

Annual Purchases Survey: technical report

Detailed information about the methodological and technical procedures used to produce the Annual Purchases Survey (APS) estimates.

Contact:
Dai Summers
ABAPS@ons.gov.uk
+44 1633 456602

Release date:
8 August 2023

Next release:
To be announced

Table of contents

1. [Introduction](#)
2. [Questionnaire design](#)
3. [Sampling procedure](#)
4. [Data collection](#)
5. [Converting respondent data into published estimates](#)
6. [Publication](#)
7. [Disclosure control and data confidentiality](#)
8. [Revisions policy](#)
9. [Product list \(as at 2023 collection\)](#)
10. [Related links](#)
11. [Cite this methodology](#)

1 . Introduction

This guide describes the methodological and technical procedures used by the Office for National Statistics (ONS) to produce the Annual Purchases Survey estimates. The report is aimed at users who want to know more about the background and history, uses and users, and concepts and statistical methods underlying the survey. It includes information about questionnaire development, sample design, data collection, results processing, publications and quality issues.

This technical guide will be revised in line with any major future survey developments.

1.1 Overview

The Annual Purchases Survey (APS) was reinstated by the Office for National Statistics (ONS) in 2015 to meet a range of user demands and to fulfil the recommendations for its reinstatement from two independent reviews, one by [Dame Kate Barker and Art Ridgeway](#) and the other by [Professor Sir Charles Bean](#). A previous survey, entitled the Purchases Inquiry, was suspended in 2006 because of insufficient quality in the data and to reduce the ONS costs and burden on UK businesses (final reference period being 2004). However, given the survey provides important information on the products that UK businesses purchase, it was decided that it should be reintroduced from 2015.

The primary aim of the APS is to provide a comprehensive picture of the goods and services used up or transformed in the production process and running of UK businesses, otherwise referred to as intermediate consumption.

The survey collects information about businesses' expenditure on energy, services, goods and materials that are used up or transformed by the business activity. It specifically excludes fixed assets or capital investment, staff costs and goods and services bought for resale without further processing.

APS questionnaires are sent by the ONS annually to approximately 33,000 businesses in the UK. In the UK, it is a compulsory survey that is administered under the statutory powers of the Statistics of Trade Act 1947 for Great Britain and under the Employment (Northern Ireland) Order 1988 for Northern Ireland.

1.2 Main uses and users of the data

Intermediate consumption is required as part of the process to set the annual level of gross domestic product (GDP), which is an essential statistic for informing fiscal and monetary policy decisions. Broadly speaking, GDP is calculated by adding up the value of the output of firms less the goods and services that are used in the production of that output. The value of these goods and services used in production is called "intermediate consumption" and is defined within the [European Systems of Accounts: ESA 2010](#) manual as follows:

"Intermediate consumption consists of goods and services consumed as inputs by a process of production, excluding fixed assets whose consumption is recorded as consumption of fixed capital. The goods and services are either transformed or used up by the production process."

The data collected from the APS feed into the supply and use framework, which is a central component of the national accounts balancing process and sets the annual level of nominal GDP.

ONS statistics on trade, intangible assets and environmental accounts also incorporate relevant parts of the APS dataset.

Externally to the ONS, users include the Department for Energy Security and Net Zero and the Department for Business and Trade, who use the energy consumption information for policy making.

Data is also used by the Northern Ireland Statistics and Research Agency, as well as the Scottish and Welsh Governments.

There are regular meetings of the "Government User Group", which includes discussion of the APS. This, along with bilateral meetings with other stakeholders, provides an opportunity for any changes or developments to the survey to be discussed directly with its government users, so that, where possible, their requirements can be met.

1.3 History

The original Purchases Inquiry (PI) ran from the 1950s to 2006. Data were collected at five-yearly intervals, then annually from 1999 onwards. The sample size and industry coverage were expanded between 1999 and 2001, from 3,000 businesses (covering partial production, distribution and sales industries) to 28,000 (covering full production, construction, distribution and services industries). In its fullest form, the PI contained around 500 variations of questionnaires and 1,400 questions, with an average of 40 questions going to each respondent.

In 2006 (reference year 2005), the PI sample size was cut by 50% to reduce costs and burden on businesses. Consequently, the results were considered to be of insufficient quality and the survey was suspended. The last dataset dates from 2005 and covers reference year 2004.

The APS was re-introduced from 2015 and runs annually. During the coronavirus (COVID-19) pandemic, while forms were dispatched, processing (response chasing, clearance and quality assurance) for the 2019 and 2020 APS was deprioritised and data for these years have not been published.

1.4 Timeline

The APS process from sample selection through to the publication of the final estimates has varied since its reintroduction. A typical year pattern would be as follows for collection of survey period year "Y":

- agree any questionnaire changes, January Y plus one
- dispatch the forms in February or March with a return date of May Y plus one
- editing and imputation, quality assurance, outliering
- provide provisional estimates to National Accounts, December Y plus one
- finalise results for National Accounts, March Y plus two

Given its experimental status, and the deprioritisation during the COVID-19 pandemic, publication of provisional or final data has not settled into a regular pattern. For the 2021 reference year, data was released in June 2023.

2 . Questionnaire design

2.1 Overview

The Annual Purchases Survey (APS) has different questionnaire types depending on the industry of the business. The survey questions were produced based on the Statistical [Classification of Products by Activity \(CPA\) version 2.1](#) structure.

The CPA is the classification of products (goods as well as services) at the level of the European Union (EU). Product classifications are designed to categorise products that have common characteristics. They provide the basis for collecting and calculating statistics on the production, distributive trade, consumption, international trade and transport of such products.

Each form starts with the request for total intermediate consumption and then asks for a further breakdown of this product, grouped into the three subtotals of energy, services and goods.

There are a set of core products considered crucial to all industries and asked on every form type. The other questions relate to products likely to be reported based on the industry. However, not wanting to restrict businesses to an assumed list of products, a list of the remaining products for each section are also specified and can be included if relevant.

Industries that fall in the manufacturing sectors are also asked an additional question about whether the business carried out work on behalf of a customer where they only provided labour and no materials.

2.2 Questionnaire development

The development of the APS was essential in meeting the regulations set out in the Eurostat Manual of Supply, Use and Input-Output Tables. Requirements from stakeholders such as national accounts, the (then) Department for Business, Energy and Industrial Strategy (BEIS), and devolved administrations were also gathered and prioritised.

All requests for inclusions and exclusions or changes to the questionnaires had to be agreed with the APS Project Board. Any such requests were assessed and costed before being agreed, including the relevant compliance costs (the costs incurred by businesses through responding to the survey). Once provisional agreement to any change had been obtained, the required changes were tested to ensure that responding businesses understood the proposed wordings and were able to supply the relevant information.

2.3 Questionnaire review

The APS is a relatively complex survey and balance needs to be struck between questions that businesses can respond accurately to and the detail required to underpin national accounts estimates.

A high-level review of the APS questionnaire was conducted in 2018, in preparation for reference year 2018 (sent out in 2019). Working closely with the Behavioural Insight Unit within the Office for National Statistics (ONS), significant changes were made to the questionnaire format. Some of the changes to note are the use of a checklist to ensure all stages of the questionnaire are completed, additional guidance on how to provide a best estimate and moving the list of additional products from the end of the questionnaire to the end of each section.

From the reference periods 2020 and 2021, minor changes were made to the collection of some of the energy related categories.

During 2022 to 2023, the survey was subject to further review to facilitate its move from a paper-based collection to an online questionnaire.

The list of products collected is included in Section 9: Product list (as at 2021 collection).

3 . Sampling procedure

3.1 Sampling frame

The Inter-Departmental Business Register

A sampling frame is a complete list of all the members of a population being studied, from which the sample is drawn. The sampling frame for the Annual Purchases Survey (APS) is the list of UK businesses on the Inter-Departmental Business Register (IDBR).

The Inter-Departmental Business Register (IDBR) is a comprehensive list of UK businesses used by government for statistical purposes. It is also an important data source for analyses of business activities. There are 2.8 million businesses on the IDBR.

The information used to create and maintain the IDBR is obtained from five main administrative sources. These are:

- HM Revenue and Customs (HMRC) Value Added Tax (VAT) – traders registered for VAT purposes with HMRC
- HMRC Pay As You Earn (PAYE) – employers operating a PAYE scheme, registered with HMRC
- Companies House – incorporated businesses registered at Companies House
- Department for Environment, Food and Rural Affairs (DEFRA) farms
- Department of Finance and Personnel, Northern Ireland (DFPNI)

As well as these five main sources, a commercial data provider, Dun and Bradstreet is used to supplement the IDBR with Enterprise Group information.

In addition, the Office for National Statistics (ONS) [Business Register Employment Survey \(BRES\)](#) and other ONS surveys supplement these administrative sources, identifying and maintaining the business structures necessary to produce detailed industry and small area statistics. It should be noted that BRES is the only source of local unit (site) information.

[Further information about IDBR sources, structure and updating for publications \(PDF, 59.32KB\)](#) is published on our website.

Reporting units

The reporting unit holds the mailing address to which survey questionnaires are sent and displays the selection criteria used by surveys. The reporting unit can cover the whole enterprise, or parts of the enterprise identified by lists of sites (called local units). A local unit is where the business activity and employees are based.

An enterprise may contain one or more local units, for example, several shops or restaurants. An enterprise may, therefore, have local units at different locations and may carry out more than one type of economic activity. Except for a minority of large or complex businesses, the reporting unit is the same as the enterprise (representing all local units attached to the enterprise). For this reason, the APS reporting unit counts are presented as enterprise counts in publications.

An enterprise can also belong to an enterprise group. This is where a set of enterprises under common ownership are linked together in groups. The ultimate owner of an enterprise group can be in the UK or overseas.

Standard Industrial Classification (SIC)

Each enterprise is classified according to the [Standard Industrial Classification of Economic Activities \(SIC\)](#) system. The UK is required by European legislation to have a system of classification consistent with the European Union's industrial classification system. The system underwent a major review in 2007. APS data have been collected and published on the SIC 2007 system.

UK SIC 2007 is divided into 21 sections, each denoted by a single letter from A to U. The letters of the sections can be uniquely defined by the breakdown to the divisions (denoted by two digits), which are then broken down into groups (three digits), then into classes (four digits), and into subclasses (five digits).

For example, in SIC 2007:

- section C manufacturing (comprising divisions 10 to 33)
- division 13 manufacture of textiles
- group 13.9 manufacture of other textiles
- class 13.93 manufacture of carpets and rugs
- subclass 13.93/1 manufacture of woven or tufted carpets and rugs

The full structure of SIC 2007 consists of 21 sections, 88 divisions, 272 groups, 615 classes and 728 subclasses.

Each local unit is assigned a single five-digit SIC code, which corresponds to the unit's principal activity. Where more than one type of economic activity is carried out by a local unit or reporting unit, its principal activity is the activity in which most of the people are employed, though this does not necessarily account for 50% or more of the total employment of the reporting unit.

For example, if a reporting unit contains four local units (for example, four shops), and each specialises in selling different things, then they would be under different SICs. If one local unit contains more employees than the others, then the local unit with the most employees would determine the SIC of the reporting unit. The proportion of employees in each local unit could be 30%, 25%, 25% and 20% of all employees, so the reporting unit would be placed under the SIC that holds 30% of the employees. There are detailed rules for determining SIC for multiple-activity economic units, including situations where measures of value added are not available.

Re-classification of a business can occur due to a relatively small change to the nature of its operation, and this can have a significant effect on APS estimates by industry. In addition, the correction of misclassification of businesses can lead to bias, particularly where there is systematic movement from one industry to another. This is because, where classification updates are identified via survey returns, it is only units in the survey sample that are updated.

All surveys that do not cover the whole business population, such as the APS, have the potential for some underestimation of output variables due to the re-classification of units moving out of the APS survey population but never into it. However, such underestimation is likely to be small. In the APS, this effect is corrected for by adjusting the weights of the businesses that remain in the sample.

The exact inclusions and exclusions of legal statuses and industries in the APS are detailed in this section.

Legal status

The APS covers the private sector only:

- legal status one (company, including building society, Limited Liability Partnership (LLP) and joint venture)
- legal status two (sole proprietor)
- legal status three (partnership, and limited partnerships)

SIC inclusions

- Part of Agriculture, forestry and fishing (part of section A – Standard Industrial Classification 01.6 to 01.7).
- Mining and quarrying (section B).
- Manufacturing (section C) – partial coverage of division 23 (Manufacturing of non-metallic mineral products) up to and including the 2022 collection (see SIC exclusions for detail).
- Electricity, gas, steam, and air conditioning supply (section D).
- Water supply; sewerage, waste management and remediation activities (section E).
- Construction (section F).
- Wholesale and retail trade; repair of motor vehicles and motorcycles (section G).
- Transport and storage (section H).
- Accommodation and food service activities (section I).
- Information and communication (section J).
- Part of Financial and insurance activities (part of section K – Standard Industrial Classification 64301 to 65202 and 66110 to 66300).
- Real estate activities (section L).
- Professional, scientific, and technical activities (section M).
- Administrative and support service activities (section N).
- Education (section P).
- Human health and social work activities (section Q).
- Arts, entertainment, and recreation (section R).
- Other service activities (section S).

SIC exclusions

- Part of Agriculture, forestry and fishing (part of section A – Standard Industrial Classification 01.1 to 01.5).
- Manufacturing (section C) – groups 23.1, 23.3, 23.4 and 23.9 were excluded up to and including the 2022 data collection – they have been included from the 2023 collection.
- Part of Financial and insurance activities (part of section K – Standard Industrial Classification 64110 to 64290 and 65300).
- Public administration and defence; compulsory social security (section O).
- Activities of households as employers; undifferentiated goods and services producing activities of households for own use (section T).
- Activities of extraterritorial organisations and bodies (section U).

3.2 Sample design

Data are collected by the ONS from approximately 33,000 businesses in the UK (excluding the Channel Islands and the Isle of Man). The area covered is in line with other business surveys produced by the ONS.

Sample selection is carried out using a stratified random sample design. Similar reporting units (businesses) are grouped into strata as defined by three variables:

- employment size band
- SIC
- geographical region

Sample selection occurs independently for each stratum. When the sample is designed, the size of the sample in each stratum is determined by an algorithm, which distributes the sample amongst the cells to give the lowest estimated variance (uncertainty). This design is significantly more efficient (that is, it gives a much more accurate estimate for the same sized sample) than a simple, unstratified random sample.

The variables defining the strata are:

- six employment size bands: 0 to 9, 10 to 19, 20 to 49, 50 to 99, 100 to 249, and 250 or more.
- industry class: four-digit SIC [UK Standard Industrial Classification 2007: SIC 2007](#).
- region: England, Scotland, Wales, and Northern Ireland.

Please note, the APS does not publish to this level of disaggregation.

All businesses with 250 or more in employment are selected every year. This is because, as they are large enterprises, they are dominant respondents to estimated total values, including all the largest enterprises significantly reduces uncertainty on the estimated total values. The APS also uses inclusion markers to select businesses with low employment but high turnover.

For businesses with fewer than 250 in employment, the survey is designed so that a proportion of a stratum will generally be selected for two years before being rotated out of the sample. The random sample selection uses permanent random numbers (PRN), a unique nine-digit identifier that is randomly assigned to each reporting unit when it is added to the IDBR. The sample for each stratum is constructed using consecutive PRNs from within that stratum until the sample size required has been reached.

For the APS, each sample is generally selected for two years, and there is a year-to-year overlap of half the sample. That is, in any year, approximately half of the sample will be newly selected, and half will have been selected in the previous year as well. This is illustrated in Table 1, for a sample of four reporting units taken from a stratum containing 10 units (note that these are not real PRNs). This design means that, for half the sample, returns are available from the same businesses in consecutive years, and this helps to maintain the quality of editing and validation, imputation and outlier detection (see Section 5: Converting respondent data into published estimates for more information).

Table 1: Example of permanent random numbers sampling method

Year 1	Year 2	Year 3
143*	143	143
290*	290	290
339*	339*	339
418*	418*	418
497	497*	497*
545	545*	545*
624	624	624*
785	785	785*
824	824	824
908	908	908

Source: Office for National Statistics

Notes

1. The asterisk (*) signifies that these units were selected in this year.

Table 1 shows that in the first year, units 143, 290, 339 and 418 were selected. In the second year, the first two units were rotated out (143 and 290) and the last two units were retained (339 and 418). Additionally, two new units were added into the sample (497 and 545). This process was then repeated for the third year. When the last PRN is reached, selection rolls around to the smallest again.

However, there are a few exceptions to this design. If a selected reporting unit then moves to another stratum, for example, by changing SIC classification (see Section 3.1: Sampling frame), then it may be selected again in the new stratum. Also, if there are fewer reporting units within a stratum because of businesses moving out, the likelihood of consecutive selection will increase. For these reasons, there is never a guarantee that a business will only be selected for two years.

A further exception arises for the strata within the smallest employment size band.

For businesses with employment of zero to nine, which are not part of an enterprise group, Osmotherly rules apply. These rules state that when a business with zero to nine employment has been selected for an annual survey, it will only be selected for a single year and it will not be reselected for at least three years (provided they complete and return the questionnaire). This is implemented to reduce the burden on small businesses, which may not have as much resource for completing survey questionnaires.

4 . Data collection

4.1 Timetable of questionnaire dispatch

Questionnaires are sent to collect information from businesses relating to the previous 12 months, which is known as the reference year. The questionnaires are required to be returned to the Office for National Statistics (ONS), in a pre-paid envelope, within six weeks.

4.2 Welsh questionnaires

The ONS Welsh Standards give an option for Welsh business respondents to request a Welsh language version of the questionnaire. This option is clearly shown and written in Welsh on the front page of the Annual Purchases Survey (APS) questionnaire.

4.3 Euro respondents

Respondents who prefer to provide their purchases values in the Euro currency are provided with a Euro questionnaire upon request.

These responses are converted to pounds sterling (£) using the [universal currency converter](#).

4.4 Reminder letters

Up to two reminder letters are sent to businesses who have not returned a completed questionnaire by the end of April. One is sent at the beginning of May, the second in June.

All non-responders with employment of 1,000 or more are sent a Chief Executive Letter (CEL), and a duplicate questionnaire in August if they have still not returned. The CEL is a stronger reminder to inform the chief executive or managing director that the business has not responded and is a reminder of the legal requirement to respond. The CEL further outlines the non-compliance penalties prior to any enforcement procedures.

4.5 Response chasing

During the data collection period, the APS response rates for returned questionnaires are monitored regularly. A manual exercise is undertaken during the data collection cycle to identify industries with low response rates.

Telephone response chasing may occur depending on resourcing priorities. It starts after the second reminders have been dispatched (start of June) and continues, if necessary, up to the final result run (November following the reference year). It is intended to encourage the completion of the questionnaire and address any respondent issues in a timely and efficient manner, which all leads to the production of a quality output.

4.6 Enforcement strategy

The APS carries out enforcement action under the [Statistics of Trade Act 1947](#). Enforcement action is used to maintain response rates, and hence the quality of the survey. It is used only as a last resort, after attempts to encourage businesses to complete the survey following telephone calls and reminder letters.

5 . Converting respondent data into published estimates

5.1 Editing and validation

Returned data is passed through a set of validation checks, for example:

- components of the questionnaire do not add to the reported totals
- the returned dates fall outside determined thresholds
- the returned data contain negative values
- data are not returned where a question is compulsory
- for businesses in the sample for two consecutive periods, the year on year estimates increase or decrease by over a certain percentage

Following these automatic editing and validation checks, manual quality assurance of the data is conducted.

5.2 Imputation

Imputation techniques are used to estimate the value of missing data due to non-response, whether partial (item non-response) or full (unit non-response). Item non-response occurs when a business returns a value for its section total (for example, total amount spent on energy, water and waste) but is unable to break the total down to a product level. Unit non-response occurs when a business does not respond to the questionnaire at all.

Imputation is designed to give better results than deletion, in which all subjects with any missing values are omitted from the analysis.

Unit non-response

For large businesses (250 employment or more) that have responded in the previous reference period, unit non-response can be imputed for. This uses the business' data from the previous reference period, and an average growth, to impute the missing data in the current reference period for the business. This method, called the "ratio of means" technique assigns each business to an imputation class (a group of similar businesses), which are based on the two-digit standard industrial classification and employment brackets below or above 100 employees. Businesses in the imputation class that have responded in both the previous and current period will be used to calculate an "imputation link".

The imputation link is calculated as follows:

$$R_{class,t} = \frac{\sum_{class} y_{i,t}}{\sum_{class} y_{i,t-1}}$$

where:

- $class$ is the set of all businesses in the imputation class that have a response at both time t and $t-1$ (the previous and current year)
- $y_{i,t}$ is the section total for business i in the current period
- $y_{i,t-1}$ is the section total for business i in the previous period

The section total for business j is imputed as follows:

$$\hat{y}_{j,t} = R_{class,t} \times y_{j,t-1}$$

The component values in the section for business j are imputed as follows:

$$\hat{y}_{j,k,t} = R_{class,t} \times y_{j,k,t-1}$$

where:

$y_{j,k,t-1}$ is the value of purchases for product k in the previous year.

It is important to note:

- only returned values are used in the calculation of $R_{class,t}$ – imputed values are excluded
- the link is applied to returned values for the previous year, or imputed values that were imputed using this method in the previous year

Item non-response

When a business has provided us with a high-level total for each of energy, goods and services, but cannot break it down in to detailed products, then the breakdown will be automatically imputed based on other similar businesses.

For a product k , and a given imputation class, the proportion of a section total that it will be assigned is calculated as follows:

$$p_{class,k,t} = \frac{\sum_{class} y_{i,k,t}}{\sum_{class} y_{i,t}}$$

where:

- $class$ is the set of businesses that have provided a breakdown for the section at time t
- $y_{i,k,t}$ is the value of purchases for product k for business i at time t
- $y_{i,t}$ is the section total for business i at time t

The imputed value of purchases for business j for product k is calculated as:

$$\hat{y}_{j,k,t} = p_{class,k,t} \times y_{j,t}$$

5.3 Estimation of totals

It is not possible to collect data on every UK business every year (a census), because:

- the burden on businesses would be too great
- the cost of running such a census would be prohibitive
- a well-designed sample survey can produce better estimates than a census with a poor response rate

Therefore, the APS collects information from a sample of the UK business population each year. The sample design is described in Section 3.2. This section describes how returns from the sample are used to estimate totals for the whole population.

Weighting

To calculate the estimates for an entire population from data collected from a sample, the APS uses standard statistical weighting methods. Specifically, the results received from the sample are multiplied by three weights:

- the design weight, known in ONS as the a-weight
- the calibration weight, known in ONS as the g-weight
- the o-weight or outlier weight

Design weight

Known in ONS as the a-weight, this accounts for the sample design so that a business' probability of selection is properly reflected. So, for example, a business with a small probability of being selected for the survey will have a large design weight.

Calibration weight

Known in ONS as the g-weight, this makes a correction for an unrepresentative selected sample. For example, in a random selection of five businesses out of a population of 10, it is possible that the five businesses selected are, by chance, towards the upper boundary of the employment size band, as opposed to an even distribution of businesses across the size band. If no correction is made, the population total would be over-estimated.

Auxiliary information, that is information not collected by the survey, which acts as a proxy for the variable of interest, is used to correct for this effect. The weight is the ratio of the actual population total for the auxiliary variable to the population total estimated from the sample's auxiliary variables is calculated. For the APS, the auxiliary variable is the employment value found on the IDBR.

Outlier weight

The o-weight, or outlier weight, identifies potential outliers in the sample, reducing them in line with comparable businesses from the same cell. For example, businesses reporting intermediate consumption much higher than similar businesses of the same size and within the same industry, would be reduced in line with these to reduce volatility in the overall population estimate. These are weighted at product level.

The weighted value is then calculated as:

$$\text{weighted value} = \text{returned value of the variable} \times \text{a-weight} \times \text{g-weight} \times \text{o-weight}$$

Estimates of population totals are then found by simply summing the weighted values across the whole sample.

Calculating the a-weights

The a-weight is calculated for each stratum in the sample, which is a group of businesses defined by their size and industry (see Section 3.2). In its simplest form, the a-weight, a , for each stratum is calculated as:

$$a = \frac{N}{n}$$

where:

- N is the total number of businesses in the cell (the population)
- n is the number of businesses in the returned sample

Calculating the g-weights

G-weights are calculated for groups of strata within the same industry, but across several size bands. Generally, regions (England, Scotland, Northern Ireland and Wales) are grouped for the calculation of g-weights. These groups are called g-weight bands.

In its simplest form, the g-weight is the ratio between the total of the auxiliary variable estimated from the sample and the actual population total for the auxiliary variable. The g-weight will therefore be greater than one when the total auxiliary estimated from the sample is less than the total auxiliary in the population, and less than one when the total auxiliary estimated from the sample is more than the total auxiliary in the population. If response is representative, all the g-weights should be close to one. The g-weight therefore helps correct for any imbalances in the selected sample that arise through random chance or non-response. This is calculated as follows:

$$g = \frac{T_{pop}}{(T_{samp} \times N/n)}$$

where:

- T_{pop} is the sum of IDBR employment over all businesses in the population
- T_{samp} is the sum of IDBR employment over all businesses in the returned sample
- N is number of businesses in the population
- n is number of businesses in the returned sample
- $T_{samp} \times N/n$ is the total for the auxiliary estimated from the sample

Calculating the o-weights

The o-weights are calculated for each product separately, to determine if it represents an outlier or not. The o-weights are calculated using L-values, which set the parameters for the returned data, setting an “upper” and “lower” limit of acceptable values, excluding extreme values through winsorisation.

The o-weights are calculated using the L-values to determine an individual weight for each product using the following formula.

For positive returns ($y_{RU} > 0$), the outlier weight for each reference unit for each question (y) should be:

$$owt_{RU,y} = \min \left[1, \left(\frac{1}{a_{RU} g_{RU,y}} \right) + \left(\frac{1}{y_{RU}} \right) \left(\beta_c x_{RU} + \frac{L_{pos}}{a_{RU} g_{RU,y} - 1} \right) \left(1 - \frac{1}{a_{RU} g_{RU,y}} \right) \right]$$

where:

$$\beta_C = \frac{\sum_{h \in C} \frac{N_h}{n_h} \sum_{RU \in h} y_{RU}}{\sum_{h \in C} \frac{N_h}{n_h} \sum_{RU \in h} x_{RU}}$$

For zero returns ($y_{RU} = 0$), the outlier weight should be calculated as:

$$owt_{RU,y} = 1$$

where:

- owt is the o-weight
- y is the question number

$owt_{RU,y}$ is the o-weight for a given RU for the returned figure for a specific question (y).

Weighted values will be calculated as the product of the value for each question and the weight for that record or variable. Standard errors will be calculated for each variable using the following formula:

$$Var_1(\hat{Y}) = \sum_{h=1}^H N_h^2 \left(1 - \frac{n_h}{N_h}\right) \frac{1}{n_h} s_h^2$$

Where

$$s_h^2 = \sum_{ieh} \frac{(y_i - R_h x_i)^2}{n_h - 1} \text{ and } R_h = \frac{\bar{y}_h}{\bar{x}_h}$$

6 . Publication

6.1 Statistical bulletin

Estimates from the Annual Purchases Survey (APS) are (normally) published on the [Office for National Statistics \(ONS\) website](#) in our statistical bulletin, [Energy, goods and services used by UK businesses](#).

Alongside the bulletin, there are [Energy, goods and services used by UK businesses datasets](#), which look at the consumption of each product by each industry. One table shows this in proportion terms: the proportion of total intermediate consumption by industry X spent on product Y. The other table gives the value in pounds sterling. However, it should be noted that this latter estimate does not currently show the APS figure. As explained in Section 3: Sampling procedure, the APS sample covers only the private sector (legal statuses one, two and three). In an attempt to represent the whole of the UK economy, the method used in the publication is to apply the APS proportions to the Annual Business Survey (ABS) estimates of intermediate consumption (by industry). ABS is not limited to the private sector. The detailed methodology for the ABS is explained in our [ABS technical report](#). There is also some discussion of the coherence between ABS and APS in our [APS Quality and Methodology Information \(QMI\)](#). Resolution of these issues and better documentation to explain them to users is part of an ongoing project.

7 . Disclosure control and data confidentiality

7.1 Confidentiality protection requirement by law and Government Statistical Service (GSS) policy

The need to keep records of individuals, businesses or events used to produce official statistics confidential is enshrined in law. However, this does not prevent the release of anonymised or aggregated data.

The [Code of Practice for Statistics](#) and the [Data and Analysis Method Review on privacy and data confidentiality methods](#) provides the Government Statistical Service (GSS) with a framework for compilation of official statistics in this regard.

Furthermore, the ONS surveys are conducted on behalf of the UK Statistics Authority, and all outputs are subject to Section 39 of the [Statistics and Registration Service Act 2007](#) (SRSA).

Confidentiality requirements that relate to published data are specified in Section 9 of the [Statistics of Trade Act 1947](#). It also states that tables disclosing any information relating to an individual business should not be published unless there is expressed consent in writing from that business. In addition, data should not be published that would reveal the exact number of respondents contributing to a cell if there are fewer than n respondents, as detailed in Section 7.4: Identifying disclosive data for the Annual Purchases Survey.

The [General Data Protection Regulation and the Data Protection Act 2018](#) determine how, when and why any organisation can process personal data.

7.2 The ONS confidentiality pledge

The confidentiality pledge is an assurance of confidentiality given to survey respondents and states:

"All the information you provide is kept strictly confidential. It is illegal for us to reveal your data or identify your business to unauthorised persons."

7.3 Statistical disclosure control and the ONS

The [Statistical Disclosure Control Policy](#) sets out the standards for safeguarding the information provided in confidence to the ONS. "Statistical disclosure control" refers to the methods that reduce the risk that confidential information is published in any official statistics. These methods are applied if ethical, practical or legal considerations require the data to be protected.

Statistical disclosure control involves modifying data so that the risk of identifying individuals is reduced, but at the same time attempts to find a balance between improving confidentiality protection and maintaining an acceptable level of quality in the published data. Statistical disclosure control is applied to the Annual Purchases Survey (APS) data before publication.

7.4 Identifying disclosive data for the Annual Purchases Survey

The design of the APS means that totals can be estimated for each industry and product. However, these totals are aggregated for publication purposes, for example, to all businesses in an industry, or to higher-level industry groups. Combining totals like this improves the statistical quality of the estimates and reduces the risk of disclosure. It is at the aggregated level that the statistical disclosure control will be carried out. The first step is to identify whether data could be disclosive, that is whether there is a risk that information about an individual business could be identified.

In the discussion in this section, a "cell" refers to an element of a published table, containing the aggregated data (as described previously), not to the sampling strata described in Section 3.2. For tables of total values published by the APS, there are two criteria that must be met for the published value to be deemed non-disclosive. These criteria are:

- minimum threshold rule: this rule states that there must be at least n reporting units (businesses) in a cell
- $p\%$ rule: this rule states that the total contribution of the m largest respondents in the cell aggregated total must be less than $p\%$ of the total in that cell

The values of n , m and p should remain confidential. Knowing these values could allow information on individual businesses to be calculated.

In this example, there are 10 businesses in a cell, of which four have returned their total purchases estimates, and n equals three, m equals one, and p equals 95%.

Table 2: Example of disclosure control

Business	A	B	C	D	E	F	G	H	I	J
Total purchases (£ thousands)	20	30		5			1500			

The following two criteria are applied to the data:

- threshold rule: there are four businesses that have reported values; the minimum threshold, n, is three, so the cell is not disclosive under this rule
- p% rule: total returned purchases of the cell equals £(20 plus 30 plus 5 plus 1,500) thousand equals £1,555 thousand, m is one, and the largest respondent is business G, with a total purchases value of £1,500 thousand

So, the percentage contribution of business G to the total purchases in the cell is calculated as follows:

$$\left(\frac{1,500}{1,555}\right) \times 100 = 96.5\%$$

In this case, 96.5% is greater than 95%, so under this rule, the cell is disclosive.

As the cell has not met both criteria, it is identified as a disclosive cell, and disclosure control methods must be applied before the data can be published.

7.5 Disclosure control methods

Standard techniques for controlling statistical disclosure are used for the release of the APS results. These are described in this section.

Cell suppression

Cell suppression is the standard method used to protect tables with disclosive cells. The disclosive cells are suppressed, that is, they are not published. This is known as primary suppression. Other, non-disclosive cells must sometimes also be suppressed, to prevent the values of the primary suppressed cells from being calculated by subtraction of all the other cells from the total. These are known as secondary suppressions. This is the method used by the APS (and other surveys) to suppress disclosive values.

Merging of cells and cell aggregation

Cells may also be combined to prevent publication of disclosive data. For example, where there are very few industries in a specific sector a higher industry classification will be used instead.

Rounding

Monetary estimates in the APS statistical bulletin dataset are rounded to the nearest £ thousand. Percentages are derived from the unrounded values and then rounded.

8 . Revisions policy

The Office for National Statistics (ONS) ensures that published estimates are as accurate as possible. However, if significant changes are made to source data after publication, then estimates will be revised. The ONS has a clear [Policy on how revisions are handled](#) across the organisation.

8.1 Planned revisions

Planned revisions usually arise from either the receipt of additional data from late responding businesses or the correction of errors to existing data by businesses responding. In general, it is anticipated that alongside figures for the current year, revised estimates will be published for the previous year's estimates. It should be noted that during the coronavirus (COVID-19) pandemic, processing for the 2019 and 2020 Annual Purchases Survey (APS) was deprioritised and data for these years have not been published.

8.2 Unplanned revisions

In addition to planned revisions to the current and previous survey years, additional unplanned revisions may be published if they are considered to be large enough and of sufficient interest to users such that a delay until the next standard release is not justifiable, or if they effect data in more than just the current and previous survey years. The timing with which these revisions are released will take into account:

- the need to make the information available to users as soon as is practicable
- the need to avoid two or more revisions (to the same data items) in quick succession, where this might cause confusion to users

All unplanned revisions will be released in compliance with the same principles as other new information.

9 . Product list (as at 2023 collection)

Energy, water and waste products

- Coal and lignite.
- Crude petroleum and tar sands.
- Natural gas, liquefied or in gaseous state.
- Wood.
- Diesel.
- Petrol.
- Gas oils.
- Kerosene.
- Liquefied petroleum gas (LPG).
- Wood products.
- Other biomass.
- Other coke and refined petroleum products not included elsewhere.
- Electricity.
- Gas supply from mains.
- Steam and air conditioning.
- Water supply from mains.
- Water treatment services.
- Sewerage services.
- Waste collection services.
- Waste treatment and disposal services.
- Material recovery services.
- Remediation services and waste management services.

The energy section of the questionnaire (particularly in relation to coke oven and refined petroleum products) has changed between 2020 and 2023. Further detail is given in the notes pages of the statistical bulletin datasets for each year.

Services

- Mining support services.
- Printing and reproduction services of recorded media.
- Repair and installation services of machinery and equipment.
- Repair and maintenance services of ships and boats.
- Repair and maintenance services of aircraft and spacecraft.

- Construction of residential and commercial buildings.
- Construction and construction works for civil engineering; roads, railways, bridges and tunnels.
- Specialised construction, demolition and site preparation; electrical and plumbing installation, building completion work.
- Repair and maintenance services of motor vehicles and motorcycles.
- Transport services via roads and pipelines.
- Passenger and freight rail transport services.
- Water transport services.
- Air transport services.
- Warehousing and transportation support services.
- Postal and courier services.
- Travel accommodation services.
- Food and beverage serving.
- Publishing services of books, newspapers, journals, advertising space in books and newspapers and software publishing services.
- Motion pictures, video and television programme production services, sound recording and music publishing.
- Programming and broadcasting services.
- Telecommunication services.
- Computer programming, consultancy and related services.
- Information services; data processing, web hosting and IT infrastructure provisioning.
- Financial services.
- Insurance and reinsurance services.
- Support services to financial and insurance activities.
- Real estate services on a fee or contract basis.
- Renting, buying or selling of property.
- Legal services.
- Accounting, bookkeeping and auditing services.
- Head office and management consulting services.
- Architectural, engineering and technical testing services.
- Scientific research and development services.
- Advertising and market research.

- Specialised design services.
- Veterinary services.
- Rental and leasing services.
- Employment services.
- Travel agency, tour operator and other reservation services and related services.
- Security and investigation services.
- Buildings and landscape services.
- Office administration, office support and other business support services.
- Public administration and compulsory social security services.
- Education and training services.
- Human health services.
- Residential care services.
- Social work services without accommodation.
- Creative, arts and entertainment services.
- Library, archive, museum and other cultural services.
- Gambling and betting services.
- Sporting, amusement and recreation services.
- Membership organisation services.
- Repair services of computers, personal and household goods.
- Personal services, dry cleaning and laundry, hairdressing and beauty, funeral and related services.

Goods

- Agriculture, hunting and related products and services.
- Forestry, logging and related products and services.
- Fish and other fishing products, aquaculture products, support services to fishing.
- Metal ores, iron, non-ferrous metal, uranium and thorium, copper, nickel, aluminium, lead, zinc and tin ores.
- Mining and quarrying products; ornamental and building stone, gravel, sand and clays, precious and semi-precious stones.
- Preserved meat and meat products.
- Processed and preserved fish products.
- Processed and preserved fruit and vegetables.

- Vegetable and animal oils and fats.
- Dairy products.
- Grain mill products, starches and starch products.
- Baked goods; bread, fresh pastry goods and cakes, rusks and biscuits and flour-based products.
- Other food products; sugar, cocoa, chocolate and sugar confectionery, vinegar, sauces, processed tea and coffee.
- Prepared animal feeds.
- Soft drinks and bottled waters.
- Alcoholic beverages, distilled and fermented; wine, cider, perry, mead, vermouth, beer and malt.
- Tobacco products; cigars, cheroots, cigarillos, cigarettes and cured, stemmed or striped tobacco leaves.
- Textile articles; carpets and rugs; preparation services of natural textile fibres; textile finishing services.
- Clothing and wearing apparel; fur, leather, work wear, underwear, hats and headgear
- Footwear and leather products.
- Wood and products of wood and cork, except furniture; articles of straw and plaiting materials.
- Paper and paper products.
- Paints, varnishes and similar coatings, printing ink and mastics.
- Soaps and detergents, cleaning and polishing preparations, perfumes and toilet preparations.
- Chemical products; explosives and related products, glues, essential oils, aromatic distilled waters.
- Dyestuffs and agro-chemicals.
- Petrochemicals; organic basic chemicals, plastics in primary form, synthetic rubber and man-made fibres.
- Industrial gases, inorganic basic chemicals and fertilisers.
- Basic pharmaceutical products and preparations.
- Rubber products.
- Plastic products.
- Glass, clay, porcelain, ceramic and refractory products.
- Cement, plaster and concrete.
- Basic, precious and non-ferrous metals and casting services.
- Basic iron, steel and ferro-alloys including tubes, pipes, hollow fittings and other products of the first processing of steel.
- Fabricated metal products.
- Weapons and ammunition.

- Computers and electronics.
- Electrical equipment, electric motors, generators, transformers and parts.
- Machinery and equipment.
- Motor vehicles, trailers and semi-trailers.
- Boats, ships and floating structures.
- Railway locomotives and rolling stock.
- Aircraft, spacecraft and related machinery.
- Military fighting vehicles.
- Motorcycles, bicycles and carriages.
- Furniture; office, shop and kitchen furniture, upholstering and finishing services of new furniture and parts of furniture.
- Medical and dental instruments.
- Toys, games and musical instruments.
- Safety wear and sports goods.
- Jewellery and decorative goods.
- Brushes, except electric motor brushes.
- Pens and writing implements.

10 . Related links

[Annual Purchases Survey \(APS\) Quality and Methodology Information \(QMI\)](#)

QMI | Revised 8 August 2023

Quality and Methodology Information for the Annual Purchases Survey, detailing the strengths and limitations of the data, methods used and data uses and users.

[Development of the Annual Purchases Survey](#)

Article | Released 15 December 2017

Article providing detail on the development and reinstatement of the UK Annual Purchases Survey, detailing goods and services used by businesses in the production process, otherwise referred to as intermediate consumption.

For further information, please contact the Annual Purchases Survey Team by email at abaps@ons.gov.uk.

11 . Cite this methodology

Office for National Statistics (ONS), released 8 August 2023, ONS website, methodology, [Annual Purchases Survey: technical report](#)