Article

Update on transformation of research and development statistics: November 2023

Improvements to the ONS research and development (R&D) statistics. Outlines changes to both the Business Enterprise Research and Development (BERD) and UK Government Expenditure on Research and Development (GovERD) surveys and methods.

Contact:
Beck Keane
RandD@ons.gov.uk
+44 1633 455463

Release date:
6 November 2023

Next release:
To be announced

Table of contents

1 . Transformation of research and development statistics

- For the 2022 reference period, the Business Enterprise Research and Development (BERD) survey has incorporated a new sample design to the uplift method for BERD statistics in 2021.

- The new sample design for BERD has seen the sample size increase to approximately 40,000 UK businesses.

- This year, data collection for BERD used electronic questionnaires, to enhance the experience of the responder and improve data quality, with automatic validation at point of entry.

- To improve the quality of data collection on BERD, questionnaire guidance has been enhanced to help businesses better interpret the questions.

- New questions have been added to the UK Government Expenditure on Research and Development (GovERD) survey to collect regional information on where purchased R&D is performed and questions on R&D expenditure on certain technology types.

- We have commissioned an academic review of international methods used for R&D statistics to better understand methods we can draw on to improve our estimates.

2 . Improvements to the business enterprise research and development sampling methodology

Previous sampling methods

The previous sample design for business enterprise research and development (BERD) statistics was based on sampling from a universe of known research and development (R&D) performers. This universe was compiled and maintained from feeder surveys and comprised approximately 40,000 businesses, who had stated they were R&D performers. This was used as the basis from which a sample of around 4,000 businesses were selected.

Businesses selected would receive either a "long form" (asking detailed breakdown of type of R&D, and more granular information) or "short form", which asks for total spend on R&D but does not capture granular detail. The level of R&D expenditure by each business was used to determine which form type a business would receive. The sample had the following design:

- 3,600 sampled businesses for Great Britain (GB) who received the short form

- the sample covered 58 BERD Product Groups (PG)

- random sample of 0 to 99 and 100 to 399 employment

- census sample of those businesses in the universe with employment of 400 and over

- census sample of 1,400 businesses in Northern Ireland

- a reference list of 400 businesses in GB, in addition to the sample, of large contributors to R&D, who receive the long form

This design resulted in 174 GB sampling strata (cells), of which 116 are randomly sampled, and 58 are census.

Challenges with our sample design

Using this sample design allowed us to target businesses most likely to do research and development (R&D). Despite targeting businesses, there were limitations to the sample design.

The use of feeder surveys (which target large businesses, but are themselves, samples, not censuses) led to undercoverage of small firms carrying out R&D, as our previous methods did not weight these samples to the population in the estimation methods. The feeder surveys captured whether businesses performed R&D at a point in time, which could change between responding to the feeder survey and being included in the BERD survey, sampling for which could take place up to 12 months later.

While the current sample design was appropriate when it was first implemented, the expectation of where R&D is happening has changed, particularly in smaller businesses. This means the universe is no longer reflective of where R&D is happening.

We looked to address this in our 2021 estimates by applying an uplift method to a set of results that represented only the known population of R&D performers. The uplift method was a way of estimating for the small- to medium-sized businesses that were missing from the R&D population at that time.

The need for granular data, and additional analysis has grown over the past decade. The existing stratification of variables of interest across region, industry and employment size band had limited quality, and is difficult to meet user needs and understanding R&D at a more detailed level.

New sampling methods

Measurement of research and development (R&D) is complex, and to better inform the sample design, the Office for National Statistics (ONS) carried out research to better understand different methods used by some countries, as well as best practice for business survey methods.

This research found that the approach for estimating R&D in businesses varies by countries: some countries sampled directly from their business register, or a combination of the business register and administrative tax data sources. Feeder surveys are used in a few countries exclusively, as we have done to date, but others combine the feeder surveys with other data sources to identify R&D performers.

To improve business enterprise research and development (BERD) estimates, we have implemented a new design that samples directly from the Inter-Departmental Business Register (IDBR), which is the main sampling frame used by the ONS for business surveys. This is to ensure we are sampling from a wider range of industries, that will help to identify those businesses not picked up in the feeder surveys under the previous approach. This approach was introduced following an assessment of approaches taken by other countries and is considered survey methodology best practice.

To get more robust measures of R&D, the sample size of the new design is 40,000 UK businesses, making the sample unbiased and internationally comparable. This will support our ability to understand sub-national and product detail available in our statistics. The sample will also contain a higher proportion of businesses selected for the "long form" (one quarter of our GB sample), which captures more detail on types of R&D and where it is carried out.

Following our comparison with administrative data sources, we have approval to use the HM Revenue and Customs (HMRC) R&D Tax Credit data for informing our sample, to include 400 claimants in the survey sample, alongside the 400 highest-performing R&D businesses from our previous sample design (the reference list). The remaining sample will be drawn from the IDBR to ensure it is inclusive of all business types and better sub-national representation.

To help meet our user needs, the stratification of the sample has been based on clear drivers and requirements of the statistics:

- understanding of small businesses and R&D – size bands align to that of the Annual Business Survey (ABS); this will allow us to have a better understanding of R&D in smaller businesses in the UK, one of the drivers behind our improvements, to make sure these are not underestimated in our statistics

- devolved authority – stratifying by devolved administration will allow us to produce accurate estimates at a country level, one of the priorities for the output to support decision making

- industry breakdowns – to ensure we can maintain the quality of the estimates by not spreading the sample too thinly, we will reduce the industrial stratification to the lowest published level, BERD-published product groups (PPG)

3 . Collecting our business enterprise research and development data electronically

Previous collection

Business enterprise research and development (BERD) has traditionally been collected via paper questionnaires. This method of collection can be time consuming and can lead to delays because of businesses having to manually fill out the information and return it. The paper questionnaire can be quite difficult to adapt and change, as our user needs for our statistics change.

New collection

This year the collection of the business enterprise research and development (BERD) statistics has been launched on an electronic questionnaire. The move to electronic question will allow:

- more flexibility to make changes to the collection in the future, such as adding and removing questions

- an enhanced user experience for BERD respondents, with easy navigation and checks to reduce burden

- automatic validation at the point of data entry, to help improve the quality of the data reported

- collection of data in a timelier manner – responses can be instantly sent to the Office for National Statistics (ONS) systems for processing

The use of the electronic questionnaire is part of the ONS commitment to bring more flexibility to our processes and how we work, as well as reducing the respondent burden.

4 . Changes to our questionnaires

Improving the guidance in our Business Enterprise Research and Development questionnaire

Research and development (R&D) is a complex topic, so we have enhanced the guidance in the questionnaire, to support respondents in reporting accurate information. Users highlighted two areas that were important to address:

- the capture of R&D from within less-known R&D industries (such as humanities, social sciences and the arts)

- capturing R&D in new and evolving technologies

Additional guidance has been developed, to be relatable to a wider range of businesses (for instance, businesses in social sciences, humanities and the arts). This guidance has been tested with a range of businesses who do and do not carry out R&D, to make sure they are understandable. These changes have been shared and reviewed with wider experts in the field, including the Department for Science, Innovation and Technology, the British Academy and government users. The questionnaire guidance still aligns to Frascati guidelines and has been complemented with additional prompts to:

- include a wider range of examples of R&D

- include examples of more recent technologies in software development, to prompt businesses to consider a wider range of activities

- make some guidance more prominent for businesses in less-known R&D areas (such as humanities and social sciences)

- explicitly reference R&D funded via venture capital, to ensure businesses capture the full extent of their R&D expenditure

Capturing new information on our Government Expenditure on Research and Development questionnaire

Our Government Expenditure on Research and Development (GovERD) survey is an annual census survey of UK government departments, the devolved administrations, UK Research and Innovation (UKRI), and higher education funding bodies, to produce expenditure estimates of research and development (R&D) performed in and funded by UK government departments and the devolved administrations.

Users highlighted the need for more granular information relating to R&D, specifically the need to understand more about the location of where R&D purchased by the UK government is being performed and the amount of expenditure on R&D by UK government linked to certain technology types.

We previously published experimental estimates of UK public-funded gross capital and non-capital R&D expenditure for FYE 2021, broken down by International Territorial Level 1 (ITL1) UK country or region to help meet these needs and help inform our transformation of the GovERD collection.

From this work, we identified more data can be collected, to help meet these needs. This year (reference year 2022), we have added additional questions to our GovERD questionnaire, to capture this new information as part of our regular data collection.

5 . Understanding international methods for research and development statistics

Reviewing our methods in line with international practices

As part of our engagement, users highlighted the need to understand comparability of methods across different countries, and understand where they are different and why.

We have commissioned an academic review of international methods for research and development (R&D) statistics, to help inform some of our improvements to the sample design for this year.

This work compares different national statistical institutes' (NSIs') methods for producing R&D estimates, and findings on the strengths and limitations of the different approaches to help us better understand comparability.

We will also look to use this to inform our future transformation and improvements to our R&D statistics.

6 . Future developments

Changing our business enterprise research and development systems and methods

As part of the work to improve the research and development (R&D) statistics, we will carry out further developments to support the new sample design and collection. The improvements will include:

- developing new methods – to support the new sample design, we are reviewing our imputation and estimation methods used to produce the business enterprise research and development (BERD) statistics

- developing systems – development of new and more flexible systems to allow more analysis on the increased sample size

- developing a back series – understanding the new estimates, and how these can inform a back series for previous years' estimates

- onward impact to national accounts – progress the work to take the redeveloped BERD statistics on in the national accounts and understand the impacts

Publishing our new estimates

Business enterprise research and development (BERD) has undergone large-scale changes, as outlined in this article. These changes are to improve the estimates we produce and to help better meet user needs. Because of the scale of these changes, we plan to publish estimates for BERD in early 2024, to analyse the data to ensure these are of high quality.

Following the publication of our BERD estimates, we will then analyse and publish our government expenditure on research and development (GovERD) estimates.

Government estimates for financial year ending 2022

With the introduction of the new questions on the Government Expenditure on Research and Development (GovERD) questionnaire, we will be looking to provide estimates for financial year ending (FYE) 2022 to complement our [experimental estimates of UK public-funded gross capital and non-capital research and development (R&D) expenditure](). This will provide an annual data time series of regional public-funded R&D expenditure from FYE 2021 onwards.

This work will look to use the new GovERD collection for 2022 and 2023, alongside the regional estimates for FYE 2021 and additional expenditure data from government departments.

7 . Related links

Options for Transformation of Business Enterprise Research and Development Statistics
Article | Released 22 November 2022
Improvements in the annual Business Enterprise Research and Development release to give better coverage of small businesses undertaking research and development.

The Power of Innovation: How the ONS is transforming R&D statistics
Blog | Released 21 September 2022
How the Office for National Statistics (ONS) is enhancing the research and development (R&D) statistics we collect, which includes improvements to methodology, the questions we ask businesses and the sources for our data.

Comparison of ONS business enterprise research and development statistics with HMRC research and development tax credit statistics
Article | Released 29 September 2022
Guidance to help users interpret two data sources of expenditure on research and development and understand the differences between them. Outlines the impact of interim methodological improvements to how the ONS BERD statistics are compiled.

Business enterprise research and development, UK: 2021
Bulletin | Released 22 November 2022
Spending and numbers employed on research and development by businesses in the UK, including data on sources of funds and regional spread.

Business Enterprise Research and Development Survey QMI
Methodology | Released 20 November 2020
Quality and methodology information for UK business enterprise research and development statistics, detailing the strengths and limitations of the data, methods used, and data uses and users.

8 . Cite this article

Office for National Statistics (ONS), released 6 November 2023, ONS website, article, Update on transformation of research and development statistics: November 2023