

Standard Report
on
Methods and Quality (v1)

for
BUSINESS EXPENDITURE ON RESEARCH AND DEVELOPMENT (BERD)
SURVEY

This documentation applies to the reporting period:

2007- 2008 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 Business Expenditure on Research and Development (BERD) Survey of enterprises measures the research and development expenditure, human resources and other research and development related topics for enterprises in Ireland. The survey is designed to be a census of all enterprises who are believed to be engaged in research and development activities in all business sectors of the economy. The survey is carried out under Commission Regulation (EC) No 753/2004 implementing Decision No 1608/2003/EC. It should be noted that Forfás previously carried out the BERD Survey but the CSO and Forfás now jointly conduct the survey under Section 11 of the Statistics Act, 1993.

2 General Information

2.1 Statistical Category

Primary Statistical Survey

2.2 Area of Activity

Science and Technology – Business Expenditure on Research and Development (BERD) Survey

2.3 Organisational Unit Responsible, Persons to Contact

Structural Business Statistics, Business Statistics Directorate

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

2.4 Objectives and Purpose; History

The Business Expenditure on Research and Development (BERD) Survey is a survey of the research and development expenditure and human resources of enterprises with the primary objective of gathering detailed information on research and development activities in Ireland. The survey was also used to collect other related research and development topics. 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 Business Expenditure on Research and Development Survey should be jointly conducted by the CSO and Forfás.

2.5 Periodicity

This postal survey is carried out biennially.

2.6 Client

The Business Expenditure on Research and Development (BERD) Survey is carried out under Commission Regulation (EC) No 753/2004 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 Business Expenditure on Research and Development (BERD) Survey is carried out under Commission Regulation (EC) No 753/2004 implementing Decision No 1608/2003/EC. The survey was jointly conducted under Section 11 of the Statistics Act, 1993, and data were collected in accordance with Section 33 of the Statistics Act 1993. The survey was carried out under the agreed set of international rules as laid out in the Organisation for Economic Co-operation and Development (OECD) Frascati manual.

3 Statistical Concepts, Methods

3.1 Subject of the Statistics

The primary subject of the BERD Survey is to collect information about research and development expenditure and research and development human resources for all business sectors of the economy.

3.2 Units of Observation/Collection Units/Units of Presentation

The survey is designed to be a census of all enterprises who are believed to be engaged in research and development activities in all business sectors of the economy. Just over 2,700 such enterprises were identified and survey forms were issued to these enterprises. The CSO's Central Business Register (CBR) was used to generate the necessary information to issue the forms to these enterprises.

Results are generally expressed as monetary amounts, numbers of persons and number of performing enterprises. In addition, some results are expressed as performing enterprises as a percentage of the enterprise population

3.3 Data Sources

- CSO Business Register data.
- CSO Structural Business Statistics data.
- CSO Balance of Payment (BOP) data.
- CSO/Forfás Community Innovation Survey (CIS) data.

3.4 Reporting Unit/Respondents

Reporting unit – Results aggregated to NACE (European Industrial Activity Classification) 2 digit sectors for all sectors engaged in research and development activity.

Respondents – Target of all enterprises with research and development activities.

3.5 Type of Survey/Process

Census of all enterprises believed to be engaged in research and development activities.

3.6 Characteristics of the Sample/Process

3.6.1 Population and Sampling Frame

The survey population is made up of all enterprises all business sectors of the economy that are believed to be engaged in research and development activities. These enterprises are identified from various sources including previous responses to the survey, existing CSO data from Structural Business Statistics, Balance of Payments data, enterprises coded in the CSO's Business Register in section 73 of the NACE Rev 1.1 Classification (section 73 is Research and Development) and Community Innovation Survey (CIS) data along with other administrative data sources. This information is used to create a register of likely performers and this register was then supplemented with additional information from the CSO's Central Business Register.

3.6.2 Sampling Design

N/A – the survey is conducted as a census of all likely research and development performers.

3.7 Survey Technique/Data Transfer

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/BERD07.pdf

3.8 Questionnaire (including explanations)

The Business Expenditure on Research and Development (BERD) Survey 2007/2008 collected information about research and development expenditure as well as human resources engaged in research and development activities primarily for the reference year of 2007. Enterprises were also asked for estimates of expected research and development expenditure in 2008. The principal variables collected relate to:

- Detailed information on research and development expenditure;
- Sources of funds for research and development expenditure;
- Detailed information on research and development personnel;
- Recruitment of PhD researchers;
- Research and development collaboration;

3.9 Participation in the Survey

Voluntary.

3.10 Characteristics of the Survey/Process and its Results

As detailed above, the Business Expenditure on Research and Development (BERD) Survey 2007/2008 collected specific data relating to the core topics of research and development related expenditure and human resources, with the principal variables listed in section 3.8.

The survey is conducted as a census of all enterprises in all business sectors of the economy that are believed to be engaged in research and development activities. Data from respondents is grossed using the CSO's Central Business Register and aggregated results are then published.

The Business Expenditure on Research and Development (BERD) Survey 2007/2008 Preliminary Findings were issued in March 2009 and is available at:

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

The Business Expenditure on Research and Development (BERD) Survey 2007/2008 publication was issued in July 2009 and is available at:

http://www.cso.ie/releasespublications/documents/information_tech/2008/berdpub0708.pdf

3.11 Classifications used

As per Commission Regulation (EC) No 753/2004 NACE REV1.1 is used for the 2007/2008 edition of the BERD Survey.

All business sectors of the economy where there are enterprises believed to be engaged in research and development activities are covered in the survey.

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

Some of the main variables collected are published by region. The regions which are used are NUTS2 (Nomenclature of Territorial Units) regions which were proposed by Government and agreed by Eurostat in 1999. The survey form asks enterprises to detail total research and development expenditure along with total persons engaged in research and development activities in each branch/plant by county. This information is then used to allocate data regarding expenditure and persons engaged in research and development to the appropriate NUTS2 region.

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 form - engaged in research and development activities/completed form - not engaged in research and development activities/blank/gone away/duplicate) 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.

The edits carried out include:

1. Where the research and development expenditure of each sub-category was not equal to the total expenditure.
2. Where the source of funds of expenditure did not equal the total expenditure.
3. Where the types of research and development activity did not sum to 100%.
4. Where the sub-categories of research and development personnel did not sum to the totals of research and development personnel.
5. Where the total research and development expenditure did not equal the total by county.
6. Where the total research and development personnel did not equal the total by county.

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. Returns are also compared to the Annual Business Survey returns in Forfás 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 to check on an ongoing basis what level of responses have been received. In addition, reports are run to estimate the number of edits outstanding.

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

Use was made of imputation for both unit and item non-response in the first instance from data held by Forfás for previous returns and from other available resources such as the Forfás Annual Business Survey data. In addition, in the absence of any other source for missing data, missing values were estimated using averages from other responding enterprises by NACE and employment size.

4.5 Grossing and Weighting

Grossing is performed using SAS software.

Grossing factors for the BERD Survey 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 2 grossing factors used: the number of enterprises along with persons engaged. 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

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. The nature of the survey is such that a low number of enterprises are responsible for a high percentage of research and development expenditure and human resources. These enterprises are identified and these returns are scrutinised in detail.

5 Quality

5.1 Relevance

The Business Expenditure on Research and Development (BERD) Survey is carried out under Commission Regulation (EC) No 753/2004 implementing Decision No 1608/2003/EC. The survey is jointly conducted under Section 11 of the Statistics Act, 1993, and data are collected in accordance with Section 33 of the Statistics Act 1993. The survey is carried out under the agreed set of international rules as laid out in the OECD Frascati manual.

The main users of the BERD Survey 2007 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 research and development indicators.

5.2 Accuracy and Reliability

5.2.1. Sampling Effect & representivity

The survey is conducted as a census of all enterprises in all business sectors of the economy that are believed to be engaged in research and development activities. Some editing takes place to remove duplicates or enterprises that have ceased trading.

5.2.2. Non-Sampling Effects

No sampling effect as survey is a census.

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

All required information is taken from the Central Business Register after the target list of enterprises has been created. No other data sources are used.

5.2.2.2 Register Coverage

The survey is conducted as a census of all enterprises in all business sectors of the economy that are believed to be engaged in research and development activities.

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.

5.2.2.3 Non-response (Unit and Item)

Unit non-response: The response rate for BERD 2007/2008 Survey was approximately 67%.

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 surveyed 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 Annual Business Survey or any other Forfás data source
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.
- (b) data collection method – forms are collected by post. This method is unlikely of itself to be a source of measurement error (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

Use is made of imputation for both unit and item non-response in the first instance from data held by Forfás for previous returns and from other available resources such as the Forfás Annual Business Survey data. In addition, in the absence of any other source for missing data, missing values are estimated using averages from other responding enterprises by NACE and employment size.

5.3 Timeliness and Punctuality

5.3.1 Provisional Results

The Business Expenditure on Research and Development Survey 2007/2008 (Preliminary Findings) release was issued in March 2009. This is T+15 months relative to the 2007 data or T+3 months relative to 2008 estimates.

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

5.3.2 Final Results

The Business Expenditure on Research and Development Survey 2007/2008 (Final Results) publication was issued in July 2009. This is T+19 months relative to the 2007 data or T+7 months relative to 2008 estimates.

http://www.cso.ie/releasespublications/documents/information_tech/2008/berdpub0708.pdf

5.4 Coherence

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 BERD Survey prior to the 2007/2008 survey. As a result, a break in series was declared between results of the BERD2007/2008 and previous results from the CIS. This was clearly referenced in both the preliminary release and the final publication:

Text in the preliminary release:

“Methodological changes have been introduced in the BERD 2007-2008 Survey. As a consequence of this, no conclusions should be drawn regarding the direction or scale of any real changes between BERD2005-2006 and BERD 2007-2008. Results for previous years are included but comparisons should be treated with caution.”

Text in the final publication:

“Methodological changes were introduced in the BERD 2007/2008 Survey. As a consequence of this, no conclusions should be drawn regarding the direction or scale of any real changes between BERD 2005/2006 and BERD 2007/2008. Summary results for previous years are included in Appendix 3 but comparisons should be treated with caution. In particular, it is not advised to make any comparisons of results at detailed sectoral level due to sectoral differences in the CSO Business Register (used for the 2007/2008 BERD Survey) when compared with the Forfás Business Register (used for previous editions of the BERD Survey). This CSO Business Register is used as the sampling frame for all CSO business surveys.”

As a result, direct year on year comparisons were difficult to establish and indeed were flagged (as above) as not being advisable.

Variables can be compared against the results of other European Union Member States but as the primary results are totals of expenditure and research personnel it is difficult to make any comparisons with regard to expected results.

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/Business_Expenditure_Research_Develop.htm

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

5.6.2 Revisions

There were minor revisions to published data in the preliminary release compared to the final publication. As an illustration, total research and development expenditure was €1,598m in the preliminary release compared to €1,603m in the final release, an increase of 0.3%.

5.6.3 Publications

5.6.3.1 Releases, Regular Publications

The Business Expenditure on Research and Development Survey 2007/2008 Preliminary Findings were issued in March 2009.

The Business Expenditure on Research and Development Survey 2007/2008 publication was issued in July 2009.

5.6.3.2 Statistical Reports

N/A

5.6.3.3 Internet

Preliminary Findings:

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

Final Results Publication:

http://www.cso.ie/releasespublications/documents/information_tech/2008/berdpub0708.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:

Preparation of results for the CSO databank is ongoing and is expected to be completed shortly.

Statcentral:

<http://www.statcentral.ie/viewStat.asp?id=240>

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