



**An  
Phríomh-Oifig  
Staidrimh**

Central  
Statistics  
Office

# **Standard SIMS Report:**

## **Business Expenditure on Research and Development**



# **Single Integrated Metadata Structure (SIMS) Report**

## **For**

# **Business Expenditure on Research and Development**

This documentation applies to the reporting period:  
**2023-2024**

Last edited:

April 2025



## 1. Table of Contents

1. Table of Contents .....	3
2. Introduction .....	5
3. Contact .....	5
4. Metadata Update .....	5
4.1. Metadata last certified .....	5
4.2. Metadata last posted .....	5
4.3. Metadata last update .....	5
5. Statistical Presentation .....	6
5.1. Data Description .....	6
5.2. Classification System .....	6
5.3. Sector Coverage .....	6
5.4. Statistical Concepts and definitions .....	6
5.5. Statistical Unit .....	6
5.6. Statistical Population .....	6
5.7. Reference Area .....	7
5.8. Time Coverage .....	7
5.9. Base period .....	7
6. Unit of Measure .....	7
7. Reference Period .....	7
8. Institutional Mandate .....	7
8.1. Legal Acts and other agreements .....	7
8.1.1. European legislation .....	7
8.1.2. National Legislation .....	7
8.1.3. Standards and manuals .....	7
8.2. Data Sharing .....	7
9. Confidentiality .....	8
9.1. Confidentiality – policy .....	8
9.2. Confidentiality – data treatment .....	8
10. Release Policy .....	8
10.1. Release Calendar .....	8
10.2. Release calendar access .....	8
10.3. User access .....	8
11. Frequency of Dissemination .....	9
12. Accessibility and clarity .....	9
12.1. News release .....	9
12.2. Publications .....	9
12.3. On-line database .....	9
12.3.1. AC 1. Data tables - consultations .....	9
12.4. Micro-data Access .....	9
12.5. Other .....	9
12.5.1. AC2. Metadata consultations .....	9
12.6. Documentation on Methodology .....	9
12.6.1. AC3 – Metadata completeness – rate .....	9
12.7. Quality Documentation .....	10
13. Quality Management .....	10
13.1. Quality Assurance .....	10
13.2. Quality Assessment .....	10
14. Relevance .....	10
14.1. User Needs .....	10
14.1.1. Main National Users .....	10
14.1.2. Principal External Users .....	11
14.2. User Satisfaction .....	11
14.3. Data Completeness .....	11
14.3.1. Data Completeness rate .....	11
15. Accuracy and reliability .....	11



15.1. Overall accuracy.....	11
15.2. Sampling Error.....	12
15.2.1. A1. Sampling error indicator.....	12
15.3. Non-sampling Error .....	12
15.3.1. Coverage error .....	12
15.3.2. Measurement error .....	13
15.3.3. Non-Response Error.....	13
15.3.4. Processing error .....	13
15.3.5. Model assumption error .....	14
16. Timeliness and punctuality.....	14
16.1. Timeliness .....	14
16.1.1. TP1. Time lag – First results .....	14
16.1.2. TP2. Time lag – Final results.....	14
16.2. Punctuality.....	14
16.2.1. TP3. Punctuality – Punctuality - delivery and publication.....	14
17. Comparability.....	15
17.1. Comparability – Geographical.....	15
17.1.1. CC1. Asymmetry for mirror flow statistics.....	15
17.1.2. General issues of comparability .....	15
17.1.3. Survey Concepts Issues.....	15
17.1.4. Deviations from recommendations .....	15
17.2. Comparability over time .....	15
17.2.1. Length of Comparable Time series.....	15
17.3. Coherence – cross domain.....	15
17.3.1. Coherence – Sub annual and annual statistics.....	15
17.3.2. Coherence with National Accounts.....	15
17.4. Coherence – internal.....	16
18. Cost and Burden.....	16
19. Data Revision .....	16
19.1. Data Revision Policy .....	16
19.2. Data Revision Practice.....	16
19.2.1. Data Revision – Average size.....	16
20. Statistical processing .....	16
20.1. Source Data .....	16
20.1.1. Population and sampling frame .....	17
20.1.2. Sampling design .....	17
20.1.3. Survey size.....	17
20.1.4. Survey technique.....	17
20.2. Frequency of data collection.....	17
20.3. Data Collection.....	17
20.3.1. Type of Survey/Process .....	17
20.3.2. Questionnaire (including explanations).....	17
20.3.3. Survey Participation .....	18
20.3.4. Data Capture .....	18
20.4. Data Validation .....	18
20.5. Data Compilation .....	18
20.5.1. Imputation (for Non-Response or Incomplete Data Sets) .....	19
20.5.2. Grossing and Weighting.....	19
20.6. Adjustment.....	19
20.6.1. Seasonal Adjustment .....	19
21. Comment.....	19



## 2. Introduction

The biennial Business Expenditure on Research and Development (BERD) Survey is a survey of 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. Forfás, now part of the Department of Enterprise, Trade and Employment (DETE), carried out the survey prior to 2007/2008.

## 3. Contact

<b>Contact Organisation:</b>	Central Statistics Office
<b>Contact Organisation Unit:</b>	Business Statistics – Analysis and Results
<b>Contact Name:</b>	Mr. Devin Zibulsky (Analysis and Results), Mr. Tony Zheng (Data Collection)
<b>Contact person function:</b>	Statistician
<b>Contact Mail address:</b>	Central Statistics Office, Skehard Road, Cork, T12 X00E, Ireland
<b>Contact email address:</b>	Business_stats@cso.ie
<b>Contact Phone Number:</b>	+353 21 4535000
<b>Contact Fax Number:</b>	

## 4. Metadata Update

### 4.1. Metadata last certified

25/04/2025

### 4.2. Metadata last posted

04/04/2025

### 4.3. Metadata last update

25/04/2025



## 5. Statistical Presentation

### 5.1. Data Description

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.

### 5.2. Classification System

The results are related to the population of all R&D performing units classified in Sections A to U of the common statistical classification of economic activities as established by Regulation (EC) No 1893/2006 of the European Parliament and of the Council (NACE Revision 2).

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

- The distribution of principal economic activity and by product field are based on Statistical classification of economic activities in the European Community (NACE Rev. 2);
- The local unit for the statistics are compiled at regional level according to NUTS 2 – Nomenclature of Territorial Units for Statistics;
- The distribution by socioeconomic objectives (SEO) are based on Nomenclature for the Analysis and Comparisons of Scientific Programmes and Budgets (NABS);
- The fields of research and development are based on Classification and distribution by Fields of Research and Development (FORD).

The BERD regional data has been produced in accordance with the latest NUTS classifications. An Information Note can be found here: <https://www.cso.ie/en/methods/revnuts23/>

### 5.3. Sector 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. Data from respondents is grossed using the CSO's Central Business Register and aggregated results are then published.

See annex I for further information

### 5.4. Statistical Concepts and definitions

The main concepts and definitions used for the production of R&D statistics are given by OECD (2015), Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities, which is the internationally recognised standard methodology for collecting R&D statistics. Statistical Unit,

See annex I for further information.

### 5.5. Statistical Unit

The statistical unit is the enterprise.

### 5.6. Statistical Population

Statistics on Business enterprise R&D (BERD) measure research and experimental development (R&D) performed in the business enterprise sector, i.e. R&D expenditure and R&D personnel. In line with this objective the target population for the national R&D survey of the business enterprise sector consists of all R&D performing units (including all R&D performers – occasional and continuous, known and unknown - in all branches and size classes) belonging to this sector.



## 5.7. Reference Area

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 in the publication.

## 5.8. Time Coverage

2011-2024

## 5.9. Base period

Not applicable.

## 6. Unit of Measure

Results are generally expressed as monetary amounts, numbers of persons, and number of performing enterprises.

## 7. Reference Period

2023-2024

## 8. Institutional Mandate

### 8.1. Legal Acts and other agreements

#### 8.1.1. European legislation

Since the beginning of 2021, the collection of R&D statistics is based on Commission Regulation (EU) Regulation (EU) No 2019/2152. The technical specifications for this regulation are enumerated in Commission Implementing Regulation (EU) No 2020/1197 of 30 July 2020. These Regulations set the framework for the collection of R&D statistics and specify the main variables of interest and their breakdowns at predefined level of detail. Please note that according to Article 12(4) of Regulation (EU) 2020/1197, the provisions of Regulation (EU) 995/2012 continue to apply for the reference years that fall before 1 January 2021.

#### 8.1.2. National Legislation

The BERD 2023/2024 survey was conducted under the Statistics (Business Expenditure on Research and Development Survey) Order 2024 (S.I. No. 190 of 2024) made under 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.

#### 8.1.3. Standards and manuals

Frascati Manual 2015, Guidelines for Collecting and Reporting Data on Research and Experimental Development

### 8.2. Data Sharing

Not applicable.



## 9. Confidentiality

### 9.1. Confidentiality – policy

All information supplied to the CSO is treated as strictly confidential. The Statistics Act, 1993 sets stringent confidentiality standards: Information collected may be used only for statistical purposes, and no details that might be related to an identifiable person or business undertaking may be divulged to any other government department or body.

These national statistical confidentiality provisions are reinforced by the following EU legislation: Council Regulation (EC) No 223/2009 on European statistics for data collected for EU statistical purposes. Further details are outlined in the CSO's Code of Practice on Statistical Confidentiality.

For more information on the CSO confidentiality policy please visit:  
<https://www.cso.ie/en/aboutus/lgdp/csodatapolicies/statisticalconfidentiality/>

### 9.2. Confidentiality – data treatment

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.

- <3 responses in a cell,
- 80% dominance of one unit, or
- 90% dominance of two units

## 10. Release Policy

### 10.1. Release Calendar

The date of dissemination of all statistics released by CSO can be found in the Release Calendar published in CSO.ie. This calendar is regularly updated.

### 10.2. Release calendar access

The release calendar can be accessed via the CSO website, [www.cso.ie](http://www.cso.ie), or directly from this link:  
<https://www.cso.ie/en/csolatestnews/releasecalendar/>

### 10.3. User access

In accordance with Principle 6 of the European Statistics Code of Practice all users of CSO statistics have equal access via the CSO website at the same time of 11 am. Any privileged pre-release access to any outside user is limited, controlled and publicised. In the event that leaks occur, pre-release arrangements are revised so as to ensure impartiality.

The CSO recognises that in very limited circumstances a business need for pre-release access may be substantiated. Any form of pre-release access is a privilege and a strict CSO pre-release access policy is adhered to for these special requests. The full pre-release access policy can be accessed at  
<https://www.cso.ie/en/aboutus/lgdp/csodatapolicies/csopolicyonpre-releaseaccess/>

The various results are published nationally in statistical release format as well as on the CSO website ([www.cso.ie](http://www.cso.ie)). Selected extracts from the results are posted on the CSO's data dissemination database, PxStat.





## 11. Frequency of Dissemination

Biennial.

## 12. Accessibility and clarity

### 12.1. News release

Not applicable.

### 12.2. Publications

The electronic releases can be found via this link

<https://www.cso.ie/en/statistics/multisectoral/businessespenditureonresearchanddevelopment/>

### 12.3. On-line database

Tables are also disseminated in CSO's dissemination database PxStat as follows:

- Nationality of Ownership <https://data.cso.ie/product/berdno>
- Size of Enterprise <https://data.cso.ie/product/berdse>
- Sector of Activity <https://data.cso.ie/product/BERDSA> • Region <https://data.cso.ie/product/berdr>
- Region <https://data.cso.ie/product/berdr>

#### 12.3.1. AC 1. Data tables - consultations

Not applicable.

### 12.4. Micro-data Access

Not applicable.

### 12.5. Other

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

Data which is submitted to Eurostat can be viewed at: <http://ec.europa.eu/eurostat/web/science-technologyinnovation/overview>

Metadata, graphs and tables are used to enhance clarity when disseminating the data. Data are presented clearly and concisely while highlighting the main findings.

#### 12.5.1. AC2. Metadata consultations

Not calculated.

### 12.6. Documentation on Methodology

Further information on the methodology used to generate the BERD output can be found in the CSO's methods page directly from this link:

<https://www.cso.ie/en/methods/multisectoral/businessespenditureonresearchanddevelopment>

#### 12.6.1. AC3 – Metadata completeness – rate

Not calculated.



## 12.7. Quality Documentation

For more information on the quality associated with the BERD outputs please visit the CSO's methods page, available directly from this link:

<https://www.cso.ie/en/methods/multisectoral/businessexpenditureonresearchanddevelopment/>

## 13. Quality Management

### 13.1. Quality Assurance

#### Quality Management Framework

The CSO avails of an office wide Quality Management Framework (QMF). This framework allows all CSO processes and outputs to meet the required standard as set out in the European Statistics Code of Practice (ESCOP). The QMF foundations are based on establishing the UNECE's Generic Statistical Business Process Model (GSBPM) as the operating statistical production model to achieve a standardised approach to Quality Management. All and any changes implemented to CSO processes and outputs require adherence to the QMF.

### 13.2. Quality Assessment

Further work to strengthen the sample frame will continue in the coming years. The collection of the BERD survey is now mandatory under a Statutory Instrument.

## 14. Relevance

### 14.1. User Needs

Results are primarily used to inform policy makers with regard to data on research and development indicators.

#### 14.1.1. Main National Users

Users' class1	Description of users	Users' needs
1	Department of Further and Higher Education, Research, Innovation and Science,	Science Strategy goals: data to support evidence-based policy at national level and planning across all sectors of the economy
1	Other government agencies and interest groups	All R&D performance
1	Eurostat, CSO	To meet requirements of Commission Regulation, data used for comparable purposes etc. Indicator Progress
1	Regional Assembly	Regional Data
1	DG Research and Enterprise, OECD	R&D Analysis
2	ICTU and other unions	All R&D performance
2	National Competitiveness Council	R&D Data and analysis for knowledge economy
2	IBEC, ISME, Small business forum	All R&D performance



3	Television, newspapers, media and access to database direct website	Act as a vehicle for the dissemination of survey results to the general public .
4	PhD students and researchers	Utilise the BERD micro-data for more in depth analysis and produce outputs that are subsequently published.

#### 1) Users' class codification

##### 1- **Institutions:**

- European level: Commission (DGs, Secretariat General), Council, European Parliament, ECB, other European agencies etc
- in Member States, at the national or regional level: Ministries of Economy or Finance, Other Ministries (for sectoral comparisons), National Statistical Institutes and other statistical agencies (norms, training, etc.), and
- International organisations: OECD, UN, IMF, ILO, etc.

##### 2- **Social actors:** Employers' associations, trade unions, lobbies, among others, at the European, national or regional level.

##### 3- **Media:** International or regional media – specialized or for the general public – interested both in figures and analyses or comments. The media are the main channels of statistics to the general public.

##### 4- **Researchers and students** (Researchers and students need statistics, analyses, ad hoc services, access to specific data.)

##### 5- **Enterprises or businesses** (Either for their own market analysis, their marketing strategy (large enterprises) or because they offer consultancy services)

##### 6- **Other** (User class defined for national purposes, different from the previous classes. )

#### 14.1.2. Principal External Users

Eurostat

#### 14.2. User Satisfaction

No user satisfaction has been performed. Despite many data requests no complaints have been received.

#### 14.3. Data Completeness

Completeness is assessed via comparison of the data delivered against the requirements of Commission Regulation (EU) No 2019/2152.

##### 14.3.1. Data Completeness rate

Not calculated.

## 15. Accuracy and reliability

### 15.1. Overall accuracy

Accuracy in the statistical sense denotes the closeness of computations or estimates to the exact or true values. Statistics are not equal with the true values because of variability (the statistics change from implementation to implementation of the survey due to random effects) and bias (the average of the possible values of the statistics from implementation to implementation is not equal to the true value due to systematic effects).

Several types of statistical errors occur during the survey process. The following typology of errors has been adopted:



1. **Sampling errors.** These only affect sample surveys. They are due to the fact that only a subset of the population, usually randomly selected, is enumerated.
2. **Non-sampling errors.** Non-sampling errors affect sample surveys and complete enumerations alike and comprise:
  - a. Coverage errors,
  - b. Measurement errors,
  - c. Non response errors and
  - d. Processing errors.

Model assumption errors are treated under the heading of the respective error they are trying to reduce.

The overall level of accuracy is deemed to be good since the level of coverage and response are high. The main sources of error might derive from non-sampling error such as measurement and processing errors.

To minimise the effect of measurement errors built-in edits e.g. check the size of expenditure against the size of the enterprise, etc. have been integrated into the process.

Processing errors at very low risk because of number of checks both manual and electronic that is undertaken.

The survey is based on a census – all known administrative data sources utilised together with good quality Central Business Register to produce population frame of all likely performers of R&D. As such any sampling errors are minimal.

## 15.2. Sampling Error

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.

That part of the difference between a population value and an estimate thereof, derived from a random sample, which is due to the fact that only a subset of the population is enumerated.

### 15.2.1. A1. Sampling error indicator

Not calculated.

## 15.3. Non-sampling Error

Non-sampling errors occur in all phases of a survey. They add to the sampling errors (if present) and contribute to decreasing overall accuracy. It is important to assess their relative weight in the total error and devote appropriate resources for their control and assessment.

### 15.3.1. Coverage error

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.

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.



#### **15.3.1.1. A2. Over coverage rate**

Not calculated.

#### **15.3.1.2. A3. Common units – proportion**

Not applicable.

### **15.3.2. Measurement error**

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

- Questionnaire - we ensure that the questionnaire is clear with definitions provided as appropriate.
- Data collection method – data are collected online
- Interviewer - field interviewers are not used by the survey
- 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

### **15.3.3. Non-Response Error**

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. There are 4 written reminders issued with each iteration of the survey and these are followed up by telephone call when non-response persists.

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. Missing values estimated using averages from other responding enterprises by NACE and employment size
4. Contact enterprise to confirm the details on the survey form

#### **15.3.3.1. Unit non-response rate**

The unweighted response rate for BERD 2023/2024 Survey was approximately 60%.

#### **15.3.3.2. Item non-response rate**

Not calculated.

### **15.3.4. Processing error**

Data Capture Errors are likely to be a low risk as survey data are captured online.

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

1. That all the relevant questions have been answered
2. 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.

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.

#### **15.3.5. Model assumption error**

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.

## **16. Timeliness and punctuality**

### **16.1. Timeliness**

Timeliness and punctuality refer to time and dates, but in a different manner: the timeliness of statistics reflects the length of time between their availability and the event or phenomenon they describe. Punctuality refers to the time lag between the release date of the data and the target date on which they should have been delivered, with reference to dates announced in the official release calendar.

#### **16.1.1. TP1. Time lag – First results**

Time lag between the end of reference period and the release date of the results: Indicator: (Release date of provisional/ first results) - (Date of reference for the data)

Preliminary results are disseminated nationally at T+3 months.

Preliminary results for in respect of Sources of Research & Development Funds and Full Time Equivalent (FTE) research personnel sent to Eurostat at T+10 months

#### **16.1.2. TP2. Time lag – Final results**

Final results are disseminated nationally at T+15 months.

Final data is sent to Eurostat at T+ 18 months.

### **16.2. Punctuality**

Punctuality refers to the time lag between the release date of data and the target date on which they were scheduled for release as announced officially.

The dissemination of data to Eurostat is within the legally defined deadline of T+10 months for provisional data and T+18 months for final data.

At national level the BERD release was disseminated in line with the date indicated in the CSO's advance release calendar.

#### **16.2.1. TP3. Punctuality – Punctuality - delivery and publication**

Punctuality of time schedule of data release = (Actual date of the data release) - (Scheduled date of the data release).

0 days for both Eurostat and national dissemination.



## 17. Comparability

### 17.1. Comparability – Geographical

Not applicable.

#### 17.1.1. CCI. Asymmetry for mirror flow statistics

Not applicable.

#### 17.1.2. General issues of comparability

No divergences from Frascati

#### 17.1.3. Survey Concepts Issues

There are no deviations of key survey concept definitions recommended by the implemented regulation of the Frascati manual. The approach to obtaining full time equivalence data is to ask time use questions in the questionnaire.

#### 17.1.4. Deviations from recommendations

With respect key methodological issues which may affect the international comparability of national R&D statistics it is to be noted that with regards the issue of cooperation with respondents we deviate from the recommendations in that time-use questions are asked for the compilation of full time equivalent. Regarding the treatment of non-response, data is weighted by employment and number of firms.

### 17.2. Comparability over time

BERD Survey 2023/2024 was classified by NACE Rev.2 (as was each survey since 2009/2010).

Prior to this, results were classified by NACE Rev 1.1. This was clearly referenced in both the relevant preliminary release and the final publication.

Regarding the compilation and dissemination of individual key variables there have been no breaks in time series since 2009.

#### 17.2.1. Length of Comparable Time series

15 years.

### 17.3. Coherence – cross domain

In general, there are no directly similar results available for comparison, and there are no directly comparable short-term results.

#### 17.3.1. Coherence – Sub annual and annual statistics

Not applicable.

#### 17.3.2. Coherence with National Accounts

Not applicable.



## 17.4. Coherence – internal

Not applicable.

## 18. Cost and Burden

Estimates of Cost and Burden can be obtained from the Response Burden Barometer

<https://www.cso.ie/en/statistics/enterprisestatistics/responseburdenbarometer/>

Survey specific information is available via CSO's dissemination database PxStat.

<https://data.cso.ie/product/RBB>

## 19. Data Revision

### 19.1. Data Revision Policy

Published statistics are subject to correction and revision for a variety of reasons. The most common reasons include the receipt of additional information (for example, late survey responses) and updated seasonal factors. Occasional revisions also occur as a result of changes to definitions, methodology, classifications and general updating of statistical series.

It is recognised internationally that the existence of a sound revisions policy maintains credibility in official statistics. The CSO General Revisions Policy, which details how revisions should be managed and communicated to users, outlines the three main types of revisions:

- Planned Routine Revisions
- Planned Major Revisions
- Unplanned Revisions.

One reason for unplanned revisions occurring can be when errors are detected after publication. The 'CSO Error Correction Policy – How to deal with Publication Errors' outlines the steps taken when these errors are detected. As required under Principle 6.3 of the European Statistics Code of Practice, errors detected in published statistics are corrected at the earliest possible date and users are informed. An important step in the process is the documentation and analysis of errors that have occurred and their causes. This allows the CSO to take measures preventing similar errors from occurring in the future and uniformity in dealing with them when they do.

The data revision policy that CSO statistics adheres to can be found via the following link:

<https://www.cso.ie/en/methods/quality/treatmentofrevisions/>

### 19.2. Data Revision Practice

Not applicable.

#### 19.2.1. Data Revision – Average size

Not applicable.

## 20. Statistical processing

### 20.1. Source Data

The main data source used for the Business Expenditure on Research & Development is extracted from questionnaires sent to all enterprises which are believed to be actively engaged in R&D activities across all business sectors of the economy. (NACE Rev. classifications used).

The most recent implementation of the survey is for 2023-2024 with actual figures for 2023 and estimates for 2024.





### **20.1.1. Population and sampling frame**

The survey population is made up of all enterprises in 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 72 of the NACE Rev 2 Classification (section 72 is Research and Development) and Community Innovation Survey (CIS)/ Innovation in Irish Enterprises (IIE) data, along with other administrative data sources.

This information is used to create a register of likely research and development performers, and this register was supplemented with additional information from the CSO's Business Register such as sectoral classification, number of persons engaged, etc.

### **20.1.2. Sampling design**

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

### **20.1.3. Survey size**

Approximately 4,500 such enterprises were identified as belonging to the population, and survey forms were issued to these enterprises.

### **20.1.4. Survey technique**

The CSO's Central Business Register (CBR) was used to generate the necessary information to issue the forms / links to web-forms (eQ) to these enterprises.

The survey is completed online. Initially, enterprises receive notification of the online process and how to access the relevant URL for completion of the questionnaire. When the return is submitted to the CSO, a response code is added and the data is loaded onto a CSV file 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 updated information.

## **20.2. Frequency of data collection**

Biennial.

## **20.3. Data Collection**

Forms are completed online by respondents. The output of this process is a CSV formatted file which is then imported into SAS for further checking, processing and analyses.

### **20.3.1. Type of Survey/Process**

Data is collected via web-based electronic questionnaire.

### **20.3.2. Questionnaire (including explanations)**

The Business Expenditure on Research and Development (BERD) Survey 2023/2024 collected information about research and development expenditure, as well as human resources engaged in research and development activities, primarily for the reference year of 2023. Enterprises were also asked for estimates of expected research and development expenditure in 2024. 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 researchers;
- Research and development collaboration.

The survey form can be viewed at:

<https://www.cso.ie/en/methods/surveyforms/businessespenditureonresearchdevelopment/>

### **20.3.3. Survey Participation**

Statutory.

### **20.3.4. Data Capture**

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 reviewed the CBR number of the enterprise is captured and linked to the activity code of the enterprise as per the CBR.

## **20.4. Data Validation**

Edit programs are run on Data Management System 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 errors and for any comment on the form that may explain the failure. A limited number of returns from the previous survey data 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 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.

## **20.5. Data Compilation**

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.



### **20.5.1. Imputation (for Non-Response or Incomplete Data Sets)**

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.

#### **20.5.1.1. A7. Imputation rate**

Not calculated.

### **20.5.2. 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 variables used for grossing: 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.

There are three different data sources used for deriving population totals:

- The CSO's Central Business Register
- The Census of Industrial Production
- The Annual Services Inquiry.

## **20.6. Adjustment**

Not applicable.

### **20.6.1. Seasonal Adjustment**

Not applicable.

## **21. Comment**