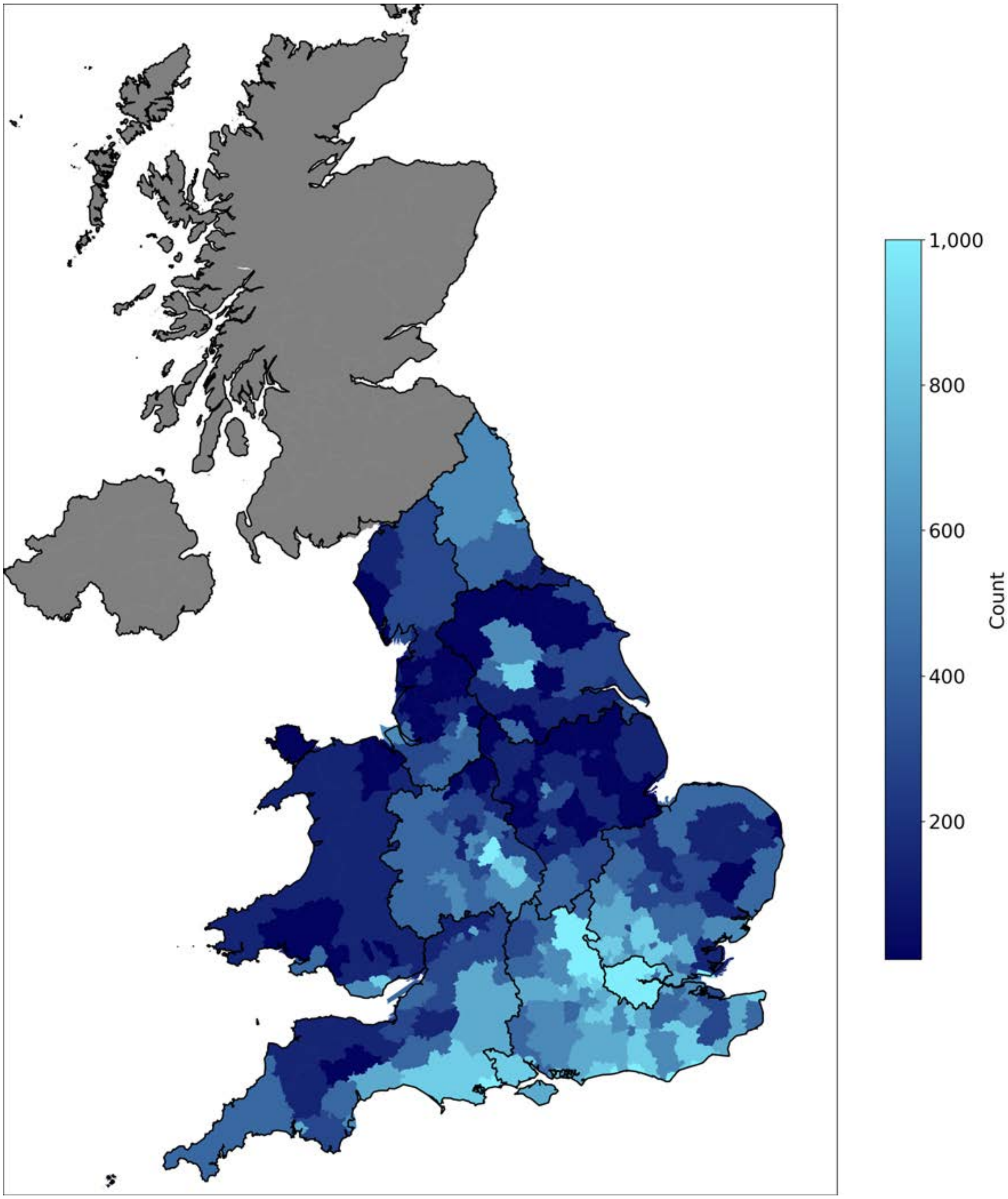
The figure presents a histogram of the number of years between purchase and extension time. The sample is all extended flats.

Figure A.8: Rent to Price (Freeholds vs Leaseholds)



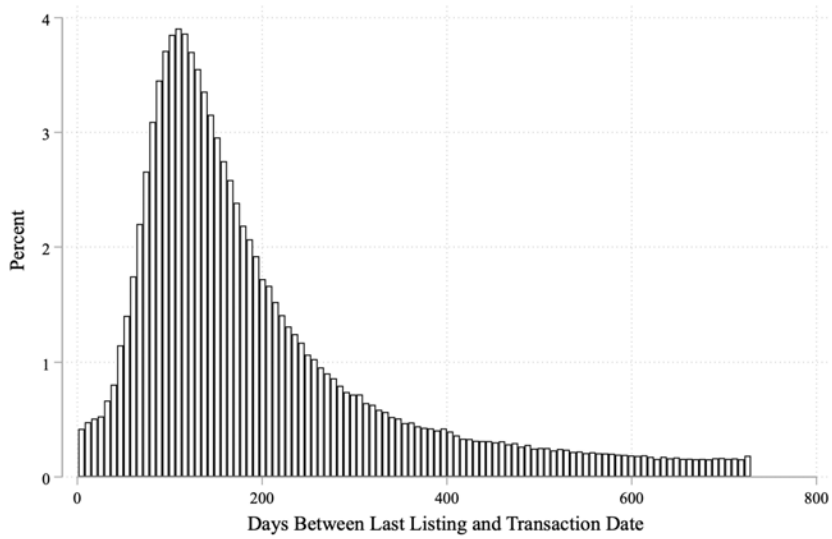
The figure shows the property-level price to rent ratio for leasehold and freehold flats. Leaseholds are subdivided into those which extend during our sample and those which do not. Property-level rental price data is collected from Rightmove and Zoopla.

Figure A.9: Heat Map of Extension Rate



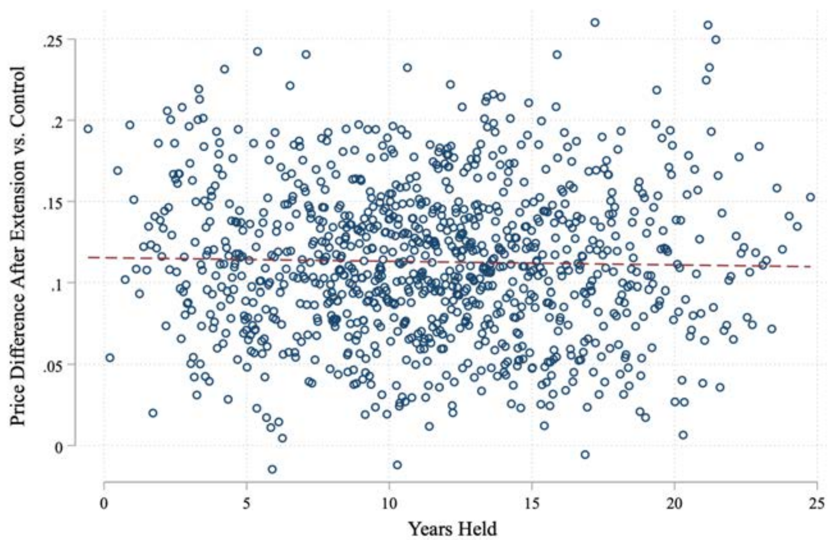
The figure shows a heatmap of the number of properties extended in each Local Authority in England and Wales.

Figure A.10: Time Elapsed Between Rightmove Listing and Land Registry Transaction Date



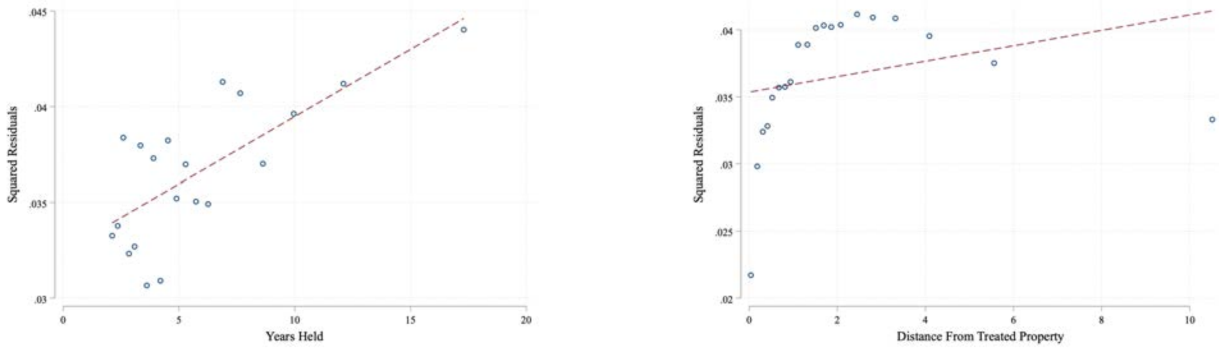
The figure shows a histogram of the amount of time (in days) between the last property listing on Rightmove and the date in which the transaction is recorded by the Land Registry. Each bin refers to one week. The sample is transactions for properties which have Rightmove listings within two years of the Land Registry transaction date.

Figure A.11: Relation Between Holding Period and Difference-in-Difference



The figure shows a binscatter of the difference-in-difference estimator, Δ_{it}, by holding period, controlling for sale year t fixed effects. The sample is all 90 year extensions.

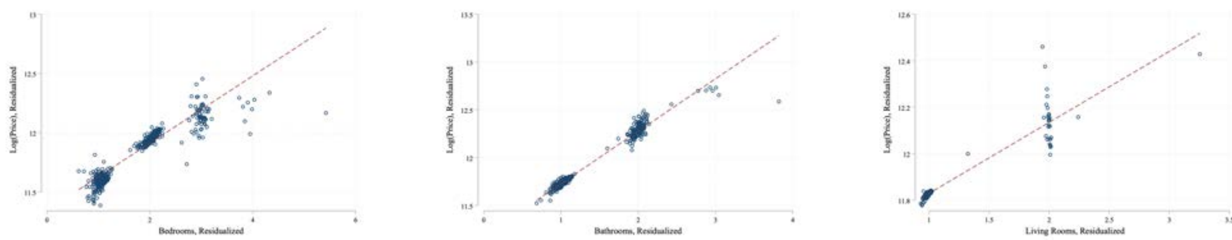
Figure A.12: Binscatter of Repeat-Sales Residuals



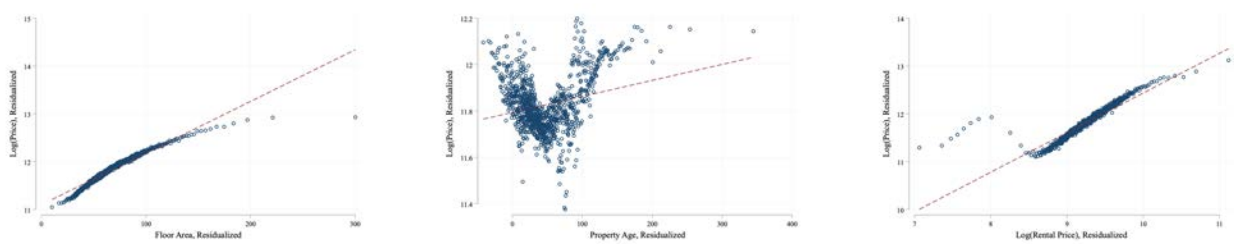
(a) Holding Period **(b)** Haversine Distance

The figures are binscatters of the squared residuals from the first step of the repeat-sales methodology against two variables which may be correlated with the variance of the estimate: (1) the time between transactions and (2) the distance between the treated and control property.

Figure A.13: Binscatter Log(Price) on Hedonics



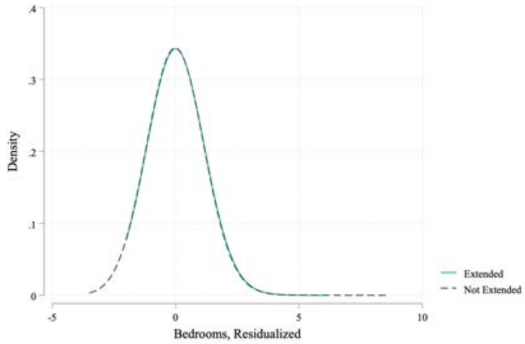
(a) Bedrooms **(b)** Bathrooms **(c)** Living Rooms



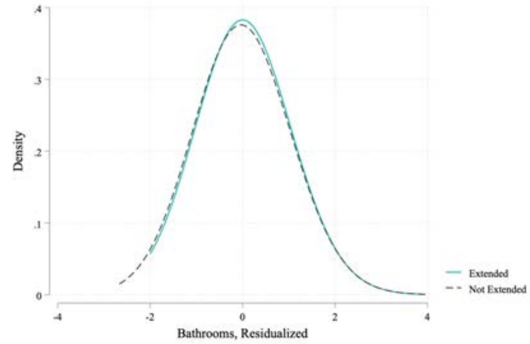
(d) Floor Area **(e)** Year Built **(f)** Log(Rent)

The figures are binscatters of log transaction price against the following hedonic characteristics: number of bedrooms, number of bathrooms, number of living rooms, floor area (sq. meters), year that the property was built, and log yearly rental price. Both the x and y-axis variables are residualized by Local Authority fixed effects, Γ_{it}. In particular, the y-axis variable is $\overline{\log(P_{it})} + \epsilon_{it}$ where $\overline{\log(P_{it})}$ is the mean log transaction price and ϵ_{it} is the residual from the following specification: $\log(P_{it}) = \Gamma_{it} + \epsilon_{it}$. The x-axis variable for each hedonic characteristic X_{it} is $\overline{X_{it}} + \eta_{it}$ where $\overline{X_{it}}$ is the mean level of X_{it} and η_{it} is the residual from the following specification: $X_{it} = \Gamma_{it} + \eta_{it}$. The sample is leasehold flats which appear at least once in the Land Registry Transaction Data Set.

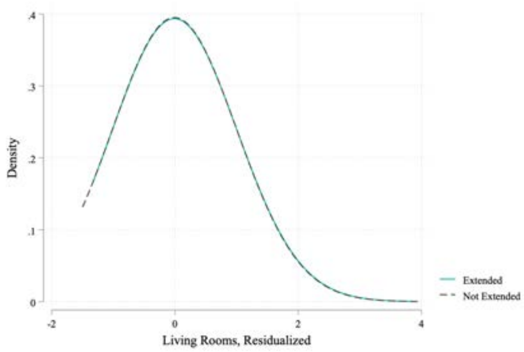
Figure A.14: Hedonics Residuals Density Curve for Extended and Non-Extended Properties



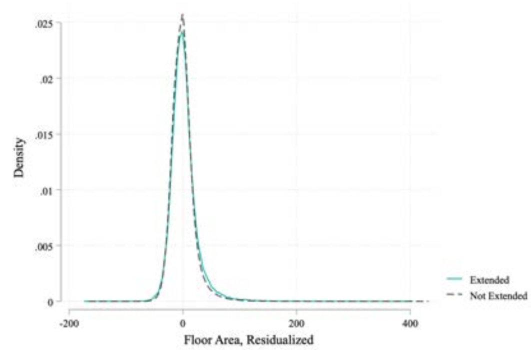
(a) Bedrooms



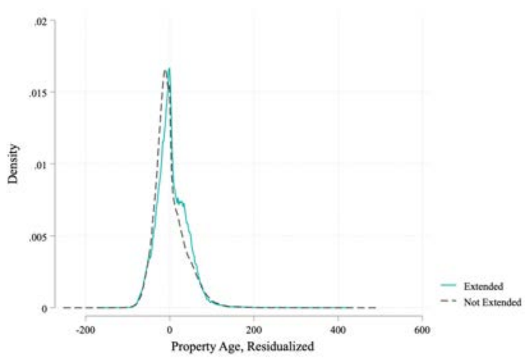
(b) Bathrooms



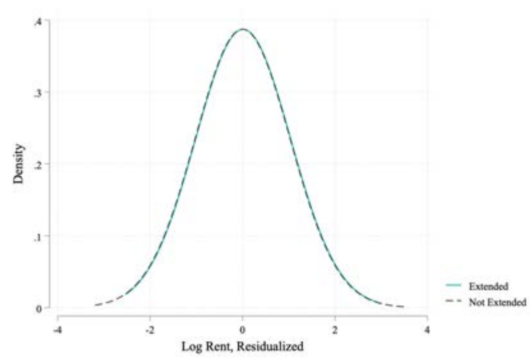
(c) Living Rooms



(d) Floor Area



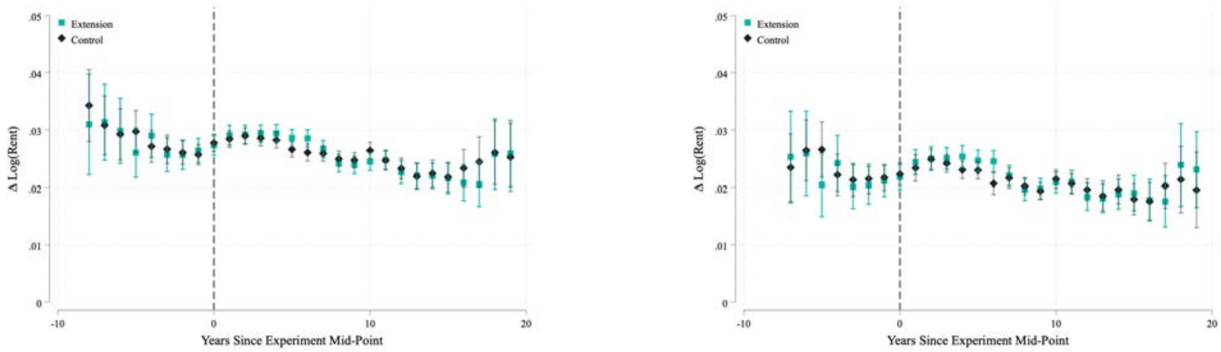
(e) Age



(f) Log(Rent)

The figures show the distribution of residuals for extended and non-extended flats after regressing hedonic characteristics on 5-year duration bin $\times$ Local Authority $\times$ year fixed effects. The sample is leasehold flats which appear at least once in the Land Registry Transaction Data Set.

Figure A.15: Rent Growth At Long Horizons Robustness

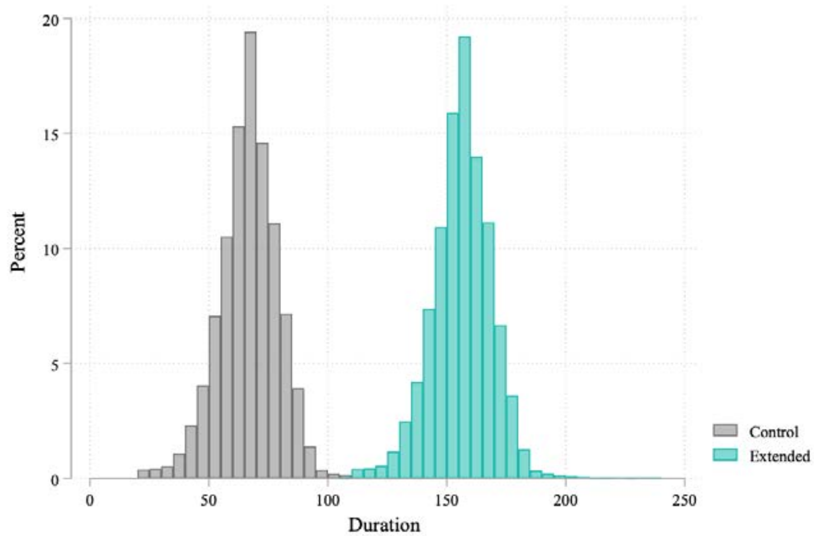


(a) Excluding Zero Rent Growth

(b) Excluding Distant Listings

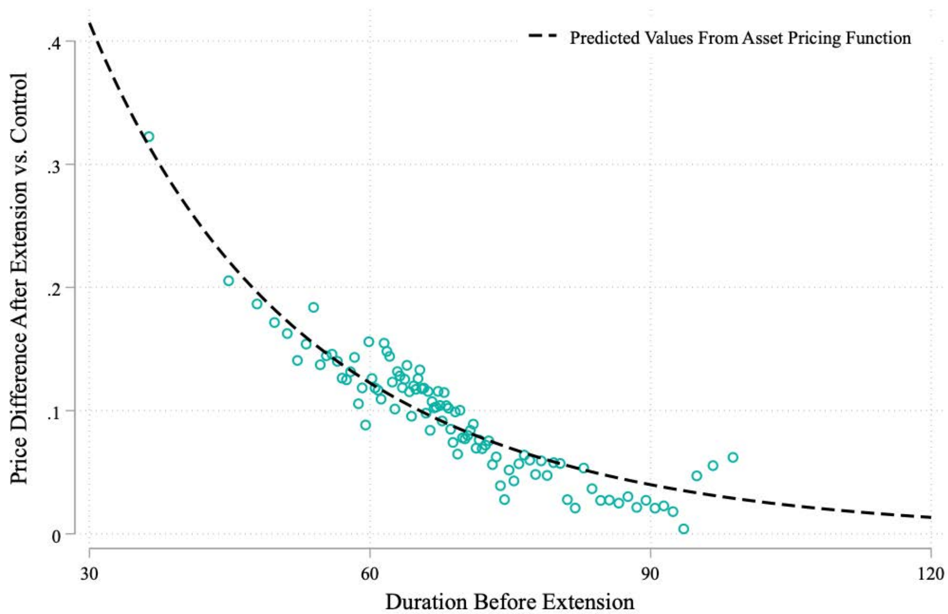
The figure presents robustness for Figure 5. Panel (a) excludes rental listing pairs for which there is zero rent growth. Panel (b) excludes rental listing pairs that are more than 5 years apart.

Figure A.16: Histogram of Remaining Lease Term At Sale



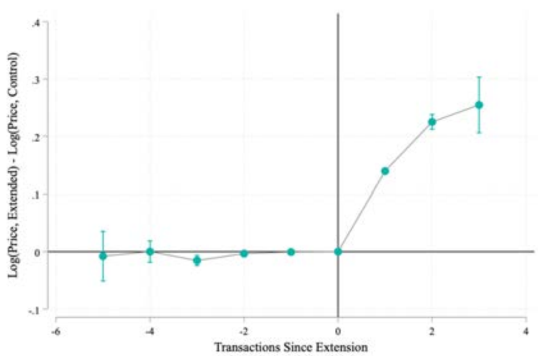
The figure shows a histogram of remaining duration at sale time t for the approximately 40 thousand experiments which were extended for 90 years. The blue-shaded histogram includes the sale duration for the controls in the experiment and the grey-shaded histogram includes the sale duration for the extensions.

Figure A.17: Duration vs. Price Gain After Extension, Leases with 700+ Years

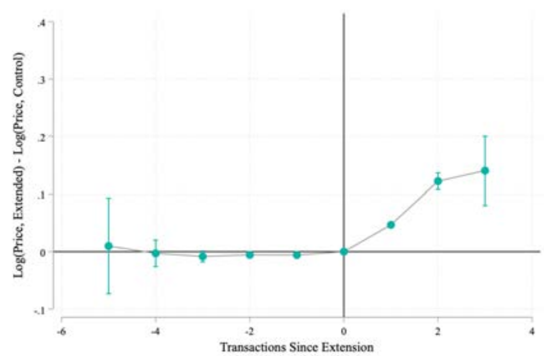


The figure is a binscatter of our difference-in-difference estimator against duration before extension, T, with 100 bins. The sample includes leases that were extended for more than 700 years. The black line shows fitted estimates of Equation (5).

Figure A.20: Event Study By Duration



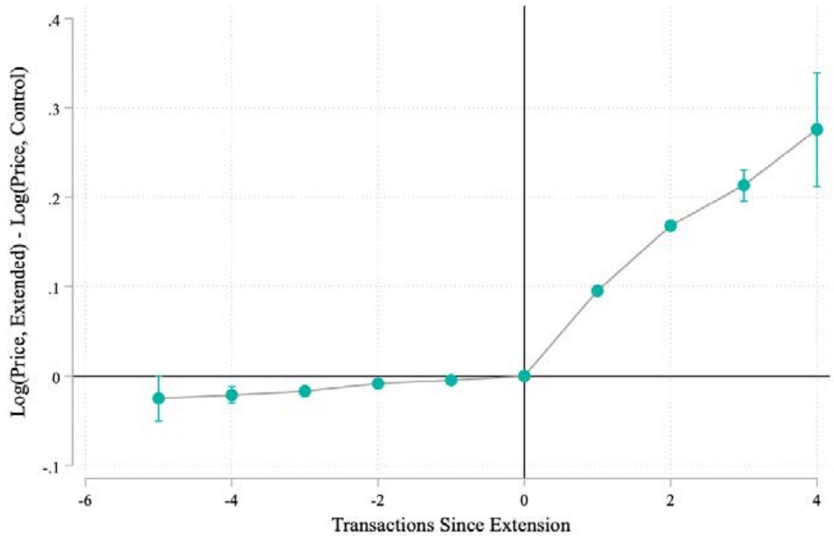
(a) $T < 70$



(b) $T \geq 70$

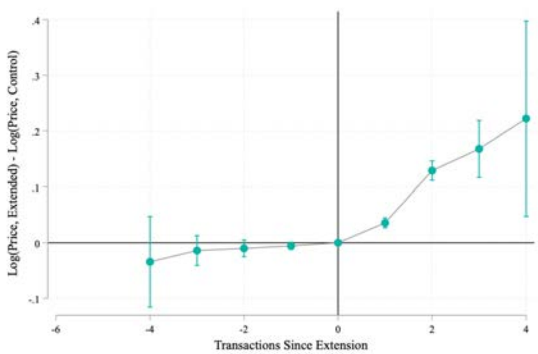
The figure replicates Figure 6 for properties that were extended with below and above median durations.

Figure A.18: Event Study Representation of Lease Extension, All Extensions

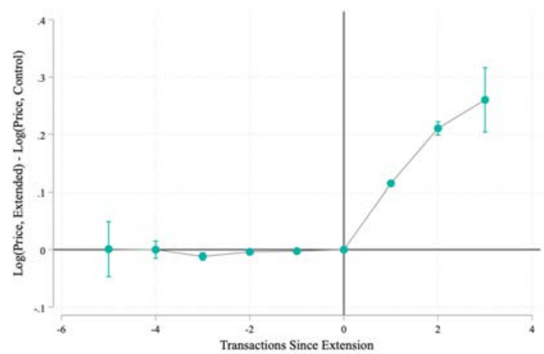


The figure replicates Figure 6 for properties of all extension amounts.

Figure A.21: Event Study By Time Period



(a) Pre 2012



(b) Post 2012

The figure replicates Figure 6 for two different sub-periods: pre 2012 (inclusive) and post 2012.

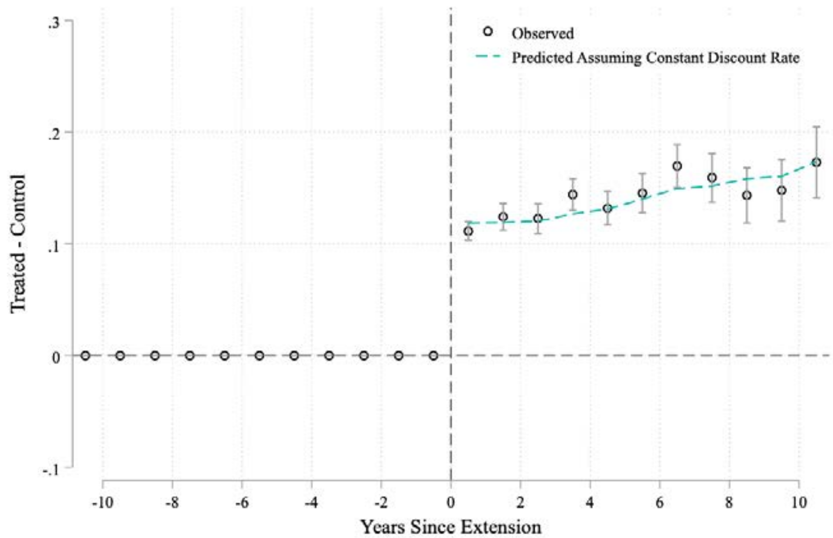
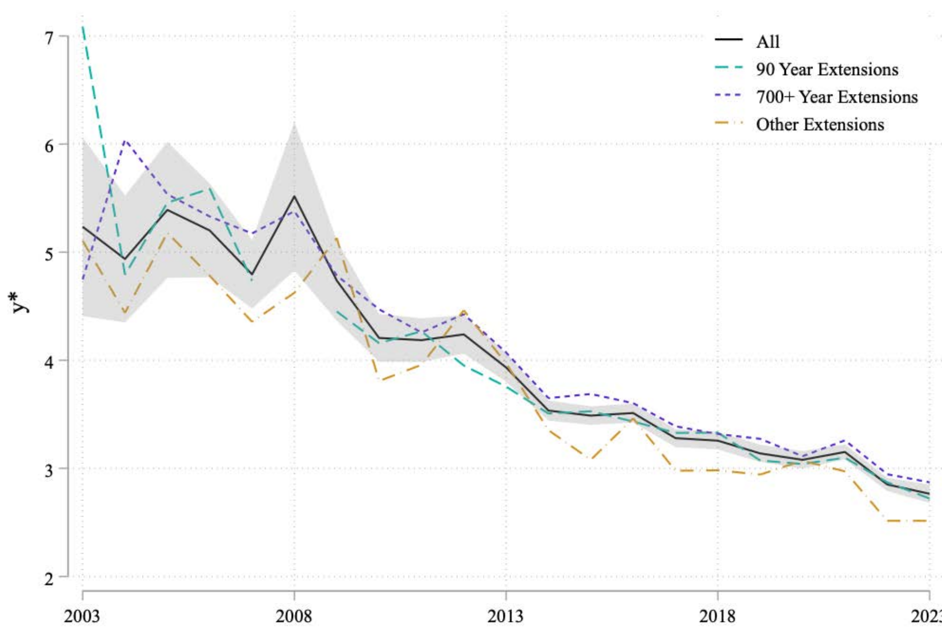


Figure A.19: Trend in Price Gain Following Extension

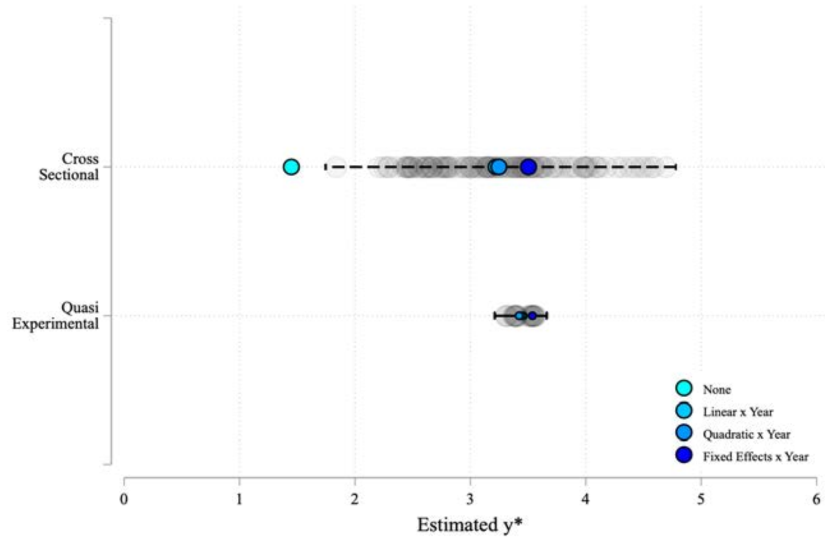
The figure unpacks the change between the point at $t = 0$ and $t = 1$ in Figure 6. We normalize the pre-extension price difference to zero and plot the mean difference in price change between time $t - h$ and t against time since extension on the x-axis. We restrict the sample to properties sold after 2018, a period during which y^* is relatively stable, so that we can isolate the effect of duration on post-extension trends. The dotted line plots the predicted price gain based on mean duration at each time.

Figure A.22: y_t^* Estimates For Various Extension Amounts



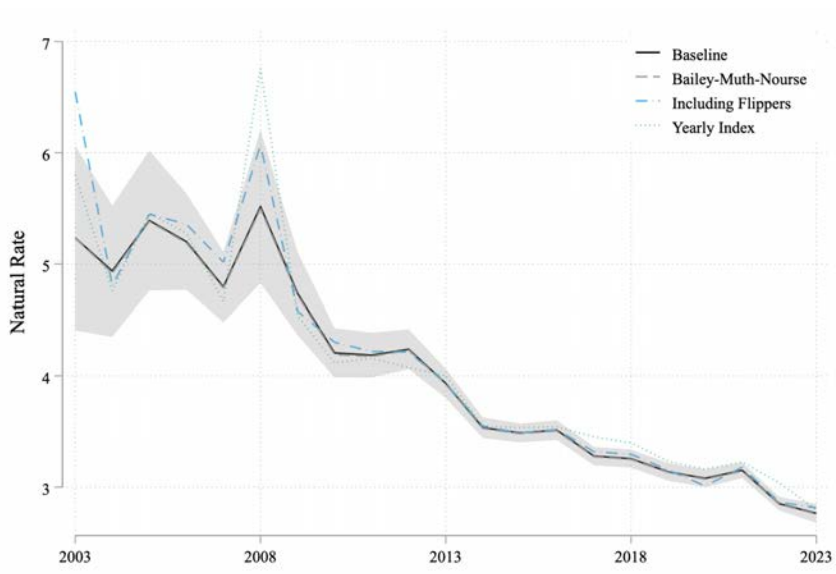
The figure presents estimates of y_t^* for every year of our sample period for lease extensions that were extended for 90 years, for more than 700 years, or for another amount, separately. The estimate of y^* for 90-year extensions in 2008 is excluded, because the mean estimated difference-in-difference is negative due to low sample size and volatility from the housing crisis, which leads to an extreme y^* estimate.

Figure A.23: Stability of y^*, Controlling for Observables



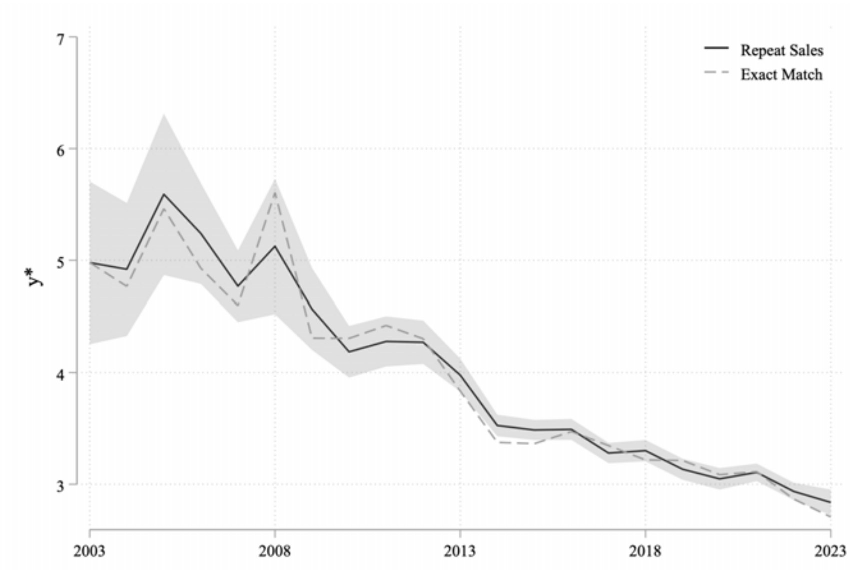
The figure replications Figure 10 controlling for hedonic interacted with time fixed effects.

Figure A.24: Robustness Checks



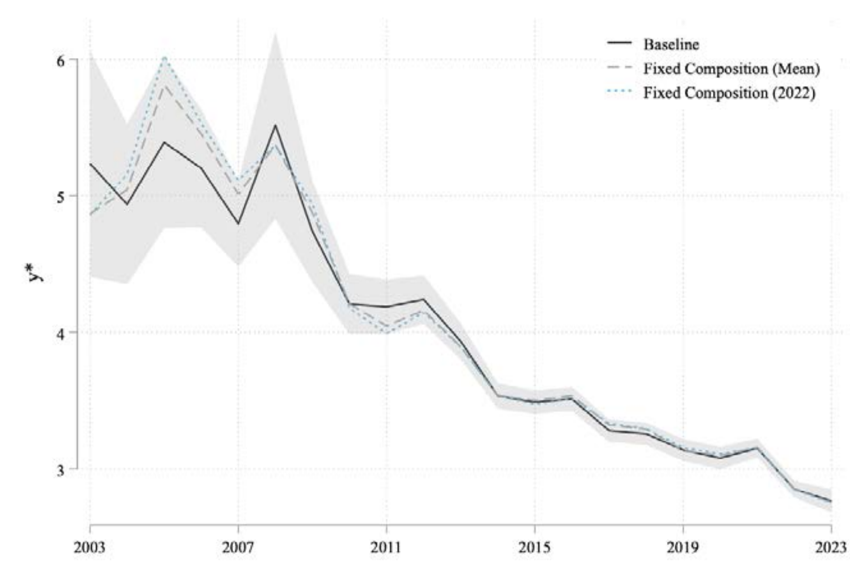
The figure presents robustness for our estimates of y_t^*. The solid black line presents our baseline estimates from Figure 1. The dashed grey line presents estimates when using the methodology from Bailey et al. (1963) to produce the repeat sales index of control properties. The dash-dot blue line presents estimates if we include "flipper" properties that were held for 2 years or less. The dotted green line presents estimates when constructing the repeat sales index at an annual rather than quarterly frequency.

Figure A.25: Restrictive Controls



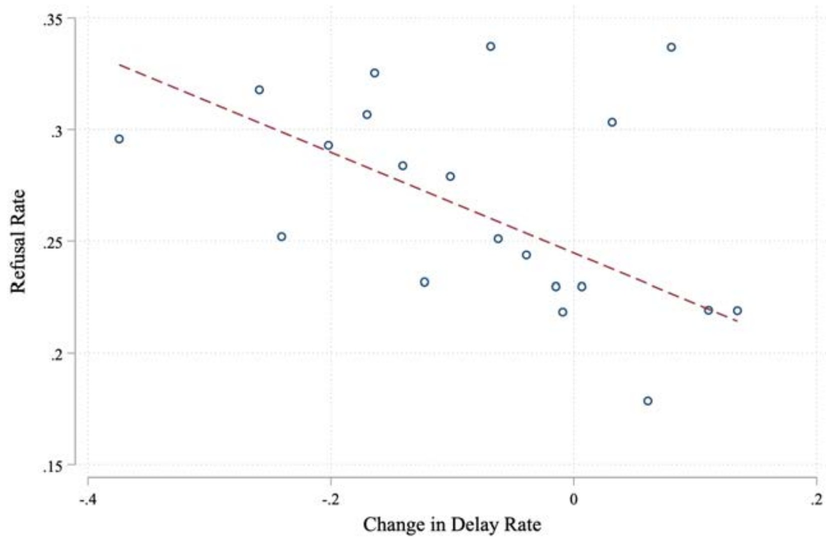
The figure presents estimates of y^* when we pick controls that transact in the same purchase and sale years as the treated property. Our difference-in-difference then becomes $\Delta_{it} = (\log P_{it}^{T+90} - \log P_{i,t-h}^{T+h}) - (\log P_{jt}^{T} - \log P_{j,t-h}^{T+h})$, where i is a treated property and j is a neighboring control property of similar duration. The black line presents our baseline estimates on the sample for which we are able to find exact controls. The dashed grey line presents estimates using exact controls. A 95% confidence interval is shaded. The sample is all lease extensions.

Figure A.26: Fixed Regional Composition



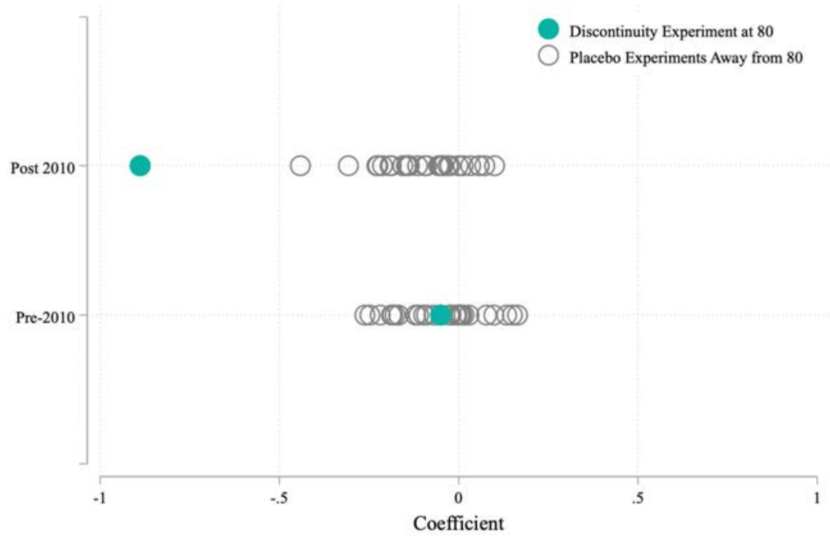
The figure presents estimates of y^* controlling for the geographic composition of the data. The solid black line presents our baseline estimates. In the dashed grey line, we fix the composition to the mean composition of Local Authorities over the sample period. In particular, we produce weights, $w_{it} = \bar{N}_i / N_{it}$, where $\bar{N}_i$ is the average number of observations in Local Authority i over the entire period and N_{it} is the number of observations in i in year t. We then run a weighted NLLS regression using the weights w_{it}. The dotted blue line presents estimates where we fix the composition of the data to its 2022 level using the same method. The 95% confidence interval of the baseline estimates is shaded. The sample is all lease extensions.

Figure A.27: Refusal Rate Instrument, First Stage



The figure presents the first stage binscatter for the IV regression in Table 5. The x-axis variable is the change in application delay rate after the 2002 policy discussed in Section 6.3. The y-axis variable is the mean refusal rate of major construction applications. The slope coefficient is -0.22, with standard error of 0.05. The binscatter has 20 bins.

Figure A.28: Test for Discontinuity at 80 Years (Placebos)



The figure presents the estimates from Table A.6 in red. In grey, we run 30 placebo tests using alternative cutoffs between 70-100. In each case, the sample includes properties within 10 years above and below each cutoff.

A.2 Additional Tables

Table A.1: Freehold vs Leasehold Statistics (English Housing Survey)

	Freehold	Leasehold
Income	29,628.73	25,653.20
	(52.95)	(138.48)
Age	53.95	51.49
	(0.03)	(0.10)
% Have Mortgage	54.82	59.07
	(0.10)	(0.28)
LTV	72.17	76.16
	(0.14)	(0.39)
N	305,135	

This table reports the mean characteristic for freehold owners, in column 1, and leasehold owners, in column 2. The standard error of the mean is in parentheses.

Table A.2: Hedonic Characteristics in Extended vs Non-Extended Flats, Only 90 Year Extensions

Panel A: Levels

	(1) Num Bedrooms	(2) Num Bathrooms	(3) Num Living Rooms	(4) Floor Area	(5) Age	(6) Log Rental Price
Extension	-0.029***	0.010***	-0.005***	-0.559***	2.483***	-0.011***
	(0.004)	(0.002)	(0.001)	(0.139)	(0.256)	(0.002)
Fixed Effects	✓	✓	✓	✓	✓	✓
Control Mean	1.73	1.18	1.03	65.02	45.93	9.23
N	2,415,295	1,945,080	1,719,663	2,006,388	1,449,578	1,401,817

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel B: Differences

	(1) Δ Num Bedrooms	(2) Δ Num Bathrooms	(3) Δ Num Living Rooms	(4) Δ Floor Area	(5) Δ Log(Rent)
Extension	-0.001	-0.002	-0.004*	-0.056	0.003
	(0.006)	(0.002)	(0.002)	(0.095)	(0.014)
Fixed Effects	✓	✓	✓	✓	✓
Control Mean	.001	0	0	-.001	.012
N	1,149,191	935,429	828,473	959,097	642,238

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel A reports level of hedonic characteristics for extended properties relative to their non-extended counterparts. It presents estimates of β for the specification, $X_{it} = \alpha + \beta \mathbb{1}(\text{Extended}_{it}) + \Gamma_{it} + \epsilon_{it}$, for several hedonic characteristics X_{it} where Γ_{it} are 10-year duration bin × Local Authority × year fixed effects. The sample is leasehold flats which appear at least once in the Land Registry Transaction Data Set. Panel B reports renovation rates in extended properties relative to their non-extended counterparts. It presents estimates from regression equation (6). For Columns 1-4, the sample includes properties with two separate Rightmove or Zoopla listings corresponding to two separate transactions. For Column 5, the sample includes properties that transact twice and have distinct rental price data around each transaction. The sample is all 90 year extensions.

Table A.3: Estimated y^* for Extensions of 90 Years vs More Than 700 Years

	(1)	(2)	(3)	(4)	(5)	(6)
$y*$	3.51***	3.56***	3.46***	3.55***	3.47***	3.53***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Extension Amount	90	700+	90	700+	90	700+
Hedonics Sample			✓	✓	✓	✓
Control For Hedonics					✓	✓
N	40,500	50,719	25,190	26,825	25,190	26,825

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table presents estimates of y^* for properties that were extended for 90 and more than 700 years separately. The price change after extension, $\Delta_i t$, is first residualized on year x local authority fixed effects to remove the effect of differences in geographic or time distribution on leases of different extension amounts. Columns (3) and (4) repeat the analysis on properties for which we have bedroom and floor area data. Columns (5) and (6) repeat the analysis using a price measure which controls for bedroom and floor area.

A.3 Equivalence of y^* and User Cost of Capital

First consider a static model in which a firm has profits $\Pi(K, X)$, where X is the cost of other inputs and K is capital. The firm maximizes profits, i.e.

$$\max_{K,X} \Pi(K, X) - R_K \cdot K$$

where R_K is the cost of renting one unit of capital. Then, the first order condition for capital is,

$$\Pi_K(K, X) = R_K$$

so the marginal product of capital is equal to the user cost of capital.

Now, consider a dynamic problem in which capital depreciates at rate δ. Suppose that the firm purchases capital at price P_K. Then, the firm faces the following intertemporal maximization problem,

$$V(K_t) = \max_{I_t} \int_t^\infty e^{-(r_t^* + \zeta_t^*)s} (\Pi(K_s) - P_{K_s} I_s) ds$$

where I_t is investment at time t and $\dot{K}_t = I_t - \delta K_t$, such that the change in the capital stock is equal to investment minus depreciation. As in the main text, $r^* + \zeta^*$ is the required rate of return of the firm, which can be separated in a risk-free rate and a risk-premium. The

Hamiltonian for this problem is,

$$\mathcal{H} = \Pi(K_t) - P_{K_t} I_t + \lambda_t (I_t - \delta K_t)$$

where the optimality conditions are,

$$P_{K_t} = \lambda_t$$

$$\Pi_K(K_t) - \delta \lambda_t = (r_t^* + \zeta_t^*)\lambda_t - \dot{\lambda}_t$$

Substituting the optimality conditions and reorganizing yields,

$$\Pi_K(K_t) = P_{K_t}\left(r_t^* + \zeta_t^* + \delta - \frac{\dot{P}_{K_t}}{P_{K_t}}\right)$$

$$R_{K_t} = P_{K_t}\left(r_t^* + \zeta_t^* + \delta - \frac{\dot{P}_{K_t}}{P_{K_t}}\right)$$

where in the last equation, we utilize the fact that the marginal product of capital is equal to the user cost of capital. Normalizing both sides by the price of capital we see that,

$$\begin{aligned} \frac{R_{K_t}}{P_{K_t}} - \delta &= r_t^* + \zeta_t^* - \frac{\dot{P}_{K_t}}{P_{K_t}} \\ &= y_t^* \end{aligned}$$

so the user cost of capital normalized by the price of capital and net of depreciation is equal to the expected long term housing yield.

A.4 Details on Merge Between Lease and Transaction Data

This section briefly describes our procedure to merge Land Registry data on lease lengths, and house price transaction data. In the UK, every property can be uniquely identified by three items: the first address number, the second address number and the 6-digit postcode. Therefore, we merge according to the following procedure: First, we conduct a perfect merge using address as our merge key. This methods accounts for 93% of our matches. Second, we conduct a fuzzy merge on all observations not matched by step 1. The fuzzy merge matches observations in which (1) all numeric elements of the address are the same, (2) all single letters (e.g. A, B, C, etc.) are the same, (3) a select set of identifying terms (e.g. first floor, second floor, basement, etc.) are the same and (4) the postcode is the same in both addresses. For example, this allows for the property "3 Swan Court 59-61 TW13 6PE" in the transaction data set to be matched to "Flat 3 Swan Court 59-61 Swan Road Feltham

TW13 6PE” in the lease term data set. This method accounts for 7% of our matches.

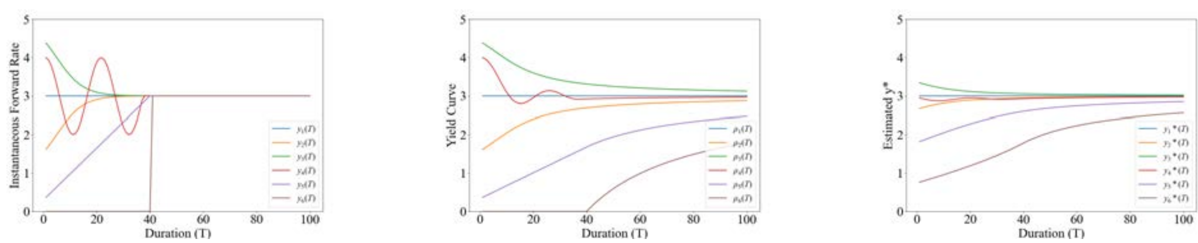
Additionally, we purchase the leasehold titles from the Land Registry for approximately 20 thousand transactions for which we are unable to identify a lease term based on the fuzzy merge. These include cases where the lease term address has typos, or has been accidentally omitted from the main public data set.

A.5 Simulation Results For Flexible Forward Curve

In Section 4.2 we presented one possible parameterization of $y(T)$. In this section, we explore other parameterizations and present several insights from our simulation. We assume that for $T \geq 40$, $y(T) = y^*$ is constant. Hence, for $T < 40$ we assume $y(\cdot)$ can have any shape as long as it asymptotes to y^* as $T \to 40$.

Figure A.29a shows several possible choices of $y(T)$, all of which asymptote to the same value. Figure A.29b presents the yield curves associated with each of these forward curves, which we denote by $\rho(T)$; and Figure A.29c shows the corresponding $\hat{y}^*(T)$ curves that we estimate by Equation (5) using simulation data. We also plot the point estimate of $\hat{y}^*$ we obtain at the median of our true distribution.

Figure A.29: Long Run Discount Rates Using Simulation Data



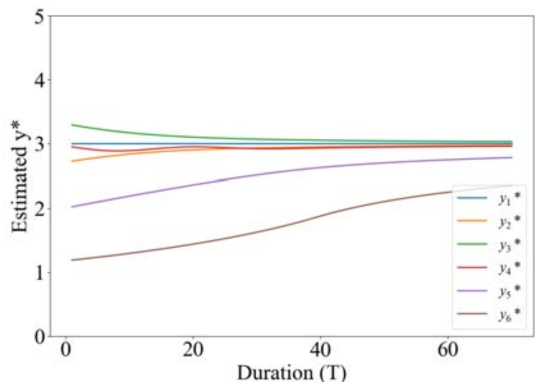
(a) $y(T)$ **(b)** $\rho(T)$ **(c)** $y^*(T)$

Panel (a) presents multiple possible parameterizations of $y(T)$, all of which asymptote to 3% by $T = 40$. Panel (b) presents the corresponding yield curve, $\rho(T)$, for each choice of $y(T)$. Panel (c) presents estimates of $y^*(T')$ obtained by NLLS corresponding to each choice of $y(T)$, where T' is the average between the control and extension sale duration. The point presents the estimate of y^* we would obtain at the median of our distribution.

These simulation results yield several key insights. First, when the yield curve is flat, $y(T) = \rho(T) = \hat{y}^*(T) = y^*$, as exemplified with the by the dark blue line in Figure A.29. When the yield curve is not flat, however, our estimate will differ from the true asymptotic value of $y(T)$ by some amount $\eta \equiv y^* - \hat{y}^*$. When the yield curve is upwards sloping, $\eta > 0$ and when the yield curve is downwards sloping, $\eta < 0$.

Notice that for $T \geq 40$, $\hat{y}^*(T)$ converges to y^* much more quickly than $\rho(T)$. The reason for this is that our difference-in-difference estimate differences out a large portion of the short-end of the yield curve. To see this, consider a property with duration T that extends by k years to a total of 160 years ($T + k = 160$). The shorter T is, the less of the short-end that will be differenced out by our estimate. We present simulation evidence for this in Figure A.30. For this reason, it is important that our experiments take the difference between two long-duration properties, as opposed to one long duration property and one short duration property.

Figure A.30: Estimated y^* When Extending From T to 160



The figure indicates the point estimate of $\hat{y}^*$ we obtain by NLLS for an extension from duration T to duration 160. As the duration before extension increases, the estimate of $\hat{y}^*$ approaches the limit of the forward curve. We repeat this for each example forward curve, $y(T)$, from Figure A.29a.

A.6 Repeat Sales Index Methodology

This subsection explains the details of how we estimate the repeat sales index for the control group associated with a given extending property p.

First, we estimate a Bailey, Muth and Nourse (1963) repeat sales index for the control group of each treated property p. We fit a regression

$$\log P_{qt'} - \log P_{qt} = \sum_{s=t_0+1}^{t_1} \gamma_s D_s + \varepsilon_{q,t,t'} \tag{8}$$

for each property q in p's control group, where D_s is a dummy variable that is equal to 1 if $s = t'$, -1 if $s = t$, and 0 otherwise. t_0 is the first date for which a control property is available and t_1 is the last date. The coefficients γ_s capture the estimated change in log price since time t_0.

As in Case and Shiller (1989), we correct for the effects of heteroskedasticity by allowing the error term to be related to the interval of time between sales, $t' - t$, and also to the

(Haversine) distance, d_q, between the treated property p and the control q. We estimate,

$$\hat{\varepsilon}^2_{q,t,t'} = \beta_0 + \beta_1(t' - t) + \beta_2 d_q + \mu_{q,t,t'} \tag{9}$$

where $\hat{\varepsilon}_{q,t,t'}$ are the predicted residuals from the regression in Equation (8). Figure A.12 presents binscatters of $\hat{\varepsilon}_{q,t,t'}$ against $t' - t$ and d_q, showing that both are positively correlated. We estimate Equation (9) for all properties together to reduce noise. Then, we calculate weights, $w_{q,t,t'} = 1/(\hat{\beta}_0 + \hat{\beta}_1(t' - t) + \hat{\beta}_2 d_q)$, equal to the inverse predicted values from Equation (9). The last step is to rerun Equation (8) using these weights.

A.7 Liquidity Premium

One institutional factor which could raise concerns about our estimator is the difficulty for owners of short leases to obtain a mortgage. If short leases have more limited access to financing than longer leases, we might worry that part or all of the observed price gain upon extension is a result of increased access to financing opportunities; in other words, we may wonder if the effect is driven by a "liquidity premium." Indeed, important lenders, such as Barclays, Halifax and The Co-Operative Bank, refuse to lend to leaseholds with less than 70 years remaining. Others have different thresholds, such as 55 years, and some have a preference for longer leases but allow for case-by-case exceptions.[40]

Reassuringly, using detailed micro-data from the English Housing Survey 1993-2019, we find that mortgage access and conditions are not vastly different for shorter and longer leaseholds, especially for those with more than 30 years remaining, as indicated in Table A.4. Approximately 60% of short (under 80) duration leaseholds were purchased with a mortgage, relative to 58% of longer (over 80) duration leaseholds. The typical mortgage length and Loan-To-Value (LTV) ratios are similar across the duration spectrum, at around 23 years and 75-80%, respectively. Additionally, short leaseholds have similar interest rate types as long leaseholds, with around 30% choosing adjustable rate mortgages (as opposed to mortgages with a fixed interest rate for a number of years, or tracker mortgages which are indexed to the Bank of England bank rate). These results suggest that financing constraints are unlikely to drive the very large extension price changes we observed in Section 5.

Another common way to test for the existence of a liquidity premium resulting from financing frictions is to use the amount of time a property was listed on the market (i.e. sale time minus the time of the first listing) as a proxy for its liquidity (Lippman and McCall, 1986; Lin and Vandell, 2007; Genesove and Han, 2012). The intuition is that properties for

[40] A comprehensive list of lease length policies for banks in England can be found in the UK Finance Lenders' Handbook For Conveyancers.

Table A.4: Mortgage Statistics For Short Leaseholds

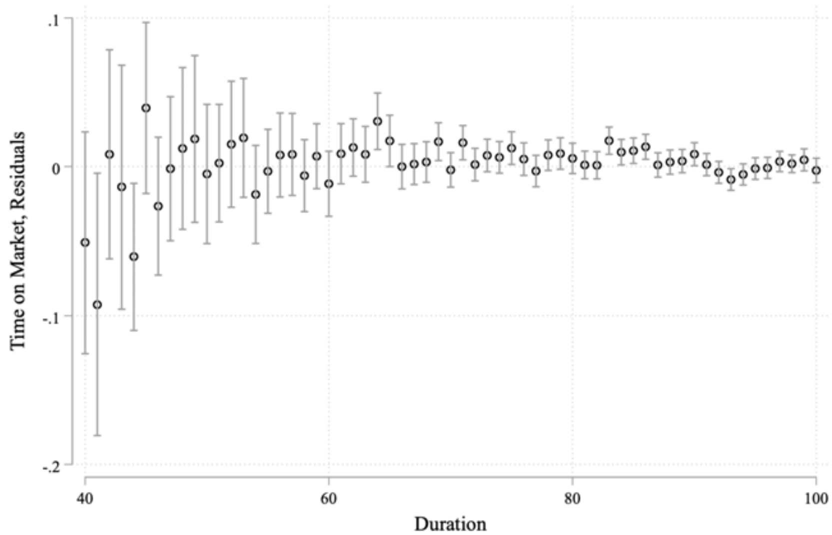
	Less Than 50 Years	50-60 Years	60-70 Years	70-80 Years	80-99 Years	100+ Years	Total
Mortgage Length	22.1	22.1	23.9	23.0	23.9	23.1	23.3
	(0.6)	(0.5)	(0.5)	(0.3)	(0.2)	(0.1)	(0.1)
LTV	76.3	80.8	81.4	77.7	73.3	76.5	76.2
	(3.3)	(2.6)	(1.8)	(1.7)	(1.0)	(0.6)	(0.5)
% Have Mortgage	59.9	60.4	62.1	58.1	63.9	55.6	58.2
	(2.4)	(2.4)	(1.6)	(1.4)	(0.8)	(0.5)	(0.4)
% Adjustable Rate	24.2	40.0	38.3	32.8	25.3	31.0	30.2
	(5.3)	(5.2)	(4.1)	(3.4)	(1.5)	(1.0)	(0.8)
N	18,292						

mean reported; standard error of mean in parentheses

The table reports several summary statistics for leases of various durations. The first row presents the average mortgage length; the second presents the average Loan-To-Value (LTV) ratio for the mortgage, calculated as the initial mortgage value divided by the market price of the property; the third row presents the percent of properties of that duration which have a mortgage; the fourth row presents the percent of properties with a mortgage that have a fully adjustable rate mortgage.

which the buyer cannot obtain a mortgage have a smaller pool of potential buyers, which ought to increase the amount of time that the property is on the market. We find limited evidence of this in the data, as illustrated in Figure A.31. After controlling for quarter $\times$ 3-digit postcode fixed effects and hedonics, the typical listing time for a 50-70 year lease is only 3-4 days longer than for a long lease of more than 100 years. This is negligible given that the average listing time is 5 months. For shorter leases, the listing period actually decreases further; all else equal, a leasehold with less than 50 years remaining will sell about a week faster than a leasehold with more than 50 years.

Figure A.31: Time on Market by Lease Duration



The figure shows the mean time on market for every duration under 125, de-meaned by quarter × Local Authority fixed effects and controls for bedroom count, floor area and property age.

To further test for the existence of a liquidity premium, we are able to model the case of discontinuous financing frictions and reject the existence of a liquidity premium using the data. Consider that under the threshold of 70 years–which is the most prominent bank mortgage cutoff duration—it becomes significantly more difficult to finance a leasehold property. Then the price of a $T > 70$ duration leasehold is given by,

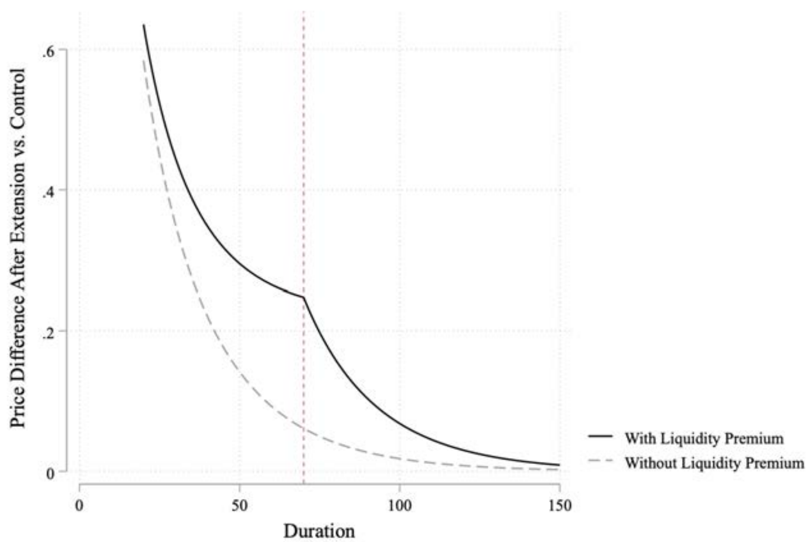
$$\begin{aligned}
P_t^T &= R_t \left[\int_t^{t+(T-70)} e^{-(y^*)(s-t)} ds + e^{-(y^*)(T-70)} \int_{t+(T-70)}^{t+T} e^{-(y^*+\sigma)(s-(t+(T-70)))} ds \right] \\
&= R_t \left[\frac{1 - e^{-(y^*)(T-70)}}{y^*} + \frac{(1 - e^{-(y^*+\sigma)\cdot 70})}{y^* + \sigma} (e^{-(y^*)(T-70)}) \right]
\end{aligned}$$

where rents below the cutoff of 70 are discounted at an additional rate σ. One way to think about σ is as the difference between the return on housing and outside investment options. Therefore, our difference in difference will yield the following equation:

$$\begin{aligned}
\log P_t^{T+90} - \log P_t^T = \log &\left[\frac{1 - e^{-(y^*)(T+90-70)}}{y^*} + \frac{1 - e^{-(y^*+\sigma)70}}{y^* + \sigma} (e^{-(y^*)(T+90-70)}) \right] \\
- \log &\left[\frac{1 - e^{-(y^*)(\max\{0,T-70\})}}{y^*} + \frac{1 - e^{-(y^*+\sigma)(\min\{70,T\})}}{y^* + \sigma} (e^{-(y^*)(\max\{0,T-70\})}) \right]
\end{aligned} \tag{10}$$

When $\sigma > 0$, the price change for extending a lease will have a kink at 70 as shown in Figure A.32. The reason for this is that as $T \to 70$, there are two incentives to extend: first is the value of 90 additional years of discounted rents and the second is the value of postponing the liquidity discount, σ. Once $T < 70$, the first incentive continues to grow, as we have discussed in earlier sections, but the second incentive becomes increasingly weaker, since there are less periods at which rents will be discounted at rate σ. Moreover, as $T \to 0$, $T + 90$ grows closer to 70, so the value of the extended lease also starts to decrease as it approaches the liquidity premium cutoff. This kink is not present in the data, which can be verified visually in Figure 7, and the existence of a discontinuous liquidity premium at $T = 70$ is rejected by estimating Equation (10) by NLLS.

Figure A.32: Liquidity Premium Example



The figure shows the effect of a liquidity premium on the value of extending for 90 years ($P_t^{T+90} - P_t^T$). The dark line plots Equation (10) when $\sigma = 1\%$ and the dashed light line plots the same equation absent a liquidity premium, i.e. $\sigma = 0\%$. The liquidity premium is assumed to start at $T = 70$. We can see that when there is a liquidity premium, the value of extension will exhibit a kinked shape, with a kink at $T = 70$.

A.8 House Price Seasonality

An important feature of the UK housing market, as well as in other countries including the United States, is that it is highly seasonal, with systematically higher sale prices in the second and third quarter (the "hot months") and lower prices in the first and fourth quarters (Ngai and Tenreyro, 2014). This may be associated with a number of factors, including mortgage conditions and market tightness, which we would not expect to affect the long-run and therefore ought to be differenced out by our estimator.

To determine whether our estimates are seasonal, we must select control properties which transact in the same quarter as extended properties. Then, we can estimate y_t^* at a quarterly frequency, as described in the main text. We then test for seasonality by regressing our time series of y_t^* on quarter dummy variables, controlling for year fixed effects. The regression results are presented in Table A.5, which indicate that there is no statistically significant difference in the level of y_t^* across quarters.

Table A.5: Test for Estimate Seasonality

	(1)
2nd Quarter	0.06
	(0.08)
3rd Quarter	-0.05
	(0.06)
4th Quarter	-0.00
	(0.06)
Year FE	✓
N	84

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table indicates the regression results for a regression of quarter dummies on y_t^*. We control for year fixed effects and weight by the inverse variance of the y_t^* estimate. Standard errors are heteroskedasticity robust.

A.9 Components of y^* for UK Housing—A VAR Analysis

This section uses a VAR to estimate long run risk premia and expected capital gains for UK housing, in order to show that neither of these can account for the trend fall in the long-term yield for UK housing that we have documented. We stress that this exercise is tentative, given the inherent uncertainty of any VAR based procedure.

To estimate risk premia and capital gains, we follow the standard practice of Campbell and Shiller (1988). The inputs into the VAR are: the growth rate of rents, the rent to price ratio, and the GDP growth rate. To estimate long run expected capital gains, in every quarter, we estimate forecasts of the rental growth for the subsequent 30 years and take the mean, recalling that long run expected rent growth is equal to long run expected capital gains. To estimate long run housing risk premia, we calculate the 30-year ahead predicted Rent-to-Price ratio (again using forecasts from the VAR). Then $\zeta = [\text{Predicted RTP}] + g - r$, where r is the 10 year 15 year real forward rate.

Figure A.33 displays the results. The estimates for long run expected capital gains, g_P^*, are stable. Estimates of the long run housing risk premium, ζ^* have been rising in the post-Great Recession era. Therefore the estimates suggest that neither capital gains nor risk premia for UK housing can account for the trend fall in the long-term housing yield, $y^* \equiv r^* + \zeta^* - g_P^*$, that we have documented.

Figure A.33: Risk Premia and Capital Gains for UK Housing in the Long Run

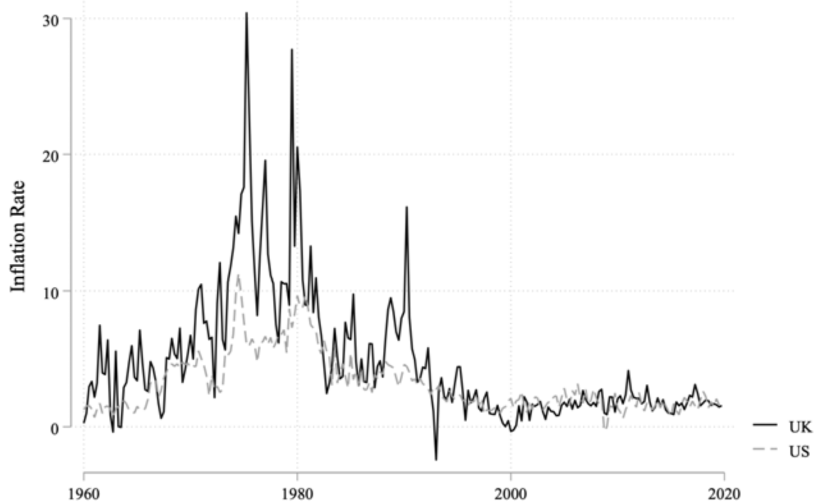
The figure presents estimates of long-run housing risk premia and capital gains estimated using a VAR. Estimates are produced using data on the UK rent-to-price ratio from the OECD Housing Price Index, as well as the mean level of house prices from Land Registry data and the mean level of rents from the Valuation Office Agency. UK GDP data is obtained from the St. Louis Fed, Federal Reserve Economic Data and long-run UK interest rates are obtained from the Bank of England.

A.10 Comparison with Holston et al. (2017) (HLW)

Our natural experiment and microdata approach to estimating long-run yields complements the structural time series approach to estimating the natural rate of interest in real-time. The most well-known approach to estimate r^* is the method proposed by Laubach & Williams (2003) and Holston, Laubach & Williams (2017). This method assumes that the output gap is an autoregressive, and natural output is a random walk with drift. This output gap is then linked to the real natural rate of interest via an Euler equation and a Phillips curve.

When estimated using US data, HLW estimate that r^* has fallen by 2.9pp between 2000 and 2020, with a mean standard error on r^* of 1.35pp per quarter. When estimated with UK data, the decline of r^* over this time period is only 0.81pp, with average standard error of 4.3pp, as seen in Figure A.35a. The reason why the UK estimate has much larger standard errors has to do with higher inflation volatility in the UK, which is a small open economy and is therefore more responsive to exchange rate fluctuations (Figure A.34).

Figure A.34: Inflation Rate In US and UK

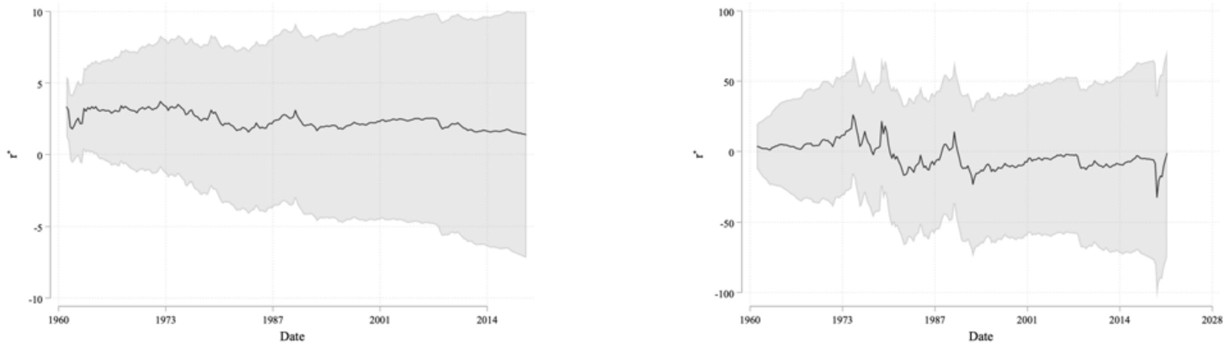


The solid dark grey line shows inflation in the UK between 1960-2020. The dashed light grey line shows inflation in the US between 1960-2020. The data is the same as that used by HLW.

Our estimates are a useful addition to those of HLW for several reasons. First, our estimate relies on minimal structure and is largely model-free. Moreover, our standard errors are on average 0.2 percentage points over our sample period — an order of magnitude smaller than the HLW r^* estimate for the UK.

Most importantly, our estimates persist throughout the pandemic and post-pandemic era and provide valuable insight on the growing wedge between discount rates of housing and of long-term government bonds. In contrast, when post-2020 data is included, the HLW methodology is de-stabilized by the steep decline in GDP in the second quarter of 2020. In the case of the US, the model is incapable of converging when post 2020 data is included. In the UK, inclusion of post 2020 data results in implausible estimates; ranging from -32.2% in the third quarter of 2020 to 25.9% in the second quarter of 1975. This error not only affects the pandemic era, but also propagates to previous to previous decades (Figure A.35b).

Figure A.35: HLW estimate of r^*



(a) Pre COVID-19 Pandemic

(b) Post COVID-19 Pandemic

Panel (a) shows HLW's estimates of r^* in the UK using pre-2020 data. Panel (b) shows HLW's estiamtes of r^* using post-2020 data. The shaded area represents the 95% confidence interval.

A.11 Calculating the Extension Hazard Rate

In this section, we explain how we calculate the extension hazard rate, shown in Figure A.3 and Figure A.4. We define the conditional probability that a property i extends given that it has duration T as $\theta_i(T) = P(\text{Extended At } S|\text{Duration} = S)$. To get the total cumulative probability that a T duration property i will extend over the course of its lifetime, we must convert our conditional probabilities to unconditional probabilities as follows,

$$\begin{aligned}\pi_i(S) &= P_i(\text{Extended At } S) \\ &= P_i(\text{Extended At } i|\text{Duration} = S)P_i(\text{Duration} = i) \\ &= \theta_i(S)\prod_{U=S+1}^{T}(1-\theta_i(U))\end{aligned}$$

The cumulative probability that property i extends over its lifetime is then given by $\Pi_i(T) = \sum_{S=1}^{T}\pi_i(S)$. In Figure A.4, we scale the hazard rate up by a factor of 1.17. This is because our method to identify lease extensions does not capture extensions that have no transactions before extension. We estimate that there are about 17% more extensions that have been extended but do not have pre-extension transaction data.

Then, the price of a T duration property at time t is given by the following recursive formula,

$$P_{i,t}^{T} = \frac{R_{i,t+1} + \theta_i(T)(P_{i,t+1}^{T+90-1} - \kappa_{t+1}^{T-1}) + (1-\theta_i(T))P_{t+1}^{T-1}}{1 + r^* + \zeta^*} \tag{11}$$

Intuitively, the price of a T duration asset is the discounted dividends it yields next period, plus with probability $1-\theta_i(T)$, the price of a $P_{i,t+1}^{T-1}$ asset, and with probability $\theta_i(T)$, the price of a $P_{i,t+1}^{T+90-1}$ duration asset minus the cost of extending, all appropriately discounted.

A.12 Difference-in-Differences Estimator with Option Value

This section derives our differences-in-differences estimator of y^* in the presence of option value from lease extensions. Let Π_{Tt}^H be the likelihood that a lease, with $T > 80$ years of duration remaining, extends before its duration reaches 80 years. Let Π_{Tt}^L be the likelihood that a lease with $T \leq 80$ years of duration remaining is extended at some point before expiration. Assume a constant discount rate y^*, and also assume that the event of extending is uncorrelated with the stochastic discount factor of the extender. Then the price of a leasehold is the present value of its cashflows, i.e.

$$P_t^T = \int_0^T e^{-y^* s} R_{t+s} ds + \Pi_{Tt}^H \left(1-\alpha_t^H\right) \int_T^{T+90} e^{-y^* s} R_{t+s} ds + \Pi_{Tt}^L \left(1-\alpha_t^L\right) \int_T^{T+90} e^{-y^* s} R_{t+s} ds.$$

In this equation, the first term is the present value of the first T years of service flow. The second term is the next 90 years of service flow, scaled by the share going to the leasholder, $\left(1-\alpha^H\right)$; and the likelihood that the lease extends at any time before it falls below 80 years remaining, Π_H^T. The third term is the analogous option value if the lease extends with less than 80 years remaining. Rearranging this expression implies

$$\begin{aligned}
P_t^T &= \int_0^T e^{-y_t^*} R_{t+s} ds + \Pi_{Tt}^H \left(1-\alpha_t^H\right) \int_T^{T+90} e^{-y_t^* s} R_{t+s} ds + \Pi_{Tt}^L \left(1-\alpha_t^L\right) \int_T^{T+90} e^{-y_t^* s} R_{t+s} ds \\
&= \int_0^T e^{-y_t^* s} R_t ds + \Pi_{Tt}^H \left(1-\alpha_t^H\right) e^{-y_t^* T} R_t \int_0^{90} e^{y_t^* s} ds + \Pi_{Tt}^L \left(1-\alpha_t^L\right) e^{-y_t^* T} R_t \int_0^{90} e^{-y_t^* s} ds \\
&= \int_0^T e^{-y_t^* s} R_t ds + e^{-y_t^* T} R_t \int_0^{90} e^{y_t^* s} ds \left[\Pi_{Tt}^H \left(1-\alpha_t^H\right) + \Pi_{Tt}^L \left(1-\alpha_t^L\right)\right] \\
&= \int_0^T e^{-y_t^* s} R_t ds + e^{-y_t^* T} R_t \int_T^{T+90} e^{y_t^* s} ds \left[\Pi_{Tt}^H \left(1-\alpha_t^H\right) + \Pi_{Tt}^L \left(1-\alpha_t^L\right)\right] \\
&= \frac{1-e^{-y_t^* T}}{y_t^*} R_t + e^{-y_t^* T} R_t \frac{1-e^{-y_t^* 90}}{y_t^*} \left[\Pi_{Tt}^H \left(1-\alpha_t^H\right) + \Pi_{Tt}^L \left(1-\alpha_t^L\right)\right].
\end{aligned}$$

Then for two properties i and j with identical service flow growth, where property i extends and j does not, we have

$$\begin{aligned}
\Delta_{it}^T &= \log\left(\frac{1-e^{-y_t^*(T+90)}}{y_t^*}\right) - \log\left(\frac{1-e^{-y_t^* T}}{y_t^*} + e^{-y_t^* T} \frac{1-e^{-y_t^* 90}}{y_t^*} \left[\Pi_{Tt}^H \left(1-\alpha_t^H\right) + \Pi_{Tt}^L \left(1-\alpha_t^L\right)\right]\right) \\
&= \log\left(1-e^{-y_t^*(T+90)}\right) - \log\left(\left(1-e^{-y_t^* T}\right) + e^{-y_t^* T}\left(1-e^{-y_t^* 90}\right)\left[\Pi_{Tt}^H \left(1-\alpha_t^H\right) + \Pi_{Tt}^L \left(1-\alpha_t^L\right)\right]\right),
\end{aligned}$$

which is the expression in the main text.

A.13 Measuring the Option Value of Lease Extension

So far, we have assumed that there is no option value from lease extension, because leaseholders pay the market value of extension to freeholders when they extend. This section studies option value and its consequences for estimates of y^*, in three steps. First, we present a simple framework that encompasses a key feature of the UK law on extensions, namely discontinuities in the cost of lease extensions when leaseholds have 80 years remaining. Second, we use this framework to derive discontinuity based tests about whether there is option value. These tests indicate that the baseline estimate of the fall in y^*, which ignores option value, is a lower bound for the true fall in y^*. Third, we generalize our the difference-in-differences estimator to point identify the size of option value. We find that our baseline estimate of y^*, which ignores option value, is an excellent approximation.

A.13.1 A Framework for Lease Extension Costs

We now summarize the institutional framework of lease extensions and develop a simple model of this framework.

Tribunals. Leaseholders are legally entitled to a 90-year extension by the 1993 Leasehold Reform, Housing and Urban Development Act. According to the act, lease extensions ought to be priced at their market value, the present value of service flows from the lease. However the leaseholder and freeholder negotiate the size of the market value, by independently hiring surveyors. If there is no agreement, a Residential Property Tribunal determines the value of the extension after a costly legal process. Tribunals require that the extension is 90 years long, and price extensions using a two part formula, which requires an estimate of *reversion value* and *marriage value.*

Reversion value. The reversion value is the value of the lease extension according to a yield assumed by the tribunal. Therefore reversion value of property i at time t satisfies the formula

$$RV_t^T = \frac{R_{it}}{r_{RV}} \left(e^{-r_{RV}T} - e^{-r_{RV}(T+90)} \right), \tag{12}$$

which is the value of the lease extension according to a Gordon Growth model, with a discount rate of r_{RV} and a current service flow R_{it}. In practice, the service flow R_{it} is imputed from the price of an observably similar freehold property. The discount rate is fixed by the tribunal at $r_{RV} = 5\%$.

Marriage value. The marriage value is the tribunal's estimate of the market value of the lease extension, given by $MV_{it}^T = P_{it}^{T+90} - P_{it}^T$, which is the difference in price between property to be extended, with duration T, and the price of the same property with duration $T+90$. The price of the property with $T+90$ duration is imputed from the transacted prices of observably similar properties. Provided that the tribunal imputes correctly, the marriage value is the market value of the lease extension, and satisfies

$$MV_{it}^{T} = \frac{R_{it}}{y_t^*}\left(e^{-y_t^* T} - e^{-y_t^*(T+90)}\right). \tag{13}$$

Here, we have written the marriage value as the market value of a lease extension according to a Gordon Growth model, where the natural rate expected by the market, y_t^*, enters the equation for market value.

Reversion value vs. marriage value. Our approach makes extensive use of the following observation. Compare Equation (12) and Equation (13) for reversion and market value. Reversion value calculated by the tribunal is greater than market value, if and only if y^* is greater than the 5% yield assumed by the tribunal.

Discontinuity at 80 years duration. There is a discontinuity in the tribunal assessed cost of the lease extension when 80 years remain on the lease. For leases with more than 80 years remaining, the tribunal dictates that only the reversion value must be paid, whereas for leases with less than 80 years remaining, the cost is the average of the reversion value and the marriage value.

Tribunal costs to leaseholder. There are significant costs to the leaseholder of appealing to the tribunal. These costs include time and information costs, uncertainty from the outcome of the tribunal, costs of hiring a survey to value the lease, and legal costs associate with the tribunal. Moreover the law dictates that leaseholders must cover all of the freeholder's legal and surveyor fees associated with extension. We will denote these costs for property i by γR_{it}, which we assume for simplicity scales with the current service flow R_{it}.

Leaseholder's participation constraint. The leaseholder also has a participation constraint—if the tribunal assessed costs of extending are greater than the value of extending, the leaseholder can opt not to extend. Therefore the freeholder should reduce costs down to the market value of extension, so that a transaction can occur. The converse is not true: if the tribunal associated costs are less than the market value of extending, the leaseholder will choose the tribunal costs instead of paying the freeholder the market value of extension.

Total extension costs. We can summarize the cost of extension κ_{it}^{T}, of a property i with T years remaining at time t, as

$$\kappa_{it}^{T} = \begin{cases} \min\left\{RV_{it}^{T} + \gamma R_{it}, MV_{it}^{T}\right\} & T \geq 80 \\ \min\left\{\frac{RV_{it}^{T}+MV_{it}^{T}}{2} + \gamma R_{it}, MV_{it}^{T}\right\} & T < 80 \end{cases} \tag{14}$$

Equation (14) recognizes that the cost of extension κ_{it}^T is the minimum of the market value of extension from the present value of rents; and the tribunal recommended value of extension plus the costs of a tribunal settlement. For simplicity, we have equated the market value of extension to MV_{it}^T, the marriage value. The tribunal recommended settlement varies discontinuously at 80 years, provided that the market value and the reversion value are not equal.[41]

Price of a T duration property. We assume property prices satisfy a simple no arbitrage assumption. Therefore the price of a property P_{it}^T, that has not extended, must equal the price of an otherwise similar property that has extended, after deducting lease extension costs. That is, prices satisfy

$$P_{it}^T = P_{it}^{T+90} - \kappa_{it}^T. \tag{15}$$

We denote $\alpha_t^T = \kappa_{it}^T / MV_{it}^T \leq 1$ as the share of the extension value that is lost by the leaseholder, noting that under our assumptions α does not depend on the service flow i.

Remark on option value. When $\alpha_t^T = 1$ we say there is *zero option value* and when $\alpha_t^T < 1$ there is *positive option value.* This terminology acknowledges that when $\alpha_t^T = 1$ for all T, then all of the gains of lease extension are lost to the leaseholder. As a result, the option to extend the lease has no value to the leaseholder ex ante.

A.13.2 A Discontinuity Based Test for Option Value

We now use our framework for lease extension costs to derive a discontinuity based test of whether there is positive option value. We summarize our predictions in the following proposition.

Proposition A.1. *There exists some value $\bar{r}_K \leq r_{RV}$ such that:*

1. *If y^* satisfies $y_t^* \geq \bar{r}_K$ then: (i) there is zero option value at all years of duration remaining, that is, $\alpha_t^T = 1$ for all T; (ii) the price of a leasehold is continuous in duration as the property's duration falls below 80 years, so $\lim_{T \to 80^-} P_{it}^T = \lim_{T \to 80^+} P_{it}^T$.*

2. *If y^* satisfies $y_t^* < \bar{r}_K$ then: (i) there is positive option value above 80 years in duration, that is, $\alpha_t^T < 1$ for all $T > 80$ and option value discontinuously falls at 80 years, so that α_t^T discontinuously increases at $T = 80$; (ii) the price of a leasehold discontinuously falls as the property's duration falls below 80 years, so $\lim_{T \to 80^-} P_{it}^T < \lim_{T \to 80^+} P_{it}^T$.*

[41] For exposition, we assume that the tribunal, the leaseholder and the freeholder all agree on the market value of the extension. Our qualitative and quantitative results are not affected by this assumption.

This proposition, which we prove in Appendix A.14, has two implications. First, we should expect zero option value at all durations, including above 80 years remaining, only if y^* is relatively high. Second, we can test for the presence of zero option value by searching for discontinuities in the price of leaseholds at 80 years.

Part (1) of the proposition shows that when y^* is high, there is zero option value at all durations. Intuitively, suppose that y^* is greater than the yield assumed by the tribunal to calculate the reversion value of the extension. Then, the reversion value of the lease extension calculated by the tribunal is greater than the market value. The freeholder will only require the leaseholder to pay the market value, given their participation constraint. Beneath 80 years, the tribunal assessed value remains above the market value of the lease extension—again, by the participation constraint, the freeholder can only force the leaseholder to pay the market value. Part (1) of the proposition also shows how to detect whether the economy has zero option value everywhere—in that case prices are continuous around 80 years of duration. Note that we can extrapolate from the 80 year threshold to conclude that there is no option value at any durations, because we know the functional form of lease extension costs.

Part (2) of the proposition shows that when y^* is low, there will be positive option value when lease durations are greater than 80 years. Suppose that y^* is significantly lower than the yield used by the tribunal to calculate the reversion value of the lease extension. Then the cost paid by leaseholders to extend via the tribunal, if more than 80 years remain, is less than the market value of extension plus time and legal costs—meaning positive option value. Beneath 80 years, the tribunal assessed cost of lease extension discontinuously increases, since the tribunal assessed extension cost is now a weighted average of the reversion value and market value plus additional costs, and market value is greater than reversion value. As a result, the price of leases must discontinuously fall. Importantly, Part (2) of the proposition does not rule out full holdup for leases with less than 80 years remaining.

We use Proposition A.1 to test for whether there is option value. The long-term housing yield seems to have been declining over time. As a result, our proposition suggests that there should be zero option value at all durations, only in the early part of the sample. Later on, there should be positive option value, at least for long duration leases. Our discontinuity based test confirms these predictions.

We test for holdup by estimating whether there is a discontinuity in prices at 80 years, before and after 2010. To determine whether there is a discontinuity in property price at $T = 80$, we can estimate,

$$\frac{\Delta_h \log P_{it}^T}{h} = \alpha + \beta \cdot \mathbb{1}\text{Crossed } 80 + \Gamma_{i,t,t-h} \tag{16}$$

where the left hand side variable is the annualized log change in price of a property i between time $t-h$ and t and the right hand side variable is a dummy which checks whether i fell below 80 between time $t-h$ and t. More precisely, we say that a property crossed 80 if at time t, $T < 80$ and at time $t-h$, $T > 82$. We choose a cutoff of 82 because leaseholders cannot extend through the tribunals during the first two years of ownership, so any property purchased with less than 82 years remaining must pay the marriage value upon extension. $\Gamma_{i,t,t-h}$ are Purchase Year x Sale Year x Local Authority fixed effects. We restrict the sample to properties with duration between 70 and 90, to get the local effect around 80.

Table A.6 reports the estimated coefficient from Equation (16). The first column presents estimates before 2010. There is no statistically significant discontinuity in price at 80. In fact, Appendix Figure A.28 shows that in the pre-2010 period, the estimated coefficient is approximately in the middle of a distribution of "placebo" experiments using 30 placebo cutoffs between 70-100. Column (2) of Table A.6 presents estimates after 2010. There is a significant discontinuity in the price of properties that fall below 80 years duration remaining. This fall is much greater in magnitude than any placebo test with cutoffs ranging from 70-100, shown in Appendix Figure A.28. Taken together, the price discontinuity results show that there is zero option value at all durations before 2010, and positive option value for long duration leases after 2010. However the results do not pin down whether there is full holdup for leases with less than 80 years remaining.

Table A.6: Test for Discontinuity at 80

	(1)	(2)
Crossed Cutoff	-0.05	-0.89***
	(0.09)	(0.08)
Sale Year x Purchase Year x LA FE	✓	✓
Period	Pre 2010	Post 2010
N	64,695	12,322

Standard errors in parentheses

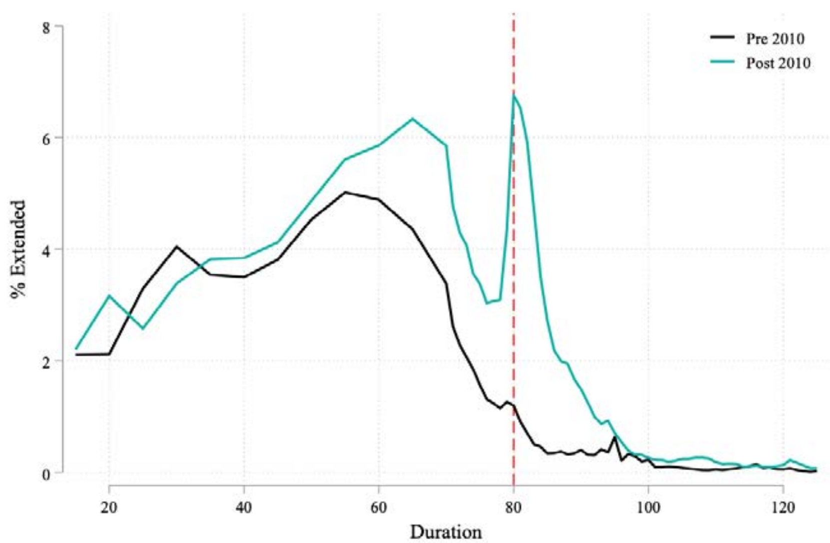
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table provides an estimate of the discontinuity in log price at $T = 80$. The sample includes properties with duration between 70 and 90. The first column is run on the pre-2010 data, and the second column is run on post-2010 data. Standard errors are clustered at the purchase year, sale year and Local Authority (LA) level.

Proposition A.1 also suggests there should be time varying bunching of lease extensions. When y^* is relatively high, there is no gain to extending slightly before 80 years remaining; whereas when the yield is low there can be large gains to extending before 80 years. Therefore lease extensions should bunch in a time varying fashion—leases should be more likely to extend shortly before 80 years remain, only if y^* is relatively low.

Figure A.36 shows precisely this pattern of bunching. In the figure, we plot the likelihood that a lease extends when it has T years remaining, separately before and after 2010. Before 2010, the likelihood that a lease extends smoothly increases as leases cross the 80 year threshold. After 2010, the likelihood jumps just before 80 years, and there is a missing mass after 80 years. This time varying bunching strongly suggests that there is a difference in option value above versus below 80 years, only when y^* is low—consistent with Proposition A.1. Our results on option value so far imply an informative bound—the estimates of

Figure A.36: Hazard Rate of Extension, Before and After 2010



The figure shows the conditional probability of extension for each T. The black line shows the probability before 2010 and the blue line shows the probability after 2010. We exclude the post-2020 pandemic period due to disruptions to the lease registration process, which resulted in lower extension rates than usual.

the fall in y^* from the main analysis of Section 5 , which assumes zero option value at all durations and times, are a lower bound on the magnitude of the true fall. Intuitively, if there is positive option value then the *true* duration of non-extended leases is higher than their *notional* duration. As such the price gain from lease extension is associated with a smaller increase in duration after lease extension. Therefore incorrectly assuming zero option value biases estimates of y^* upward. Given that option value emerges later in the sample, this bias increases over time, meaning the estimates that ignore this source of bias will under-estimate the fall in y^*. We now introduce more structure to show that this bias is small.

A.13.3 A Difference-in-Differences Estimator of Option Value

This subsection directly estimates the degree of option value before and after 2010, both above and below the 80 year threshold, and explores the implications for y^*. To do so, we introduce more structure by embedding the framework for lease extension costs into our difference-in-differences estimator of y^*. Doing so lets us point identify option value, using a different source of variation from the discontinuity based tests of the previous subsection.

In order to incorporate option value and the threshold in a simple fashion, we take as given the probability that a lease of duration T extends at time t. We also assume that the share of extension value lost by leaseholders is piecewise constant in duration and discontinuous at 80 years remaining, which captures the discontinuities imposed by the tribunal. Therefore we have $\alpha_t^T = \alpha_t^H$ if $T > 80$ and $\alpha_t^T = \alpha_t^L$ otherwise. In this case, Appendix A.12 shows that the difference-in-differences estimator of the price gain upon lease extension becomes

$$\begin{aligned}\Delta_{it}^T = &\log\left(1 - e^{-y_t^*(T_{it}+90)}\right) - \log\Big(\left(1 - e^{-y_t^* T}\right) \\ &+ \left[\Pi_{Tt}^H\left(1 - \alpha_t^H\right) + \Pi_{Tt}^L\left(1 - \alpha_t^L\right)\right] e^{-y_t^* T_{it}}\left(1 - e^{-y_t^* 90}\right)\Big) \qquad (17)\end{aligned}$$

For simplicity, this derivation makes the additional assumption that leases extend only once.[42] In this equation, Π_{Tt}^H is the probability that a lease with $T > 80$ years remaining extends with more than 80 years remaining. Π_{Tt}^L is the probability that a lease extends with less than 80 years remaining. The cumulative probability of extension is derived from the observed extension hazard rate, as shown in Figure A.36. Equation (17) is the same as the baseline estimator Equation (4) either when $\alpha_H = \alpha_L = 1$, or $\Pi_H^T = \Pi_L = 0$. In either case, there is no option value from lease extension and the final term in square brackets vanishes.

We jointly estimate y_t^*, α_t^H and α_t^L, by using the difference-in-differences estimator with option value, Equation (17), as an input into a non-linear least squares estimation similar to the baseline procedure. Relative to the baseline estimation, we also add information on the probabilities of extension Π_H^T and Π_L, which helps to identify α_t^H and α_t^L.

The estimated y^*, α_t^H and α_t^L parameters are presented in Table A.7. In the estimation, we constrain the α parameters to lie between zero and one. The results suggest, once again, that $\alpha = 1$ for all T in the pre-2010 period. In the post-2010 period, there is positive option value when leases have more than 80 years remaining, but there is zero option value below 80 years remaining.

Our difference-in-difference estimates of option value in this subsection are consistent

[42]The value of subsequent extensions, in the very far future, is quantitatively small but complicates the algebra.

Table A.7: Estimating Alpha

	(1)	(2)
	4.85***	3.30***
	(0.08)	(0.01)
α_t^H	1.00***	0.50***
	(0.01)	(0.06)
α_t^L	1.00***	1.00***
	(0.00)	(0.08)
Period	Pre 2010	Post 2010
N	18,832	105,191

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table presents estimates of α_t^H and α_t^L, estimated jointly with y^* from a nonlinear least squares estimate of Equation (17). Estimates of α_t^L and α_t^H are constrained to lie within [0,1]. Standard errors are bootstrapped.

with the discontinuity-based results from the previous subsections—even though the two subsections use different sources of variation. In both subsections, there is zero option value for leases with less than 80 years remaining, at all times; and positive option value after 2010 for leases that extend with more than 80 years remaining. Our estimates of the change in option value before versus after 2010 for leases above 80 years, is also similar across the two subsections. In this section, we estimate a change of 0.5. In the previous subsection we estimate a change of 0.44.[43]

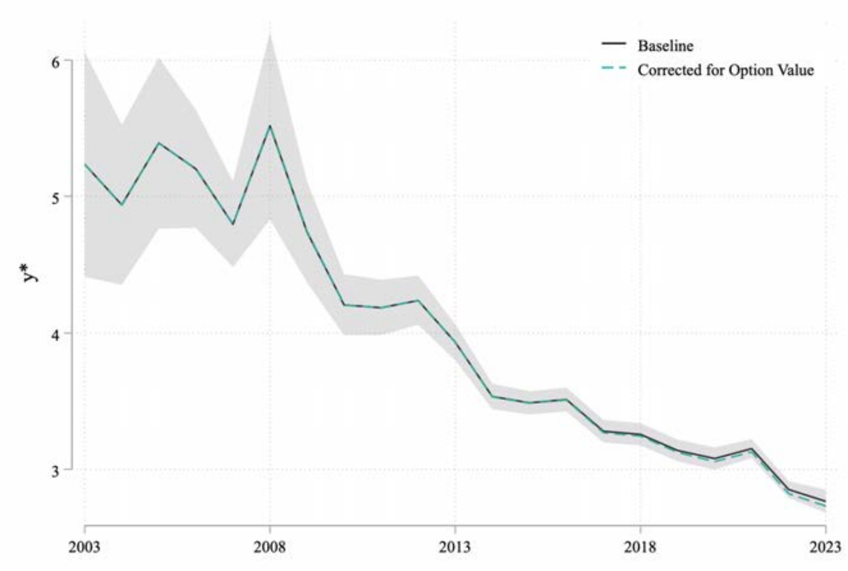
A.13.4 Estimates of y^* Corrected for Option Value

Finally, we present estimates of y^* that correct for option value using the estimates of option value. Figure A.37 presents a version of our y^* timeseries which corrects for $\alpha_t^H = 50\%$ in the post-2010 period. The solid line is the estimate of y^* from Section 5, using our baseline assumption of zero option value at all times and durations. The dashed line is the estimate of y^* using our estimates of the degree of option value.

The estimate of y^* that corrects for option value is very similar to the baseline estimates that assume no option value. The reason is that most leases do not extend with more than 80 years remaining, as Figure A.36 shows. Therefore the possibility of extending with more than 80 years remaining has a small effect on equilibrium prices, meaning departures from zero option value are quantitatively small.

[43] The details of this calculation are presented in Appendix A.13.5.

Figure A.37: Estimate of y^* Correcting for Option Value



The figure presents a corrected y^* timeseries, adjusting for the fact that $\alpha_t^H = 0.69$ for the post-2010 period. The black line presents the unadjusted estimates from Section 5.3.

A.13.5 Estimating Change in Option Value Using Discontinuities

As before, assume that $\alpha_t^H = \alpha_t^T$ for $T \geq 80$ and $\alpha_t^L = \alpha_t^T$ for $T < 80$, such that the share of holdup in a given time period t is fixed above and below 80, separately. In this section, we aim to estimate $\alpha_t^L - \alpha_t^H$ for the post-2010 period, using the discontinuity in prices at $T = 80$ observed in Table A.6. From the preceding subsection, the price of a property i with duration T is

$$P_{it}^T = \frac{R_{it}}{y^*}\left(1 - e^{-y^*T} + \left[\Pi_{Tt}^H\left(1-\alpha_t^H\right) + \Pi^L\left(1-\alpha_t^L\right)\right] e^{-y_t^* T_{it}}\left(1 - e^{-y_t^* 90}\right)\right) \tag{18}$$

To condense notation, denote the option value term

$$\Omega(T) \equiv \left[\Pi_{Tt}^H\left(1-\alpha_t^H\right) + \Pi^L\left(1-\alpha_t^L\right)\right] e^{-y_t^* T_{it}}\left(1 - e^{-y_t^* 90}\right).$$

Then, the difference in change in price between time $t-h$ and t of properties i and j is,

$$\begin{aligned}\Delta \log P_{it}^T - \Delta \log P_{jt}^T = &\log(1 - e^{-y^*T_i} + \Omega(T_i)) - \log(1 - e^{-y^*(T_i+h)} + \Omega(T_i + h)) \\ &- (\log(1 - e^{-y^*T_j} + \Omega(T_j)) - \log(1 - e^{-y^*(T_j+h)} + \Omega(T_j + h_j))).\end{aligned}$$

This equation acknowledges that option value will discontinuously change around the 80 year threshold, via changes in the Ω terms. We can then estimate α_t^H by nonlinear least squares on the same sample as for regression Equation (16), by setting $\alpha_t^L = 1$, which is what we estimate in Table A.7, and y^* to its mean for that period. We obtain an estimate of $\alpha_t^H = 0.56$ in the post-2010 period.

A.13.6 The 2006 Sportelli Case

In Appendix A.13, we have presented the asset pricing formula used by British property tribunal courts to price lease extensions, which utilizes a court-determined discount rate that we denote r_{RV}. Following the September 2006 Cadogan v. Sportelli case, and the subsequent October 2007 appeals case in the Court of Appeals, British property courts decided to set a fixed discount rate of $r_{RV} = 5\%$. Therefore, the decline in our estimator in the subsequent period cannot be attributed to arbitrary changes in the tribunal discount rate, and instead reflects changes in the market's expectations of long-run interest rates. Before the Sportelli decision, freeholders and leaseholders would negotiate the discount rate r_{RV} in court, which was not fixed at any particular value, but on average was approximately 6% in years immediately before the Sportelli case. Given our estimates of $\alpha_t^T = 1$ from Appendix A.13, which imply that there was no option-value from extending for leaseholders during this period, the Sportelli case should not have had an effect on leasehold market prices, and should therefore have no effect on our estimates.

We can verify this empirically by examining the price and extension rate of leaseholds before and after the 2006 Sportelli case. In particular, we will compare the price and extension rate around the cutoff of $T = 80$, which as we saw in Appendix A.13, is informative as to the difference between the market value of lease extension and the court dictated outside option. First, in Figure A.38, we plot the extension hazard rate before and after 2006. There is no bunching around 80 in either of these two periods, suggesting that the Sportelli case had no effect on the option-value from extension for leaseholders. Moreover, Table A.8 indicates that there is no significant discontinuity in price around $T = 80$ both before and after 2006, again implying that the Sportelli case did not significantly affect the actual cost of extension. In both Figure A.38 and Table A.8, we include the post-2010 period for reference, which does experience significant bunching and a discontinuity in price at $T = 80$.

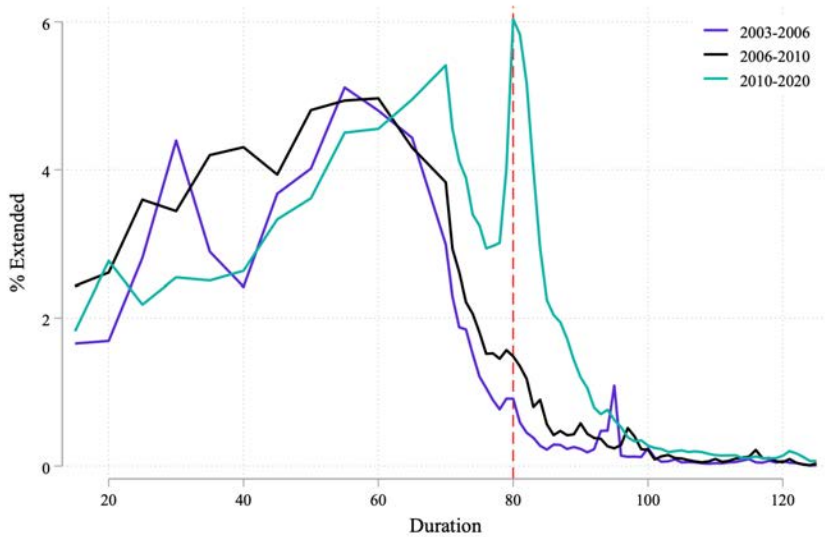


Figure A.38: Hazard Rate of Extension, Three Periods

The figure shows the conditional probability of extension for each T. The purple line shows the probability from 2003-2006, the black line shows the probability from 2006-2010 and the blue line shows the probability after 2010. We exclude the post-2020 pandemic period due to disruptions to the lease registration process, which resulted in lower extension rates than usual.

Table A.8: Test for Discontinuity at 80, Three Periods

	(1)	(2)	(3)
Crossed Cutoff	-0.11	0.20	-0.89***
	(0.08)	(0.13)	(0.08)
Sale Year x Purchase Year x LA FE	✓	✓	✓
Period	2003-2006	2006-2010	2010-2023
N	50,471	14,224	12,322

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table provides an estimate of the discontinuity in log price at $T = 80$. The sample includes properties with duration between 70 and 90. The first column is run on 2003-2006 data, the second column is run on 2006-2010 data, and the third column is run on 2010-2023 data. Standard errors are clustered at the purchase year, sale year and Local Authority (LA) level.

The small effect of the Sportelli case on leasehold prices and extension rates around the $T = 80$ cutoff provides additional evidence that in the pre-2010 period, the full value of extension was held by the freeholder. The tribunal discount rate fell from an average of about 6% before the Sportelli decision to 5% after. Therefore, by Equation (12) and Equation (14), we would expect for the Sportelli case to increase α_t^T, resulting in a shift in bunching and prices, if and only if $\alpha_t^T < 1$. The only scenario in which bunching and prices would not change as a result of a decline in r_{RV} is if leaseholds were already priced at their market value in the first place. The lack of a significant change in either bunching or prices in the aftermath of Cadogan v. Sportelli therefore suggests that $\alpha_t^T = 1$ for this entire period.

A.14 Proofs

Proposition A.2. *There exists some value $\bar{r}_K < r_{RV}$ such that:*

1. *If y^* satisfies $y_t^* \geq \bar{r}_K$ then*

   (a) *There is zero option value at all years of duration remaining, that is, $\alpha_t^T = 1$ for all T.*

   (b) *The price of a leasehold is continuous in duration as the property's duration falls below 80 years, so*
$$\lim_{T \to 80^-} P_{it}^T = \lim_{T \to 80^+} P_{it}^T.$$

2. *If y^* satisfies $y_t^* < \bar{r}_K$ then*

   (a) *There is positive option value above 80 years in duration, that is, $\alpha_t^T < 1$ for all $T > 80$ and option value discontinuously falls at 80 years, so that α_t^T discontinuously increases at $T = 80$.*

   (b) *The price of a leasehold discontinuously falls as the property's duration falls below 80 years, so*
$$\lim_{T \to 80^-} P_{it}^T < \lim_{T \to 80^+} P_{it}^T.$$

Proof. From Equation (14) and Equation (15), the price of a property is

$$P_{it}^T = \begin{cases} P_{it}^{T+90} - \min\left[RV_{it}^T + \gamma R_{it}, MV_{it}^T\right] & T \geq 80 \\ P_{it}^{T+90} - \min\left[\frac{RV_{it}^T + MV_{it}^T}{2} + \gamma R_{it}, MV_{it}^T\right] & T < 80. \end{cases} \tag{19}$$

Recall the definitions of reversion value and marriage value

$$MV_{it}^T = \frac{R_{it}}{y_t^*}\left(e^{-y_t^* T} - e^{-y_t^*(T+90)}\right) \tag{20}$$

$$RV_{it}^T = \frac{R_{it}}{r_{RV}}\left(e^{-r_{RV} T} - e^{-r_{RV}(T+90)}\right) \tag{21}$$

We will define $\bar{r}_K$ as the value of y_t^* such that $RV_{it}^T + \gamma R_{it} = MV_{it}^T$, that is, the tribunal costs for a lease above 80 years are exactly the market value. The value of $\bar{r}_K$ satisfies

$$\frac{R_{it}}{r_{RV}}\left(e^{-r_{RV}T} - e^{-r_{RV}(T+90)}\right) + \gamma R_{it} = \frac{R_{it}}{\bar{r}_K}\left(e^{-\bar{r}_K T} - e^{-\bar{r}_K(T+90)}\right)$$

$$\implies \frac{e^{-r_{RV}T} - e^{-r_{RV}(T+90)}}{r_{RV}} + \gamma = \frac{e^{-\bar{r}_K T} - e^{-\bar{r}_K(T+90)}}{\bar{r}_K} \tag{22}$$

where in the first line we have substituted in the definitions of marriage value (Equation (20)) and reversion value (Equation (21)). The right hand side of Equation (22) is strictly decreasing in $\bar{r}_K$. Therefore there is a unique value of $\bar{r}_K$ satisfying the equation.

Now we will prove part (1) of the proposition, in which $y_t^* \geq \bar{r}_K$. Equation (22) implies that for all $y_t^* \geq \bar{r}_K$ we must have

$$RV_{it}^T + \gamma R_{it} \geq MV_{it}^T. \tag{23}$$

Equation (23) implies

$$\begin{aligned} RV_{it}^T + \gamma R_{it} &\geq MV_{it}^T \\ \implies RV_{it}^T + \gamma R_{it} + MV_{it}^T &\geq 2MV_{it}^T \\ \implies \frac{RV_{it}^T + \gamma R_{it} + MV_{it}^T}{2} &\geq MV_{it}^T \\ \implies \frac{RV_{it}^T + MV_{it}^T}{2} + \gamma R_{it} &\geq MV_{it}^T \end{aligned} \tag{24}$$

Therefore for $y_t^* \geq \bar{r}_K$, prices satisfy

$$P_{it}^T = \begin{cases} P_{it}^{T+90} - MV_{it}^T & T \geq 80 \\ P_{it}^{T+90} - MV_{it}^T & T < 80, \end{cases} \tag{25}$$

where we have substituted Equation (23) and Equation (24) into Equation (19) for $y_t^* \geq \bar{r}_K$. Recall the definition of α_t^T as the ratio of the lease extension cost to MV_{it}^T. Equation (25) shows that $\alpha_t^T = 1$ for all t, which proves part (1a) of the proposition. Since the top and bottom of Equation (25) are equal at $T = 80$, prices are continuous at $T = 80$, which proves part (1b) of the proposition.

Now we will prove part (2) of the proposition, in which $y_t^* < \bar{r}_K$. Equation (22) implies that for all $y_t^* < \bar{r}_K$ we must have

$$RV_{it}^T + \gamma R_{it} < MV_{it}^T. \tag{26}$$

Then by Equation (19), the price of a property with more than 80 years duration remaining is

$$P_{it}^{T} = P_{it}^{T+90} - \left(RV_{it}^{T} + \gamma R_{it}\right).$$

The price of a property with less than 80 years remaining is

$$P_{it}^{T} = P_{it}^{T+90} - \min\left[\frac{RV_{it}^{T} + MV_{it}^{T}}{2} + \gamma R_{it}, MV_{it}^{T}\right]. \tag{27}$$

Also, note that

$$RV_{it}^{T} + \gamma R_{it} < \min\left[\frac{RV_{it}^{T} + MV_{it}^{T}}{2} + \gamma R_{it}, MV_{it}^{T}\right], \tag{28}$$

since $RV_{it}^{T} + \gamma R_{it} < MV_{it}^{T}$ by inequality (26) and also by inequality (Equation (26)) we have

$$\begin{gathered}
RV_{it}^{T} + \gamma R_{it} < MV_{it}^{T} \\
\implies RV_{it}^{T} < MV_{it}^{T} \\
\implies 2RV_{it}^{T} < RV_{it}^{T} + MV_{it}^{T} \\
\implies RV_{it}^{T} < \frac{RV_{it}^{T} + MV_{it}^{T}}{2} \\
\implies RV_{it}^{T} + \gamma R_{it} < \frac{RV_{it}^{T} + MV_{it}^{T}}{2} + \gamma R_{it}.
\end{gathered}$$

Equation (27) and Equation (28) imply that for $T < 80$

$$P_{it}^{T} < P_{it}^{T+90} - \left(RV_{it}^{T} + \gamma R_{it}\right)$$

Therefore prices discontinuously fall when T falls below 80 which proves part (2b) of the proposition. Since lease extension costs rise when T falls below 80, α_t^T also discontinuously rises when T falls below 80, which is part (2a) of the proposition.

□