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

Verónica Bäcker-Peral
Princeton

Jonathon Hazell
LSE

Atif Mian
Princeton and NBER

April 12, 2024

Abstract

Every month, a fraction of UK property leases are extended for another 90 years or more. We build a new dataset of thousands of these natural experiments from 2000 onwards to estimate the expected long term housing yield, y^*. Starting from a level of 5.3%, y^* starts to fall during the Great Recession, reaching a low of 2.8% in 2023. Real time data shows y^* has not risen since 2021, despite rising shorter term yields. Cross-sectional estimates show that y^* is higher in areas with more housing risk, and falls by more in areas with more inelastic housing supply.

*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. Bäcker-Peral: veronicabp@princeton.edu; Hazell: j.hazell@lse.ac.uk; Mian: atif@princeton.edu.

1 Introduction

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

This paper exploits a natural experiment from the UK in order to estimate the expected long term yield for housing. In the UK, most apartments are bought and sold as long duration leases, or "leaseholds", with initial lease terms that typically exceed 100 years. Periodically, a fraction of leases have their duration extended by 90 years or more.[1] We assemble a new administrative data set on all such lease extensions between 2000-2023, with over 130,000 extensions in total. The data contains market prices of leaseholds both before and after extension. We develop an empirical methodology to estimate the real-time dynamics of the expected long term housing yield, using the gain in leasehold market value due to duration extension. We have made our data and code public, and provide real time updates, so that other researchers can track and estimate y^* in real time.

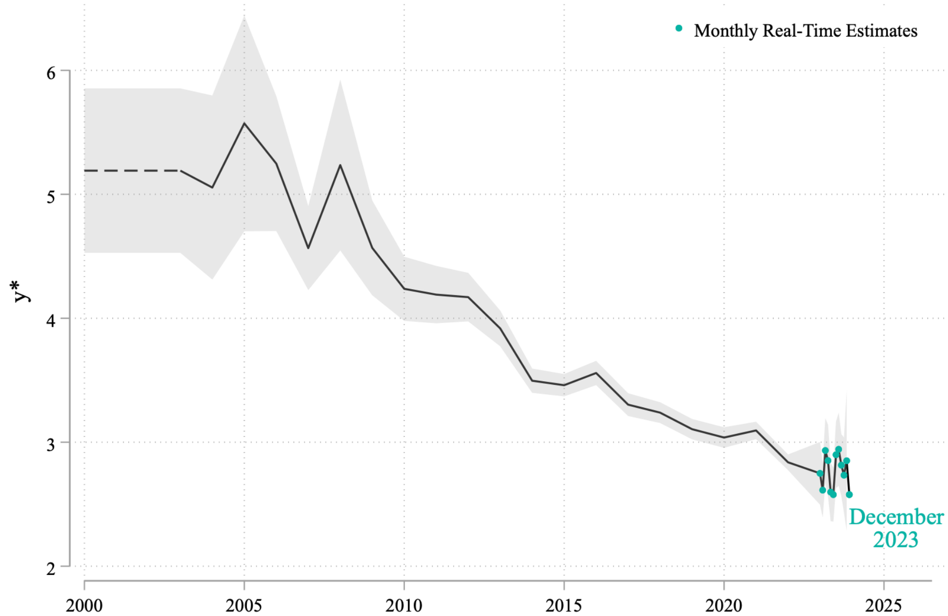
Formally, y^*_t is the market's expectation of the long-term housing yield at a point in time, i.e. $y^*_t = \lim_{u\to\infty} E_t\left[\frac{R_{t+u}}{P_{t+u}}\right]$, where $\frac{R}{P}$ is the rent-price ratio. y^*_t can equivalently be written as, $y^*_t \equiv r^*_t + \zeta^*_t - g^*_t$, where r^*_t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

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

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

Figure 1 plots the estimated average y^* from 2000 to 2023, and the shaded area represents the 95% confidence interval. There are two key findings. First, though y^* stays stable around 5.3% between 2000 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. The sample size expands over time, resulting in more precise estimates in recent years.

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



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

An important advantage of our data set is that it is not only public, but also updated in real time on monthly basis. Since y^* contains information about important macroeconomic parameters such as r^*, our methodology and real-time estimates should be useful for policy-relevant questions. There are close to a thousand lease extension natural experiments reported every month, allowing for precise real time estimates. Figure 1 illustrates real-time estimates, plotting separately the twelve monthly updates for 2023.[2]

Our natural experiment and microdata based approach for estimating the expected long-term yield has three main advantages. First, the approach "differences out" the shorter

[2]See the paper's website for data, code, real-time estimates and replication instructions.

end of the yield curve. As such, we can identify the long-term yield y^* without making hard-to-verify assumptions about the medium run shocks affecting the economy. Intuitively, the price before extension measures short term value, whereas the price after extension measures both the short and long term value of the same property. Hence the difference in prices depends almost exclusively on the long term yield — short and medium term shocks, such as monetary tightening, do not contaminate our estimate of y^*. Our estimator uses very long term information about prices, almost a century in the future—meaning even persistent medium term shocks will not matter. By contrast, forward yields from financial assets are typically available at most 10-30 years in the future.

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

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

The key assumption behind our methodology is "parallel trends", i.e. the service flow of housing must grow similarly for extended properties and their control group. We support this identification assumption in several 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, the estimator is not sensitive to controlling for observed heterogeneity, suggesting bias from unobserved characteristics is small (Altonji, Elder and Taber, 2005; Oster, 2019). Fourth, and most importantly, using rents as an observable proxy for service flows, we show that there is no difference in rental growth between extended and control properties either in the short run after extension, or the long run. One violation of parallel trends could arise if extended properties are more likely to be renovated before sale, which raises their service flow. However, using textual analysis of property listings, we find no evidence that extended properties renovate at a higher rate — consistent with our finding that rents evolve similarly compared to control properties.

In the final part of the paper, we turn to the macroeconomic implications of our estimates and arrive at three conclusions. First, while our estimate of y^* is from UK housing, long-term yields seem to have behaved similarly in other markets. For example, long-term forward bond yields closely track y^* for most of our sample. Similarly, long run trends in the yields of other assets such as global housing also track y^*. 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. Given that the decline in long term yields 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).

Second, the gap between current yields and long-term expected yields can be a useful real-time indicator of whether shorter or longer term factors are the main determinant of current asset price fluctuations. For instance, one important question is whether the expected long term yield on safe assets, 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. A second important question is whether current housing valuations are driven by transitory or more persistent factors. We can answer this question by studying the gap between current rent-price ratios and y^*, which is the expected long-term rent price ratio. We find that current rent-price ratios were similar to y^* in 2000 and from 2010 onward but deviated during the early 2000s. Therefore the behavior of y^* suggests a temporary house price boom in the 2000s, consistent with abundant other evidence. However the market expects current housing valuations to be more persistent.

Third, cross sectional heterogeneity in the dynamics of y^* reveals information about the role of inelastic land supply and housing risk in the determination of the long-term yield. We find that y^* falls by more in areas with inelastic land supply, i.e. areas with more constrained housing supply experience stronger rise in valuation as demand grows. Thus the large fall in aggregate y^* reflects in part the inelastic nature of land supply in the UK, consistent with past work (e.g. Miles and Monro, 2019). We also find that y^* is higher in areas with greater exposure to housing risk. We measure housing risk as the covariance between annual housing returns and a proxy for annual consumption growth, separately for each local authority in the U.K. Our results suggest that the level of y^* incorporates a substantial risk premium for housing in the long run.

Related Literature. 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. (2015) 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 develop a new bunching estimator to measure this option value in the Appendix.

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

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

[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. (2015) because our sample of lease extensions have different durations from the cross section of properties that they study.

[5] Related papers include Kiley (2015), Lubik and Matthes (2015), Johannsen and Mertens (2016), Crump

have discussed, the structural approach may be more vulnerable to mis-specification than our method.

Outline. The rest of the paper is structured as follows. Section 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 Housing Yield

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

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

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

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

This is also equivalent to the long-run rental-price ratio, i.e.

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

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

The challenge in empirical identification of y^* is that we need to strip out any effects of short run, transient shocks on housing yields, such as monetary policy, credit booms, short run bubbles and adjustment costs. It is also difficult to observe the true rental value, or

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).

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

"service flow" for housing. Section 4 explains how our empirical approach addresses these challenges in detail.

3 Data & Lease Extension Details

This paper builds a new administrative data set on all lease extensions in the UK from 2000 onwards. We are also making this data set public, along with its real time updates. This section describes the construction of the dataset.

3.1 Data

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

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

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

We also want to build our public lease extension data set in a way that it is automatically updated in real time every month. Doing so enables us to estimate y_t^* in real time going forward – something that we hope will be of use to academics, central bankers and other policy professionals. In order to track every new lease extension going forward, we start with the full Lease Register of May 2023, and update it every month with the latest lease register vintage to identify new lease extensions. As such our public data website keeps getting updated with new real-time monthly estimates of y^*. As of the time of writing, we have identified approximately 315,000 leasehold extensions since 2000. The distribution of extension times and lengths is presented in Table A.1.

HMLR also publishes data on all property transactions in England and Wales starting in 1995 and updated monthly. The data set includes the exact date, price and address for each transaction. We were provided with exact transaction IDs to merge transactions to closed lease titles. The open lease register does not include transaction IDs, so we conduct a fuzzy merge based on provided addresses to match open leases to transactions, as we detail in Appendix A.3.[7]

Private Hedonics Data

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

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

The second data set of housing characteristics and rental prices is from Zoopla, Inc. and 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.

[7] 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.

3.2 Lease Extension Details

This section provides institutional details about the process through which leases can be extended. The distinction between freeholds and leaseholds dates to medieval England, during which permanent ownership of land and property, known as "freehold" ownership, was available only to feudal nobility.[8] 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.

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

The lease length at the time of its issuance is referred to as "initial lease term", and the lease length at any future time is referred to as "remaining lease term". The distribution of leases can be divided into two groups; about 70% 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. Lease registration became mandatory in October 2003, so after this date, we capture all extensions of leasehold properties.

In principle, the freeholder and leaseholder can mutually agree to extend the existing length of a lease at any point in time for an agreed upon price. A 1993 Leasehold Reform, Housing and Urban Development Act (1993 Act) granted flat leasehold owners the right "to acquire a new lease" 90 years longer than the original lease, conditional on a one-off negotiated payment to the freeholder that reflects the market value of the gain in leasehold value from extension.[10] The legal recourse for lease extension is seldom used as it is costly for both parties to hire qualified surveyors, and the legal process can be lengthy. The law

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

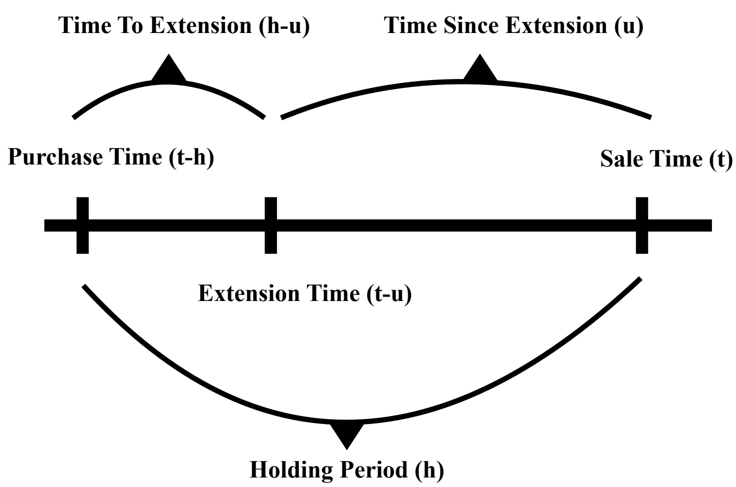
[9] This distribution is illustrated in Appendix Figure A.1.

[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.

acts more as an outside option for negotiating parties.

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

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

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

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

We restrict our sample to properties that were extended by more than 30 years, since leases that are extended by very short amounts may be unique. We also exclude properties which were extended with very short remaining durations, since we are focused on the long-end of the yield curve. Finally, we remove outliers, defined as the 1% of data that experienced highest/lowest price growth over the extension window relative to their controls. However, our main results are robust to including these properties.

Table 1 shows the distribution of lease extension length over time for our treatment pool of lease extensions. These are extended leases in our sample that have a transacted market price before and after lease extension. There are 138,154 such lease extensions. 90 years is the most common lease extension length, accounting for about 30% of all extensions, and there is also a high fraction of extensions that are for 700+ years. These extensions effectively convert short leaseholds into long leaseholds.

Table 1: Number of Extensions

Extension Amount	90	700+	Other	Total
2000-2005	804	2,606	2,619	6,029
2006-2010	3,436	6,812	5,005	15,253
2011-2015	11,718	15,379	10,297	37,394
2016-2020	16,767	18,796	11,610	47,173
2021-2023	12,900	13,366	6,039	32,305
Total	45,625	56,959	35,570	138,154

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

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

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

3.3 Transaction Lags

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

We can estimate the amount of time between the date in which the buyer and seller agree on a price and the date in which the price is recorded by the Land Registry by using property listing data from Rightmove. If we assume that sellers stop posting house listings before agreeing on a price with a buyer, then the last date that a listing is posted must precede the date in which the buyer and seller agree to a price. The median amount of time elapsed between the last property listing on Rightmove and the date recorded by the Land Registry for transactions which have an associated Rightmove listing is 3.7 months, with a mean of about 5 months. The full distribution of the time elapsed between the listing date and the Land Registry date recorded is presented in Figure A.9. Therefore our real time estimate is based on information with a lag of around 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 use the example of a lease that extends by 90 years as our illustrative example. The price of a leasehold P_t^T, with T years until expiration, and an option to extend by 90 years on expiration, is

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

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

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

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

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

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

4.1 A DiD-NLLS Estimator of y^*

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

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

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

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

In this equation, the first two terms represent discounting of the extended lease versus its control property. The final term is the difference between the growth rate of the service flow of housing, for the treatment versus the control group, over the length of the holding period. Equation (4) shows that identification of y from the difference-in-differences estimate $\underline{\Delta}_{it}$ requires the following "parallel trends" assumption: 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 (4) vanishes.

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

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

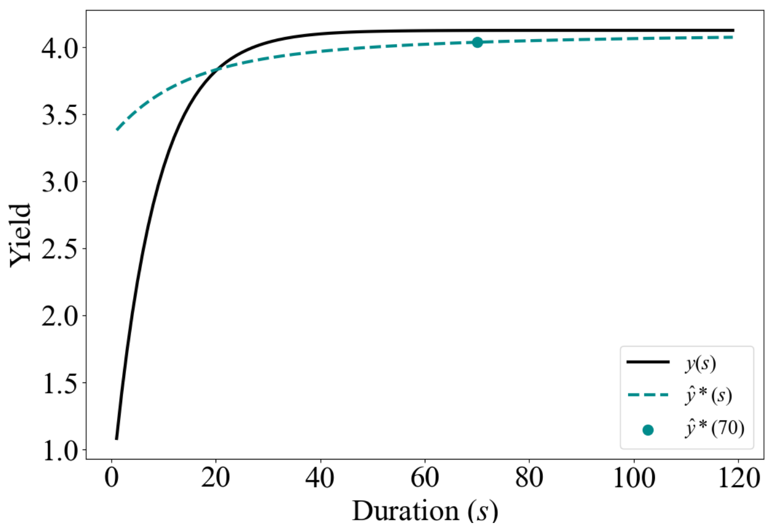
Equation (5) can be estimated using non-linear least squares (NLLS) to estimate y^*.

4.2 Properties of the DiD-NLLS estimator

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

The second advantage of our estimator is that it differences out the effect of short term yields on asset prices, and hence isolates the long-end forward of interest. We illustrate this feature of our estimator by numerically simulating a wide range of yield curves that all have the same long-term y^*, but differ greatly in the shape of yield curve over short to medium run horizon. The black solid line in Figure 3 presents one possible parameterization of the forward yield $y(s)$, where the forward yield curve $y(s)$ flattens out to y^* for $s \geq 40$ years, with y^* equal to around 4.0 pp.[11]

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 1 year gilts, the 10 year gilts and the 25 year gilts, averaged over the 2010-2020 period. $\hat{y}^*(s)$ is our estimator of y^* for each T, described below.

Given our parameterization of $y(s)$, we can solve for $\log P_{it}^{T+90} - \log P_{it}^{T}$ for all T. Then, for each T, we can solve for our estimator of y^* as a function of T numerically. The resulting values of our estimator, which we term $\hat{y}^*(T)$, are plotted in blue in Figure 3. Our estimator $\hat{y}^*(T)$ closely approximates the true long-run rate y^* for durations T after which the forward curve has flattened. We also plot the point estimate of $\hat{y}^*$ at $T = 70$, which is approximately the median duration of leaseholds at extension.

[11] We choose a flexible functional form $y(s) = \beta_1 - \beta_2 \cdot \beta_3^{-\beta_4(s-\beta_5)}$ and estimate the β parameters such that $y(0)$ is equal to spot rate on 1-year bonds, $y(10)$ is equal to the 10-year gilt yield, and the average of $y(s)$ for $10 \leq s \leq 25$ is equal to the 10 Year 15 Year gilt forward yield, averaged for the 2010-2020 period. For all the bond yields, we use the mean yield for our sample period. We present other possible parameterizations of $y(s)$ in Appendix A.6.

In fact, 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.[12]

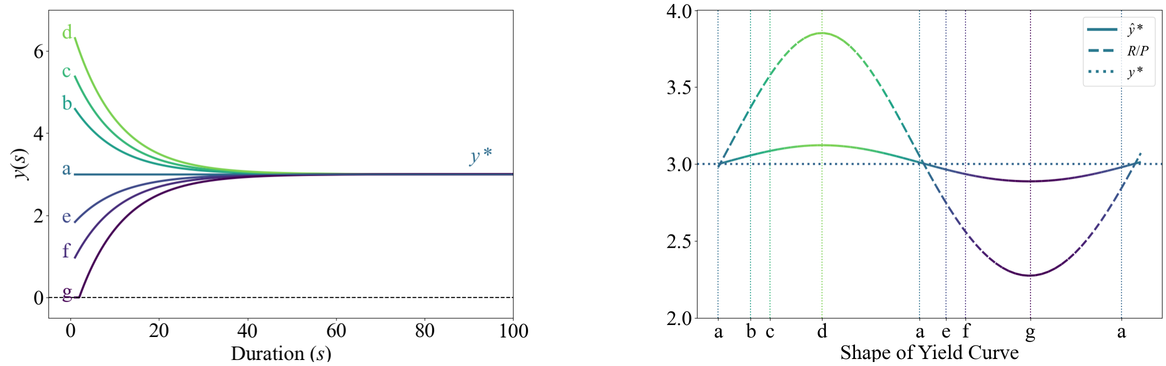
The reason why our estimator is able to provide a close approximation of y^* is because T is large, which effectively differences out most of the yield curve $y(s)$ for $s < T$. As T becomes smaller, the effect of the short-end on $\hat{y}^*(T)$ increases, meaning estimates of y^* become increasingly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

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

[12] In Appendix A.6 we show that we can use $\hat{y}^*(s)$ to bound the approximation error of our estimator, and therefore get a lower and upper bound for y^*.

Figure 4: "Differencing Out" the Short End



(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 $P^{\infty}_{it}/R_{it} = \int_t^{\infty} e^{-\int_t^s y(u)du} ds \equiv \frac{1}{R/P}$.

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".[13]

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

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

Here, we have generalized the expression of the estimator to allow a time varying expected long term yield, and to acknowledge that the duration of the property before extension, T_{it}, can vary. Equation (6) shows that we can estimate y_t^* by nonlinear least squares. The estimator is valid at any point in time, hence we can estimate the dynamics of y_t^*. Two statistics inform y_t^* in the estimator. First, the difference-in-difference $\underline{\Delta}_{it}$ can be calculated for every property i, as the difference in price growth between the extending property and its control. Second, the covariance between $\underline{\Delta}_{it}$ and the duration before extension T_{it} also

[13] 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.

helps to identify y^*. Regarding inference, we cluster standard errors at the level at which treatment is assigned, following standard practice (Abadie et al., 2023). Since treatment is assigned at the level of each extending property i, heteroskedasticity-robust standard errors without clustering suffice.

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

To construct the repeat sales control, we consider a treated property i, which was purchased at time $t-h$, sold at time t, and extended for 90 years at some time $t-h < t-u \leq t$.[14] Suppose this property has duration $T+h$ at purchase and duration $T+90$ at sale. The set of control properties, specific to property i, is those properties that (i) do not extend between $t-h$ and t; (ii) have a duration within 5 years of property i; and (iii) are within d km of property i. d is the smallest possible Haversine distance such that it is feasible to construct a repeat sales index for the counterfactual price of property i between $t-h$ and t. We discard property i if there are not enough controls within a 20 km radius. This procedure automatically adjusts for the different density of housing in urban and rural areas. Finally, we produce a repeat sales index for property i's control group, using a three-part procedure similar to Case and Shiller (1989) (see Appendix A.7 for details). In Appendix A.8, we show that these controls are similar to their extended counterparts on all of the main hedonic measures we observe in the Rightmove and Zoopla data. We sometimes denote the set including property i and its control properties as experiment i.

There are two advantages from using repeat sales methods. First, the identification assumption underlying our difference-in-differences estimator is essentially unchanged. Between $t-h$ and t, the growth of service flows for the treated property must be the same as for the control properties that constitute the treated property's repeat sales index. Second,

[14]In our primary analysis, we measure t at a quarterly frequency. We show robustness to using other frequencies in Figure A.19.

the repeat sales index method lets us construct controls for many more extending properties. With this method, we are able to construct controls for 126,018 of our 138,154 lease extension experiments.[15,16]

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 on average. This assumption could be violated if those who extend the lease are also more likely to renovate the property before sale. If this were the case, the extended properties will have higher rent growth before and after extension relative to the control properties. Since we have already excluded lease extensions that are likely to be renovated, extended and flipped within a couple of years of purchase, this concern should have been mitigated. Nonetheless we provide direct tests of the equivalent housing service flow growth assumption in this section.

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

In any event, we take the endogenous renovation concern seriously and provide two different tests for checking the parallel trends assumption.[17] As a first test of parallel trends, which directly focuses on renovations, we show that extending properties are no more likely to make home improvements than controls. First, we identify properties that were recently renovated based on textual analysis of Rightmove listing descriptions. In particular, we create a renovation dummy that is equal to one if a property's listing contains certain key

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

[16] 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.

[17] Additionally, in Appendix A.8 we provide evidence that extended and control properties are similar on observable characteristics.

words associated with renovation such as "renovated," "refurbished", or "improved." About 9% of all listings make reference to renovation. To verify that this is a valid measure of home improvement, we regress it against change in bedroom count for the renovated flat from the listing before renovation to the listing after renovation. The regression coefficient is positive and statistically significant; the probability that a property's bedroom count increases is more than twice as high if the property was marked as renovated in that year. Moreover, properties that were purchased and immediately resold ("flippers"), have much higher renovation rates and are far more likely to increase their bedroom count, as shown in Figure A.2. This is suggestive that both our text-based measure and the change in bedroom count are good indicators of renovations.[18] Our other main hedonic measures—bathroom count, living room count, and floor area—experience nearly zero growth across all flats, which is consistent with the UK's strict regulation of construction.

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

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

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

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

where $\Delta \log R_{i,j,t',t''}$ is growth in rents for property j between periods t' and t'' for which there are observed rents for both the treatment and control, $\alpha_{i,t',t''}$ are experiment $\times$ rental

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

Table 2: Renovation in Extended vs Non-Extended Flats

	(1) Renovation Rate	(2) Δ Bedrooms	(3) Δ Bathrooms	(4) Δ Living Rooms	(5) Δ Floor Area
Extension	-0.001	-0.003	0.000	-0.001	-0.051
	(0.001)	(0.002)	(0.000)	(0.000)	(0.034)
Experiment FE	✓	✓	✓	✓	✓
Control Mean	.091	.043	.002	.001	.279
N	154,394	42,926	32,592	31,472	34,030
N. Experiment	77,197	21,463	16,296	15,736	17,015

Standard errors in parentheses

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

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

transaction years fixed effects. We control for whether the property was transacted between rental listings, since properties are more likely to experience greater rental price growth when they change owners. The regression measures whether rental growth was higher for extended properties than for control properties that have rental transactions in the same years. Table 3 presents the results and finds small and insignificant differences in the behavior of rents between the treatment and control group. The first column is the baseline regression.

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

Column (3) re-estimates regression equation (7), where growth in control rents is now given by a repeat sales index of rent growth in each extending property's control group. Each observation is a rental cycle for which we have two extended property rental transactions, and we are able to construct a control repeat sales rent growth index. On average, there are approximately two rental cycles within each purchase-sale window. The resulting estimates are small and statistically insignificant.[19]

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

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

Table 3: Within-Experiment Window Rent Growth

	Δ log(Rent)		
	(1)	(2)	(3)
Extension	0.0010	-0.0010	0.0002
	(0.0007)	(0.0006)	(0.0007)
Experiment $\times$ Rent Years FE	✓		✓
Experiment FE		✓	
Annualized		✓	
RSI			✓
N	3,874,527	35,474	72,558
N. Experiment	16,131	17,737	18,447

Standard errors in parentheses

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

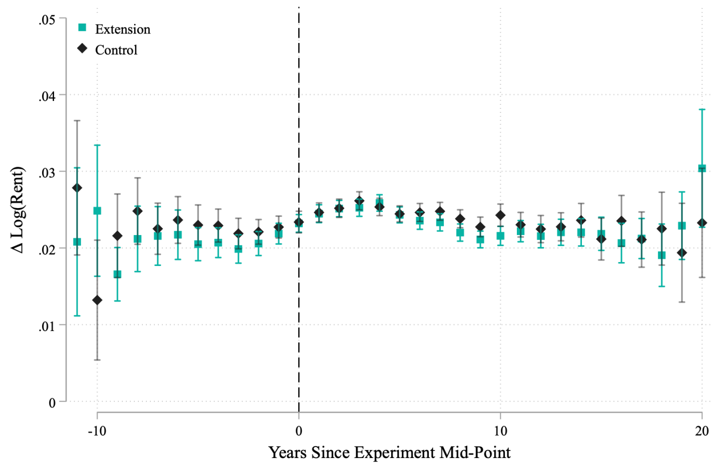
The table reports rent growth during the experiment window of $t - h$ to t for extended properties relative to control properties. The first column is the estimated coefficient from equation (7). We use mean annualized rent growth as the LHS variable, and do not control for time fixed effects. The third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

the long-run growth rate of rents, g^*, is the same for treatment and control properties.

We assess the assumption regarding long-run rental growth by expanding the window over which we study rent growth. Figure 5 plots annual rent growth for extension and control properties against time. We plot annual rent growth starting from the midpoint between the purchase and sale time of the extending property. We construct rent growth for the control properties associated with each extending property, using our baseline repeat sales method. We interpolate across years with missing rent data.[20] The results show that rent growth is similar for extending and control properties at all horizons including the long term, and is never statistically different. The results in this section should be reassuring regarding the identification assumption: extended properties do not have a higher propensity to get renovated, and their rental growth is no different from control properties.

[20]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.

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

5.1 Event Study Analysis of The Effect Of Lease Extension

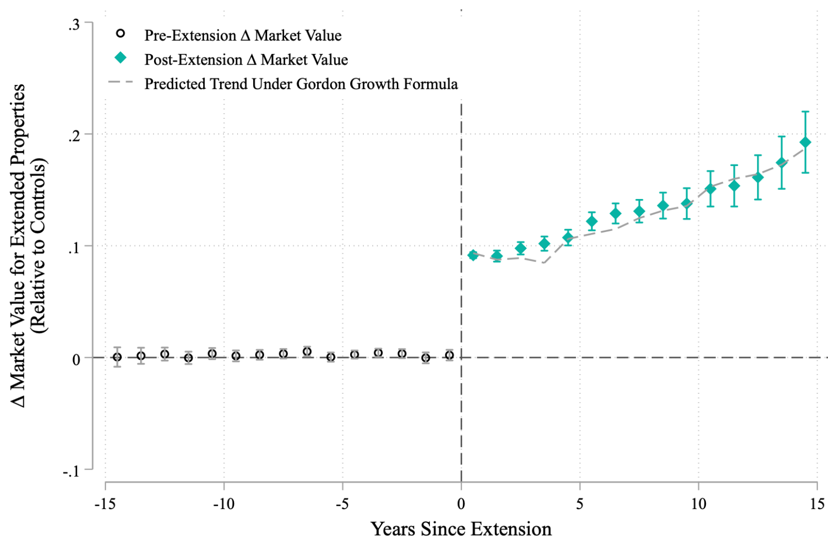
We develop intuition for our estimator of y^* through an event study representation of the methodology. Consider the modal leasehold extension in our sample that extends its duration from T to $T+90$ years. These leaseholds have a median duration at sale of 157 years, meaning that had they not extended, they would have had a duration of 67 years at sale.[21] The difference-in-differences estimate $\underline{\Delta}_{it}$, from equation (3), represents the gain in log market value at time t, of a leasehold that extends u periods earlier.

In Figure 6, we plot in blue the estimate $\underline{\Delta}_{it}$ against time between extension and sale, u. For instance, the first entry on the positive x axis is the mean DiD estimate of the increase in market value for leaseholds that sell one year after extension. One year after extension, the difference in log price between extended properties and their controls jumps by about 0.1 log points. Afterwards, the difference continues to grow exactly as predicted by asset pricing equation (2), so that leaseholds that sell 10 years after extension increase their market value by almost 0.2 log points. The increase in market value rises with the time since extension because (1) the value of the control leasehold that did not get extended falls more for each

[21] Appendix Figure A.11 displays the histogram of remaining lease term at sale for the extended and control properties.

passing year due to its shorter duration, which varies non-linearly with log price, and (2) y^* falls over our sample period. The dashed grey presents the post-extension price gain at each value of u predicted by our simple asset pricing model, given the mean duration and estimated y^* level for properties transacted in that period. The predicted trend aligns closely with the observed estimates of $\underline{\Delta}_{it}$.

Figure 6: Event Study Representation of Lease Extension



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

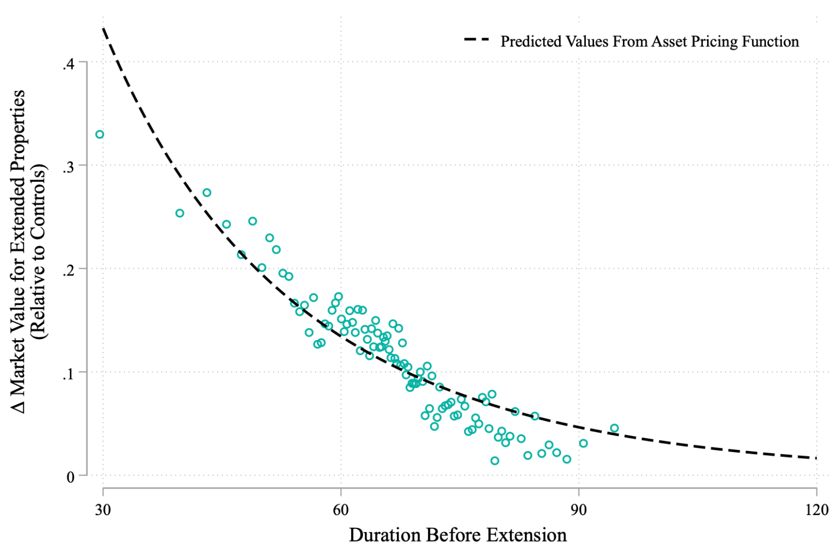
We also use the event study to investigate pre-trends. On the negative x-axis, we apply the difference-in-differences estimator to the pre period—that is, to the price growth of treated vs. control properties over the two transactions prior to the extension. In doing so, we can estimate whether there are pre-trends in price growth before extension. We plot the difference-in-differences estimate in the pre period against the number of years between the final transaction before the extension, and the extension itself. There are small and statistically insignificant pre-trends — extended properties experience similar growth to controls in the pre-period before extension, regardless of the number of years between transaction and extension.[22]

[22] Figure A.13 presents the event study for only 90-year extensions.

5.2 Estimating Average y^*

The difference-in-differences estimate, $\underline{\Delta}_{it}$, of the gain in market value as a result of duration extension is used in the non-linear equation (6) to estimate y^*. Figure 7 graphically illustrates how the estimation procedure works, using 90-year extensions as example. The figure binscatters $\underline{\Delta}_{it}$ estimates for various durations T_i. The dashed black line is the best-fit prediction from our NLLS estimation of equation (6). Figure 7 shows that the DiD estimate of the increase in market value after extension aligns remarkably well with the predicted increase according to the simple asset pricing formula of equation (6). Our simple asset pricing formula predicts that the percent valuation gain from a 90 year extension should be larger for extensions with shorter duration at the time of extension. Consistent with the formula, the percent gain in property value as a result of extension is decreasing in the duration before extension—ranging from only 7% for properties with 90 years remaining at extension, to more than 30% for properties with duration of 40 years at 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 (6).

More generally, there are three sources of variation used to identify y^* according to our NLLS DiD estimation procedure. First, for a given remaining duration T at extension, and a given extended duration $T + k$, the valuation gain $\underline{\Delta}_{it}$ is larger when y^* is lower. Second, as shown in Figure 7, the gain from extension increases as the duration before extension T falls. Moreover this increase is stronger when y^* is lower.[23] Third, the valuation gain is

[23]We can explicitly reject a linear relationship between $\underline{\Delta}_{it}$ and T_{it} by running the following quadratic

larger when the size of the lease extension amount amount k is bigger, with a larger gain when y^* is lower.

Table 4 presents the y^* estimates implied by Figure 7. Column (1) shows that y^* estimate for properties extended by 90 years is 3.5% over the full sample period. As discussed in Section 4.2, our estimator assumes that the long term forward yield curve for housing is flat, meaning a unique value of y^*. We now test if the assumption of a flat long term forward yield curve is accurate, using two sources of variation: in the duration T at which a leasehold is extended, and the size k of the increase in lease duration at extension.

Table 4: Estimated y^*

	Constant y^*	Flexible y^*				Constant y^*
	(1)	(2)	(3)	(4)	(5)	(6)
	$k = 90$	$T = 50$	$T = 60$	$T = 70$	$T = 80$	$k \geq 700$
y^*	3.47***	3.43	3.46	3.49	3.52	3.50***
	(0.023)					(0.020)
N	41,885					52,615
t-stat (700+ vs. 90)						1.11

Standard errors in parentheses

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

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

We estimate the degree to which estimates of the long term yield vary with duration, by parameterizing the estimate of y^* as $y(T) = \alpha + \beta \cdot T$, where T is the duration of the leasehold before extension. If the long term yield curve is sloped, then β will be different from zero. We exploit variation in duration after extension to estimate both α and β. We control for time of experiment by residualizing $\underline{\Delta}_{it}$ on year fixed effects before estimating y^*, to make sure spurious variation in when certain extensions are popular does not drive our result. Columns (2) through (5) of Table 4 present estimates of $y(T)$ under this more flexible form. The results show that $y(T)$ varies little over the range we observe it, even at very long durations, suggesting a flat long term yield curve.[24]

As another test of the slope of forward yield curve, we estimate y^* using the sample of leaseholds that extend by more than 700 years. These leaseholds are affected by much longer horizon yields than the baseline sample, which extend by 90 years. Column (6) of Table 4

regression: $\underline{\Delta}_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 T_{it}^2 + \epsilon_{it}$. The estimated β_i's are all statistically significantly different from zero.

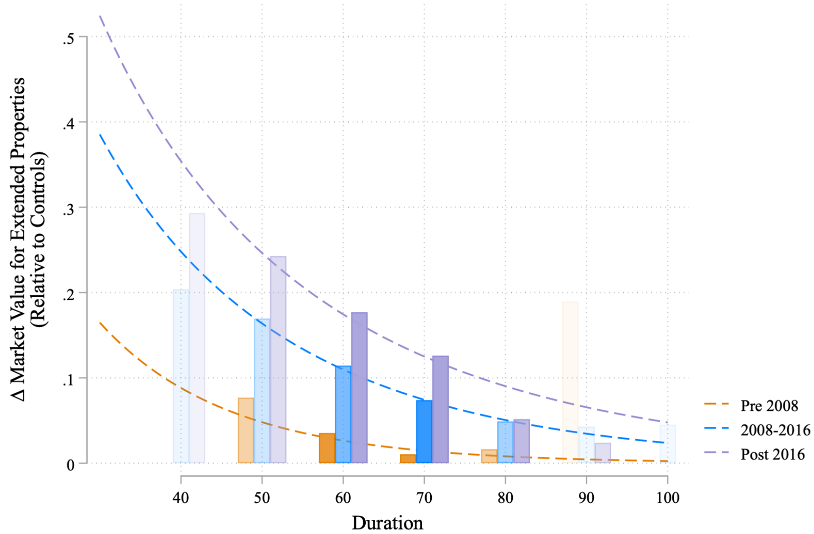
[24]In Figure A.14 we estimate the event study of Figure 6 separately for above and below median duration properties, which also documents the lack of pre-trends for both.

presents estimates of y^* using only extensions of more than 700 years. Estimates of y^* are remarkably similar for both groups, and statistically indistinguishable according to a t-test, the t-stat of which is presented at the bottom of Column (6). Therefore y^* seems to vary little at very long horizons, consistent with a flat yield curve.[25]

One concern about our estimate may be that short duration leaseholds have an additional liquidity premium, because banks might be less willing to issue a mortgage against shorter duration leaseholds. Appendix A.9 investigates this concern formally and finds little evidence of liquidity effects in our sample. For example, data from English Housing Survey shows that shorter duration leaseholds are as likely to have a mortgage, with similar mortgage characteristics in terms of maturity, LTV and interest rate. Shorter duration leaseholds in our sample are also no different in terms of trading liquidity (e.g. time on the market). Finally, incorporating discontinuities around the leasehold durations where mortgage lending might change does not affect our estimates of y^*. [26]

5.3 Estimating 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 $\underline{\Delta}_{it}$ for each bin and the dashed lines shows fitted estimates. The sample includes only 90-year extensions. Bars are shaded proportionally to the number of observations that make up the bar.

[25] In Figure A.12 we present a binscatter using properties that were extended for more than 700 years.

[26] In Appendix A.10, we investigate possible seasonality of our estimate of y^* as well. While spot UK house prices are highly seasonal (Ngai and Tenreyro, 2014), our estimate of y^* is not, consistent with its interpretation as a long run object.

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

There is a clear decline in y^* over our sample period. Figure 8 illustrates this by plotting the average gain in market value, $\underline{\Delta}_{it}$, for 90-year extensions against duration at extension, T, for three different time periods. The opacity of bars in the bar graph reflects the number of observations behind each bar. The gain in market value due to 90-year extension of a leasehold has increased by about 10pp between the beginning and end of our sample on average. This is a large increase, which reflects a significant fall in y^* over our sample period.[27]

We now move to the full data, with all extensions, so we can estimate the full dynamics of y_t^* with precision. In total, we have 126,018 leasehold extension natural experiments between 2000 and December 2023. Since there are fewer natural experiments in the beginning of our sample period, we pool the experiments together for the first 4 years and estimate a single y_t^* for 2000-2003. From then on, we estimate y_t^* separately for each year. Since lease registration became mandatory after 2003, we capture all lease extensions after this time.

Figure 1 in the introduction presents the estimates of y_t^* at annual frequency. The shaded region reflects 95% confidence interval. The estimates get more precise over time as the sample size gets larger.[28] y^* is relatively stable around 5.3% from 2000 until 2006, when it starts to fall persistently. In total y^* falls from around 5.3% to 2.8% in 2023, an almost 50% decline. The magnitude of this decline is large, corresponding to a doubling of the long run expected price-rent ratio. Notably, y^* estimate remains stable during the 2020 pandemic, despite considerable volatility in shorter term asset prices during this period.

5.4 Real Time Dynamics of y^*

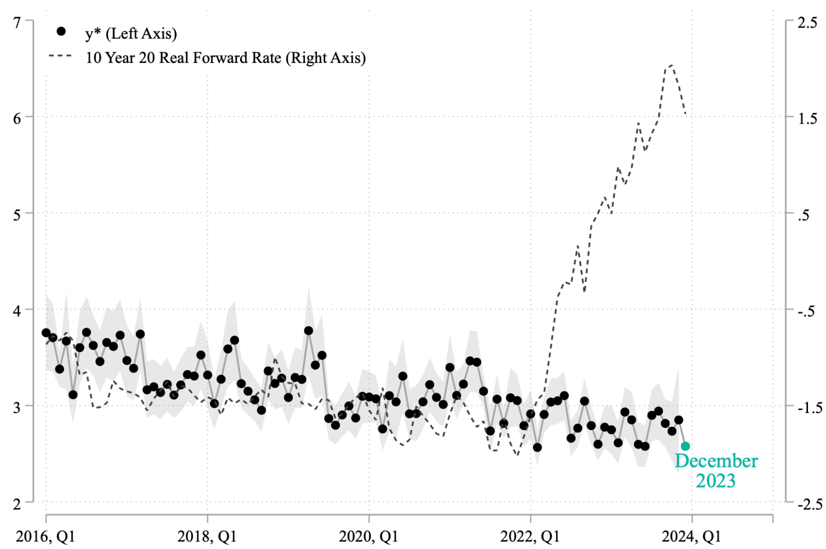
As mentioned earlier, one useful feature of our methodology is that it can be implemented in real-time, with publicly available data. Figure 9 plots our estimate of the expected long-term yield of housing, at monthly frequency until December 2023, the last month for which UK housing data are currently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

[27] In Figure A.15 we plot the event study of Figure 6 separately for each period, which also documents the lack of pre-trends at all times.

[28] Appendix Figure A.16 plots the time varying estimate y_t^*, separately for short and long extensions.

a reasonably large sample size of around 900 lease extensions per month.As discussed in Section 3.3, y_t^* estimates include new information with a lag of about 5 months.

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 December 2023. 95% confidence interval is shaded.

Two observations from the monthly time-series are worth mentioning. First, 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 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

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

The DiD estimator uses a weighted repeat sales index as in Case and Shiller (1989) to estimate price growth for control properties. An alternative would be to use the simple repeat sales index introduced by Bailey et al. (1963) (BMN). Appendix Figure A.19 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.20 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.21 and are similar to the baseline estimates.

5.6 The Advantage of Within-Property Variation in Duration

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. (2015) 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 another we allow price to vary quadratically with these same controls.[29] 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.[30] The other variations include all possible subsets of these seven characteristics.

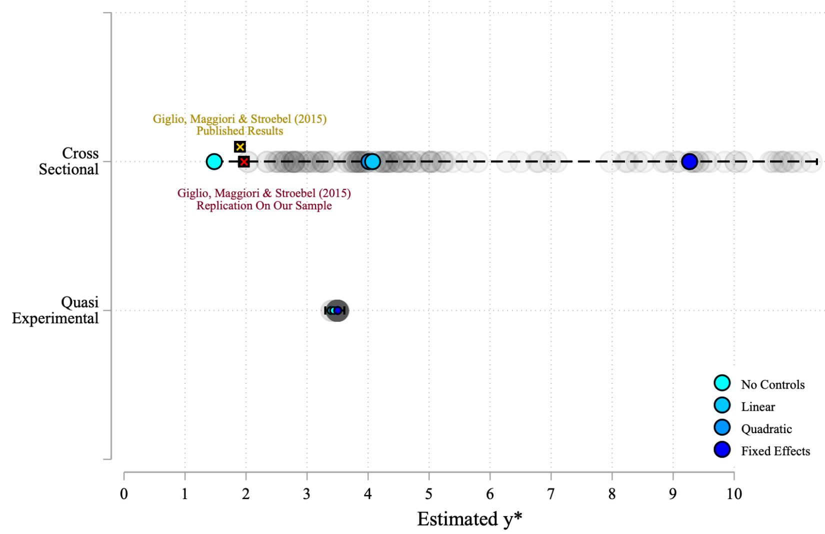
We then plot our estimates of y^* under each variation for both the quasi-experimental and cross-sectional methodologies in Figure 10. Under the cross-sectional approach, the estimates vary tremendously from 1.3% (in the case of no hedonic controls) to more than 10%. In contrast, our quasi-experimental 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.

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

[29] Figure A.17 indicates that most hedonic controls vary relatively linearly with log price.

[30] 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 2000-2023.

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.18, which again indicate that the effect of hedonic controls on the quasi-experimental design is minimal.

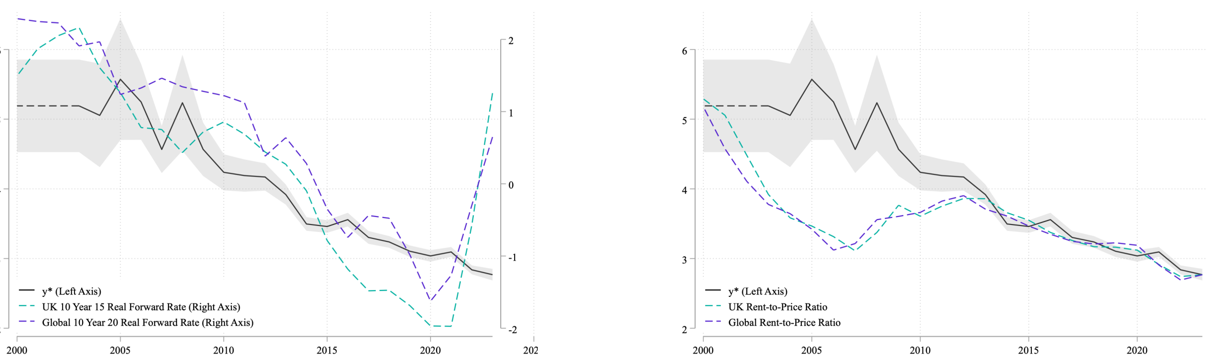
6 Macroeconomic Implications of y^*

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

6.1 The Secular Decline in Expected Long Term Yields

Our estimate of expected long-term yield comes from the UK housing market. Are these estimates useful for other asset classes? Obviously one cannot estimate expected long term forward yields on other assets since other assets do not have the unique natural experiment that we use. However, we can approximate the dynamics of expected long term yields for other assets at low frequency, using the long run trend of yields. The long run trend "averages out" short term shocks to yields, and therefore tracks low frequency movements in the long-term expected yield.[31] 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, at low frequency.

Figure 11: Comparing Other Asset Yields to y^*



(a) Real Risk-Free Forward Rate

(b) Rent-to-Price Ratio

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 and the weighted average rent-to-price ratio for 22 OECD countries for which it is available. The rent-to-price ratio is calculated from a rent-to-price index published by OECD, and normalized to the 2023 y^* level.

Figure 11 shows that the long run trend of various other yields has been similar to y^*, particularly prior to 2020. In panel (a) we plot long run safe asset yields for the UK and worldwide. In Panel (b) we plot rent-price ratios for the UK and a subset of OECD countries. The series is produced from rent-to-price index ratios published by the OECD, and normalized in 2023 to have the same level as y^*. The long run trend decline for these

[31] 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.

other asset yields is quantitatively similar to y^*, particularly before 2020. For instance, in Panel (a), UK and global real forward rates experience a decline of roughly 4pp between 2000 and 2022, compared to a roughly 3pp fall in y^* for UK housing over the same period. In Panel (b), peak-to-trough fall in rent to price ratios, either in the UK or globally, has been virtually identical to y^*. We tentatively conclude that the dynamics of long run expected yields are similar for housing and other assets. Many previous papers study long run trends in asset yields, and find declines before 2020 (e.g. Farhi and Gourio, 2018; Reis, 2022). However our measure of expected long term yields is also available at high frequency and in real time.

Therefore the decline in $y^* \equiv r^* + \zeta^* - g^*$ seems to be common across various assets, and not specific to housing. This finding suggests in turn that the fall in y^* was caused by a fall in r^*, the expected long term safe asset yield. The reason is that r^* is common to all assets, whereas movements in risk premia ζ^* and dividend growth g^* are specific to housing. Consistent with this logic, in Appendix A.11, we present suggestive evidence that the decline in y^* is due to a decline in r^*. In particular, we estimate long run housing risk premia and expected dividend growth using a standard Vector Autoregression approach, and show that neither long run UK housing risk premia, nor dividend growth for housing, can account for the trend decline of y^* for housing. Instead, the cause seems to be a fall in long run expected safe asset yields, r^*. Our finding is consistent with the large literature that has identified declines in r^* through different methods (Holston et al., 2017). We stress that this exercise is tentative due to the uncertainty of the VAR based procedure.

6.2 Comparing Current Yields to y^*

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

6.2.1 Dynamics of r^* after 2020

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

(Blanchard, 2023).

The stability of y^* after 2020 suggests that r^* has also been stable, despite the rise in medium term yields. We stress that $y^* \equiv r^* + \zeta^* - g^*$ and r^* are different objects. However, y^* contains a great degree of information about r^*, since a large rise in r^* would likely lead to a large rise in y^*.[32]

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

6.2.2 Dynamics of House Prices During the Early 2000s

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

The divergence during 2000-06 suggests that the housing boom during this period was driven by a decline in the short-end of the housing forward curve, while the long-end yield remained high. Since risk-free yields remained relatively largely stable during this period as well, once could conclude that the divergence between y^* and R/P is driven by a medium-term shift in expectations about the housing risk premium or housing dividend growth, which might prove to be transient. In contrast, the decline in rent-to-price ratios after 2012 tracks the decline in y^* very closely, suggesting that the recent increase in house prices has been driven by a shift in the long-end, and hence might be more persistent.

Through this perspective, the high housing valuations of 2000-6 and 2015 onward are quite different. In 2000-6, valuations were temporarily higher than their long run values, and likely to fall; whereas current housing valuations seem to be more persistent and less likely to revert. Going forward, our estimates can be used to understand whether movements in house prices are driven by transitory or more persistent factors.

[32] 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.

6.3 Cross-Sectional Variation in y^*

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

6.3.1 Δy^* and supply elasticity

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

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

Consistent with this logic, we find that y^* falls by more in areas with more inelastic housing supply. We define areas as Local Authorities (LAs) and use two measures of supply constraints from Hilber and Vermeulen (2016). The first is the share of major construction applications refused by that Local Authority, averaged over a nearly 30-year period. The second is the share of developable land developed in 1990.

One problem is that the refusal rate of construction applications is endogenous; developers are less likely to submit construction applications in more restrictive LAs that are unlikely to accept them, which in turn reduces the refusal rate of more restrictive LAs. For example, as we discuss in Appendix A.12, the application refusal rate in places like Central London is surprisingly low, most likely because construction firms know that applications will be refused and do not submit them. This implies that the measured application refusal rate is an underestimate of what the refusal rate would be if the application submission rate was randomly assigned.

To address this endogeneity concern, Hilber and Vermeulen (2016) develop an instrument based on a 2002 policy reform, which incentivized LAs to make decisions on major construction project applications in a timely fashion. This policy should have especially lowered the application delay rate for inelastic LAs, which were more likely to delay construction applications before the reform. Column (5) of Table 5 presents the first stage for this instrument, indicating that LAs which lowered their application delay rate after 2002 on average had higher refusal rates.[33] In Appendix A.12 we elaborate further on the validity of this instrument.

Table 5: Cross-Sectional Heterogeneity in Δy^*

	Δy^*				Refusal Rate	Δy^*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	OLS	OLS	OLS	1st Stage	IV	IV
Refusal Rate	-4.30**	-6.08***	-5.54***			-10.32**	-11.02**
	(1.28)	(1.30)	(1.55)			(3.84)	(3.47)
Share Developed		-1.03*	-0.69				-1.69**
		(0.39)	(0.50)				(0.60)
Change in Delay Rate				2.35***	-0.20***		
				(0.67)	(0.05)		
Region FE			✓				
N	119	119	119	119	119	119	119
R2	0.13	0.20	0.27	0.12	0.09	-0.12	0.07

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). 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.

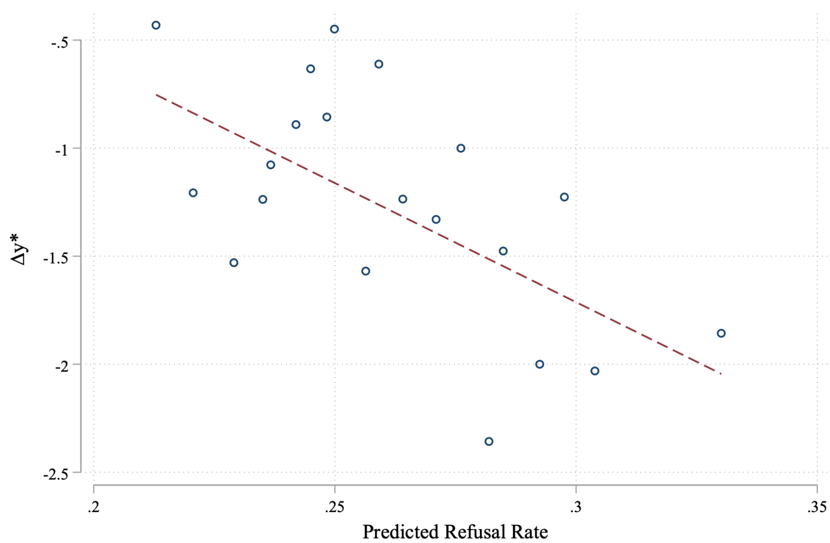
Table 5 shows that Local Authorities with more inelastic housing supply have greater declines in y^*. We measure the decline, Δy^*, as the long-difference between the pre-2009 period to the post-2022 period, although our results are robust to choosing other cutoff years. Regressions are weighted by the inverse of the sum of the variances of the pre and post y^* estimates.[34] 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 fixed effects for the 9 regions of the England. In Column (4), we present a reduced-form regression of Δy^* against the delay rate instrument from Hilber and Vermeulen (2016). In Column (5) we instrument for the refusal rate using the delay rate instrument and in Column (6) we repeat this regression, but also include the share developed as a co-variate. After instrumenting for the refusal rate, the share developed control has a smaller effect on the refusal rate coefficient, consistent with the refusal rate measure being endogenous. In all cases, the results indicate

[33] Appendix Figure A.22 presents a binscatter of the first stage.

[34] In all the regressions, we utilize the variation of $\underline{\Delta}_{it}$ built using an annual repeat sales index (RSI), as opposed to a quarterly repeat sales index. This is because we can match the annual RSI to more experiments, which therefore yields a larger sample size. Figure A.19 verifies that estimates using the annual RSI are very similar to the quarterly RSI.

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^* indicates 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.

6.3.2 y^* and The Housing Risk Premium

In this section we show that the housing risk premium is priced into y^*. This suggests that region-specific housing risk is expected to persist at very long horizons.

We define local housing risk as the correlation between local house prices and consumption. More risky Local Authorities are those for which consumption growth and housing gains are very correlated, and less risky Local Authorities are those which enable consumption smoothing. We proxy consumption with Local Authority-level earnings data for full-time male workers, which spans the period of 1974-2021.[35] We then estimate,

$$\log P_{it} = \alpha + \beta_i \times \log E_{it} + \gamma_i + \delta_t + \epsilon_{it} \tag{8}$$

where P_{it} is a house price index for each Local Authority, i, in each year, t. E_{it} is average weekly earnings for full-time male workers, γ_i are LA fixed effects and δ_t are year fixed effects. The estimated local housing betas, β_i, represent the sensitivity of house prices to changes in earnings for each Local Authority i.[36]

Table 6: Cross-Sectional Heterogeneity in Risk

	Housing β	y^*		Δy^*	
	(1)	(2)	(3)	(4)	(5)
Housing β		1.65***	1.69***	0.32	0.24
		(0.27)	(0.27)	(0.90)	(0.95)
Predicted Refusal Rate	1.06**		-1.18		-11.75***
	(0.36)		(1.05)		(3.42)
N	119	119	119	119	119
R2	0.07	0.28	0.28	0.00	0.12

Standard errors in parentheses

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

The table presents cross-sectional regressions of the local housing beta, β_i. Column (1) presents estimates of a regression of the instrumented refusal rate of major construction applications on β_i. The refusal rate is instrumented with the delay rate instrument discussed in Section 6.3. Columns (2) and (3) present estimates of β_i on the average y^*, with and without controlling for the instrumented refusal rate. Columns (3) and (4) present estimates of β_i on Δy^*, with and without controlling for the instrumented refusal rate. Columns (2) and (3) are weighted by the inverse variance of the estimated y^* and columns (4) and (5) are weighted by the inverse of the sum of the variances of y^* for the pre and post period. We restrict to the sample of 119 LAs for which we can estimate Δy^*.

We find that our measure of housing risk premia is a strong determinant of the level of r^*. Column (1) of Table 6 shows the results of a regression of the local housing betas against the instrumented refusal rate from Section 6.3. More inelastic areas have a higher local housing beta.[37] Next, we regress the local housing betas against the estimated average level of y^* for each LA. Columns (2) and (3) of Table 6 show that riskier areas have higher y^*, even when

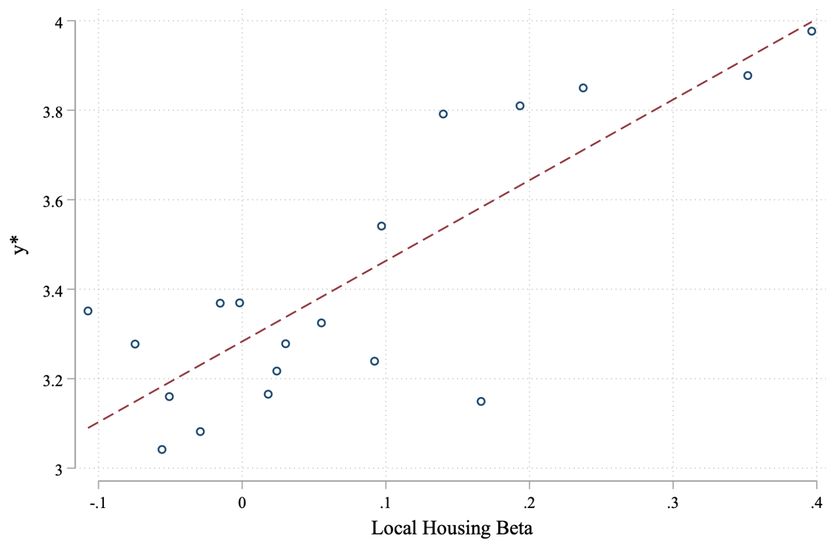
[35]Earnings for 1974-1996 is from the New Earnings Survey (NES) and is obtained from the Hilber and Vermeulen (2016) replication package. Earnings data for 1997 to the present is from the Annual Survey of Hours and Earnings (ASHE), which replaced the NES.

[36]The LA-level house price index is based on data from Hilber and Vermeulen (2016), originally from the Council of Mortgage Lenders (1974-1995) and the HM Land Registry (1995-2008). We expand the index to the present using the same methodology and data from the HM Land Registry.

[37]This finding is consistent with research showing that housing supply elasticity mitigates house price growth during boom periods (e.g. Hilber and Vermeulen, 2016).

controlling for local housing supply elasticity. The level of risk, however, does not explain the decline in y^*, as indicated by Columns (4) and (5). Figure 13 plots a binscatter of average y_i^* against the local housing betas, indicating again that riskier LAs had significantly higher average long-run yields.

Figure 13: Areas With Higher Local Housing Beta Have Higher y^*



Our measure of housing risk premia is calculated at annual frequency whereas y^* is identified from the long term. Therefore our results suggest that there is a substantial risk premium associated with housing even in the long run.

7 Concluding Remarks

This paper estimates the expected long-term yield of housing and its dynamics from 2000 to present for the UK property market. We exploit a natural experiment — extensions of long duration property leases in the British property market. Our findings show that y^* fell from 5.3% before the Great Recession to 2.8% in 2023. An important goal of this paper was to assemble an administrative data set on lease extensions and make it publicly available for free. Our data will be updated in real time going forward. We hope that both policy makers and academics will find the data and empirical methodology useful to estimate the market's expectation of long-term yield in real time.

Long-run yields are valuable because they "look through" the short term factors affecting asset prices in real time. Our results show that there has been a growing gap between long-run yields and real forward yields 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.

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.

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.

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 : journal of the Econometric Society*, May 2016, *84* (3), 1047–1091.

Goetzmann, William N. and Matthew Spiegel, "Non-Temporal Components of Residential Real Estate Appreciation," *The Review of Economics and Statistics*, 1995, *77* (1), 199–206.

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.

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.

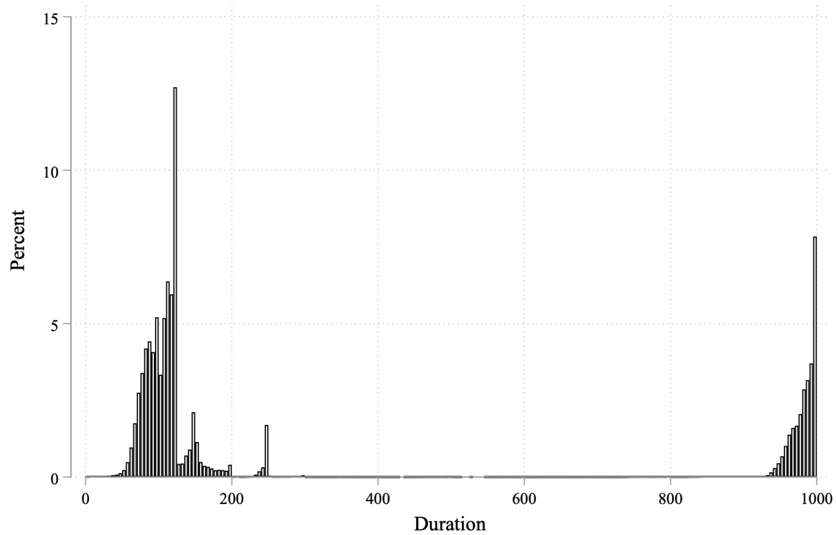
Vissing-Jorgensen, Annette, ""Has Monetary Policy Cared Too Much About a Poor Measure of r* ?" Discussion," May 2022.

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

A Appendix

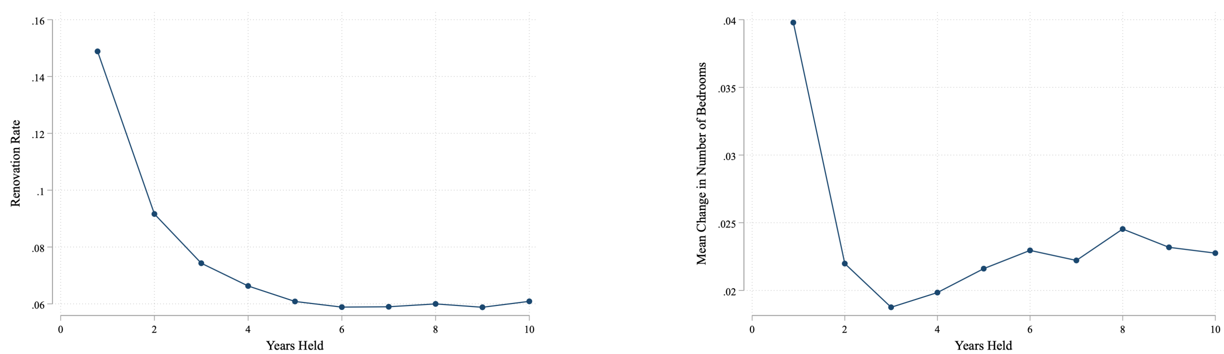
A.1 Additional Figures

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.

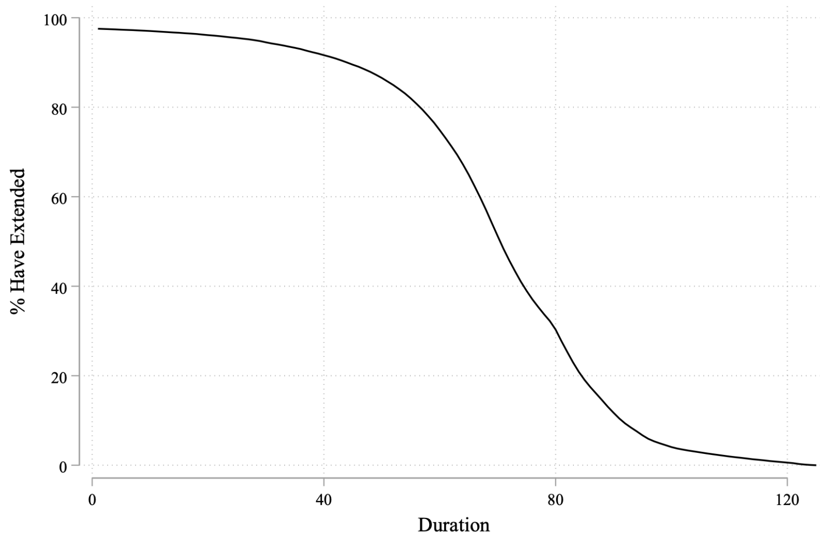
Figure A.2: Renovations By Holding Period



(a) Renovation Rate **(b)** Change in Bedrooms

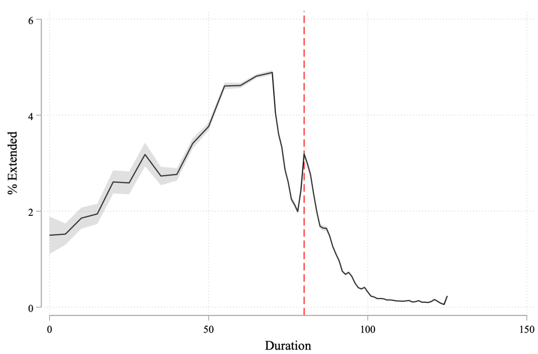
The figure shows the renovation rate and change in number of bedrooms reported relative to property holding period. The renovation rate is measured as the share of properties which have Rightmove listings that mention renovation. We see that very short holds have a disproportionately high renovation rate according to both measures. 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.4: Cumulative Hazard Rate

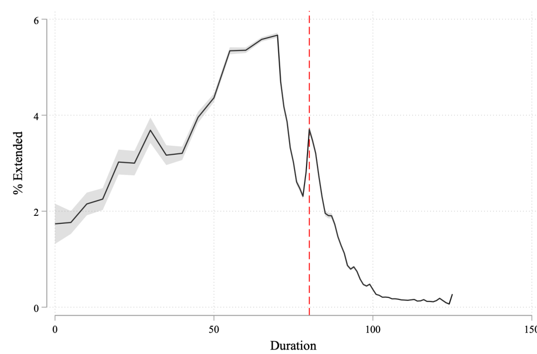


The figure shows the cumulative probability of extending over a property's lifetime. The sample includes all leases with at least one transaction and covers the 2003-2020 period. We exclude the pandemic period due to abnormally low extension rates.

Figure A.3: Hazard Rate of Lease Extension



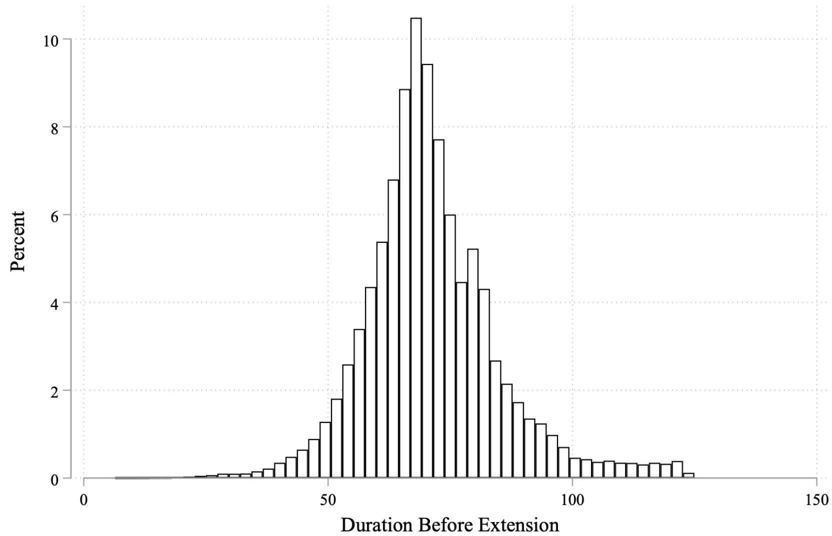
(a) Raw



(b) Corrected

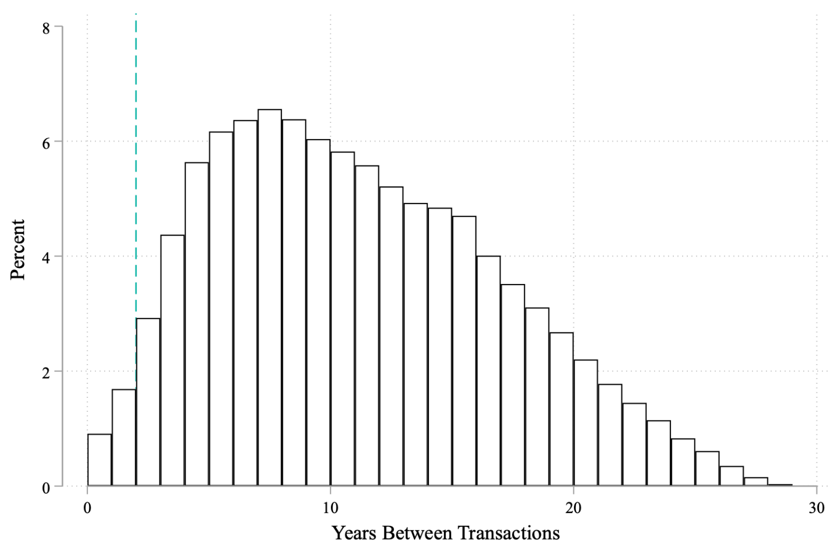
The figure shows the conditional probability of extension, $\theta(T)$ given that a property has duration T. In the first panel, the conditional probability of extension is given by $\theta_1(T) = \frac{N_T^{Ext}}{N_T}$ where N_T^{Ext} is the number of properties which extended with duration T and N_T 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 $\gamma = 1.16$ 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.5: Histogram of Duration Before Extension



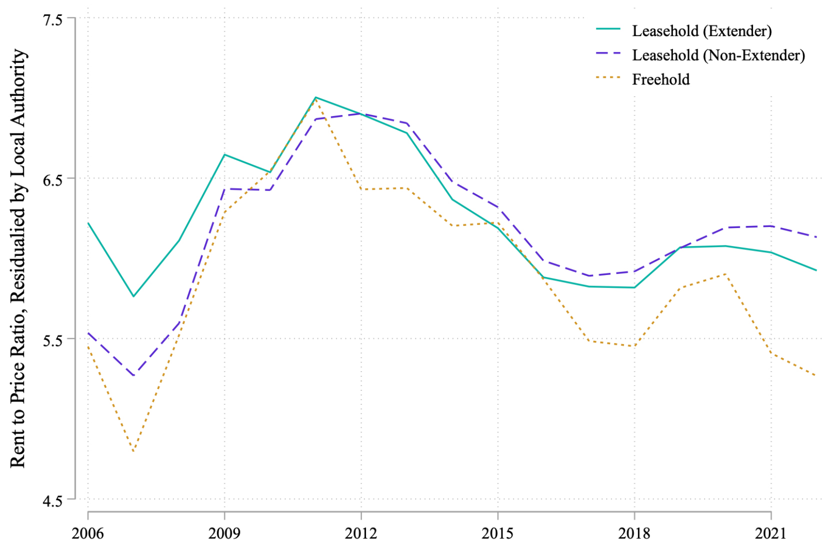
The figure presents a histogram of lease duration immediately before extension. The sample is extended flats.

Figure A.6: Histogram of Years Between Transaction



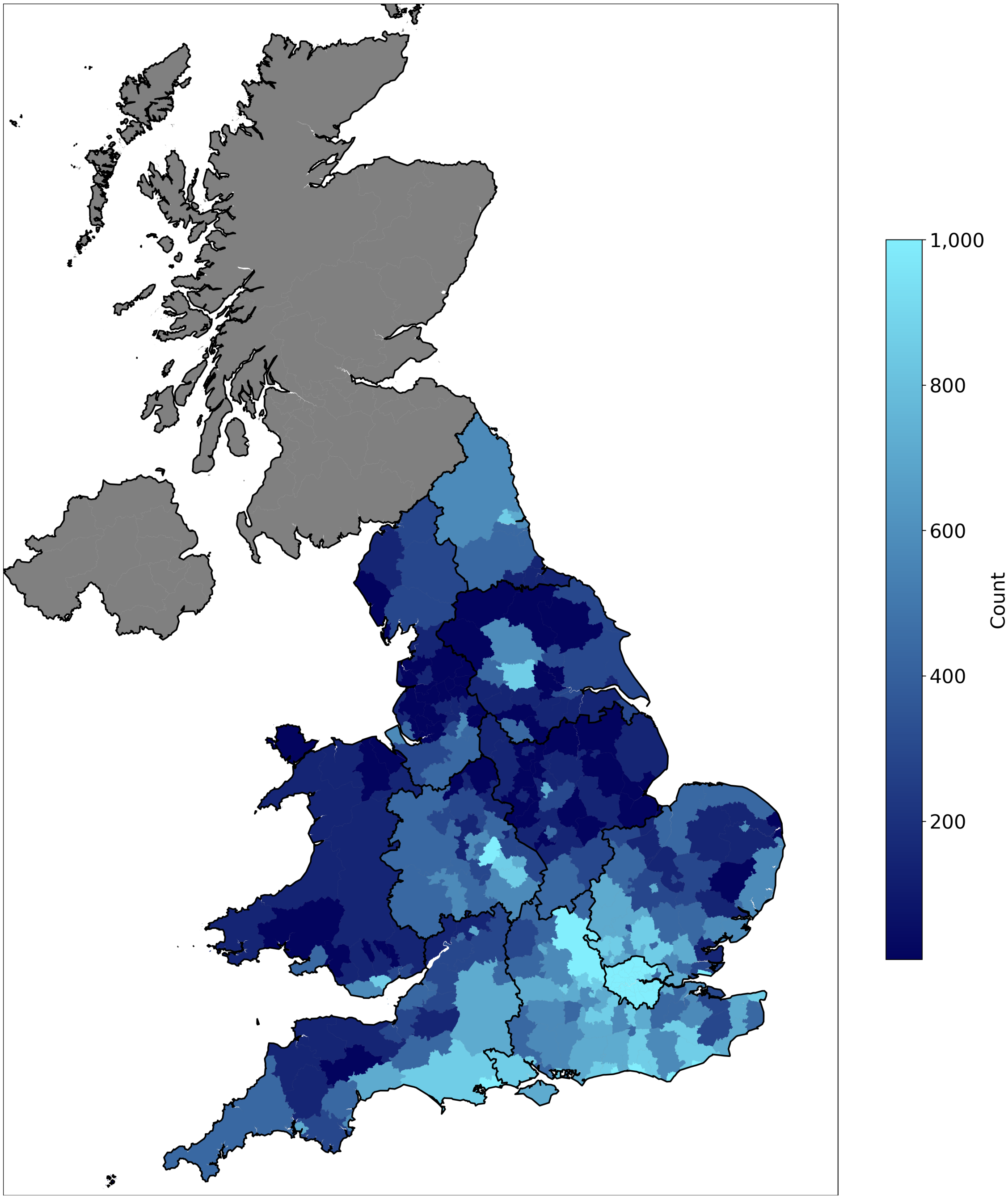
The figure shows a histogram of the holding period, h, for lease extensions which have a recorded transaction before and after extension. The dotted line shows the $h = 2$ cutoff; properties below the cutoff are not included in our primary sample. The sample is all extended flats.

Figure A.7: Rent to Price (Freeholds vs Leaseholds)



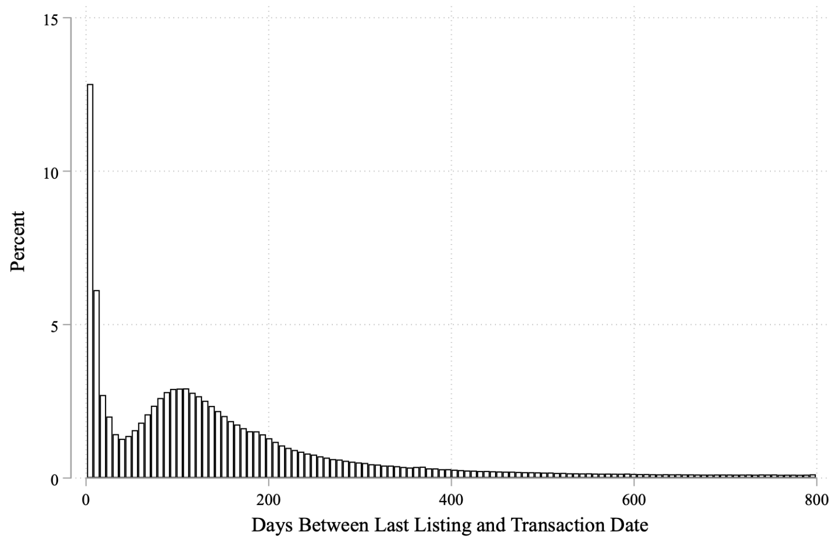
The figure shows the property-level price to rent ratio for leasehold and freehold flats. Leaseholds are subdivided into those which extend during our sample and those which do not. Property-level rental price data is collected from Rightmove and Zoopla.

Figure A.8: Heat Map of Extension Rate



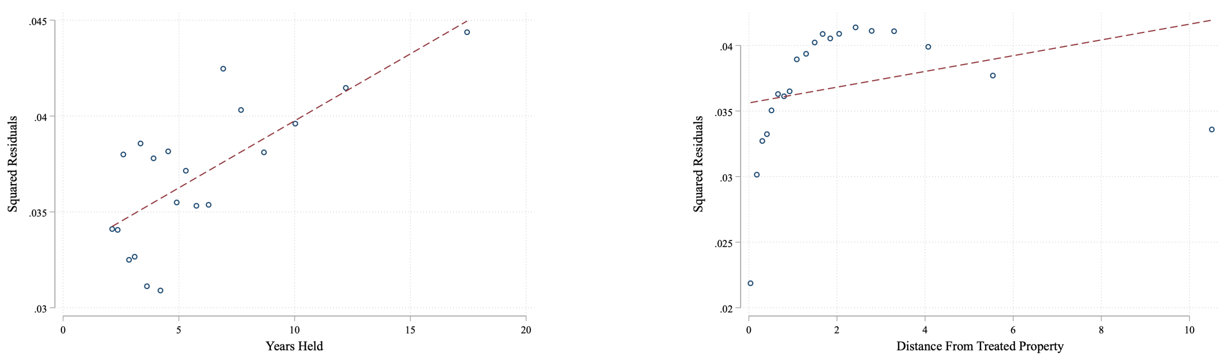
The figure shows a heatmap of the number of properties extended in each Local Authority in England and Wales.

Figure A.9: 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.10: Binscatter of Repeat-Sales Residuals

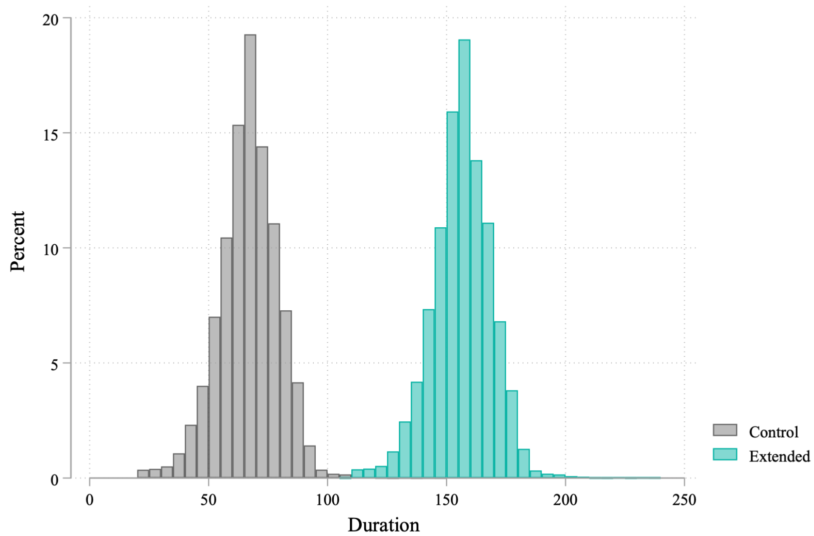


(a) Holding Period

(b) Haversine Distance

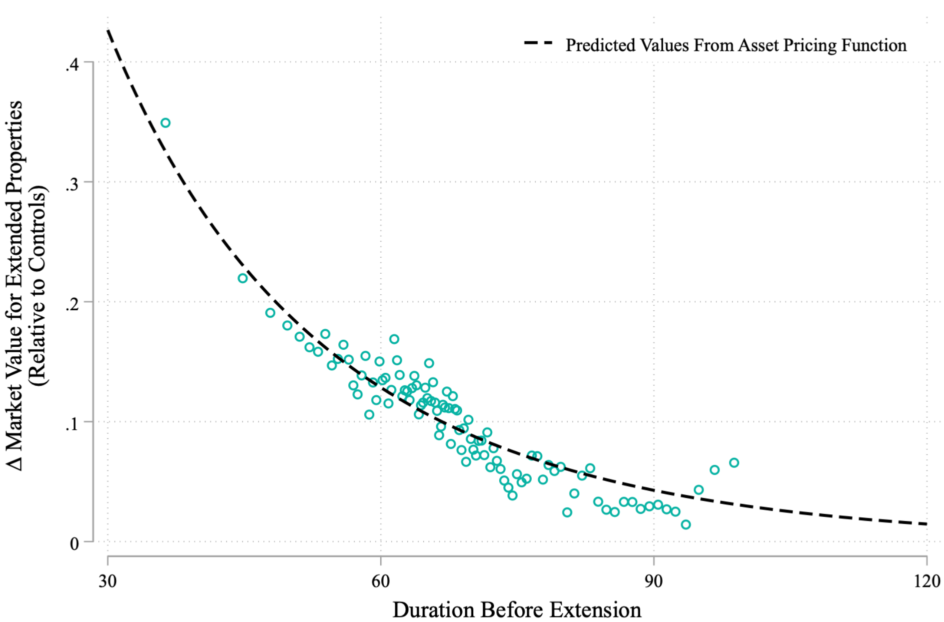
The figures are binscatters of the squared residuals from the first step of the repeat-sales methodology against two variables which may be correlated with the variance of the estimate: (1) the time between transactions and (2) the distance between the treated and control property.

Figure A.11: Histogram of Remaining Lease Term At Sale



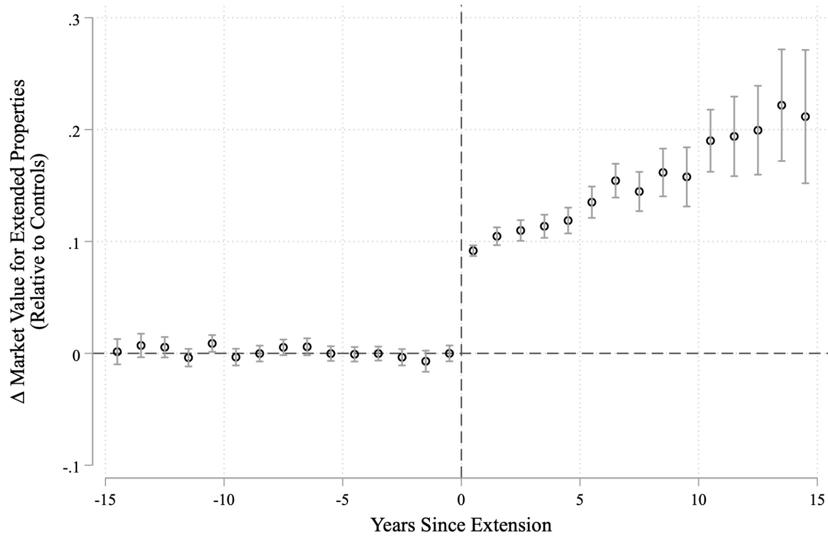
The figure shows a histogram of remaining duration at sale time t for the approximately 40 thousand experiments which were extended for 90 years. The blue-shaded histogram includes the sale duration for the controls in the experiment and the grey-shaded histogram includes the sale duration for the extensions.

Figure A.12: Duration vs. Price Gain After Extension, Leases with 700+ Years



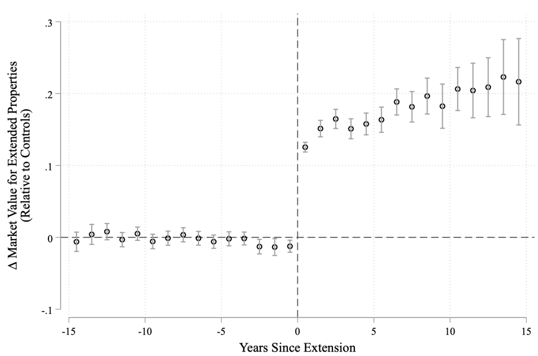
The figure is a binscatter of our difference-in-difference estimator against duration before extension, T, with 100 bins. The sample includes leases that were extended for more than 700 years. The black line shows fitted estimates of equation (6).

Figure A.13: Event Study Representation of Lease Extension, 90-Year Extensions

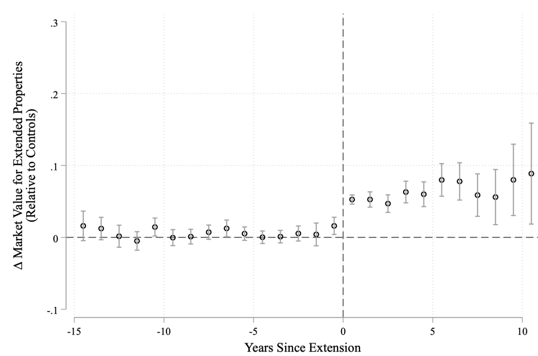


The figure replicates Figure 6 for properties extended by 90 years.

Figure A.14: Event Study By Duration



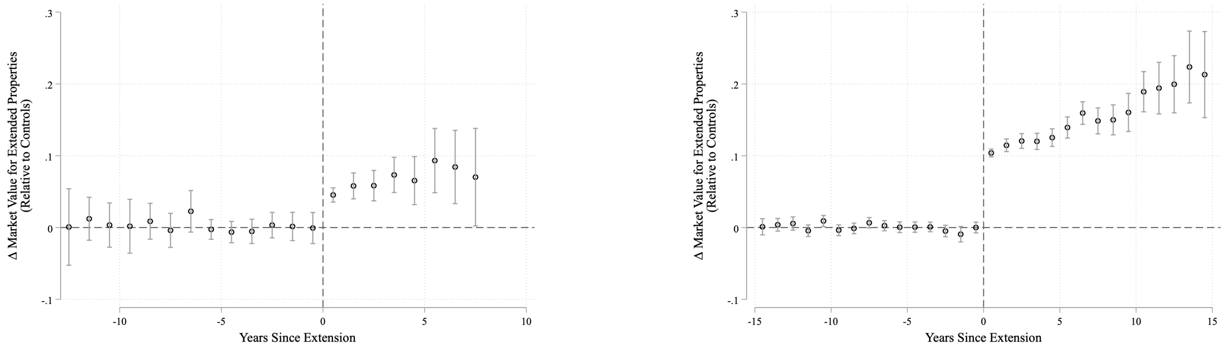
(a) $T < 70$



(b) $T \geq 70$

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

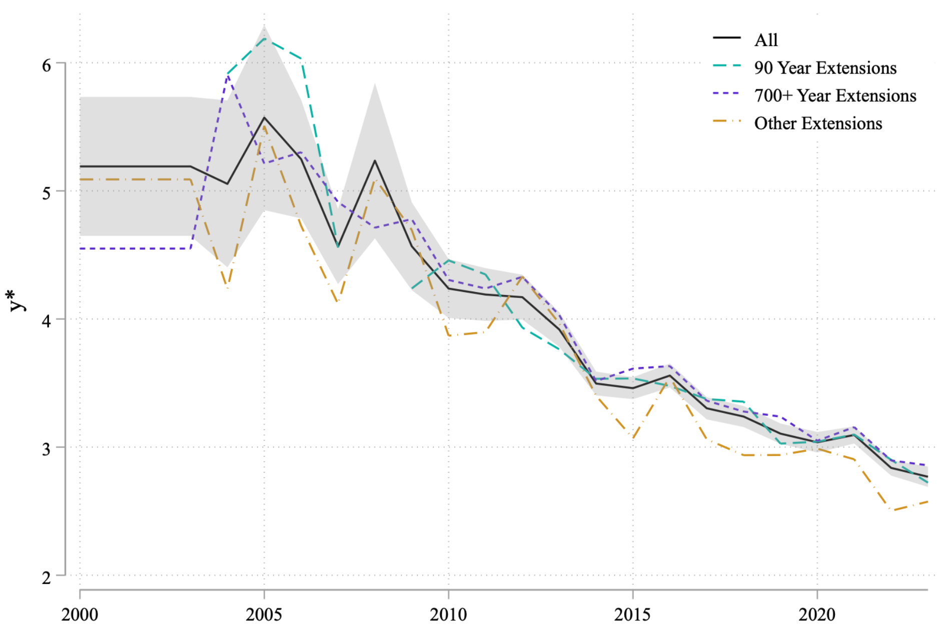
Figure A.15: Event Study By Time Period



(a) Pre 2012 **(b)** Post 2012

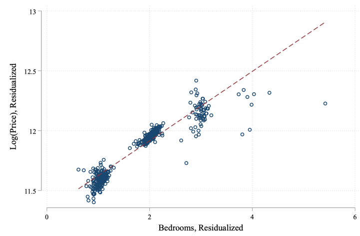
The figure replicates Figure 6 for two different sub-periods: pre 2012 (inclusive) and post 2012.

Figure A.16: y_t^* Estimates For Various Extension Amounts

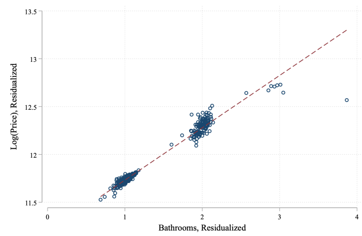


The figure presents estimates of y_t^* for every year of our sample period for lease extensions that were extended for 90 years, for more than 700 years, or for another amount, separately. The estimate of y^* for 90-year extensions in 2008 is excluded, because the mean estimated difference-in-difference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

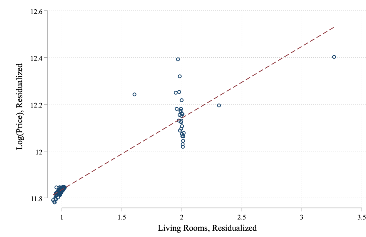
Figure A.17: Binscatter Log(Price) on Hedonics



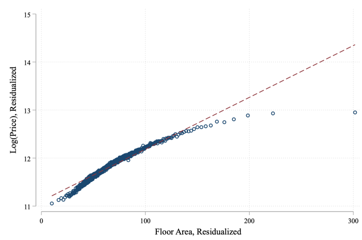
(a) Bedrooms



(b) Bathrooms

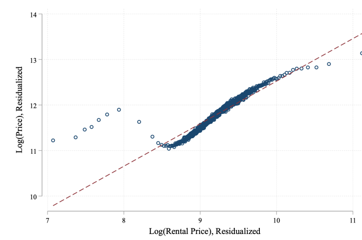


(c) Living Rooms



(d) Floor Area

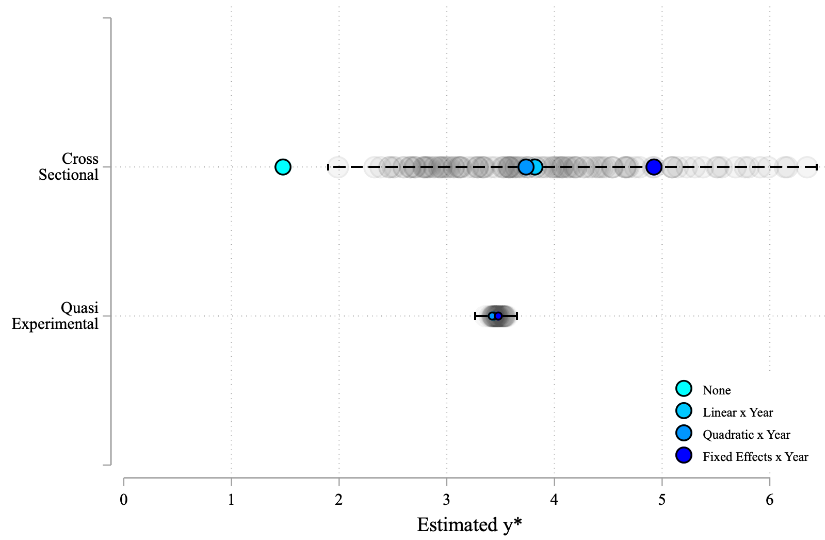
(e) Year Built



(f) Log(Rent)

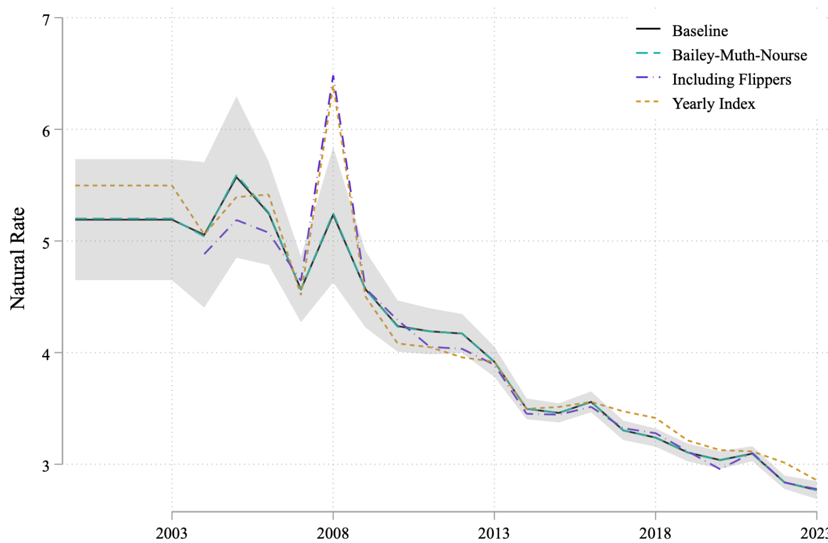
The figures are binscatters of log transaction price against the following hedonic characteristics: number of bedrooms, number of bathrooms, number of living rooms, floor area (sq. meters), year that the property was built, and log yearly rental price. Both the x and y-axis variables are residualized 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.18: Stability of y^*, Controlling for Observables Interacted with Time



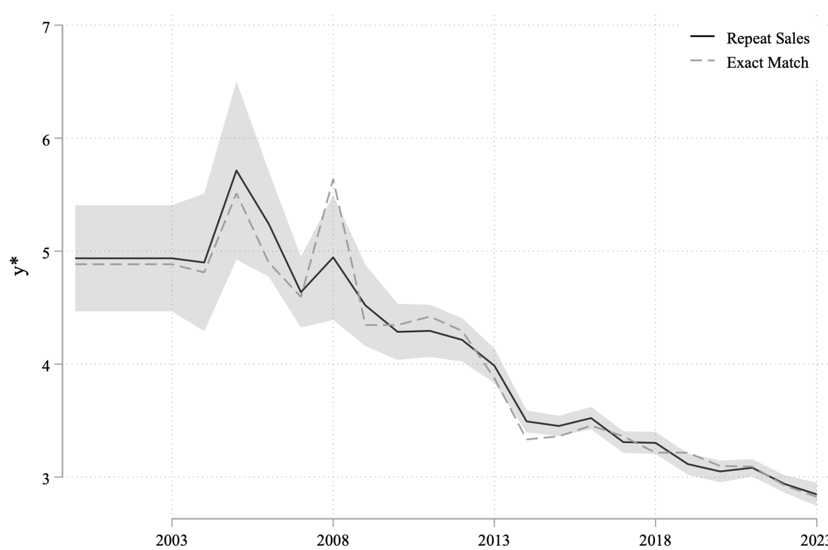
The figure replications Figure 10 controlling for hedonic interacted with time fixed effects.

Figure A.19: Robustness Checks



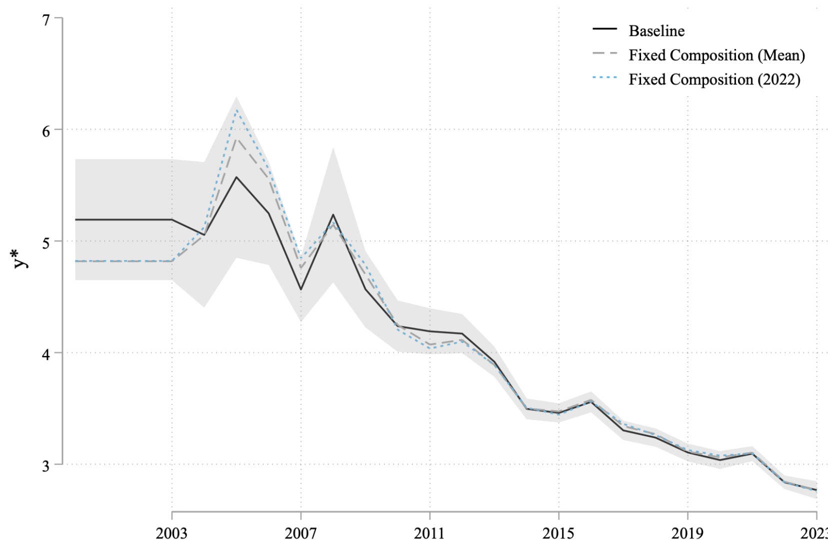
The figure presents robustness for our estimates of y_t^*. The solid black line presents our baseline estimates from Figure 1. The dashed grey line presents estimates when using the methodology from Bailey et al. (1963) to produce the repeat sales index of control properties. The dash-dot blue line presents 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.20: Restrictive Controls



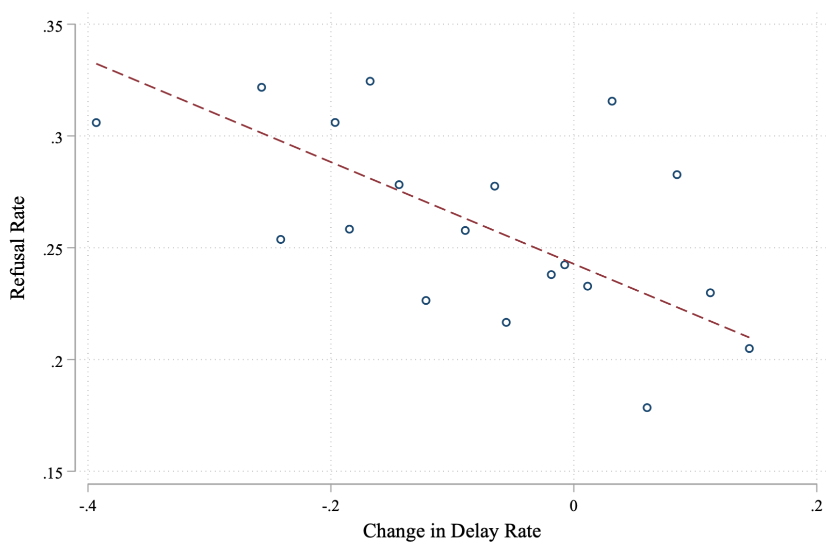
The figure presents estimates of y^* when we pick controls that transact in the same purchase and sale years as the treated property. Our difference-in-difference then becomes $\underline{\Delta}_{it} = (\log P_{it}^{T+90} - \log P_{i,t-h}^{T+h}) - (\log P_{jt}^{T} - \log P_{j,t-h}^{T+h})$, where i is 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.21: 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.22: 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.

A.2 Additional Tables

Table A.1: Total Number of Extensions

Extension Amount	90	700+	Other	Total
2000-2005	4,036	14,141	10,396	28,573
2006-2010	13,761	26,115	13,398	53,274
2011-2015	32,648	38,695	19,662	91,005
2016-2020	45,867	48,774	23,155	117,796
2021-2023	9,655	11,114	3,590	24,359
Total	105,967	138,839	70,201	315,007

The table reports the number of extended leases that have transaction data before extension time and may or may not have transaction data after extension time. The first column includes 90 year extensions, the next column includes 700+ year extensions, and the third column includes others, which are almost all non-90 under 200 year extensions. The first column denotes the time of extension.

Table A.2: 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.3: Estimated y^* for Extensions of 90 Years vs More Than 700 Years

	(1)	(2)	(3)	(4)	(5)	(6)
y^*	3.51***	3.56***	3.45***	3.54***	3.42***	3.49***
	(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	41,713	52,348	29,248	31,304	29,334	31,424

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, $\underline{\Delta}_{it}$, 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 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.4 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 1.16. 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.[38]

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

$$P_{i,t}^{T} = \frac{R_{i,t+1} + \theta_i(T)(P_{i,t+1}^{T+90-1} - \kappa_{t+1}^{T-1}) + (1-\theta_i(T))P_{t+1}^{T-1}}{1 + r^* + \zeta^*} \tag{9}$$

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.5 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,

[38] To obtain this estimate, we utilize the fact that the most common lease extension amount is for 90 years and the most common initial lease terms are for 99 and 125 years. Therefore, almost all leases with recorded terms of 189(=99+90) and 215(=125+90) have almost certainly been extended. As such, we infer that the share of extensions missing pre-transaction data is equal to the share of 189/215 leases missing pre-transaction data.

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.5.1 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

$$RV_t^T = \frac{R_{it}}{r_{RV}} \left(e^{-r_{RV}T} - e^{-r_{RV}(T+90)} \right), \tag{10}$$

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

$$MV_{it}^T = \frac{R_{it}}{y_t^*} \left(e^{-y_t^* T} - e^{-y_t^*(T+90)} \right). \tag{11}$$

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 (10) and equation (11) 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

$$\kappa_{it}^T = \begin{cases} \min\left\{RV_{it}^T + \gamma R_{it}, MV_{it}^T\right\} & T \geq 80 \\ \min\left\{\frac{RV_{it}^T + MV_{it}^T}{2} + \gamma R_{it}, MV_{it}^T\right\} & T < 80 \end{cases} \tag{12}$$

equation (12) 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.[39]

[39]For exposition, we assume that the tribunal, the leaseholder and the freeholder all agree on the market

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

$$P_{it}^T = P_{it}^{T+90} - \kappa_{it}^T. \tag{13}$$

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.5.2 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* $\lim_{T\to 80^-} P_{it}^T < \lim_{T\to 80^+} P_{it}^T$.

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

value of the extension. Our qualitative and quantitative results are not affected by this assumption.

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,

$$\frac{\Delta_h \log P_{it}^T}{h} = \alpha + \beta \cdot \mathbb{1}\text{Crossed } 80 + \Gamma_{i,t,t-h} \tag{14}$$

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.4 reports the estimated coefficient from equation (14). The first column presents estimates before 2010. There is no statistically significant discontinuity in price at 80. In fact, Appendix Figure A.23 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.4 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.23. 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.4: Test for Discontinuity at 80

	(1)	(2)
Crossed Cutoff	-0.04	-0.84***
	(0.09)	(0.11)
Sale Year x Purchase Year x LA FE	✓	✓
Period	Pre 2010	Post 2010
N	54,685	9,965

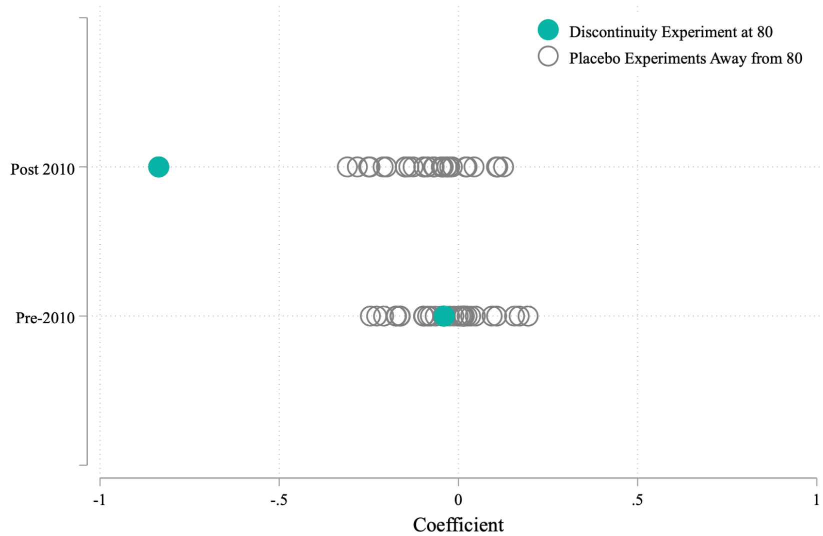
Standard errors in parentheses

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

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.23: Test for Discontinuity at 80 Years (Placebos)



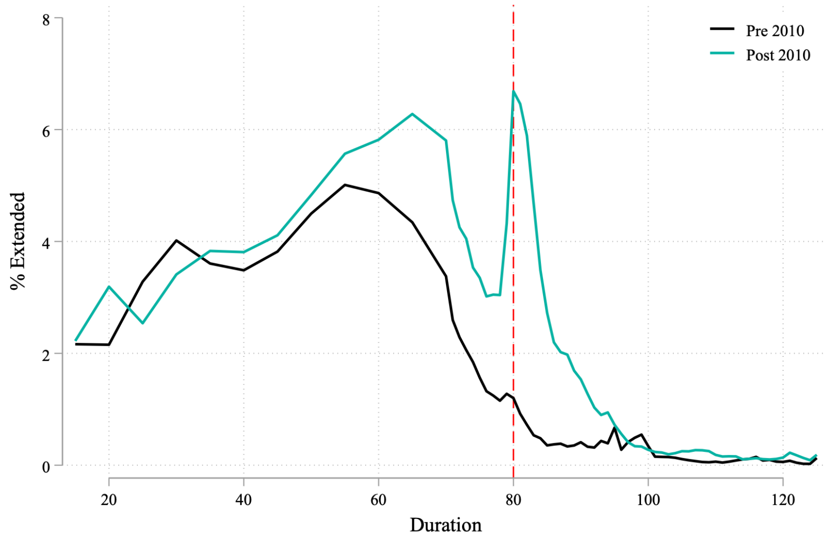
The figure presents the estimates from Table A.4 in red. In grey, we run 30 placebo tests using alternative cutoffs between 70-100. In each case, the sample includes properties within 10 years above and below each cutoff.

Figure A.24 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 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.5.3 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

Figure A.24: 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.

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.13 shows that the difference-in-differences estimator of the price gain upon lease extension becomes

$$\underline{\Delta}_{it}^T = \log\left(1 - e^{-y_t^*(T_{it}+90)}\right) - \log\Big(\left(1 - e^{-y_t^* T}\right) + \left[\Pi_{Tt}^H\left(1-\alpha_t^H\right) + \Pi_{Tt}^L\left(1-\alpha_t^L\right)\right] e^{-y_t^* T_{it}}\left(1 - e^{-y_t^* 90}\right)\Big) \quad (15)$$

For simplicity, this derivation makes the additional assumption that leases extend only once.[40] 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

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

the observed extension hazard rate, as shown in Figure A.24. Equation (15) is the same as the baseline estimator equation (5) 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 (15), as an input into a non-linear least squares estimation similar to the baseline procedure. The two additional parameters, α_t^H and α_t^L are estimated from the covariance between T_{it} and $\underline{\Delta}_{it}^T$. 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.5. 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.

Table A.5: Estimating Alpha

	(1)	(2)
y^*	4.81***	3.25***
	(0.09)	(0.01)
α_t^H	1.00***	0.53***
	(0.01)	(0.06)
α_t^L	1.00***	1.00***
	(0.00)	(0.02)
Period	Pre 2010	Post 2010
N	18,064	106,478

Standard errors in parentheses

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

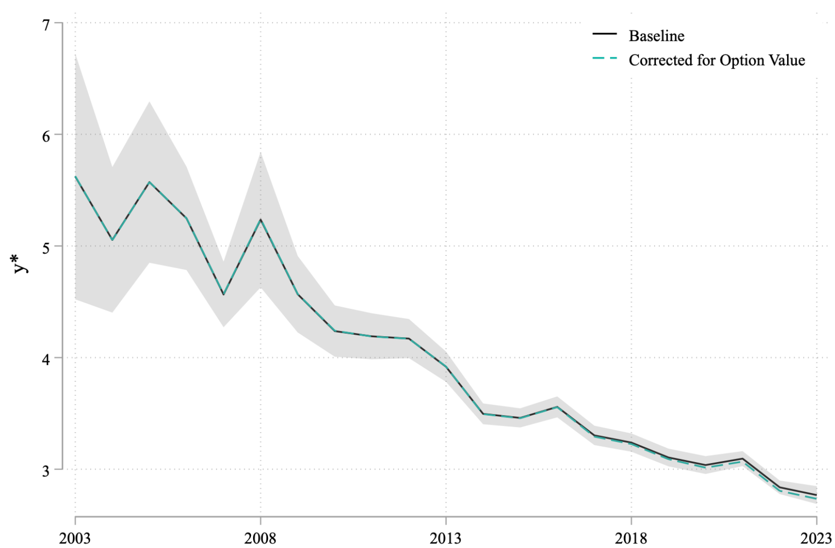
The table presents estimates of α_t^H and α_t^L, estimated jointly with y^* from a nonlinear least squares estimate of equation (15). Estimates of α_t^L and α_t^H are constrained to lie within [0,1]. Standard errors are bootstrapped.

Our difference-in-difference estimates of option value in this subsection are consistent 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 0.47.[41]

A.5.4 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.25 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.

Figure A.25: 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.

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.24 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.

A.5.5 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,

[41]The details of this calculation are presented in Appendix A.5.5.

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.4. From the preceding subsection, the price of a property i with duration T is

$$P_{it}^T = \frac{R_{it}}{y^*} \left(1 - e^{-y^*T} + \left[\Pi_{Tt}^H \left(1 - \alpha_t^H\right) + \Pi^L \left(1 - \alpha_t^L\right)\right] e^{-y_t^* T_{it}} \left(1 - e^{-y_t^* 90}\right)\right) \tag{16}$$

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 (14), by setting $\alpha_t^L = 1$, which is what we estimate in Table A.5, and y^* to its mean for that period. We obtain an estimate of $\alpha_t^H = 0.53$ in the post-2010 period.

A.5.6 The 2006 Sportelli Case

In Appendix A.5, 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.5, 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.5, is informative as to the difference between the market value of lease extension and the court dictated outside option. First, in Figure A.26, 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.6 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.26 and Table A.6, we include the post-2010 period for reference, which does experience significant bunching and a discontinuity in price at $T = 80$.

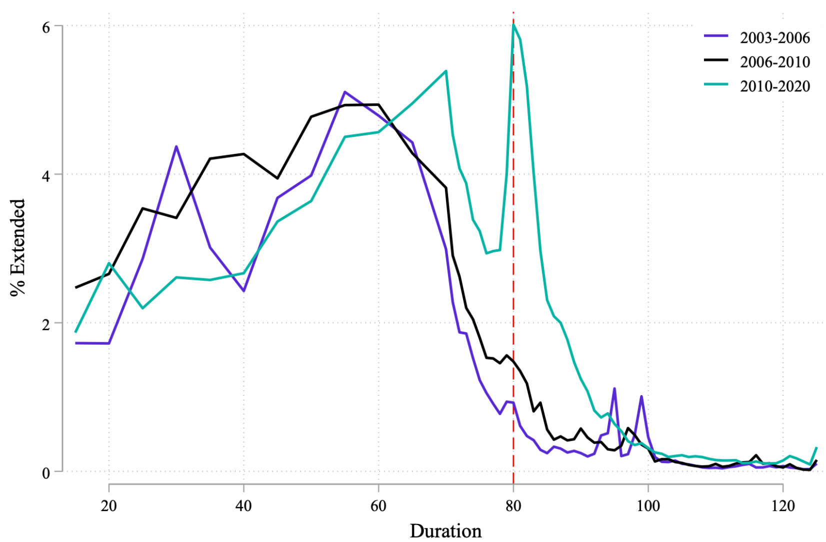


Figure A.26: Hazard Rate of Extension, Three Periods

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.6: Test for Discontinuity at 80, Three Periods

	(1)	(2)	(3)
Crossed Cutoff	-0.09	0.19	-0.84***
	(0.09)	(0.12)	(0.11)
Sale Year x Purchase Year x LA FE	✓	✓	✓
Period	2003-2006	2006-2010	2010-2023
N	41,370	13,315	9,965

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 (10) and equation (12), 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.

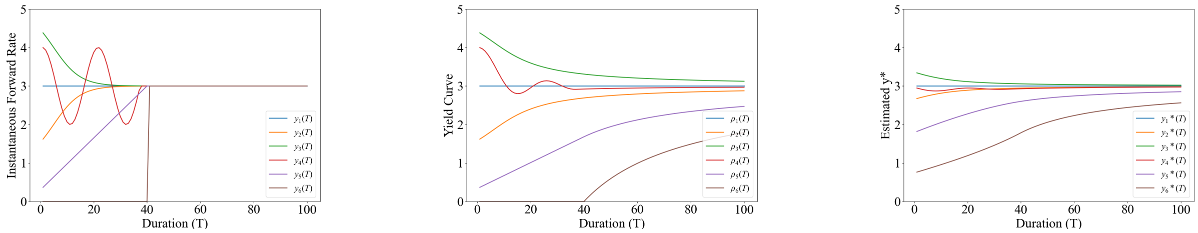
A.6 Simulation Results For Flexible Forward Curve

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.27a shows several possible choices of $y(T)$, all of which asymptote to the same value. Figure A.27b presents the yield curves associated with each of these forward curves, which we denote by $\rho(T)$; and Figure A.27c shows the corresponding $\hat{y}^*(T)$ curves that we estimate by equation (6) using simulation data. We also plot the point estimate of $\hat{y}^*$ we obtain at the median of our true 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.27.

Figure A.27: Long Run Discount Rates Using Simulation Data



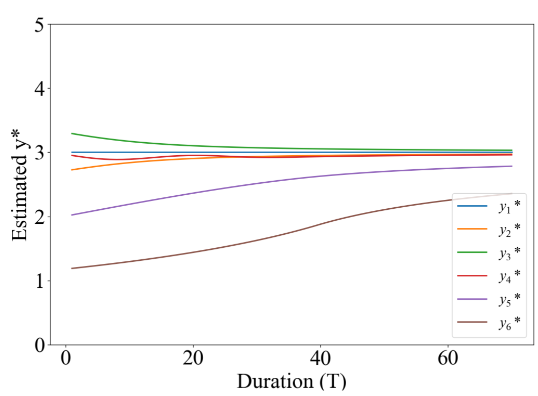
(a) $y(T)$ **(b)** $\rho(T)$ **(c)** $y^*(T)$

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.

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.28. 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.28: 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.27a.

A.7 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 i.

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

$$\log P_{jt'} - \log P_{jt} = \alpha + \sum_{s=t_0+1}^{t_1} \gamma_s D_s + \varepsilon_{j,t,t'} \tag{17}$$

for each property j in i'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. We include a constant, α, in the estimation equation as in Goetzmann and Spiegel (1995).

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_j, between the treated property i and the control j. We estimate,

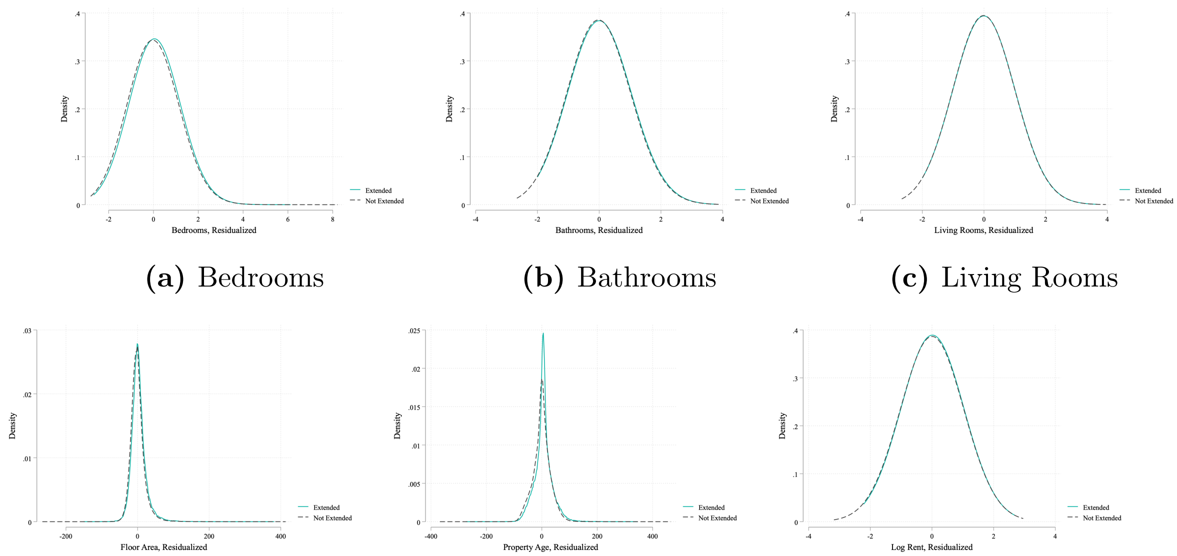
$$\hat{\varepsilon}^2_{j,t,t'} = \beta_0 + \beta_1(t' - t) + \beta_2 d_j + \mu_{j,t,t'} \tag{18}$$

where $\hat{\varepsilon}_{j,t,t'}$ are the predicted residuals from the regression in equation (17). Figure A.10 presents binscatters of $\hat{\varepsilon}_{j,t,t'}$ against $t' - t$ and d_j, showing that both are positively correlated. We estimate equation (18) for all properties together to reduce noise. Then, we calculate weights, $w_{j,t,t'} = 1/(\hat{\beta}_0 + \hat{\beta}_1(t' - t) + \hat{\beta}_2 d_j)$, equal to the inverse predicted values from equation (18). The last step is to rerun equation (17) using these weights.

A.8 Balance Test

In this subsection, we show that extended and control properties are balanced on observable characteristics. First, we regress our main hedonic variables—bedroom number, bathroom number, living room number, floor area, age, and rental price—on an experiment × year of listing fixed effect. We plot the density distribution of the residuals for extended and control properties separately in Figure A.29. The distributions are visually very similar.

Figure A.29: 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 a experiment $\times$ year of listing fixed effect. The sample is leasehold flats which appear at least once in the Land Registry Transaction Data Set.

The average differences between treatment and control are also very small. For instance, the mean difference in number of bedrooms between extended and control properties is 0.08 bedrooms, a tiny difference to the average bedroom count of 1.8. The mean difference in floor area is 3.5 sq. meters, compared to an average floor area of 67 sq. meters. As a whole, rental prices, which ought to measure the quality of the property, are about 2% higher in extended properties, or about £20 per month. These differences are even smaller if we consider properties extended for 90 years. Importantly, what matters for our identification assumption is whether extended properties *change* their hedonic characteristics more than their controls, which we test for in Section 4.4. Therefore, even if there were economically significant in differences in the level of these hedonic variables across extended and control properties, our identification assumption still holds.

A.9 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.[42]

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.7. 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.

Table A.7: 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

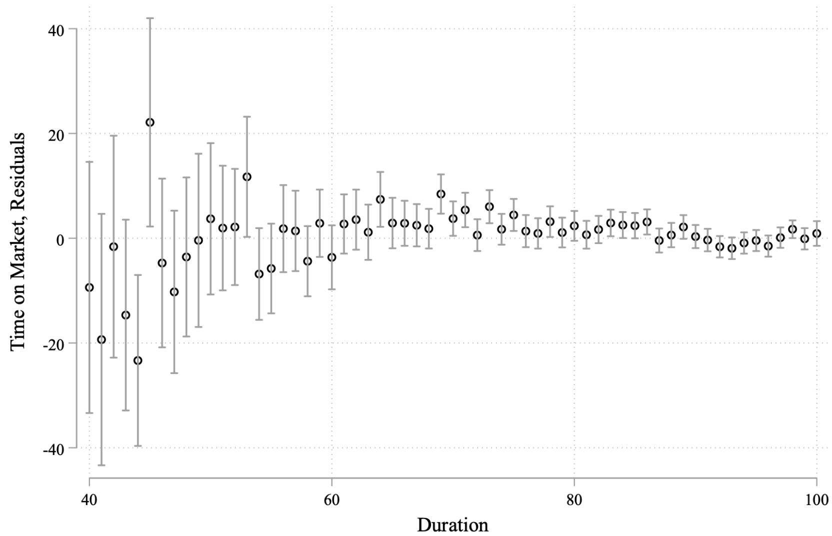
The table reports several summary statistics for leases of various durations. The first row presents the average mortgage length; the second presents the average Loan-To-Value (LTV) ratio for the mortgage, calculated as the initial mortgage value divided by the market price of the property; the third 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.

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 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

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

evidence of this in the data, as illustrated in Figure A.30. After controlling for quarter × 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.30: 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,

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

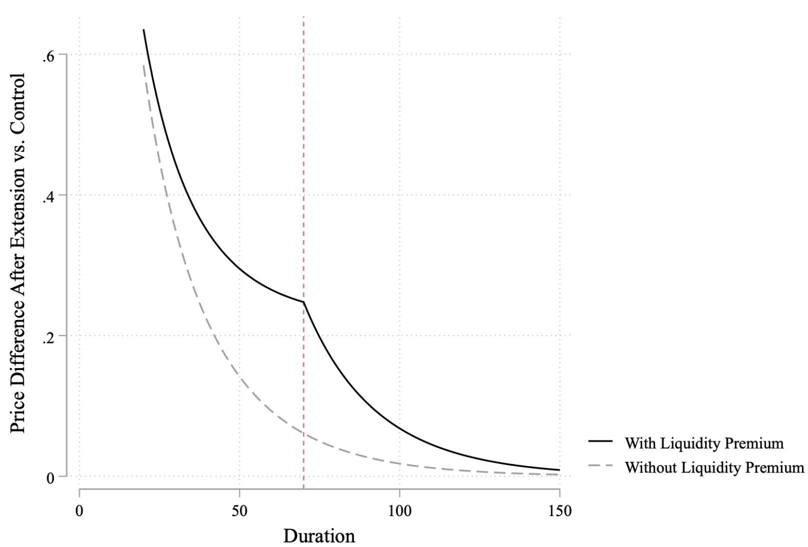
where rents below the cutoff of 70 are discounted at an additional rate σ. One way to think about σ is as the difference between the return on housing and outside investment options.

Therefore, our difference in difference will yield the following equation:

$$\log P_t^{T+90} - \log P_t^T = \log\left[\frac{1-e^{-(y^*)(T+90-70)}}{y^*} + \frac{1-e^{-(y^*+\sigma)70}}{y^*+\sigma}\left(e^{-(y^*)(T+90-70)}\right)\right]$$
$$- \log\left[\frac{1-e^{-(y^*)(\max\{0,T-70\})}}{y^*} + \frac{1-e^{-(y^*+\sigma)(\min\{70,T\})}}{y^*+\sigma}\left(e^{-(y^*)(\max\{0,T-70\})}\right)\right] \tag{19}$$

When $\sigma > 0$, the price change for extending a lease will have a kink at 70 as shown in Figure A.31. 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. The existence of a discontinuous liquidity premium at $T = 70$ can be formally rejected by estimating equation (19) by NLLS, which uses variation in $\underline{\Delta}_{it}^T$ for different values of T to test for the existence of a liquidity premium.

Figure A.31: 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 (19) 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.10 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.8, which indicate that there is no statistically significant difference in the level of y_t^* across quarters.

Table A.8: Test for Estimate Seasonality

	(1)
2nd Quarter	0.06
	(0.08)
3rd Quarter	-0.08
	(0.06)
4th Quarter	-0.03
	(0.05)
Year FE	✓
N	83

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.

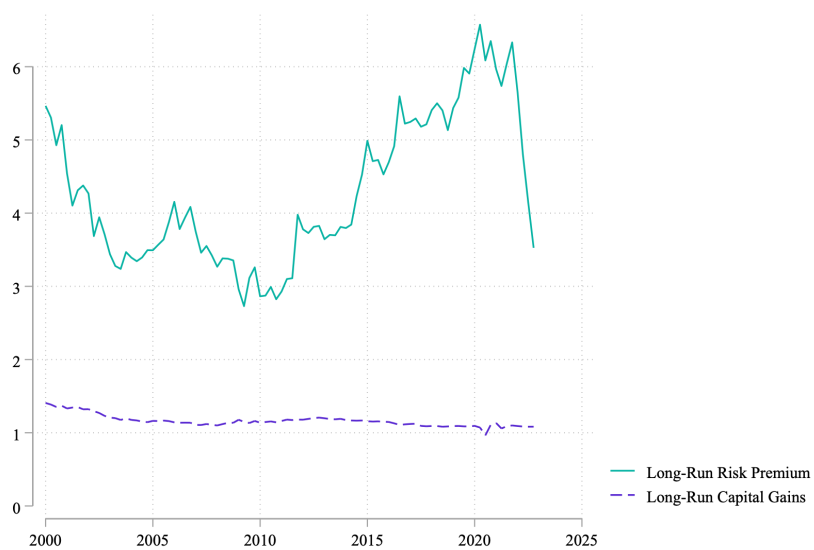
A.11 Components of y^* for UK Housing—A VAR Analysis

This section uses a VAR to estimate long run risk premia and expected dividend growth 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 dividend growth, we follow the standard VAR approach, with growth rate of rents, the rent to price ratio, and the GDP growth rate as inputs. To estimate long run expected dividend growth, 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 dividend growth. 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.32 displays the results. The estimates for long run expected dividend growth, 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 dividend growth 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.32: Risk Premia and Dividend Growth for UK Housing in the Long Run

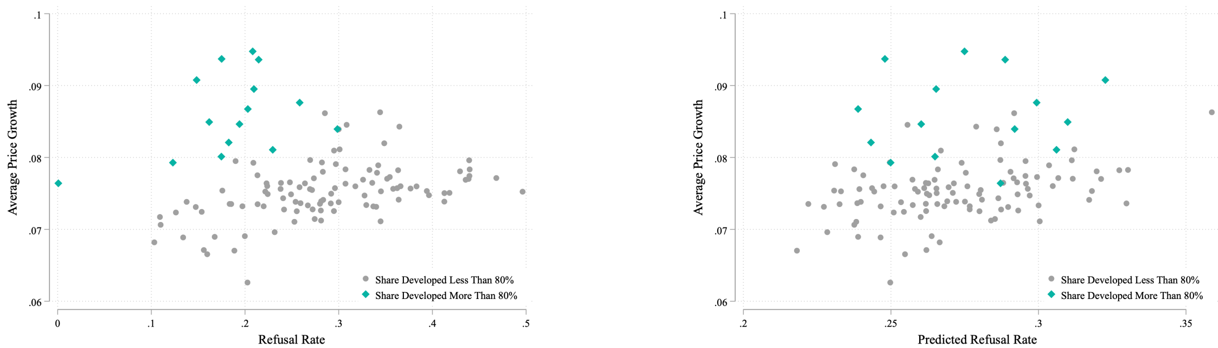


The figure presents estimates of long-run housing risk premia and dividend growth estimated using a VAR. Estimates are produced using data on the UK rent-to-price ratio from the OECD Housing Price Index, as well as the mean level of house prices from Land Registry data and the mean level of rents from 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.12 Housing Supply Elasticity Instrument

In this section, we provide more details on the housing supply elasticity instrument developed by Hilber and Vermeulen (2016). As described in the main text, Hilber and Vermeulen (2016)'s primary measure of supply inelasticity is the average refusal rate of major construction applications for each Local Authority, average over a nearly thirty year period. The problem with this measure is that the housing application rate is endogenous. More restrictive Local Authorities will on average receive less applications if developers expect that their applications will be refused. This in turn lowers the measured refusal rate.

Figure A.33: Endogeneity of Application Refusal Rate



(a) Not Instrumented **(b)** Instrumented

This problem is evident in Figure A.33a, where we plot average price growth over the period of 1974-2022, a measure which we expected to be very correlated with housing supply elasticity, against the average refusal rate measures. In blue, we mark the Local Authorities for which more than 80% of land is developed, all of which are in the three most central Tube zones of London. More than 60% of these LAs are in "Central Activities Zone," London's very dense touristic and cultural center. The application refusal rates in these LAs are somewhat below average, probably because many developers in these areas do not take the time to file applications which almost certainly will be refused. On aggregate, the implied relationship between the refusal rate and house price growth is weakly positive and statistically insignificant. However, if we instrument the refusal rate using the delay rate instrument introduced by Hilber and Vermeulen (2016), the slope coefficient is nearly an order of magnitude larger and statistically significant.

A.13 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^H_{Tt} be the likelihood that a lease, with $T > 80$ years of duration remaining, extends before its duration reaches 80 years. Let Π^L_{Tt} 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

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

which is the expression in the main text.

A.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 (12) and equation (13), the price of a property is

$$P_{it}^T = \begin{cases} P_{it}^{T+90} - \min\left[RV_{it}^T + \gamma R_{it}, MV_{it}^T\right] & T \geq 80 \\ P_{it}^{T+90} - \min\left[\frac{RV_{it}^T + MV_{it}^T}{2} + \gamma R_{it}, MV_{it}^T\right] & T < 80. \end{cases} \tag{20}$$

Recall the definitions of reversion value and marriage value

$$MV_{it}^T = \frac{R_{it}}{y_t^*}\left(e^{-y_t^* T} - e^{-y_t^*(T+90)}\right) \tag{21}$$

$$RV_{it}^T = \frac{R_{it}}{r_{RV}}\left(e^{-r_{RV} T} - e^{-r_{RV}(T+90)}\right) \tag{22}$$

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

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

$$\implies \frac{e^{-r_{RV} T} - e^{-r_{RV}(T+90)}}{r_{RV}} + \gamma = \frac{e^{-\bar{r}_K T} - e^{-\bar{r}_K(T+90)}}{\bar{r}_K} \tag{23}$$

where in the first line we have substituted in the definitions of marriage value (equation (21)) and reversion value (equation (22)). The right hand side of equation (23) 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 (23) implies that for all $y_t^* \geq \bar{r}_K$ we must have

$$RV_{it}^T + \gamma R_{it} \geq MV_{it}^T. \tag{24}$$

equation (24) implies

$$
\begin{aligned}
& RV_{it}^T + \gamma R_{it} \geq MV_{it}^T \\
\implies & RV_{it}^T + \gamma R_{it} + MV_{it}^T \geq 2MV_{it}^T \\
\implies & \frac{RV_{it}^T + \gamma R_{it} + MV_{it}^T}{2} \geq MV_{it}^T \\
\implies & \frac{RV_{it}^T + MV_{it}^T}{2} + \gamma R_{it} \geq MV_{it}^T \qquad (25)
\end{aligned}
$$

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

$$
P_{it}^T = \begin{cases} P_{it}^{T+90} - MV_{it}^T & T \geq 80 \\ P_{it}^{T+90} - MV_{it}^T & T < 80, \end{cases} \qquad (26)
$$

where we have substituted equation (24) and equation (25) into equation (20) 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 (26) shows that $\alpha_t^T = 1$ for all t, which proves part (1a) of the proposition. Since the top and bottom of equation (26) 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 (23) implies that for all $y_t^* < \bar{r}_K$ we must have

$$
RV_{it}^T + \gamma R_{it} < MV_{it}^T. \qquad (27)
$$

Then by equation (20), 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

$$
P_{it}^T = P_{it}^{T+90} - \min\left[\frac{RV_{it}^T + MV_{it}^T}{2} + \gamma R_{it}, MV_{it}^T\right]. \qquad (28)
$$

Also, note that

$$
RV_{it}^T + \gamma R_{it} < \min\left[\frac{RV_{it}^T + MV_{it}^T}{2} + \gamma R_{it}, MV_{it}^T\right], \qquad (29)
$$

since $RV_{it}^T + \gamma R_{it} < MV_{it}^T$ by inequality equation (27) and also by inequality equation (27) we have

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

$$\implies RV_{it}^T < MV_{it}^T$$

$$\implies 2RV_{it}^T < RV_{it}^T + MV_{it}^T$$

$$\implies RV_{it}^T < \frac{RV_{it}^T + MV_{it}^T}{2}$$

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

equation (28) and equation (29) 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.

□