

Income estimates for small areas QMI

Background information and methodology for income estimates for small areas, England and Wales, detailing the strengths and limitations of the data, methods used, and data uses and users.

Contact:
Financial Well-being team
Economic.Wellbeing@ons.gov.
uk
+44 1329 447767

Release date:
25 June 2024

Next release:
To be announced

Table of contents

1. [Output information](#)
2. [About this QMI report](#)
3. [Important points](#)
4. [Quality summary](#)
5. [Quality characteristics of the data](#)
6. [Methods used to produce the data](#)
7. [Other information](#)
8. [Related links](#)
9. [Cite this methodology](#)

1 . Output information

- Accredited official statistic: yes
- Data collection: survey data from the Department for Work and Pensions Family Resources Survey (FRS) and administrative data from HM Revenue and Customs and the Department for Work and Pensions
- Frequency: every two years
- How compiled: survey data modelled using administrative data
- Geographic coverage: Middle-layer Super Output Areas (MSOAs) in England and Wales
- Related publications: Income estimates for small areas bulletin and technical note, financial year ending 2020
- Related publications: [Income estimates for small areas, England and Wales](#) and [Income estimates for small areas in England and Wales, technical report](#)

2 . About this QMI report

This quality and methodology report contains information on the quality characteristics of the data (including the [European Statistical System 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 data

3 . Important points

- Middle-layer Super Output Area (MSOA)-level estimates of income (unequalised total annual household income, unequalised net annual household income before household costs, and equalised net annual household income before and after housing costs) are produced by modelling survey variables using administrative data.
- Responses across England and Wales from the Department for Work and Pensions (DWP) Family Resources Survey (FRS) form the basis of the income estimates for small areas.
- As around two-thirds of MSOAs in England and Wales will contain no FRS responses, it is not possible to create MSOA-level estimates directly from the surveys, and a statistical model is created to estimate mean incomes dependent on available (administrative) information about each MSOA.
- A regression-based model is created for each of the four income measures, with random effects to account for the sampling structure using the log of FRS-based incomes as the dependent variable and known MSOA-level administrative data.
- A comprehensive set of model and data validation processes are carried out to ensure the robustness of the models and the accuracy of the estimates produced from these.

4 . Quality summary

Overview

[Income estimates for small areas, England and Wales](#) (SAIE) are [accredited official statistics](#) released around every two years since the mid-2000s. The statistics include estimates for each Middle-layer Super Output Area (MSOA) in England and Wales of:

- total annual household income (unequalised)
- net annual household income before housing costs (unequalised)
- net annual household income before housing costs (equalised)
- net annual household income after housing costs (equalised)

The source income data for SAIEs comes from the Department for Work and Pensions (DWP) [Family Resources Survey: background information and methodology](#) and its [accredited official statistics](#) publication [Households below average income \(HBAI\) statistics](#). The first of the listed income measures, total annual household income, comes from the FRS while the remaining three measures come from HBAI.

Although the FRS is a large and representative survey, the overall sample of approximately 14,000 responses in England and Wales is insufficient to provide accurate income results below regional International Territorial Level (ITL1) level. Approximately two-thirds of all of the 7,201 MSOAs in England and Wales will contain no survey responses.

A single regression-based statistical model covering the whole of England and Wales is instead created to provide "synthetic estimates" of local income, based on information that is known about each MSOA, such as deprivation levels, the proportion of individuals receiving benefits, and house prices. The results reflect our best estimate of incomes at small geographic areas.

Uses and users

There has been extensive interest in income data produced at the very local geographical level from a broad range of users including central government, local authorities, academia, the commercial sector and independent policy researchers. Users are keen to identify deprived and disadvantaged communities for the development and evaluation of policy (particularly related to the coronavirus (COVID-19) response, the cost of living and poverty, and differences between local areas) information for practitioners and geographic profiling.

Strengths and limitations

The main strengths of the output include:

- the output and release are well established, used and recognised by a broad range of stakeholders as an accredited official statistic in the area of incomes and for informing policy
- there are no other sources of government statistics of household incomes at the very localised level, and developed based on distributional data
- the methodology is robust, with extensive model validation and the results calibrated to published survey-based statistics making best use of their underlying survey data
- incomes are predicted from a wide range of explanatory variables from different sources, which explain 85% to 90% of variability in income levels between MSOAs

The main limitations include:

- levels of precision of the income estimates for most MSOAs are fairly wide; the median [coefficient of variation](#) for the net income after housing costs model is 5.7%
- the model is optimised to produce mean incomes for each MSOA and does not support the generation of other distributional measures including median incomes, which are more commonly reported
- the model does not support the disaggregation of incomes below MSOA level, nor the aggregation to any level (including to local authorities) apart from ITL1 region or nation
- caution should be applied when interpreting trends over time as the methodology used to produce the estimates is optimised for point-in-time estimates (with separate models containing separate subsets of significant independent variables) and not for estimating change

5 . Quality characteristics of the data

This report provides a range of information that describes the quality of the data and identifies the issues that should be noted when using the output. We have developed [guidelines for measuring statistical quality](#) based on the European Statistical System's five dimensions of quality. This report addresses the quality dimensions and important quality characteristics, which are:

- relevance
- accuracy, reliability and output quality
- coherence and comparability
- concepts and definitions
- geography
- accessibility and clarity
- timeliness and punctuality

More information is provided about these quality dimensions in the following sections.

Relevance

Income estimates for small areas (SAIE) involve a robust and objective methodology for mean household incomes across each Middle-layer Super Output Area (MSOA). Equivalised and unequivalised mean incomes are produced for a range of income measures, including gross income, incomes net of taxes and social contributions, and incomes net of taxes and social contributions and housing costs.

As badged [accredited official statistics](#), these are the main source of estimates of household incomes at the very localised level and inform local and subnational policy. They are of substantial topical interest, particularly given coronavirus (COVID-19) and the rise in the cost of living.

The SAIE method is well established; having been first implemented by the Office for National Statistics (ONS) in 2005, enabling broad comparisons between local areas. The SAIE findings align with those from the widely-used [Department for Work and Pensions \(DWP\)](#) Family Resources Survey (FRS) and Households below average income (HBAI) publications.

Accuracy, reliability and output quality

The SAIE model and results are underpinned by a robust, large-scale official survey; the FRS, which in turn is based on a large, nationally representative sample. To ensure that its findings represent the population of residents and households, steps are taken with both the issued and the achieved FRS sample, to ensure that the households within the sample are representative of the country as a whole. This includes stratification of the issued sample (set out in detail later) and also the weighting of the achieved sample by several control factors, including age and Council Tax band.

The SAIE methodology itself is robust and an extensive suite of model validation processes including simulation techniques are employed and the results are calibrated to published FRS survey-based statistics at regional level. Household incomes are predicted from a wide range of explanatory variables from different sources and these explain a high proportion (85% to 90%) of the variability in income levels between MSOAs.

Accurate measures of precision ([confidence intervals](#)) are provided for each income estimate for each MSOA to enable users to determine whether geographical or temporal differences in incomes are likely to reflect underlying differences that go beyond those because of [sampling error](#).

Coherence and comparability

SAIE are disseminated at MSOA level, while the DWP FRS and HBAI statistics are published at regional level (and for inner and outer London). As SAIEs are calibrated to the FRS and HBAI estimates at region and nation level, they fully align to these estimates.

The ONS's [gross disposable household income \(GDHI\)](#) also presents local area incomes within the system of the national accounts framework. They will not be exactly comparable with SAIE, because unlike SAIE, GDHI includes intangible "household sector" components of income such as assets produced by households for households. Further details are available in [Income statistics: coherence and comparison information \(PDF, 324KB\)](#).

The MSOA-level income estimates are subject to statistical variation arising from the sampling and the modelling process and any observed year-on-year change in an estimate may not necessarily represent a true change in income across all households within an MSOA. This is because the methodology used to produce the estimates is optimised for point-in-time estimates (with separate models containing separate subsets of significant independent variables) and not for estimating change. [Confidence intervals](#) are presented with each estimate and large year-on-year differences in comparison with the widths of their respective confidence intervals are generally indicative of [statistically significant](#) changes.

Accessibility and clarity

Our recommended format for accessible content is a combination of HTML webpages for narrative, charts and graphs, with data being provided in usable formats such as CSV, XML and Excel. The ONS website also offers users the option to download the narrative in PDF format. In some instances, other software may be used or may be available on request. Available formats for content published on our website, but not produced by us or referenced on our website but stored elsewhere, may vary. For information regarding conditions of access to data, please refer to the following:

- [terms and conditions \(for data on the website\)](#)
- [accessibility](#)

Timeliness and punctuality

The Department for Work and Pensions (DWP) published the Family Resources Survey (FRS) and Households below average incomes (HBAI) findings for the financial year ending 2020 on 25 March 2021, 12 months after the end of the reference period. Income estimates for small areas were published on 11 October 2023 reflecting the complexity of the modelling process and availability of the data used as covariates in the model.

For more details on related releases, the [release calendar](#) is available online and provides 12 months advance notice of release dates. If there are any changes to the pre-announced release schedule, public attention will be drawn to the change alongside full explanation of the reasoning behind it, as set out in the [Code of Practice for Statistics](#). This itself has been recently updated, with a greater focus on statistical context and recommended usage.

Concepts and definitions (including list of changes to definitions)

Local area

Local areas within this report refer to areas called [Middle layer Super Output Areas \(MSOAs\)](#). MSOAs have a mean population of 7,200 and a minimum population of 5,000. They are built from groups of [Lower layer Super Output Areas \(LSOAs\)](#) and constrained by the local authority boundaries used for 2011 Census outputs. For consistency with previous publications, incomes for 2020 have been modelled on 2011 MSOA boundaries, rather than the more recent one presented in the 2021 Census outputs.

Average (mean) income

The average (mean) income is the equivalent of adding every household income together and dividing by the number of households.

Disposable (net) household income

The sum of the disposable (net) income of every member of the household, that is, all income (from wages and salaries, self-employment, pensions, investments, benefits) minus Income Tax, National Insurance, Council Tax, maintenance or child payments deducted through pay, and contributions to occupational pensions.

Equivalised

Equivalised income considers the household size and composition and makes it easier to compare income across households. It acknowledges that, for example, two people do not need double the income of one person to have the same living standards. Like other Office for National Statistics (ONS) equivalised income data, these estimates use the Organisation for Economic Co-operation and Development (OECD) equivalisation scale.

Confidence intervals

This represents a range of values that a measure can take, based on statistical uncertainty and the fact that the data were derived from a sample of households across the country. For further details, please see our [Uncertainty and how we measure it for our surveys guide](#).

For more definitions and concepts, please refer to our [Income estimates for small areas in England and Wales, technical report: financial year ending 2020](#) and [Income and earnings: glossary of terms](#).

Geography (including list of changes to boundaries)

Household estimates of income are produced by the ONS, for each Census 2011 MSOA in England and Wales. While previous income estimates for small areas releases (for example, 2018) used the 2011 Census data, this release uses the new 2021 data (and 2011 boundaries as described previously). Consequently, almost all of the census covariates were updated. In 2011, 7,201 MSOA units existed in England and Wales, and this increased to 7,264 in 2021.

For consistency with previous releases and the structure of the other covariates from 2019 to 2020, the 2011 MSOAs were reconstructed in terms of census data from the 2021 MSOAs. This meant that we could obtain the census profiles for each 2011 MSOA so that their contained census profiles would be up to date for 2021. This was carried out by weighting according to the numbers of postcodes contained within each MSOA. Such transitions only affected a minority of MSOAs whose borders had moved between the censuses.

The model does not support the disaggregation of incomes below MSOA level, nor the aggregation to any level (including to local authorities). The exception is at International Territorial Level (ITL1) region and nation level, where SAIEs are calibrated to the FRS and HBAI at these levels. They will fully align with accredited official statistics published by the Department for Work and Pensions.

Why you can trust our data

The ONS is the UK's largest independent producer of statistics and its national statistical institute. Our [data policies](#) detail how data are collected, secured and used in the publication of statistics. We treat the data that we hold with respect, keeping it secure and confidential, and we use statistical methods that are professional, ethical and transparent.

Income estimates for small areas are designated [accredited official statistics](#) by the [Office for Statistics Regulation \(OSR\)](#) in accordance with the [Statistics and Registration Service Act 2007](#). This designation signifies compliance with the [Code of Practice for Statistics](#), which has recently been updated and focuses on trustworthiness of data in greater depth.

6 . Methods used to produce the data

How we collect the data, main data sources

Survey data

The survey data underlying the dependent variable (income) used in the model were obtained from the [Family Resources Survey \(FRS\): financial year 2019 to 2020, published on the GOV.UK website](#).

The FRS was chosen as the source for survey data on the basis of it being the largest sample that includes suitable questions on income. This allows four survey variables to be modelled and the average is used as the summary variable. For example, the estimates produced are values of average Middle-layer Super Output Area (MSOA) income for the following four income types, the first of these from the Family Resources Survey and the other three from the Department for Work and Pensions publication, [Households below average income](#) (HBAI), which uses the FRS:

- total annual household income (unequalised)
- net annual household income before housing costs (unequalised)
- net annual household income before housing costs (equalised)
- net annual household income after housing costs (equalised)

Sample size and survey data file

The FRS uses a stratified clustered probability sample drawn from the Royal Mail's Postcode Address File (PAF). The survey selects 1,417 postcode sectors with a probability of selection that is proportional to size and 28 addresses within each selected postcode sector. In the financial year ending March 2020, the achieved sample size (for the UK) was 19,244 households. More information on the FRS methodology is contained within the [FRS background note and methodology report on the GOV.UK website](#).

The requirement for this release is to produce MSOA-level estimates of average household income (four types) for England and Wales. The survey data file used contained 14,408 households from 1,170 postcode sectors in the financial year ending March 2020. The final survey data file for England and Wales contained cases in 2,551 different MSOAs out of a total of 7,201.

Definitions from Family Resources Survey data

Although all the survey data used in the modelling process are obtained from the FRS, three of these income types (net weekly household income: unequivalised, and equivalised both before and after housing costs) is defined and calculated in the [HBAI report, published on GOV.UK](#). The HBAI dataset is a cut-down and modified version of the FRS data with slightly different grossing factors. Details of and reasons for these modifications are covered in the HBAI report.

How we process the data

FRS and HBAI-based income; the dependent variables in the model

The dependent variable (income) information used was obtained from the FRS and HBAI. Certain processing steps were needed to convert the four income variables into ones useable in the regression models.

Firstly, for each different income type, a minority of records (302 of 14,408, or 2% for total annual household income) were found with values of income less than or equal to £1. These were removed from the sample dataset. Additional records with extremely high total income values were removed as they would have had an unduly large influence on the model. These households either had a total weekly household income that equated to over £1,000,000 per year, or a total weekly household income over £15,000, and were the only households sampled in a MSOA.

For the net weekly (unequivalised and equivalised) income, records were removed where the net income was greater than the total income. The net equivalised weekly income excludes households containing a married adult whose spouse is temporarily absent. This is because the data for net equivalised weekly incomes come from another FRS-based dataset, [Households below average income data \(HBAI\), published on GOV.UK](#). This is a record-level dataset maintained by the Department for Work and Pensions (DWP).

Covariate datasets

The methodology requires covariates data to be available at a geographic level compatible with MSOAs. A range of data sources were used in the modelling process presenting variables that may be related to household income. They are:

- Census 2021: a wide range of variables relating to the MSOA where each FRS respondent is located; examples include the proportion of adults involved in managerial and professional work, and the proportion of households who are defined as deprived in terms of health dimension
- Department for Work and Pensions benefit claimant counts, August 2019 (provided as counts)
- Valuation Office Agency (VOA) Council Tax Bandings, March 2019 (provided as counts)
- Office for National Statistics, House price statistics for small areas, Quarter 1 (January to March) 2020 (in addition to counts of the number of dwelling sales, data contain measures of house prices (median price) for sales that took place)
- Department of Energy and Climate Change, Energy consumption data, 2019
- HM Revenue and Customs, Pay as You Earn data, 2019
- regional or country identification variable

For further details, refer to our accompanying technical report, [Income estimates for small areas in England and Wales, technical report: financial year ending 2020](#).

Data preparation

Variables of interest from the survey dataset, such as the weekly household income, were classified into MSOAs according to the respondents' postcodes. The covariate dataset comprises MSOA covariates along with the corresponding MSOA identifiers. These two datasets are matched by the MSOA codes.

While previous releases (for example, 2018) used the 2011 Census data, this release uses the new 2021 data. Consequently, almost all of the census covariates were updated. In 2011, 7,201 MSOA units existed in England and Wales, and this increased to 7,264 in 2021.

For consistency with previous releases and the structure of the other covariates from 2019 to 2020, the 2011 MSOAs were reconstructed in terms of census data from the 2021 MSOAs. This is described under the "Geography" sub-heading of [Section 5: Quality characteristics of the data](#).

The resulting matched dataset, containing the survey variable along with associated covariates and MSOA and postcode sector (that is, the FRS "Primary Sampling unit") identifiers, became the analysis dataset. The analysis dataset is required for the modelling and the full covariate dataset is required to produce the final estimates once the modelling has been performed.

As with the modelling for previous publications, where missing values existed for any of the covariates, the England and Wales mean of the variable in question was used to impute the missing value.

How we analyse and interpret the data

Linear models were developed for estimates of incomes for England and Wales. These models relate the survey variable of interest (measured at household level) to the covariates related to the small area in which the household is located. They were fitted as multilevel models and can be used to produce estimates of the target variable at the small-area level. These models can be used to produce MSOA-level estimates of average weekly household income and calculate [confidence intervals](#) for the estimates.

For all four types of income, the response variable "weekly household income" was positively skewed (the largest values differ from the mean more than the smaller values do). By using the natural logarithm of the appropriate type of income as the response variable, this skewness was reduced. It is then assumed for the analysis, that the transformed variable follows a normal distribution.

The models were fitted using the statistical software SAS with postcode sectors at the higher level and households at the lower level. Region and country indicator terms are forced into the model (whether significant or not) and then the method of stepwise forward selection is used to identify the significant covariates to be included in the models from the set of covariates.

All the appropriate covariates (those expressed as percentages or proportions) were transformed onto the logit scale and both the transformed and original covariates were considered for inclusion in the models. The covariates were centred by subtracting the corresponding means for England and Wales.

Initially, [statistically significant](#) (at the 5% level) covariates were selected using a stepwise method for inclusion in the models. Then with these significant covariates, interaction terms were created, tested for significance and where appropriate, included in the models. Note that there were some covariates that did not maintain significance at the 5% level once the interaction terms were applied. Where this was the case, these covariates were, nevertheless, included in the model containing the interaction terms.

After modelling, adjustments are made to the modelled estimates to ensure they were consistent with the direct survey estimates at regional level for England and country level for Wales (this is known as "calibration"). The FRS data are used to calculate direct estimates of income at these higher geographical levels (estimates at this level are considered robust).

The model-based MSOA estimates of income were aggregated to region and country level, and comparisons made between the two sets of estimates. The ratio of direct survey estimate to aggregated model estimate at the region and country level was used to scale all of the modelled MSOA-level estimates and their confidence intervals. More details on this calibration and benchmarking methodology, and aspects of the modelling methodology are given in our previously published methodology, [Income estimates for small areas technical report: financial year ending 2016](#).

How we quality assure and validate the data

Once a model has been selected, an assessment of the quality is made using several diagnostics, in order to assess the appropriateness of the models developed. The diagnostic checks employed here are those developed by the Office for National Statistics (ONS) for small area estimation and published in our article, [Evaluation of small area estimation methods - an application to unemployment estimates from the UK Labour Force Survey \(LFS\), on the Research Gate website](#), as well as some additional ones. The analysis shows that in general, the models are well specified, and the assumptions are satisfied. This provides confidence in the accuracy of the estimates and the confidence intervals produced from the models.

More detail about the diagnostic tests and why they are performed, along with quality finding for the financial year ending March 2020 models can be found in our previously published methodology, [Income estimates for small areas technical report: financial year ending 2016](#). The following paragraphs describe some of the diagnostic tests performed on the data.

Residual compared with model estimates diagnostic plot

A plot of model estimates against model residuals both at the household and the area level is a method of checking that the model assumptions are satisfied, and the model accurately describes the population. Here we are testing for two things: model misspecification and non-constant variance of the residuals (heteroscedasticity). If any pattern remains in the residuals, this implies model misspecification. For example, a covariate influential to income may have been left out of the model.

Constant variance in the area-level residuals is required, since this will have an impact on the calculation of the confidence intervals. Model estimates are calculated at the household and plotted against the household-level residuals. The standard errors can be used to determine whether the constant and linear terms are significantly different from zero.

Model compared with sample estimates diagnostic plot

A plot of direct survey estimates (y-axis) against model-based estimates (x-axis) for MSOAs, for which there is a sample, is one method of assessing whether the relationship between the target variable and the covariates has been specified properly. For good model-based estimates, the direct estimates will be randomly distributed around the estimates and the regression line between the two will be very close to the line "y equals x". A misspecified relationship between the direct and model-based estimates would present as a curve or scattered round a straight line different from the "y equals x" line.

An important assumption when using this diagnostic is that the direct estimates are unbiased. The technique for calculating direct survey estimates at an MSOA level is described in our previously published methodology, [Income estimates for small areas technical report: financial year ending 2016](#), along with further detail about this diagnostic test. Freedom from bias are evidenced:

- if in quadratic fit, neither the quadratic term nor the intercept are significant
- if in the linear fit, the intercept term is not significantly different from zero and the slope term is not significantly different from one

Coverage diagnostic

The purpose of this diagnostic is to examine the validity of the confidence intervals for the model-based estimates. For those MSOAs in the sample, there will be direct survey estimates with associated 95% confidence intervals. The diagnostic measures the overlap between the direct confidence intervals and the corresponding model-based estimate confidence intervals. For example, it measures the percentage of MSOAs for which the model and direct confidence intervals overlap.

However, the overlap between two independent 95% confidence intervals for the same quantity is higher than 95%, therefore it is necessary to modify the nominal coverage levels (that is, narrow the width) of the confidence intervals being compared to ensure a 95% overlap. Further details of the modification and this test are available in our previously published methodology, [Income estimates for small areas technical report: financial year ending 2016](#). Any significant deviation from a 95% overlap indicates that the model-based confidence intervals are generally too wide or too narrow.

Wald statistic

This diagnostic test assesses the assumptions underlying the model by using a Wald goodness-of-fit statistic to test whether there is a significant difference between the expected values of the direct estimates and the model-based estimates. Typically, small area-level model-based and direct survey estimates will be approximately correlated and there should be a non-significant p-value associated with the Wald statistic.

Stability analysis

This diagnostic test analyses the stability of the model's predictive power. The data are split into two datasets similar in size and MSOA representation. The model is fitted to one-half of the data to obtain regression coefficients.

In a similar way, the other half of the data are used in the model to obtain the regression coefficients. These two sets of regression coefficients are then used to obtain two sets of comparable model-based estimates for all MSOAs. This process is repeated 10 times and for each repetition, the difference between the two sets of estimates is measured to evaluate the stability of the model. A relative root mean square error (RRMSE) is also used as a measure of how close the two sets of model-based estimates are. A small RRMSE indicates that the differences between the two sets of estimates are not significant.

How we disseminate the data

A dataset containing [estimates of the four types of income for the financial year ending March 2020](#) are available for each MSOA in England and Wales, along with their confidence intervals. These can be downloaded in Excel format. A [statistical bulletin](#) and a [technical report](#) accompanies the publication. The underlying data for the charts and tables in the bulletin can also be downloaded.

Most queries can be answered from the website datasets or supporting methods documents. Any additional enquires regarding the household projections can be made by emailing economic.wellbeing@ons.gov.uk. It may be possible to meet additional data requests, but these may be chargeable depending on the time required to produce the additional data requested. Metadata describing the limitations of the data for more detailed tables are provided with each individual request.

How we review and maintain the data processes

The modelling process has in most cases, been carried out every two years. Each time, the most recent data that relates most closely to the reference period are used for both the dependent variables in the model (DWP FRS or HBAI income) and the independent variables (as listed previously), such as census data. While small incremental improvements in the methodology and model validation will take place between releases, the method has largely remained unchanged since its first release in 2005.

7 . Other information

User feedback and assessment of user needs and perceptions

There continues to be substantial and increasing stakeholder interest in household income levels at very localised levels of geography. We are also aware of demand to further improve the detail and utility of outputs, for example, to generate distributional estimates (such as medians) and the proportion of households that fall below various "poverty" thresholds. There is also interest in providing these data at lower-level geographies, such as Lower-layer Super Output Area.

Interest also exists in extending the models to include Scotland and Northern Ireland and to make further use of administrative data, reducing future dependency on surveys.

Early feasibility work to explore the use of administrative data to directly determine small area incomes rather than as covariates has been carried out and our article, [Exploring the use of admin data to derive small area income estimates, England and Wales](#) published on 30 April 2024. It showed that available admin-based income statistics data sources go a long way towards producing small area estimates of gross and new incomes before housing cost deductions, but there are still significant gaps to fill regarding housing costs. There is additional work needed to capture the missing components by including additional data sources in the methodology.

8 . Related links

[Income estimates for small areas, England and Wales: financial year ending 2020](#) Bulletin | Released 11 October 2023

Small area model-based income estimates covering local areas called Middle-layer Super Output Areas (MSOAs) in England and Wales.

[Income estimates for small areas in England and Wales, technical report: financial year ending 2018](#)

Technical report | Released 11 October 2023

Methods used to produce small area income estimates for local areas in England and Wales, known as Middle-layer Super Output Areas (MSOAs), including information on the quality of the models and estimates.

[Exploring the use of admin data to derive small area income estimates, England and Wales](#) Article |

Released 30 April 2024

Comparison of admin-based small area income (official statistics in development) statistics (ABIS) with official accredited modelled estimates and how they meet the Canberra Handbook criteria.

[Admin-based income statistics, England and Wales: tax year ending 2018](#) Article | Updated 20 December 2022

Brings together data from the Pay as You Earn and benefits systems to drive experimental estimates of gross and net income.

[Average household income, UK: financial year ending 2022](#)

Bulletin | Released 25 January 2023

Final estimates of average household income in the UK, with analysis of how these measures have changed over time, accounting for inflation and household composition.

[Household income inequality, UK: financial year ending 2022](#)

Bulletin | Released 25 January 2023

Main estimates of household income inequality in the UK and how these measures have changed over time accounting for inflation and household composition.

9 . Cite this methodology

Office for National Statistics (ONS), published 25 June 2024, ONS website, Quality and Methodology Information report, [Income estimates for small areas QMI](#)