

Article

VAT turnover implementation into National Accounts: November Update

Provides details of how value added tax returns will be used in quarterly national accounts from December 2017 onwards

Contact:
Mark Stephens
vatdev@ons.gsi.gov.uk
+44 (0)1633 456387

Release date:
20 November 2017

Next release:
To be announced

Table of contents

1. [Introduction](#)
2. [Background](#)
3. [How will we use VAT?](#)
4. [Data Methods](#)
5. [National Accounts methods](#)
6. [Which industries will use VAT?](#)
7. [Next steps](#)
8. [Authors](#)

1 . Introduction

Office for National Statistics has previously announced its plans to use Value Added Tax (VAT) returns in the calculation of its quarterly national accounts releases by the end of 2017. This article gives details about exactly how VAT data will be used.

Initially, ONS will begin to use VAT data for small & medium-sized businesses for all industries covered by the monthly business surveys excluding retail, which cover around 25% of monthly output. In addition, VAT will initially only be used to estimate growth rates, with the overall level of output still derived from the Annual Business Survey.

Over time the majority of turnover will be recorded using VAT returns, with monthly surveys only used to cover the largest businesses.

2 . Background

Office for National Statistics (ONS) is committed to addressing the strategic recommendations made in the [Bean Review of UK Economic Statistics](#), in line with our [Economic Statistics and Analysis Strategy \(ESAS\)](#). As part of this commitment we will use VAT turnover data within the quarterly national accounts for the first time in December 2017.

The use of the VAT turnover dataset is one of the first steps towards transforming the way that we use large externally-collected administrative data instead of to data collected via ONS survey. With the development and implementation of the VAT turnover dataset, we are helping to realise the potential of new technologies and creating methods that could deliver useful statistical data from a range of other externally collected sources, including those from administrative data sources.

In December 2017, VAT turnover will be used to estimate parts of the output measure of gross domestic product (GDP(O)) for the first time.

Following an internal review of our methodology and consultation with stakeholders, academic associates and international experts, we have agreed to combine output estimates from both the Monthly Business Survey (MBS) and the newly developed VAT turnover dataset. MBS data pertaining to the largest businesses will continue to be used as a timely and effective measure of these entities. In addition to this, VAT turnover data will initially be employed to improve coverage of smaller businesses for selected industries. In light of lessons learnt through delivering this, VAT turnover data is being considered for further strategic deployment across relevant statistics within the national accounts but also in the [transformation of short-term turnover statistics](#).

3 . How will we use VAT?

The following [implementation principles for using VAT turnover in the national accounts](#) were published in June 2017

VAT turnover data will be used to supplement Monthly Business Survey measured industries only.

The Monthly Business Survey for UK Production and Great Britain Services (MBS); the Retail Sales Inquiry (RSI) and the Monthly Business Survey for Construction and Allied Trades all collect a variation of turnover as their main variable. Thus these are more comparable with VAT turnover data than other sources used as a proxy in measuring output.

VAT turnover data will be used for selected employment size bands only.

VAT returns do not always directly correspond with single reporting units as defined by the Inter- Departmental Business Register (IDBR). In fact, multiple VAT returns may be submitted for an enterprise consisting of multiple reporting units. This leads to difficulty allocating turnover to different industries covered by the same return.

More than 90% of reporting units have a one-to-one correspondence with a single VAT return. However, as turnover is more concentrated in businesses with more complex business structures, only 20 to 40% of turnover is reported by “simple” reporting units with a single VAT return. Most commonly, these businesses will belong to a small employment size band. With this in mind, we have taken the decision to use VAT to replace the small employment size bands only.

Only relative changes of VAT turnover over time will be used

Analysis of VAT turnover data across industries shows the nominal level to be different from both the MBS and Annual Business Survey (ABS). As such, the quarter on previous quarter growth rates from the VAT data at size-band level will be applied to the MBS nominal level.

Short term output indicators indices will be constrained to quarterly national accounts datasets With the exception of RSI.

The revisions policy for short term indicators will be revised to minimise the frequency of revisions and ensure consistency with the accounts.

To minimise the number of revisions to Output in the construction industry and Index of Production, a new revisions policy will be introduced. The changes to how the back series of these two economic indicators will be revised are detailed in Table 1.

Table 1: Summary of the proposed changes to the national accounts revisions policy, UK, 2017

Date	Publications	Data sources	Current revisions periods	New revisions periods (from December 2017)
8 December 2017	Index of Production (Oct 2017)	Survey data	Revisions open January 2016 to September 2017	Closed back series
	Index of Construction (Oct 2017)			Only October 2017 data is provided
22 December 2017	Index of Services (Oct 2017)	Survey data and VAT data	Revisions open January 2016 to September 2017	Revisions open January 2016 to September 2017
	Quarterly National Accounts (July to Sept 2017)			
10 January 2018	Index of Production (Nov 2017)	Survey data and VAT data	October to November 2017 open to revision	October to November 2017 open to revision
	Index of Construction (Nov 2017)			Back series constrained to quarterly national accounts

Source: Office for National Statistics

Quarterly national accounts published on 22 December will include data from Index of Services based upon survey and VAT returns. It would also normally include data for the Production and Construction sectors consistent with their 8 December publication but these would be based only on survey returns. However, the data for Production and Construction published on 22 December will instead be based upon survey and VAT returns and so will differ from that published on 8 December. On 10 January the Production and Construction bulletins will publish data consistent with that published on 22 December.

4 . Data Methods

In order to process VAT data a new statistical processing platform has been developed with new methodologies to support.

Matching and linking

VAT data is matched against a corresponding IDBR snapshot each month using a unique identifier – the VAT registration number of each business. This process enriches the source data with features from the IDBR that can be used in later processing of VAT data. Matching rates are over 99% each month.

Apportionment

As has already been discussed, VAT returns can often be matched to multiple reporting IDBR Reporting Units. To apportion the VAT return to reporting units, frozen employment is used as the auxiliary variable and assumes a linear relationship between VAT turnover and employment.

Content editing

This cleaning of VAT data in itself provides a challenge in processing the data, in that we cannot follow the same methods used for MBS data and contact the business to confirm their return. As such we have developed the following cleaning and suspicious turnover rules that provide us with the opportunity to take evidence-based action should this be needed.

Thousand pounds rule

Businesses should report turnover to HM Revenue and Customs in pounds, but some report in thousand pounds. This rule calculates the ratio between the current return and the previous return. The flagged VAT unit turnover is corrected by multiplying the value by 1,000.

Quarterly pattern rule

Identifies quarterly reporters (stagers 1, 2, and 3) not following a “true” quarterly pattern. There are three variations of this rule:

- VAT unit reports the same turnover value in any three consecutive quarters followed by a different value in the fourth quarter (x,x,x,y)
- VAT unit reports the same turnover value in any four consecutive quarters (x,x,x,x)
- VAT unit reports a turnover value of zero in any three quarters followed by a positive value in the fourth quarter (0,0,0,x)

Correct any turnover that follows a quarterly pattern by redistributing the turnover for that year in the same seasonal pattern to other records from the same industry. The quarterly medians of the values from the same industry are used for this correction.

First returns

Cleaning can be complicated where no previous return exists as an accepted comparator. There are two types of first returning businesses:

- VAT units, where additional return data for later periods are available
- VAT units, where no additional return data for later periods are available

The code flags the largest 5% of first returners for further investigation.

Suspicious turnover rule

This calculates whether the return is 33% larger or smaller than the previous returned turnover. Such returns are flagged but not adjusted.

Estimation

Ratio estimation using the IDBR frozen turnover variable is the estimation method selected for those cells that do not have a 100% response rate. This method is already used within current ONS surveys by:

- calculating the amount of real turnover that has arrived in each cell, relative to the proportion of IDBR frozen turnover for those businesses that have provided a return in the cell
- the ratio is multiplied by turnover value for the cell to provide an estimate for real turnover missing from the cell for the target period

The estimation is assessed by applying the method as if it were a set point in the past and comparing the calculated estimate with the known “truth” when looking at the data in the present day.

Calendarisation

Seasonal and trading days are used to calendarise the data. This method proportionally allocates quarterly and annual returns to months. The proportions are cell-specific and are derived from seasonal and trading day components estimated from cell-level MBS time series.

VAT returns adopt a standard period through this approach, enabling a subsequent monthly estimation process.

After processing the VAT returns at a micro level, data is aggregated using the Standard Industrial Classification (SIC) and size band structure used for MBS data. This aggregate data is then passed to the Central ONS Repository for Data (CORD) for processing into the national accounts.

5 . National Accounts methods

In order to use VAT within the national accounts, the following steps have been taken:

- bring together Monthly Business Survey (MBS) and VAT raw data by size band , non-seasonally adjusted on a pounds millions basis; the level of industrial detail mirrors the current processing systems with five-digit Standard Industrial Classification (SIC) for construction, most commonly three-digit for production and two-digit for services
- aggregate the MBS and VAT raw data so that both are on a quarterly and annual basis calculate annual growth rates for both the VAT and MBS data
- create a new “VAT consistent” annual series that equates to growth in the MBS prior to the join year and VAT growth in the year (and subsequent years) that VAT data are taken on
- create a quarterly “VAT consistent” series, which is the quarterly VAT data benchmarked (using BenchCD) to the annual “VAT consistent” series; in a case where quarterly VAT data are available for an incomplete calendar year, growth in the quarterly tail equates to VAT growth
- create a monthly “VAT consistent” series, which is the MBS data benchmarked (again using BenchCD) to the quarterly “VAT consistent”; if MBS data are available for a later time period than VAT, growth in the monthly tail equates to MBS growth
- use an industry selection matrix to choose either MBS or VAT data for any given industry by size-band combination
- run data through the remaining part of the monthly gross domestic product, Index of Production, Index of Services and construction systems as per usual

6 . Which industries will use VAT?

All industries covered by the Monthly Business Survey (MBS) were deemed to be in scope for using aggregate VAT data. At this stage in the development of the data source we have not considered using VAT to cover non-MBS industries.

The MBS is split into five sampling stratum based on the registered employment of each business on the [Inter-Departmental Business Register \(IDBR\)](#). Band 1 represents the smallest businesses in a given industry and the sample is selected at random from the universe of businesses contained on the IDBR for the sampling period. Bands 2 and 3 similarly represent small and medium-sized businesses, which are selected at random although for a few small industries Band 3 can also be a census of activity. Band 4 represents the largest businesses in a given industry and the selection constitutes a census of businesses in excess of the employment cut-off for that industry. In addition, businesses below the census cut-off with turnover in excess of £60 million can represent a significant part of an industry and so these undergo a forced selection at Band 5. There are relatively few of these businesses.

It should be made clear that although Band 1 selections for many industries are confined to the 0 to9 employment size-band, for some industries Band 1 covers 0 to4 employment. Similarly Bands 2 and 3 can differ in make-up between different industries in terms of the employment size-bands used to delineate the Bands.

At this stage in the development of the VAT data source we have adopted a principle that the large Band 4 businesses would not be in scope. This is based on the acceptance that MBS data is available more quickly than VAT.

In assessing the VAT data we have based our analysis on the following criteria.

Fit

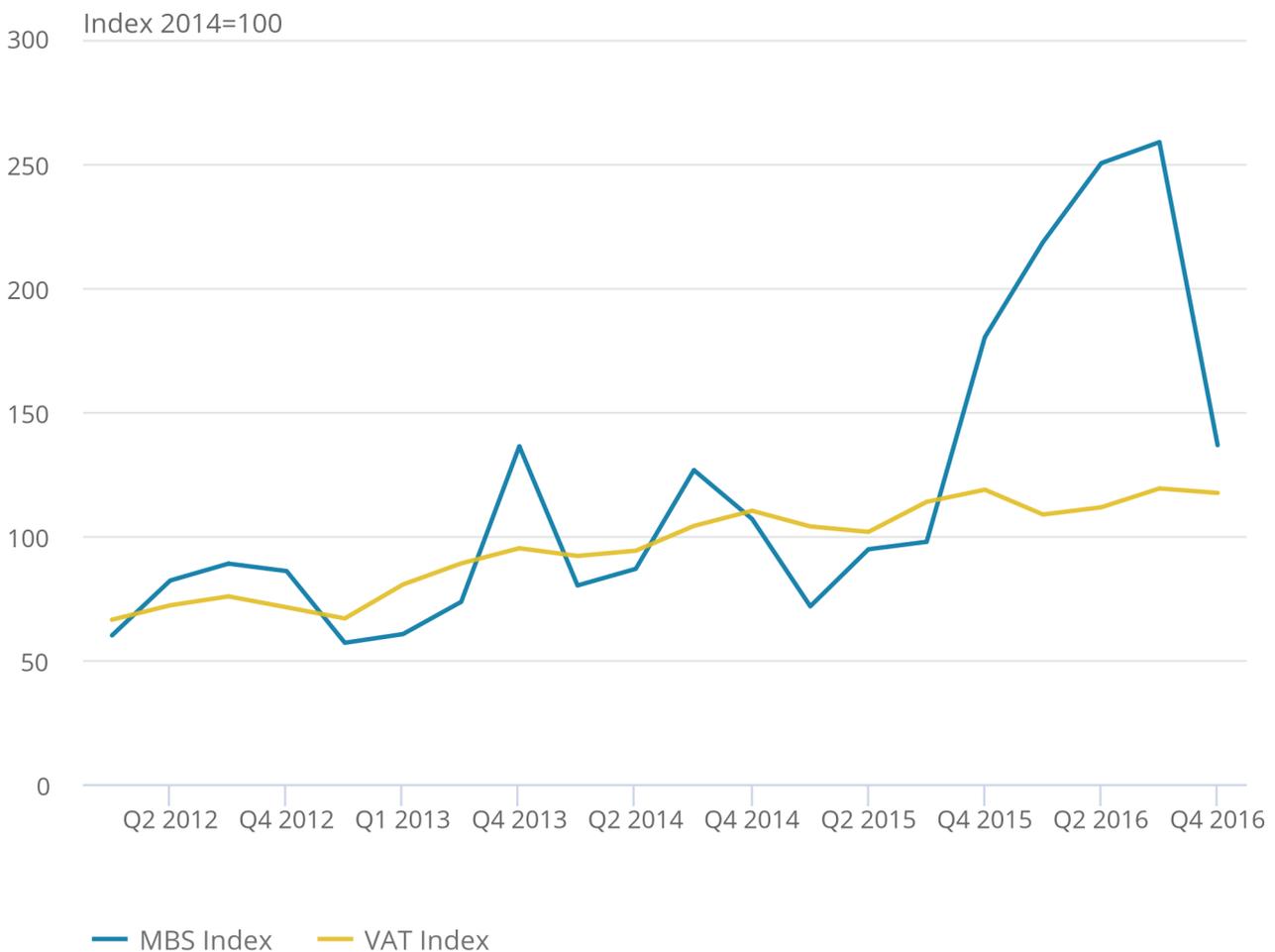
How do the VAT and MBS data correlate at a cell level? A degree of caution was needed when working towards this criteria as a weak correlation between VAT and MBS data meant in-depth investigation into the reasons why the two sources were different was needed. This investigation highlighted areas of concern with both sources, with the main one being the volatility seen in the MBS data largely as a result of sample rotation. Figure 1 provides an example of sample rotation, where the MBS data experiences a level shift for several months as a result of one respondent and then returns to its original level as the respondent is rotated out of the sample.

Figure 1: Example of rotation effects within Monthly Business Survey data

(Demolition and site preparation)

Figure 1: Example of rotation effects within Monthly Business Survey data

(Demolition and site preparation)



Source: Monthly Business Survey - Office for National Statistics and Value Added Tax Returns - Her

Source: Monthly Business Survey - Office for National Statistics and Value Added Tax Returns - Her Majesty's Revenue and Customs

Impact

This analysis focused on the impact to published aggregates when VAT data was used. In some cells, VAT had been chosen as a suitable source to measure output, but after aggregation with MBS data, deflation and seasonal adjustment the impact on these published aggregates was significant and could not be explained in enough detail to justify the use of VAT.

Revisions performance

This assessment has involved analysing the size of the revisions each quarter as a result of implementing VAT, their direction and frequency.

Data journey processing statistics

Here we have assessed what processes the data has been through, for example, what cleaning rules have been used, the extent and impact of estimation and has any apportionment been applied?

Microdata

Our final assessment has been to examine individual responses, how they compare with MBS responses if available, how they are distributed within a cell and the back history of the company's returns.

In working through these criteria and particularly around assessing impact, it was apparent that the conceptual differences between VAT turnover and retail sales turnover resulted in significant changes to the gross value added (GVA) of this industry. As a result we have taken the decision to not use VAT in the measurement of retail output in the quarterly national accounts.

Extent to which VAT data is being used to measure production, construction and services

Table 2 provide a summary of the number of cells and industries where VAT will be used, further it provides an estimate of the gross value added (GVA) coverage. Table 3 provides a summary of the number of VAT units that will be used to supplement monthly business survey (MBS) data.

Table 2: Industry coverage through the use of VAT data

	Number of cells where VAT will be used	Number of cells available	Number of SIC 07 industries where VAT will be used	Number of SIC 07 industries available	GVA coverage (parts per thousand)	GVA Coverage Available
Production	139	240	64	80	21.4	34.8
Construction	28	56	9	14	8.7	32.5
Services	77	192	35	64	45.9	132.9
Total	244	488	108	158	76.0	200.2

Source: Office for National Statistics

The [reference table VAT selection matrix](#) provides details of the industries and sizebands where VAT will be used.

Table 3: Change in sample sizes as a result of using VAT

	Sample size in available cells	VAT records available
Production	1,900	81,000
Construction	2,500	86,250
Services	3,800	436,500
Total	8,200	603,750

Source: Office for National Statistics

7 . Next steps

A significant amount of VAT data has yet to be assessed and there are areas where the comparison between VAT and survey data must be more fully explained. In that sense we are still developing our knowledge of the dataset and its applicability across further cells and industries. Similarly we should be clear that the new VAT processing system is in its early stages of development. Here too there are areas such as apportionment where we have identified the need to improve upon our current methods.

We will continue to assess and quality assure VAT data, both for the cells and industries selected for inclusion in December and those where we have currently decided we will not proceed. The scope for inclusion will be reviewed in time for Blue Book 2018 (BB18) with a view to expand our scope.

In addition we are also committed to review the use of VAT outside the monthly business survey (MBS). Table 4 lists industries where the use of VAT turnover data may be helpful to improve the quality of our estimates. Our initial interest will be in considering agriculture, forestry and fish; commercial property; financial auxiliaries; rail, sea and air transport; iron and steel. The remaining industries are of lesser importance.

Table 4: Industries under review for future use of VAT

Industry	Current Source	GVA weight
Agriculture, forestry and fish	Agricultural indicators and annual harvest estimates; annual forestry accounts and deep sea fishing data	0.70%
Social Care	Government expenditure and ONS jobs data	1.90%
Mining and quarrying	BEIS volume data and ONS jobs data	1.60%
Energy	BEIS volume data	1.50%
Commercial property	Investment Property Databank	1.40%
Financial auxiliaries	Financial market indicators; output of the Insurance and pension funds industry	1.30%
Rail, sea and air transport	Office of Rail and Road; Chamber of Shipping and Department for Transport; Civil Aviation Authority	1.20%
Others		1.30%
- Gambling		
- Non-state schools		
- Oil refineries	BEIS	
- Iron and steel	International Steel Statistics	

Source: Office for National Statistics

Finally, we have completed a discovery stage for the transformation of the Retail Sales Index. Within this transformation we plan to use VAT to measure monthly retail sales alongside a rationalised MBS for the retail, wholesale and motor trades industry. More information on this discovery can be found in the article [Transformation of short-term turnover statistics](#).

8 . Authors

Mark Stephens, John Allcoat, Kate Davies, Office for National Statistics