

Standard Report
on
Methods and Quality (v1)
for
COMMUNITY INNOVATION SURVEY (CIS)

This documentation applies to the reporting period:

2004 – 2006 inclusive

Last edited: 15/October/2009

CENTRAL STATISTICS OFFICE

Skehard Road, Cork

021 4535000

www.cso.ie

Table of Contents

1 Overview

2 General Information

- 2.1 Statistical Category
- 2.2 Area of Activity
- 2.3 Organisational Unit Responsible, Persons to Contact
- 2.4 Objectives and Purpose; History
- 2.5 Periodicity
- 2.6 Client
- 2.7 Users
- 2.8 Legal basis

3 Statistical Concepts, Methods

- 3.1 Subject of the Statistics
- 3.2 Units of Observation/Collection Units/Units of Presentation
- 3.3 Data Sources
- 3.4 Reporting Unit/Respondents
- 3.5 Type of Survey/Process
- 3.6 Characteristics of the Sample/Process
 - 3.6.1 Population and Sampling Frame
 - 3.6.2 Sampling Design
- 3.7 Survey Technique/Data Transfer
- 3.8 Questionnaire (including explanations)
- 3.9 Participation in the Survey
- 3.10 Characteristics of the Survey/Process and its Results
- 3.11 Classifications used
- 3.12 Regional Breakdown of Results

4 Production of the Statistics, Data Processing, Quality Assurance

- 4.1 Data Capture
- 4.2 Coding
- 4.3 Data Editing
- 4.4 Imputation (for Non-Response or Incomplete Data Sets)
- 4.5 Grossing and Weighting
- 4.6 Computation of Outputs, Estimation Methods Used
- 4.7 Other Quality Assurance Techniques Used

5 Quality

- 5.1 Relevance
- 5.2 Accuracy and Reliability
 - 5.2.1. Sampling Effects, Representativity
 - 5.2.2. Non-Sampling Effects
 - 5.2.2.1 Quality of the Data Sources used
 - 5.2.2.2 Register Coverage
 - 5.2.2.3 Non-response (Unit and Item)
 - 5.2.2.4 Measurement Errors
 - 5.2.2.5 Processing Errors
 - 5.2.2.6 Model-related Effects
- 5.3 Timeliness and Punctuality
 - 5.3.1 Provisional Results
 - 5.3.2 Final Results
- 5.4 Coherence
- 5.5 Comparability
- 5.6 Accessibility and Clarity
 - 5.6.1 Assistance to Users, Special Analyses
 - 5.6.2 Revisions
 - 5.6.3 Publications
 - 5.6.3.1 Releases, Regular Publications
 - 5.6.3.2 Statistical Reports
 - 5.6.3.3 Internet
 - 5.6.4 Confidentiality

6 Additional documentation and publications

1 Overview

The biennial Community Innovation Survey (CIS) of enterprises measures the level of innovation activity among enterprises in Ireland. The survey covers industry and selected services sector enterprises with 10 or more persons engaged. The survey is carried out under Commission Regulation (EC) No 1450/2004 implementing Decision No 1608/2003/EC. The 2004-2006 edition of the survey collected the information with relevance to this three-year period. It should be noted that Forfás previously carried out the CIS but the CSO and Forfás now jointly conduct the survey.

2 General Information

2.1 Statistical Category

Primary Statistical Survey

2.2 Area of Activity

Science and Technology – Community Innovation Survey (CIS)

2.3 Organisational Unit Responsible, Persons to Contact

Structural Business Statistics, Business Statistics Directorate

Carol Anne Hennessy 021 – 453 5161 CarolAnne.Hennessy@csso.ie

2.4 Objectives and Purpose; History

The Community Innovation Survey (CIS) is a survey of the innovation activities of enterprises with the primary objective of gathering information on rates of innovation for enterprises. The survey collects information about product and process innovation as well as organisational innovation and other key variables. The 2004-2006 CIS was the fifth edition of the survey. Previous editions of the survey were carried out by Forfás. However in the National Statistics Board Strategy for Statistics 2003-2008, the Boards articulated a medium-term strategy to support the development of Ireland's statistical system. A key pillar of the Board's strategy was for the CSO to work with Government Departments and Agencies to maximise the use of administrative data to generate statistics. In particular, the Board in conjunction with a CSO Expert Group produced a 'Policy Needs for Statistical Data on Enterprises' (link: http://www.nsb.ie/pdf_docs/policy_needs.pdf) report in November 2005. This report included a specific recommendation (recommendation 13) that the Community Innovation Survey should be jointly conducted by the CSO and Forfás.

2.5 Periodicity

This postal survey is carried out biennially.

2.6 Client

The Community Innovation Survey (CIS) is carried out under Commission Regulation (EC) No 1450/2004 implementing Decision No 1608/2003/EC.

2.7 Users

- Eurostat
- Forfás
- Department Of Enterprise, Trade and Employment
- Other Departments
- Other agencies and interest groups
- General public

2.8 Legal basis

The Community Innovation Survey (CIS) is carried out under Commission Regulation (EC) No 1450/2004 implementing Decision No 1608/2003/EC. Data is collected in accordance with Section 33 of the Statistics Act 1993, and the survey is carried out under the agreed set of international rules as laid out in the Organisation for Economic Co-operation and Development (OECD) Oslo manual.

3 Statistical Concepts, Methods

3.1 Subject of the Statistics

The primary subject of the Community Innovation Survey (CIS) is to collect information regarding innovation statistics.

3.2 Units of Observation/Collection Units/Units of Presentation

The survey is issued to those enterprises with 10+ persons engaged in NACE (European Industrial Activity Classification) Revision 1.1 sectors 10-37, 51, 60-67, 72, 74.2 and 74.3. A total of 4,150 forms were issued in the 2004-2006 edition of the survey. The CSO's Central Business Register (CBR) is used to generate the list of enterprises in the survey.

Results are generally expressed as performing enterprises as a percentage of the enterprise population. In addition, monetary amounts are used to present results related to research and development expenditure.

3.3 Data Sources

- CSO Business Register data.
- CSO Structural Business Statistics data.
- Forfás Business Expenditure on Research and Development Survey data.

3.4 Reporting Unit/Respondents

Reporting unit – Results aggregated to NACE 2 digit sectors according to the coverage detailed in section 3.2.

Respondents - Enterprises with 10+ persons engaged.

3.5 Type of Survey/Process

Census – 50+ persons engaged

Stratified random sample – Less than 50 persons engaged (It should be noted that a census is conducted for certain strata in the less than 50 persons engaged category where the number of enterprises in the strata is low.)

3.6 Characteristics of the Sample/Process

3.6.1 Population and Sampling Frame

The survey population is made up of all enterprises in NACE Rev1.1 sectors 10-37, 51, 60-67, 72, 74.2 and 74.3 with 10+ persons engaged.

The sampling frame is the CSO's Central Business Register.

3.6.2 Sampling Design

The core target population is as per under Commission Regulation (EC) No 1450/2004 – namely those enterprises with at least 10 persons engaged in NACE Rev1.1 sectors 10-37, 51, 60-67, 72, 74.2 and 74.3. A full list of all relevant enterprises is generated for this coverage using the CSO's Central Business Register. This population is then subdivided into strata which were 2 digit NACE by employment size class: namely enterprises with 10 to 49 persons engaged, 50 to 249 persons engaged and 250+ persons engaged.

A census is conducted of all enterprises with 50 or more persons engaged while a census is also conducted for certain strata in the less than 50 persons engaged category where the number of enterprises in the strata is low. Finally a sample of enterprises is taken for those remaining enterprises with less than 50 persons engaged who are not included in the census. Neyman allocation is used to select the sampling fraction in each of these remaining strata and the average sampling fraction is approximately 40%. The actual enterprises selected are selected using random selection.

3.7 Survey Technique/Data Transfer

1. Survey forms are printed in the Printing Section of the CSO and posted out to respondents for self-completion. When a survey form is returned to the CSO the form is scrutinised to resolve any substantial errors prior to scanning. When a batch is large enough (circa 20 forms) the forms are scanned and verified before being edited.

The failed edits that arise are scrutinised and resolved if possible. If an edit cannot be easily resolved, then the enterprise is contacted for clarification and the edit is subsequently cleared based on the clarification.

The survey form can be viewed at:

http://www.cso.ie/surveysandmethodologies/surveyforms/documents/industry/pdf_docs/Forfas_form.pdf

2. The 2004-2006 edition of the survey was also available for completion on the CSO website as an e-form.

3.8 Questionnaire (including explanations)

The Community Innovation Survey (CIS) collects information about product and process innovation as well as organisational innovation during a three year reference period. The principal variables collected relate to:

- Product Innovation;
- Turnover from new to market and new to firm product innovations;
- Process Innovation;
- Ongoing and abandoned innovation;
- Innovation expenditure;
- Innovation co-operation;
- Innovation hampering factors;
- Organisational Innovation

3.9 Participation in the Survey

Voluntary.

3.10 Characteristics of the Survey/Process and its Results

As detailed above, the Community Innovation Survey (CIS) 2006 collected specific data relating to the core topics of product, process and organisational innovation, with the principal variables listed in section 3.8.

The survey is conducted as a census of those enterprises with 50+ persons engaged and as a random stratified sample of enterprises with less than 50 persons engaged. Data from respondents is grossed using the CSO's Central Business Register and aggregated results are then published.

The Community Innovation Survey (2004 – 2006) First Findings were issued in June 2008 and is available at:

http://www.cso.ie/releasespublications/documents/information_tech/current/comminnfirst.pdf

The Community Innovation Survey (2004 – 2006) publication was issued in April 2009 and is available at:

http://www.cso.ie/releasespublications/documents/information_tech/2006/comminn0406.pdf

3.11 Classifications used

As per Commission Regulation (EC) No 1450/2004 NACE REV1.1 is used for the Community Innovation Survey (CIS) Survey.

The relevant sectors from this classification are:

Mining and Quarrying enterprises – NACE 10-14

Manufacturing enterprises – NACE 15-37

Electricity, gas and water supply enterprises – NACE 40-41

Selected services sector enterprises – NACE 51, 60-67, 72, 74.2, 74.3

Please see the following link for details of NACE coding:

http://www.cso.ie/surveysandmethodologies/classification_indus_act.htm

3.12 Regional Breakdown of Results

Results are not broken down by region. The results (at 2 digit NACE level) are available at a national level only.

4 Production of the Statistics, Data Processing, Quality Assurance

4.1 Data Capture

Once the Survey forms are returned to the Business Statistics Integration Section, they are scrutinised and are categorised (completed/blank/gone away) in batches for scanning and verifying using teleforms scanning software. The output of this scanning process is a CSV formatted file which is then imported into SAS for further checking, processing and analyses

4.2 Coding

The CSO's Central Business Register (CBR) is used to activity code the surveyed enterprises. No questions are included on the form with regard to enterprise activity. When the returned form is scanned the CBR number of the enterprise is captured and linked to the activity code of the enterprise as per the CBR.

4.3 Data Editing

Edit programs are run on SAS software to perform checks on consistency of returns between separate questions and for an examination of returns with monetary amounts for research and development related expenditure.

The edits carried out includes:

1. Where an enterprise indicates that it was engaged in product innovation but did not indicate who developed these product innovations.
2. Where an enterprise indicates that it was engaged in process innovation but did not indicate who developed these process innovations.
3. Enterprises who indicates that they carried out some aspect of research and development but who did not indicate their expenditure on research and development.

A listing of all the failed edits is sorted by enterprise number and printed. The image of the form is called up on the PC screen and the reason for the edit failure is identified on the form. Each failed edit is checked for any scanning errors and for any comment on the form that may explain the failure. A limited number of returns from the previous survey data held by Forfás are manually compared to some returns for consistency purposes. It should be noted that to reduce response burden automatic updates are used in general but if the edit failure cannot be resolved locally using automatic updates, then the enterprise is contacted to attempt to resolve the issue. Once all errors are amended a copy is made of the clean dataset.

Reports are run during the survey period recording the number of forms returned by NACE group and employment size and to ensure that there is a response in every strata used for grossing so that it will be possible to gross all strata. In addition, reports are run to estimate the number of edits outstanding.

4.4 Imputation (for Non-Response or Incomplete Data Sets)

There is no formal imputation for non-response. However, use is made of imputation for both unit and item non-response on a manual basis from data held by Forfás for previous returns and from other available resources such as the Forfás Business Expenditure on Research and Development Survey data. In the absence of any other source for missing monetary amounts in the case of research and development expenditure, missing values are estimated using averages from other responding enterprises by NACE and employment size. The grossing factors are verified by comparing the grossed values for turnover and employment to the results of the 2006 Census of Industrial Production and Annual Services Inquiry.

4.5 Grossing and Weighting

Grossing is performed using SAS software.

Grossing factors for the Community Innovation Survey (CIS) are calculated using the inverse of the number of valid respondents divided by the relevant population from the CBR. Grossing is calculated for

strata of 2-digit NACE by employment size class. There are 3 grossing factors used: the number of enterprises, persons engaged and turnover. The factors which are calculated are then applied to the survey results in order to gross the data up to the relevant populations.

Weighting is not used for the derivation of results.

4.6 Computation of Outputs, Estimation Methods Used

Published results are aggregated to National level and also a limited number of tables are produced at 2-digit NACE sectoral level or at groupings at 2-digit NACE sectoral level in situations where confidentiality would be breached.

Estimation is not used for the derivation of results.

4.7 Other Quality Assurance Techniques Used

The sample is selected from the latest version of the Central Business Register (CBR). Duplicate enterprises, companies that have ceased trading or who have stated that they are unable to provide the necessary data are removed from the sample.

5 Quality

5.1 Relevance

The Community Innovation Survey (CIS) is carried out under Commission Regulation (EC) No 1450/2004 implementing Decision No 1608/2003/EC. Data are collected in accordance with Section 33 of the Statistics Act 1993, and the survey is carried out under the agreed set of international rules as laid out in the OECD Oslo manual.

The main users of the CIS 2006 results are:

- Eurostat
- Forfàs
- Department Of Enterprise, Trade and Employment
- Other Departments
- Other agencies and interest groups
- General public

Results are primarily used to inform policy makers with regard to data on innovation indicators.

5.2 Accuracy and Reliability

5.2.1. Sampling Effect & representivity

A representative sample is taken of all enterprises with 10 to 49 persons engaged and a full census of enterprises with 50+ persons engaged is also taken. Some editing of the sample takes place to remove duplicates or enterprises that have ceased trading. Eurostat determines the target population. The core NACE categories generally remain consistent across each survey. The Community Innovation Survey (CIS) 2004-2006 was the first time CSO ran the survey but it would be intended to rotate the sample for those enterprises with 10-49 persons engaged in order to reduce the response burden on these smaller enterprises.

Every effort is made to ensure that there are a sufficient number of responses in each cell to be representative of the sample and to ensure that confidentiality is maintained.

5.2.2. Non-Sampling Effects

No sampling effect as the survey is a mixture of census and sample.

5.2.2.1 Quality of the Data Sources used (other than survey register)

The random stratified sample is taken from the Central Business Register. No other data sources are used.

5.2.2.2 Register Coverage

The sample for this survey is taken from the Central Business Register (CBR). Forms are issued to all enterprises with at least 50 persons engaged and a stratified random sample is conducted for those enterprises with less than 50 persons engaged. It should be noted that a census is conducted for certain strata in the less than 50 persons engaged category where the number of enterprises is low.

Coverage errors – there may be some multiple listings (duplicates) and ceased companies when the sample is first taken from the Central Business Register. They are removed from the sample when discovered. Therefore, coverage is restricted to the required NACE categories only.

5.2.2.3 Non-response (Unit and Item)

Unit non-response: The response rate for the Community Innovation Survey (CIS) 2004-2006 was approximately 48%.

The main reasons for unit non-response are as follows:

1. Company Gone-Away from the address on the survey form
2. Duplicate – company receive two forms with different reference numbers
3. Company has ceased trading
4. Form returned blank
5. There is a query about the survey form
6. Company refuses to respond

The section investigates these issues and resolves them as appropriate e.g. ringing the company for confirmation of information, removing the company from the survey etc.

Item non-response: Item non-response is not formally measured in the survey. Where a part of the form is not completed the following is done:

1. An attempt is made to automatically update the form using edit checks
2. A manual check to see if the variable was captured in the last running of the survey was the particular enterprise in question and if the answer can be reused
3. Reference is made depending on the variable in question to see if information is available from the Forfás Business Expenditure on Research and Development Survey
4. Missing values estimated using averages from other responding enterprises by NACE and employment size
5. Contact enterprise to confirm the details on the survey form

5.2.2.4 Measurement Errors

Measurement Error is not formally calculated for the survey. However, the following should be noted:

- (a) questionnaire - we ensure that the questionnaire is clear with definitions provided as appropriate. The majority of questions are tick box style questions which makes the form easier to complete.
- (b) data collection method – there are 2 data collection method: by post and by electronic form (e-form). Neither method is unlikely of itself to be a source of measurement error and because of validation which are part of the e-form this method of return is likely to reduce potential errors (telephone interviewing is not used).
- (c) interviewer - field interviewers are not used by the survey.
- (d) respondent – survey data can contain sensitive and confidential information. The introduction on the form makes reference to Section 33 of the Statistics Act, 1993 guaranteeing the confidentiality of all data returned to ensure that respondents are not reluctant to provide accurate information.

5.2.2.5 Processing Errors

(A) Data Capture Errors are likely to be a low risk as survey forms are scanned and verified.

(B) Data Editing - Edits are run which check for the following:

- That all the relevant questions have been answered
- There is internal consistency in answers within the form

If the edit cannot be resolved locally, then the enterprise is contacted to confirm the returned data. The return is manually changed if required.

(C) NACE Code and NACE division are linked to the enterprise number as part of the original register information. This means that the possibility of a NACE misclassification error is very low.

5.2.2.6 Model-related Effects

There is no formal imputation for non-response. However, use is made of imputation for both unit and item non-response on a manual basis from data held by Forfás for previous returns, and from other available resources such as the Forfás Business Expenditure on Research and Development Survey data. In the absence of any other source for missing monetary amounts in the case of research and development expenditure, missing values are estimated using averages from other responding enterprises by NACE and employment size. The grossing factors are verified by comparing the grossed values for turnover and employment to the results of the 2006 Census of Industrial Production and Annual Services Inquiry.

5.3 Timeliness and Punctuality

5.3.1 Provisional Results

The Community Innovation Survey 2004 – 2006 (First Findings) release was issued in June 2008. This is T+18 months relative to the reference period.

http://www.cso.ie/releasespublications/documents/information_tech/current/comminnfirst.pdf

5.3.2 Final Results

The Community Innovation Survey 2004 – 2006 (Final Results) publication was issued in March 2009. This is T+27 months relative to the reference period.

http://www.cso.ie/releasespublications/documents/information_tech/2006/comminn0406.pdf

5.4 Coherence

Turnover and employment figures generated by the grossing exercise are compared to results from the Census of Industrial Production and the Annual Services Inquiry (for those NACE codes where results were available in the Structural Business results). Where results deviated from expected results data is reviewed and updated if required. In general, there are no directly comparable results available for comparison, and there are no directly comparable short term results.

5.5 Comparability

Forfás carried out the Community Innovation Survey (CIS) prior to the 2004-2006 survey. As a result, a break in series was declared between results of the CIS2004-2006 and previous results from the CIS. This was clearly referenced in both the preliminary release and the final publication:

“Methodological changes were introduced in the CIS 2004-2006 and as a consequence, no conclusions should be drawn regarding the direction or scale of any real changes between CIS 2002-2004 (CIS4) and CIS 2004-2006.” As a result, direct year on year comparisons were difficult to establish.

In addition, as a standardised survey is carried out across all member states of the European Union, results can be compared and cross referenced against other states.

5.6 Accessibility and Clarity

5.6.1 Assistance to Users, Special Analyses

Survey information and the survey form are available to users on the CSO website:
http://www.cso.ie/surveysandmethodologies/surveys/industry/com_inn_survey.htm

Specific user requests are acceded to where possible, and where confidentiality issues do not arise.

5.6.2 Revisions

There were no revisions to published data – some data queries were raised by Eurostat upon submission of data but these matters were restricted to keying errors on the data submission worksheet.

5.6.3 Publications

5.6.3.1 Releases, Regular Publications

The Community Innovation Survey (2004 – 2006) First Findings: June 2008.

http://www.cso.ie/releasespublications/documents/information_tech/current/comminnfirst.pdf

The Community Innovation Survey (2004 – 2006) Publication: April 2009.

http://www.cso.ie/releasespublications/documents/information_tech/2006/comminn0406.pdf

5.6.3.2 Statistical Reports

N/A

5.6.3.3 Internet

First Findings:

http://www.cso.ie/releasespublications/documents/information_tech/current/comminnfirst.pdf

Final Results Publication:

http://www.cso.ie/releasespublications/documents/information_tech/2006/comminn0406.pdf

5.6.4 Confidentiality

In line with Section 33 of the Statistics Act, 1993 all data are treated as strictly confidential. In order to ensure confidentiality, results are produced at NACE sectional level. Appropriate confidentiality checking is run on data and as per usual CSO requirements the following rules are applied:

6 Additional documentation and publications

CSO Databank:

<http://www.cso.ie/px/pxeirestat/database/eirestat/Community%20Innovation%20Survey/Community%20Innovation%20Survey.asp>

Statcentral:

<http://www.statcentral.ie/viewStat.asp?id=202>

EUROSTAT Data:

Data which is submitted to Eurostat can be viewed at:

http://epp.eurostat.ec.europa.eu/portal/page/portal/science_technology_innovation/introduction