Analysis of Publications Arising from Research Funded by the Engineering and Physical Sciences Research Council

Prepared by Elsevier's Analytical Services

Executive Summary

Analysis of publications arising from research funded by the Engineering and Physical Sciences Research Council

Commissioned by the Engineering and Physical Sciences Research Council (EPSRC), this report aims to give insight into the output and impact characteristics of publications in core subject areas, in the period between 2010 to 2015, arising from research funded by EPSRC.

Across all subject areas, EPSRC-funded research produced over 53, 000 publications between 2010 and 2015, with a peak count in 2014 of 10,107 publications. The growth in the number of publications between 2010 and 2015 was strongly positive, at 3.6% CAGR. In keeping with the remit of the EPSRC as a funding body, the most prolific subject areas funded by EPSRC are the core Engineering and Physical Sciences subject areas. Materials Science, Chemistry, and Computer Science each accounted for over 10,000 publications in their respective subject areas. Physics & Astronomy publications numbered over 18,000 and Engineering over 14,000. Energy research produced over 3,000 publications and the Mathematics and Chemical Engineering subject areas produced over 8,000 and 5,000 publications respectively. Output in each of the core subject areas was prolific, with overall counts commensurate with subject area publishing trends.

In terms of overall impact, EPSRC-funded research produced publications that are not only well above the world average in terms of Field-weighted Citation Impact (FWCI), but also above the UK and EU28 averages, for the same subject areas. Engineering research produced very high impact publications, with a FWCI of 2.37, versus a world average of 1.0. Energy, with a FWCI of 2.06 and Computer Science at 2.01, were also well above the world average and Chemistry publications produced an FWCI of 1.72.

Across all subject areas, counts of publications included in the world top citation percentiles were well above the world average. Overall, EPSRC-funded publications in the top 1%, 5% and 10% of cited publications, were over-represented in all brackets. The core subject areas, for example, Energy and Engineering, were over-represented by a factor of three in all top citation percentiles. Almost all core subject areas had higher shares of top 1%, versus shares of top 5% and top 10% cited publications, emphasising again that research articles funded by the EPSRC are of the very highest quality. Compared to the UK and EU28, EPSRC-funded output was higher or even, in some cases, much higher than either comparator for the same top citation percentiles.

Overall, 44% of EPSRC-funded publications between 2010 and 2015 were international collaborations. At over 35% in all core subject areas, the level of international co-authorship of publications arising from EPSRC funding demonstrates that researchers funded by EPSRC collaborate internationally to a significant extent; it is also notable that these publications have even higher impact (FWCI), on average, than publications without international co-authors. Corporate collaboration shares ranged from 3% for Mathematics to 7% for Energy, and associated FWCI values were mostly far higher than the world average for all subject areas.

Projects were differentiated into two categories, *Longer Larger*, i.e., projects that last at least 42 months in duration and that receive at least 700,000 GBP in funding, and the self-explanatory *All Others* category. Longer Larger projects were generally more productive than the All Others category; they also produced research with greater impact and their publications were more likely to be included in a top cited percentile. International collaboration shares were almost identical between the two project categories, but Longer Larger projects had a slightly higher share of corporate collaboration.

Key Findings of EPSRC-funded research

HIGH VOLUME OF OUTPUT FROM PROJECTS

Between 2010-2015, EPSRC-funded research produced over 53,000 publications

HIGH FIELD-WEIGHTED CITATION IMPACT

Overall, EPSRC-funded research has higher relative impact (2.06) than UK (1.55), EU 28 (1.23) and World (1.0) output

HIGH RELATIVE SHARE OF WORLD TOP CITED PERCENTILE PUBLICATIONS

Engineering publications perform 4 times higher than world average in top citation percentiles

INCREASING INTERNATIONAL COLLABORATION

44% international collaboration overall, 47% in Materials Science

EFFECTIVE FUNDING DECISIONS

‘Longer, Larger’ grants produce high impact publications with average FWCI of 2.35

Contents

Chapter 1
Research Output, Impact, and Excellence

This section assesses, per subject area, EPSRC-funded research in terms of its publication output, field-weighted citation impact, and expected highly-cited article shares.

1.1 Output and impact by subject area

Scholarly outputs, as measured by the publication of journal articles, reviews, and conference proceedings papers, are a traditional indicator of research intensity. Figure 1.1 presents the number of such publications over the period of 2010-2015, per subject area[1], resulting from research funded by EPSRC (hereafter referred to as EPSRC's output) and the citations received by those publications,. When interpreting these numbers, one should keep in mind that absolute volumes of publication and citations are known to reflect differences in disciplinary practice; certain research fields tend to publish and/or cite more frequently than others. In line with EPSRC's goal of funding research and training in engineering and the physical sciences[2], the chart shows high outputs in those subjects.[3]

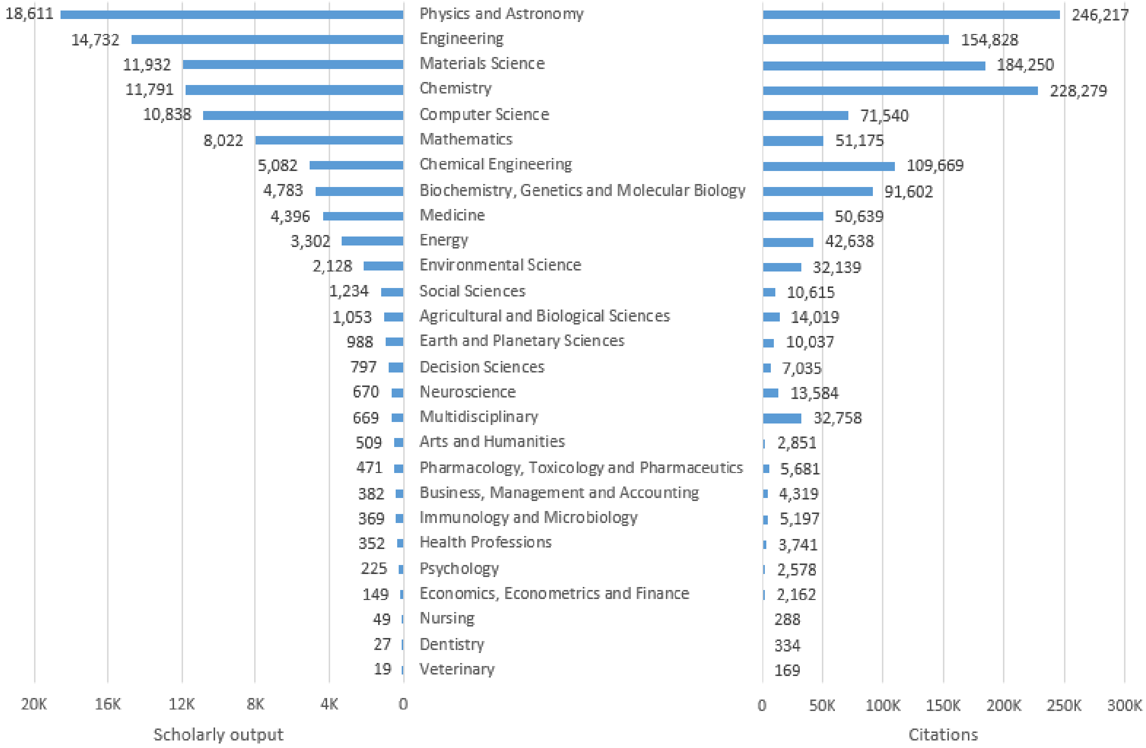


Figure 1.1— Publications by EPSRC-funded authors and citations received by those publications, per subject, 2010-2015.

[1] Publication can be assigned to multiple subject areas. As such, summing publication counts across subject areas will always result in a higher count that the actual total.

[2] Owing to this focus, the remainder of this report will focus on the following subjects (hereafter referred to as Core Subjects): Chemical Engineering, Chemistry, Computer Science, Energy, Engineering, Materials Science, Mathematics, Multidisciplinary, and Physics & Astronomy. Although they are among the top 10 subjects in terms of scholarly output, Biochemistry, Genetics & Molecular Biology and Medicine are excluded from the analyses as they are outside of EPSRC's typical scope. Despite having a much smaller output, Multidisciplinary has been included in the analyses as a special case on account of its high citation rate.

[3] Our analyses make use of full counting rather than fractional counting. For example, if a publication has been published in a journal that is classified as belonging to the subjects Chemistry and Engineering, then that publication counts once towards EPSRC's output in Chemistry and once towards its output in Engineering. Total counts for each subjects are the unique count of publications.

Research in Physics & Astronomy is the most prolific by a considerable margin of over 4,000 publications and, as may be expected, therefore also receives the highest number of citations. However, owing to differences in disciplinary practice, a high output does not always correlate with a high citation rate, as demonstrated here by Engineering: despite producing close to 3,000 more publications than Materials Science or Chemistry in 2010-2015, it receives far fewer citations than the outputs of either of those subjects (± 30,000 to 74,000, respectively). The subject area 'Multidisciplinary' produces relatively few publications, but receives a high number of citations. Aside from the indication that those publications represent high quality research, this is partly because Multidisciplinary consists of publications in multidisciplinary journals, among which are highly prestigious titles such as *Science* and *Nature*.

Citations, however, are a lagging indicator of impact: newly-published articles need to be read, after which they might influence studies that will be carried out, which are then written up in manuscript form, peer-reviewed, published and finally included in a citation index such as Scopus. Only after these steps are completed can citations to the earlier article be systematically counted. Thus, recent publications nearly always have fewer citations than older ones, and might mistakenly give the impression of a decrease in impact. While still useful on a per-subject basis, absolute counts do not compare well across subjects. Putting aside the fact that more publications likely means more citations, different subject fields also have different practices in terms of the frequency at which they cite other publications, as well as the type of publications they publish and cite.

To take these characteristics and behaviours into account, in this report, we use field-weighted citation impact (FWCI) as the main indicator of citation impact. This metric compares the actual number of citations received with the average number of citations for a publication of the same subject, document type, and publication year. It therefore accounts for differences in citation practices between subjects, and is benchmarked against the world average, set at 1.00. For instance, an FWCI of 1.24 means that a publication in this subject is cited 24% more often than expected compared to the world, while a value of 0.90 would mean the publication is cited 10% less than the global average.

Figure 1.2 compares the total 2010-2015 scholarly output in the Core Subjects with their average FWCI over the same period. This clearly shows that while the differences in scholarly output between subjects is significant, they achieve a very similar – and high – FWCI. Four out of the nine subjects are cited more than twice as often as the world average. Even the subject with the lowest FWCI (Chemistry, at 1.72) is still cited 72% more often than the world average. Multidisciplinary clearly stands out for having both the fewest publications and the highest FWCI.

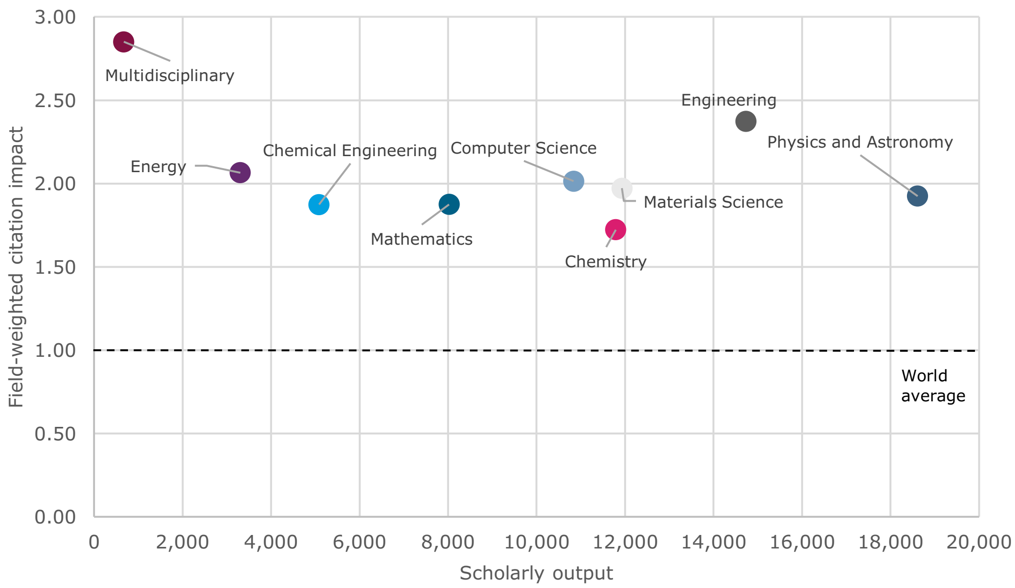


Figure 1.2— Scholarly output and field-weighted citation impact of EPSRC-funded research in Core Subjects, 2010-2015.

Another way of measuring the impact of research is by looking at citation percentiles, i.e., the number of publications that belong to the world's top x% of most cited publications. Figures 1.3 to 1.5 present such numbers for the top 1%, 5% and

10% most cited publications, shown as an Expected Output Index (EOI).[4] This metric compares, per subject area, the percentage of EPSRC's output in a percentile to the world's share of top cited publications in the same subject. For example, if 3% of EPSRC's publications in Mathematics are included in the world's top 1% of cited papers in Mathematics, then it has an EOI of 0.03/0.01 = 3.00 in that subject, the world average being an EOI of 1.00. This allows for an easy comparison of EPSRC's relative performance in the different percentiles.

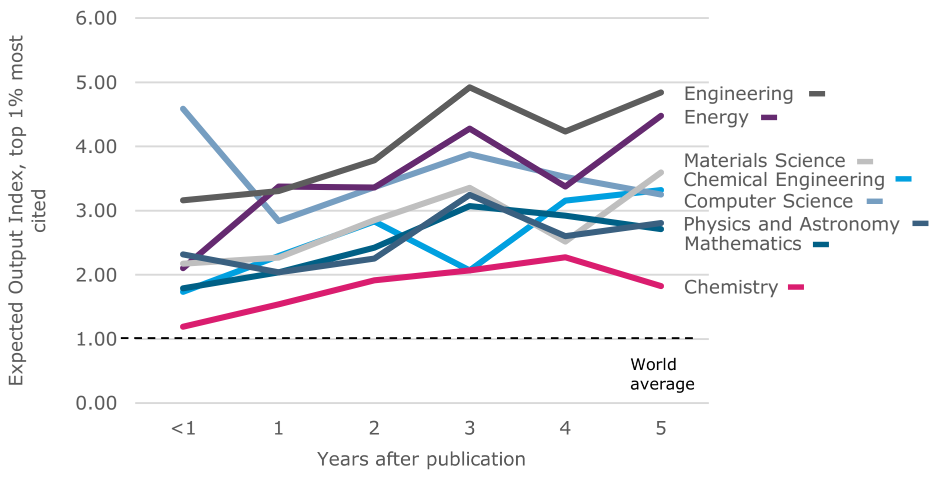


Figure 1.3— Expected Output Index of Core Subjects' output in the top 1% most cited publications, per year, 2010-2015.

For most of the 2010-2015 period, Engineering is the leading subject in the top 1% and 5% cited percentiles, with an EOI of 4.1 and 3.0, respectively. This equates to having an average of 4.1% and 15.1% of its output among these percentiles. Computer Science also does well in the top 1% and 5% percentiles, and surpasses Engineering in 2015 thanks to a steep increase in the most recent year. However, in the top 10% cited percentile, its EOI drops along with that of most other subjects. The reverse is true for Chemical Engineering, which is among the lowest in terms of EOI in the top 1% and 5%, but shows the highest EOI among the Core Subjects for the top 10% most cited publications in 2015.

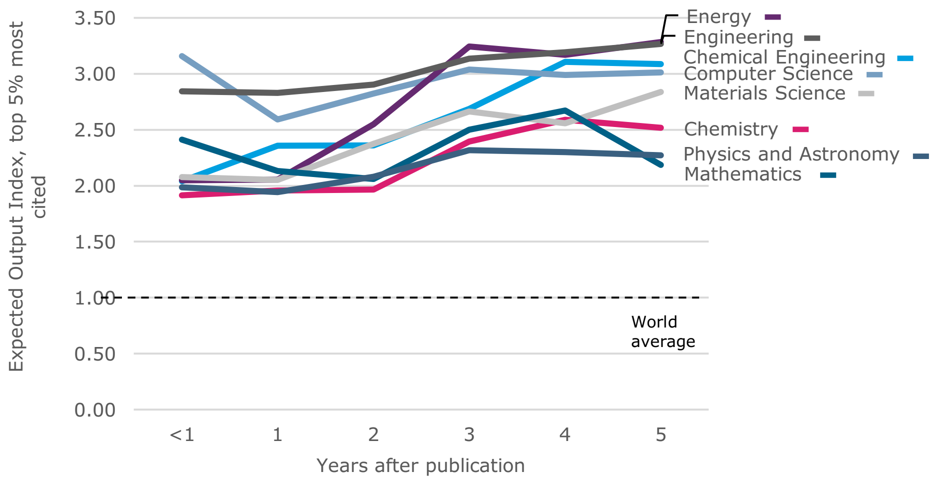


Figure 1.4— Expected Output Index of Core Subjects' output in the top 5% most cited publications, per year, 2010-2015.

[4] Because these are based on citations, a slight decrease in recent years is to be expected. Multidisciplinary has been excluded from these figures, as its EOI in top cited output is 2 to 5 times higher than that of other subjects, skewing the charts. In 2010-2015, 15.4% of its publications belong to the world's top 1% most cited, 38.6% to the top 5%, and 53% to the top 10%.

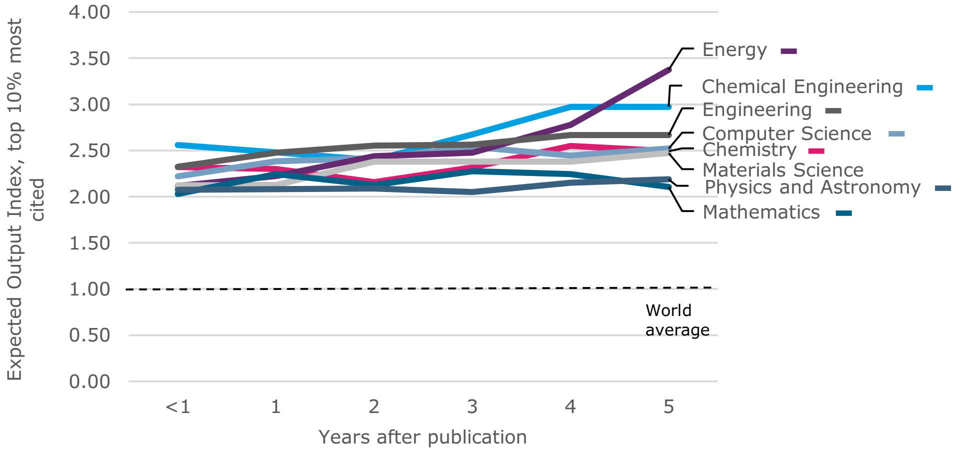


Figure 1.5— Expected Output Index of Core Subjects' output in the top 10% most cited publications, per year, 2010-2015.

Overall, the differences between the subjects' expected output in top percentiles are more pronounced in the top 1% percentile than in the top 5% and 10%. In the former, the 2015 EOIs range from about 1.2 in Chemistry (indicating 20% more publications in this percentile than the world average) to nearly 4.6 in Computer Science. In the latter two, all subjects score between 1.9 to 3.2, and 2.0 to 2.6, respectively.

It is noteworthy that in each percentile analysed, most core subjects show a rising EOI, indicating that the already high impact of publications arising from EPSRC funding actually increases relative to other publications with the passage of time.

To synthesise the previous figures, we provide the EOI for EPSRC's total output for each of the three studied percentiles over the entire period 2010-2015 (Figure 1.6). The most striking result is that EPSRC's share of publications in all top percentiles is at least twice as high as the world average across nearly all percentiles and subjects. Its greatest relative performance is seen in the top 1%, which scores the highest EOI for 6 out of 8 Core Subjects.

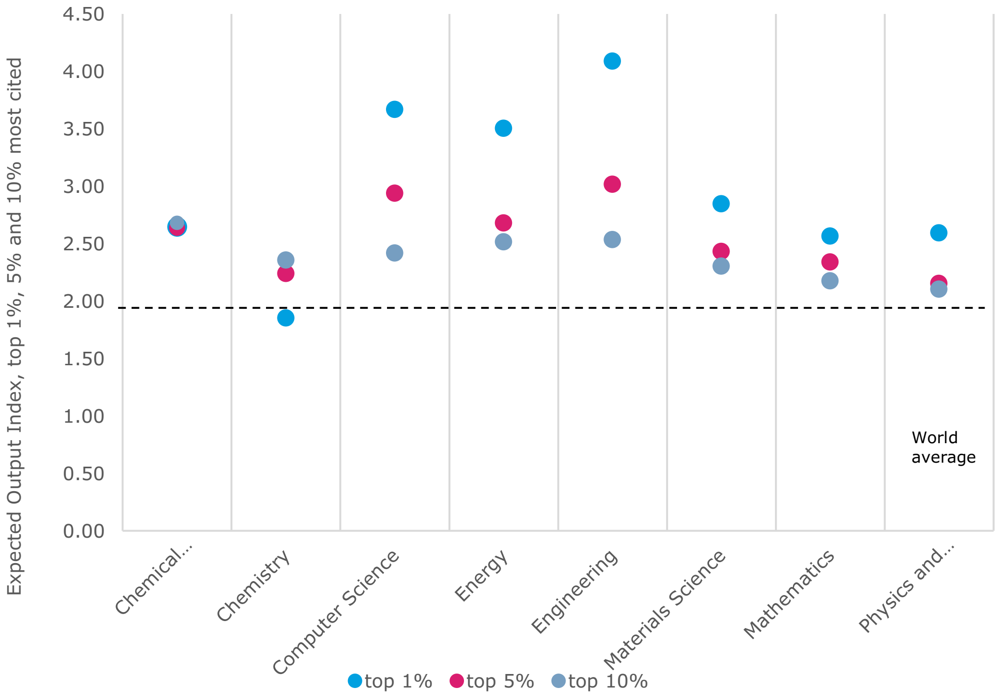


Figure 1.6— Expected Output Index of EPSRC's total output in the top 1%, 5%, and 10% most cited publications in the period 2010-2015.

These findings are testament that EPSRC-funded research is among the most impactful in the world. Especially striking are the subjects Engineering, Computer Science and Energy, in which EPSRC's publications end up being at least 3.5 times more in the top 1% cited percentile than would be expected based on the world average.

1.2 EPSRC and its comparators

This section attempts to provide a different perspective on some of the data in the previous section by contrasting EPSRC-funded research with research in the United Kingdom (UK) and the European Union (EU28).[5] Looking at the research focus of EPSRC and comparators in the selected Core Subjects, there are some small differences in terms of focus. Compared to both the UK and EU28, EPSRC shows a relatively greater proportion of publications in Chemistry, Materials Science, and Physics & Astronomy, as well as Chemical Engineering and Energy. On the other hand, it has a lower output share in Computer Science, Engineering, Mathematics, and Multidisciplinary.

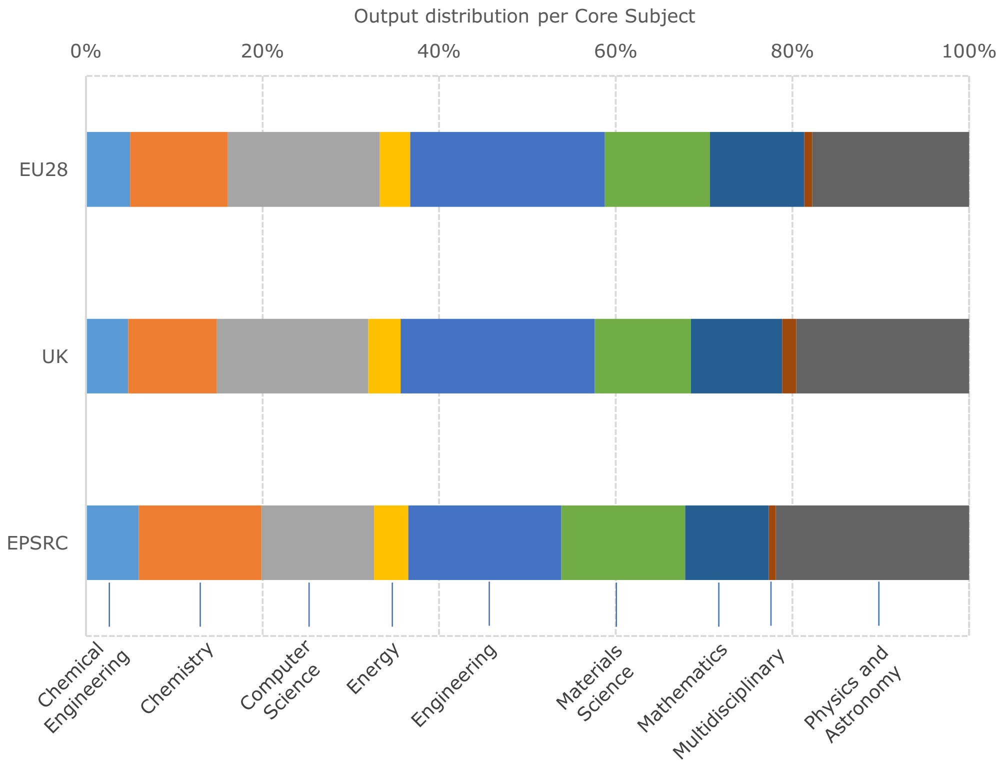


Figure 1.7— Publication shares for EPSRC and comparators, per Core Subject, 2010-2015. Percentages are normalised to make their sum 100%.

In Figure 1.2 we saw that the citation impact of EPSRC-funded research is situated around 2.00 FWCI for all Core Subjects. This means that EPSRC's publications are cited twice as often as the world average in the subjects they belong to. Figure 1.8 shows that such a citation impact is exceptional, also when compared to the UK and EU28. In all of the Core Subjects EPSRC scores well above the citation impact of the UK and, to a greater extent, EU28. The differences are especially striking for the subjects of Engineering (+0.88 index points for EPSRC) and Energy (+0.71 index points).

[5] In general, constituents are not removed from larger entities when making comparative analyses. Therefore, the UK and EU publication data here include EPSRC-funded publications.

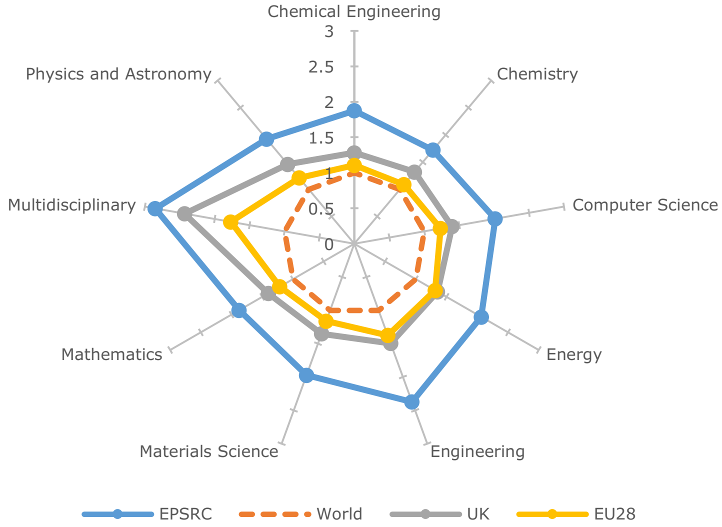


Figure 1.8— Field-weighted citation impact (FWCI) for EPSRC and comparators, per Core Subject, 2010-2015.

When we look at the expected output in top 1% cited percentile (Figure 1.9), the same picture emerges: EPSRC-funded research clearly outperforms that of both comparators for all Core Subjects. As might be expected from Figure 1.6, the subjects that perform most strongly are Engineering, Computer Science, and Energy.

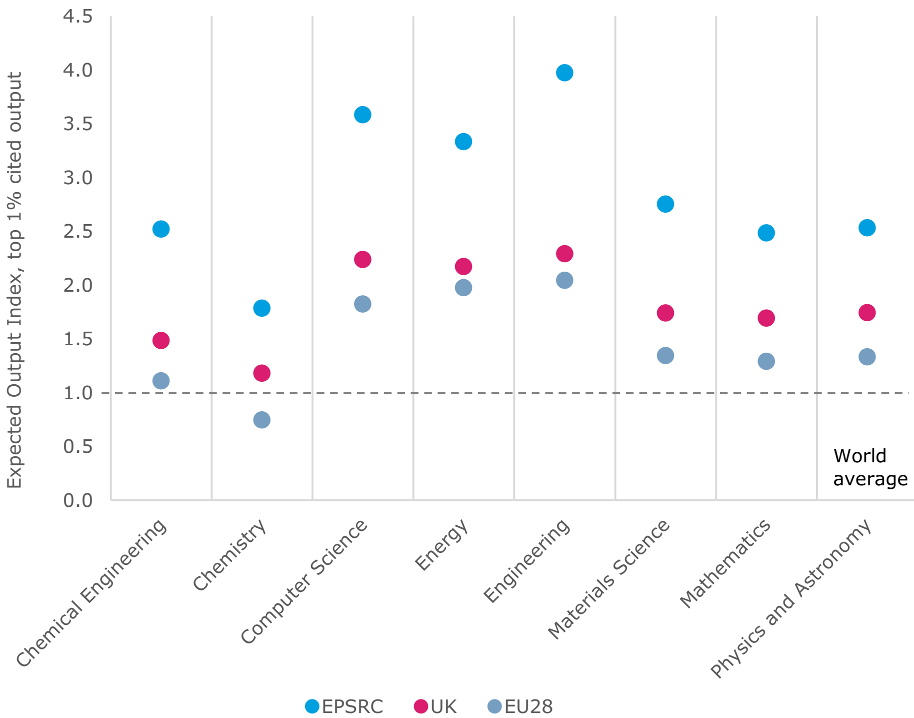


Figure 1.9— Expected Output Index of top 1% cited publications for EPSRC and comparators, per Core Subject, 2010-2015.

Chapter 2
Research Collaboration

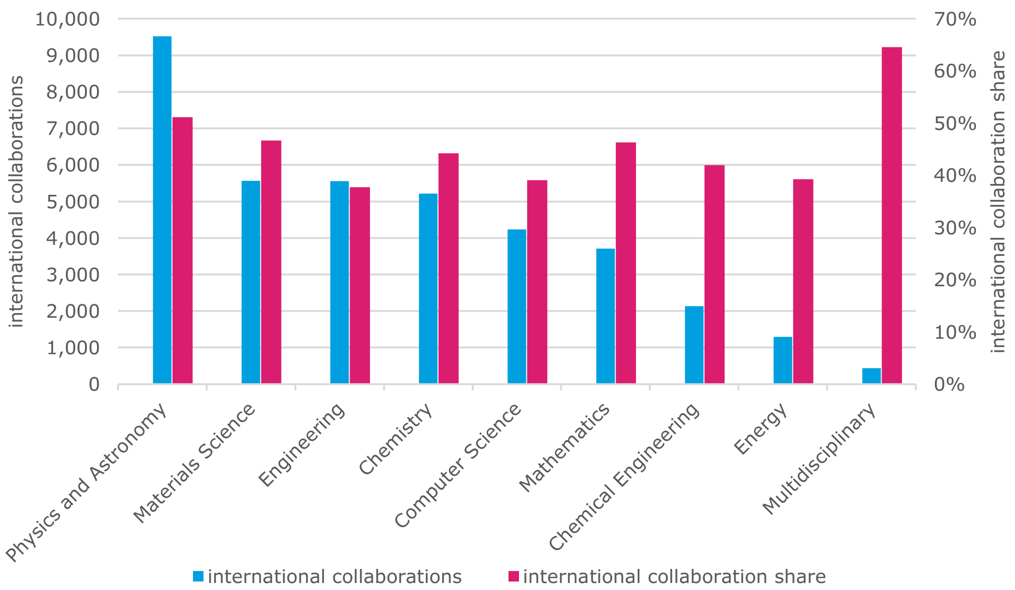
This section assesses EPSRC-funded publications in terms of their collaboration. Specifically, we look at the output and impact of articles that have co-authors from abroad, from the corporate sector, and from different scientific disciplines.

2.1 International collaboration

Research is becoming more and more international. Researchers collaborate every day with others at universities around the world – discussing joint research, organizing conferences, sharing seminar papers. International collaboration provides researchers with the opportunity to work with the best collaborator in a specialized field and gain access to expanded resources, data, and facilities. Single-authored articles are becoming less and less common.

This report measures research collaboration through analysing patterns in the co-authorship of publications. In general, collaboration is defined as the set of publications with at least two co-authors (as opposed to single-authored publications). International collaboration occurs if an article has at least two different countries listed in the authorship byline. If an article has only one author affiliated with institutions in two different countries though, this article is not counted as an internationally-collaborated article but as a single-authored article.

In the period 2010-2015, EPSRC's internationally-collaborated output grew from almost 3,000 articles to nearly 4,300 in 2015 (7.8% CAGR). Looked at as a proportion of its overall scholarly output, EPSRC's international collaborations increased from 40.8% in 2010 to 45.2% five years later. Analysed by the Core Subjects, it appears they are ranked in the same manner they are for their overall scholarly output, with Physics & Astronomy having the largest output and Multidisciplinary the smallest (see Figure 2.1).



+

Figure 2.1— Number and share of publications with an international co-author, per core subject, per period 2010-2015.

Due to a decline in 2014 and 2015, Physics & Astronomy demonstrated the lowest growth for the period, 1381 publications in 2010 to 1581 in 2015 (2.7% CAGR) (see Figure 2.2). However, this decrease in scholarly output in the latest year is not unique to ESPRC's publications. As is the case for Mathematics and Computer Science output, these trends can also be seen for the EU28 and – to a lesser extent – the UK. Ignoring Multidisciplinary because of its relatively small scholarly output, the subjects demonstrating the highest relative growths are Energy (15.2% CAGR) and Engineering (12.1% CAGR).

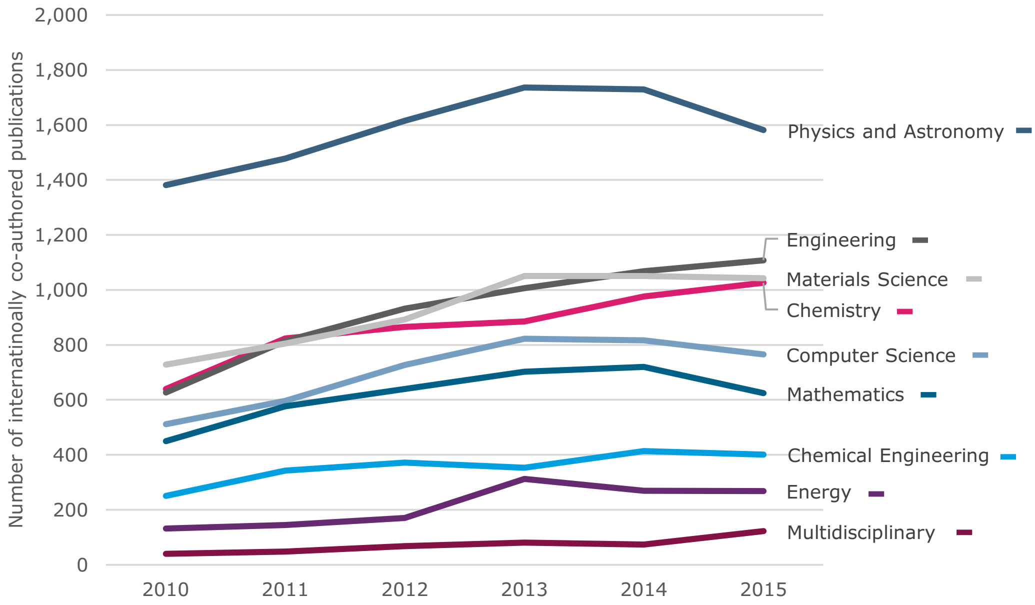


Figure 2.2— Number of publications with an international co-author, per core subject, per year 2010-2015.

A different picture emerges when we look at the percentage of internationally-collaborated papers within a subject and contrast it with its field-weighted citation impact (FWCI). Where we found Multidisciplinary to be the least international subject in terms of absolute output, Figure 2.3 shows it is by far the most international in terms of the proportion of its research that is internationally-collaborated (64.6%) and its citation impact (3.36 FWCI) over 2010-2015. The other subjects occupy places within the interval of 37.7% international collaboration (Engineering) and 51.2% (Physics & Astronomy). Their FWCIs range from 1.89 for Chemistry to 2.94 for Engineering.

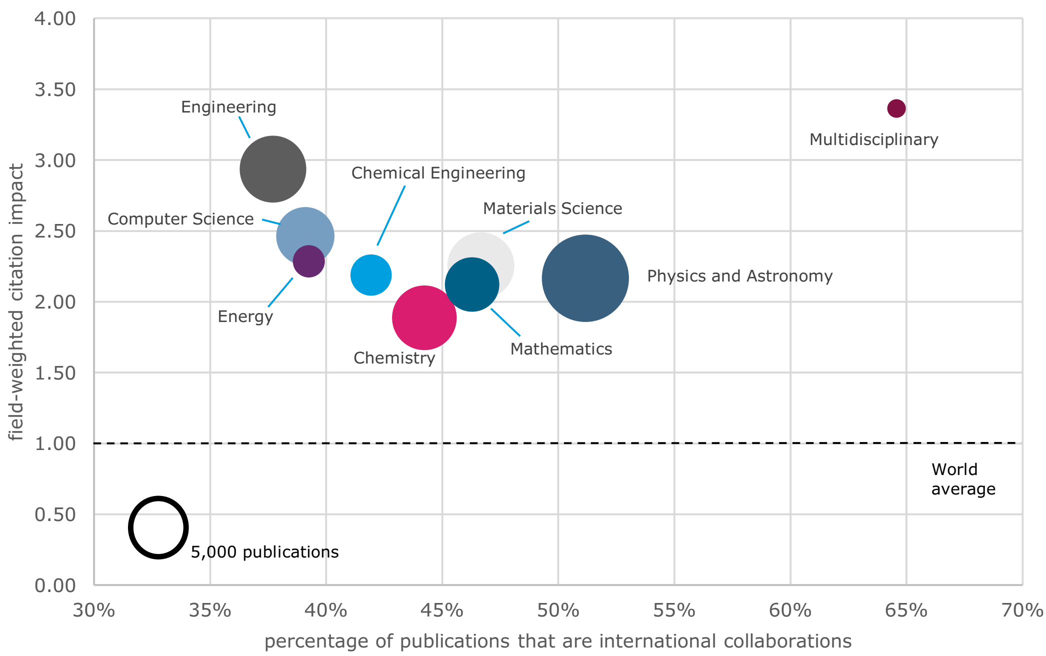


Figure 2.3— Percentage and FWCI of publications with an international co-author, per core subject, 2010-2015. Bubble area is proportional to a subject's volume of publications arising from EPSRC funding.

It is a common finding in bibliometric studies that internationally-collaborated scholarly output is more impactful than other output. This is confirmed in Figure 2.4, where we contrast the FWCI of EPSRC's international collaborations (vertical axis) and the FWCI of its overall output (horizontal axis) per Core Subject over 2010-2015. Since all subjects are plotted on the left side of the diagonal line segment, it is the case for every subject that its international research is more impactful than all its publications taken together. The distance from the diagonal indicates to what extent the international research in that field is more impactful. The uplift is greatest for the subjects Multidisciplinary (3.36 FWCI for international collaborations vs. 2.85 FWCI overall), Engineering (2.94 vs. 2.37), and Computer Science (2.46 vs. 2.01).

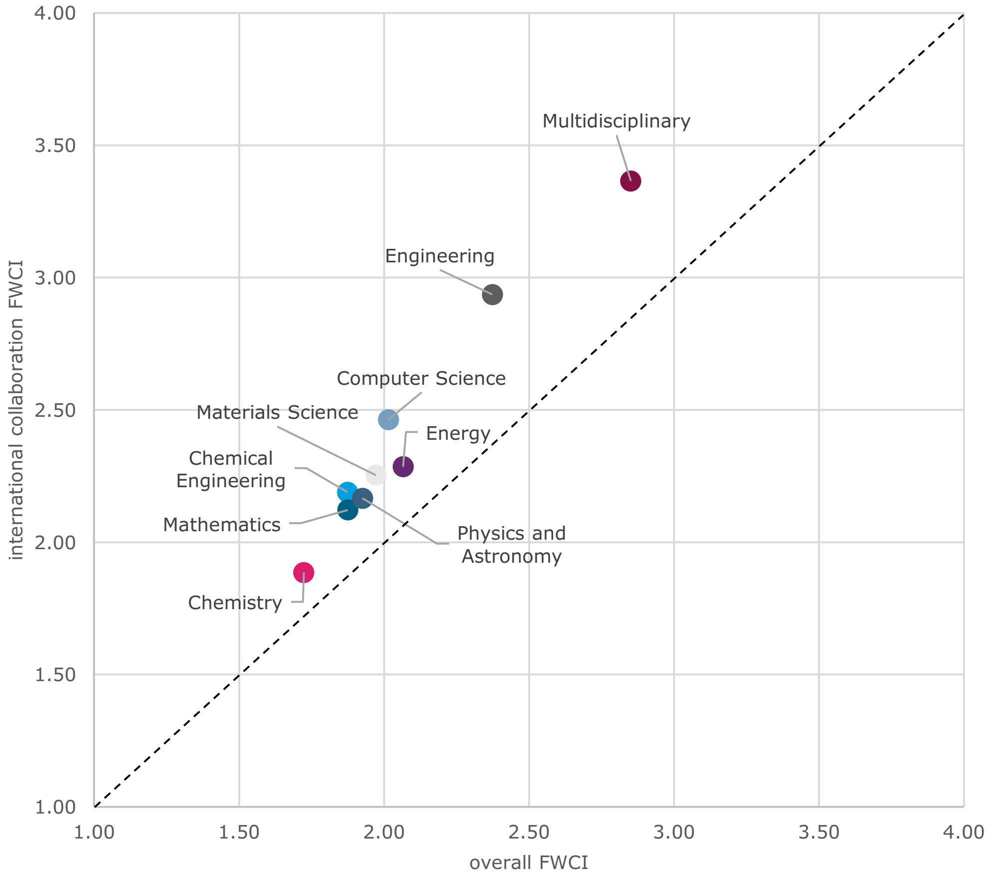


Figure 2.4 — FWCI of publications with an international co-author (vertical axis) and FWCI of all publications, per core subject, 2010-2015.

When we compare the international publication output shares for EPSRC with those of the United Kingdom (UK) and the European Union (EU28) it appears that the UK consistently has a higher percentage of its publications being international collaborations (see Figure 2.5).[6] This may well be partially explained by the significant levels of funding that UK researchers receive from the European Union, which can have associated requirements for international collaboration. EPSRC does score higher than EU28 in every subject area, but it should be noted that for EU28 only collaborations with countries outside of the European Union count as international collaborations, so a lower share is to be expected. The citation impact of the international collaborations corresponds closely to the citation impact we saw for all publications in Figure 1.8 (see Figure 2.6), with EPSRC's publications having a considerably higher impact than those of the UK and EU28.

[6] In general, constituents are not removed from larger entities when making comparative analyses. Therefore, the UK and EU publication data here include EPSRC-funded publications.

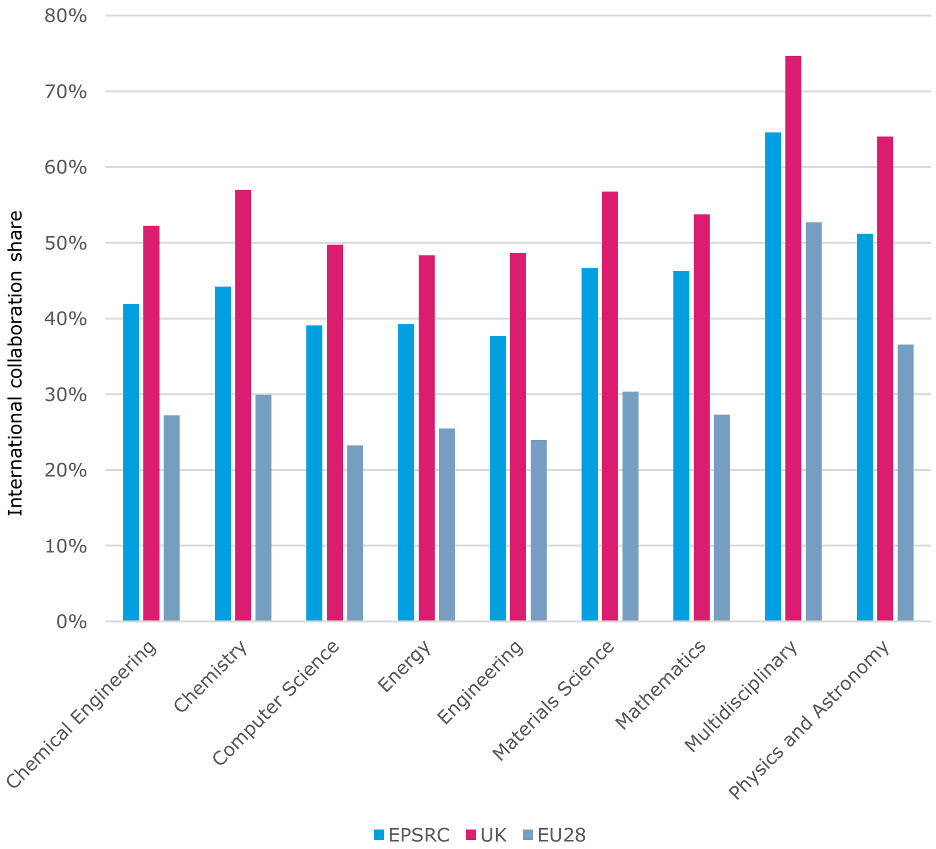


Figure 2.5 — Percentage of publications with an international co-author for EPSRC and comparators, per core subject, 2010-2015.

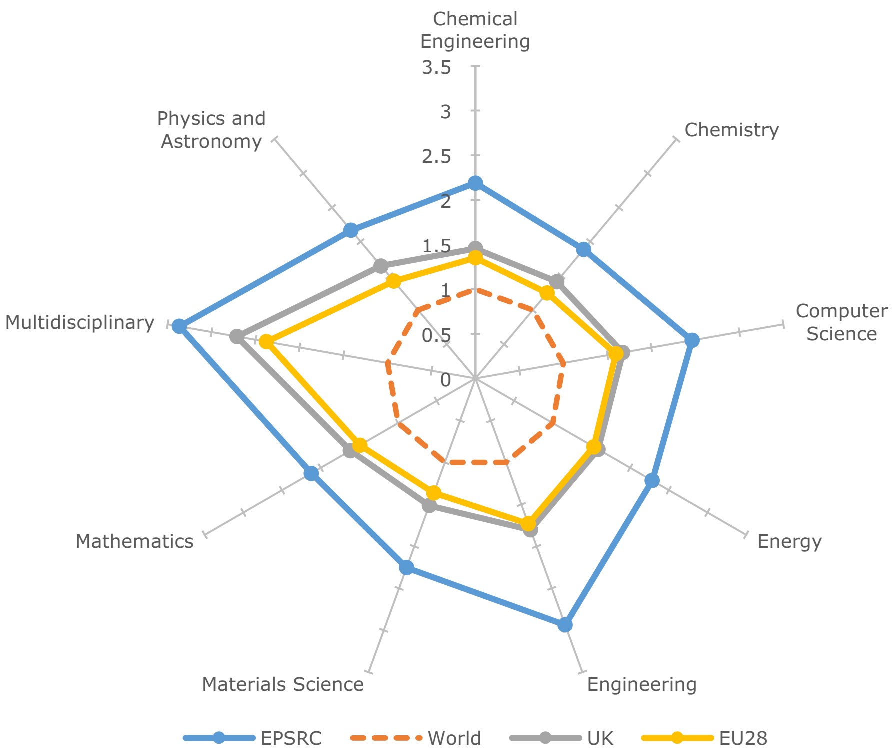


Figure 2.6 — FWCI of publications with an international co-author for EPSRC and comparators, per core subject, 2010-2015.

2.2 Corporate collaboration

In order to assess the relationship between EPSRC's investments in research and the transfer of knowledge to the private sector and its eventual commercialisation, we measure the amount of collaboration with the corporate sector among EPSRC's output. In order to do so, we define as corporate collaborations those articles with at least one author with an affiliation in the corporate sector. The corporate sector is defined to include companies and law firms. Universities, colleges, research institutes, medical schools, hospitals, government agencies, military organisations, policy institutes and NGOs are not considered part of the corporate sector.

Looking at the number and proportion of corporate collaborations in EPSRC's scholarly output over 2010-2015 (see Figure 2.7), the subjects Engineering and Physics & Astronomy rise up as the fields producing the highest number of corporate collaborations at 784 and 764 papers respectively. However, when we correct for the size of the subjects and look at the percentage of research that is corporately-collaborated, we learn that the subjects Energy (6.5%) and Chemical Engineering (5.8%) are, relatively speaking, more connected with the corporate world.

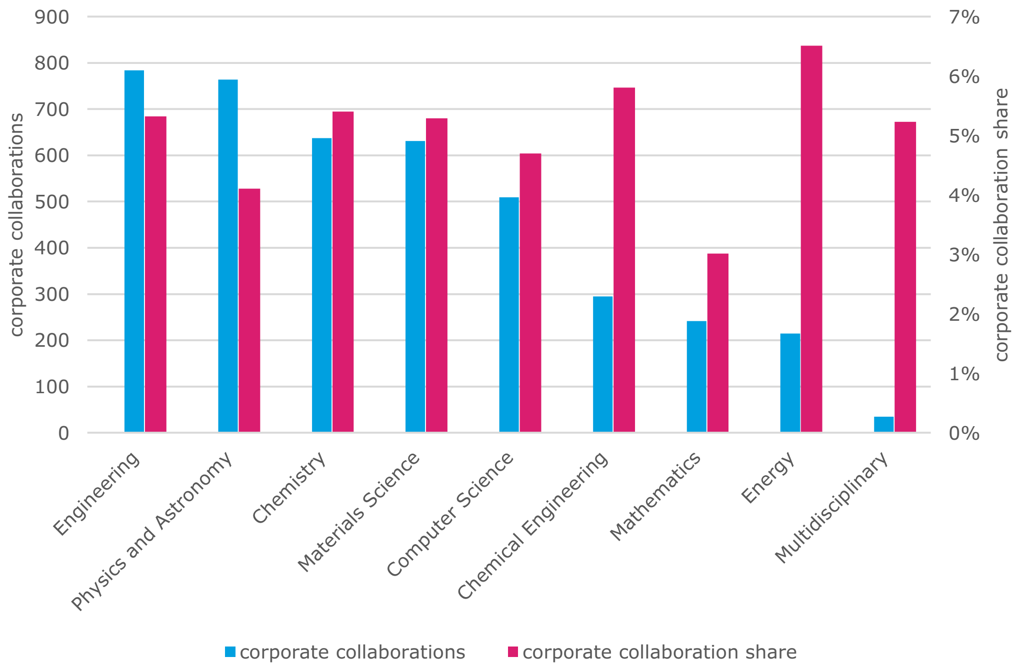


Figure 2.7 — Percentage and FWCI of publications with a corporate co-author, per core subject, 2010-2015.

When we compare the percentage of corporate collaborations within a subject and its field-weighted citation impact, we again see that the FWCI of the relatively small subject Multidisciplinary is the greatest among the Core Subjects at 3.56 (see Figure 2.8). There does not appear to be a correlation between the share of corporate collaborations within a field and the citation impact of those publications. The corporate collaborations in the field with the least corporate co-authorships, Mathematics, have in fact a greater citation impact than those of the field with the most corporate co-authorships, Energy.

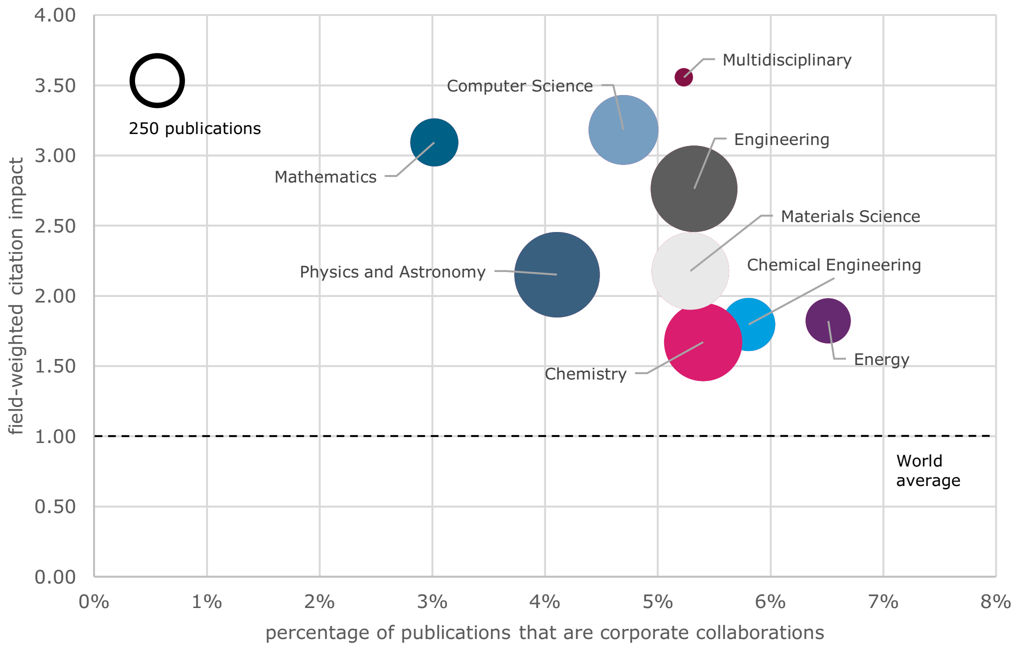


Figure 2.8 — Percentage and FWCI of publications with a corporate co-author, per core subject, 2010-2015. Bubble area is proportional to a subject's scholarly output.

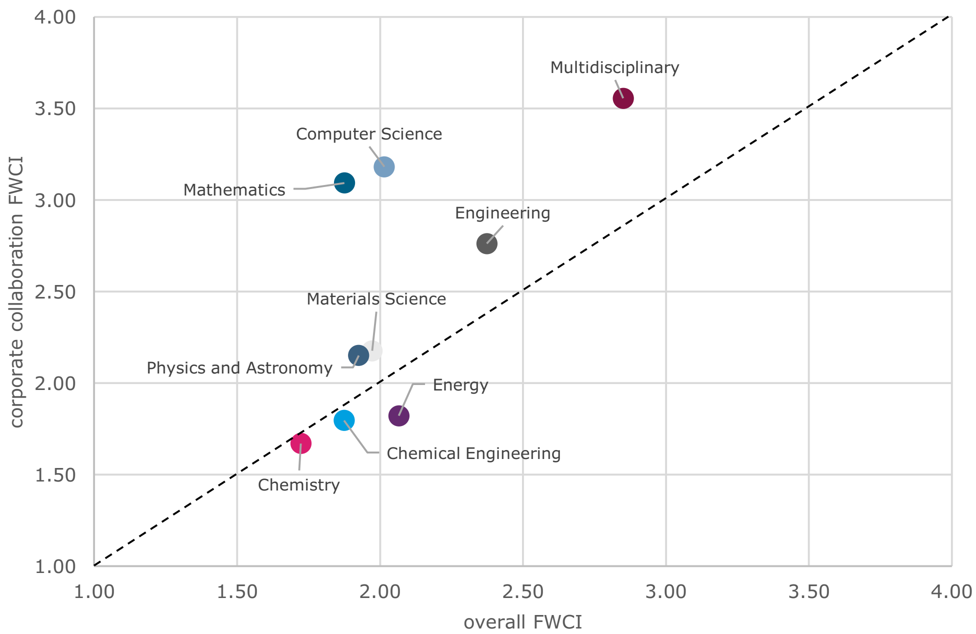


Figure 2.9 — FWCI of publications with a corporate co-author (vertical axis) and FWCI of all publications, per Core Subject, 2010-2015.

Unlike the international collaborations in Figure 2.4, we see in Figure 2.9 that corporate collaborations do not necessarily result in publications that are more impactful compared to the average in a field. When we contrast the FWCI of EPSRC's corporate collaborations (vertical axis) and the FWCI of the overall output (horizontal axis) over 2010-2015 per subject, subjects fall on different sides of the diagonal line segment. For the majority of the subjects corporate collaborations are more impactful – especially Multidisciplinary, Computer Science, and Mathematics. However, this is not the case for the fields Energy, Chemical Engineering, and Chemistry; their corporate collaborations have a lower citation impact than the average publication in the field. This may be due to the field-dependent differences in publication output, and citation behavior, of each of the core subject fields.[8]

When we compare the corporate publication output shares for EPSRC with those of the United Kingdom (UK) and the European Union (EU28), it appears that the UK has a larger corporate share than EPSRC for most Core Subjects (see Figure 2.10). The subjects for which EPSRC ranks first are Multidisciplinary and Physics & Astronomy; EU28 ranks first in Engineering. The citation impact of the corporate collaborations largely corresponds to the citation impact we saw for all publications in Figure 1.8 (see Figure 2.11), with EPSRC's publications having a considerably higher impact than those of the UK and EU28 (Multidisciplinary being the exception).

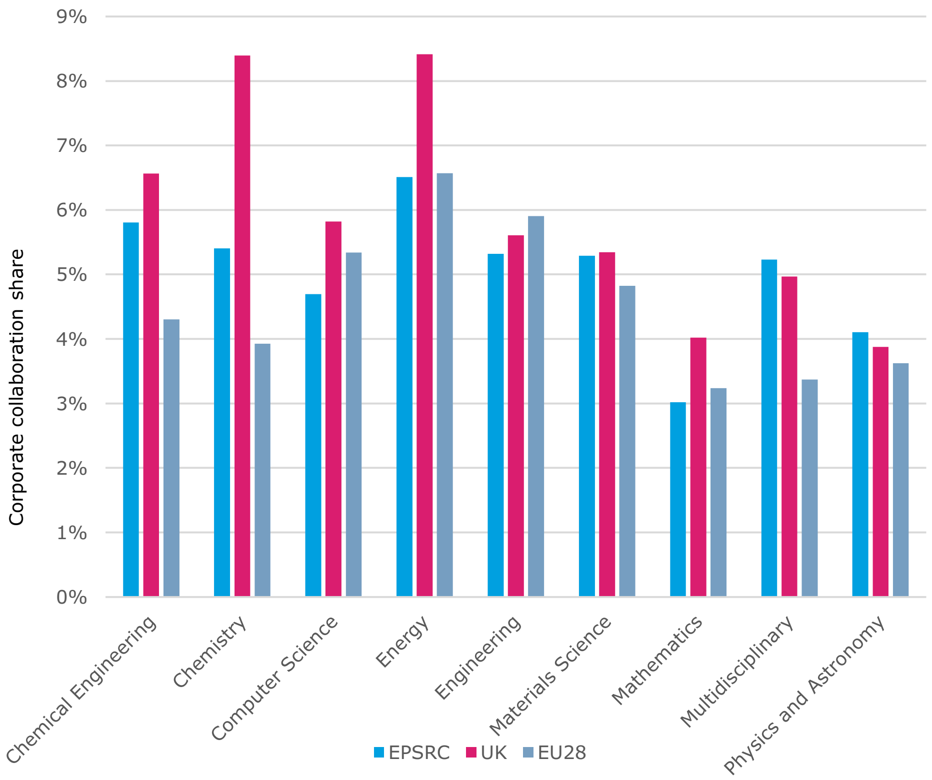


Figure 2.10 — Percentage of publications with a corporate co-author for EPSRC and comparators, per core subject, 2010-2015.

[8] De Moya-Anegon et al. (2014). How to interpret the position of private sector institutions in bibliometric rankings of research institutions. ***Scientometrics***, Vol. 98, pp 283-298.

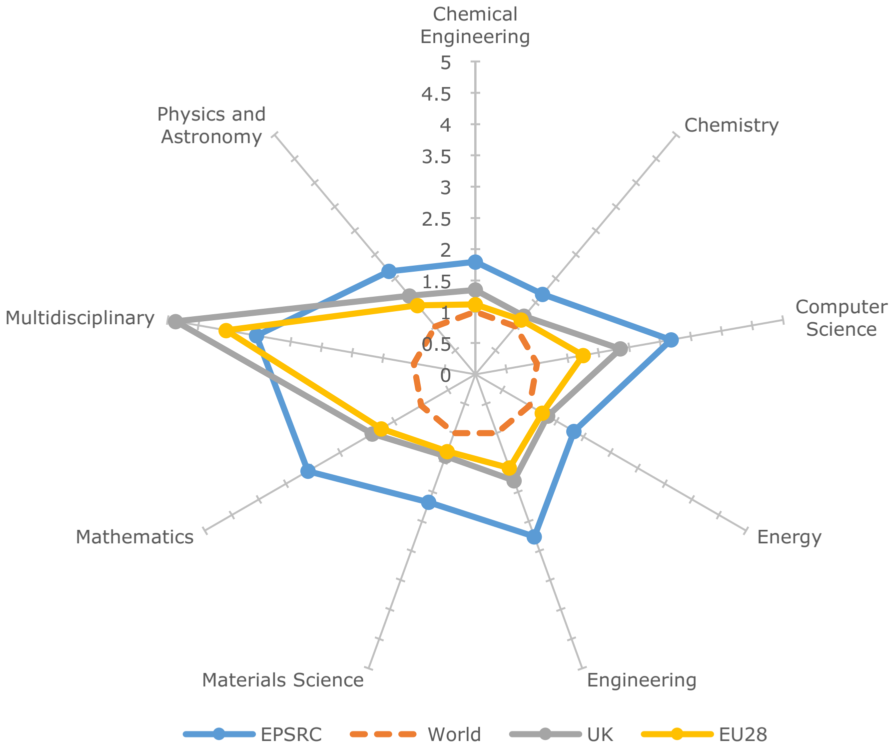


Figure 2.11 — Field-weighted citation impact (FWCI) of publications with a corporate co-author for EPSRC and comparators, per core subject, 2010-2015.

Chapter 3
Analysis of awarded projects

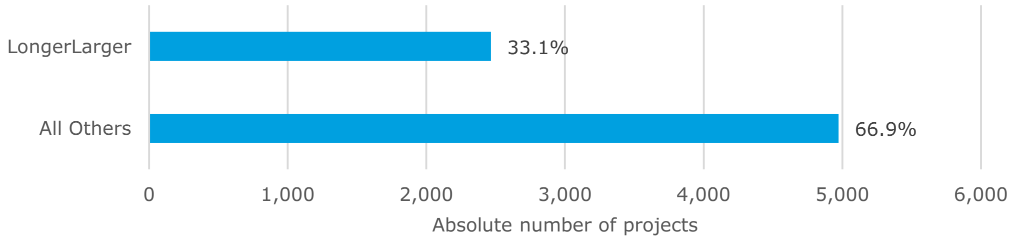
This chapter compares publications arising from typical research projects with those arising from projects of significantly longer duration and higher cost than the EPSRC norm

Project categories

EPSRC is the UK's main agency for funding research in engineering and the physical sciences, investing around 800 million GBP a year in research and postgraduate training. The research projects it funds display a great diversity; they vary from short-term research projects, receiving relatively small amounts of funding, to research programmes that last for many years and receive millions of pounds in funding. This section focuses on the scale (duration and cost) of the projects[9] to which publications in the dataset have been attributed, and the impact (FWCI) of publications arising from these projects. Projects have been categorised into those that are significantly ' longer and larger', i.e., projects that last at least 42 months in duration and that receive at least 700,000 GBP in funding (named "LongerLarger" in the charts below), and the remainder, which constitute the majority of EPSRC-funded projects (named "All Others" in the charts below).

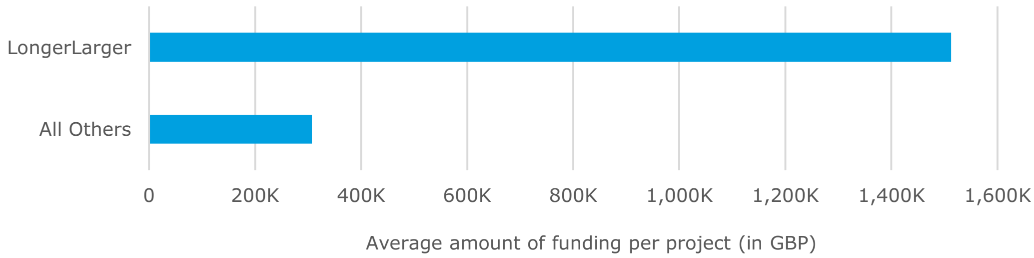
Inputs and impact

Figures 3.1a to 3.1c provide an initial characterisation of the two categories of projects by presenting an overview of the number of projects referenced in the dataset, the average received funding, and the average project length found in each category. Longer-larger projects account for almost exactly one third of those referenced, with 2,466 projects (see Figure 3.1a). The 4,974 projects in the All Others category make up the remaining two-thirds. The average amount of funding per project for longer-larger projects, at over 1.5 million GBP, is nearly five times as much as that for the remaining projects (see Figure 3.1b). And while the average project length for projects in All Others is 33 months, for projects in Longer Larger it is 52 (see Figure 3.1c).



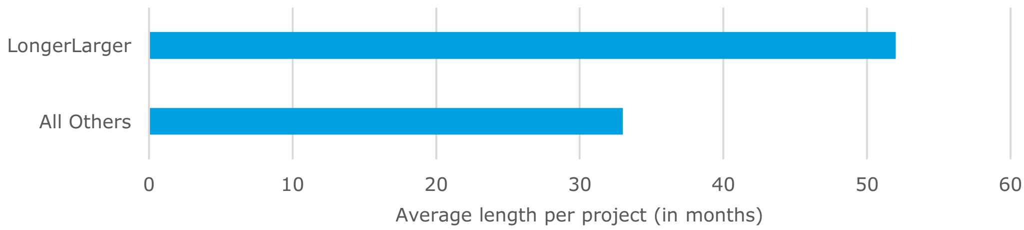
+

Figure 3.1a — Number of awarded projects, per project category.



+

Figure 3.1b — Average values per project (in GBP), per project category.



+

Figure 3.1c — Average duration (in months) of projects, per project category.

[9] Excludes projects funded in reponse to EPSRC calls for 'Equipment Business Cases', 'Grant Balances' and 'Experimental Equipment'.

While the longer-larger projects make up one-third of the projects, they actually contribute 61.1% of the entire publication output. Of the total 53,429 publications linked to the project between 2010 and 2015, the remaining projects are responsible for just over 20,000[10]. By their very nature, longer-larger projects generally yield more publications per project than those in the All Others category - differing by a factor of three.

Figure 3.3 displays the average field-weighted citation impact (FWCI) of the publication outputs of each project category. Both the projects in Longer Larger and All Others generate research that is considerably more impactful than the world average. Longer-larger projects generate publication output that is, on average, 135% more likely to be cited than the world average, whilst the remaining projects produce output that is 110% more likely to be cited than the world average.

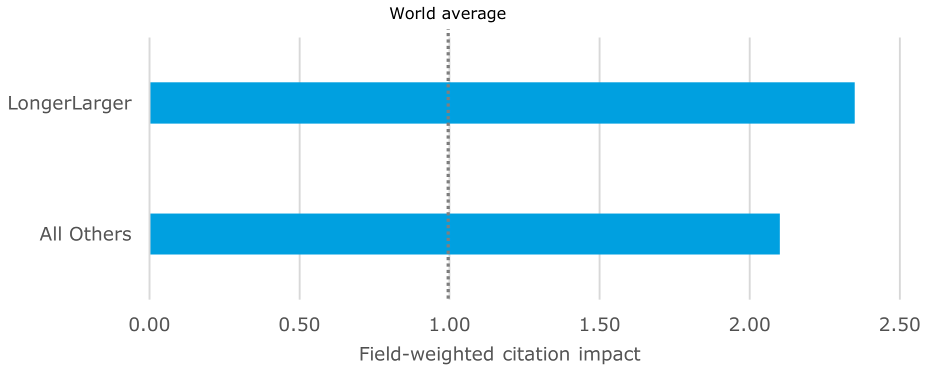


Figure 3.3 —Field-weighted citation impact, per project category.

Figure 3.4 provides an overview of each category's expected output share in the top 1%, 5% and 10% cited publications. Both categories produce more articles in the top citation percentiles than expected compared to the global baseline. However, whilst the primary differentiating factors between the two categories could previously be ascribed to the difference in number of projects, the EOI of longer-larger projects is much higher than those arising from the remaining projects, across all percentiles.

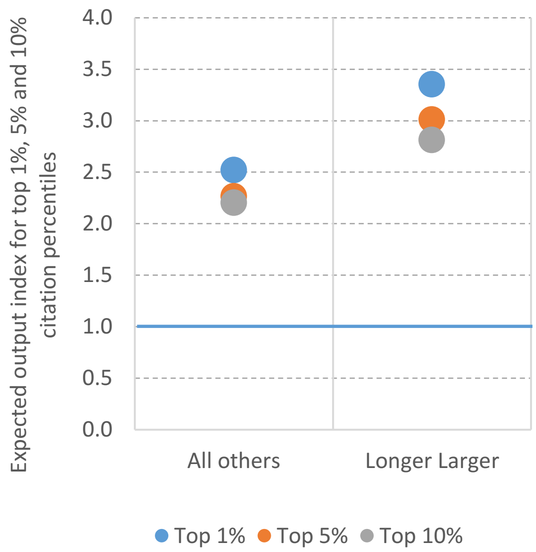


Figure 3.4 — Expected Output Index for top 1%, 5% and 10% most cited publications, per project category, 2010-2015.

[10] It should be noted that publication analyses in Chapters 1 and 2 include publications from all types of projects and calls for the period 2010-2015 (See Appendix for matching statistics). The publications and attributed projects analysed in this chapter are from the period 2010-2015, and have been filtered by length, value and call type.

Collaboration

This chapter examines the international and inter-sectoral collaborative output of each project category. Figure 3.5 displays, for each project category, the number of international collaborations and the proportion of these publications among the project category's publication output. The longer-larger projects produce by far the most internationally co-authored publications (16,460 internationally co-authored publications equating to 7.2 internationally co-authored publications, on average, per project). The remaining EPSRC-funded projects have 10,008 internationally co-authored publications with just over 2, on average, per project. Their international collaboration publication share is relatively close (50.4% for longer-larger projects vs. 48.2% for the remaining projects).

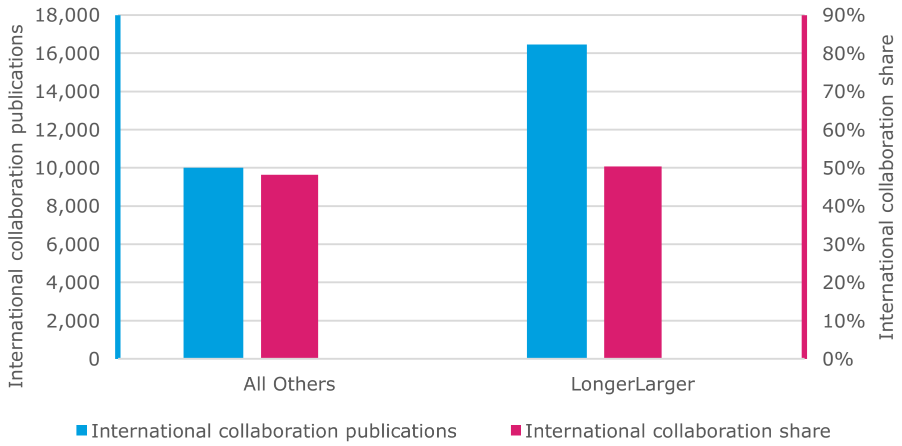


Figure 3.5 — International collaboration publication count (left vertical axis) and international collaboration share (right vertical axis) by project category, 2010-2015.

Figure 3.6 shows for both project categories the collaborations with the corporate sector (left vertical axis) and the proportion of these publications among their total output (right vertical axis). The longer-larger projects produce almost twice the number of corporate collaborations of the remaining projects (1909 publications vs. 1001). Accordingly, longer-larger projects also have a higher average corporate collaborations per projects (0.83) than the remaining projects (0.23). Their corporate collaboration share differs one percentage point, with longer-larger projects at 5.8% and the remaining projects at 4.8%.

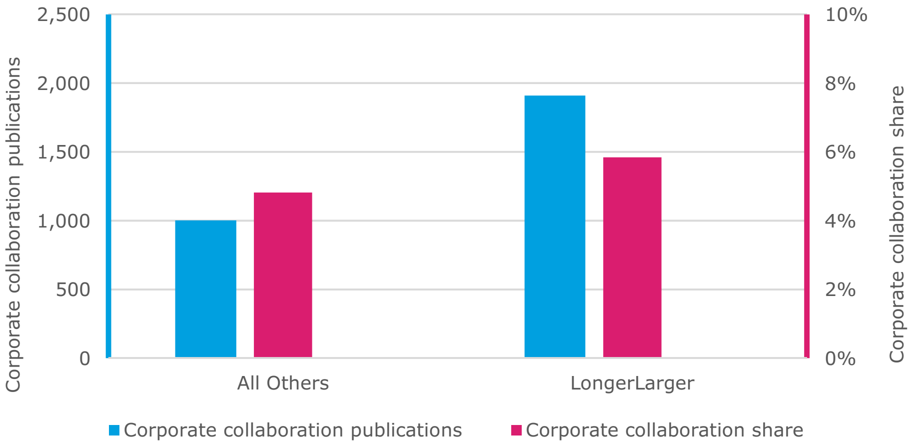
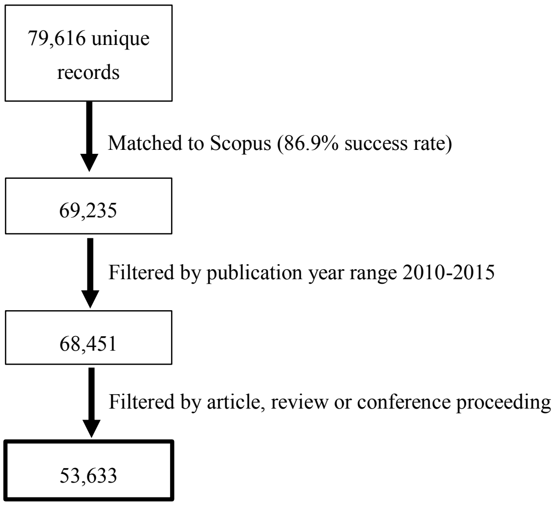


Figure 3.6 — Corporate collaboration publication count (left vertical axis) and corporate collaboration share (right vertical axis) by project category, 2010-2015.

Appendix A
Methodology and data sources

Data

Publication data provided by EPSRC contained 79,616 unique records, published between 1997 and 2016, as confirmed in researchfish® by research grant Principal Investigators. Records were matched by (in order) Scopus Identifier, PubMed ID, DOI, ISSN + Volume + Page, Title + Year, Author + Year + Page, with a 86.9% success rate, resulting in 69,235 documents. For analyses in Chapters 1 and 2, filters were applied to only include those published between 2010 and 2015, resulting in 68,451 documents. These were further filtered to only include either articles, reviews or conference proceedings with the resultant, final document set containing 53,633 documents.



Publication year	Number of documents
2010	7,222
2011	8,239
2012	8,957
2013	9,618
2014	10,107
2015	9,490
Total	**53,633**

For analyses in Chapter 3, filters in project value, length and call type resulted in a total of 53,429 publications (articles, reviews and conference proceedings), from the period 2010-2015. The grants which the publications in the data sets are attributed to may or may not be currently active and inevitably represent a subset of all grants awarded by EPSRC. However, the volumes (of different grant types and of publications) in the dataset are considered sufficiently large to allow robust conclusions to be drawn.

Methodology and Rationale

Our methodology is based on the theoretical principles and best practices developed in the field of quantitative science and technology studies, particularly in science and technology indicators research. The *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems* (Moed, Glänzel and Schmoch, 2004)[11] gives a good overview of this field and is based on the pioneering work of Derek de Solla Price (1978),[12] Eugene Garfield (1979)[13] and Francis Narin (1976)[14] in the USA, and Christopher Freeman, Ben Martin and John Irvine in the UK (1987)[15], and in several European institutions including the Centre for Science and Technology Studies at Leiden University, the Netherlands, and the Library of the Academy of Sciences in Budapest, Hungary.

The analyses of bibliometric data in this report are based on recognized advanced indicators (e.g., the concept of relative citation impact rates). Our base assumption is that such indicators are useful and valid, though imperfect and partial measures, in the sense that their numerical values are determined by research performance and related concepts, but also by other, influencing factors that may cause systematic biases. In the past decade, the field of indicators research has developed a best practices which state how indicator results should be interpreted and which influencing factors should be taken into account. Our methodology builds on these practices.

---

[11] Moed, H. F., Glänzel, W., & Schmoch, U. (Eds.) (2005). *Handbook of Quantitative Science and Technology Research*. Dordrecht: Kluwer Academic Publishers. doi:10.1007/1-4020-2755-9

[12] Price, D. J. de S. (1977). Foreword. In *Essays of an Information Scientist* (pp. v–ix).

[13] Garfield, E. (1979). Is citation analysis a legitimate evaluation tool? *Scientometrics*, 1(4), 359–375. doi:10.1007/BF02019306

[14] Pinski, G., & Narin, F. (1976). Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. *Information Processing & Management*, 12(5), 297–312. doi:10.1016/0306-4573(76)90048-0.

[15] Irvine, J., Martin, B. R., Abraham, J., & Peacock, T. (1987). Assessing basic research: Reappraisal and update of an evaluation of four radio astronomy observatories. *Research Policy*, 16(2-4), 213–227. doi:10.1016/0048-7333(87)90031-X

Year range

All analyses in this report are based on data that range from 2010 to 2015. To measure trends in publication output over time, it is customary to group publications (and other indicators based on publication outputs, such as citations or co-authorships) based on the calendar year in which they were published.

Article types

For all bibliometric analysis, only the following document types are considered:

- Article (ar)
- Review (re)
- Conference Proceeding (cp)

In bibliometric studies, these article types are generally considered to be article types with scholarly content that has been peer-reviewed. That is, such article types have been scrutinized by experts in the same field and were determined by said experts to be suitable for publication. In contrast, our analyses exclude document types such as letters, notes, editorials, etc. that are also published in journals and other serials titles, but are not necessarily peer-reviewed.

Counting

Our analyses make use of full counting rather than fractional counting. For example, if a publication has been published in a journal that is classified as Chemistry and Engineering, then that publication counts once towards EPSRC's output in Chemistry and once towards its output in Engineering. Total counts for each subjects are the unique count of publications.

Field-weighted citation impact (FWCI)

Field-weighted citation impact (FWCI) indicates how the number of citations received by an entity's publications compares with the average number of citations received by all other similar publications in the data universe: how do the citations received by this entity's publications compare with the world average? A field-weighted citation impact of 1.00 indicates that the entity's publications have been cited exactly as would be expected based on the global average for similar publications; the FWCI of "World", or the entire Scopus database, is 1.00. An FWCI of more than 1.00 indicates that the entity's publications have been cited more than would be expected based on the global average for similar publications; for example, 2.11 means 111% more cited than world average. An FWCI of less than 1.00 indicates that the entity's publications have been cited less than would be expected based on the global average for similar publications; for example, 0.87 means 13% less cited than world average.

The Field-Weighted Citation Impact (FWCI) for a set of N publications is defined as:

$$FWCI = \frac{1}{N}\sum_{i=1}^{N}\frac{c_i}{e_i}$$

c_i = citations received by publication i

e_i = expected number of citations received by all similar publications in the publication year plus up to following 5 years.

Citation percentiles

Citation percentiles are a measure of research excellence, as they indicate how many very highly cited publications an institution has produced. Since different institutions may have different subject foci and thus different rates at which their publications accrue citations, this metric is field-weighted, meaning it accounts for the differences in citation rates between subjects, publication years and document types. Also presented in this report is the Expected Output Index (EOI) wherein the actual share of a citation percentile is divided by the citation percentile to produce a multiplier index – the EOI.

Collaboration

A publication is considered collaborative when there is more than one author in the authorship byline in Scopus. When it comes to collaboration across borders, we distinguish between four collaboration types:

- International collaboration, whereby at least one co-author is affiliated with an institution from a different country or region.
- National collaboration, whereby at least one co-author is affiliated with a different institution, but all authors are affiliated with institutions from the same country or region.
- Institutional collaboration, whereby all authors are affiliated with the same institution.
- Single author publications. These are technically not collaborations, and are used as a baseline for comparison of other collaboration types.

These collaboration types are mutually exclusive. For example, if 9 out of 10 authors are affiliated with same institution and the same country, and 1 author is affiliated with an institution from a different country, their publication is considered internationally-collaborative.

Note that due to the rounding of numbers, the sum of each of these categories' share of an institution's publication output may not add up to 100% exactly.

When it comes to collaborations across sectors, we define corporate collaborations as those article with at least one author with an affiliation in the corporate sector. The corporate sector is defined to include companies and law firms. Universities, colleges, research institutes, medical schools, hospitals, government agencies, military organisations, policy institutes and NGOs are not considered part of the corporate sector.

Data Sources

Scopus is Elsevier's abstract and citation database of peer-reviewed literature, covering 62 million documents published in over 22,500 journals, book series and conference proceedings by some 6,000 publishers. Reference lists are captured for nearly 39 million records published from 1996 onwards, and the additional 23.3 million pre-1996 records reach as far back as the publication year 1823.

Scopus coverage is multi-lingual and global: approximately 15% of titles in Scopus are published in languages other than English (or published in both English and another language). In addition, more than half of Scopus content originates from outside North America, representing many countries in Europe, Latin America, Africa and the Asia Pacific region. The database contains titles from more than 120 different countries and over 50 languages in all geographic regions. Scopus covers approximately 18,000 titles from Europe, 10,500 from North-America, 3,600 from Asia-Pacific, 900 from Central and South America, and 1,050 titles from the Middle East and Africa.

Scopus coverage is also inclusive across all major research fields, with 11,700 titles in the Physical Sciences, 12,900 in the Health Sciences, 6,300 in the Life Sciences, and 9,800 in the Social Sciences (the latter including some 3,200 Arts & Humanities related titles). Titles which are covered are predominantly serial publications (journals, trade journals, book series and conference material), but considerable numbers of conference papers are also covered from stand-alone proceedings volumes (a major dissemination mechanism, particularly in the computer sciences). Acknowledging that a great deal of important literature in all fields (but especially in the Social Sciences and Arts & Humanities) is published in books, Scopus has begun to increase book coverage in 2013, and currently covers more than 121,000 books.

More information can be found at
www.elsevier.com/online-tools/scopus.

Scival is Elsevier's ready-to-use solution to visualize research performance, benchmark relative to peers, develop collaborative partnerships and analyse research trends. It offers quick, easy access to the research performance of 7,500 research institutions and 220 nations worldwide. It uses advanced data analytics super-computer technology, to instantly process an enormous amount of data to generate powerful data visualizations on-demand, in seconds.

More information can be found at
https://www.elsevier.com/solutions/scival

Authors

This study was commissioned and funded by the Engineering and Physical Sciences Research Council (EPSRC), and conducted and written by Thomas Gurney, Steven Scheerooren, and Anton Jumelet. We also acknowledge the help of Eleonora Palmaro (Elsevier), Ben Ryan (EPSRC), Sue Smart (EPSRC) (in alphabetical order).

THOMAS GURNEY

Thomas Gurney is Analytical Services Product Manager with Elsevier. As a scientometrician he specialises in adapting bibliometric and scientometric analytics for use by corporate R&D. Thomas holds a PhD from VU Amsterdam in knowledge transfer and dynamics, an MSc in Science and Technology Studies from the University of Amsterdam and BScs in Chemistry, and Oceanography and Climatology from the University of Cape Town. He has published on scientometrics, bibliometrics, alt-metrics, science parks, large-scale research infrastructures and academic entrepreneurship.

STEVEN SCHEEROOREN

Steven Scheerooren is a Content Analytics Support Manager in Elsevier's Analytical Services team. Having graduated from the University of Leiden as a apanologist (BA), Steven joined Analytics in 2013 to work on staff evaluation projects for Japanese universities. In his current position he conducts comparative research performance analyses for academic and government institutions.

ANTON JUMELET

Anton Jumelet is a Content Analytics Support Manager in Elsevier's Analytical Services Team. He received his BA in Philosophy from the University of Groningen in the Netherlands. Large research projects he has been involved in include the assessment of the research bases of India and the UK, and an analysis of the research performance of cities around the world.

About

This report was prepared by Elsevier. Its Research Intelligence portfolio of products and services serves research institutions, government agencies, and funders. Whether your institution is conducting research or funding it, Elsevier's Research Intelligence provides the objective and analytical insight needed to improve your ability to establish, execute, and evaluate national and institutional research strategy.

For more information about Research Intelligence, please visit:
elsevier.com/research-intelligence

Notes