Compositionally adjusted labour input (CALI) QMI

Quality and methodology information for compositionally adjusted labour input (CALI), detailing the strengths and limitations of the data, methods used, and data uses and users.

Contact:
Productivity team
productivity@ons.gov.uk
+44 1633 582563

Release date:
27 May 2025

Next release:
To be announced

Table of contents

1 . Output information

- Statistical designation: official statistics in development
- Data sources: Labour Force Survey (LFS), Annual Survey of Hours and Earnings (ASHE), and other surveys
- Data collection: combined from survey data from households and businesses
- Frequency: quarterly and annual
- Geographic coverage: UK
- Related publications: Annual multi-factor productivity, market sector, UK: Quarter 4 2024 bulletin; Multi-factor productivity estimates, UK articles

2 . About this quality and methodology information report

This quality and methodology information (QMI) report contains information on the quality characteristics of the data (including the European Statistical System's five dimensions of quality) as well as the methods used to create it.

The information in this report will help you to:

- understand the strengths and limitations of the data
- learn about existing uses and users of the data
- understand the methods used to create the data
- help you to decide suitable uses for the data
- reduce the risk of misusing the data

3 . Important points

- Compositionally adjusted labour input (CALI) was previously referred to as quality-adjusted labour input (QALI); CALI accounts for changes in the composition (or "quality") of the employed workforce and changes in hours worked.
- CALI weights hours worked by different types of workers by their relative income share; this reflects their contribution to economic production.
- The two main sources of data for CALI are our Labour Force Survey (LFS) and our Annual Survey of Hours and Earnings (ASHE).
- CALI is an input into estimating multi-factor productivity (MFP).

4 . Quality summary

Compositionally adjusted labour input (CALI) measures labour input by accounting for differences in workers' skills, experience, and productive capacity. CALI adjusts for workforce composition, unlike traditional metrics like total hours worked. It recognises that an hour worked by a highly skilled worker contributes more to output than an hour by a less-skilled worker.

Within the calculation, we categorise workers by identifiable characteristics such as age, gender, industry of employment, and level of education. We weight the changes in the hours worked of each worker type by their share of total labour income.

The rationale for this approach is that economic theory suggests that different factors of production (different categories of workers and different types of capital assets) will be paid according to their productivity under competitive markets. As a result, relative shares of labour income provide a proxy for the relative productivity or "quality" of different types of workers.

We first calculate the quarterly and averaged four quarters (annual) series of hours and earnings by age, gender, industry of employment, and level of education from the Labour Force Survey (LFS). We then benchmark total hours and earnings against those from the Annual Survey of Hours and Earnings (ASHE) for each worker type. This benchmarking improves accuracy, as the estimates rely on two data sources. Hours are weighted by their share of total income; weights are based on the geometric mean over two consecutive periods. This allows us to construct Törnqvist indexes for hours worked by each group, as described in Ulrich Kohli's Trading Gains and Productivity paper (PDF, 774KB). We compute both quarterly and annual estimates at the whole economy and the market sector. We then publish CALI indices for the different age groups, industry sectors, genders, and education groups.

For a further explanation of the underlying framework of CALI, please see the Section 3: Quality-adjusted labour input (now known as compositionally adjusted labour input) in our A simple guide to multi-factor productivity.

Uses and users of CALI

CALI provides a volume series of labour input that can be used in a growth accounting framework, such as for multi-factor productivity (MFP) estimates. CALI estimates measure changes in the volume of services from labour, rather than just changes in hours worked.

We provide an example of the informative nature of CALI estimates in our Analysis of compositional changes in hours worked in the UK article. We discuss how the growth in hours worked following the 2008 economic downturn until the coronavirus (COVID-19) pandemic was disproportionately caused by workers who are 50 years or older, and by workers with degrees.

CALI is used as an input into the MFP framework and the volume index of capital services (VICS). This allows changes in economic output to be decomposed into changes in services from labour, changes in services from capital, and for the unaccounted growth to be multi-factor productivity; this is also known as "disembodied technical change" or the "Solow residual".

Strengths and limitations of CALI

There is no universally agreed method of accounting for different types of labour input into production. The method used in CALI is subject to continuous development. The theoretical framework behind CALI is based on the concept of services flowing from labour and is therefore open to criticism.

CALI is computed according to a methodology under review by the Organisation for Economic Co-operation and Development (OECD). It allows a breakdown of the labour input by different characteristics. This makes CALI informative at a multidimensional level.

A recent OECD review of 21 countries supports the composition choice we are currently using for CALI, as discussed in their The composition of labour input working paper. However, the OECD review demonstrates low explanatory power from industrial components, which are headline estimates for both CALI and MFP. Different components, definitions and data sources makes international comparison difficult.

Granularity

CALI uses estimates of employment, actual hours worked, and pay at low levels from the quarterly LFS. LFS response rates have fallen since the pandemic, which has decreased the responding cell counts for some industries, education, genders, and age groups. Improvements to the LFS are ongoing, as discussed in our latest LFS quality update: May 2025 article.

The reduction in LFS response rates increases CALI estimate variability. However, it is not uniform. Proportionally, CALI categories with small counts are taking longer to recover to pre-pandemic levels, compared with CALI categories where response was higher.

LFS response rates have improved with the sample boost from Quarter 1 (Jan to Mar) 2023. Annual CALI estimates use a four-quarter average of components, which provides more reliable, but less timely estimates of labour composition. Improvements were made to the LFS highest qualification question from Quarter 1 2022, as described in our LFS User Guide Volume 4a - Updates to the LFS Education Derived Variables (PDF, 474KB). This makes pre- and post-pandemic CALI comparisons difficult to interpret.

The CALI theoretical framework relies on the economic assumption that workers receive their marginal products. This means that earnings can be used to accurately represent different workers' relative contributions (or their marginal additions) to output (the product). This may not always hold true. In non-market sector industries like non-profit institutions or the public sector for example, this assumption may be broken because there are no market prices for non-market outputs.

Because of this theoretical difference, users should not try to calculate non-market sector estimates as a residual between the estimates for the whole economy and those for the market sector. However, whole economy and market sector estimates are comparable for industries that are wholly market sector.

Another point where this economic assumption may not hold is around the gender variable. If the persistent pay gap that exists between males and females reflects discrimination, then the assumption that workers are paid their marginal product is violated. This results in hours growth being weighted incorrectly and the quality adjustment will carry a downward bias.

CALI estimates are still labelled as official statistics in development, given the innovative nature of growth accounting. They therefore are not accredited official statistics. Official statistics in development are newly developed or innovative statistics. They are published so that users and stakeholders can be involved in the assessment of their suitability and quality at an early stage.

CALI provides additional information to our Labour productivity statistics. The main strength of CALI is that it can account for different types of workers. Because it can differentiate between hours worked within the framework of growth accounting, the labour index is more informative and precise.

For a detailed explanation of the CALI framework and its more informative nature when compared with other methods, please refer to our A simple guide to multi-factor productivity.

We have calculated market sector CALI estimates from a basis of 19 industries since April 2018. This was done to address the limited level of detail that CALI estimates were providing and to make them more informative at the industry level.

5 . Quality characteristics of compositionally adjusted labour input data

This section provides a range of information that describes the quality and characteristics of the data and identifies issues that should be noted when using the output.

Relevance

Relevance is the degree to which statistical outputs meet users' needs.

External researchers, including academics, non-governmental bodies, and other government departments have used compositionally adjusted labour input (CALI) data as an input into their own growth accounting estimates. Other researchers have used CALI for broader purposes, such as analysis of the productivity puzzle by using CALI for labour productivity estimates.

We work alongside the Economic Statistics Centre of Excellence (ESCOE) to make sure our estimates and the methods to obtain them are at the frontier of economics and statistics research.

CALI estimates can also give insight on the productivity puzzle. Labour composition was the only input to multi-factor productivity (MFP) to show a positive contribution to growth between the 2008 economic downturn and the economic downturn resulting from the coronavirus (COVID-19) pandemic. For a better overview of the possible uses of CALI within the analytical framework of productivity growth, please see Section 3: Quality-adjusted labour input (now known as compositionally adjusted labour input) section of our MFP estimates, UK: July to September 2019 article.

Accuracy and reliability

Accuracy

Accuracy is the degree of closeness between an estimate and the true value.

CALI estimates represent changes in an intangible concept, so their accuracy is somewhat unobservable. However, we can evaluate the degree of accuracy for CALI by comparing with labour productivity estimates, where we assess whether CALI and labour productivity are coherent and provide similar economic insights. CALI and its use in a growth accounting framework are subject to various theoretical assumptions about the operation of employment markets and corporate behaviour.

The degree of accuracy of CALI has increased with improved methodologies and data sources. For example, we have implemented the benchmarking process of the Labour Force Survey (LFS) to the Annual Survey of Hours and Earnings (ASHE). These are two very different surveys that are addressed to different respondent categories.

The LFS is the largest survey of households in the UK, and ASHE collects data from UK employers. The LFS provides information about personal characteristics of the labour force, such as educational attainment and the occurrence of multiple job holding. These are missing in ASHE. Because of the importance of self-employment and zero hours contracts in the UK economy, we believe that the LFS provides the best estimate of overall labour input into the economy. The LFS is also more reliable for variables such as age and actual hours worked.

In household surveys, individuals self-define the industries in which they work. Analysis shows that household surveys are less reliable for detailed industry analysis than business surveys. ASHE data avoids this problem of inaccurate industry categorisation and provides more accurate data on pay (such as employers' contributions) and contracted hours.

For a more detailed explanation on the quality of LFS and ASHE data, and the methodologies used for their collection, please refer to our LFS quality and methodology information (QMI) and our ASHE QMI.

Combining data from these two data sources increases the accuracy of the resulting estimates. Using ASHE data alongside LFS data also helps achieve more granular industry splits because of the larger datasets. For a more thorough explanation on the implementation of ASHE data for CALI estimates, please refer to our Developing improved estimates of quality adjusted labour inputs using the ASHE: a progress report article.

Reliability

Reliability is the closeness of the initial value to the subsequent estimated measure.

In terms of reliability, CALI estimates are normally subject to minor variations because of revisions in the data sources. Users should keep in mind that such small variations are expected, especially during the first and the third quarters when ASHE and LFS data are published.

Changes to benchmarks used can also cause revisions to published series. We reference and explain any major revisions or reclassifications brought about by national accounts changes (normally with our annual Blue Book publications that present a full set of economic accounts for the UK) or from other sources in our MFP estimates, UK articles.

We have various procedures in place to ensure that errors are minimised. The data go through thorough quality assurance processes, and outputs are peer reviewed before being published. If errors are found in the data after publication, a notice will be attached to the publication to inform users and datasets will be revised in line with the Code of Practice for Statistics.

Coherence and comparability

Coherence is the degree to which data that are derived from different sources or methods, but refer to the same topic, are similar. Comparability is the degree to which data can be compared over time and domain, for example geographic level.

CALI is coherent with our other productivity statistics. This is because the sources used to aggregate CALI hours and earnings estimates are the same as those used elsewhere in productivity systems. For example, productivity hours, which are used to scale the hours results from the CALI systems to industry totals, are used in our Productivity flash estimate and overview, UK articles.

Concepts and definitions

Concepts and definitions describe the legislation governing the output, and a description of the classifications used in the output.

CALI concepts and definitions are produced in line with international standards, most notably those laid down by the Organisation for Economic Co-operation and Development (OECD).

Accessibility and clarity

Accessibility is the ease with which users can access the data, also reflecting the format in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the release details, illustrations and accompanying advice.

We publish CALI estimates each November as a statistical release. We also publish CALI tables as part of our MFP article series. All datasets are available on our website.

Our recommended format for accessible content is a combination of HTML web pages for narrative, charts and graphs, with data provided in usable formats such as CSV and Excel. Our website also offers users the option of downloading the text in PDF format. In some instances, other software may be used throughout the computation of the estimates, and such files may be available on request.

For further information please email productivity@ons.gov.uk.

Timeliness and punctuality

Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the gap between planned and actual publication dates.

The main constraint of CALI production is the publication of the GDP quarterly national accounts bulletins, from which the CALI income constraints are constructed. As a result, CALI is usually published one week after the publication of the second quarter national accounts and around 14 weeks after the reference quarter.

For more details on related releases, our Release calendar is available and provides 12 months' advance notice of release dates. In the unlikely event of a change to the pre-announced release schedule, we will publicly announce and explain the reasons for the change, as set out in the Code of Practice for Statistics.

6 . Methods used to produce compositionally adjusted labour input data

How we collect the data, main data sources, and accuracy

The two main data sources for compositionally adjusted labour input (CALI) are the Labour Force Survey (LFS) and the Annual Survey of Hours and Earnings (ASHE).

The LFS sample is made up of approximately 30,000 responding UK households containing about 50,000 persons in the 16 to 74 years age group (2024 quarterly dataset). Respondents are interviewed for five successive waves at three-monthly intervals, and 20% of the sample is replaced every quarter.

The ASHE is based on a 1% sample of employee jobs (around 180,000) taken from HM Revenue and Customs's (HMRC's) Pay As You Earn (PAYE) records. Information on earnings and hours is obtained from employers and treated confidentially.

The LFS provides us with quarterly datasets for hours and corresponding earnings of different worker types. LFS data are then benchmarked to ASHE estimates both to add an extra level of reliability on LFS-derived estimates and to allow for a higher level of detail. This is particularly important for pay estimates, because employers are more reliable in reporting gross pay figures.

Hours estimates are further benchmarked to productivity hours, which are computed from a combination of LFS and Short-Term Employment Survey data. The final source used in the process is the national accounts labour income data, which form the income constraints on employee and self-employed compensation levels. These are derived from our UK National Accounts, The Blue Book income data and quarterly national accounts data.

How we process the data

First stage: construction of Labour Force Survey hours and earnings time series

We code the raw datasets to stratify the workforce in four ways. These combinations result in 684 categories. These are three age groups, two genders, six educational classifications, and 19 industry groups:

Gender

- female
- male

Age groups

- 16 to 29 years
- 30 to 49 years
- 50 years and over

Educational attainment

- higher degrees level 7 and 8 qualifications, or equivalent
- undergraduate and other degrees, or equivalent
- higher education
- A levels or trade apprenticeships, or equivalent
- GCSEs or equivalent
- no qualifications

Industry

- A
- B
- C
- D
- E
- F
- G
- H
- I
- J
- K
- L
- M
- N
- O
- P
- Q
- R
- STU

Industry classification is indicated according to the UK Standard Industrial Classification (SIC) 2007. Average hours and earnings are computed for each of these categories in each time period and formatted as time series. Separate series are produced for first and second jobs, and for employees and self-employed.

Labour characteristics

The choice of labour characteristics involves a trade-off between comprehensiveness and data availability. None of the following characteristics represents labour quality in and of themselves, but only as dimensions of the income-share weights.

Age

Age is included as a proxy for work experience. It is not a complete measure because it does not account for periods of unemployment or inactivity. The assumption is that older workers tend be more productive, because they are more experienced, and that they therefore receive more compensation for their labour.

However, it has been suggested that younger workers may be more dynamic and innovative than their older counterparts, as discussed in the Bank of England's A quality-adjusted labour input series for the UK working paper. If this is true in some cases, these workers will be paid their marginal product, and growth in hours will be weighted accordingly, if labour markets are competitive.

Gender

Gender is chosen because of the persistent pay differential that exists between males and females, even after holding other factors constant. It is not a element of quality change itself. However, it may represent hidden characteristics, such as an increased tendency for women to take career breaks for childcare or to fill part-time posts that are not as well paid.

This complements, or improves, the use of age as a proxy for work experience. It also helps to explain the pay differential. However, if the pay differential is a result of discrimination, then the assumption that workers are paid their marginal product is violated. This results in hours growth being weighted incorrectly and the quality adjustment will carry a downward bias.

Education

Education is measured as the highest qualification attained. It is used as a proxy for skills. Qualifications either indicate a person's ability to employers, or provide the knowledge for specific job requirements. This characteristic is the main component of the index. There is an increasing prevalence of higher degrees and their National Vocational Qualification (NVQ) level 7/8 equivalents. Higher degrees are also increasingly associated with higher pay, so they are included as a separate category.

Industry

Industry is mainly included so we can observe industry trends. We also use of ALI in industry-level multi-factor productivity (MFP). This category also helps to capture inherent differences in skills and productivity trends that exist between industries.

The industry categories chosen are broad. This is partly because respondents self-report the industry in which they work in the Labour Force Survey (LFS), making this variable less reliable. Since incorporating the Annual Survey of Hours and Earnings (ASHE) dataset into CALI estimates, we have been moving to more granular analysis at the 19-industry level.

For the 19-industry split, we publish log changes (and therefore growth rates) of CALI, hours, and labour composition indices on a quarterly basis. However, at the 19-industry level we only publish annual estimates for relative hourly remuneration and shares of hours worked.

Second stage: constraining total hours and earnings

CALI is consistent with other data series, particularly the national accounts, and our headline productivity measures. It is important for its use in growth accounting. To achieve this external consistency, various components of CALI are scaled and constrained to our aggregates. Specifically, we scale:

- gross weekly pay to national accounts industry-level "labour income", which is compensation of employees plus the share of labour income from self-employed
- both hours worked and jobs to industry level, so they are consistent with estimates used in labour productivity

For each industry, we add compensation of employees (CoE) data to an estimated value of self-employed labour income from self-employed mixed income, which is taken from quarterly national accounts data. Self-employed labour income is derived by assuming that in mixed income, the ratio between the return on capital and the return on labour is the same as the relationship between CoE and gross operating surplus. This produces a component that can be compiled with CoE to give total labour income by industry for scaling.

Third stage: computation of indices

The CALI indices are computed using a Törnqvist index. The Törnqvist index uses data from the current and previous period. Weighting the change in hours makes them more current or representative measures.

The index is widely used in economic analysis, particularly in quality-adjusted labour measures, as described in the Bank of England's A quality-adjusted labour input series for the UK working paper.

Hours worked by the different worker types contribute to total labour input L through a function g.

$$L = g(h_1, h_2, \ldots, h_n)$$

Then, for each worker type, income weights are calculated as an average of the current and preceding period share in total income

$$\Delta \ln L_t = \ln\left(\frac{L_t}{L_{t-1}}\right) = \sum_i \left[\left(\frac{w_{i,t} + w_{i,t-1}}{2}\right) \ln\left(\frac{h_{i,t}}{h_{i,t-1}}\right)\right]$$

where $w_{i,t}$ is the share of total labour income paid to group i in period t, while $h_{i,t}$ is the share of total hours worked by group i in period t. The weight used is the average of $w_{i,t}$ and $w_{i,t-1}$, and the income shares sum to one. The logarithms of the changes are then used to create an index.

Log changes are used in CALI calculations because they are additive, which makes it possible to calculate contributions to growth. Calculating quarter-on-quarter natural log changes also means that the magnitude of change remains the same, regardless of which way the quarters are compared. The only thing that changes is the sign.

The difference between the Törnqvist index and an unadjusted index of hours is referred to as "labour composition".

An example for calculating CALI estimates and more information on log changes can be found in our A simple guide to multi-factor productivity.

How we analyse the data

Our Multi-factor Productivity team, within our productivity branch, analyse data on a continuous basis. The focus is to monitor the variation of the CALI index for the different variables and their relations throughout time. We do this to make sure that any important change has a coherent interpretation, in terms of concepts from economic theory. We also conduct ad hoc analysis for specific categories.

Our Multi-factor productivity estimates articles always include one section on the CALI index and labour composition analysis.

7 . Other information

User needs

We invite user feedback on CALI estimates to productivity@ons.gov.uk. Any proposals or comments from users will be considered when deciding on future development work on CALI outputs.

8 . Cite this methodology

Office for National Statistics (ONS), released 27 May 2025, ONS website, quality and methodology information report, Compositionally adjusted labour input (CALI) QMI