

# Price Index of Private Rents detailed methodology

Detailed description and insight into the development of the ONS's new Price Index of Private Rents methodology and comparison with the ONS's old Index of Private Housing Rental Prices methodology.

Contact:  
Housing Market Indices team  
hpi@ons.gov.uk  
+44 1633 456400

Release date:  
27 October 2025

Next release:  
To be announced

## Table of contents

1. [Overview](#)
2. [Background](#)
3. [Previous methodology \(IPHRP\)](#)
4. [Current methodology \(PIPR\)](#)
5. [Discussion](#)
6. [Related links](#)
7. [Cite this methodology](#)

# 1 . Overview

This article provides a detailed description of the Price Index of Private Rents (PIPR) methodology and expands upon the higher-level methods summary in our [Price Index of Private Rents quality and methodology information \(QMI\)](#).

The PIPR measures inflation in the price of renting residential property from private landlords and letting agents in the UK. The [PIPR](#) measure replaced the previous [Index of Private Housing Rental Prices \(IPHRP\)](#) measure in March 2024 (for Great Britain) and March 2025 (for Northern Ireland).

The old IPHRP method used a matched pairs approach, with private rents data split between a sample pool and substitution pool. The IPHRP indices were calculated using the privately rented property data in the sample pool. The substitution pool enabled "replacement" of property records in the sample pool over 14 months old (that is, properties without updated data collected for them in the last 14 months) with the most recently collected comparable property record. This enabled the sample pool to represent a pseudo-fixed basket of privately rented properties, which was used to monitor price changes throughout the year.

The new PIPR calculates an index using a hedonic regression double imputation approach. One of the benefits of the new approach is that all the private rents data are used to calculate PIPR-based estimates, because there is no need to split the data into two pools (which was needed in IPHRP). This means the PIPR-based estimates are based on an increased sample size than IPHRP-based estimates, allowing for more granular statistics.

In addition, the hedonic approach used in PIPR does not require property updates to be collected within 14 months for them to be used in the index calculation. So, in contrast to IPHRP methodology, all rents data collected in a given month are included in the PIPR-based estimate calculations. This reduces bias against new let inflation in PIPR-based estimates compared with IPHRP-based estimates.

The article also describes how we decided upon our methods for some of the other detailed methodological choices made, such as regression model specification and imputation method for property characteristics. This part of the article mainly draws on material already published in the Advisory Panel on Consumer Prices technical paper [APCP-T2114-Rents-Development \(PDF, 1.53MB\)](#), and which we expand upon with updated and more complete information.

We have previously assessed the impact of moving from IPHRP to PIPR methodology for [Great Britain](#) and for [Northern Ireland](#). These impact analyses showed that IPHRP and PIPR inflation broadly demonstrated similar trends, but that PIPR was more responsive to market movements.

In this article we provide a more technical description and insight into the development of the PIPR methodology and compare this with the previous IPHRP methodology.

## 2 . Background

The Office for National Statistics (ONS) previously published two private rental prices statistical outputs: the [Index of Private Housing Rental Prices \(IPHRP\)](#), [UK statistical bulletins](#) and the [Private rental market summary statistics in England \(PRMS\) statistical bulletins](#). The IPHRP was published monthly and provided a rental price index and its annual percentage change for the UK, its countries and English regions. The PRMS was published twice a year and provided median monthly rental prices for England, English regions and English local authorities. These have now been replaced by a new single monthly publication for rental price statistics, the [Price Index of Private Rents \(PIPR\)](#).

The new PIPR measures inflation in the price of renting residential property from private landlords and letting agents, including new and existing tenancies. It is published as a series of price indices and levels covering the UK, its constituent countries, English regions, local authorities in England and Wales, and broad rental market areas (BRMAs) in Scotland and Northern Ireland. It is also published by property size for each geographical level (for example, number of bedrooms by local authority and by region). There are also further breakdowns by property type (for example, flat or detached house) for every geographical level.

PIPR aims to reflect price inflation for all private rental properties in the UK. However, as housing is a devolved topic, there are differences in the data collection methods used to collect rental prices for the different countries across the UK:

- England and Wales data are for achieved rents (the actual price a tenant pays to rent the property) for a mix of new and existing tenancies
- Scotland data are mainly based on advertised new lets (the advertised rental price, which may be different to the actual achieved rent price for a tenancy which may start at a later date) and some achieved rents (new and existing tenancies)
- Northern Ireland data consist of entirely advertised new lets

PIPR uses these data to approximate the rental stock price inflation and produce statistics aiming to reflect price changes for the average private rental property at a given point in time. For more information about the data collection differences between the different UK countries see our [Price Index of Private Rents quality and methodology information \(QMI\)](#).

The new PIPR uses a more complex methodology than the IPHRP method it replaced. PIPR indices are calculated using a hedonic double imputation approach, which is an internationally recognised method documented in Chapter 5: Hedonic Regression Methods of the [Eurostat handbook on residential property price indices \(PDF, 8.29MB\)](#) (page 54). This is similar to the approach we use to compile the [UK House Price Index](#) but using rental price rather than house price data.

The devolved nations collect rental prices as part of their statutory functions, relating to the setting of Local Housing Allowance rates for the administration of housing benefits. This is done by rental officers or agencies in each UK country.

The data are a purposive sample covering approximately 10% of the overall private rental property market in the UK (see our [Quality assurance of administrative data used in the Price Index of Private Rents](#)). These rents price data have been routinely used by the ONS to produce our regular publications of rental price statistics for many years. In late 2019, we started to explore using record-level data because of gaining access to record-level data for England, which ultimately enabled us to use a hedonic regression model approach.

Hedonic models assume that product prices are determined by their characteristics. The hedonic model used in PIPR is an ordinary least squares regression model fitted on the natural log rental prices of each month's sample of the rented stock at that time. This model fit is used to predict (impute) the monthly rental prices of a set of properties that is fixed at the beginning of the year. This fixed set of properties is called the fixed basket. It is updated annually every January to represent the changing stock of rental properties.

The predicted prices of the fixed basket calculated using the monthly hedonic models are then used to calculate price relatives between the current month and the base month (January) for each property. Elementary aggregates (or strata-level indices) are formed from an unweighted geometric average of these price relatives; in index number theory this type of index is referred to as a Jevons price index. These elementary aggregates are aggregated into higher-level aggregates, or published indices, using separately calculated expenditure weights.

The expenditure weights are calculated annually every February using the latest period available for data on the privately rented dwelling stock. These expenditure weights are used to account for potential non-representativeness in the fixed basket of properties (arising from rents price data collection being from a purposive sample). Use of expenditure weights in this way is referred to as a Lowe price index, which describes PIPR's published indices.

Both the weights and the fixed basket of properties are updated annually, resulting in a different unlinked price inflation series each year (January to January). Chain-linking with an overlap period in January is then used to tie the unlinked series together, creating a continuous price index series, allowing us to track changes in private rental property prices over time.

Like any other price index time series, the PIPR indices carry uncertainty because of both sampling and non-sampling errors. Our analysis has shown that the new PIPR method is an improvement on the IPHRP method it replaced. We are also continuously working on further improving our outputs and how we communicate the quality of our statistics, as detailed in the [Private rental prices development plan, UK](#).

In the next section we summarise the previous method and explain why a new method was used.

### 3 . Previous methodology (IPHRP)

As with the Price Index of Private Rents (PIPR), our previous methodology, the Index of Private Housing Rental Prices (IPHRP), measured price inflation of the privately rented stock, including inflation for new lets and for existing lets. However, a limitation of the matched pairs approach meant that IPHRP-based estimates were biased against fully capturing new let inflation, leading to slight over-reflection of existing let inflation.

Since inflation tends to be generally lower for existing rents than new lets, the matched pairs approach meant that IPHRP's stock measure was likely to slightly underestimate stock inflation during periods of rapid new let inflation and be less responsive to recent market inflationary pressures.

In addition, the IPHRP was published as a series of price indices covering the UK, its constituent countries and the English regions. However, it did not have indices at a more granular level, such as by local authority or broad rental market area, and rent price levels were not available.

Private rents data are collected separately in England, Wales, Scotland and Northern Ireland by rent officers who collect rental prices from letting agents, online sources (for Scotland the majority comes from these) and landlords who are willing to provide data (see [Linking the input data section](#) for more details about the sources of rental data). The sample is purposive, but our data suppliers set targets to ensure collection is representative of the private rental market (although for Scotland the aim is to reflect the advertised rental market).

In addition, we use expenditure weights to ensure that aggregated headline data are representative. Northern Ireland data have been included in IPHRP since 2015. They were excluded before this because of their infrequency and lack of coverage.

In an ideal world, updated rent price data would be available for all privately rented properties, every month. However, current rent data collection processes mean these are not available. To measure the stock of rental prices, we make use of the market behaviour that rent price tends to remain unchanged for a period of time, balanced with data collection operational processes, by assuming a collected rent price remains valid for up to 14 months. It is considered likely that the rent price has remained unchanged for this period of time. This is known as the 14-month validity period modelling assumption, and was used in IPHRP methodology as well as in PIPR methodology.

The method that was used in IPHRP was a matched pairs approach. Each January, a stratified random sample of half of all the records collected by rent officers over the previous 14 months was selected, called the "sample pool". The number of rent records in the sample pool remained fixed throughout the year and the sample pool was used to calculate a stock measure of rental price inflation for both new and existing rentals.

The remaining records were stored in a substitution pool. The size of this substitution pool fluctuated from month to month as old property records (older than 14 months) dropped out and newly collected rents (for properties not in the sample pool) were added.

To calculate a rental price index, the sample pool of properties was monitored for price changes throughout the year. The price of a property record in the sample was updated when a match was identified in the data collected in the latest month and the price change was within acceptable tolerance levels. The exact tolerance levels were:

- new price less than previous price multiplied by 1.49995
- new price greater than previous price multiplied by 0.6667

This meant that if the price rose by 50% or more, or fell by 33% or more, then the new price was not accepted. These tolerances were used because historical market behaviour shows that price changes above or below these thresholds within the space of 14 months are unlikely and is more likely to be an error and not a valid price update. For more information, see the article [Improvements to the measurement of owner occupiers' housing costs and private housing rental prices \(PDF, 2.48MB\)](#).

If a valid update was received, the entry date and rent price for that property record was updated and the rent for that record is considered valid for up to 14 months. Properties that did not have their price updated within 14 months were dropped from the sample pool and replaced with the most recent comparable replacement property record from the substitution pool.

A 14-month validity period was used as it balances typical contract lengths (which tend to be either 6, 12, 18 or 24 months) against operational practices. The [Goodlord Rental Index August 2020 \(PDF, 307KB\)](#) report showed that average tenancy term for the UK was 10 months, with the regional averages ranging between 9 and 11 months in all regions, apart from London where the average term was 14 months. The 14-month validity period modelling assumption reflects the behaviour of most private rent prices remaining constant for several months and not typically changing month to month.

Annex D in the [Improvements to the measurement of owner-occupiers' housing costs and private housing rental prices \(PDF, 2.48MB\)](#) article provides further information on the rationale for using a 14-month validity period. The 14-month validity period modelling assumption was reviewed in the [ONS's 2021 paper in Annex D \(PDF, 1.53 MB\)](#) and in the [ONS's 2023 paper \(PDF, 447KB\)](#) presented to the Technical Advisory Panel on Consumer Prices (APCP-T), focusing on Scotland. APCP-T concluded that there was [not sufficient evidence to justify a change to the 14-month validity period in ONS's rents methodology](#).

If a property record was last updated within the latest 14 months but was not updated in the latest month, its price was assumed to remain valid and so rent price inflation between the date of last update and the current month would be 0% for that property. If an update attempt was received but did not pass the above tolerance check, then the update was not accepted as valid for that property, the sample pool record remained unchanged, and the new record would be excluded from IPHRP data.

Data in the latest month's data collection that were not for any properties in the sample pool were added to the substitution pool to be used later in the processing. If a property record dropped out of the sample pool because of having no price update (or revisit) within 14 months, then the outgoing sample pool property record would be replaced with the most recent comparable property record (that is, for a property that has similar characteristics) from the substitution pool.

IPHRP directly derived price relatives for individual properties in the sample pool by taking ratios of collected prices in the current period to the base period (January). This meant we were comparing like for like, and annual weights were used to mix-adjust for compositional changes in the stock of rental properties by assigning fixed shares to properties with a given set of characteristics.

Mix-adjustment is a procedure used to reduce the effect of changes in the mix (composition) of the sample of rented properties on the rental price index, ensuring we are tracking pure price changes in rental property prices rather than changes in characteristics. The substitution pool was used to mitigate the effect of not receiving a property update within 14 months.

Both PIPR and IPHRP measure inflation of the privately rented stock: new lets and existing rents. New let inflation was captured in IPHRP-based estimates where a new tenancy had started since the previous data collection (within the previous 14 months) for a given property in the sample pool, and where a recent property record for a new let was used to "replace" a 15-month-old comparable property record in the sample pool. However, where time between property updates was longer than 14 months, or data collected was for a property not in the sample pool, then this record would have remained in the substitution pool and not used in calculation of IPHRP-based estimates unless the record was used to replace a 15-month old comparable property record in the sample pool.

New lets were more likely to be affected by this (for example, a property being privately rented for the first time ever, or the first time in at least 15 months, could not already be in IPHRP's sample pool or substitution pool) than existing lets, which may have been included in the last 14 months of data collection. This biased the IPHRP-based estimates against fully capturing new let inflation and slightly over-reflected existing let inflation.

The matched pairs approach meant that while the majority of collected rents data (new lets and existing rents) were used in IPHRP calculations, the matched pairs sample pool and substitution pool methodology meant that by definition not all rents data could be used in IPHRP index calculations. This reduced sample was one main limitation of IPHRP, and contrasts with the new PIPR method, which uses all the rents data available for each period.

Compared with IPHRP, PIPR's methods improvement increases the sample data volumes used in index calculations, improving the accuracy and stability of the PIPR-based estimate and allowing us to produce more geographically granular indices (for example, at local authority level).

Details on the quality characteristics of the data and further details on the old methodology can be found in our [IPHRP's quality and methodology information \(QMI\)](#).

## 4 . Current methodology (PIPR)

The Price Index of Private Rents (PIPR) shares the same rental price data source as the Index of Private Housing Rental Prices (IPHRP), namely a monthly data collection that is added to an existing sample of rents. However, since 2019 we have obtained access to property-level microdata for private rents data and property attributes data across the UK (see Data section in [The redevelopment of private rental prices statistics, intended methodology](#)). With richer record-level data we are now able to perform a regression model to predict prices for privately rented properties even when a price has not been collected this month for that given property.

For PIPR we use the full sample of collected rental prices (roughly 500,000 per year), rather than only part of it, as the basis for monitoring price changes throughout the year. We can do this because PIPR does not require a substitution pool from which to replace properties that drop out of the sample pool during the year. This increased sample size brings some important benefits:

- all available data for each period are used in constructing the index, improving the accuracy and stability of the estimates
- the sample is large enough to produce more geographically granular indices (such as local authority (LA) level) than IPHRP

One of the challenges we have addressed with implementing a hedonic regression is the difficulty in specifying an accurate model for the prices; a poor model will result in lower accuracy predicted prices. Because of the large number of interacting factors that determine the rental price of a property, it can be difficult to specify a model that is both accurate and robust.

Some of the information that would explain the rental price of a property might not be present in the data available, and there could be complex non-linearities that are not captured by the model. However, this risk is mitigated by the index being constructed from price relatives, which are a ratio of two predicted prices; if there is a potential bias in the predicted prices in a consistent direction, they are likely to cancel out when the index is calculated. More information on the model selection and assumptions is available later in Fitting a hedonic regression model.

In this section we describe in more detail how we use our source data to produce the published indices using the new PIPR method. The material presented here expands upon [Section 6: Methods used to produce PIPR data of the Price Index of Private Rents quality and methodology information \(QMI\)](#).

The main steps of PIPR's statistical process, which are described in detail in this section, are:

- input data are cleaned and linked together using a mix of manual and automatic data validation checks and linkage methods
- each January, for Great Britain (March, for Northern Ireland because of a 2-month lag in data availability), an annual fixed basket of properties is created using all rents data collected in the previous calendar year
- every month, the monthly dataset used by the hedonic regression model will be updated with the latest month's newly collected price records, and records over 14 months old are dropped
- automatic data cleaning checks are carried out on the fixed basket and on the monthly dataset
- missing attributes for properties in the fixed basket and in the monthly dataset are imputed
- a hedonic regression model is fitted to the monthly dataset and the coefficients from the model are used to calculate predicted rental prices for every property in the fixed basket for the current month
- annual expenditure weights are calculated by combining information on the stock of rental properties with the average observed rental price, to be later used when aggregating the strata-level indices
- elementary aggregates (or strata-level indices) are produced using a Jevons index
- strata-level indices are aggregated with expenditure weights into a Lowe index and then chain-linked annually (using January as the linking month) to produce a rental price index time series for a range of geographies and breakdowns
- to reduce volatility, indices below region level have a 3-month moving average applied, after chain-linking
- re-referencing and growth rate calculations for these granular level indices are performed after this
- price levels (or average rental prices) are calculated using the predicted prices fixed basket from a reference period
- a price level series is then calculated by extrapolating the reference period price levels using the growth rates of the referenced indices (this ensures that the price level series is consistent with the published index; at launch in March 2024, the reference period for PIPR was January 2023)

## Linking the input data

The first step in the new PIPR process is to link our administrative data sources together; the rental price data, which contains price data as well as some property attributes, and a few other sources that contain additional property attributes data.

Rental price data are collected by rent officers from agencies in each of the devolved nations of the UK:

- Valuation Office Agency (VOA) lettings information
- Welsh Government private rental data
- Scottish Government private rental data
- Northern Ireland Housing Executive (NIHE) private rental data (data are delivered with a 2-month lag)

The rental price data also include data on the number of bedrooms and property type for all countries, furnished status for all countries except Northern Ireland, and property age for Scotland.

Data on floor area and age of property come from VOA Council Tax for England and Wales, and from Land and Property Services (LPS) domestic valuation list for Northern Ireland. Data on floor area are not used for Scotland, because the address information in the Scotland rental price data was not of sufficient quality to obtain a high-quality linkage between price data and attribute property data on floor area. For all countries, the rental price data are linked to these additional attributes data via the Unique Property Reference Number (UPRN).

For many of our data sources we use the Office for National Statistics (ONS) Address Index Matching Service (AIMS) to link the supplied address data to a matched UPRN. AIMS is a tool used to index addresses collected through any data source, or manually entered by users, against the AddressBase database. See [ONS working paper series number 17 - Using data science for the address matching service](#) and [AIMS-API](#) for more details.

The following additional data sources are used by PIPR:

- [CACI Acorn](#), geo-demographic segmentation at the postcode level
- National Statistics Postcode Lookup (NSPL)

Acorn and NSPL data are linked to the rental data using postcodes.

Acorn uses geographical and demographic information about an area consumers live in and classifies postcodes into three levels, from least to most granular: category (seven categories), group (22 groups) and type (65 types). We use the group version of the Acorn classification in PIPR, as described in [How Acorn Works](#). The category version of the Acorn classification only has seven categories, which were too broad for our needs, while the Acorn type classification would have been too granular for use in our regression model.

The NSPL data are used to link property records to their higher-level geographies, for example, their LA. For Scotland we additionally map from LA to broad rental market area (BRMA) using a mapper supplied by the Scottish Government. For Northern Ireland, the postcode from the rental data is mapped to a BRMA using a mapper supplied by the NIHE.

Before delivery to the ONS, the data we use are quality assured and cleaned by our data providers who check for outlier values (for example, rents that are too high or too low), duplicates and incorrect column names.

Before records are linked, we carry out data cleaning steps, including ensuring all rental prices are recorded as monthly rent in pounds per month, and create derived variables, such as bedroom category (a categorical variable for the number of bedrooms).

More detailed information about the data sources used in PIPR and their quality can be found in our [Quality assurance of administrative data used in the PIPR](#). Average linking rates from January 2015 to December 2024 can be found in Table 1 of [PIPR's quality and methodology information \(QMI\)](#) for VOA Council Tax, LPS valuation list and Acorn data.

## Creating a fixed basket of properties

PIPR is mix-adjusted to account for the changing composition of rental properties being sampled in different months, and to account for distribution changes in the private rental sector stock over time. The process of mix-adjustment requires that in each January a fixed basket of properties is created using all the rental properties sampled in the previous calendar year.

This annual fixed basket of properties is used to produce predicted rental prices for the current calendar year before a new basket is constructed in the subsequent year. For example, the 2024 fixed basket for Great Britain was created in January 2024 and is made up of all privately rented properties for which data was collected in 2023, and this basket was used throughout 2024 to create the index series for 2024. In January 2025, the new year's fixed basket for Great Britain was prepared using 2024 rents data. As data for Northern Ireland are delivered with a 2-month lag, the annual fixed basket (January to December rents data) for Northern Ireland is constructed two months later, in March.

When creating the fixed basket, if any rental property is identified to have been visited more than once in the year, only the most recently collected data for that property would be used. Deduplication of properties in the fixed basket is performed, see [Data cleaning of extreme values and duplicates](#) for further details about data cleaning and duplicates.

In a 12-month period, over 500,000 rents are collected across the UK and used in PIPR calculations. Our [Quality assurance of administrative data used in the PIPR](#) shows that the annual rents data collection results in collection of over 450,000 rents for England, around 30,000 rents for Wales, up to 40,000 rents for Scotland and around 10,000 rents for Northern Ireland.

## Creating the monthly dataset for the model

PIPR is used to reflect price changes for all private rental prices currently being rented, or the stock of rental prices. This will include both properties currently being rented (existing tenancies) and those that are newly rented (new tenancies). In addition to the fixed basket (which is fixed for the entire year and contains January to December data from the previous calendar year), a "monthly dataset" of rental prices representing the current month's rental stock is used in PIPR methodology.

As for IPHRP, ideally updated rent price data would be available for all privately rented properties every month, but these are not currently available, so PIPR uses the same 14-month validity period modelling assumption used in IPHRP to make use of the market behaviour to not change rent price every month. More detail is in the [Previous methodology \(IPHRP\) section](#).

The monthly dataset is created using all the rental properties sampled in the most recent 14 months and keeping only the most recent rent record for each property in the sample. That is, the monthly dataset is updated every month with newly collected prices, and we only use the most recently collected rent in our PIPR calculations. This means the monthly dataset contains rents data from more recent periods than the fixed basket.

If a property is visited multiple times over a year, we will use only the most recently collected price, provided that this new price is within a pre-specified tolerance level. Any properties revisited with a valid new price will have all their property characteristics in the monthly dataset updated. Deduplication of properties in the monthly dataset is performed, see the Data cleaning of extreme values and duplicates section for further details about data cleaning and duplicates.

Our monthly dataset is used to construct a hedonic price regression model, each month. The hedonic model uses the supplied rents data to determine, for the current month, the relationship between rent price and main property characteristics. This monthly relationship is then used to predict the current month's price for all properties in the fixed basket. The following month, the monthly dataset is updated, and the hedonic regression is run again on the updated data, which produces a different model prediction of rent price for the following month for the fixed basket properties.

By using a fixed basket of properties, we can predict rent price for the same set of properties for each month of the year, which removes effects of composition change. This allows us to estimate pure price change for an average privately rented property in the stock of all private rents, which is the aim of our private rental price index.

Once a month we update our rents sample using newly collected rents data using the following process.

If a property has already had data collected for it in the previous 14 months, then the price is updated and the collection date is updated, provided that the price change is within the acceptable tolerance levels (if it fails the tolerance levels then the existing price is carried over and the collection date is assumed to be unchanged). We use the same tolerance levels as IPHRP (see [Section 3. Previous methodology \(IPHRP\)](#)).

If a property is new to the sample, the rent record is added to the new sample with its collection date recorded.

The 14-month validity period modelling assumption is applied:

- for the remaining properties in the sample, rent price is considered to remain unchanged
- any property that has not received a price update for over 14 consecutive months is dropped from this updated sample, since it is considered likely that the rent price may have changed since data were last collected for that property

This approach means we only ever use the latest month's price for a property in our index calculations, which is either the price received in the latest month's data collection or the previous month's price copied forward (up to a maximum of 14 months, to reflect market behaviour tendency to retain the same rent price for a period of time). We do not take a 14-month average of all prices we have for a given property in the validity period. See [How we measure rental price inflation](#) for a worked example of price updates.

The [PIPR data volumes dataset](#) shows that, between January 2024 and December 2024, the average number of rents (rounded to the nearest thousand) used by the regression model each month was 496,000 rents for England, 27,000 rents for Wales, 43,000 rents for Scotland and 11,000 rents for Northern Ireland.

# Data cleaning of extreme values and duplicates

## Data cleaning checks

Automatic data cleaning checks are carried out on the fixed basket and on the monthly dataset to remove extreme observations and ensure reported values are plausible.

We check that a property's floor area is plausible by using the [legal requirement](#) for houses of multiple occupation (HMOs). We reject floor area values that are below 4.64 square metres multiplied by the number of bedrooms (which would be below the legal requirement) or greater than 800 square metres. Where a floor area is rejected this way, it is replaced with an imputed value (see Imputation of missing property characteristics).

There is also a data cleaning check applied to the number of bedrooms recorded for a property: if it is a multiple of 11, then it is set to its value divided by 11. This is because the types of properties covered in the collected data are for residential dwellings that are privately rented and not managed accommodation with large numbers of bedrooms, such as student halls.

During development of PIPR, our assessment of the rents data found (by linking the rents price data with other sources of attributes data to verify values have been entered correctly) that on the rare occasion the number of bedrooms has been recorded as a multiple of 11, this was most likely a data entry error where the correct single-digit number has been mistakenly entered twice; for example, a bedroom number of 4, which has been mistakenly entered as 44. Therefore, we assume that a recorded bedroom number that is a multiple of 11 is a data entry error and we correct the value by dividing by 11 to remove the duplicate digit (for example, replace 44 with the correct value of 4).

Prices to rent a single room in a HMO are also removed from the collected data because rooms are excluded from PIPR calculations (the previous IPHRP measure also excluded rooms). Where they can be identified, properties where the tenants are in receipt of housing benefit are removed from the rents data provided by our data suppliers and are excluded from PIPR. See our [Quality assurance of administrative data used in the PIPR](#).

Since data volumes are not high enough to produce reliable estimates for studio properties, studio properties and non-studio one-bedroom properties are combined to produce the one-bedroom category breakdown published by PIPR.

Additionally, any properties where we are unable to identify the local authority (for England and Wales), or the broad rental market area (for Scotland and Northern Ireland), are excluded from rents data used in PIPR calculations.

## Treatment of duplicate cases

There may be multiple records with the same property ID but with different information, such as address. This may happen when data are collected for a given property multiple times within a single month, or if a property is incorrectly assigned the wrong property ID number. This leads to some duplication of property records within the same month in the rents data supplied to the ONS.

We therefore undertake a deduplication process on each monthly delivery of rental prices data to ensure that each property in this month of data will have a unique property ID, allowing the new data to be appended to the existing monthly dataset (or during the creation of the fixed basket) and any matching properties correctly linked and their rent updated.

## Imputation of missing property characteristics

Both the fixed basket and monthly dataset have missing values for some of the property characteristics. The average missingness rates by property characteristic are reported in Table 1 for the fixed basket. Missingness rates for the monthly dataset are very similar to those for the fixed basket.

Table 1: Average imputation rate from the fixed basket between January 2015 and December 2024

	<b>England &amp; Wales</b>	<b>Scotland</b>	<b>Northern Ireland</b>
<b>Property type</b>	0.00%	0.17%	0.36%
<b>Number of bedrooms</b>	0.00%	0.00%	0.00%
<b>Furnished status</b>	0.00%	0.00%	N/A
<b>Floor area</b>	10.84%	N/A	13.53%
<b>Property age</b>	10.30%	0.24%	19.96%
<b>Acorn category</b>	0.05%	0.01%	1.09%

Source: Price Index of Private Rents from the Office for National Statistics

### Notes

1. Table 2 in [PIPR's quality and methodology information](#)

Imputation is carried out before the monthly regression is performed. We primarily use [decision trees](#), which are a non-parametric supervised learning method used to solve classification and regression problems. The goal of decision trees is to create a model that predicts the value of a target variable by learning decision rules inferred from data features. An illustrative example is provided on page 6 in the Advisory Panel on Consumer Prices technical paper [APCP-T2114-Rents-Development \(PDF, 1.53MB\)](#). For a more detailed description of tree-based methods, including regression and classification trees, as well as a comparison with linear models, see Chapter 8 of [An introduction to statistical learning](#).

For the fixed basket, a univariate decision tree approach is used to impute missing values in all price-determining characteristics. Specifically, a decision tree regressor is used to impute missing values in the continuous variable (the natural logarithm of floor area) and a decision tree classifier is used to impute for the categorical variables.

Univariate decision trees were selected for the following reasons:

- they are fast to implement and re-train
- they perform similarly to the best performing algorithms after hyperparameter tuning
- they are simpler to define and explain than other alternatives
- they are generally easy to interpret because of the ability to visualise the tree and view the decision rules used
- decision trees are used in more complex algorithms (for example, random forests and gradient boosted trees) and popular imputation tools (for example, [multivariate imputation by chained equations](#)); this means that a different tool can be used in the future (as part of future improvements) without compromising methodology consistency

For the monthly dataset, a missing indicator approach is used to handle missing values in categorical variables, with missing values being set to a "missing" category value. The missing indicator approach is less appropriate for continuous variables, so the natural logarithm of floor area is imputed using the same decision tree regressor approach used by the fixed basket imputation process.

For the monthly dataset, unlike for the fixed basket, we do not use a decision tree approach to impute missing values in all price-determining characteristics. Our analysis showed increased volatility in the index and growth rate when decision tree imputation was used on all characteristics for the monthly dataset. We also considered the option of excluding all records with missing values in price-determining characteristics from the monthly dataset before the regression is performed. But this approach would have added representation bias, as the types of property that have missing property attribute data tend to be of a particular type (for example, flats in large cities).

## Fitting a hedonic regression model

Once property characteristics are imputed, a hedonic regression model is performed on the monthly dataset every month to estimate the contribution of each characteristic to the log rental price of a property; this is performed separately for each country. As we use the natural logarithm ( $\ln$ ) of the rental price as the dependent variable, it is a log-linear model. This approach was chosen because prices are log-normally distributed and it reduces the risk of heteroskedasticity (non-constant variance of the errors). See [Eurostat handbook on residential property price indices \(PDF, 8.29MB\)](#), page 50, paragraph 5.3, for more information on this in the housing context.

The coefficients from the model are applied to the fixed basket of properties to predict the rental price for each property in the basket for every month of the year. This makes PIPR methodology a hedonic double imputation approach, as collected rental prices in the fixed basket are replaced with predicted prices from the model.

PIPR uses an unweighted log-linear ordinary least squares regression model (OLS), with the following mathematical formulation:

$$\ln(p_i) = k + \sum_j \beta_j x_{ij} + e_i$$

where:

- $p_i$  is the rental price of property  $i$
- $k$  is a constant term
- $\beta_j$  is the coefficient associated with characteristic  $j$
- $x_{ij}$  is either a continuous variable in the case of floor area, or an indicator variable for all other characteristics (which are categorical); for the latter group, it indicates whether property  $i$  has the characteristic  $j$  (such as detached property); if so, it takes the value 1, otherwise it takes the value 0
- $e_i$  is the statistical error term for property  $i$

The price-determining characteristics that we use in the hedonic regression model are:

- number of bedrooms
- natural log of floor area (used only for properties in England, Wales and Northern Ireland)
- property type (detached, semi-detached, terraced, and flat or maisonette)
- furnished status (used only for properties in England, Wales and Scotland)
- geo-demographic segmentation (ACORN)
- local authority (LA) district in England and Wales, and broad rental market area (BRMA) in Scotland and Northern Ireland
- property age bracket

The natural log of floor area is calculated from the raw floor area values (in square metres) provided by our data suppliers, and this is used in the regression model. This was required because of the natural log of rental price being used as the dependent variable.

A separate regression model is run for each of England, Wales, Scotland and Northern Ireland. There are some differences in the regression model variables between the country models because of differences in data collection and availability across the UK.

For Scotland, natural log of floor area is excluded from the regression model for Scotland, for the reason mentioned in Linking the input data. For Northern Ireland, furnished status is excluded from the regression model for Northern Ireland, because furnished status is not available in the collected rental price data supplied to the ONS.

For Scotland and Northern Ireland, LA district is not used in the regression model. Instead, we use BRMA, which is broader than LA. This was recommended by the Scottish Government and the Northern Ireland Housing Executive (NIHE) because their data collection processes use BRMA, with collection counts by LA too low. There are 18 BRMAs in Scotland and eight in Northern Ireland.

Before running the regression, the bedroom number variable in the monthly dataset is capped at five. We use a cap of five bedrooms to improve the performance of our regression model as only 2% of all properties have five or more bedrooms and only 1% of all properties have six or more. Any properties with a missing or nil rental price are dropped from the monthly dataset. Additionally, where there are very low numbers of properties with a specific Acorn group in our monthly dataset, we combine this group with an adjacent group for the purposes of regression modelling.

We run our OLS regression twice, first to identify outliers and then to make predictions. It is run initially to calculate internally studentised residuals (defined later); any property with a studentised residual (in absolute value) greater than or equal to four are considered an outlier and are dropped from the monthly dataset. It is then re-run a second time on the monthly dataset after the outliers are dropped (note that outliers will not be dropped after the second OLS), and the coefficient estimates obtained from this second regression are those applied to the fixed basket to calculate predicted prices.

Internally studentised residuals,  $t_i$ , are defined for each observation,  $i = 1, \dots, n$ , as the residual obtained from the regression, divided by an estimate of its standard deviation and an influence measure:

$$t_i = \frac{\epsilon_i}{\hat{\sigma}\sqrt{1 - h_{ii}}}$$

- $\epsilon_i$  is the  $i^{\text{th}}$  residual
- $h_{ii}$  is the leverage, the  $i^{\text{th}}$  diagonal element of the hat matrix  $H = X(X^T X)^{-1} X^T$ , where  $X$  is the data matrix and  $X^T$  its transpose (see [hat matrix and leverage](#))
- the estimate of the standard deviation of residuals is

$$\hat{\sigma} = \sqrt{\frac{1}{n - m} \sum_{i=1}^n \epsilon_i^2}$$

where  $n$  is the number of observations and  $m$  the number of estimated parameters.

We can show how regression coefficients from the second OLS are used to predict the rental price with the following mathematical formula.

For a given year  $y$ , let the basket be the set of properties  $B_y$ ; for every property  $i$  in this basket, the price predicted from the OLS model of ln rental price in month  $m$  is derived as:

$$\ln(\hat{p}_{i(y,m)}) = \hat{k}_{(y,m)} + \sum_j \hat{\beta}_{j(y,m)} x_{ij}$$

where  $x_{ij}$  is the value of explanatory variable  $j$  for property  $i$ ,

$$\hat{\beta}_{j(y,m)}$$

is the estimated coefficient for variable  $j$  in year  $y$  and month  $m$  (the value of that coefficient in the model derived at that time), and

$$\hat{k}_{(y,m)}$$

is the constant term estimated from the regression model in year  $y$  and month  $m$ . This predicted ln price is calculated for all properties  $i$  in the fixed basket  $B_y$  for every month in that year.

It is worth highlighting that the set of explanatory variables included in the regression model are fixed (that is, they do not vary every time a new model is run each month). Using a fixed set of explanatory variables is not uncommon in regression analysis when a model is run within a monthly production pipeline and the model is used for predictive purposes, rather than estimating effects. It is also important to use a consistent model specification when calculating price relatives from predicted prices generated by a hedonic model.

The current explanatory variables were selected after extensive research and testing, as discussed in the following section. Each month test statistics are reviewed to ensure the hedonic regression model has run correctly and achieves the expected fit - see section "How we quality assure and validate data" in our [PIPR quality and methodology information \(QMI\)](#). The model will be reviewed every five years to ensure it is still the best model for the data we are receiving.

## How we selected our regression model

Differing model approaches were compared before choosing OLS. These approaches included weighted least squares (WLS), with and without interaction terms, and random forests. These methods were tested on data for England and Wales only, which accounts for approximately 90% of the UK private rents, because not all required data for Scotland and Northern Ireland were available when we started conducting this analysis. The suitability of the method was re-assessed at later stages during the project's development phase, including when Scotland and Northern Ireland were incorporated, to verify the selection.

To evaluate and compare these methods, statistical accuracy and performance of each was assessed. Quality aspects such as transparency and timeliness were also considered. For a detailed breakdown of the models tested and their results see Advisory Panel on Consumer Prices technical paper [APCP-T2114-Rents-Development \(PDF, 1.53MB\)](#).

Each model's ability to predict out of sample rental prices was assessed using a technique called K-fold cross-validation. K-fold cross-validation is a process for testing a model where the data are divided randomly into several "folds" (K). The model is fitted on the data in K-1 folds and tested on the remaining 1, and this process is repeated until every fold has been used for testing. For each test fold the root mean-squared error (RMSE), coefficient of determination ( $R^2$ ), and standardised residuals were computed. These metrics were then examined across all test folds (10 folds were chosen for this analysis) and compared across different models.

We found that both the OLS and WLS models had better average RMSE and  $R^2$  across 10 folds than the random forest-based method. Comparison of the distribution of the residuals showed that most of the residuals for OLS and WLS were centred around or close to zero. While the random forest was also mainly centred around or close to zero, it was slightly positively skewed.

Although the use of random forests is established in academic research, random forests are less commonly used for production of official statistics compared with general linear models and may therefore be considered more experimental. Furthermore, it is difficult to interpret contributions to the rental price from different property characteristics in random forests, because of a lack of regression coefficients being output and the increased complexity of the approach.

The WLS model did not improve on the OLS model but is more complex since observation weights must be configured. We also tested the effect of using interaction terms in the model. We found the WLS model with interaction effects yielded similar results to the WLS model without interaction effects, suggesting that a less complex model without interaction terms would be suitable.

## Weights

Because of the rents price data collection across the UK being a purposive sample of the private rental market stock and not a stratified random sample, the set of properties in the fixed basket we calculate predicted prices on needs to have weighting applied at the elementary level. This would ensure PIPR estimates are representative of the composition of the UK private rental stock of properties. For example, the rents data collected tend to under-collect flats as a proportion of the total dwelling stock compared with other property types.

To account for sampling biases in the collected rents data, PIPR uses expenditure weights when aggregating the strata-level indices (described in Stratification and elementary aggregates), to produce representative rents statistics for publication. These expenditure weights are derived by combining information on the stock of rental properties with the average rental price. These expenditure weights are calculated annually in February using mostly statistical surveys or census, with the average rental price derived from our rental data sources.

In PIPR, the UK private rental property sector is stratified by local authority (England and Wales) or broad rental market area (Scotland and Northern Ireland) combined with:

- property type and by furnished status, or
- bedroom category

Each stratum (combination of geography and property features as per the above) has a weight, defined as the estimated UK expenditure share on privately rented properties for the relevant stratum. To calculate expenditure, we multiply the latest estimate of the dwelling stock for each stratum with its average observed rental price.

If we are in year  $t$ , the weights for the current year are based on estimated expenditures in the previous year  $t-1$ .

We can use a subscript notation  $fri$  to describe the strata - let furnished status be  $f$ , property type be  $r$  and LA /BRMA be  $i$ . The weight for stratum  $fri$  at time  $t$ ,  $w_{fri,t}$ , is calculated using a simple expenditure share method:

$$w_{fri,t} = 100 \times \frac{\bar{p}_{fri,t-1} s_{fi,t-1} s_{ri,t-1} c_{i,t-1}}{\sum_{all\ l} \sum_{all\ m} \sum_{all\ n} \bar{p}_{lmn,t-1} s_{ln,t-1} s_{mn,t-1} c_{n,t-1}}$$

where:

- $p_{fri,t-1}$  is our best estimate of the average price of a property in stratum  $fri$  in the previous year
- $s_{fi,t-1}$  is the estimated share (or proportion) of properties in LA/BRMA  $i$  with furnished status  $f$  in the previous year
- $s_{ri,t-1}$  is the estimated share (or proportion) of rented properties in LA/BRMA  $i$  of property type  $r$  in the previous year
- $c_{i,t-1}$  is the estimated dwelling stock count of privately rented properties in LA/BRMA  $i$  in the previous year
- the denominator is the sum of the estimated expenditures across all strata

A similar approach is applied to calculate weights used for PIPR's bedroom category indices, where the share of properties of a particular bedroom category is used instead of the share of furnished properties combined with the share of properties of a particular type.

To calculate our weights, we use the latest available official statistics from a range of data sources.

Estimates for average rental prices are produced using the rental price data supplied by:

- Valuation Office Agency (VOA)
- Scottish Government
- Welsh Government
- Northern Ireland Housing Executive

Dwelling stock estimates for privately rented properties come from:

- The [ONS's subnational estimates of dwellings and households by tenure, England](#)
- [Welsh Government's dwelling stock estimates by local authority and tenure](#)
- [Scottish Government's housing statistics: stock by tenure](#)
- [Northern Ireland Department of Finance's annual housing stock statistics](#)

Northern Ireland dwelling stock data are published broken down by property type and we use data from the [Northern Ireland House Condition Survey](#) to produce dwelling stock estimates by tenure for Northern Ireland.

Dwelling stock estimates are split by the proportion of property types rented privately in Wales, Scotland and the nine regions of England using data from:

- English Housing Survey
- Scottish Housing Conditions Survey
- census (for properties in Wales)

Dwelling stock estimates are split by property furnished status, using the national-level split estimated from the Living Costs and Food Survey for Great Britain. For Northern Ireland, dwelling stock estimates are not split by furnished status.

Lastly, to calculate bedroom category weights we use bedroom category distributions estimated from the Family Resources Survey.

For more information about the data sources used to calculate our weights see [Expenditure weights: Quality assurance of administrative data used in the PIPR](#).

An [indicative weights summary](#) is published every March alongside our bulletin. Weights will always reflect our best understanding of the housing stock at the time they are calculated for use in PIPR. If some of these housing stock data are revised at a later date, following improved data or methods, it does not affect PIPR because we do not revise our weights in PIPR.

While we use the best available data to produce PIPR's expenditure weights, there are limitations in the available data, such as granularity and timeliness.

Sample sizes are not always large enough to break down compositions at the LA level. Data sources for property type and bedroom category breakdowns only produce reliable estimates at the region level. While our weights are calculated at the LA level, we apply region-level estimates, which creates increased uncertainty, for example, in regions that are dominated by large cities (property type composition is likely to be very different in cities compared with rural LAs). Data sources for furnished status breakdowns only produce reliable estimates at the national level.

There is a lag between publication date and the time period covered for our data sources for dwelling stock count, property type, bedroom category and furnished status. We use the timeliest data available, although these data sources can have between a 2- and 5-year lag. This is common in statistical production, and PIPR methodology aligns with standard practice within the ONS.

It is commonly accepted practice to use the timeliest data available to construct expenditure weights if the data are not available for the current year. It is for this same reason that national statistical institutes (NSIs) use the Lowe index formula with regular updates to the representative basket of goods and expenditure weights (see for example, [A practical introduction to index numbers by Jeff Ralph, Joe Winton and Rob O'Neill from Open Library](#), pages 137 to 138). An alternative could be to use a Young index instead, but this has inferior axiomatic properties compared with the Lowe index (see [Producer Price Index Manual: Theory and Practice. International Monetary Fund, Statistics Department \(2004\) \(PDF, 6.01MB\)](#), page 432).

We impute expenditure data where there are low dwelling counts (fewer than five properties in that stratum for rent price data) for a given stratum. This increases uncertainty for that stratum's weight when calculating expenditure values by multiplying price and stock data.

Although these data limitations increase uncertainty in our weights, weighted estimates are more reliable than an unweighted index, because an unweighted index would be strongly influenced by variations in the composition of the monthly sample. It is very challenging to make a numerical estimation of the error bounds for the weights, since the error term includes factors beyond the purposive sampling of the rents price data, such as the issues listed under these data limitations, which we are unable to measure.

PIPR is also used to measure the owner-occupiers' housing costs (OOH) component of the Consumer Prices Index including owner-occupiers' housing costs (CPIH). To measure OOH, only unfurnished properties are used, and the strata are weighted by the owner-occupier stock.

Owner-occupiers' housing costs are measured using the rental equivalence approach (see our [Consumer Prices Indices Technical Manual](#), Section 4.2). To derive the owner-occupier weights, we use rental prices but combined with stock and type share data pertaining to privately owned properties (instead of privately rented properties) from the same data sources we used for the PIPR weights.

The sources used to estimate rental stock and rental property type shares have information on whether the sampled property is owner-occupied or rented, allowing us to calculate the corresponding type shares for owner-occupied housing. The methods are the same as for PIPR and bedroom category weights, except that we count properties and calculate shares sub-setting by the owner-occupied property category instead of the privately rented property category. More information about how OOH is calculated in the CPIH is available in our Consumer Prices Indices Technical Manual.

## Index creation

The published indices are all created following a two-stage approach. First, price indices are compiled at the stratum level, which are then aggregated across all the strata by different groupings to give the indices for the breakdowns and geographical regions mentioned in [Section 2: Background](#). The price indices compiled in the first stage are called elementary aggregates, and those derived from the second stage are the final indices that we publish. This approach is also documented in Chapter 4 of the [Eurostat handbook on residential property price indices \(PDF, 8.29MB\)](#).

For a more general introduction to index numbers theory and applications, including the various types of index numbers formula mentioned in this section, we refer the reader to [A practical introduction to index numbers by Jeff Ralph, Joe Winton and Rob O'Neill from Open Library](#).

## Stratification and elementary aggregates

The predicted prices calculated from the regression model's coefficients allow for the calculation of a price relative for each property, achieved by dividing its predicted price in the current period by its predicted price in the base period, January.

Before price relatives are calculated, the fixed basket with predicted prices is stratified to reduce sample bias, as mentioned in Weights. There are two types of stratification that we use for the creation of elementary aggregates in PIPR: strata defined as the combination of the variables furnished status, local authority (LA) or broad rental market area (BRMA) and property type, and strata defined as the combination of bedroom category and LA (or BRMA).

The first step in calculating the indices is to generate an index for each stratum. These are derived using a Jevons index formula, that is, an unweighted geometric mean of price relatives between the current month and base month of all properties within a stratum. The formula is:

$$A_{s(y,m)} = \left[ \prod_{i \in s(y,m) \cap s(y,1)} \frac{\hat{p}_{i(y,m)}}{\hat{p}_{i(y,1)}} \right]^{\frac{1}{n_{s(y,m) \cap s(y,1)}}}$$

where:

- $A_{s(y,m)}$  is the elementary aggregate for a stratum  $s$  at year  $y$  month  $m$
- $p_{i(y,m)}$  is the predicted rental price of property  $i$  at year  $y$  month  $m$
- $p_{i(y,1)}$  is the predicted rental price of property  $i$  in the base period for year  $y$
- $n_{s(y,m)s(y,1)}$  is the number of properties in strata  $s$  with predicted prices in both current period and base period

Some strata in the annual fixed basket may contain only a few properties because of low numbers of privately rented properties in the rented stock. This is most likely to be seen for strata for detached furnished properties in some LAs in central London, because of very low numbers of rented detached furnished properties from which to collect data.

The elementary aggregates derived directly from the hedonic model for these strata with few properties are less likely to be representative of those strata because of their low property count. Thus, we impute for the elementary aggregates of any strata where the index has less than five properties, using a LA-level parent donor. This is where elementary aggregates for low count strata are imputed using the index of the LA to which they belong and aggregated into higher-level groupings as an imputation donor. The LA donor is calculated by aggregating the elementary aggregates of all strata in that LA that do not have low counts.

## Within-year unlinked indices and chain-linking

Within-year unlinked indices are calculated using a Lowe index formula as a weighted arithmetic average of the elementary aggregates after low count strata imputation. This is the process of aggregating the elementary aggregates upwards into higher level groupings. Let the higher-level grouping at year  $y$ , month  $m$  be  $g(y,m)$ . The unlinked index is calculated as:

$$U_{g(y,m)} = 100 \left[ \frac{\sum_{s(y,1) \in g(y,m)} A'_{s(y,m)} w_{s(y,1)}}{\sum_{s(y,1) \in g(y,m)} w_{s(y,1)}} \right]$$

where:

- $U_{g(y,m)}$  is the index for the higher-level grouping  $g$  in year  $y$ , month  $m$
- $A'_{s(y,m)}$  is the elementary aggregate post low count imputation for strata  $s$  at year  $y$ , month  $m$
- $W_{s(y,1)}$  is the expenditure weight for strata  $s$  in year  $y$
- $\frac{W_{s(y,1)}}{\sum_{s(y,1) \in g(y,m)} W_{s(y,1)}}$  is a rescaling factor used to make the adjusted weight sum to 1 within the higher-level grouping  $g(y,m)$

Once we have the within-year unlinked indices, we link these together into a single continuous series using standard annual chain-linking methods. Section 12.4.2 in our [Consumer Prices Indices Technical Manual](#) describes the standard chain-linking methodology used in our price statistics, including in PIPR, Retail Prices Index, and the UK House Price Index.

The overlap period between unlinked series in adjacent years is January. The overlapping indices for years  $y$  and  $y-1$  are  $U_{g(y,1)}$  and  $U_{g(y-1,13)}$ , respectively.

The chain-linked index at an arbitrary time  $(y,m)$  expressed in terms of the unlinked indices is as follows:

$$C_{g(y,m)} = U_{g(y,m)} \prod_{z=y_0}^{z=y-1} \frac{U_{g(z,13)}}{100}$$

where:

- $U_{g(y,m)}$  is the unlinked index for the higher-level grouping  $g$  at year  $y$ , month  $m$
- $U_{g(z,13)}$  is the unlinked index for the higher-level grouping at year  $z$ , month 13
- $y_0$  is the first year of the series
- the summation is over all years  $z$ , from the start of the series,  $y_0$ , until year  $y-1$

## 3-month moving average and growth rates

Small sample sizes at the LA level make the LA-level coefficient in the monthly hedonic model and hence the LA-level index volatile. To reduce this volatility, a 3-month moving average is applied to the chained LA-level indices for all periods.

The moving average is derived as follows:

$$C_{LA(t)}^{MA} = \begin{cases} \frac{1}{3} \sum_{k=t-2}^{k=t} C_{LA(k)}, & t \geq 4 \\ C_{LA(t)}, & t = 2 \text{ or } t = 3 \end{cases}$$

where:

- $C_{LA(t)}$  is the chain-linked LA level index at time  $t$
- $t$  is the condensed time notation for the chained index, equal to  $12y + m$ , where  $y = 0, 1, 2, 3, \dots$  and  $m = 1, 2, \dots, 11, 12$
- the time series starts at  $t = 2$ , because the first month of the first year is  $m = 2$

PIPR calculates 12-month (annual) and month-on-month (monthly) growth rates. For non-LA level indices, we use the chained index to calculate these, while for LA-level indices, we use the 3-month moving averaged indices.

Using the simplified time notation, the monthly growth rates are derived as follows:

$$G_{g(t)}^{monthly} = \begin{cases} 100 \left( \frac{C_{g(t)} - C_{g(t-1)}}{C_{g(t-1)}} \right) & \text{if } g \text{ is not a LA level group} \\ 100 \left( \frac{C_{g(t)}^{MA} - C_{g(t-1)}^{MA}}{C_{g(t-1)}^{MA}} \right) & \text{if } g \text{ is a LA level group} \end{cases}$$

similarly, for annual growth rates:

$$G_{g(t)}^{annual} = \begin{cases} 100 \left( \frac{C_{g(t)} - C_{g(t-12)}}{C_{g(t-12)}} \right) & \text{if } g \text{ is not a LA level group} \\ 100 \left( \frac{C_{g(t)}^{MA} - C_{g(t-12)}^{MA}}{C_{g(t-12)}^{MA}} \right) & \text{if } g \text{ is a LA level group} \end{cases}$$

## Re-referencing and the published index

The chain-linked index described in this section has the reference period set to the start of the series, February 2014, so the chain-linked index value in this period is 100. However, for the published index, the reference year is usually set to a recent year. This is done so that the reference basket used for price levels is representative.

At the time of writing, the reference year for PIPR is set to 2023, so the published index has a value of 100 in January 2023. The choice of reference year only affects the levels of the index and price levels, but none of the growth rates. We will update the reference period every five years.

The PIPR index creation methodology for Northern Ireland is the same as that for the other UK countries. However, Northern Ireland rental price data are delivered with a 2-month lag. To produce timely estimates, we impute Northern Ireland indices for the latest two months.

We do this by applying the monthly average of the latest available 2-month inflation rate for the Northern Ireland headline to the latest available index value for that series. Between March 2024 and March 2025, Great Britain headline growth rates were used to impute the missing months, but were found to be a poor predictor for Northern Ireland data so the method was improved in March 2025 (see [Advisory Panel on Consumer Prices \(Stakeholder and Technical\) minutes: 2 December 2024 - UK Statistics Authority](#)).

The imputed index values for Northern Ireland are then aggregated with the corresponding data for Great Britain, to produce provisional UK estimates for the latest two months to obtain a UK series (including UK-level breakdowns). All imputed Northern Ireland indices and UK estimates are revised using newly delivered Northern Ireland data every month, creating a 2-month revision period for the UK series in PIPR. For more information about the creation of a UK index and PIPR's 2-month revision policy, including a worked example, see [Section 6 in our PIPR quality and methodology information \(QMI\)](#).

## Price levels

One of the advantages of PIPR is that we are now able to publish estimates of average rental prices, the price levels; these are also broken down by LA (or BRMA) and by property type.

Price levels directly measure the absolute level of prices rather than changes in prices. In IPHRP we did not publish price levels, while in the PRMS we published medians and quartiles of monthly rents. Price levels are calculated as a weighted geometric average of the predicted prices, while the index is a weighted arithmetic average of predicted price relatives.

To derive price levels, we use:

- the results from the regression model at the reference period to provide an estimate of the geometric average price for each stratum for the basket in the reference year; these are the reference prices
- a type of nearest neighbour donor imputation is used to replace the estimated prices for strata with low property counts in the reference period

After low count imputation, these estimated strata reference prices are then aggregated using a weighted geometric mean with the expenditure weights of the reference period. This provides estimates of the average prices for higher-level groupings at the reference period.

Once we have the reference price for each required grouping at the current time,  $P_{g(t_r)}(t)$ , we can finally calculate the price levels,  $R_{g(t)}$ . For this, we apply the growth rates calculated from the indices for group  $g$  to the reference price, to extrapolate the price at any given time.

However, it can be mathematically shown that the cumulative monthly growth rate from the reference period to any period is equal to the referenced index, denoted by  $I_{g(t)}$ , rescaled such that it is equal to 1 at the reference period -

$$\frac{I_{g(t)}}{100}$$

Therefore, the price levels can be derived as follows:

$$R_{g(t)} = P_{g(t_r)}(t) \times \frac{I_{g(t)}}{100}$$

where:

- $P_{g(t_r)}(t)$  is the reference price estimate for grouping  $g$  at the reference time  $t_r$ , calculated at the current time  $t$
- $I_{g(t)}$  is the published index for grouping  $g$  at the current time  $t$

## Historical series

Data from PIPR are only available from January 2015 but there is a user need to understand the long-term inflation in private rents. The best source of private rents inflation before 2015 comes from the previous method, IPHRP.

To produce a [Price Index of Private Rents, UK: historical series](#) we have chain-linked IPHRP with PIPR in January 2015. This involved re-referencing the IPHRP indices so that the IPHRP index value matches the corresponding PIPR index value in January 2015. For more details see section "Creation of a historical series" in our [PIPR quality and methodology information \(QMI\)](#).

## 5 . Discussion

The Index of Private Housing Rental Prices (IPHRP) relied on splitting the rents data collected into a sample pool and substitution pool. Since the Price Index of Private Rents (PIPR) methodology has no such requirement and uses all rents data, the number of rents used to calculate PIPR estimates each month is larger than that for IPHRP. This reduces the effect of sampling variance on the PIPR index and makes PIPR estimates more robust than IPHRP estimates.

IPHRP measured rental price changes seen in properties directly, but only for properties that are updated (revisited) in the sample pool, or for properties brought in from the substitution pool to replace out-of-date data. This could lead to a higher representation of long-term tenancies, which are likely to experience lower rental inflation than new tenants (see for example [Rent Indices for Housing in West Germany 1985 to 1998, Johannes Hoffmann and Claudia Kurz](#)). This means IPHRP was likely to underestimate average UK rental price inflation because of not fully capturing new let inflation.

The hedonic approach used in PIPR does not require properties to be recollected in the sample pool for them to be used to track price changes, and thus PIPR is a more accurate measure of inflation of both existing tenancies and new tenancies than IPHRP.

Since IPHRP's methodology limitation restricting capture of new let inflation is not applicable in PIPR's methodology, new let inflation is better reflected in PIPR than in IPHRP, increasing the responsiveness of our private rental inflation measure to market conditions. This is a consequence of not requiring updated data for the same property for the data to be incorporated into index calculations.

In IPHRP, price changes were calculated from rents in the sample pool and if property record age exceeded the validity period then they were replaced with similar properties' records from the substitution pool. This means that IPHRP's sample pool was a pseudo fixed basket, not a true fixed basket. In contrast, PIPR's basket of properties is fixed over the year, improving the mix-adjustment in PIPR relative to in IPHRP. This means PIPR is more effective at stripping out influences of changing property composition from the price inflation measure, and producing an accurate measure of pure price inflation.

## Acknowledgements

The article was produced by the Economic Statistics Methods Hub and the Housing Market Indices team at the Office for National Statistics.

## 6 . Related links

### [Price Index of Private Rents quality and methodology information \(QMI\)](#)

Methodology | Released 20 March 2024

Quality and Methodology Information (QMI) for Price Index of Private Rents, detailing the strengths and limitations of the data, methods used and data uses and users.

### [Quality assurance of administrative data \(QAAD\) used in the Price Index of Private Rents](#)

Methodology | Released 20 March 2024

Quality assurance of administrative data (QAAD) used in monthly production of the Price Index of Private Rents (PIPR).

### [Index of Private Housing Rental Prices quality and methodology information \(QMI\)](#)

Methodology | Released 9 November 2018

Quality and Methodology Information (QMI) for the Index of Private Housing Rental Prices, detailing the strengths and limitations of the data, methods used and data uses and users.

### [Private rent and house prices. UK](#)

Statistical bulletin | Released monthly

The Price Index of Private Rents (PIPR) measures private rent inflation for new and existing tenancies. The UK House Price Index measures house price inflation.

### [Redevelopment of private rental prices statistics, impact analysis, UK: January 2025](#)

Article | Released 23 January 2025

We continue to redevelop our rents statistics and intend to improve Northern Ireland data from 2025. This article assesses the impact of this redevelopment.

### [Redevelopment of private rental prices statistics, impact analysis, UK: December 2023](#)

Article | Released 1 December 2023

We are redeveloping our private rental price statistics, intending to use them in our measurement of consumer prices from 2024. This article provides research indices using these data.

## 7 . Cite this methodology

Office for National Statistics (ONS), released 27 October 2025, ONS website, methodology, [Price Index of Private Rents detailed methodology](#)