

Quality assurance of administrative data used in the Price Index of Private Rents

Quality assurance of the administrative data used in the monthly production of the Price Index of Private Rents.

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

Release date:
20 March 2024

Next release:
To be announced

Table of contents

1. [Introduction to the Price Index of Private Rents](#)
2. [Quality assurance of administrative data assessment summary](#)
3. [Valuation Office Agency lettings information](#)
4. [Welsh Government private rental data](#)
5. [Scottish Government private rental data](#)
6. [Northern Ireland Housing Executive private rental data](#)
7. [Valuation Office Agency Council Tax](#)
8. [CACI Acorn](#)
9. [Expenditure weights](#)
10. [Related links](#)
11. [Cite this article](#)

1 . Introduction to the Price Index of Private Rents

The Price Index of Private Rents (PIPR) is a monthly release that publishes private rental price indices and price levels for the United Kingdom, Great Britain, its countries, English regions, English and Welsh local authorities, and Scottish broad rental market areas. The index is calculated using rental data collected by rent officers operating for the Valuation Office Agency (VOA), Welsh Government, Scottish Government, and Northern Ireland Housing Executive (with further data collected for Northern Ireland from propertynews.com).

Incorporating so many different data sources into any statistics involves a certain degree of risk. Administrative data may be collected and compiled by third parties outside the Code of Practice for Official Statistics.

This publication is part of an ongoing process of dialogue with our data suppliers to increase our understanding of any quality concerns in the source data. Through this document, we aim to provide information and assurance to users that the sources used to construct the PIPR is sufficient for the purposes for which it is used. We will, therefore, review this document regularly.

For further information relating to quality and methodology for the Price Index of Private Rents data, see our [Price Index of Private Rental Quality and Methodology Information \(QMI\)](#).

2 . Quality assurance of administrative data assessment summary

Office for Statistics Regulation toolkit

The assessment of our administrative data has been carried out in accordance with the [Office for Statistics Regulation \(OSR\) administrative data quality assurance toolkit](#).

The toolkit's risk and profile matrix is used to evaluate each administrative data source used in the production of the Price Index of Private Rents (PIPR). The matrix allows us to assign a level of risk to data quality and the public interest profile of the statistics. For more information on the matrix, see the [Administrative data quality assurance toolkit matrix](#).

The [UK Statistics Authority Administrative Data Quality Assurance Toolkit \(PDF, 243KB\)](#) outlines the assurance level framework for four specific data quality practice areas and the rest of this report will focus on these areas in turn. These areas are:

- operational context and administrative data collection
- communication with data supply partners
- quality assurance (QA) principles, standards and checks applied by data suppliers
- producer's QA investigations and documentation

Assessment and justification against the risk and profile matrix

In the assurance of our data source, we have chosen to give a separate risk and profile matrix score for each of the datasets. This will allow us to focus our investigatory efforts on areas of particular risk to the statistics. The sources are scored on their:

- level of risk of quality concerns
- public interest profile

Each administrative data source ranked as follows:

- Valuation Office Agency (VOA) lettings information: medium level of risk; higher public interest profile; justified by a high weight in the PIPR
- Welsh Government lettings information: low level of risk; medium public interest profile; justified by lower weight in the PIPR
- Scottish Government lettings information: low level of risk; medium public interest profile; justified by lower weight in the PIPR
- Northern Ireland Housing Executive: low level of risk; medium public interest profile; justified by lower weight in the PIPR
- VOA Council Tax: low level of risk; lower public interest profile; justified by the use for property attributes only
- CACI Acorn: low level of risk; lower public interest profile; justified by the use for property attributes only

All scoring was carried out by the prices division at the Office for National Statistics (ONS) based on the level of risk of the data and interest of our users.

We considered a high, medium, or low risk of data quality concerns based on:

- the importance of the data source in calculating the PIPR; what would we do if we did not have this data
- the use of the PIPR to also construct the owner-occupiers' housing component of the CPIH
- the complexity of the data source; for example, whether it is compiled from a number of different sources
- the existing contractual and communication arrangements currently in place
- other considerations, such as any existing published information on data collection, methodology or quality assurance, or mitigation of high-risk factors with the data

We considered a high, medium, or low public interest profile based on:

- the level of media or user interest in the PIPR or its sub-components
- the economic or political importance of the PIPR and its use in inflation statistics
- any additional scrutiny from commentators, based on particular concerns about the data

Assessment and justification against the quality assurance matrix

In line with the OSR expectations for QA of data sources, we have assessed the QA levels of each source, measuring them against:

- category one: operational context and administrative data collection
- category two: communication with data supply partners
- category three: QA principles, standards and checks applied by data suppliers
- category four: producers' QA investigations and documentation

Each of the four criteria are scored as follows:

- no assurance
- basic
- enhanced
- comprehensive

This is how each of the administrative data sources were assessed:

- VOA lettings information: comprehensive for operational context; comprehensive for communication with data suppliers; enhanced for QA principles by data suppliers; comprehensive for producers' QA investigations
- Welsh Government lettings information: comprehensive for operational context; comprehensive for communication with data suppliers; enhanced for QA principles by data suppliers; comprehensive for producers' QA investigations
- Scottish Government lettings information: comprehensive for operational context; comprehensive for communication with data suppliers; enhanced for QA principles by data suppliers; comprehensive for producers' QA investigations
- Northern Ireland Housing Executive: enhanced for operational context; comprehensive for communication with data suppliers; enhanced for QA principles by data suppliers; comprehensive for producers' QA investigations
- VOA Council Tax: enhanced for operational context; enhanced for communication with data suppliers; enhanced for QA principles by data suppliers; comprehensive for producers' QA investigations
- CACI Acorn: basic for operational context; basic for communication with data suppliers; no assurance for QA principles by data suppliers; basic for producers' QA investigations

For items collected as part of our Consumer Price Index (CPI), discussed in our [Consumer price inflation bulletins](#), quality management systems comply with ISO 9001 accreditation and the division is audited regularly by [Certification International](#) to ensure our systems are operating effectively and there is continued compliance with the standard.

3 . Valuation Office Agency lettings information

Operational context and administrative data collection

Valuation Office Agency (VOA) rent officers compile and maintain a database of private market rents in England. Prices are collected by rent officers from a variety of sources, including landlords and letting agents using a purposive sample, with the aim to collect approximately 10% of the market.

Monitoring tools are used to provide information on the volumes of collected data against the census figures, and rent officers apply their local knowledge to attempt to achieve a good distribution of rental prices.

There is no legal obligation for landlords and letting agents to share data with the VOA, and so all data shared is because of goodwill. There are no data agreements in place and no data are paid for. Voluntary data provision has resulted in an annual sample of over 450,000 rents.

The main statutory duties to collect these data are for housing benefit (Rent Officers (Housing Benefit Functions) Order 1997, as amended), local housing allowance and universal credit (Rent Officers (Universal Credit Functions) Order 2013 as amended) as well as Fair Rents (Rent Act 1977).

Collected rental price data are based on agreed rents (not advertised rents that are subject to negotiation), that is rent paid for the tenancy. The data comprises rents agreed because of letting to a new tenant, a renewal agreement with an existing tenant, or a rent increase during a periodic tenancy. The current database does not allow us to distinguish between these three categories. However, the collection processes ensure that every effort is made to capture the renewals and rent increases so that the data represents both the flow and stock of the market.

Rent officers are directed by law to “assume that no-one who would have been entitled to housing benefit had sought or is seeking the tenancy” when making their housing benefit-related determinations. Therefore, a tenancy identified as supported either wholly or partly by housing benefit is not knowingly added to the database. Local authorities are legally obliged to provide the Rent Officer with information pertaining to local housing allowance claims to assist with data validation.

Rent officers use a variety of collection methods, including:

- contact with letting agents, corporate landlords, or individual private landlords; these can be accompanied by evidence from recent letting lists, or their own previously-collected data as a prompt
- reports from letting agent or management company software systems, as many larger businesses use one of a number of standard property management software which can be used to generate a report which details all lettings made within a defined period; once the template is set up, this method allows Rent Officers to collect all the lettings an agent has completed on a regular basis, resulting in high-volume digital collection
- trade events; these range from local authority landlord forums to national exhibitions and professional conferences, where Rent Officers can proactively engage with many important influencers in the rented sector

Rent Officers follow regular collection cycles, and are alerted, through a quarterly report, of properties that are about to reach a 12-month anniversary since their last update. However, there is no formal procedure for VOA rent officers to revisit previous properties. As a result, some properties may drop out of our monthly dataset.

Communication with the data supplier

We have regular monthly meetings with statisticians and rent officers at the VOA to identify and resolve emerging collection or data issues. A formal service level agreement for the provision of sharing microdata is in place. This lists the legal basis for data supply, the transfer process, the schedule for data provision, as well as the data protection required. The service level agreement is reviewed annually. Under this agreement, we require three months' notice prior to any changes in data collection to ensure sufficient time for amendments, testing, and sign off.

Separately to regular meetings with the VOA, we also hold quarterly private rental market working group meetings with all rental data suppliers.

The VOA work closely with us to ensure their data collection practices are of high quality. In 2021, the VOA used our methods advisory service (MAS) to obtain advice and a review of their current data collection processes. The MAS provided important findings and recommendations to the VOA, which have been used to improve VOA's data collection and processing.

Quality assurance principles, standards and checks by the data supplier

The VOA has strong confidence in the core data attributes variables used for Local Housing Allowance, Universal Credit, and the statistical work with us. These are:

- address
- tenancy start date – month of year (90% and more capture)
- property type
- furnished status
- number of bedrooms
- type of tenancy
- rent payable
- periodicity of rent
- services

Other data attributes are collected to support comparable valuations for housing benefit and Fair Rent purposes and to enrich the dataset and understanding of the market:

- age
- number of living rooms
- kitchen
- bathrooms
- condition
- text free comments

It is not feasible or necessary for the VOA to include these for every record, and some, such as condition, are subjective and relative to the location. The Price Index of Private Rents (PIPR) does not use these variables.

As the data are entered onto the VOA system, there are a range of quality assurance and validation processes that take place. These include:

- cross checks against housing benefit and Fair Rents records, which allows VOA to make a judgement about the nature of the tenancy
- high-low rent checks to remove exceptional or outlier rental values within the context of each broad rental market area
- removal of duplicates, which is a two-part process; initially to view a list of possible duplicates, and secondly to remove any records added to the system in error
- 23 automated quality assurance checks, which include property and dwelling type validation, and various validations against property characteristics; typically, about 10% of the data entered during the month is subject to further challenge and scrutiny

Data issues identified as part of this process are reported back to the responsible rent officer. Rent Officers validate data collected by following up with sources where there is any question relating to rent, property attributes or validity. Ultimately, data are only used when the Rent Officer is satisfied it is genuine.

Whether or not a property is furnished is not always available. If the rent officers do not know a property's furnished status, rent officers input the property as being unfurnished. There is no way for us at the Office for National Statistics (ONS) to know whether the property had an unknown furnished status.

The PIPR uses information collected about the unique property address, rental price, number of bedrooms, property type, and furnished status.

Data collection also undergoes a range of regular auditing including hard copy data matching, accompanied visits and follow-up phone calls. Monitoring tools enable collection to be tracked.

Producers' quality assurance investigations and documentation

We check metrics detailing information on the sample size, the number of price updates, and the number of properties that have not been recollected within 14 months of the last collection. We also investigate the proportion of successful links to the VOA Council Tax data, which any reduction could determine issues with the property address.

Within the PIPR methodology, we remove outliers as determined by the studentised residuals calculated from the ordinary least squares regression model. Further information on our methods can be found in our [Quality and Methodology Information \(QMI\) document](#).

The final outputted indices are analysed and compared with other producers of rental price estimates. If any unexpected movements within the series are attributed to the raw data, these can be queried with the VOA who liaise directly with rental officers.

A strength of the data source for measuring rental prices is the large volume of collected data points. However, low collection rates in some local authorities and the use of a purposive sample can lead to volatility at local levels.

4 . Welsh Government private rental data

Operational context and administrative data collection

Rent Officers Wales, which is part of the Housing and Regeneration Division of the Welsh Government provides rental data to construct the Wales estimate. Prices are collected by rent officers from landlords and letting agents, with the aim to collect approximately 15% of the market using a purposive sample, using census data to establish a baseline for the sample. Data are collected for the use within Local Housing Allowance.

There are currently three Rent Officers and one line manager responsible for data collection within the 22 broad rental market areas (BRMAs) in Wales. Their main collection methods include visits, telephone calls and emails to letting agents, looking at letting lists and websites, and forums and surveys. Data are provided on a voluntary basis, and all data are based on current open market (achieved) rents. Voluntary data provision has resulted in an annual sample of around 30,000 rents.

The information is captured electronically in the Rent Officers Wales Lettings information database. Checks are carried out at the point of entry to ensure that any housing benefit funded tenancies are identified.

Landlords who only let one or two properties are contacted once or twice a year to obtain details, whereas agents and those landlords that have large portfolios are contacted frequently for new additions or changes to their letting portfolio. Rent Officers also monitor websites and follow up contacts with agents to obtain details as properties are let or removed from the sites.

Communication with the data supplier

We have regular monthly meetings with statisticians and Rent Officers at Welsh Government, to identify and resolve emerging collection or data issues. A formal service level agreement for the provision of sharing microdata is in place. This lists the legal basis for data supply, the transfer process, the schedule for data provision as well as the data protection required. The service level agreement is reviewed annually. Under this agreement, we require three months' notice prior to any changes in data collection to ensure sufficient time for amendments, testing and sign off.

Separately to regular meetings with Welsh Government, we also hold quarterly private rental market working group meetings with all data suppliers.

Quality assurance principles, standards and checks by the data supplier

As the data are entered onto the Welsh Government system, there are a range of quality assurance and validation processes that take place. These include:

- cross checks against housing benefit and Fair Rents records, which are flagged as necessary
- property and dwelling type validation
- various validations against property characteristics
- high-low rent checks identify and correct input errors
- duplicate properties are identified on input using unique property reference number (UPRN) matching
- AddressBase data lookup to ensure address data is consistent and accurate
- regular auditing using a random sample check against collected source material

The Price Index of Private Rents (PIPR) uses information collected about the UPRN, property address, rental price, number of bedrooms, property type, and furnished status.

Producers' quality assurance investigations and documentation

We check metrics detailing information on the sample size, the number of price updates, and the number of properties that have not been recollected within 14 months of the last collection. We also investigate the proportion of successful links to the VOA Council Tax data, which any reduction could determine issues with the UPRN supplied by the Welsh Government.

Within the PIPR methodology, we remove outliers as determined by the studentised residuals calculated from the ordinary least squares regression model. Further information on our methods can be found in our [Quality and Methodology Information \(QMI\) document](#).

The final outputted indices are analysed and compared with other producers of rental price estimates. If any unexpected movements within the series are attributed to the raw data, these can be queried with the Welsh Government who liaise directly with Rental Officers.

A strength of the data source for measuring rental prices is the large volume of collected data points. However, low collection rates in some local authorities and the use of a purposive sample can lead to volatility at local levels.

5 . Scottish Government private rental data

Operational context and administrative data collection

Rent Service Scotland, who are part of the Communities Analysis Division of the Scottish Government, supply rental data to construct the Scotland indices.

Rent Service Scotland currently collect rental data for a sample of properties at “Broad Rental Market Area” (BRMAs) level. BRMAs are defined in UK legislation and used by the Department for Work and Pensions (DWP) to set Local Housing Allowance rates. There are 18 BRMAs in Scotland.

The BRMA data collection is a sample of up to 40,000 individual rents each year, representing about 12% of all private rented dwellings. This is a mostly desk-based exercise to collect a representative sample of advertised rents from the internet and other available data sources, such as letting agents and private landlords through Rent Officer engagement activity and voluntary [market evidence returns](#). According to [Scottish Government's Private sector rent statistics](#), the data predominantly (around 85%) covers advertised rentals (when a property is newly advertised on the market) rather than achieved rents for new tenancies. Of the 15% of records based on achieved rents, many of these will be based on recently advertised properties.

The current sample provides a representative sample of newly advertised tenancies at BRMA level by bedroom size. It is not representative of newly advertised rental values at a local authority (LA) level, hence the Price Index of Private Rents (PIPR) publishes by BRMA. The sample is also not representative of existing tenancies; Rent Service Scotland has no power to mandate landlords provide rent data. Any information on existing tenancies is voluntarily provided by letting agencies or landlords.

Where possible, with the information collected, the database excludes any rents related to social housing, mid-market rents, halls of residence and private tenancies known to be the subject of housing benefit and regulated tenancies.

The data collection includes a minimum level of address, property attributes and tenancy details.

Landlord registration data and census data are used as a baseline for establishing and monitoring the total sample proportion that is aimed to be achieved. Although, currently, only the 2011 Census is available in Scotland to do this. Therefore, Rent Service Scotland uses annual data from the Scottish Household Survey on the size of the private rented sector (PRS) sector, by bedroom size at BRMA level. Rent Officers also monitor local market activity and take every opportunity to acquire feedback from landlords, letting agents and tenants. This market intelligence means that Rent Officers are able to continually evaluate the composition of the collected data and divert resources where necessary to address any perceived weakness in the data.

There are important policy differences in Scotland compared with the rest of the UK:

There have been rent caps for price increases within a tenancy, as a result of the Cost of Living Act (Tenant Protection) (Scotland), October 2022. The timeline for the main policy changes regarding the caps is as follows:

- from October 2022 until March 2023, most in-tenancy private rent increases were capped at 0%, and landlords could apply for increases of up to 3% to cover certain increases in cost
- from March 2023 until March 2024, most in-tenancy private rent increases will continue to be capped at 3% for any 12-month period, and landlords can apply for increases of up to 6% to cover certain increases in cost

From 1 April 2024 onwards, subject to parliamentary approval, the process for rent adjudication will temporarily be modified for one year. A taper approach is proposed, and is summarised on the [Scottish Government website](#).

This rental price cap only applies to in-tenancy rent increases, with no restriction on rent increases for new lets.

In Scotland, rents data are predominantly for advertised new lets (which are not subject to the price cap), with only a small proportion based on existing lets data. Data collection procedures do not involve actively seeking to re-collect data for previously collected properties. Therefore, price changes for existing tenancies are largely estimated for Scotland.

In PIPR, assumptions on average periods between rent price increases are used to measure price inflation for the stock of rents. We assume that rent price remains constant for up to 14 months if updated rents data for that property are not available. Records more than 14 months old dropped in PIPR.

Caution is advised when comparing Scotland's estimates with other areas in England and Wales and within Scotland. This is because of differences in data collection and housing policy (in-tenancy rent price increases are currently capped in Scotland) across the UK.

Communication with the data supplier

We have regular meetings with statisticians at the Scottish Government to identify and resolve emerging collection or data issues. A formal service level agreement for the provision of sharing microdata is in place. This lists the legal basis for data supply, the transfer process, the schedule for data provision, as well as the data protection required. The service level agreement is reviewed annually. Under this agreement, we require three months' notice prior to any changes in data collection to ensure sufficient time for amendments, testing, and sign off.

Separately to regular meetings with Scottish Government, we also hold quarterly private rental market working group meetings with all data suppliers.

Quality assurance principles, standards, and checks by the data supplier

Rent Service Scotland use local evidence to ensure the advertised rental data are representative (of size and type of property) of each broad rental market area.

There is a range of quality assurance and data validation processes that take place as the data are entered onto the system. These include:

- cross checks against housing benefit and Fair Rents records, which allows Rent Service Scotland to make a judgement about the nature of the tenancy
- high-low rent checks to remove exceptional or outlier rental values within the context of each broad rental market area
- removal of duplicates, which is a two-part process; initially to view a list of possible duplicates, and secondly to remove any records added to the system in error

The processes also include automated quality assurance checks, which involves identifying records with:

- zero bedrooms
- non-self-contained properties that are not “rooms”
- number of living rooms exceeding the number of bedrooms
- property type “studio” with more than one room
- self-contained “room” lettings
- number of bedrooms exceeding seven
- rents listed as including gas and electricity costs but without a value deducted from the rent for this component

The PIPR uses information collected about the property’s postcode, rental price, number of bedrooms, property type, property age and furnished status.

Data collection also undergoes a range of regular auditing including the team leader performing random sample checking, accompanied visits and follow up phone calls (where managers make phone calls to data sources to check the quality of interaction with the source). Monitoring tools enable collection to be tracked.

Producers’ quality assurance investigations and documentation

We check metrics detailing information on the sample size, the number of price updates, and the number of properties that have not been recollected within 14 months of the last collection.

The final outputted indices are analysed and compared with other producers of rental price estimates. If any unexpected movements within the series are attributed to the raw data, these can be queried with the Scottish Government, who liaise directly with rental officers.

Within the PIPR methodology, we remove outliers as determined by the studentised residuals calculated from the ordinary least squares regression model. Further information on our methods can be found in our [Quality and Methodology Information \(QMI\) document](#).

A strength of the data source for measuring rental prices is the large volume of collected data points. However, low collection rates in some broad rental market areas and the use of a purposive sample can lead to volatility at local levels.

Another limitation of the data source is that the Scotland rents data are mainly for advertised new lets.

6 . Northern Ireland Housing Executive private rental data

Operational context and administrative data collection

Rental data for Northern Ireland are provided by the Northern Ireland Housing Executive (NIHE), working in partnership with Ulster University and propertynews.com. These data are collected to inform the Housing Executive's ongoing research into house prices, rents, and affordability in Northern Ireland, and is a combination of two data sources. The first provides data originally collected by the NIHE to calculate Local Housing Allowance (LHA) rates for the administration of private sector Housing Benefit. The LHA data are combined with rental data from propertynews.com to provide coverage across Northern Ireland. Both datasets are based on advertised, rather than achieved, rents.

Data collection results in around 10,000 rental prices collected annually across Northern Ireland.

Data are provided at a six-week lag to other data sources used in the Price Index of Private Rents (PIPR). This is because of the data cleaning and combining of data sources undertaken by Ulster University and the NIHE.

The NIHE, in conjunction with Ulster University, publishes a bi-annual report on the Performance of the Private Rental Market in Northern Ireland, which is based on analysis of the combined LHA and propertynews.com data. The NIHE, in conjunction with Ulster University, also published a short overview of rental affordability in 2017, using this database. All of these reports can be accessed on the [NIHE website](#).

Communication with the data supplier

We have regular meetings with the research team at the NIHE to identify and resolve emerging collection or data issues. A formal information sharing agreement for the provision of sharing microdata is in place. This lists the legal basis for data supply, the transfer process, the schedule for data provision as well as the data protection required. The information sharing agreement is reviewed annually. Under this agreement, we require as much notice as is reasonably practicable prior to any changes in data collection to ensure sufficient time for amendments, testing, and sign off.

Separately to regular meetings with NIHE, we also hold quarterly private rental market working group meetings with all data suppliers.

Quality assurance principles, standards and checks by the data supplier

On receipt of the LHA data, the research unit at the NIHE perform quality assurance checks, including:

- checking all data is for the correct month
- check for missing data and misspellings in the address information (including postcode)
- checking high and low rents, verifying these against advertisements where possible
- checking the number of bedrooms and reception rooms, particularly any high numbers

The LHA data are then combined with data from propertynews.com. This combined dataset is quality assured by both Ulster University and the NIHE research unit. They check for removing outliers, invalid observations, multiple entries, and anomalies, such as:

- duplicate Unique Property Reference Numbers (UPRNs), including separate but adjacent flats which are kept in the file, or true duplicates (such as the same property captured in both datasets with slightly different spellings); in this case, the LHA record will be removed
- property type – ensuring terminology is correct, and the overall distribution by property
- display address – check for duplicates (which can identify spelling mistakes)
- town – check for naming consistency and field-appropriate data (such as the name of county instead of town)
- postcode – check for no missing data, and full postcodes
- monthly rent – check of high and low rents and verifying where possible
- bedrooms – check high number of bedrooms
- receptions – check high number of reception rooms

Producers' quality assurance investigations and documentation

We check metrics detailing information on the sample size, whether the rents are within the boundaries assigned by the user, and whether there are any duplicates in the file. We have import errors which include records that fail to import; for example, they might not contain rental values or have dates earlier or later than expected. We also have change queries, which are records of addresses that already exist in the repository. However, the attributes of the address may have changed from the last collection.

The final outputted indices are analysed and compared with other producers of rental price estimates. If any unexpected movements within the series are attributed to the raw data, these can be queried with the NIHE, who liaise directly with rental officers.

7 . Valuation Office Agency Council Tax

Operational context and administrative data collection

The Council Tax Valuation list is maintained by the Valuation Office Agency (VOA). The VOA is responsible for banding properties for Council Tax purposes. It is their duty to act fairly and impartially to make sure that each domestic property is correctly assessed and placed in the right band. As of 31 March 2023, there are 26.8 million properties with a Council Tax band in England and Wales.

Data are collected by billing authorities in England and Wales as part of the process of billing Council Tax, and from other sources including:

- the planning portal
- building developers
- taxpayers
- inspections

Data are updated when information comes to VOA's attention that a valuation list entry might be inaccurate, or there is a new build, demolition, or alteration.

Further information on updates to properties can be found within the [VOA property attribute data: quality assurance of administrative data \(QAAD\) used in Census 2021](#).

Communication with the data supplier

Initial meetings between us at the Office for National Statistics (ONS) and VOA took place to ensure the sharing of data was both reasonable and proportionate as required under the [Commissioners for Revenue and Customs Act 2005](#). An Information Sharing Protocol (ISP) was written up in 2017 as the data sharing agreement between VOA and the ONS. It covers details such as the legal gateway and the data sharing procedures to be followed.

The legal gateway to share the data to us was created through the [Statistics and Registration Service Act 2007](#). We received a snapshot of data in July 2016, which covered the period of 1 April 1993 to 1 July 2016. We now receive monthly update files that are encrypted and transferred through secure software.

Quality assurance principles, standards, and checks by the data supplier

In principle, VOA will review property attributes (including the number of rooms variable) whenever new information comes to light, and an improvement indicator is added if a change has taken place. If the change was splitting the property, merging properties, new builds, or a change of use (that is, to non-domestic), the property attribute data would be updated without a sale taking place. For other changes, the property attribute data are not updated until the case is cleared. This is triggered by notification that the property has been sold. If these changes would change the Council Tax banding, this change will be actioned once sold too.

The floor space and type of property are especially important in valuing property and assigning the Council Tax band, more so than the number of rooms. However, every effort is made to ensure the accuracy of all attribute data and multiple sources of information are used to do so.

VOA case workers are assigned to handle reports and collect the property information needed to update the valuation list. Where information received is incomplete or verification is required, this can be sought from a variety of sources such as:

- planning portal
- estate agents
- property websites (for example, Rightmove)
- Google StreetView
- digital mapping
- physical inspections
- communications with taxpayers

Information sources that can be accessed and utilised remotely are preferred and physical inspections are used when information is difficult to source by other means or requires additional verification.

Following collection of data by VOA case workers, the data attributes are input manually into the VOA central operational database; 5% of records input are checked by managers weekly (also see "Oversight of quality checks by VOA" in our [VOA property attributes data used in Census 2021 publication](#)).

Specific checks are completed on the data for missing and incompatible codes. For example, the number of rooms should always be greater than the number of bedrooms (except studio or bedsit). Whenever VOA interact with a taxpayer, inspect a property, or receive or research information about a property, all property attributes currently held by VOA should be checked and any missing attributes should be completed, if possible.

The VOA has been continually working to update and improve the accuracy of the Council Tax Valuation Lists, including work to limit the amount of missing property attribute data on its systems. As of March 2023, missing or "unknown" property type data has reduced to less than 1%. A more detailed breakdown can be found in Section 3.4 of the VOA's [Background Information Document of the Council Tax Valuation Lists](#).

The minimum standard is that on clearance of any report, proposal or appeal, the dwelling will have the eight primary property attributes completed:

- group (broadly a signal of architectural style)
- type (detached, semi-detached)
- age
- floor area (meaning size in metres squared)
- number of rooms
- number of bedrooms
- number of bathrooms
- number of floors

For the Price Index of Private Rents (PIPR), we use the property's unique property reference number (UPRN), floor area, and age.

Producers' quality assurance investigations and documentation

The production of statistics is distributed across different teams within the ONS, so basic quality checks are done by a centralised team when data are received. Data are then delivered to research teams to produce statistical outputs. These teams will carry out more in-depth quality checks specific to their statistical output.

Common processing is carried out by the ONS Data Engineering Team. An initial snapshot of VOA property attributes was taken in 2016 and supplied to the ONS. Since then, the ONS receives updates to these data from VOA every month in the form of six text files for changes, additions, and deletions for England and Wales separately.

The VOA data files are processed to create current and historical data. We add record start date, record end date (if applicable), a live record flag and a change event flag (new, old, updated, deleted). There is no additional cleaning, but records received are checked against those already held. Additions (new cases) are appended to the existing dataset and these records are checked to make sure they are whole new records rather than an update to an existing record. For deletions and changes, it is implied that the record is already held.

8 . CACI Acorn

Operational context and administrative data collection

The Acorn classification is produced by [CACI](#). It is a geo-demographic segmentation of residential neighbourhoods in the UK. It provides a general understanding of the attributes of a neighbourhood by classifying postcodes into a category, group, or type. It is used as an explanatory variable within the Price Index of Private Rents (PIPR).

Acorn draws on a wide range of data sources, both commercial and public sector Open Data and administrative data. These include HM Land Registry, Registers of Scotland, commercial sources of information on age of residents, ethnicity profiles, benefits data, population density and data on social housing and other rental property. In addition, CACI have created proprietary databases, including location of prisons, traveller sites, age-restricted housing, care homes, high-rise buildings, and student accommodation. Traditional data sources such as the Census of Population and large-volume lifestyle surveys are also used.

The Acorn type allocated to a postcode is predominantly calculated through an algorithm which is calculated using the data sources mentioned above. In some instances, a manual allocation of Acorn type is applied. Examples of this are for traveller sites, student halls of residence, and prisons.

Communication with the data supplier

We, at the Office for National Statistics (ONS), have a rolling three-year license with CACI for the use of the Acorn classification in the calculation of rental and house price indices. The licence also provides us with access to an account manager and technical support to answer any data queries that occur during the annual update. This provides a clear and established point of contact to discuss any issues or quality concern regarding the annual provision of Acorn data. No regular meetings are scheduled with CACI, although meetings are established when needed, and usually coincide with the renewal of the license agreement.

The Acorn dataset is provided to us on an annual basis. As such, newly created postcodes will not be incorporated into the Acorn dataset until its next annual update.

Quality assurance principles, standards, and checks by the data supplier

CACI Acorn has many methods of controlling, ensuring, and maintaining the quality of the input data to their models. These include:

- evaluating quality information published alongside the data sources used
- cross-checks against other data sources at the record and aggregate level
- manual checks (internet searches, maps, Zoopla)
- internal consistency checks

In evaluating the resulting model and output, the main method used to evaluate and monitor the segmentation of postcodes into classification types is based around the calculation of gains scores – specifically Gini scores – which measure the effectiveness of Acorn in discriminating across a wide range of variables.

Producers' quality assurance investigations and documentation

Acorn is updated annually for the PIPR purposes and then subsequently used monthly in the PIPR production process by matching the Acorn data to the latest rental data using the postcode attribute.

Internal (within the ONS) quality assurance takes place initially on the annual update of Acorn, to assess the latest delivery of Acorn data in comparison with previous versions. The main quality assurance at this stage is to assess the distribution of postcodes within each category. This distribution is compared with previous years and any substantial changes are investigated with the account manager at CACI for clarification. A further series of spot checks are carried out on postcodes that have disappeared from the latest delivery (meaning they were in the previous year's Acorn but are missing from the current). These cases usually relate to changes to postcode boundaries.

Further quality assurance then takes place monthly to ensure the matching of Acorn to property data takes place successfully. Further checks are then done within the modelling process.

The PIPR modelling process used can also account for those records where a match cannot be made between the CACI Acorn data and rental data. Any missing property attributes are allocated to a missing category. This process allows the use of all property transaction data in the calculation of average rental prices each month, even though some attribute data could be missing.

Coverage is comprehensive with data available at a postcode level for matching and is found to explain some of the price of a rental property, which is why it is included in the model. While it is acknowledged that the classification for every postcode may not be exact, this is not a requirement given how the data is used within the model. As the Acorn dataset is an annual release, new postcodes are not updated until the next release.

Overall, this data source is judged to be of adequate quality for the use to which it is being put in the PIPR.

9 . Expenditure weights

Expenditure weights are used to ensure the aggregated data are representative of expenditure on rental properties in the UK. They are derived by combining the information on the stock of rental properties with the average rental price.

Most of the data used to calculate the expenditure weights are from statistical surveys or census. The prices used are from our rental data sources, as described above.

The stock of rental properties is from our [Subnational estimates of dwellings and households by tenure, England: 2021 article](#), [Welsh Government's dwelling stock estimates by local authority and tenure](#), and [Scottish Government's housing statistics: stock by tenure](#). The England and Wales data are calculated using census data, supplemented with social survey data from the Annual Population Survey. The method of compiling the statistics is likely to produce less reliable estimates for the years further away from the census base. While the Scotland data are based on the 2001 Census, National Records of Scotland dwelling counts and Scottish Household Survey tenure splits.

England Housing Survey data, Scottish Housing Conditions Survey, and census (for Wales) are used to split the dwelling stock by property type. The Living Costs and Food Survey are used to split the dwelling stock by furnished status. The Family Resources Survey is used to split the dwelling stock by bedroom category.

This method enables us to calculate lower-level strata, which means we can produce more representative data. However, a limitation is that these dwelling stock data are not updated frequently and can be two or more years out of date.

10 . Related links

[Price Index of Private Rents Quality and Methodology Information \(QMI\)](#)

Methodology article | Released 20 March 2024

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

[Private rent and house prices](#)

Bulletin | Released 20 March 2024

The Price Index of Private Rents (PIPR) produces rent price and inflation statistics for the UK, tracking prices paid for new and existing tenancies in the private rental sector. Includes headline UK House Price Index statistics.

11 . Cite this article

Office for National Statistics (ONS), released 20 March 2024, ONS website, article, [Quality assurance of administrative data used in the Price Index of Private Rents](#)