Quality Assurance of Administrative Data used in Trade Statistics

An assessment of administrative data sources for ONS trade statistics in accordance with the UK Statistics Authority Quality Assurance of Administrative Data (QAAD) toolkit.

Table of contents

1 . Summary of the quality of UK trade statistics

1.1 Overview of UK trade

UK trade shows the extent of import and export activity and is a main contributor to the overall economic growth of the UK.

The UK trade balance is the headline figure in the bulletin and is calculated as total exports less imports. The trade balance reflects the overall net position of the UK, describing whether the UK exports more than it imports (a trade surplus) or whether it imports more than it exports (a trade deficit).

Trade in services covers import and export activity across 12 categories: transport, travel, telecommunication /computer/information, construction, insurance, financial, intellectual property, other business, personal/cultural /recreational, government, manufacturing and maintenance and repair services. As with trade in goods, monthly movements are shown via the statistical bulletin, with increasing detail on a quarterly basis and a complete geographical breakdown released annually.

The compilation of UK trade aggregate estimates requires data that are collected from around 40 feeder sources including a variety of administrative sources. There is a challenge to establish and maintain quality assurance processes and relationships with these businesses or suppliers.

Trade in services in particular collate data from a large number of suppliers, many of whom provide the data on a voluntary basis. Trade in goods has one main administrative data supplier; HM Revenue and Customs (HMRC), which contributes approximately 90% of overall trade in goods.

In November 2014, following a series of errors (see annex B) the UK Statistics Authority decided to suspend the National Statistics status of UK trade, a reflection of an assessment against the Code of Practice for Official Statistics. Taken as a whole, the Code aims to ensure that official statistics meet the needs of users; are produced, managed, and disseminated to a high standard; and that statistics are well explained. In this case the Authority concluded that this overall objective of the Code was not being met at that time.

It was apparent that some of the errors originated from administrative source data supplied to UK trade. Therefore, the implementation of this quality assurance document is primarily to address the issues that concern the quality of administrative sources of data for UK trade.

The primary aim of this document is to provide clear information about the quality assurance measures carried out on administrative data that are used in the compilation of UK trade.

1.2 Our assessment of the quality of UK trade

Following the issues surrounding the quality of our administrative data sources the UK trade team set out to establish the strengths and weaknesses of the administrative data inputs.

Using the risk and profile matrix set out in the UK Statistics Authority toolkit, the UK trade team have concluded that UK trade outputs are fit for purpose; this judgement is based on quality assurance process and actions carried out by the UK trade team and on our evaluation of the level of risk of quality concerns and public interest in the data sources and outputs of UK trade.

Table 1 of Information on quality of trade statistics gives a more detailed summary of the quality assurance of administrative data in trade statistics.

The following sections outline how our processes meet the criteria required for the basic, enhanced and comprehensive levels of quality assurance, drawing on the four practice areas as set out by the UK Statistics Authority toolkit: operational context and data collection; communication with data suppliers; quality assurance principles; standards and checks; and quality assurance investigations carried out for enhanced and comprehensive assurance.

The quality assurance process carried out draws on the practice models developed by the UK Statistics Authority to ensure that the assessment of data sources is proportionate to the level of public interest in UK trade and the scale of the risk over the quality of data.

2 . How we have ensured the quality of UK trade

2.1 Overview of survey data sources

Survey data is information collected specifically for our purpose and we have established practices in place for the collection, production and dissemination of these estimates. A high level of quality assurance measures are maintained for these data and service level agreements (SLAs) are in place between UK trade and the survey data suppliers. The SLA provides us with a certain level of assurances of the data and covers a range of expectancies such as maintaining response rates, achieving delivery targets and providing accompanying briefings. The relationship built between survey outputs and National Accounts means we have a responsive culture of resolution enabling us to query dubious data following delivery as well as the ability to input into methodological and questionnaire changes. Close communication with the data suppliers ensures that the strengths and weaknesses of the underlying data sources are well understood. In some cases, this has led to reduction to the burden on suppliers, where it was discovered extra data was being supplied that wasn't being used in the UK trade estimates.

2.1.1 Trade in Goods

Trade in goods statistics are considered to be one of the strongest data sources across economic statistics.

Statistics of the UK's overseas trade in goods have been collected for over 300 years by HMRC, formerly HM Customs and Excise (HMCE). Since 1993 these data comprise statistics of UK imports from, and exports to, countries outside the EU, and statistics on trade with other EU member states collected via the Intrastat survey. Data are compiled from declarations made to HMRC by importers, exporters or their agents, and statistics of UK arrivals (imports) from and dispatches (exports) to other member states of the EU compiled from the Intrastat returns submitted by traders or their agents to HMRC.

Information on trade with EU countries

The Intrastat system has applied since 1993, with minor variations in all EU member states. In the UK all VAT registered businesses are required to complete 2 additional boxes on their VAT returns, which are normally submitted quarterly. These show the total value of exports of goods to customers in other member states (dispatches) and the total value of imports of goods from suppliers in other member states (arrivals).

Traders whose annual value of arrivals and/or dispatches that exceed the given thresholds are required to provide an Intrastat declaration each month showing full details of their arrivals and dispatches during the month. These thresholds are reviewed annually. For the calendar years 2010, 2011 and 2012 the threshold for arrivals has been set at £600,000, to capture a reduced coverage of 95% of trade by value. The threshold for dispatches has been set at £250,000 for this period to maintain coverage at 97%.

Information on trade with non-EU countries

In general the figures for trade with non-EU countries show: the trade as declared by importers and exporters or their agents, and for which documentation has been received and processed by HMRC during the calendar month.

Importers are usually required to present a Customs declaration before they can obtain Customs clearance and remove the goods. The great majority of imports are cleared immediately by a computerised system. Import statistics may include documents received by HMRC up to the third working day after the end of the month. Therefore, the import figures correspond fairly closely to goods actually imported during the calendar month.

All information on the statistics, data collection and quality assurance is published on the HMRC dedicated trade statistics website.

Information on trade in non-monetary gold

Trade in gold (that is, gold bullion, gold coin, unwrought or semi-manufactured gold and scrap) is reported to HMRC but it is excluded from the statistics of total exports and imports published in the Overseas Trade Statistics. However, trade in ores and concentrates and finished manufactures of gold (for example, jewellery) are included in total exports.

For balance of payments purposes, all trade in non-monetary are included under trade in goods. Non-monetary gold is defined as all gold not held as reserve assets (monetary gold) by the authorities. Non-monetary gold can be sub-divided into gold held as a store of value and other (industrial) gold.

The Bank of England provides data from a survey of the 6 gold clearers in the London Bullion Market. These data show net exports of nonmonetary gold (and other precious metals) that are held as a store of wealth in allocated accounts. To avoid disclosing the activity of individual gold clearers, the raw data cannot be presented. Therefore a smoothing method is applied. However, briefing is obtained from Bank of England on both the value and volume to cross check with the estimates provided.

Trade in non-monetary gold can be volatile and had an impact on the monthly trade balance and has an equivalent or offsetting impact on change in valuables (see explainer). We also produce a trade less erratic series to try and show the underlying trends in visible trade.

2.1.2 Trade in Services

Trade in services outputs are primarily based on 4 survey inputs, details shown in table 2. These are:

- the International Trade in Services (ITIS) survey

- the International Passenger Survey (IPS)

- Financial Services supplied by the Bank of England

- the Chamber of Shipping members

More detail on each of these survey sources can be found in Table 2 of Information on quality of trade statistics.

2.2 UK trade administrative data sources

Administrative data refers to information collected primarily for administrative reasons (not research). This type of data is collected by government departments and other organisations for registration, transactions and record-keeping, usually when delivering a service. Administrative data are often used for operational purposes and their statistical use is secondary.

We recognise that there are limitations with administrative data and that these can create complications when compiling official statistics. However their use is central to the production of official statistics and the existence of these challenges places a premium on proactive quality assurance to investigate the data, manage identified issues, and clearly communicate any limitations to users.

Our administrative data inputs can be categorised into 8 groups, Trade in goods, Government, Transport, Travel, Financial, Personal, Cultural and Recreational, Other business services and Insurance.

With the exception of trade in goods, it is important to note that the administrative data sources contribute a very small proportion of the overall trade in services outputs; this is because the survey data, as set out in the previous section provides the majority of the trade in services figures.

The details of the administrative data sources used in trade statistics can be found in Table 3 of Information on quality of trade statistics.

3 . Assessing the risk of quality concerns and public interest profile for trade statistics sources

3.1 Risk and profile matrix

To determine the quality assurance level of each admin data source, UK trade team used the risk and profile matrix set out in the UK Statistics Authority Administrative Data Quality Assurance Toolkit (also availble as Table A1 in Information on quality of trade statistics).

3.2 Trade public interest profile

Based on our assessments of the inputs and outputs of UK trade, we have identified the public interest as:

- high public interest in the Trade in Goods input and outputs, this is based on the contribution being more than 90% for a market sensitive monthly release

- low public interest in the Transport Services inputs and outputs, this is based on the individual suppliers contributing a very small proportion to the overall figures that are published in the annual pink book and the quarterly national accounts

- medium public interest in the Financial Services inputs and outputs, one data source from the Baltic exchange is published in the annual pink book, in its own right

- medium public interest in the Government Services inputs and outputs, data are published in the annual pink book and the quarterly national accounts in their own right

- medium public interest in the Other business Services, legal services data is published in its the annual pink book and the quarterly national accounts in own right

- medium public interest in the Personal, Cultural and Recreational Services, data from the gambling commission is published in the annual pink book in its own right

- medium public interest in the Travel Services inputs and outputs, HMRC data on personal imports of cars is published in the annual pink book and quarterly national accounts

- medium public interest in the Insurance Services inputs and outputs, data are extracted from the Lloyds annual report and data are published in the annual pink book and quarterly national accounts

3.3 Level of risk of quality concerns over administrative data for UK trade

Based on our initial assessment of admin data sources, we identified the following:

- trade in goods is the only group to have a higher risk level of quality concern

- travel and personal, cultural and recreational services have a low level of risk of quality concern

- all other data sources have a medium level of risk of quality concern; details of which are set out below

Trade in Goods

Trade in Goods have a higher risk level of quality concern. HMRC administrative data represents over 90% of overall estimates for trade in goods. HMRC collect import and export of physical goods data from UK businesses on a monthly basis and also calculate estimates for non-response and those businesses not registered for VAT. This data is defined as Overseas Trade Statistics (OTS), which records the actual movement of goods and is delivered to ONS on a Combined Nomenclature (CN) classification basis. The HMRC data is mapped in the UK trade statistical system from CN to Standard Industrial Trade Classification (SITC) and adjusted to meet the Balance of Payments (BoP) standards as defined by the International Monetary Fund. We have to make various adjustments to certain transactions which are not reported to HMRC and also to exclude certain transactions that do not involve a change of ownership. Estimates of the freight and insurance costs of transporting the goods to the UK are also deducted.

The remainder is made up from other data sources such as the Bank of England, International Passenger Survey and International Trade in Services survey. Service level agreements (SLA) are in place, regular meetings are maintained with suppliers and data is received on time.

The risk of higher is based on the following:

- data represents a large contribution to the overall trade in goods outputs.

- data quality issues encountered in the past. (See annex B)

Transport Services

Transport Services have a medium level of risk of quality concerns. Individually the inputs represent a small contribution and even at an aggregate level, still carry a small weighting. Although there is a mixture of medium and high concerns over data quality, it is important to note that this is offset by the small contribution that each of the outputs carry towards the overall output of transport services.

- The data are received on time, but then filter through a variety of spreadsheets (including a pre-processing method to extract un-required elements of source data) increasing chances of keying errors.

- All data suppliers have voluntary agreements to provide the data.

- Limited knowledge on how the data is collated and expertise of processors, due to length of time between establishing agreements and review.

- Estimations have replaced actual data for a number of suppliers, as an interim measure until review meetings can be held.

- Data is extracted from government statistics publications, for example, DFT freight statistics.

Data supply is at risk of being discontinued due to changes in structure or supplier systems. This will have implications for the air account and how we formulate our annual data and subsequent quarterly grossing factors. Ongoing process improvement stemming from a Lean Six Sigma review last year will tie into assessing the current method of grossing airline data and shape any requirements for new potential agreements.

Financial Services

Financial Services have a medium risk level of quality concern, brokerage figures from one data source has been given this risk based on the initial following observations:

- the source data contributes a very small weighting to the overall financial services figures

- data are provided on a voluntary basis

- there is limited knowledge on how the data is collated

- often a time lag in the data being available

Government Services

Government Services have a medium risk level of quality concern, there are 3 data sources and all have been assessed at medium risk, this is based on the following initial observations:

- source data contributes a very small weighting to the overall government services figures

- data are supplied on a voluntary basis

- estimates often part of existing government statistical publication with methodological information

Other business Services

Other business Services have a medium risk level of quality concern; there is only one administrative data supply contributing a small proportion of the overall figures, the medium risk was based on the following observations:

- source data contributes a very small weighting towards the legal services category and the overall other business services figures

- regular and timely estimates received from supplier, however data are provided on a voluntary basis

Personal, cultural and recreational services

Personal, cultural and recreational services have a low risk level of quality concern. Data are lifted from the Gambling Commission Industry statistics report. This report provides a half-yearly update on the previously published figures. The report provides good briefing and background into how the data are compiled and this is why the low risk level has been given. Source data contributes a very small weighting towards the overall cultural and recreational services figures.

Travel Services

Travel Services have a low risk level of quality concern; coverage of personal imports of cars from HMRC, the risk level has been based on the following observations:

- source data contributes a very small weighting towards the overall travel services figures

- regular and timely contributions received from supplier, and we have good knowledge of how data is collated

Insurance Services

Insurance Services have a medium risk level; data is derived from the Lloyds annual register.

- export and import data are calculated by ONS and supplement ONS survey data

- regular and timely contributions received from the supplier.

- good communication with the supplier.

3.4 Overall assessment of assurance level required for trade data sources

Combining the level of risk and public interest profile leads to an overall assurance level shown in the table below:

UK trade administrative data sources assurance level

Group	Assurance Level
Trade in goods	A3
Government services	A2
Transport services	A2
Travel services	A2
Financial services	A2
Personal, recreational services	A1
Other business services	A2
Insurance services	A2

4 . Trade quality management actions

4.1 Summary of quality management of administrative data sources

Table 5 of Information on quality of trade statistics gives examples of quality actions completed by the trade team.

Internal quality assurance checks

The data processing steps and data outputs are subject to a series of internal quality assurance checks by the team. This are split into automatic system checks, quality assurance graphical checks on moments or trends in the data and peer review "by-eye" checks. Checklists are provided and sign-off procedures in place.

Roles and responsibilities of the UK trade team and data suppliers

Clear communication between the people involved was an important part of ensuring the quality of the final outputs. Regular contact with the data suppliers helped understand the strengths and weaknesses of the different input data sources and any techniques used.

4.2 Further examples of quality and recommended actions

Trade in Goods have an overall assurance level A3. To mitigate this we have carried out the following:

- established quarterly face-to-face meetings

- Trade in Goods team carried out shadowing of HMRC for two days, HMRC the shadowed the Trade in Goods team, better knowledge of data collection, quality assurance measures and expertise of the compiler

- data goes through ONS quality assurance procedures, consistency checking and comparisons with previously published data

- HMRC collects the raw data and automated validation checks are run, these checks include identifying invalid or new codes and comparisons between the returned value and VAT; there is also sense checking of the data by HMRC compilers

- HMRC provides a comprehensive briefing to accompany their data deliveries and these detail what data they have delivered by commodity, we can then make comparisons to ensure the data received is as expected

- there are a comprehensive number of automated consistency checks built into the ONS trade in goods statistical system, these recognise any large movements in the data, check components add to totals and identify erroneous negative values

- outside of the trade in goods statistical system all outputs are run through excel graphing tools; these create visual comparisons of new and previously published data, highlighting any gaps, large movements and revisions

- a new practice that has been introduced is an earlier exchange of publication documents between HMRC and ONS, this will highlight any differences in HMRC and ONS "stories" at an earlier opportunity; shadowing at ONS and HMRC was also a successful new idea and will be arranged in the future if there are changes to staff or processes and by exchanging newly drafted quality assurance process maps and accompanying documentation we were able to target any additional quality assurance needed; these can be exchanged in the future as they evolve, we have also exchanged a full list of up-to-date telephone and email contact information to ensure any urgent communication can be dealt with promptly, if our usual contacts are unavailable

- conducted a "As-is" and "To-be" process review to ensure quality assurance is strengthened going forwards

- HMRC have published a range of methodology and quality documentation on their [website](#)

For Transport services the overall assurance level is A2. To mitigate this we have carried out the following:

- an internal project to review the transport process, the intention will be to reduce the number of spreadsheets and find a better way of processing the data

- the review will also improve our understanding of the strengths and limitations of the transport data sources. We will also investigate alternative administrative data sources for airlines as recommended in the [Independent Review of Economic Statistics](#)

- data goes through ONS internal consistency checking and quarter on quarter, year on year comparisons; this also includes usage of checking spreadsheets which will identify aggregation errors caused by incorrect task parameters

- a workshop for transport suppliers was run to ascertain the strengths and weaknesses of all data sources and to gain a better understanding of quality assurance measures taken

- commenced a line of communication with current suppliers to refresh relationships and revise contacts lists

- pledging to contact each supplier at least annually, preferably quarterly (outside of standard correspondence to request data)

- have begun to instigate "quick win" improvements, such as NATS changing their delivery format which reduces risk for us (less keying) at no penalty to them

- cross training of skills within Trade in Services now allows for QA to be undertaken by a different person than the processor, reducing risk of errors going unnoticed

Financial services have an overall assurance level A2. To mitigate this we have carried out the following:

- procured a new contact and ongoing discussions taking place about the importance of their data to minimise the need to estimates

- consistency checks are run in systems and quarter on quarter and year on year movements are carried out internally

- annual meetings to be carried out to maintain supplier relationship

Government services have an overall assurance level A2. To mitigate this we have carried out the following:

- agreed that regular estimates and briefing are received from supplier

- consistency checks are run in our systems and quarter on quarter and year on year movements are carried out internally

- face to face meeting held to ascertain the strengths and weaknesses of the data and to establish quality assurance measures

- quarterly meetings to be carried out to maintain supplier relationship

Other business services have an overall assurance level A2. To mitigate this we have carried out the following:

- consistency checks to ensure data is calculated correctly are run in systems and quarter on quarter and year on year movements are carried out internally

- quarterly teleconferences to be carried out to maintain supplier relationship

- attempts were made to arrange face to face meeting arranged to ascertain the strengths and weaknesses of the data and to establish quality assurance measures

Travel services have an overall assurance level A2. To mitigate this we have carried out the following:

- meeting held to ascertain the strengths and weaknesses of the data and to establish quality assurance measures

- consistency checks are run in systems and quarter on quarter and year on year movements are carried out internally

Insurance services have an overall assurance level A2. To mitigate this we have carried out the following:

- carry our regular face-to-face meetings

- provide more information on what is collected and how this reflects the real world as well as information on the method used to compile trade in insurance services estimates in ONS

- action to put in place and SLA

4.3 Closing Remark

As discussed we have concluded that the UK trade outputs are fit for purpose. The exercise of reviewing the administrative data used in trade has strengthened our procedures and also highlighted areas that could be improved further. It is anticipated that this document will be updated to reflect these actions.

5 . Annex A: Quality assurance of administrative data toolkit

More information on UK Statistics Authority Administrative Data Quality Assurance Toolkit can be found on their website.

6 . Annex B: Errors encountered leading to the de-designation of UK trade

B.1

Following a query from BIS on large revisions recently introduced on the exports of goods to the USA, investigations identified an element of double counting from 2009 onwards. Prior to 2009, the USA and Puerto Rico were recorded separately in the UK trade system and then aggregated to form the published USA estimate. From 2009 onwards only one series was supplied for the USA but 2 series continued to be used with zeroes applied to the Puerto Rico series. However both were still aggregated to form the published USA estimate. Following implementation of the new methods to comply with the new Balance of Payments Manual 6 and European System of Accounting standards, the system replaced the zeros with a forecast using earlier estimates for the USA. This issue has now been fixed within the build of the system. There are only 2 countries within trade in goods that are formed using this method (the other country is France, which now also includes St Barthélemy, which has very little reported trade), so the probability of a reoccurrence of the problem is very low.

B.2

The April 2014 UK trade bulletin was subject to a late review caused by a data correction from HMRC. Due to a human error, HMRC omitted a value of roughly £700 million from their trade in oil statistics at the time of delivery to the ONS. As oil is an extremely volatile commodity, large movements in the data are not unusual and therefore the missing data was not identified through our normal quality assurance processes. Although the announcement came too late to include the corrected oil figure, the decision was made to publish the April UK trade bulletin with an accompanying note stating the size of the omission; falling in line with the code of practice regarding open /honest statistics. The UK Stats Authority then asked us to correct the statistics, bulletin and associated datasets to include the additional £700 million retrospectively. Strong relationships already existed between ONS and HMRC staff, with quarterly liaison meetings and regular telephone and email communication. In October 2015 members of the trade in goods and HMRC trade statistics teams' job shadowed each other's quality assurance processes and quality process maps were drawn up by both HMRC and ONS. This was to identify any gaps in the quality assurance processes and to inform on where additional quality assurance should be targeted.

B.3

In June 2014 UK trade bulletin was subject to a correction from the International Passenger Survey (IPS). IPS methodology changes meant that a large proportion of data was omitted. Errors were identified in the expenditure estimates for UK residents visits abroad and overseas residents' visits to the UK within the Overseas Travel and Tourism release (which feed into the imports and exports of services) and in the geographic breakdown of the exports of goods respectively. The errors in the expenditure estimates affected data between January 2014 and September 2014 and the UK trade releases over the same period. A full review was conducted, the results of which are yet to be published (the Dodd review).