Article

Options for Transformation of Business Enterprise Research and Development Statistics

Improvements in the annual Business Enterprise Research and Development release to give better coverage of small businesses undertaking Research and Development.

Contact:
Abbe Williams

+44 1633 456629

Release date:
22 November 2022

Next release:
To be announced

Table of contents

1 . Main points

- We have introduced improvements in the annual Business Enterprise Research and Development (BERD) publication to give better coverage of small businesses undertaking Research and Development (R&D).

- We investigated a range of approaches to improve coverage of small businesses; the final approach was validated using wider data sources and peer reviewed by experts.

- The new figures provide the current best estimate of BERD at the UK level and have been validated against other available data.

- Uncertainty currently exists in the estimates below UK level where there are fewer data for validation; to maintain the quality of the statistics, our BERD publication will feature a reduced number of detailed breakdowns.

- While the existing National Statistics designation for the total BERD figure will remain, given they provide the current best estimate of the level of research and development (R&D) in the UK, to help further convey the uncertainty below the UK level, the Office for National Statistics (ONS) has asked the Office for Statistics Regulation (OSR) to temporarily pause the National Statistics status of the detailed breakdowns until the further redevelopment takes place and there is more certainty on the distribution below the total level.

- These estimates will be further refined as part of a wider redevelopment of all our R&D statistics due to conclude in 2024 (but with some improvements delivered in 2023 when we aim to reintroduce more detailed breakdowns), although we expect this first improvement to be the most substantial change to the levels of business enterprise R&D.

- Upcoming developments of the BERD statistics include improvements to sampling to give improved coverage of small businesses, improved methodological processes, improved systems used to produce the statistics and the format of the survey itself.

2 . Overview

In September, we published an [article summarising work to understand differences between two estimates of business expenditure on research and development](#) (R&D). These two data sources were Office for National Statistics (ONS) Business Enterprise Research and Development (BERD) statistics, and HM Revenue and Customs (HMRC) R&D Tax Credit statistics.

The article concluded that one important difference identified between the two data sources was the coverage of small businesses. The article then outlined improvements being made to the BERD methodology to address this.

This article:

- sets out the international comparison work that highlighted further investigation was sensible and was one of the catalysts for initiating the development work

- provides more detail on the work undertaken to improve the estimates, including other approaches that were considered, and an indication of the range of possible estimates

- details how the final approach was validated

- provides information on the wider redevelopment taking place, which is due to conclude in 2024 (but with some improvements delivered in 2023)

Figures presented in this article are those available when the analysis was undertaken and align with those in our Comparison of ONS business enterprise research and development statistics with HMRC research and development tax credit statistics article. Since then, revisions have been incorporated to the underlying BERD Survey returns back to 2018, as part of our regular survey processing. These revisions are reflected in the BERD and Gross Expenditure on Research and Development (GERD) releases published on 22 November, so estimates presented there differ slightly.

3 . International comparisons

Measurement of research and development (R&D) is complex and making comparisons internationally is challenging because of different economic structures and varied R&D tax relief policies.

Many countries have estimates of business spend in R&D via an R&D tax credit system and a survey to businesses.These two sources fundamentally measure different things and have their own purposes.

The Organisation for Economic Cooperation and Development's (OECD) Frascati manual sets out the definitions that should be used in the collection and use of R&D statistics to allow international comparability between survey estimates. While there is guidance on how countries may want to approach collection, there is flexibility in the methods countries use to collect and produce estimates.

Despite country differences in survey methodology and R&D tax relief policy, research carried out by the OECD (Figure 16) indicates a general pattern across most countries (PDF 2.3MB), that the business R&D survey indicates greater expenditure on R&D than activity used as a basis for tax credits. This is because the two sources measure different things and so using tax credit data alone would not pick up businesses that carry out R&D but do not claim tax credits. As a result, the administrative source would be unlikely to provide the full picture of BERD spending (see Section 4, Options considered to improve Business Enterprise Research and Development (BERD), for more detail on the relationship between the tax data and BERD data in the UK). This pattern was not seen in the previously published UK data, where recently the HMRC's R&D Tax Credit statistics indicated spending substantially above the BERD Survey.

We undertook research to better understand the methods used by some other countries. While information on methods used in different countries was limited, the information found indicated that they sampled directly from their business register or a combination of the register and the administrative tax data source. This suggested that the UK methodology, using feeder surveys to build an R&D population frame from which the sample was then selected, was out of line with other countries. The method we have implemented to address this in the short-term (see Section 4, Options considered to improve Business Enterprise Research and Development BERD) brings us as close to those countries that sample from the business register as possible, while being implemented retrospectively.

Over the coming months, we will further engage internationally to share the lessons learned from this work and build further understanding of international approaches to inform the longer-term redevelopment of R&D statistics in the UK. This may include sampling directly from the business register, using an administrative source to inform our sample, or a combination of the two. This work will bring us even closer to international best practice.

4 . Options considered to improve Business Enterprise Research and Development (BERD)

The best way to produce robust estimates of research and development (R&D), which better cover small businesses would be to collect information directly from more businesses. However, for historic data and for 2021, that was not feasible, as the potential improvements to coverage were identified too close to the launch of the 2021 survey. As such, several alternative approaches, using a wide range of methods, were investigated based on existing information.

These can be grouped into three main options:

- Option 1 – uplifting BERD estimates to the Inter Departmental Business Register (IDBR) independent of the HMRC R&D tax credit statistics (implemented)

- Option 2 – uplifting BERD estimates using the HMRC R&D tax credit statistics (discounted)

- Option 3 – pausing the BERD Survey or leaving BERD estimates unchanged until the wider redevelopment is completed in 2024 (discounted)

Option 3 (pause or leave as is) was discounted given the scale of the under-coverage and the importance of users being aware of this sooner than 2024.

Option 2 (HMRC uplift) was discounted as while theoretically the concepts covered in HMRC R&D tax credit statistics and BERD are broadly the same, there are some important differences, which include:

- error and fraud (E&F) – HMRC R&D tax credit statistics will include some claims for R&D tax relief that are fraudulent or include errors

- overseas coverage – HMRC R&D tax credit statistics include claims for some overseas expenditure on R&D, which would not align with the requirements for national accounts purposes and should not be considered as part of the BERD statistics

- sector coverage – HMRC R&D tax credit statistics cover the Financial and Insurance activities sector, which, while included under the Frascati definition, are largely not covered by BERD, and which the further redevelopment work will look to address

- claims coverage – recent questions on the [Business Insights and Conditions Survey](#) (BICS) indicate that around 30% of all businesses that performed R&D in 2021 do not plan to, or are unsure if they will, claim tax credits - perhaps owing to being unaware of the schemes, ineligible (arts, social sciences, and humanities, and so on) or perhaps having other funding sources

- eligibility – HMRC R&D tax credit statistics are affected by changes in R&D tax credit policy; a fundamental shift in the eligibility for tax credit schemes, such as changing scope to include more or fewer categories of qualifying costs, could increase the amount of qualifying expenditure while the underlying amount of R&D undertaken remains the same so it could skew the figures or require a method amendment

While discounted, indicative estimates based on Option 2 were produced to help validate the implemented approach (Option 1).

The implemented approach was developed by a team of methodology experts and validated by a wider set of internal and external experts not involved in the work. The work was also quality assured by one of the Office for National Statistics (ONS) Economic Expert Fellows.

Option 1 (implemented) detail

The implemented uplift method preserves previously published growths at UK level and preserves nation (England, Scotland, Wales and Northern Ireland) growths as much as possible. All other growths should be treated with caution as they were not preserved and are subject to uncertainty.

To implement this preferred option, a range of approaches was investigated by methodology experts and assessed alongside the R&D production team in ONS to conclude on the best implementation.

- Method 1a – uplift 2018 (because this is the most recent year unaffected by the coronavirus (COVID-19) pandemic) and move growth forward by published UK BERD growth.

This was implemented to improve stability and coherence over time.

- Method 1b – uplift each year independently.

This was discounted because of a drop in response rates in 2019 and 2020 (years affected by the pandemic), which may overstate or understate growth.

- Method 1c – uplift 2018, assuming that businesses from the IDBR who are selected for the filter survey, but fail to respond, perform R&D, and move growth forward by UK BERD growth.

This was discounted because during the pandemic, evidence showed there were important differences between respondents and non-respondents.

- Method 1d – uplift each year independently assuming that non-respondents to the Annual Business Survey (ABS) perform R&D.

The ABS is the largest ONS feeder survey used to identify R&D performers. This approach was discounted because during the pandemic, evidence showed there were important differences between respondents and non-responders.

- Method 1e – uplift using a ratio based on estimated business employment rather than estimated business counts.
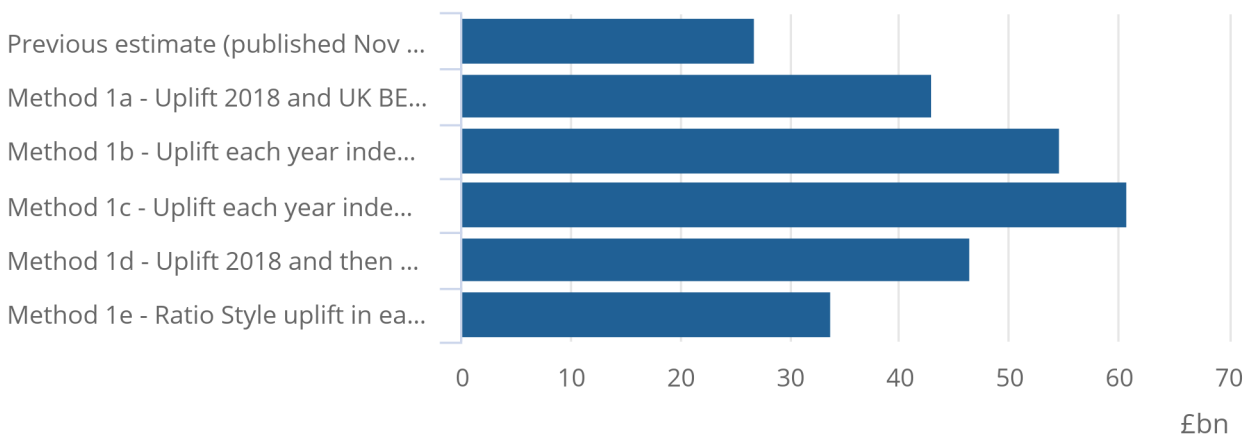
This was discounted because evidence showed limited correlation between R&D and employment.

Indicative estimates from each approach for 2020 can be seen in Figure 1. It illustrates that the subtly different methods result in a range of estimates of UK R&D, although all methods would have resulted in a substantial increase in R&D relative to the current approach. While there were valid reasons to discount the alternative methods, they provide a range of possible estimates for UK R&D.

Method 1a was the preferred method as it was the methodologically most sound approach and minimised artificial volatility resulting from data quality issues during the pandemic. More detail on the uplift approach can be found in the Uplift factors section, with the uplift factors for each of the 68 combinations of product group and size band captured in the BERD statistics.

Figure 1: Outcome of Option 1 uplift approaches, 2020

Figure 1: Outcome of Option 1 uplift approaches, 2020

Source: ONS calculations using unpublished ONS data

Notes:

1. Estimates for Methods 1b-1e are indicative as only the agreed method (1a) was refined to publication standard and do not include any data revisions to 2020 BERD which would be reflected in the 2021 BERD publication. Calculation is based on 2020 ONS data.

5 . Validation of options using HM Revenue and Customs (HMRC) and research and development (R&D) tax statistics

We were able to further inform and validate the options using HMRC R&D tax credit statistics. HMRC and Office for National Statistics (ONS) securely shared the R&D tax credit and Business Enterprise Research and Development (BERD) microdata to allow research and comparisons to take place. As detailed, Option 1, which was implemented did not use the HMRC microdata in the approach.

Understanding of the tax credit scheme and the BERD questionnaire as well as analysis of the microdata confirmed:

- the Frascati definition is used in both HMRC and BERD statistics - other than a few exceptions both sources collect Frascati R&D (exceptions include overseas R&D (in HMRC statistics but not in BERD); R&D in arts, social sciences, and humanities (in BERD but not HMRC statistics))

- matched businesses have similar levels - looking at the microdata for those businesses in both sources, the level of R&D reported was fairly similar, while looking across all businesses in this group in recent years, the HMRC tax credit estimate has been above that from BERD perhaps because of overseas activity and error and fraud

- BERD only businesses - this was a sizeable group of businesses in BERD only but not HMRC; these are likely to be businesses that perform R&D but do not claim tax credits, perhaps because of eligibility, awareness, or alternative funding sources

- HMRC only businesses - in 2019 more than 52,000 businesses were in the HMRC R&D tax credit statistics but not in the BERD universe, including many small to medium enterprise (SME) scheme R&D claimants who had not been sampled by the ONS feeder surveys; this further supported the under-coverage of small businesses in BERD, indicating that the majority should have been estimated for within BERD

Using the learning from the microdata analysis, indicative estimates for Option 2 were calculated at the macro level by using the HMRC statistics to amend the previously published BERD figures, while accounting for the known differences between these data.

The calculation was:

Uplifted BERD equals BERD

plus HMRC only R&D tax credit statistics

minus HMRC R&D tax credit expenditure in finance and insurance activities sector (to give the same sectoral coverage as current BERD)

minus HMRC R&D tax credit expenditure error and fraud

minus HMRC R&D tax credit expenditure overseas R&D

Two methods were considered. Method 2a - assumes that all the error and fraud, and overseas R&D are from "HMRC only businesses" in the HMRC R&D tax credit statistics; this means that the last two lines of the calculation are as stated.
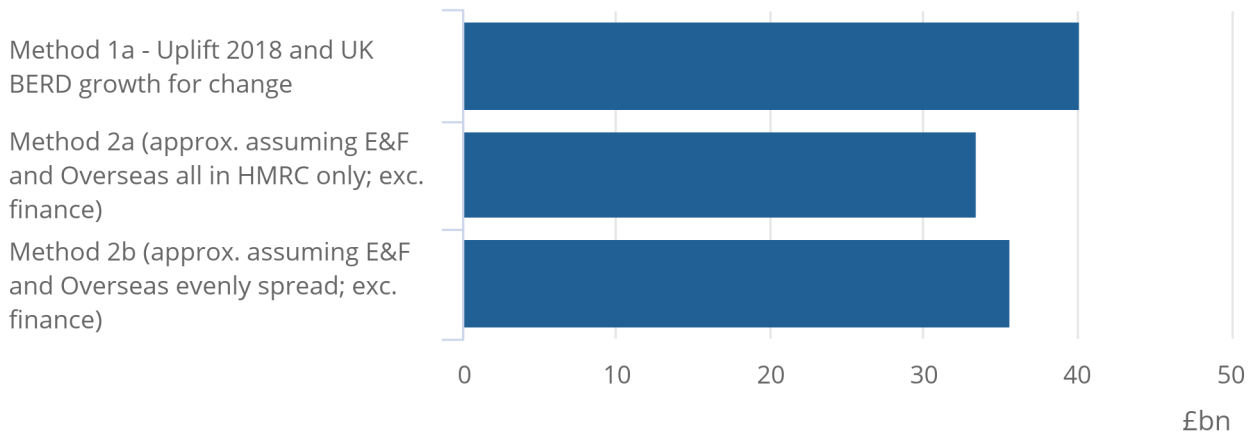
Method 2b - assumes that error and fraud and overseas R&D explain some of the differences seen in businesses that are in both HMRC and BERD, where in recent years the HMRC estimate for these businesses has been higher on aggregate than those in BERD; this means we adjust the last two lines of the calculation to only account for the proportion of error and fraud and overseas R&D contributed by businesses that are only in the HMRC R&D tax credit statistics.

Both methods provide conservative indications of total BERD, as they presume that there is no overlap between error and fraud, and overseas R&D.

Once these adjustments have been made, we would expect the BERD estimate from Method 1a to be above the estimates based on HMRC R&D tax credit statistics. This is because the BERD estimate also includes businesses that perform R&D but do not claim tax credits (30% of all businesses in 2021 either do not plan to, or are unsure if they will claim, as reported in our Business Insights and Conditions Survey results (BICS)). Figure 2 confirms this is the case and gives confidence that estimates from the agreed uplift approach (Method 1a) are of a similar magnitude.

Figure 2: Outcome of various uplift approaches for 2018

Figure 2: Outcome of various uplift approaches for 2018

Method	£bn
Method 1a - Uplift 2018 and UK BERD growth for change	~40
Method 2a (approx. assuming E&F and Overseas all in HMRC only; exc. finance)	~33
Method 2b (approx. assuming E&F and Overseas evenly spread; exc. finance)	~35

£bn

Source: ONS calculations using unpublished HMRC and ONS data

Notes:

1. While adjustments have been made to make ONS and HMRC data as comparable as possible (removing the finance sector; removing E&F; removing overseas R&D) differences will remain e.g. BERD should cover those not planning to or unable to claim tax credits. Calculation is based on 2018 ONS data and 2018-19 HMRC data.

2. Figure 2 covers 2018/2018-19 given this was the most comprehensive microdata for both sources available at the time analysis was conducted.

6 . Data quality of preferred option

The new figures provide the current best estimate of Business Enterprise Research and Development (BERD) at the UK level and have been validated against other available data. The evidence to support the preferred approach was strong. However, as with all survey estimates the improved BERD estimates are subject to uncertainty. As the improvements to account for under-coverage of small businesses have required retrospective methods application to already collected data there is a larger degree of uncertainty.

The results of the Option 1 uplift approach provide an indicative range in which the actual value of BERD may sit and the validation against an "HMRC approach" (Option 2) indicates the uplifted UK estimate using method 1a appear to be of the right magnitude. Users are advised that the statistics will be further refined in the future as development work is ongoing to implement further methodological improvements.

Uncertainty in the estimates increases below the UK level where there are also fewer data for validation. As such, to ensure the quality of data provided at lower levels, the 2022 BERD and Gross Expenditure on Research and Development (GERD) publications provide reduced breakdowns, including basic splits by size-band, sub-national and broad product groups.

The existing National Statistics designation for the total business enterprise research and development (BERD) figure will remain as this is the current best estimate of the level of research and development (R&D) in the UK. However, to help convey the uncertainty below the UK level, the Office for National Statistics (ONS) has asked the Office for Statistics Regulation (OSR) to temporarily pause the National Statistics status of the detailed breakdowns until the further redevelopment takes place and there is more certainty on the distribution below the total level. You can read this request in the letter, Darren Morgan to Ed Humpherson: Business Enterprise Research and Development (BERD) statistics, published on the OSR website. The OSR confirmed their agreement to this approach in their letter published on 21 November 2022.

Improved estimates of more of the splits of BERD spend across different breakdowns will be available once the further development work concludes in 2023 and 2024.

7 . Uplift factors

Table 1 provides the uplift factors used to adjust the Business Enterprise Research and Development (BERD) estimates for under-coverage. The uplift factors are the ratio of two estimates: the estimated number of R&D businesses on the Inter Departmental Business Register (IDBR); and the estimated number of R&D businesses on the BERD reference list. As these are estimates, they have variation around them because of sample sizes.

<div align="center">Table 1: Calculated Uplift Factors</div>

Detailed product group	Uplift factor	
	0-99 emp	100+ emp
Agriculture, hunting and forestry; Fishing	13.1	1.0
Extractive Industries	7.0	1.1
Food products and beverages; Tobacco products	3.8	0.9
Textiles, clothing and leather products	4.3	1.3
Pulp, paper and paper products; Printing; Wood and straw products	5.5	1.1
Refined petroleum products and coke oven products	0.6	0.9
Chemicals and chemical products	1.8	1.0
Pharmaceuticals	1.1	1.1
Rubber and plastic products	3.6	1.0
Other non-metallic mineral products	7.2	0.9
Casting of iron and steel	7.0	0.8
Non-ferrous metals	2.1	0.9
Fabricated metal products except machinery and equipment	3.7	0.9
Machinery and equipment	2.5	0.9
Computers and peripheral equipment	1.3	0.9
Electrical equipment	2.6	1.0
Consumer electronics and communication equipment	4.4	0.9
Precision instruments and optical products; photographic equipment	1.7	0.9
Motor vehicles and parts	4.6	1.0
Other transport equipment	0.6	1.0
Shipbuilding	3.5	1.0
Aerospace	1.1	0.9
Other manufactured goods	2.7	0.9
Sewerage, waste management, remediation activities	2.5	1.0
Electricity, gas and water supply	23.4	0.8
Construction	16.0	1.0
Wholesale and retail trade	11.7	1.1
Transport and storage, incl. postal and courier activities	7.0	1.2
Telecommunications	6.0	0.8
Miscellaneous business activities; Technical testing and analysis	13.6	1.1
Computer programming and information service activities	11.5	1.2
Software development	11.7	1.3
Research and development services	1.0	1.0
Public administration	15.3	5.1

As an example, an uplift factor of 9/6 equals 1.5 might be more appropriately presented as (8 to 10)/(5 to 7). As the two estimates are independent, it is equally likely to see a value of 8/7 as 10/5 as 9/6. So, the uplift factor would be presented as 1.14 to 2, rather than 1.5.

To continue the example, for the same uplift factor of 9/6 equals 1.5, with more breakdowns and smaller sample sizes, ranges around the estimates would increase, and even an insignificant increase in range to (7 to 11)/(4 to 8) would mean 7/8 and 11/4 were equally likely, and the uplift factor would be 0.875 to 2.75, rather than 1.5. The uplift factors were calculated for 68 breakdowns of the UK figure, to reflect different under-coverage by industry and employment, however, the sample sizes in each breakdown were small, which means there is uncertainty in the estimates. At the total level there is less uncertainty because of the large effective sample size, however, the increased uncertainty below the total level is the reason a reduced set of breakdowns below the total is being made available.

The implemented uplift method preserves previously published growths at UK level and preserves nation (England, Scotland, Wales and Northern Ireland) growths as much as possible. All other growths should be treated with caution as they were not preserved and are subject to uncertainty. Users should be aware that the November publication follows the standard data revisions policy, so changes to growth rates are possible.

8 . Further developments

The interim uplifted Business Enterprise Research and Development (BERD) figures, along with the inclusion of a new data source to measure research and development (R&D) activity within the higher education sector (called Transparent Approach to Costing) were used in the compilation of the most recent gross expenditure on research development (GERD) release. This marks the end of the first phase of the R&D redevelopment work. The second phase of redevelopment is due to conclude by 2024, with interim improvements throughout 2023 and includes:

- comparability – work with Organisation for Economic Cooperation and Development (OECD) countries to better understand methodological approaches for BERD

- methods – work to improve the coverage of small businesses in BERD at the sampling stage and to review imputation and estimation more widely

- coverage – consideration of expanding BERD estimates to cover the full economy (the main gap is the Finance and Insurance sector)

- detailed breakdowns – reviewing approaches for producing sub-national and product group data (of both business and government R&D)

- questionnaire – reviewing the question guidance on what constitutes R&D to ensure it is understood by businesses and aligns to Frascati and improvements to government R&D survey

- systems – moving the R&D surveys to an online collection and developing new, more flexible processing systems

- onward impact to National Accounts – outlining the work needed to take the second phase redevelopment data on into National Accounts

9 . Related links

[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.

[Gross domestic expenditure on research and development, UK: 2020](#)
Bulletin | Released 22 November 2022
Annual estimates of research and development performed and funded by business enterprise, higher education, government, UK Research & Innovation and private non-profit organisations.

[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.

10 . Cite this article

Office for National Statistics (ONS), released 22 November 2022, ONS website, article, [Options for Transformation of Business Enterprise Research and Development Statistics: 22 November 2022](#)