

Article

Transforming short-term turnover statistics: October 2017

An overview of the proposed transformation of short-term turnover statistics.

Contact:
Jayne Edwards
DCTP@ons.gsi.gov.uk
+44 (0)1633 455184

Release date:
23 October 2017

Next release:
To be announced

Table of contents

1. [Introduction](#)
2. [Transformation of short-term turnover statistics](#)
3. [Planned research](#)
4. [Business spine and partitioning](#)
5. [Use of administrative data](#)
6. [Features of the statistical design and data collection](#)
7. [Outputs](#)
8. [Timeliness](#)
9. [Next steps](#)

1 . Introduction

Official statistics play an important role in informing and monitoring policy decisions, providing a firm evidence base for decision making within and outside government. Recognising the need for a responsive statistical system that can keep pace with rapid changes in society, “Better Statistics, Better Decisions” sets out the UK Statistics Authority strategy for a statistics system exploiting the potential of a range of rich data sources, technology and methods developments to continue to meet the information and evidence need now and in the future.

Alongside the continued production of existing statistical outputs, Office for National Statistics (ONS) is taking forward a programme of transformation to deliver improvements to the UK's economic statistics. Aligned with the strategic direction and wider transformation agenda, this programme aims to make use of alternative administrative and commercial data, rebalancing data collection activities significantly toward wider, more integrated use of non-survey sources. Delivery will be underpinned by systems that make use of common architecture for collection, processing and analyses, with remaining survey operations becoming more efficient through rationalised designs and a move to online collection. New methods and innovative approaches will be applied to enhance existing outputs delivering better information about the UK economy.

Transformation will improve quality and timeliness in outputs and ensure we continue to meet the ongoing user needs for relevant and more detailed information and greater insight into economic statistics, complying with the recommendations of the [Bean Review](#). Other benefits include a reduction in burden on businesses responding to ONS survey requests and in operating costs.

2 . Transformation of short-term turnover statistics

A focus of early research and development in the transformation of economic statistics is short-term turnover outputs where the potential to redesign existing methods, replacing some direct data collection with use of administrative data, is being explored. In the case of short-term turnover statistics we are exploring the potential use of HM Revenue and Customs (HMRC) Value Added Tax (VAT) data. This builds on work underway to integrate aggregate ¹ VAT data into the national accounts for December 2017.

Research will extend this work to assess the potential to integrate unit level ² administrative data via a statistical “business spine”. Short-term turnover statistics will provide an initial application of an approach that may be applied more widely to other economic sectors and outputs to enhance the production of existing statistical outputs and enable new statistical analysis and insights.

Figure 1: Transformed short-term turnover statistics design

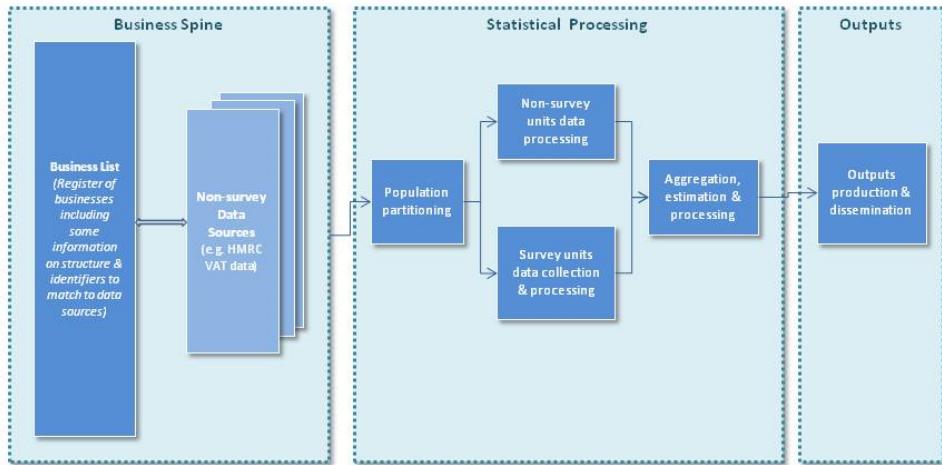


Figure 1 outlines the proposed design for transformed short-term turnover statistics. The business population will be selected from the Statistical Business Register (SBR), which defines the population of business units. The SBR will also hold identifiers against each business in the population to facilitate the identification and linking of records held on available data sources (both survey and non-survey sources) to this business population. We will then link VAT turnover for the relevant period to the spine. This business population will then be partitioned into a subset of businesses for which VAT data will be used as the basis of estimates – non-survey partition, and those for which survey data will continue to form the basis of outputs – survey partition.

Research will be carried out to identify appropriate methods for partitioning the population and for producing estimates for the non-survey population partition. Further research will explore the potential to improve the survey design and information collected for the survey partition of the population. Data for the survey and non-survey population partitions will be combined and processed using new methods to produce transformed statistical outputs. This article discusses early thinking on some of the main features of the transformed design and outlines the associated challenges to be explored through the planned research activities.

A staged approach to the transformation of short-term statistics is proposed with an initial focus on the retail trade sector (retail, wholesale and motor trades), to learn from this initial application of a transformed design before implementation across other sectors. Subject to fit with strategic priorities, we aim for an initial implementation for the retail trade sector in late 2018. Alongside work to transform turnover outputs, we are also researching the potential to make use of alternative administrative sources in the production of short-term employment outputs. This work is at an early stage and will be reported on separately as progress is made.

There are a number of benefits associated with an early focus on transformation of short-term turnover statistics and with a staged approach to implementation. These include cost and burden reductions, improved quality and early value.

Cost and burden reductions

Short-term surveys collect information from around 46,000 businesses each month (11,000 within the retail, wholesale and motor trades sectors) and there is potential for reduction in both respondent burden and operating costs from a reduction in survey sample sizes associated with use of administrative data and rationalisation of survey designs.

Improved quality

Move to use of administrative data for part of the population will increase coverage, reducing the volatility in estimates associated with survey sample rotation. Building on work already being carried out to integrate [HMRC VAT data into the National Accounts](#) estimates in December 2017, development of a business spine and integration of administrative data at unit level has the potential to support improved output delivery, for example, alternative breakdown of outputs or better integration with other data sources for improved economic analysis.

Early value

A staged approach to implementation supports delivery of early value with the initial implementation in the retail trade sector before wider application, while also providing opportunities for iteration and improvement ahead of wider implementation. This approach also aligns with the Framework Regulation Integrating Business Statistics (FRIBS) regulation to align the timing of retail trade outputs for retail, wholesale and motor trades, and extend coverage to UK rather than Great Britain estimates.

There are also a number of associated challenges both with a change in methods and with a staged implementation. In particular, it will be important to understand the impact on outputs, whilst a staged implementation requires transition arrangements to ensure continuity in existing outputs for other sectors and to avoid unnecessary additional burden on respondents. These issues are within scope of the planned research.

Notes for: Transformation of short-term turnover statistics

1. Aggregate data – administrative records that have been grouped together to provide totals for statistical use.
2. Unit level data – use of administrative records for individual businesses for statistical purposes.

3 . Planned research

An initial phase of research is now underway to explore important features of the proposed design and assess the viability of some of the proposed changes. In particular this early stage of research will consider the following:

- business spine and partitioning – initial research will consider linking Value Added Tax (VAT) units to create the Business Index to populate the Statistical Business Register (SBR) and create a business spine; partitioning methods to determine when it will be appropriate to replace survey data with VAT records will also be explored
- use of non-survey data – a focus of this work will be development of methods to support use of VAT data in place of survey data for a partition of the population; the approach to creation of a monthly series from a largely quarterly data source will be explored along with development of estimation methods
- features of the statistical design and data collection – early research into the potential to improve the quality of the statistical design and associated outputs through some changes to the sampling approach and to information collected for the survey partition of the population will also be explored

All of these changes have the potential to impact on outputs produced and this is discussed further in the “Outputs” section of this article.

4 . Business spine and partitioning

Statistical Business Register

We are developing a new Statistical Business Register (SBR) as a replacement for the current Inter-Departmental Business Register (IDBR). The SBR will be populated using the Business Index, which links Value Added Tax, Pay As You Earn and Companies House data to provide a definitive list of the UK business population legal units and SBR statistical units and is a central point for linking business records together across and for use by government.

As a maintained and shared data source containing a complete list of businesses, the SBR forms the basis of the business spine and will also hold identifiers against each business that will allow records from a range of available data sources to be associated with individual business units on the SBR. This approach of unit level linking has the potential to provide a rich and detailed source of information for use in the production of statistical outputs and analysis. HM Revenue and Customs (HMRC) Value Added Tax (VAT) data will be integrated into this business spine, using the identifiers held on the SBR and datasource. Other potential sources for integration include Pay As You Earn (PAYE) Real Time Information (RTI), providing employment information, and commercial data sources providing financial information.

SBR will also be used for the creation of statistical sampling frames, with this integrated approach supporting the strategy of designing surveys to supplement use of non-survey sources. SBR will be maintained on a continuous basis, via data feeds from internal and external sources. Transformed retail trade will be the first output to select units for sampling via the SBR.

Population partitioning

An important area of research will be to determine criteria to identify the non-survey and survey partitions of the population. For the VAT partition of the population, where a VAT return provides the business turnover information, it will be important to ensure that there is a direct match between the VAT record and the business listed on the SBR, so it is likely that this partition will be limited to businesses with relatively simple structures.

Depending on the outcome of analysis, partitioning will differ between industries, due to the size, complexity and activities being carried out by businesses classified to the sector. Turnover information reported on VAT returns will also be more limited than information that can be captured by a survey, so it will also be important to ensure that partitioning ensures assumptions hold in cases where it is necessary to derive model-based estimates for the VAT partition using survey sample responses (when estimating internet sales, for example, where information is required for administrative purposes and so is not recorded on the datasource).

5 . Use of administrative data

Businesses that meet the Value Added Tax (VAT) threshold must provide data to HM Revenue and Customs (HMRC) in a VAT return, where they complete the “Total value of sales and all other outputs excluding VAT” value. This is comparable to turnover data collected on the Monthly Business Survey. Taken as a whole, the VAT source is a census of all businesses registered to pay VAT (at April 2017, VAT threshold is £85,000). While VAT and survey data are comparable, there are challenges associated with using VAT in place of direct survey data collection.

Timeliness

The UK has a legal obligation under the EU Short-term Business Statistics (STS) regulation to provide turnover data for retail estimates one month after the reference period month ($t+1$), however, VAT data are not available within this timeframe as Table 1 shows.

Table 1: Timing of submission of Value Added Tax returns from HM Revenue and Customs

Time from VAT reference period (t)	Average number of VAT returns received for the 6 months January 2016 to June 2016
t+1	1.2%
t+2	60.6%
t+3	95.9%
t+4	97.8%
t+5	98.6%

Note: "t" refers to the VAT period reference month

Source: Office for National Statistics

As a significant amount of VAT data are unavailable when we process Month 1 ($t+1$) and Month 2 ($t+2$) retail estimates and it is not until Month 3 that a high proportion of VAT returns are available, we need a method of estimating for missing data. At this early research stage we are researching and developing statistical methods, which will be used to estimate for the output production for the most recent periods (latest two months) when we will have an almost complete VAT dataset by Month $t+3$, that is, the third month when estimates for the reference period will be produced.

Calendarisation

VAT returns can be submitted to HMRC on a monthly, quarterly or annual basis with quarters and years covered by returns starting in any month with no specified reporting pattern for annual or quarterly returns. This means that monthly turnover values are not available for most units in the VAT population partition. Calendarisation is the process of converting the quarterly and annual returns to monthly estimates and methods could be applied at either a unit level or an aggregate level. Calendarisation enables the estimation process for production of monthly estimates for the VAT population partition. Table 2 shows the percentage of VAT returns submitted by time period.

Table 2: Percentage of Value Added Tax returns submitted by periodicity

Time period	Percentage of VAT returns received
Monthly	10%
Quarterly	90%
Annual	0.2%

Source: Office for National Statistics and HMRC

6 . Features of the statistical design and data collection

Coverage

The retail sales output is compiled using data from the Monthly Business Survey – Retail Sales Index (MBS-RSI) and covers businesses engaged in retail sales operating in Great Britain (England, Scotland and Wales); it does not include Northern Ireland, the Isle of Man or the Channel Islands. We are proposing to move from the production of a retail sales index to a retail trade index, combining the retail sales, wholesale and retail trade of motor vehicles sectors. This change in output coverage will fulfil the legislative requirement of the Eurostat Framework Regulation Integrating Business Statistics (FRIBS) project. This change will also bring the UK in line with other countries such as the US who already publish a monthly [Retail Trade Report](#).

It is the intention to also extend geographic coverage to include Northern Ireland enterprises, creating a UK index. There are a number of issues to be considered, specifically the output, respondent and operational impacts of moving retail from Great Britain to UK coverage.

The aim of the transformation is to produce outputs by sector, which give us an early estimate of the short-term picture of total UK economic activity on a monthly basis. Collecting data at enterprise level with a breakdown by business activity will enable us to provide a more complete picture as we will have a single response for each business (enterprise) setting out both total activity and the breakdown of this total by types of activity.

Enterprise sampling

The statistical unit for current short-term (and many other) surveys is a reporting unit. For the vast majority of businesses the reporting unit is equivalent to the enterprise. In other words, the respondent supplies information for the whole business. However, in some cases reporting units relate to a portion of the enterprise, supporting businesses preferences to report separately for different parts of the enterprise activity.

In a change to current methods we are researching application of sampling, data collection and estimation methods at enterprise level. This approach supports collection of total activity (turnover) for the business alongside a breakdown of this total into types of business activity from the enterprise in line with our output requirement. There are several potential quality benefits of using this approach, in particular:

- one selection with full coverage of units and outputs meaning that all activity, regardless of range of activities of the business, will be collected on a single survey response
- mitigates the risk of double-counting as the business will only receive one request for turnover information
- provides a better breakdown of activity as businesses will be required to provide details of all activity within the enterprise

The move to enterprise sampling also ensures full compliance with the Eurostat Short-term Business Statistics (STS) regulation.

However, we recognise there are challenges in adopting this approach, which will be explored during this research phase. In particular, it will be important to ensure that the design of a data collection strategy and online questionnaire facilitate business reporting of the information required. Consideration will also need to be given to managing the move from reporting unit and enterprise unit selection and reporting alongside plans for a staged implementation of a transformed design for short-term statistics. There is potential for under-recording of retail trade activity during this transition phase, associated with loss of reporting units from the retail sector if the main activity of the enterprise is not retail. The need for an approach that manages this potential discontinuity will be considered.

7 . Outputs

The retail sales output provides a short-term measure of the changes in volume and value of sales of goods by retail businesses in Great Britain. The main output measures include an index measure of value and volume estimates, in both seasonally adjusted and non-seasonally adjusted forms. The value estimates reflect the total turnover that businesses have collected over a standard period, while the volume estimates are calculated by taking the value estimates and adjusting to remove the impact of price changes. Implementation of the proposed changes in design is likely to result in a discontinuity in the output series.

We provide a split of small and large businesses as part of the retail sales output, however, as part of the move to enterprise sampling, the definition of a small and large business split could differ slightly and will be informed as the methods work on population partitioning progresses. While the use of HM Revenue and Customs (HMRC) Value Added Tax (VAT) data to replace direct data collection for a portion of the population increases overall coverage, it is likely to also result in a reduction in the level of detail available to publish for the latest period (i.e. Month 1) due to the time lag in provision of VAT returns. Current proposals are for top level estimates for the main store types to continue to be available at Month 1. These are shown in Table 3. A list of all stores types can be found [here](#).

Table 3: Retail sales top level store types and their weight

Publication category	Percentage of All Retailing	RSI Aggregate
Predominantly food stores	39.84	Agg1
Predominantly non-food stores	42.07	Agg12
Non-specialised stores	8.83	47.19
Textile, clothing and footwear stores	11.79	Agg5
Household goods stores	8.44	Agg7
Other non-food stores	13.02	Agg13
Non-store retailing	8.59	Agg14
Automotive Fuel	9.5	47.3

* Based on aggregated annual data from the Retail Sales Inquiry, 2016.

** UK Standard Industrial Classification 2007

Source: Office for National Statistics

In addition to the main retail store types, we will also publish data for the wholesale and motor trades sectors at Month 1. Detailed estimates for all industries will be available by the time we publish, that is at Month 3. This may include an additional split of motor trades not currently available, between sale of motor vehicles and sale, maintenance and repairs of motorcycles and related parts and accessories.

We understand this will be a significant reduction in the detailed data currently available at the time of first publication (Month 1); however, this approach is also likely to result in a more stable, less volatile output. Retail sales estimates for small businesses can show volatility and this can be due, in part, to selection of a relatively small subset of businesses and the impact of sample rotation when there is significant variation in business activity and turnover, depending on the goods they sell. The use of VAT data for simple businesses will result in a more stable estimate at Month 3 of this partition of the population. That is not to say we will not sample any smaller or simple businesses, but we will endeavour to use VAT data wherever possible for these units, with survey coverage focussed on validating assumptions made in the development of methods and to support the production of estimates for outputs not covered by VAT returns (for example, value of internet sales).

We are at a relatively early stage of thinking on changes to outputs and plan to consult more fully on this at a later date.

8 . Timeliness

The Retail Sales Index is released 18 or 19 days after the reference period and there are no plans to change this publication timetable. Data for wholesale and motor trades sectors, however, are published as part of the Index of Services around eight weeks after the end of the reference month, aligned with publication of gross domestic product (GDP). A move to production of a retail trade output will mean initial estimates for these sectors will be available earlier.

9 . Next steps

This article has outlined a proposed design for a future short-term turnover statistics system and set out some important areas of research. This initial phase of development will continue until late 2017, at which time we will have a firmer view on features of a transformed design and a timetable for implementation, allowing us to move forward with a clear steer on the direction of travel into the next phase of the project.

It is our intention, depending on priorities within the organisation that transformation of retail trade will happen in late 2018. We will continue to provide updates on developments and consult on changes to outputs as progress is made.