

User guide to earnings and employment from Pay As You Earn Real Time Information

A comprehensive overview of the background and methodology of the Pay As You Earn (PAYE) Real Time Information (RTI) statistics.

Contact:
ONS Labour Market team and
HMRC RTI Statistics
labour.market@ons.gov.uk;
rtistatistics.enquiries@hmrc.gov.
uk
+44 1633 455400

Release date:
25 July 2025

Next release:
To be announced

Table of contents

1. [Overview](#)
2. [Calendarisation](#)
3. [Imputation of missing payment submissions](#)
4. [Industry classifications](#)
5. [Other adjustments to the data](#)
6. [Additional considerations](#)
7. [Access to data for accredited researchers](#)
8. [Data on Real Time Information](#)
9. [Glossary](#)
10. [Strengths and limitations](#)
11. [Related links](#)
12. [Cite this user guide](#)

1 . Overview

Earnings and employment from Pay As You Earn (PAYE) Real Time Information (RTI) monthly statistics covers the UK and contains estimates of payrolled employees and their pay. The statistics can be used to understand general trends and patterns in the UK's labour market and earnings, and to inform policy decisions and research. Stakeholders include but are not limited to HM Treasury, the Office for Budget Responsibility, the Bank of England, plus a variety of think tanks and academics.

The data for this publication come from HM Revenue and Customs's (HMRC's) PAYE RTI system, drawn from payment submissions made by employers. Payment submissions are digital documents that employers need to submit to HMRC every time they pay their employees. Employers inform HMRC of the employees' details, pay, and deductions. Data is received in real time, with submissions data being uploaded to analytical datasets overnight and made available to HMRC analysts the day after receipt.

The data cover the whole payrolled employee population, and are used to produce granular statistics on pay and employment. These statistics are jointly released by HMRC and the Office for National Statistics (ONS). This includes the monthly [Earnings and employment from Pay As You Earn Real Time Information bulletin](#) that highlights trends, as well as [accompanying datasets](#) that contain breakdowns of the data, allowing users to conduct their own analyses.

The RTI statistics have several strengths. They are timely and frequent, with a "flash" estimate first published two to three weeks after the reference period. This includes coverage based on data for of all payrolled employees across the UK, allowing distributional analysis and estimates to be produced for small sub-groups of the payrolled population. As they are not sample based, they are not affected by variations in response rate. However, these statistics also have limitations. The scope is employees and pay received through PAYE payroll, so this excludes those self-employed, and non-payrolled employee earnings. For example, this includes but is not limited to, those in benefits-in-kind as well as a small minority of employers who do not have to report payments through PAYE. Within the earliest "flash" estimate, around 15% of the data is imputed, meaning it is subject to revisions when more data becomes available.

This user guide provides a comprehensive overview of the background and methodology of the PAYE RTI statistics. It explores the methodologies involved in producing estimates of employment and earnings, including imputing data for more recent time periods. In instances where the data can be incomplete at the time of extraction, adjustments are made to make it more suitable. This methodology is designed to align with international guidelines for labour market statistics.

2 . Calendarisation

What calendarisation is and why it is necessary

Pay As You Earn (PAYE) Real Time Information (RTI) data consist of records of payments employers make to their employees, including the payment amount, payment date, and the start and end dates of employments. However, it does not contain the dates of the working period the employee is being paid for.

Other Office for National Statistics (ONS) labour market statistics have a framework to describe the major concepts that exist within the labour market and their relationship to each other. This approach aligns with international level, including the International Labour Organisation. This framework cannot be used precisely when using PAYE data in its raw form. Because these data have been collected primarily for tax purposes, measuring pay as the simple sum of payments within a period would be problematic when aggregating across monthly and weekly paid jobs from the PAYE RTI data.

The number of weeks in a month can vary as well as the number of weekdays falling in each month. For example, an employee paid each Friday could be paid four times in one month and five times in another. Aggregating pay across months would show their pay as being higher in the latter month, although their pay rate would not have changed.

In the PAYE RTI statistics, we adopt a method of calendarisation (instead of aggregating reported payments), calculating daily pay and averaging across the month. This method enables us to produce statistics that better reflect labour market conditions, reducing distortions caused by differences in the number of paydays and payments made each calendar month.

The method of calendarisation allows the alignment of the data to the period in which work was done. This brings the treatment of the data more in line with the [European System of Accounts \(ESA\) 2010 recommendations](#), including that wages and salaries are recorded in the period during which work is done. This brings better alignment to other labour market data sources, converting the data from a dataset of payments to a dataset of jobs and earnings. Consequently, counts of monthly jobs, employees and earnings can be produced more accurately, particularly for those paid weekly, bi-weekly or four-weekly.

The calendarisation methodology

To convert from period of payment to the estimated period of work, the payment amount is transformed into a daily pay rate. This is calculated by dividing the payment by the average length of the work period. For example, weekly pay is divided by seven, while monthly pay is divided by 30.4 (for example, 365 divided by 12).

This is done by dividing payments by the length of the work period before that payment. It is assumed that payments are made in arrears, that the payment date is the last day of the work period, and that the work period begins the day after the previous payment. The purpose of calculating this daily rate is to ensure that pay can be aggregated across employments with different pay frequencies.

The calculation of pay rates and work periods creates a daily dataset of all those currently employed, along with their respective rate of pay. This is then aggregated daily to calculate employment totals and pay statistics, and subsequently averaged over the calendar month to provide monthly aggregates.

For the first payment of an employment, the reported start date will be considered when calculating the beginning of the work period. If the start date generates a work period that is not too long or short according to the job's reported pay frequency, it will be used as the first day of the work period, with the payment date as the final day. Similarly, where a leaving date is recorded, this will be used as the end of the work period. A weighting methodology is used to handle jobs that begin or end part way through a month. If an employment begins halfway through a month, for example, it will be counted as half a job for the whole month. This will not affect the calculated pay rate for the employment.

For example, if a job pays £1000 a month and begins on 16 April, the pay rate for this employment will still be £1000 a month when calculating average pay for April. The shorter work period will be reflected in the employment being given a weight of a half when calculating average pay in April.

3 . Imputation of missing payment submissions

Why we need to impute some data

To produce timely statistics, the Real Time Information (RTI) data used in the estimation are generally extracted within a week of the end of the reference month. This will mean the dataset is incomplete at this stage as, for some individuals, payments relating to work done in that month are yet to be received. To produce reliable, unbiased statistics on a timely basis, imputation is required for recent periods to account for these reporting habits.

Two different approaches are taken to impute missing data. First, for those employments where there is a previous record of submissions but the next submission is not received and no leaving date is recorded, we need to determine whether there is a continued employment or not. Second, we also need to properly account for "new" employments not currently in the data (for example, a person has started a new job in the reference month, but their first payment submission has not been received). For both, we then need to calculate the pay rate for that imputed employment.

Estimating employment using probabilistic imputation

Probabilistic imputation is a method for handling missing data where missing values are replaced with plausible values drawn from a probability distribution. To impute for employees who have a previous employment record, but no current submission is received, the probability that employment continues is calculated based on historical data from similar cases. Two main variables are used: length of time since the previous payment date, and pay frequency. The former is important as, broadly speaking, the longer the amount of time since the last payment was received, the less likely it is that another payment will be received. In this context, it is also important to account for pay frequency. While there may be a high probability that a monthly paid job will receive another payment when the previous payment was made 25 days ago, there might be a lower probability of return if the job is paid weekly.

To calculate these likelihoods, historical data are used. We use them to calculate the probability that, at t days after the payment date, given a leaving date has not been submitted, and conditional on the pay frequency of the job, a next payment would be received. The imputation process involves imputing upcoming payments for all relevant active employments. This probabilistic approach provides a set of weights that can be assigned to the imputed payments to produce accurate aggregate estimates.

To summarise, for jobs that have previously submitted a payment to HM Revenue and Customs (HMRC) but have not submitted a leaving date, their last payment is "carried forward". This can be done several times depending on pay, frequency and historic probabilities, but with the calculated weight reducing in each subsequent imputation. The most recent payment must be within 14 months of the publication run date.

To calculate pay for these imputed employments, an average payment growth rate is calculated from the historic data and applied to the previous payment, for each job being imputed.

Changes to probabilistic imputation following labour market volatility

The probabilistic imputation model was improved in July 2022. This led to a reduction in the size of our initial payrolled employee revisions; particularly those associated with the end of tax-year period.

Firstly, the model was expanded to better account for annual patterns in the probability of employments remaining active in Real Time Indicators (RTI). Specifically, these patterns show that later in the tax year, employments are less likely to receive further payments following periods of inactivity. This improvement was achieved by creating month-specific sub-models that track changes at a more granular level. The outputs from these sub-models can then be better targeted at the relevant sample of current employments.

Secondly, the model was updated to incorporate newer data alongside established historical trends. This was implemented following analysis, which showed that recent revisions had a systemic bias caused by changes in the wider labour market. Specifically, the probabilities of employments remaining in RTI over time had decreased, leading to an over-estimation of payrolled employees in our initial figures. This new methodology calibrated the previously discussed annual trends in probabilistic weights with data from recent months. This improved reliability by combining established and contemporary data sources.

From the July 2022 publication, a refinement was made to the model to make it more responsive to changes in the labour market, by incorporating more recent data. Though the probabilities described above are still calculated using historical data, each month these historic trends are now compared with similar probabilities calculated from recent data from the last three months. This creates a ratio of how much recent months diverge from historic trends. The weight from historic data is still the main probability weight used, but it is adjusted up or down based on the average divergence ratio from the last three months of data. This better aligns the data with current trends, while still allowing weights to be built on a long-term annual cycle. More information on these changes and how they affected estimates can be found in the [Impact of imputation changes in employment statistics from PAYE RTI methodology](#).

Estimating employment and payment amounts for new jobs

New jobs, by definition, do not have a history of payments. The lack of information on past payments that could be "carried forward" to use in the imputation methodology (described in the previous subsection) means that a different approach is needed to impute first payments for new jobs.

There are two ways to impute first payments for new jobs. One way is to use those submissions relating to new jobs that we do have data for in the latest data extract. These submissions are treated as a biased sample of the total new jobs in the reference month we expect to receive data. These are grossed up to the total new jobs expected, based on historic data.

Historical data are analysed to estimate average bias based on the receipt date of payments relative to their payment date (controlling for pay frequency and the day of the month on which the payment falls). This is to account for potential bias in average pay levels from submissions already received when the data is extracted. These are then used to calculate bias adjustments for pay in the periods under imputation, which are applied to the data for the purposes of calculating average payment amounts.

In some cases, the percentage of submissions that have already been received for new employments may be too small to use as a sample and grossing based on the methodology outlined above. This is calculated based on a combination of payment month and payment frequency of the submissions. For these, a "nowcast" based on previous growth trends in the aggregated data is designed to impute new jobs. If the analysis of historic submission rates shows that responses for a particular pay frequency would be below 5% for a particular period, then this macro-based nowcast and forecast is incorporated to calculate new-job payments and average payment amounts for the period.

Depending on the submission rates, summary statistics are generated for all payment frequencies using one of the above methods. Once summary statistics are produced, they are used to derive "synthetic data" from these totals. The synthetic microdata is used alongside real receipts to produce the overall summary statistics and replicate other characteristics of the data. The purpose of this is to account for dimensions of our data (such as the distribution of pay), which:

- are not accounted for by the grossing methodology
- might differ between those who file their RTI submissions on a more timely basis, and those who do not

4 . Industry classifications

The industrial sectors in this bulletin are based on the UK Standard Industrial Classification (SIC) codes, as defined by the Office for National Statistics (ONS). These codes have been determined from both the most recent quarterly Inter-Departmental Business Register (IDBR) extract and data from Companies House for each Pay As You Earn (PAYE) enterprise.

Large enterprises that cover multiple SIC codes are classified into a single SIC code, based on the code that represents the highest number of employees. Changes to the proportion of employees across SIC codes in large enterprises can result in the enterprise being reclassified to a different SIC code.

During a standard monthly cycle, we incorporate the latest data from both of our sources, unless it would overwrite an existing assigned code. These more comprehensive codes are then applied to the recent data.

Once a year, when we refresh the data for the whole series, the IDBR link is refreshed using the most recent version available for all enterprises (including those that require overwriting). These codes are then applied across the entire time-series and form the basis for our industry data moving forward.

This means that breaks between recent and older data are less likely to be distorted by changes in employer classification relating to small changes at the lower unit level.

As a result, although the timeseries may be revised between publications, most employers should maintain consistent sectoral classifications across time within publications. As previously mentioned, this method should minimise discrepancies in the data caused by reclassifications and should more easily allow the tracking of job movements between sectors.

5 . Other adjustments to the data

Pay frequencies

Employments vary in how often employees are paid, with options including weekly, fortnightly, four-weekly, monthly, quarterly, or even annually or bi-annually.

Occasionally, the gap between payments for an employment differs from the stated pay frequency. This may reflect circumstances such as unpaid leave or other legitimate cases. This does not affect the operation of Pay As You Earn (PAYE) if payments have otherwise been recorded correctly. However, amending the pay frequency recorded in the statistical data to reflect the actual pay frequency observed in the data can improve the functioning of the calendarisation methodology.

We check if the time between payments matches the reported pay frequency. If it aligns for three consecutive periods, the pay frequency is adjusted accordingly.

Missed and double payments

For these statistics, "employment" is paid, filled employee jobs. A missed payment is interpreted as a missed period of work. In some instances, the data can be interpreted that a missed period represents an unusual payment situation, which does not correspond to a missed period of work.

Also, there may be double payments in the data. A missed payment is followed by a payment that is nearly double an employment's usual amount, followed by a subsequent payment which returns to near the usual amount. In this case, this pattern reflects a delayed payment instead of a period of missed work. The payment that is nearly double the normal amount can be used to cover two work periods.

For example, an employment that is usually paid in arrears is paid in advance for a single month. This would be processed as if the payment corresponds to two work periods instead of one.

Seasonal adjustment

All the key series within the PAYE Real Time Information (RTI) statistics publication are provided both non-seasonally adjusted and seasonally adjusted. The seasonally adjusted data aims to remove cyclical trends in the data because of seasonal patterns or other similar regular influences on the labour market.

Seasonal adjustment is performed using the X13-ARIMA method, in line with best practice. This involves applying bespoke models calculated from historic data by the ONS, with models for the main series being reviewed annually. The annual review is typically implemented in Autumn, but this could be subject to change. Where bespoke models are unavailable, the X13-ARIMA method is still utilised, but using a generic model.

6 . Additional considerations

Throughout the Pay As You Earn (PAYE) Real Time Information (RTI) statistics, we use median monthly pay to represent earnings. Median monthly pay shows what a person in the middle of all employees' pay distribution would earn each month. The median pay is generally considered to be a more accurate reflection of the "average wage" because it discounts the extremes at either end of the scale.

We also publish estimates of mean monthly pay in the supporting datasets, which is total pay divided by the number of employees. Mean pay can be more useful when trying to examine how average pay is affected by all of those within the population, including those at the higher and lower end of the pay distribution, rather than just the middle earner.

To provide further insight, the supporting datasets include estimates of the median of pay growth as well as monthly flows into and out of RTI. During periods of labour market instability, care needs to be taken when interpreting these statistics. Median pay growth can be influenced by the pay levels of those entering and leaving the labour market. So volatility in median pay growth may reflect unusual inflows and outflows in certain sectors.

7 . Access to data for accredited researchers

HM Revenue and Customs (HMRC) has a [Datalab](#), which allows approved researchers from contracted institutions to access de-identified data in an on-site secure environment. The Datalab facility is part of HMRC's support for the [National Data Strategy](#), which aims to use data securely, to encourage innovation and productivity across the UK, enhancing the delivery of public services to improve people's lives.

8 . Data on Real Time Information

[Earnings and employment from Pay As You Earn Real Time Information, non-seasonally adjusted](#)

Dataset | Released monthly

Earnings and employment statistics from Pay As You Earn (PAYE) Real Time Information (RTI), UK, NUTS 1, 2 and 3 areas and local authorities, monthly, non-seasonally adjusted. These are official statistics in development.

[Earnings and employment from Pay As You Earn Real Time Information, revision triangle](#)

Dataset | Released monthly

Revisions of earnings and employment statistics from Pay As You Earn (PAYE) Real Time Information (RTI), UK, monthly. These are official statistics in development.

[Earnings and employment from Pay As You Earn Real Time Information, seasonally adjusted](#)

Dataset | Released monthly

Earnings and employment statistics from Pay As You Earn (PAYE) Real Time Information (RTI), UK, NUTS 1, 2 and 3 areas and local authorities, monthly, seasonally adjusted. These are official statistics in development.

9 . Glossary

Earnings

Earnings are defined as the gross money people receive in return for payrolled work. This includes the following if paid through Pay As You Earn (PAYE):

- regular pay
- bonuses, overtime, shift premium
- allowances
- arrears
- employees on trainee or junior rates of pay
- employees whose earnings were affected by absence
- payrolled redundancy payments, signing-on fees, expenses

This excludes:

- self-employed income, which is reported through self-assessment
- non-payrolled stock options
- employer National Insurance contributions
- employer pension contributions
- non-payrolled benefits in kind
- earnings for members of schemes where no employee earns above the Lower Earnings Limit for National Insurance or has another job

Individual Payment Submission

When someone is paid through Pay As You Earn (PAYE) Real Time Information (RTI) (either for a job or an occupational pension), a record of that payment is submitted to HM Revenue and Customs (HMRC) by their employer or payroll operator. This record is called an Individual Payment Submission (IPS) and contains information on a variety of pay and employment variables.

Full Payment Submission

When an employer or pension provider makes a submission to HM Revenue and Customs (HMRC), the Individual Payment Submissions (IPs) are delivered "within" a Full Payment Submission (FPS).

National Insurance and PAYE Service (NPS)

This service is a system that maintains control over data related to National Insurance and Pay As You Earn. Our primary interaction with NPS is through a quarterly list of PAYE schemes that they provide, and which has employee or occupational pension counts.

National Minimum Wage and National Living Wage

The National Minimum Wage (NMW) is a minimum amount per hour that most workers in the UK are entitled to be payrolled. See [current and previous rates for the NMW and NLW](#) on the GOV.UK website.

Pay As You Earn Real Time Information

Pay As You Earn (PAYE) is the system employers and pension providers use to deduct Income Tax and National Insurance contributions before they pay wages or pensions to employees and pensioners.

Under PAYE RTI, employers and pension providers will tell HM Revenue and Customs (HMRC) about tax, National Insurance contributions (NICs) and other deductions when or before payments are made, instead of waiting until after the end of the tax year.

Self-Assessment

Income that is reported through Self-Assessment (SA), for example, income for self-employed work, is not included in our Pay As You Earn (PAYE) figures.

10 . Strengths and limitations

Strengths

The main strengths of using Real Time Information (RTI) admin data to estimate the level of employment and earnings include the following.

The data cover the whole of the payrolled population, rather than a sample of people or companies. This means that detailed breakdowns, which include cross-tabulations and granular breakdowns of the population, and sub-totals, can be calculated from the data without being affected by sample bias in the same way survey-based data would be affected. Therefore, this allows us to be able to do more in-depth analysis.

From the RTI data we use this advantage to produce estimates for geographic areas and other more detailed breakdowns of the population. It also allows additional analysis such as decomposing pay growth into components, tracking flows of employees over time, and conducting distributional analysis.

The statistics provide more timely estimates than alternative measures, as the data are received at or before the point that payments are made to employees. This allows the production of a "flash" estimate two to three weeks after the end of the reference period.

Limitations

Limitations of the RTI data and estimated statistics are the following.

Although the admin data have full coverage of the payrolled population, they do not include income for self-employed work or those who are remunerated through company dividends.

The industry data that are derived by the Inter-Departmental Business Register (IDBR) are based on company sector and not employment occupation. They only have Pay As You Earn (PAYE) data (so not Self-Assessment, savings, rental income and so on), which could have distributional limitations.

Our published statistics exclude occupational pensions, therefore income in that category is not included.

To provide the mostly timely estimates, RTI data used in each monthly release are extracted in the weeks following the end of the latest reference month. For some individuals this means payments relating to work done in recent reference months are yet to be received. Rather than wait until all payment returns have been received, we impute values for missing returns as explained in [Section 3: Imputation of missing payment submissions](#). This means for the latest reference months, around 15% of the data are imputed. This figure is referred to as the "flash" or "early" estimate in the statistics publication. This figure is subject to revisions as payment returns are received and the imputed data are replaced with real data, meaning the timeliest estimate is not always as accurate as later estimates.

Other estimates as well as the "flash" estimate are also subject to revisions as we can continue to receive payment submissions for the remainder of the tax year and can even see corrections to earlier tax years. However, we see lower levels of revisions to the non-flash months.

The seasonal adjustment models applied follow best practice, but this approach assumes any seasonal patterns remain broadly consistent over time. When seasonal patterns change in strength, this will be seen as greater volatility in the seasonally adjusted figures. This also appears in the revisions to the seasonally adjusted datasets.

11 . Related links

[Quality assurance of administrative data used in earnings and employment from PAYE RTI](#)

Article | Released 10 January 2025

Investigation of the administrative data used in production of ONS and HMRC earnings and employment from Pay As You Earn (PAYE) Real Time Information (RTI) data.

[Comparison of labour market data sources](#)

Article | Released 27 April 2022

The strengths and weaknesses of the main data sources we use to produce the labour market figures, including the advantages of new administrative data sources and limitations of some of our published figures.

[Earnings and employment from Pay As You Earn Real Time Information: UK](#)

Bulletin | Released monthly

Monthly estimates of payrolled employees and their pay from HM Revenue and Customs's (HMRC's) Pay As You Earn (PAYE) Real Time Information (RTI) data. This is a joint release between HMRC and the Office for National Statistics (ONS). These are official statistics in development.

[Payroll information to report to HMRC](#)

GOV.UK web page | Released 12 June 2014

Find out what to put in your Full Payment Submission (FPS) and Employer Payment Summary (EPS) if you're paying employees through PAYE.

[PAYE and payroll for employers: Introduction to PAYE](#)

GOV.UK web page

Guide for employers on PAYE.

12 . Cite this user guide

Office for National Statistics (ONS), released 25 July 2025, ONS website, user guide, [User guide to earnings and employment from Pay As You Earn Real Time Information](#)