



Department of
**Enterprise, Trade
and Investment**
www.detini.gov.uk

Economy

Statistics Report

Guidelines for Measuring Statistical Quality -

NI Census of Employment

March 2006

Stimulating
innovation
enterprise and
competitiveness

A NATIONAL STATISTICS PUBLICATION

National Statistics are produced to high professional standards set out in the National Statistics Code of Practice. They undergo regular quality assurance reviews to ensure that they meet customer needs. They are produced free from any political interference.

For general enquiries about National Statistics, contact
The National Statistics Public Enquiry Service on **0845 601 3034**
minicom: 01633 812399
E-mail: info@statistics.gov.uk
Fax: 01633 652747
Letters: Room 1015, Government Buildings, Cardiff Road, Newport,
Gwent, South Wales NP10 8XG

You can also find National Statistics on the Internet – go to
www.statistics.gov.uk

Department of Enterprise,
Trade and Investment

Guidelines for Measuring Statistical Quality - NI Census of Employment

March 2006

Contents

1

Introduction.

Pages 1-3

2

Quality Measurement
Guidelines.

Page 4

3

Data Collection.

Pages 5-22

4

Data Processing.

Page 23-30

5

Data Analysis.

Pages 31-38

6

Dissemination.

Pages 39-43

Introduction

1

These guidelines provide a checklist of quality measures and indicators for consideration when measuring and reporting on the quality of statistical outputs. The guidelines are not a National Statistics protocol but represent best practice for measuring quality at various stages of the data collection – output cycle. This document contains information on the NI Census of Employment.

The quality measures and indicators themselves have been designed with the needs of the users of statistical outputs in mind. Many of the measures and indicators will already be addressed by those involved with producing statistical outputs - for example, reporting on standard errors. Others are newly developed. The purpose of this document is to provide a single checklist of measures and indicators to promote a more standardised approach to measuring and reporting on quality.

The focus of this version of the Guidelines is on quality measurement for survey outputs. Future versions are likely to be extended to include quality measurement requirements for administrative data and outputs derived from other sources, for example National Accounts.

Background

The Government Statistics Service (GSS) has long recognised the need to provide users with information about the quality of statistics and about the analytical techniques used to derive the figures. This commitment is demonstrated in the National Statistics Code of Practice, which states that:

“...processes and methods used to produce National Statistics will be sufficiently detailed to allow users to assess fitness for particular purposes..”

The sentiment is echoed in the National Statistics Quality Strategy, which states that:

“The quality measures for National Statistics will be systematically reported, alongside results and will enable [the user] to judge the suitability of their application for their intended uses.”

These guidelines replace the GSS Statistical Quality checklist and the Office for National Statistics (ONS) Quality Measurement and Reporting Framework, which have been integrated to provide a single set of guidelines for quality measurement.

Aim and purpose of the guidelines

The overall aim of the guidelines is to outline best practice for measuring and reporting on quality of GSS outputs. In particular, the guidelines have been developed so that information is made available to users of statistical outputs to help them understand:

- the context in which the data were assembled and analysed;
- methods adopted and limitations they impose;
- the reliability of the figures; and
- the way they relate to other available data on the same subject.

The measures and indicators have also been designed to assist producers of official statistics in monitoring data quality for purposes of continuous improvement.

Quality Measurement

Quality can have many different meanings, for example; luxury, good value for money and convenience. Its meaning will often depend upon the context in which it is used. Quality, when referred to in terms of statistical outputs can generally be thought of as the degree to which the data meet user needs, or simply put, the degree to which the data are 'fit for purpose'.

Quality has often been associated with accuracy and timeliness. But even if statistics are accurate and timely, they cannot be deemed to be good quality if they are not based on concepts, which are meaningful and relevant to the users. In addition, different users will have different needs.

Quality measurement and reporting for statistical outputs is therefore concerned with providing the user with sufficient information to judge for themselves whether or not the data are of sufficient quality for their intended use(s).

It is recommended that the quality of a statistical output should be determined by its performance against a range of attributes that together can be used to assess whether an output meets users' quality criteria.

The ONS have adopted the data quality attributes defined for the European Statistical System (ESS), which are shown in Table 1.

Quality measure or quality indicator?

Quality of data can rarely be explicitly 'measured'. For example, in the case of accuracy, it is almost impossible to measure non-response bias, as the characteristics of those who do not respond can be difficult to ascertain. Instead, certain information can be provided to help 'indicate' quality.

The quality measures and indicators in the remainder of this document were initially developed around the six data quality dimensions. A good summary of quality will contain a blend of quality measures from each of the six ESS quality dimensions outlined in Table 1.

Dimensions of Quality
Table 1

Definition	Key components
1. RELEVANCE	
The degree to which the statistical product meets user needs both in coverage, content and detail.	Any assessment of relevance needs to consider: <ul style="list-style-type: none"> • who are the users of the statistics? • what are their known needs? • how well does the output meet these needs?
2. ACCURACY	
The closeness between an estimated result and the (unknown) true value.	The main sources of error are sampling and non-sampling error, where non-sampling error includes: <ul style="list-style-type: none"> • coverage error; • non-response error; • measurement error; • processing error; and • model assumption error.
3. TIMELINESS AND PUNCTUALITY	
Punctuality refers to the time lag between the actual delivery date of the data and the target date when the data should have been delivered. Timeliness is the degree to which data produced are up to date, published frequently and on time.	An assessment of timeliness and punctuality should consider the following: <ul style="list-style-type: none"> • production time; • frequency of release; and • punctuality of release.
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 additional advice provided.	Specific areas where accessibility and clarity may be addressed include: <ul style="list-style-type: none"> • needs of analysts; • assistance to locate information; • clarity; and • dissemination.
5. COMPARABILITY	
The degree to which data can be compared over time and domain.	Comparability should be addressed in terms of comparability over: <ul style="list-style-type: none"> • time (for data derived for the same purpose); • spatial domains (sub-national/national/international); and • domain or sub-population (industrial sector, household type).
6. COHERENCE	
The degree to which data that are derived from different sources or methods, but which refer to the same phenomena are similar.	Coherence should be addressed in terms of coherence between: <ul style="list-style-type: none"> • provisional and final statistics; • sources and data produced at different frequencies; • other statistics in the same socio-economic domain; • sources and outputs.

Quality Measurement Guidelines

2

Quality Measurement Guidelines

The section has been arranged according to the order in which opportunities for measuring and reporting on quality are likely to occur in the data collection – output cycle. There are four sub-sections: data collection, data processing, data analysis and dissemination. Within each section, quality measures and indicators are grouped into tables according to which specific elements of the data collection – output stage they relate.

The tables provide a list of quantitative and qualitative quality measures and indicators for consideration, together with the following supplementary information:

- description of measure/indicator and notes on use;
- type of output where measure is applicable;
- likely frequency of production for the measure (for example, each survey cycle);
- an example of the type of information you may wish to record when addressing each measure. For qualitative measures, examples have been taken where possible from recently published documents and articles. For quantitative measures, suggested formulae are shown instead;
- the corresponding ESS quality dimension; and
- the letters, 'M' or 'I' to denote whether the item is a quality 'measure' or quality 'indicator'.

Data Collection

3

B1. Decision to undertake a collection or analysis

Ref.	Notes	Outputs applicable to	Frequency of production	Example	Relevance
B1.1	<p>Describe and classify key users of output. (I)</p> <p>The Census of Employment is widely used by economists and economic commentators both within and outside Government in assessing the progress of the Northern Ireland economy. The data are also supplied to a range of academics, market research companies, colleges of further and higher education, libraries and students for a wide range of research purposes.</p>	All outputs.	For any data output.		
B1.2	<p>Describe needs of key users and uses of output. (I)</p> <p>The Census of Employment is used primarily as a source of trend information on employee jobs in Northern Ireland, as a basis for economic research and as a source of information on employee jobs. The Census allows changes in the structure of employment in Northern Ireland to be tracked over time by counting the number and type of jobs as distinct from the number of persons with a job. Results can be produced at detailed industrial level and for small geographical areas, such as wards. As every employer is contacted it is very accurate information.</p>	All outputs.	For any data output.		

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B1.3	<p>Describe uses supported and where possible how the data relate to needs of users.</p> <p>(I)</p>			
	<p>The Census provides detailed employee jobs estimates for Northern Ireland as a whole, sub-NI level data (i.e. District Council Area, Parliamentary Constituency and ward level) and detailed industrial activity breakdowns (5-digit SIC and above).</p>	<p>All outputs.</p>	<p>At design stage and for output.</p>	

B2. Collection design

Ref.	Notes	Outputs applicable to	Frequency of production	Example	Relevance
B2.1	<p>Describe key statistical concepts e.g. statistical measure; variable; population; units; domains and time reference. (I)</p> <p>Census of Employment is a full count of the number of employee jobs (self employed are not included) in all industries except for agriculture. The units to be surveyed were drawn from the IDBR (Inter Departmental Business Register), which is a register held by the Office for National Statistics (ONS), combining information from the VAT based business register and the Inland Revenue's computerised PAYE system. The Census of Employment is carried out biennially with the reference date being the first Monday in September. The survey counts the number of jobs rather than the number of persons with jobs.</p>	All outputs.	For any data output.		
B2.2	<p>Describe any gaps between measured statistical concept and user concept of interest. (M)</p> <p>Not applicable</p>	Any output where something is measured.	For any data output.		Relevance
B2.3	<p>Describe classifications. (I)</p> <p>The Standard Industrial Classification of Economic Activities 2003 (SIC) is used for classification in producing the Census of Employment.</p>	All outputs.	For any data output.		

Ref.	Notes	Outputs applicable to	Frequency of production	Example	
B2.4	Provide a statement of the national/ international agreed definitions and standards used. (I)				Comparability
	The Census of Employment uses the SIC 2003 to All outputs. For any data output. code business activities. The use of SIC03 is obligatory in the UK where it is required to transmit to the European Commission statistics on economic activity. All branch procedures are governed by GSS Codes of Practice.				
B2.5	Provide a statement of international regulations that apply and any laws which have been repealed or abolished. (I)				Comp
	As B2.4 above.	All outputs.	For any data output.		
B2.6	Describe any deviations from national/ international agreed definitions and standards, with reasons for deviations. (I)				Relevance
	None	All outputs.	For any data output.		
B2.7	Describe key sources, including known purpose of the data collection and the known merits and shortcomings of the data. (I)				Relevance
	The Quarterly Employment Survey is conducted on Outputs employee jobs and allows comparison to the based on Northern Ireland Census of Employment. other Information collected by the QES on firms with one sources, local unit is shared with the Census of e.g. Employment. This data is shared in a move to National decrease the burden on businesses. Accounts. When a form hasn't been received from a respondent, and they cannot be contacted, data is estimated. Employment is estimated on the basis of previous returns and/or information from the Inland Revenue's PAYE system. This only happens in a small percentage of cases.				

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B2.8	Time lag: reference date/period to release of provisional output. (M)			
	The Census of Employment does not issue provisional results.	All outputs with provisional versions.		For any data output.
B2.9	Time lag: reference date/period to release of final output. (M)			
	The Census of Employment is produced and published on a biennial basis. Results are published in December in the year following the reference date of the Census of Employment.	All outputs.		For any data output.
B2.10	Time lag: end of data collection and publication of first results. (M)			
	Data collection ends in July while the first results are published in December.	All outputs apart from those based on other sources.		For any data output.
B2.11	For new and ad-hoc surveys, time lag: commitment to undertake the survey and release date. (M)			
	Not applicable.	Surveys.		For any data output.

Timeliness

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B2.12	<p>Frequency of publication. (I)</p> <p>The Census of Employment is produced and published on a biennial basis, with results in the December of the year after the reference date, i.e. the results of the September 2005 census will be published in December 2006.</p>	All outputs.	At dissemination strategy planning stage and where there are subsequent re-designs of the dissemination strategy or of the survey.	
B2.13	<p>Timetable for release of data. (I)</p> <p>As above for main publication. Other publications (5-digit, District Council, Parliamentary Constituency, ward etc follow).</p>	All outputs.	At dissemination strategy planning stage and where there are subsequent re-designs of the dissemination strategy or of the survey.	

Timeliness

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B2.14	<p>Describe processing systems and quality control. (I)</p> <p>The Census of Employment system has a number of checks built-in which ensure the quality of the data. The system is outlined fully in the user manual.</p> <p>Returned forms are checked thoroughly by administrative staff to ensure that all data is correct and that any significant changes from the previous quarter have been accounted for. As forms are input, any output, which looks unreasonable, is checked and edited as necessary. This may involve contacting the respondent to verify the information. A further final edit check is then carried out. In addition, electronic validation checks are carried out after data entry. This highlights any large changes in proportion or numbers of employee jobs.</p> <p>Returns to the Census of Employment are compared with QES returns, to ensure respondent's responses are consistent. Any discrepancies are investigated as above.</p>	All outputs.	For any data output.	

Accuracy

B3. Sample design and implementation

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B3.1	<p>What type of sample design was used? (I)</p> <p>The Census of Employment is a statutory survey and covers all employees in Northern Ireland. Every employer (except in Agriculture) in NI will be sent a Census form. The units to be surveyed are drawn from the IDBR (Inter Departmental Business Register), which is a register held by the Office for National Statistics (ONS). All units considered to be live at September in the census year are selected.</p>	All outputs.	For any data output.	
B3.2	<p>Describe sampling frame, including source, date of creation and date of last update. (I)</p> <p>Companies are selected using the Inter-Departmental Business Register (IDBR), which is a list of all UK businesses that is maintained by ONS. It is used for selecting samples for surveys of businesses, to produce analyses of business activity and to provide lists of businesses. It is based on inputs from three administrative sources: traders registered for Value Added Tax (VAT) purposes with HM Customs and Excise (HMCE); employers operating a Pay As You Earn (PAYE) scheme, registered with the Inland Revenue (IR); and incorporated businesses registered at Companies House (CH). IDBR is a live database; therefore it is updated daily to take account of the births, deaths, mergers, etc. of companies based on information from the above sources.</p>	All outputs.	For any data output.	

Accuracy

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B3.3	<p>Has the frame been updated to take account of births, deaths and other relevant changes in the study population? For example, changes in SIC over time. (I)</p> <p>As before, IDBR is a live database; therefore it is updated daily to take account of the births, deaths, mergers, etc. of companies based on information from various sources, primarily HM Customs and Excise (VAT) and Inland Revenue (PAYE). As the amount of firms on IDBR change daily, two top-up selections are taken from IDBR in October and February. This is to ensure that all firms that are open on the census date will receive a form.</p>	All outputs.		
B3.4	<p>What method was used to select the sample? (I)</p> <p>All live LU's on IDBR are selected.</p>	All outputs.	For any data output.	
B3.5	<p>If the sample design involves stratification, how are the strata defined? (I)</p> <p>Not applicable.</p>	All outputs.	For any data output.	

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B3.6	What were the target and achieved sample sizes? (I)			
	Not applicable.	All outputs.		For any data output.
B3.7	Define and compare the target population and the study population, for the population as a whole and for significant sub-populations. (I)			
	Not applicable.	All outputs.		For each survey cycle.
B3.8	For a continuing survey, have there been any changes over time in the sample design methodology? (I)			
	No. All live LU's always selected.	All outputs.		For each survey cycle.
B3.9	Record the design effects for key estimates. (M)			
	Not applicable.	All outputs based on a sample.		For each survey cycle.

Accuracy

B4. Implementing collection

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B4.1	Statement that estimates are subject to non-sampling error and an indication of the main sources of non-sampling error. (I)			
	Estimates are entered for non-responders. All outputs are based on data received in previous years. If no previous or comparable information is available, data is estimated from PAYE information on IDBR or other internal sources.	All outputs	For each survey cycle.	
B4.2	Coverage error. (M)			
	Not applicable.	All outputs based on a sample.	For each survey cycle.	
B4.3	Define coverage rate and identify it as a source of error. (I)			
	Not applicable.	All outputs based on a sample.	For each survey cycle.	
B4.4	Estimate rates of under coverage, duplication, ineligibility and misclassification. (I)			
	Not applicable.	All outputs based on a sample.	For each survey cycle.	

Accuracy

Ref.	Notes	Outputs applicable to	Frequency of production	Example	
B4.5	Describe methods used to deal with coverage issues. (I)				
	Not applicable.	All outputs based on a sample.	For any data output.		
B4.6	Assess the likely impact of coverage error on key estimates. (I)				
	Not Applicable.	All outputs based on a sample.	For each survey cycle.		
B4.7	Non-response error is defined and identified as a source of error. (I)				
	Not Applicable.	All outputs.	For each survey cycle.		
B4.8	Unit response rate by sub-groups: Household surveys. (I)				
	Not Applicable.	All outputs.	For each survey cycle.		
B4.9	Unit response rate by sub-groups: Business surveys (I)				
	Overall response rate generally around 98%. This level has been maintained historically.	All outputs.	For each survey cycle.		

Accuracy

Accuracy

Accuracy

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B4.10	<p>Bias due to unit non-response for key estimates. (M)</p> <p>Non-response occurs when it is not possible to collect information from a sample unit. Non-respondents may differ from respondents, which leads to non-response bias. Imputation on a company-by-company basis will counter this to a certain extent. The remaining error would then be negligible.</p>	All outputs.	For each survey cycle.	
B4.11	<p>Rate of complete proxy responses. (I)</p> <p>Not Applicable.</p>	All outputs apart from those based on other sources.	For each survey cycle.	
B4.12	<p>Rate of partial proxy response. (I)</p> <p>Not Applicable.</p>	All outputs apart from those based on other sources.	For each survey cycle.	

Accuracy

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B4.13	Key item response rates. (I)			
	Not Applicable.	All outputs.	For each survey cycle.	
B4.14	Bias due to item non-response for key estimates. (M)			
	Not Applicable.	All outputs.	For each survey cycle.	
B4.15	Differences between responders and non-responders are investigated. Differences which are known or are discovered are described. (I)			
	Not Applicable.	All outputs apart from those based on other sources.	For each survey cycle.	
B4.16	Scanning and keying error rates (where the required value cannot be rectified). (I)			
	Forms are keyed by an external contractor and validation reports are produced. Error rates are low at around = 0.2%.	All outputs apart from those based on other sources.	For each survey cycle.	

Accuracy

Ref.	Notes	Outputs applicable to	Frequency of production	Example	
B4.17	Proportion of respondents with difficulty answering individual questions. (I)				Accuracy
	The inquiries that make up the Census of All outputs are short and questions are apart from relatively simple as well as easy-to-understand. This is confirmed by customer / user feedback as part of the Quinquennial Review Process.	those based on other sources.	For each survey cycle.		
B4.18	Describe any differences between domains. (I)				Comparability
	Not Applicable.	All outputs.	For any data output.		
B4.19	Are there any items collected in the survey for which the data are suspect? (I)				Accuracy
	Not Applicable.	All outputs based on a survey.	For each survey cycle.		
B4.20	Assess differences due to different modes of collection. (I)				Accuracy
	There are no major differences in responses due to modes of collection. The majority of where there responses are by post, with a small number by fax. Towards the end of the field period figures are given over the telephone. Care is taken to ensure that these are recorded accurately.	All outputs collection mode.	For each survey cycle.		

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B4.21	Interviewer variance. (I)			
	Not Applicable.	All outputs using data collected by an interviewer.	For each survey cycle.	
B4.22	Number of attempts to interview (for contacts and non-contacts). (I)			
	Not Applicable.	All outputs apart from those based on other sources.	For each survey cycle.	
B4.23	Define measurement error and specify the main sources of measurement error for this output. (I)			
	Not Applicable.	All outputs.	Produced once for all outputs and revised where changes to the survey design necessitate a different statement.	
B4.24	Describe processes employed to reduce measurement error, including questionnaire development, interviewer training. (I)			
	Not Applicable.	All outputs.	For any data output.	

Accuracy

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B4.25	Variance and bias due to measurement error. (M)			
	Not Applicable.	All outputs.	For each survey cycle.	
B4.26	Define processing error and specify sources of processing error. (I)			
	Processing errors can occur during the entering of data phase. The likelihood of significant processing data entry errors is reduced by the binder validation report checking system. All variances are examined and where necessary checked with the data originator.	All outputs.	Produced once for all outputs and revised where changes to the survey design necessitate a different statement.	

Accuracy

Data Processing

4

B5. Editing, imputation and validation, derivation and coding

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B5.1	<p>Key item imputation rates. (I)</p> <p>Imputation can be used to compensate for non-response bias, in that any known characteristics of non-responders can be used to predict values for missing units. The imputation rate is an indicator of possible non-response bias, in that the imputation strategy will not perfectly compensate for all possible differences between responders and non-responders. Imputation takes place on a company-by-company basis, looking at previous patterns, similar companies in the sector and by using other internal surveys (QES).</p>	All outputs.	For each survey cycle.	
B5.2	<p>Unit imputation rate. (I)</p> <p>This is defined as the number of units imputed divided by the number of eligible units. Per suggested formula the rate is approximately 1% those (483 companies / 43,087)</p>	All outputs apart from those based on other sources.	For each survey cycle.	<p>Suggested formula:</p> $\frac{\text{Number of units imputed}}{\text{Total number of units}}$
B5.3	<p>Total contribution to key estimates from imputed values. (I)</p> <p>As above, the imputation rate is very small. This will have a negligible effect on the output. Historically, it is usually only small companies that are imputed.</p>	All outputs.	For each survey cycle.	

Accuracy

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B5.4	Assess the likely impact of non-response/imputation on final estimates. (I)			
	As above	All outputs apart from those based on other sources.	For each survey cycle.	
B5.5	Edit failure rate. (I)			
	Validation and discrepancy checks would detect approximately 300 (ICS) errors per census. This is very low and will have no effect on the output. Returns from each unit are checked and the company may be contacted if data is not within acceptable parameters and confirmed. Failure rate is estimated at 0.1%.	All outputs apart from those based on other sources.	For each survey cycle.	
B5.6	Editing rate (for key items). (I)			
	Due to the multiple validation checks there are very few units that require editing.	All outputs apart from those based on other sources.	For each survey cycle.	
B5.7	Total contribution to key estimates from edited values. (I)			
	Negligible impact, as above.	All outputs apart from those based on other sources.	For each survey cycle.	

Accuracy

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B5.8	<p>Proportion of responses requiring adjustment due to data not available as required, e.g. non-standard reference periods.</p> <p>(I)</p> <p>All data is provided in the required format.</p>	Business survey outputs.	For each survey cycle.	
B5.9	<p>Coding error rates.</p> <p>(I)</p> <p>There will be a small coding error rate for this inquiry. The main check, which is undertaken regularly, is the SIC code. This will allow changes to be identified and rectified.</p>	All outputs apart from those based on other sources.	For each survey cycle.	
B5.10	<p>Coder variance.</p> <p>(I)</p> <p>Coder variance is the error introduced by variation in the way different coders assign codes to responses. This is relevant where coding is carried out manually or with computer-assisted coding. SIC codes used for the Census of Employment are internationally standard and clearly defined. This reduces the chance of any variance.</p>	All outputs apart from those based on other sources.	For each survey cycle.	

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B5.11	<p>Processing bias and variance. (M)</p> <p>Processing bias is systematic error introduced by processing systems e.g. an error in programming coding software leads to the wrong code being consistently applied to a particular class. Processing variance is the random error introduced by processing systems, which across replications would cancel each other out – e.g. random keying errors in inputting data. Processing error can be introduced via a number of mechanisms, including keying, coding, editing, weighting and tabulating. Detecting processing bias and variance is not always possible, however. The affect of this bias on the Census of Employment is minimal, given the multiple validation checks.</p>	All outputs.	For each survey cycle.	

B6. Weighting and estimation

Ref.	Example	Outputs applicable to	Frequency of production	Notes
B6.1	<p>Sampling Error. (M)</p> <p>The sampling error is the difference between a population value and an estimate based on a sample. This difference comprises (systematic error caused by observing a sample that is not representative of the population) and variance (non-systematic error arising from random differences between samples). This measure informs users as to how accurately, for a given value, the survey sample estimates population values. It is an ideal measure, as it is not often possible to calculate. In practice, the standard error is referenced as an indicator of possible sampling error.</p> <p>As all LU's in the sampling frame are selected there is no sampling error.</p>	All outputs based on a sample.	For each survey cycle.	
B6.2	<p>Statement that estimates are subject to sampling error. (I)</p> <p>Not Applicable.</p>	All outputs based on a sample.	Produced once for all outputs and revised where changes to the survey design necessitate a different statement.	

Accuracy

Ref.	Example	Outputs applicable to	Frequency of production	Notes
B6.3	Estimate standard error (or coefficients of variation) for key estimates. (I)			
	Not Applicable.	All outputs based on a sample.	For each survey cycle.	
B6.4	Summarise methods used to derive or approximate standard errors. (I)			
	Not Applicable.	All outputs based on a sample.	For any data output.	
B6.5	Reference/link to documents containing detailed standard error estimates. (I)			
	Not Applicable.	All outputs based on a sample.	For any data output.	
B6.6	Describe variance estimation method including factors taken into account, e.g. misclassifications, non-response. (I)			
	Not applicable.	All outputs based on a sample.	For any data output.	
B6.7	If the survey was based on a sample, what method of sample weighting was used to calculate the estimates contained in the report? (I)			
	Not applicable.	All outputs based on a sample.	For each survey cycle.	

Ref.	Example	Outputs applicable to	Frequency of production	Notes
B6.8	Describe models used and the selection procedure. (I)			
	Not applicable.	All outputs derived from modelling.	For any data output.	
B6.9	Description and assumptions underlying models used in each output. (I)			
	Not applicable.	All outputs based on models.	For any data output.	
B6.10	Evaluation of whether the assumptions do/are likely to hold. (I)			
	Not applicable.	All outputs based on models.	For any data output.	
B6.11	Bias and variance introduced by models. (M)			
	Not applicable.	All outputs based on models	For output	

Accuracy

Data Analysis

5

B7. Analysis of primary outputs

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B7.1	Identify known gaps between key user needs, in terms of coverage and detail, and current data. (l)			
	The last Quinquennial Review findings confirmed that customers were happy with the information supplied from the Census of Employment publications. NISRA Customer Satisfaction Survey.	All outputs.	For each survey cycle.	
B7.2	Describe reasons for lack of completeness. (l)			
	Not applicable.	All outputs.	For each survey cycle.	
B7.3	Describe plans for meeting needs arising from lack of completeness. (l)			
	Not applicable.	All outputs.	At design or re-design stage of survey.	
B7.4	Describe results of user satisfaction assessments. (l)			
	The Census of Employment this is an essential economic dataset for anyone requiring analysis of the NI economy at below NI levels of detail. Results can be produced at detailed industrial level and for small geographical areas.	All outputs.	For subsequent evaluation.	

Relevance

Ref.	Notes	Outputs applicable to	Frequency of production	Example	Relevance
B7.5	<p>Describe any measures and actions taken to improve relevance based on customer feedback/satisfaction surveys.</p> <p>(I)</p> <p>All information available from the census subject to disclosure issues is provided on the DETI website. Further complex analysis is provided as and when it is requested subject to disclosure constraints. The most recent Quinquennial review recommended that the Census of Employment should continue to produce output at the same level of detail as at present.</p>	All outputs.	For subsequent evaluation.		
B7.6	<p>Present any reasons for lack of relevance.</p> <p>(I)</p> <p>Not applicable.</p>	All outputs.	For subsequent evaluation.		
B7.7	<p>Describe any standardisation applied to outputs, e.g. seasonal adjustment, age/sex standardisation, and equalisation.</p> <p>(I)</p> <p>Not applicable.</p>	Survey outputs; time-series; index numbers.	For any data output.		

Time series analysis

B8. Time series analysis

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B8.1	Provide an estimate of discontinuity due to change. (M)			
	Not applicable. Definitions have remained the same since the beginning of Census.	Time series.		Whenever there is a discontinuity caused by a change in the statistical methodology or real- world change.
B8.2	All discontinuities are flagged and explained. (I)			
	Not applicable.	Time series.		For any data output.
B8.3	Describe and assess impact on the data of any changes in concepts, definitions, classifications and methods over time; description and assessment of impact on the data of real-world events. (I)			
	Changes in classifications such as SIC updates with any impact upon the comparability of the series; data are explained in the Notes to Editors longitudinal section of the Census of Employment Statistics studies. Bulletin. The Census of Employment was coded by SIC 80 in the years up to and including 1991. In 1993 the Census was dual coded by SIC 80 and SIC 92. Since then, the Census has been coded by SIC 92. The Census data coded under SIC 80 is not comparable with data coded by SIC 92.	Time series.		For any data output.

Time series analysis

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B8.4	<p>Describe back series available.</p> <p>(1</p>			
	<p>The back series contains data that will allow users to assess changes over time.</p> <p>Census data has been stored electronically since 1995 and as such is readily available for analysis, subject to disclosure constraints.</p> <p>Prior to this the Census reports and data are stored in paper format.</p>	Time series.	For any data output.	

B9. Further analysis

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B9.1	For outputs revised on a regular basis: estimate the likely revision between provisional and final estimates and include known reasons for differences. (M)			
	The Census of Employment only produces final estimates.	Time series; systems of accounts; survey outputs.		For any data output.
B9.2	Flag data which are subject to revision and data which have already been revised. (I)			
	Not applicable.	Time series; systems of accounts; survey outputs.		For any data output.
B9.3	Provide timetable for revisions. (I)			
	Not applicable.	Time series; systems of accounts; survey outputs.		At planning stage.

Coherence

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B9.4	For ad hoc revisions, detail revisions made and provide reasons. (l)			
	As mentioned above any changes, revisions or reclassifications are detailed in the Notes to series; Editors section of the Statistics Bulletin.	Time series; systems of accounts; survey outputs.		For any data output.
B9.5	For published estimates that are used as sources to other outputs, a statement comparing the estimates and the reasons for any differences. (l)			
	Not applicable.	All outputs which may feed into other outputs.		As required.
B9.6	Provide a statement identifying any known similar estimates from other sources. (l)			
	Quarterly Employment Survey - The QES provides a count of employee jobs. It covers all public sector employers, all employers with 25 or more employees and a representative sample of smaller firms. It provides employee jobs estimates by gender, working pattern (full / part-time) and by Standard Industrial Classification 2003 (SIC03) for Northern Ireland as a whole. Labour Force Survey - The Labour Force Survey (LFS) is a quarterly sample survey carried out by interviewing people about their personal circumstances and work. The LFS provides information on, labour market structure, employment, unemployment, and economic activity groups within the labour market.	Survey outputs; outputs based on admin data; outputs based on other sources; census output.		For any data output.

Coherence

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B9.7	<p>Compare estimates from other sources, and include any known reasons for the differences. (l)</p> <p>The Quarterly Employment Survey is compared to the Census at a Reporting Unit level as well as a Northern Ireland level. Any discrepancies are investigated and amended where appropriate.</p>	<p>Survey outputs; outputs based on administrative data; census, other sources.</p>	<p>For any data output.</p>	
B9.8	<p>Reference/link to detailed revisions analyses are available. (l)</p> <p>Not applicable.</p>	<p>Systems of accounts.</p>	<p>For any data output.</p>	
B9.9	<p>What procedures were followed to prevent disclosure of confidential information? (l)</p> <p>There are a number of features in place that act to prevent disclosure of information. The databases are held on a secure network and are only available to section staff as required. All PC's require specific log-on codes and access is monitored by ISU. All files and returned forms are held in locked cabinets and keys are kept in a safe at night. Any Press Releases and Statistics Bulletin releases are in accordance with the National Statistics Protocols as are pre-release ministerial briefings.</p>	<p>All outputs based on a sample.</p>	<p>For any data output.</p>	

Relevance

Disemmination



B10. Dissemination

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B10.1	Describe methodology used to collect and compile the output (to include reference to how the data were derived, e.g. modelling, matching, projection). (I)			
	As outlined above, all live cases on the IDBR framework are selected.	All outputs.	For any data output.	
B10.2	Document availability of measures of accuracy of the data. (I)			
	As outlined above.	All outputs.	For any data output.	
B10.3	Are data accompanied with commentary including text, graphs, maps, etc? (I)			
	Statistics Bulletins contain commentary as well as text and graphs. This improves that clarity of the information and makes it easier to understand.	All outputs.	For any data output.	
B10.4	Describe key user needs for timeliness of data and how these needs have been addressed. (I)			
	The Census of Employment is a high profile publication and a key economic indicator of long-term trends in NI employee jobs. Only the Census of Employment can provide data on employee jobs for small geographical areas and industrial codes in NI. A more frequent survey would substantially increase the burden to business.	All outputs.	For any data output.	

Accessibility

Timeliness

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B10.5	Statistics are released to a pre-announced timetable with access to all at the same time. (I)			
	In accordance with National Statistics Protocols all release dates are pre-announced. The publications are released by the Press Office at the appropriate time and publications are also available on the website at 9.30am on the release day. (www.statistics.detini.gov.uk). Ministers and a small number of official have pre-release access to the data in accordance with the National Statistics Protocol on Release Practices.	All outputs.	Once at dissemination strategy planning stage and where there are subsequent re-designs of the dissemination strategy or of the survey.	
B10.6	Time lag: actual release date and scheduled release date. (M)			
	The Census of Employment is released on the pre-announced release day.	All outputs.	For any data output.	
B10.7	For customised data requests from existing sources, time lag: receipt of request for data and delivery date. (M)			
	Customised datasets and other ad-hoc requests from customers are dealt with as soon as possible and usually within 10 days. This can vary depending on the information requested.	All outputs.	For any data output.	
B10.8	Are there links to metadata? (I)			
	Yes, metadata links are written into the DETI Statistics Research Branch website.	All outputs.	As required.	

Timeliness

Accessibility

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B10.9	<p>Highlight availability of outputs and explain search and navigation tools available within web-based outputs.</p> <p>(I)</p> <p>Users can download publications from the Statistics Research Branch website. This website contains links, search and navigation tools to aid users.</p>	All outputs.	For any data output.	
B10.10	<p>Reference outputs in catalogues and other relevant documents.</p> <p>(I)</p> <p>The ONS publications Economic Trends and Labour Market Trends, and the NISRA Annual abstract contains figures from the Census of Employment. Publications are also available in Government libraries.</p>	All outputs.	For any data output.	
B10.11	<p>Describe procedures to request access to confidentialised public-use micro data sets, including costs.</p> <p>(I)</p> <p>Under the statistics of Trade & Employment Order 1988, other public bodies may request disclosive information by means of a direction issued by the Head of Department.</p>	All outputs apart from those based on other sources.	Once at dissemination strategy planning stage and where there are subsequent re-designs of the dissemination strategy or of the survey.	

Ref.	Notes	Outputs applicable to	Frequency of production	Example
B10.12	Describe procedures to request access to non-published data, including costs and average time to provide simple tables of non-published data. (l)			
	Non-published data can be provided – subject to the usual confidentiality constraints. These are requested by contacting the Branch or the statistician responsible directly. Contact details are provided on all publications. The time required for these requests would depend on a number of factors and is determined on an individual basis.	All outputs apart from those responsible directly. Contact based on sources.	Once at dissemination strategy planning stage and where there are subsequent re-designs of the dissemination strategy or of the survey.	
B10.13	A statement advising users on other statistics produced on a coherent basis. (l)			
	As mentioned in B9.6 above, other commentaries on the Northern Ireland economy are available.	All outputs.	For any data output.	
B10.14	Provide a statement advising users on other statistics that are not produced on a coherent basis. (l)			
	Publications are available from a number of economic commentators but these are not produced with input from the department and care should be used in comparisons.	All outputs.	For any data output.	
B10.15	Number of web hits. (l)			
	The number of web hits for the Census of Employment section of the website is not yet available, but will be available in the future.	Outputs on the web.	Every time a new data set is added to the NS web site.	
B10.16	Detail contact points for further information, including technical information. (l)			
	The statistician responsible puts details on all publications for the information of users. This includes an e-mail address.	All outputs.	For any data output.	

Coherence

Accessibility