

Summary Quality Report for E-commerce

1 Introduction

This report is part of a rolling programme of quality reports being introduced by the Office for National Statistics (ONS). The full programme of work being carried out on [Statistical Quality](#)¹ is available on the ONS website. Summary Quality Reports are overview notes which pull together key qualitative information on the various dimensions of quality as well as providing a summary of methods used to compile the output.

This report relates to the e-commerce survey. The e-commerce survey is an annual survey designed to measure the extent of the use of Information Communication Technologies (ICTs) (ie computers, Internet connection, websites) by businesses, and the extent of electronic trading by businesses (ie trading on-line). This report aims to provide users with guidance to assess the quality and usability of the e-commerce estimates.

The Survey had experimental status from the 2002 survey results to the 2005 survey results. Following a review, the Survey was deemed to meet the criteria for National Statistics status, and the estimates for 2006 and subsequent years were published as National Statistics.

2 Summary of Quality

2.1 Relevance

The degree to which the statistical product meets user needs for both coverage and content.

The Survey collects data on ICT usage and e-commerce from businesses in the UK. There are separate questionnaires for the financial and non-financial sectors, with the non-financial sector excluded from estimates relating to the proportions of businesses buying and selling over ICTs, and the values of their sales. It is the only survey that measures e-commerce and ICT activity within UK businesses.

Coverage and Content	E-commerce
What it Measures	E-commerce and ICT activity of businesses in the UK such as ICT systems and their usage, use of the Internet and other electronic networks and also electronic business processes.
Frequency	Annual
Sample Size	8,000 since the 2006 survey.
Sample Frame	Inter-Departmental Business Register (IDBR)
Sample Design	Stratified random sample where the strata are defined by Standard Industrial Classification (SIC) currently SIC 2003 ² and number of employees in business (10-49, 50-249, 250-999, 1000+).
Weighting	Each sampled business represents a number of similar businesses, and is weighted to reflect the number of employees it has and the number of businesses within its stratum as recorded on the IDBR.
Estimation	Number raised estimation is used for tick box questions (ie yes or no), whilst a ratio estimator is used for the value data questions

Outliers	Post stratification is applied. A small number of respondents who are atypical for their stratum and who have a disproportionate effect on the estimates are manually treated as outliers.
----------	--

Data from the e-commerce outputs are used by a wide range of users, including:

- [Eurostat](#)³ (Key User) - The UK provides statistics to [Eurostat](#)³ measuring e-commerce activity in accordance with the [Regulation \(EC\) No 808/2004 of the European Parliament and of the Council](#)⁴, e-commerce results are used to provide information that is consistent with other EU member states, to enable eEurope benchmarking to be achieved
- ONS Economic Interpretation Division - e-commerce results enable productivity analyses to be performed and enable a better understanding of the adoption of specific technologies and e-business processes and their contribution to productivity.
- Department of Business, Innovation and Skills (BIS) - e-commerce results are used for benchmarking carried out by ICT policy makers, for international comparisons) and also for Briefing/Parliamentary Questions (PQs)
- Scottish Government (SG) and Welsh Assembly Government (WAG) - e-commerce results are used to gauge how businesses and the economy are changing in Scotland/Wales with the availability of new technologies, and to also identify areas and develop policies that aid the adoption of e-commerce practices
- Department of Enterprise, Trade and Investment Northern Ireland (DETINI) - e-commerce results are used to measure and monitor the take-up of broadband services in Northern Ireland as a measure of the success of DETINI's policies and actions

A [Quinquennial Review](#)⁵ of the Survey was published in June 2006, which provides more information on the Survey itself, including user needs.

Gaps in User Needs

SG, WAG and DETINI would all prefer to have a regional breakdown of e-commerce results. Regional estimates have been researched and considered a number of times and there are a number of issues (such as the relatively small sample size, selection at enterprise and not local unit level plus definitions of geographical regions), that need to be investigated before regional estimates can be considered. However, it is unlikely that the e-commerce survey will produce regional statistics in the near future.

Conceptual issues

Until the 2002 survey, financial sector businesses received the same e-commerce questionnaire as all other businesses. It became apparent that there are fundamental differences in how financial sector businesses report their sales and purchases. This meant that the same questions about monetary values of purchases and sales over the Internet and other ICTs were not appropriate for financial and non-financial businesses. Therefore, financial sector businesses are excluded from the estimates of the proportions of businesses buying and selling electronically and the values of their sales and purchases. The financial sector is included in the estimates for ICT usage.

E-commerce is a vast and complex subject. There are a number of terms and definitions that need to be defined for respondents, so that they are aware of what exactly is being asked and what they need to include/exclude in their e-commerce estimates. The definition of e-commerce used is that agreed by the Organisation for Economic Cooperation and Development (OECD) It is the method by which the order is placed, which determines whether a transaction is e-commerce - not the payment or delivery channel.

2.2 Accuracy

The closeness between an estimated result and the (unknown) true value.

Estimates from the e-commerce survey, as with other surveys, are subject to sources of error. Total error consists of two elements, the sampling error and the non-sampling error.

Sampling Error

This occurs because the Survey is based on a sample of businesses, selected from the IDBR, rather than being a census. Due to timing, cost and resource issues, a census would not be possible so a suitably designed sample of businesses is used. The e-commerce survey

provides standard errors and coefficients of variation to [Eurostat](#)³ and standard errors were also included for the first time in the 2005 release.

The sample size since the 2006 survey has been 8,000.

Non Sampling Error

Non-sampling errors are not easy to quantify and include errors of coverage, measurement, processing and non-response.

The Survey uses the IDBR to select the sample of businesses that receive the questionnaire. The IDBR is regularly updated, which minimises the selection of mis-classified businesses and dead reporting units. This also aims to ensure adequate coverage of newly established businesses. The sample design takes into account previous years e-commerce returns, to ensure key respondents are included in the sample each year.

Response rates provide an indication of the likely impact of non-response error on estimates. For the e-commerce survey, the response target is 80 per cent and the target for these responses that have clean data is 95 per cent, giving an overall target of 76 per cent of the sample that should have clean data and be included in final results.

Expert opinions and advice regarding the e-commerce questionnaire are regularly sought. This includes reviewing areas such as the clarity and usefulness of the questions themselves, the overall layout and style of the questionnaire and the definitions used in the glossary of terms. A clear and concise questionnaire will help to minimise both measurement and responder bias/error.

The e-commerce publication includes the current and previous years' e-commerce results. There is a revisions policy in place so that the previous two years data may be revised, and will be published with the letter 'r' next to the estimate.

Investigating and reviewing editing rates and the total contribution to key estimates from imputed values are other ways in which the e-commerce survey is trying to maximise the accuracy of the results.

2.3 Timeliness and Punctuality

Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the time lag between the actual and planned dates of publication.

[Eurostat](#)³ regulations include a deadline for estimates from member states to be provided, and this is by October each year. This deadline is met by the UK. Detailed results are then published each year on the ONS website approximately ten or eleven months after the reference year.

The annual e-commerce releases have consistently met target publication deadlines. The [UK National Statistics Publication Hub](#)⁶ provides twelve months advance notice of releases. In the unlikely event of a change to the pre-announced release schedule, public attention will be drawn to the change. The reasons for the change will be explained fully at the same time, as set out in the [Code of Practice for Official Statistics](#)⁷.

2.4 Accessibility and Clarity

Accessibility is the ease with which users are able to access the data, also reflecting the format(s) in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the metadata, illustrations and accompanying advice.

E-commerce data are published in [E-commerce Statistical Bulletins](#)⁸ which summarises the key results. A set of detailed tables is also published on the ONS' website. E-commerce data are weighted in two ways; by number of businesses and number of employees. The data in the tables are broken-down by number of employees (in size bands) and industry.

The [E-commerce Statistical Bulletins](#)⁸ can be used in conjunction with the [Internet Access – Households and Individuals Statistical Bulletin](#)⁹ to give an indication of various aspects of the UK's Internet usage.

2.5 Comparability

The degree to which data can be compared over time and domain.

A standard [Eurostat³](#) model questionnaire has been developed for the consistent measurement of e-commerce and ICT activity across the EU. The UK questionnaire is similar to the [Eurostat³](#) standard except where changes have been made to account for differences in interpretation and to improve understanding of the questionnaire for respondents. Consultation with data collection experts as well as internal and external users has resulted in a set of agreed definitions for e-commerce activity, which are provided as a glossary with the questionnaire.

The Survey covers selected parts of SIC, currently [SIC 2003²](#), and has collected information on a consistent basis from these SICs since 2000. There have been minor adaptations where [Eurostat³](#) have required more detail in certain SICs. The key variables used to optimise the e-commerce sample have remained the same since 2003, ensuring the sample design does not radically change year on year.

From the 2005 survey, sent out in 2006, businesses with between one and nine employees were not surveyed due to a lack of funds. This meant a discontinuity between previous UK results and the 2005 results (published in October 2006). However, this was only the case for the whole UK results, as the e-commerce publication includes detailed employment breakdowns.

In the 2008 survey there have been changes in the measurement of e-commerce. Until the 2007 survey, the definition of e-commerce sales and purchases included transactions over manually typed email. From the 2008 survey onwards, this changed and email transactions are now excluded. This means that e-commerce sales over non-Internet ICTs in 2008 are also not directly comparable with the 2007 estimates. Similarly, purchases have been combined into purchases over all ICTs, so these estimates are not directly comparable with previous years.

The Internet/non-Internet split of transactions that had been included since the survey began will be discontinued after the 2008 survey. This will be replaced with a breakdown of website and non-website transactions. The 2008 survey results contain estimates on both bases to allow comparison between the end of the old series and the start of the new one.

2.6 Coherence

The degree to which data that are derived from different sources or methods, but which refer to the same phenomenon, are similar.

There are no other publications that measure e-commerce activity within the UK. Therefore, it is not possible to compare the annual e-commerce results with another publication.

3 Summary of Methods Used to Compile the Output

Sample Coverage

The Survey samples businesses in the UK from all SICs, except for:

- Agriculture, Hunting and Forestry (Section A)
- Fishing (Section B)
- Mining and Quarrying (Section C)
- Public Administration and Defence; Compulsory Social Security (Section L)
- Education (Section M).

The sample is selected from the IDBR, and is stratified by SIC and number of employees (10-49, 50-249, 250-999 and 1,000+). Respondents with more than 1,000 employees are sent the questionnaire every year; whilst respondents with less than 1,000 employees are randomly sampled and will therefore not necessarily receive the questionnaire every year. The sample allocation uses an algorithm and optimises the sample on key variables. The sample is designed to minimise the coefficient of variation of these variables, and to provide a representative sample of the population of interest in each stratum.

Processes and Methodology

The e-commerce sample is selected in the November of the reference year. From the 2006 survey, questionnaires were despatched in the following January. This is a month earlier than

in previous years. Reminders are sent in March to those who have not responded. Returned data are processed by ONS between despatch and July/August. This includes ensuring the forms pass the validation checks and contacting respondents if required.

The results are checked from June each year, with the first final results run in July. As it is not possible to select every business in the population, to obtain estimates of the population it is necessary to weight the data. A number raised estimation method is used for tick box questions, whilst a ratio estimator is used for value data questions. The dataset is finalised in September, with results sent to [Eurostat](#)³ and published on the ONS website in October or November.

Statistical disclosure control methodology is applied to e-commerce data. This ensures that information attributable to an individual or individual organisation is not identifiable in any published outputs. The [Code of Practice for Official Statistics](#)⁷, and specifically the Principle on Confidentiality set out practices for how we protect data from being disclosed. The Principle includes the statement that ONS outputs should “ensure that official statistics do not reveal the identity of an individual or organisation, or any private information relating to them, taking into account other relevant sources of information”. More information can be found in [National Statistician’s Guidance: Confidentiality of Official Statistics](#)¹⁰ and also on the [Statistical Disclosure Control Methodology](#)¹¹ page of the ONS website.

4 References

	Title of Reference	Website location
1	Statistical Quality	http://www.ons.gov.uk/about-statistics/methodology-and-quality/quality/index.html
2	SIC 2003	http://www.statistics.gov.uk/statbase/Product.asp?vlnk=14012
3	Eurostat	http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home
4	Regulation (EC) No 808/2004 of the European Parliament and of the Council	http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:143:0049:0049:EN:PDF
5	Quinquennial Review	http://www.statistics.gov.uk/STATBASE/Product.asp?vlnk=14524
6	UK National Statistics Publication Hub	http://www.statistics.gov.uk/hub/release-calendar/index.html
7	Code of Practice for Official Statistics	http://www.statisticsauthority.gov.uk/assessment/code-of-practice/index.html
8	E-commerce Statistical Bulletins	http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=6645&Pos=2&ColRank=1&Rank=358
9	Internet Access – Households and Individuals Statistical Bulletins	http://www.statistics.gov.uk/statbase/Product.asp?vlnk=5672
10	National Statistician’s Guidance: Confidentiality of Official Statistics	http://www.statisticsauthority.gov.uk/national-statistician/guidance/confidentiality-of-official-statistics.pdf
11	Statistical Disclosure Control methodology	http://www.statistics.gov.uk/about/data/methodology/general_methodology/sdc.asp

Last Updated – November 2009