

Experimental methodology for producing UK interregional trade estimates

Experimental methodology for estimating the value of UK interregional trade in goods and services at International Territorial Level 1.

Contact:
Chris Goldsworthy
subnational.trade@ons.gov.uk
+44 1633 580158

Release date:
28 July 2023

Next release:
To be announced

Table of contents

1. [Overview](#)
2. [Data sources](#)
3. [Methodological approach](#)
4. [Creating coherent trade survey data](#)
5. [Imputing data gaps through data linkage](#)
6. [Mobilising novel data sources](#)
7. [Limitations and recommendations](#)
8. [Glossary](#)
9. [Related links](#)
10. [Cite this methodology](#)

1 . Overview

This article summarises the methods developed to produce experimental and provisional UK interregional trade in goods and services estimates at the [International Territorial Level 1 \(ITL 1\)](#), which covers the countries of Northern Ireland, Scotland and Wales, and the nine English regions.

Interregional trade is trade in goods and services between regions in the UK, measured on a consistent geographical basis. This differs from subnational trade, which consists of international trade in goods and services with UK regions, and intraregional trade, which is trade in goods and services within a defined UK region.

Where possible, interregional trade is estimated on an economic ownership basis; in most cases change in economic ownership occurs at the same time as change in legal ownership. However, the Trade Survey for Wales (TSW) also includes provision of goods and services to other parts of the same business outside of the region within their survey. For example, goods moved from a branch of a business in Wales to a branch of the same business in England would be captured even though no exchange of economic ownership took place. There are further challenges to consistently capturing change in economic ownership that will be detailed in the [Limitations and recommendations section](#).

The Economic Statistics Centre of Excellence (ESCoE) produced early [estimates of interregional trade between Northern Ireland, Scotland and Wales](#) (PDF, 980KB) for 2015, which led to the development of [a framework for interregional trade data collection and estimation](#). This framework made a series of recommendations to improve the quality of and coherence between existing trade survey data sets developed by Northern Ireland, Scotland, and Wales. In the absence of additional data collection, the framework also recommends leveraging other data sources to supplement the trade surveys from the devolved administrations and to fill known data gaps.

The bottom-up survey-hybrid methodology for developing UK interregional trade estimates summarised in this article is based upon the recommendations given in the ESCoE framework for interregional trade data collection and estimation. This includes how the quality of and coherence between existing trade surveys is improved and how other data sources are mobilised to fill known data gaps.

2 . Data sources

Experimental UK interregional trade in goods and services at the International Territorial Level 1 (ITL 1) and aggregated Standard Industrial Classification (SIC) level will be produced by combining survey microdata provided by each of the devolved administrations with business survey microdata produced by the Office for National Statistics (ONS) and novel administrative and survey datasets produced by other government departments and commercial companies for the years 2019 to 2020. The datasets utilised are:

- the [Northern Ireland Annual Business Inquiry \(NIABI\)](#), which is primarily used in the compilation of the [Northern Ireland Economic Trade Statistics \(NIETS\)](#) and is produced by the Northern Ireland Statistics and Research Agency (NISRA)
- the [Global Connections Survey \(GCS\)](#), which informs the compilation of the [Exports Statistics Scotland \(ESS\)](#) produced by the Scottish Government
- the [Trade Survey for Wales \(TSW\)](#) produced by the Welsh Government

The ONS business survey datasets that are utilised include:

- the [Regional Annual Business Survey \(ABS\)](#)
- the [Annual Purchases Survey \(APS\)](#)
- the [Quarterly Welsh Business Survey \(QWBS\)](#)
- [Inter-Departmental Business Register \(IDBR\)](#)

While other government department and commercial datasets include:

- the Department for Transport [Continuing Survey of Road Goods Transport \(CSRGT\)](#)
- [HM Revenue and Customs \(HMRC\) Overseas Trade Statistics](#) data, linked with the [Inter-Departmental Business Register](#)
- aggregated and anonymised payments data

3 . Methodological approach

The experimental estimates for UK interregional trade in goods and services will be calculated using a bottom-up approach. This approach consists of, where possible, using primary data relating to businesses operating locations within the geographies reported on at the International Territorial Level 1 (ITL 1) level, which covers the nine English regions and the countries of Northern Ireland, Scotland and Wales. This approach provides more direct regional representation than top-down approaches, where national estimates are apportioned to smaller geographies using proxy variables or modelling techniques.

While effort is taken to maintain a consistent bottom-up approach, a top-down approach is adopted in the use of some existing Office for National Statistics (ONS) survey data, such as the [regional Annual Business Survey \(ABS\)](#) data, which is apportioned to produce regional results. Regional economic indicators produced by the ONS tend to adopt a top-down approach as it enables constraining to regional or national totals to ensure consistency and coherency.

4 . Creating coherent trade survey data

As highlighted in the Economic Statistics Centre of Excellence (ESCoE) framework for interregional trade data collection and estimation, there is currently a lack of coherence between each of the devolved administration trade surveys as well as specific limitations such as a lack of coverage of certain parts of the business population or certain trade flows. This methodology aims to make improvements to coherence.

Conceptual changes

There are important definitional differences between the three devolved administration trade surveys drawn upon, particularly regarding business industry, employment, and treatment of [local units](#). In all data sources, the industry of [reporting units](#) is classified according to Standard Industrial Classification (SIC) 2007 and can be obtained from the Inter-Departmental Business Register (IDBR).

Changes to industry classification

Each of the surveys are designed to provide a robust coverage and representation of regional economic activities with a focus on local unit activities within the region. While some similarities exist, the surveys take different approaches to defining reporting units that are sampled, and consequently the assignment of their industrial classification:

- the Northern Ireland Annual Business Inquiry (NIABI) defines the SIC of the business as the SIC of the reporting unit; only reporting units in Northern Ireland (NI) with local units in Northern Ireland are sampled, reporting units are asked to report on the activities of the Northern Irish parts of the business only
- the Global Connections Survey (GCS) and Exports Statistics Scotland (ESS) defines the SIC of the business as the SIC that is used by the largest number of employees across all Scottish local units; a reporting unit is created by summing up employees and turnover for each Scottish local unit assigned to a common reporting unit, and in GCS these businesses are asked to report on their economic activity in Scotland
- the Trade Survey for Wales (TSW) defines the SIC of the business as the SIC of the reporting unit as reported on the IDBR; reporting units are sampled from anywhere in the UK as long as they have local units located in Wales, and businesses are asked to report data relating to the Welsh parts of the business only

The GCS-ESS approach introduces an inconsistency with the SIC classifications on the IDBR, which consider the activity of all local units, rather than only those in Scotland. Approximately 2.5% of SICs in the 2019 GCS-ESS disagree with the IDBR. [ESCoE recommendations](#) support the Scottish approach, suggesting that by defining the reporting unit industry as the industry of the Scottish local unit with the highest employment, a more realistic representation of the regional presence of a business is produced.

While the Northern Irish data are consistent with the IDBR, conceptually it takes a similar approach to Scotland. Northern Irish reporting units on the IDBR encompass Northern Irish local units only, so the SIC of businesses sampled on the NIABI relates to Northern Irish activity only.

As the TSW collates information at a Great Britain (GB) reporting unit level it considers the industry of the reporting unit, which can be anywhere in the UK, rather than the industries of the Welsh local units. This means that the industry may not be reflective of Welsh economic activity.

To improve consistency, the interregional estimates we produce will define the SIC of reporting units based on the primary activity of the local units in the region of interest. This is consistent with the GCS-ESS definition, which is based on the activity of the largest number of employees across all Scottish local units, and with the NIABI definition, which is based on the principal activity across all NI local units.

The SIC definition for businesses in the TSW data will therefore be changed to reflect the activity of the local unit with highest employment. This decision has been taken in consultation with Office for National Statistics (ONS) Methodology and devolved administrations. It is standard practice to reassign businesses to new SICs if they have changed SIC between being sampled and responding.

Analysis of the impact of this change to the 2019 TSW data shows that only 5% of the TSW sample will have their SIC changed to reflect their Welsh activity. This change is conducted at the 5-digit SIC level. Within these 5% that are changed, 29% remain within the same 2-digit SIC industry division, while another 19% remained within the same SIC industry sector.

Changes to employment

Whilst each of the devolved administration surveys use business employment information for reporting units and local units held in IDBR, the definition for business employment also varies between surveys:

- GCS-ESS defines reporting unit employment as the sum of employment of Scottish local units
- NIABI defines employment of a reporting unit as the sum of its local unit employment; this is representative of Northern Irish activity as NI reporting units contain NI local units only
- TSW uses GB reporting unit employment to sample businesses and define [weighting](#) and [imputation](#) cells, meaning that employees located outside Wales contribute to this figure

For the classification and grouping of businesses, we use four employment size bands, which are consistent with those used to breakdown ONS subnational trade statistics. They include:

- micro (0 to 9)
- small (10 to 49)
- medium (50 to 249)
- large (250 and over)

This is primarily done for analysis purposes, as there is currently no plan to publish estimates at an employment size band level.

To increase consistency, we will define business employment as the sum of all local unit employment within regions and prioritise the use of employment data from IDBR, rather than employment figures collected via the devolved administration surveys.

Non-response imputation

The TSW and the GCS are both voluntary surveys and as such have historically low response rates, for 2019, they were 16% and 13% respectively. Both surveys address non-response in their estimation methodology, with the TSW calculating a non-response weight and scaling up responses to account for non-responders, and the GCS supplementing responses with [other data sources including the Annual Business Survey \(ABS\), IDBR, HMRC Overseas Trade Statistics and a number of other sources.](#)

The imputation carried out by the Scottish Government increases the number of businesses reported in the 2019 GCS-ESS from 764 to over 120,000, and the representation of industries covered at the 2-digit SIC industry level from 69 to 78. Further detail regarding this imputation process can be found in the [Government Analysis Function user guide to regional trade](#). In contrast, the NIABI has a reasonably high response rate (approximately 60%) and imputes responses for high employment businesses, or businesses with low employment but high turnover.

TSW's non-response will be further addressed by supplementing responses with other data sources. Sales and purchases from regional ABS data can be used to fill non-response gaps and boost the TSW sample. Where businesses were sampled by the TSW and did not respond, but did respond to the ABS, trade values will be inferred from unweighted ABS turnover and expenditure data.

To prevent double counting, the existing non-response weights for the TSW will be adjusted to account for the additional businesses added. The design weights will also be updated, as described in the following Weighting section.

Treatment of outliers

Outliers, when not correctly treated, can lead to estimates that are highly variable. Estimates can become skewed based on a handful of extreme values, which are unrepresentative of the wider population. We propose a systematic approach to identify and treat outliers in the devolved administration data using an upper quantile technique. To be consistent, we apply the same methodology across all devolved administration surveys to identify and treat outliers. Rest of UK (RUK) exports and imports, which describes trade to the UK nations which are not the focus of the survey in question, go through the outliering process separately, as a business may only be an outlier with regards to exports or imports, but not necessarily both.

Outlier detection takes place at the 2-digit SIC level for the four employment bands. A minimum of 10 businesses are required in a cell for outliering to take place. Groups containing fewer than 10 businesses are merged within the same SIC sector (A to U). If groups are still too small, they will be merged with other groups containing 10 and over businesses in the same industry sector. Any unmatched groups will be manually reviewed.

Businesses with trade values above the 99th percentile are considered outliers. These businesses weights are set to 1 and are removed from the weighting strata but the values are kept. By treating outliers, rather than deleting them, we minimise the loss of information and ensure that abnormal responses are not taken to represent a large number of other businesses in the population.

Weighting

Each of the devolved administration surveys also differ in terms of weighting methodologies used to obtain population level estimates. The NIABI follows the ONS ABS and uses a design weight based on the number of businesses in the population as found in the IDBR. The design weight accounts for the fraction of the population in a particular group that the sample represents for that group.

For example, if one business out of every five is selected in a particular group, each selected business will have a design weight of five, as it "represents" five businesses in the population. This is combined with a calibration weight based on IDBR selected employment. The calibration weight corrects for imbalances in the sample, for example, in a random selection of five businesses out of a population of 10, it is possible that the five businesses selected have, by chance, higher employment values than the non-sampled businesses. If no correction is made, the population total would be over estimated because of the variability in the population. The TSW uses a design weight based on the IDBR turnover of businesses in the population, combined with a non-response weight to account for the low response rate for the survey.

The original weighting methodologies will be maintained to preserve the original survey designs, which were tailored to the regional business population and developed to minimise estimate variation.

Changes to the weighting methodology still need to be made to accommodate for identified outliers. Outliers identified will be treated through changing their design, calibration and non-response weights to ensure that outliers represent themselves only and they are not taken to be representative of the population.

Coverage

While the industries covered by the devolved administration surveys aim to ensure a good representation of regional economic activities, some industries remain uncaptured because of either exclusion from the sampling frames or because of non-response. Main industry exclusions include:

- SIC 86: Human health activities, excluded by NIABI and TSW
- SIC 84: Public administration, excluded by NIABI and TSW
- SIC sector T: including SIC 97: activities of households as employers; and SIC 98: undifferentiated goods-and-services-producing activities of households for own use, excluded by TSW and GCS-ESS
- SIC sector K: including SIC 64: financial service activities; SIC 65: insurance, reinsurance and pension funding; and SIC 66: activities auxiliary to financial services, excluded by NIABI, whereas TSW exclude only part of SIC 64
- SIC 01: crop and animal production, mostly excluded by NIABI
- SIC 799: including activities of tourist guides and other reservation services, excluded by TSW
- SIC 06: extraction of crude petroleum and natural gas, excluded by GSS-ESS; estimates of Scottish exports do not include any exports of oil and gas extracted from the UK continental shelf

It is not unusual for these industries to be excluded from business surveys. For example, the ONS [International Trade in Services Survey](#) also excludes travel, banking and other financial institutions.

Of the three devolved administration trade surveys only the TSW requires businesses to split their reported sales and purchases to the Rest of UK (RUK) to England, Scotland and Northern Ireland. The GCS-ESS and NIABI only collect data relating to RUK as a whole. These RUK figures will be disaggregated to country level estimates by utilising bilateral observations from the TSW data and through the mobilisation of additional data sources as will be seen in the following Mobilising novel data sources section. From these country level trade flows, initial estimates for country level interregional imports and exports for England will be constructed from the residual.

The GCS-ESS does not collect information on Scottish imports; how this limitation is overcome will be defined in the [Imputing data gaps section](#).

5 . Imputing data gaps through data linkage

Regional Annual Business Survey (ABS) data will be used to impute values for the Trade Survey for Wales (TSW), increasing coverage where there are low response rates at the industry and employment band level.

To estimate Rest of UK (RUK) goods and services exports using regional ABS data, international sales of goods and services will be subtracted from total sales of goods and services at the business level:

- international exports of services are captured within the regional ABS data
- to break down total sales of goods to international exports, RUK exports and intraregional sales, the ABS data will be linked to the HMRC Inter-Departmental Business Register (IDBR) linked dataset to identify and remove international goods exports
- from this we obtain a residual of RUK exports and intraregional sales; proportions of RUK and intraregional trade will be calculated for similar businesses in the same industry-employment band within the TSW data
- the median of these industry-employment band proportions will then be applied to the residual trade to obtain estimates for RUK and intraregional trade
- data from the Quarterly Welsh Business Survey (QWBS) will be annualised and used in much the same way as the regional ABS data

To estimate Scotland's imports using the Annual Purchases Survey (APS) data, we will:

- link APS data with the IDBR by reporting unit reference number to obtain local unit and employment information
- purchases data from APS will then be apportioned by Scottish local unit employment to represent regional purchasing activity
- to break down the value of purchases businesses make within the UK to the domestic and RUK level, we will calculate the proportions of domestic and RUK trade for similar businesses in the same industry-employment band within the Global Connections Survey (GCS) -- Export Statistics Scotland (ESS) data and apply this to APS RUK purchases; however, as GCS-ESS only captures exports an assumption of similarity between the proportional split between intraregional purchases and RUK will be made
- correspondence with RUK imports captured within Scottish Supply and Use tables will be made to validate this assumption

6 . Mobilising novel data sources

As recommended by the Economic Statistics Centre of Excellence (ESCoE) [framework for interregional trade data collection and estimation](#), road freight activity data provided by the Department of Transport through the Continuing Survey of Road Goods Transport GB and NI (CSRGT) will be used in the estimation of interregional trade in goods. Information on the volume of goods lifted, commodity type, point of origin and point of destination from the CSRGT will be used to disaggregate Rest of UK (RUK) trade in goods to the country level, these data will also be used to disaggregate trade in goods for England down to its constituent nine ITL 1 areas.

CSRGT data are converted from Standard goods classification for transport statistics (NST 2007) to Standard Industrial Classification 2007 (SIC 07) to ensure consistency across datasets using Eurostat correspondence tables. Following this, freight volume data are converted to values using national trade data. Proportions for interregional exports are then calculated by dividing the value of trade by SIC from each International Territorial Level 3 (ITL3) area by the total value of trade by SIC from all areas within the corresponding country. These estimated proportions are then used to disaggregate RUK exports. This will produce estimates for bilateral interregional exports of goods between Northern Ireland, Scotland and Wales.

For the accurate estimation of interregional trade in goods, road freight activity is separated between transport that is used to move domestically produced goods for consumption in other regions, and regional transit that involves movement of goods that do not originate in a region, nor are they intended for use in that region.

Regional transit is not identified separately in the CSRGT, but in order to accurately measure interregional trade, transshipment hubs will be identified using CSRGT data and the size of regional transit estimated. Regional transit will then go through an estimation process to identify the original supply location and final destination. This is estimated by tracing onward journeys of single vehicles within a specified reporting timeframe at the SIC level. This will allow us to reconcile interregional trade in goods estimates to account for regional transit.

An important limitation to these data is that they capture the physical movement of goods and not change in economic ownership, and so correlations between proportions provided by the CSRGT and interregional trade should be treated with caution.

Payments data will be used to break down RUK trade in goods and services data reported within the devolved administration trade survey data sets to the country level. Payments data will be used to create industry and geography proportions, which will be applied to the devolved administration survey RUK trade in goods and services estimates to disaggregate estimates into bilateral trade with other UK regions at the ITL 1 level.

These data have the benefit of capturing economic exchange at the SIC industry level and so are consistent with other data used in the estimation of UK interregional trade in goods and services. However, a significant limitation is the presence of the 'Head-Office Effect' in the data whereby a large proportion of transactions are recorded between head offices of businesses. In addition, the payments data capture economic transactions between businesses, not all of which would be classed as trade. For example, a financial institution may act as an intermediary between two businesses, yet this is unclear from the data.

7 . Limitations and recommendations

Because of the bottom-up approach developed to produce UK interregional trade in goods and services estimates, the estimates produced will not be consistent with national trade estimates or regional economic indicators that are apportioned from national estimates. As such these estimates will not be constrained to national estimates.

This offers the benefit of potentially producing interregional trade estimates that have a more granular representation of regional business characteristics than would be possible by taking a top-down approach. However, this means that the quality of the estimates cannot be benchmarked against other outputs. This limitation is mitigated by triangulating observations across multiple datasets where possible.

The output will be limited to estimating UK interregional trade in goods and services only for those industries that are reported on within the devolved administration trade surveys and so full industry coverage will not be achieved. Future developments will aim to improve industry coverage.

While effort has been taken to ensure conceptual consistency where possible, there remain inconsistencies in the basis on which trade is captured. As mentioned, the TSW captures the movement of goods and transfer of services between sites within the same business structure where no economic exchange has occurred. Equally the CSRG data captures the physical movement of goods, not change in economic ownership.

We also exclude non-resident flows from our data. The ESCoE framework recommends the use of tourism survey data to capture these flows, unfortunately because of the lack of availability of granular tourism data for the years 2019 to 2020, this is not possible for this period.

8 . Glossary

ITL

The International Territorial Levels (ITL) is a hierarchical classification of administrative areas used for statistical purposes. ITL1 are major socio-economic regions, while ITL2 and ITL3 are progressively smaller regions. In the context of the UK, the ITL1 areas are Wales, Scotland, Northern Ireland and the nine regions of England.

Local unit

The local unit is a statistical unit used to define an individual site (for example, a factory or shop) in an enterprise.

Reporting unit

The reporting unit is the observation unit of a business or enterprise held on the Inter-Departmental Business Register (IDBR) and holds the mailing address to which survey questionnaires are sent. This is usually based upon the businesses' main registered address in the UK.

Weighting

Weighting describes the method by which we produce estimates about the characteristics of the population as a whole from the responses given by those people and businesses selected in the sample survey. Those responses from the sample are weighted to ensure they represent the entire population without bias and produce good quality outputs.

Imputation

Imputation is the process of replacing missing data with substitute data by using statistical methods.

A full [Glossary of economic terms](#) is available.

9 . Related links

[International trade in UK nations, regions and cities: 2021](#)

Article | Released 28 June 2023

Experimental estimated value of exports and imports of goods and services for 2021 for International Territory Levels 1, 2 and 3, and city regions. Includes EU and non-EU split along with data for top 20 partner countries.

[International trade in UK nations, regions and cities QMI](#)

Methodology | Last revised 28 June 2023

Quality and Methodology Information (QMI) report for our Subnational trade estimates release, detailing the strengths and limitations of the data, methods used, and data uses and users.

[A user guide to regional trade](#)

Web page | Released 12 July 2023

User guide outlining the various measures of regional, international trade in goods produced in the UK by the Office for National Statistics (ONS), His Majesty's Revenue and Customs (HMRC), Northern Ireland Statistics and Research Agency (NISRA), the Scottish Government and the Welsh Government.

[Improving the Quality of Regional Economic Indicators in the UK: A Framework for Interregional Trade Data Collection and Estimation \(ESCoE TR-13\)](#)

Publication | Released 9 December 2021

Research findings outlining recommendations for producing UK interregional trade estimates.

[UK Interregional Trade Estimation: Estimates of Trade between Northern Ireland, Scotland, Wales and England](#)

Publication | Released 30 June 2020

Research outlining a methodological approach to and early estimates of UK country level Interregional Trade.

[Broad Economy Sales and Exports Statistics](#)

Web page | Published 14 December 2022

Northern Ireland sales and purchases of goods and services by local businesses with markets inside and outside of Northern Ireland in the non-financial industry sectors.

[About Export Statistics Scotland](#)

Web page | Last Revised 2 December 2022

Information on export statistics produces by the Scottish Government.

[Trade Survey for Wales](#)

Web page | Last Revised 31 August 2021

Information about the Trade Survey for Wales produced by the Welsh Government.

10 . Cite this methodology

Office for National Statistics (ONS), released 28 July 2023, ONS website, methodology, [Experimental methodology for producing UK interregional trade estimates](#)