

Article

Using administrative data to create headline labour market figures

Estimates created by applying administrative data to Labour Force Survey (LFS) base numbers, in response to data quality concerns with LFS estimates.

Contact:
Bob Watson
labour.market@ons.gov.uk
+44 1633 455400

Release date:
14 November 2023

Next release:
To be announced

Table of contents

1. [Main changes](#)
2. [Overview of Labour Force Survey data quality concerns](#)
3. [Overview of methods changes](#)
4. [Adjusted data](#)
5. [Future developments](#)
6. [Related links](#)
7. [Cite this article](#)

1 . Main changes

- Labour Force Survey estimates for April to June 2023 have been used as the base levels to model from moving forward.
- Employment level estimates are modelled by applying implied growths from the Pay As You Earn (PAYE) Real Time Indicator (RTI) series.
- Unemployment level estimates are modelled by applying implied growths from the Claimant Count series.
- Economic inactivity level estimates are calculated as a residual, to sum to the estimated population.

2 . Overview of Labour Force Survey data quality concerns

Since the start of the coronavirus (COVID-19) pandemic, the Labour Force Survey (LFS) has changed collection mode to a telephone-first survey, rather than face-to-face. In addition to this, response rates to the survey have been falling over time, like many other social surveys operated by comparable national statistical institutions around the world. To counter these impacts, we have brought in changes to our weighting to deal with some of the most obvious effects, although other biases in the achieved sample are still likely to exist.

As part of the work to transform the LFS, the Office for National Statistics (ONS) has needed to run both the existing LFS and transformed surveys in parallel, and more generally balance resources across its wider survey portfolio. To sufficiently balance these resources, we reversed a sample boost to the LFS that was introduced during the coronavirus pandemic from July 2023.

Initial assessments suggested that this would have a minimal impact on the existing LFS achieved sample. However, the resulting reduction in the achieved sample size for the survey was larger than anticipated. While results for May to July 2023 were partially affected by having one month with a reduced sample size, the impact became much clearer when the June to August 2023 results became available, with two months of reduced sample size.

The combination of these changes resulted in survey estimates that looked implausible in the context of wider labour market and economic intelligence. This led to the suspension of the LFS estimates due to be published in October 2023.

To give users a more considered view of the labour market, we therefore worked towards developing an alternative method for modelling the headline estimates, taking account of available administrative data, for periods after April to June 2023.

3 . Overview of methods changes

Consideration of sources

There are relatively few other sources relevant to the labour market that we could use to estimate movements.

The transformed Labour Force Survey (LFS) is still under development, with further improvements currently being deployed and tested. Therefore, it is not possible to produce ongoing consistent headline indicators.

The Pay As You Earn (PAYE) Real Time Indicator (RTI) experimental series is an administrative measure of all payrolled employees, so is a robust measure of employees. However, it does not cover self-employment. Additionally, initial “flash estimates” are still subject to revisions as they are based on early data. For this reason, “non-flash” data were used because they are subject to far less revisions.

Workforce jobs is a measure of the number of jobs and not the number of people employed. It is not timely enough for use in modelling LFS results, and only produces quarterly estimates rather than monthly or rolling three-month periods.

The Claimant Count (CC) is an administrative measure of people claiming benefits for unemployment-related purposes. It has been experimental for a number of years. Initially, this was because the move to Universal Credit introduced many extra people into the CC, resulting in changes that were driven mainly by the administrative process rather than economic conditions. Later, during the coronavirus (COVID-19) pandemic, the benefit system was used as a way of delivering government support.

These two effects have now largely been reversed, although there is still some planned future migration to Universal Credit that could have some impact. We recognise that the current CC contains far more claimants in work than would be ideal, but movements in the CC should be far more representative of movements in unemployment than has been the case during the preceding decade.

Based on these options we looked at the suitability of using PAYE RTI as a proxy for employment, and CC as a proxy for unemployment. When compared in historical situations, both provided reasonable representation of the trends in those measures and were timely enough for use in our modelling.

Methods

The basic method used applies growth rates from the administrative data to starting levels taken from the LFS.

Using monthly administrative data, a series based on three-month averages has been calculated. Growth rates have then been calculated between consecutive overlapping three-month averages. These growth rates have been applied to the LFS seasonally adjusted data to produce a new series of overlapping three-month estimates.

For PAYE RTI, the seasonally adjusted series used were the age band employees estimates, where the age bands closely match those from the LFS. The PAYE RTI age band for those aged under 18 years has been used as a proxy for those aged 16 to 17 years, although it will contain a small number of employees aged under 16 years. These series are not published broken down by sex, therefore figures represent the total for all people.

For published seasonally adjusted series, the CC age bands are wider than LFS age bands, but they share largely consistent boundaries. Therefore, the CC age band for those aged 18 to 24 years has been used for LFS age bands for those aged 16 to 17 years and aged 18 to 24 years, and so on.

The last period of fixed LFS data that these growths are applied to is April to June 2023. Therefore, experimental figures for periods from May to July 2023 onwards have been calculated.

Although the LFS usually works with non-overlapping three-month periods, overlapping growths have been used in this case. This is because multiplying the three consecutive different non-overlapping periods all by factors close to 1 would have perpetuated any sampling variations in those three periods across all subsequent estimates.

Data have been calculated by applying PAYE RTI growths to employment levels for all age bands, and CC growths to unemployment levels for all age bands. Economic inactivity has been calculated as a residual, to sum to the estimated population.

This method has been expanded to estimates for the countries and regions of the UK in our [Labour market overview, UK: November 2023 bulletin](#). Unlike for the UK level, seasonally adjusted PAYE RTI and CC are not available for these geographies broken down by age band. Therefore, a single adjustment has been applied to both the aged 16 years and over age group, and the aged 16 to 64 years age group. The experimental series for each of the countries and regions have not been constrained back to the UK figures and so will not be wholly additive to UK totals.

Known weaknesses

It should be noted that by applying growths to existing LFS estimates, this approach does not address any concerns users may have about the level of the LFS population feeding into estimates. However, this approach of linking growths directly to existing levels does maintain the comparability of the timeseries.

The method assumes that growths in the administrative sources are reasonable indicators of growth in the headline series.

For PAYE RTI, the administrative source only reflects payrolled employees. If the growth in self-employment differs significantly from employee growth, then this divergence would not be reflected in the experimental estimates. Historically, self-employment has fluctuated between around 12 and 15 percent of total employment.

For CC, the administrative data are known to contain significant numbers of people who would not meet the definition of unemployment. It is also more prone to influence from changes in administrative processes and eligibility criteria than PAYE RTI. However, movements in CC are far more representative of movements in unemployment than has been the case during the preceding decade. In addition, the intention is only to use CC as a proxy for a short period, so there is low risk of changes in the administrative process being a problematic factor.

4 . Adjusted data

[X10: Adjusted employment, unemployment, and economic inactivity](#)

Dataset | Released 14 November 2023

Experimental labour market estimates using administrative data to produce adjusted employment, unemployment and economic inactivity measures for UK countries and regions, seasonally adjusted.

5 . Future developments

We intend for this method for producing experimental headline estimates to be required only for a short interim period, so there are no plans to develop the method further. However, we will be open to considering whether any suggestions on enhancing the method would be appropriate and practical.

6 . Related links

[Labour Force Survey performance and quality monitoring reports](#)

Methodology | Released quarterly

Response rates, sample size and quality assessment of the Office for National Statistics quarterly Labour Force Survey.

7 . Cite this article

Office for National Statistics (ONS), released 14 November 2023, ONS website, [Using administrative data to create headline labour market figures](#)