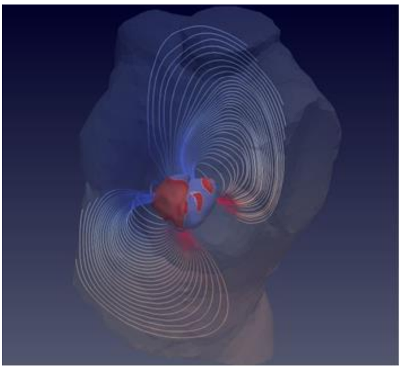
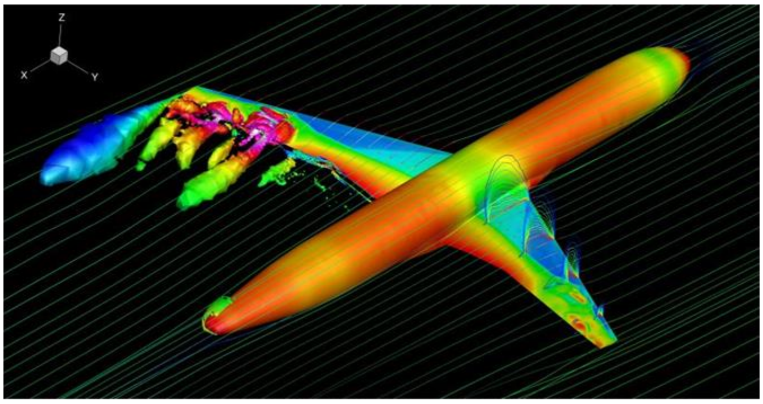
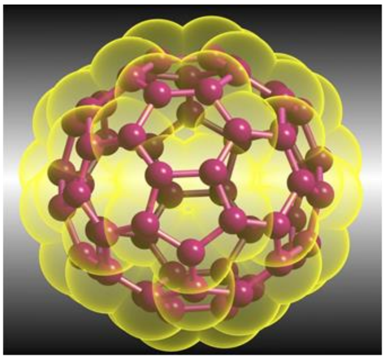
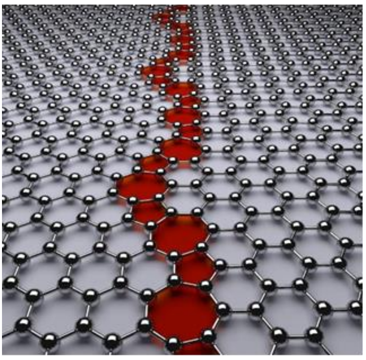
The impact of EPSRC's investments in High Performance Computing infrastructure

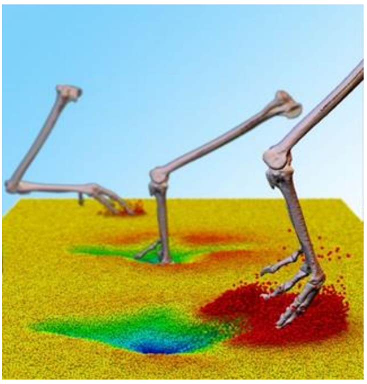
Final report

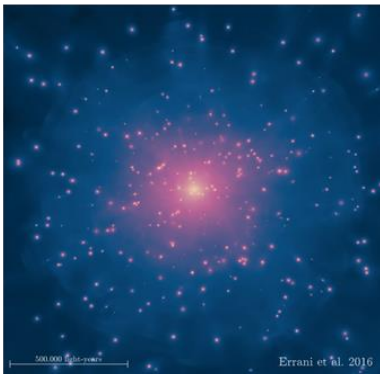




EPSRC

Engineering and Physical Sciences Research Council







LE London Economics

November 2019

About London Economics

London Economics is one of Europe's leading specialist economics and policy consultancies. Based in London and with offices and associate offices in five other European capitals, we advise an international client base throughout Europe and beyond on economic and financial analysis, litigation support, policy development and evaluation, business strategy, and regulatory and competition policy.

Our consultants are highly-qualified economists who apply a wide range of analytical tools to tackle complex problems across the business and policy spheres. Our approach combines the use of economic theory and sophisticated quantitative methods, including the latest insights from behavioural economics, with practical know-how ranging from commonly used market research tools to advanced experimental methods at the frontier of applied social science.

We are committed to providing customer service to world-class standards and take pride in our clients' success. For more information, please visit www.londoneconomics.co.uk.

Head Office: Somerset House, New Wing, Strand, London, WC2R 1LA, United Kingdom.

w: londoneconomics.co.uk e: info@londoneconomics.co.uk : @LondonEconomics
t: +44 (0)20 3701 7700 f: +44 (0)20 3701 7701

Acknowledgements

We would like to thank all individuals and organisations that contributed to the consultations for the valuable advice and information they have provided. In particular, we would like to thank the ARCHER team and scientific consortia, and the Tier-2 centres for the valuable support and data provided. We would also like to acknowledge the useful guidance and feedback provided by the Steering Committee throughout this research. Responsibility for the contents of this report remains with London Economics.

Authors

Daniel Herr, Economic Consultant, dherr@londoneconomics.co.uk, +44 (0)20 3701 7715
Charlotte Duke, Partner, cduke@londoneconomics.co.uk, +44 (0)20 3701 7705
Moritz Godel, Divisional Director, mgodel@londoneconomics.co.uk, +44 (0)20 3701 7708
Ashwini Natraj, Senior Economic Consultant, anatraj@londoneconomics.co.uk, +44 (0)20 3701 7726
Agata Makowska, Economic Analyst, amakowska@londoneconomics.co.uk, +44 (0)20 3701 7729
Rhys Williams, Economic Analyst, rwilliams@londoneconomics.co.uk, +44 (0)20 3701 7712
Ryan Perkins, Economic Analyst, rperkins@londoneconomics.co.uk, +44 (0)20 3701 7722
Carolyn Visser, Intern, cvisser@londoneconomics.co.uk, +44 (0)20 3701 7701

Cover picture credit (clockwise from top-left): Dr Alfonso Bueno Orovio, Department of Computer Science, University of Oxford. Dr Fulvio Sartor, University of Liverpool School of Engineering. Dr Sam Azadi, Imperial College London. Dr Matthieu Chavent, University of Oxford, SBCB Group, Department of Biochemistry. Mr Raphael Errani, Institute for Astronomy, Royal Observatory, University of Edinburgh. Mr Ivan Langella, University of Cambridge, Department of Engineering. Dr Peter Falkingham, Natural Sciences & Psychology, Liverpool John Moores University. Dr. Georg Schusteritsch, Department of Materials Science and Metallurgy, University of Cambridge.

Table of Contents

Page

Glossary

ARCHER	**Advanced Research Computing High End Resource** is the UK's national HPC facility which was launched in November 2013 and hosted at the EPCC at the University of Edinburgh. It is funded by EPSRC and NERC.
Archie WeSt	**Archie WeSt** is a regional Computing Centre of Excellence, based in the University of Strathclyde.
BBSRC	**Biotechnology and Biological Sciences Research Council** is a non-departmental public body, part of UK Research and Innovation, and the largest UK public funder of non-medical bioscience.
Cirrus	**Cirrus** is a Tier-2 centre run by the EPCC at the University of Edinburgh.
CSD3	**Cambridge Service for Data Driven Discovery** refers to a Tier-2 HPC centre based at the University of Cambridge. This HPC centre is also referred to as PETA-5.
Director's time	Approximately 5% of time on ARCHER is reserved for the EPCC, who host ARCHER. This time is called **Director's time**.
e-Infrastructure South	**e-Infrastructure South** was a regional Computing Centre of Excellence, a Tier-2 facility which was superseded by six new Tier-2 centres in 2016. It was run by a group of universities under the name Science and Engineering South.
eCSE	**Embedded Computational Software Engineering** is a programme with allocated funding to enhance the quality, quantity and range of science produced on ARCHER through improved software.
EPCC	**Edinburgh Parallel Computing Centre** is a HPC centre based at the University of Edinburgh. It is the host of ARCHER and has been commissioned to undertake a number of other projects by the EPSRC, including HPC outreach programmes.
EPSRC	**Engineering and Physical Sciences Research Council** is responsible for providing government funding for grants to undertake research and postgraduate degrees in engineering and the physical sciences in the UK. They have invested in HPC provisions around the country and have commissioned this report to measure the impact of these HPC investments.

HECToR	**High End Computing Terascale Resource** was the UK's Tier-1 HPC facility established in 2007 and superseded by ARCHER in 2014. Funding for Hector came from EPSRC, NERC and BBSRC.
HPC	**High Performance Computers/Computing** refers to supercomputers which have a higher level of performance relative to general purpose computers. Running on multiple cores, they can be used to conduct research in computational science, molecular modelling, physical simulations and weather forecasting, to name a few.
HPC Midlands	**HPC Midlands** is a regional Computing Centre of Excellence, a Tier-2 facility which was superseded by HPC Midlands Plus.
HPC Midlands Plus	**HPC Midlands Plus** is a supercomputing system focusing on materials science and computational fluid dynamics, based in the Midlands.
Isambard GW4	**Isambard GW4** refers to the Tier-2 HPC facility run by the GW4 group, consisting of the Universities of Bath, Bristol, Cardiff and Exeter.
JADE	**Joint Academic Data Science Endeavour** refers to a Tier-2 HPC centre, owned by the University of Oxford and hosted at the Hartree Centre, focusing on supporting research in machine learning.
kAU	**Kilo Allocation Unit** is a standard allocation unit on ARCHER and previous services. On ARCHER, 1 core hour is equal to 0.015 kAUs and 1 node hour is equal to 0.36 kAUs.
Materials Modelling Hub	**The Materials Modelling Hub** refers to the HPC facility, known as Thomas, based in UCL to perform research for materials and molecular modelling.
Midplus	**Midplus** was a regional Computing Centre of Excellence, a Tier-2 facility which was superseded by HPC Midlands Plus in March 2017.
N8 Centre of Excellence	**N8 Centre of Excellence** is a regional Computing Centre of Excellence, a Tier-2 facility formed in 2012. The member universities are Durham, Lancaster, Leeds, Liverpool, Manchester, Newcastle, Sheffield and York.
NERC	**Natural Environment Research Council** refers to the British Research Council that supports research activities in the environmental sciences. Along with EPSRC they contribute to the operational cost funding of ARCHER.

Peta-5	**Please see entry for CSD3**
PRACE	**PRACE** is the Partnership for Advanced Computing in Europe which provides access to HPC systems based in member states allowing researchers access to resources at the Tier-0 level.
Researchfish™	**Researchfish™** is EPSRC's research outcomes system which is used as a secondary evidence source on the number of collaborations, research projects, spin-outs etc.
RSE Fellow	**Research Software Engineer Fellows** play a vital role in creating efficient scientific software capable of exploiting modern HPC software.
Scientific Consortia	ARCHER **Scientific Consortia** are groups that bring together computational scientists in particular scientific areas. A list of scientific consortia can be found in Annex A3.3.
Tier-0 Centres	**Tier-0 Centres** are pan-European HPC infrastructures including the PRACE HPC.
Tier-1 Centres	**Tier-1 Centres** comprise the national supercomputing services which provide capabilities for the most complex modelling and simulations. In the UK the Tier-1 centres include ARCHER (funded by EPSRC and NERC) and DiRAC (funded by the STFC).
Tier-2 Centres	**Tier-2 Centres** refers to regional HPC centres including Peta-5/CSD3, Materials Modelling Hub, JADE, HPC Midlands Plus, Isambard GW4, Cirrus.
Tier-3 Centres	**Tier-3 Centres** are the lowest tier, the local 'work-horse' computers which enable basic computational research.

Acknowledgment of NERC funding

This study has been commissioned by the Engineering and Physical Sciences Research Council (EPSRC) to evaluate the impact of its investment in High Performance Computing for academic research; specifically its investments in ARCHER (the national Tier-1 HPC) and Tier 2 regional HPC centres funded in 2012 and in 2016.

The operational costs for ARCHER are shared by EPSRC and the Natural Environment Research Council (NERC) in a 77 : 23% partnership respectively: NERC funded academics therefore receive 23% of the allocated time on the ARCHER service in proportion to their share of the operational funding.

The focus of this study is on impacts arising from primarily engineering and physical sciences (EPS) research. Whilst the study does touch upon other (i.e. non-core EPS) disciplines, it does not provide comprehensive coverage of NERC funded research undertaken using ARCHER.

Executive Summary

London Economics was commissioned, by the Engineering and Physical Sciences Research Council (EPSRC), to undertake a study to assess the impact of EPSRC's investments in High Performance Computing infrastructure. The study sought to:

- explore the scientific, economic, societal and environmental impacts of EPSRC's investments in HPC provisions; and,
- estimate the return on EPSRC's investment in HPC.

For the purpose of this study, the scope included the UK's national (Tier-1) HPC facility, ARCHER, as well as EPSRC funded Tier-2 HPC facilities. Furthermore, whilst this study did not specifically look for impacts that continue from the UK's previous national service, HECToR, which were already captured in 'The story of HECToR'[1], it does nonetheless capture any impacts that come to the fore during the course of this study. (see Box 1, below)

Box 1 Study scope

Tier-1 comprises the national supercomputing services, providing capabilities for most complex modelling and simulations. Tier-1 includes **ARCHER** (EPSRC and NERC) and DiRAC (STFC, not included in study scope). The study scope also included impacts of ARCHER's predecessor, **HECToR**, that come to the fore during the course of the study.

Tier-2 bridges the gap between local systems and the national infrastructure, offering higher capability and power for more complex simulations. Tier-2 systems also provide a range of different architectures (such as GPUs), optimised for different computation tasks. The study scope includes the six currently active EPSRC funded Tier-2 facilities - **Peta-5**, **Materials Modelling Hub**, **JADE**, **HPC Midlands Plus**, **Isambard GW4**, and **Cirrus** - and the previous regional Computing Centres of Excellence - **Midplus**, **HPC Midlands**, **ARCHIE-WeSt**, **e-Infrastructure South** and **N8**.

Source: London Economics

High Performance Computing is a fundamental pillar of modern research and discovery. Investment by EPSRC has enabled world-leading research across a wide range of fields. These investments also contribute to the competitiveness of UK science and support research collaborations across Europe, the USA, China and Australia.

In addition to supporting world class research, HPC has supported doctoral and skills training helping to equip the next generation of UK scientists both within academia and industry. UK HPC capability also supports industrial collaborations leading to increased productivity in UK industry.

While estimating a precise monetary benefit of investment in HPC has significant challenges, evidence gathered during this study finds that, over the operational time of the services, the **overall economic benefit of HPC**[2] **is between £3.0 billion and £9.1 billion.** (Table 1)

[1] Available at http://www.storyofhector.org/

[2] This aggregate economic impact captures impacts from both EPSRC, NERC (ARCHER and HECToR benefits), BBSRC (HECToR benefits), and the Tier-2 centres. Spillover impact of HPC research on UK output are captured for EPSRC grant holders only.

This compares to a total cost of HPC to the public purse of approximately **£465.9 million** (Table 2). The **return on these investments is thus estimated to be between 6.5:1 to 19.5:1**.

In reality, the impact of HPC will lie somewhere between the low and high estimates presented here. However, due to the significant challenges and uncertainty inherent to assessments of scientific R&D investments, a point estimate is not provided.

Table 1 Aggregate economic impact of EPSRC* HPC investments in the UK (£m, 2018 prices)

Type of impact (£m in 2018 terms, over operational time of services)	Low estimate (£m)	High estimate (£m)
Impact of UK scientific research and discovery	**£ 2,021.1 m**	**£ 6,051.1 m**
Avoided cost of free HPC access for academics	£ 213.3 m	
Spillover impact of HPC research on UK output (EPSRC grant holders only)	£ 1,807.7 m	£ 5,837.7 m
Impact of direct industry access	**£ 906.7 m**	**£ 2,858.0 m**
Contribution of industry impacts to UK output	£ 906.7 m	£ 2,858.0 m
Impact of training and skills development	**£ 101.2 m**	**£ 190.3 m**
Benefits of PhD and postdoc training of students entering industry to students and the UK exchequer	£ 99.0 m	£ 188.1 m
Benefits of provision of free HPC training courses	£2.2 m	
Total economic impact	**£ 3,029.0 m**	**£ 9,099.4 m**

Note: The true benefits of improved or new software is the contribution that this software makes to scientific research & discovery. However, these benefits are already implicitly monetised when estimating the spillover benefits of EPSRC funded HPC research. Therefore, to avoid double counting, software was excluded from the calculation of total aggregate benefits. Similarly, the true benefits of skills development in HPC is the value brought to UK companies and UK science. However, these benefits are already implicitly monetised in the benefits to industry and the spillover impacts of HPC research. In addition, software has benefits on industry as well, these benefits were not quantified. All estimates are rounded to the nearest £0.1 m. * See Box 2, some of the impacts captured could be from NERC funded research. ***Source: London Economics' analysis. Icon credits: bioraven – Shutterstock***

Table 2 HPC investments (£m, 2018 prices)

Type of investment	EPSRC	NERC	BBSRC	RCUK / UK Government	Total
Investments in centres	**£ 109.6 m**	**£ 22.8 m**	**£ 3.3 m**	**£ 119.0 m**	**£ 254.6 m**
ARCHER	£ 27.3 m	£ 8.3 m		£ 46.8 m	£ 82.4 m
HECToR	£ 47.9 m	£ 14.4 m	£ 3.3 m	£ 72.1 m	£ 137.8 m
New Tier-2 centres[1]	£ 21.5 m				£ 21.5 m
Regional Tier-2 centres[1]	£ 12.8 m				£ 12.8 m
Public purse costs of EPSRC research funding	**£ 189.7 m**			**£ 21.6 m**	**£ 211.3 m**
EPSRC HPC research funding[2]	£ 189.7 m				£ 189.7 m
Further public funding of EPSRC researchers[2]				£ 21.6 m[3]	£ 21.6 m
Total	**£ 299.2 m**	**£ 22.8 m**	**£ 3.3 m**	**£ 140.5 m**	**£ 465.9 m**

Note: (1) Investment in Tier-2 centres only includes EPSRC investments; i.e. excluding CAPEX and OPEX costs to the centres themselves and hosting institutions; (2) Estimates of funding are based on EPSRC's research outcomes systems, which records quantitative information only for research activities funded directly by the EPSRC, and further funding given to EPSRC grant holders by other bodies, if these are reported to the EPSRC. Research activities using HPC systems funded by other UK Research Councils, other public bodies, or private organisations are not included in the estimate *unless* the funding was fed back into EPSRC's research outcomes systems. Only costs of further funding accruing to the public purse were counted in the calculation. This includes funding provided by UK research councils and other UK public bodies. (3) May include funding by NERC or BBSRC. ***Source: London Economics' analysis based on data provided by EPSRC***

Box 2 Note on estimated benefits and costs

While the study was focused on EPSRC's investments in HPC, the estimated economic impacts include benefits arising from funding by multiple sources including EPSRC, NERC, BBSRC, RCUK, the UK Government and the Tier-2 centres and partner institutions (see Section 1.3).

Nevertheless, the study does not provide comprehensive coverage of benefits arising from HPC research funded by non-EPSRC sources. In particular, spillover benefits of HPC research were estimated only for research funded directly by the EPSRC, and further funding given to EPSRC grant holders by other bodies, if these were reported to the EPSRC. Research activities using HPC systems funded by other UK Research Councils, other public bodies, or private organisations are not included in the estimate *unless* the funding was fed back into EPSRC's research outcomes systems.

In line with benefits, cost estimates represent the cost to the public purse, in 2018 prices, and include the total cost of ARCHER (funded by the UK Government, EPSRC and NERC); the total costs of HECToR (funded by RCUK, EPSRC, NERC, and BBSRC); EPSRC's funding of HPC research and other public purse costs of funding of EPSRC grant holders reported back to the EPSRC; and EPSRC's contribution to the Tier-2 centres.

Tier-2 centres and the hosting institutions further carry part of the capital expenditure and all of the operational expenditure themselves. This cost is assumed not to be carried by the public purse and is thus not included in the cost calculation. In reality, some costs may ultimately be covered by the public purse. Costs to the public purse may thus be higher than those reported here.

On the other hand, as discussed above, only benefits of research by EPSRC grant holders were included in the calculation of benefits. Moreover, benefits of software to industry were also not monetised (see discussion in Section 6).

Source: London Economics

HPC Usage

ARCHER has over **4,100 registered UK academic users** from more than 65 institutions along with over **150 industry users** who accessed ARCHER over its lifetime. Furthermore, many more industrial users benefit from HPC capabilities via collaborations with academic users. These numbers represent nearly double the number of users who accessed HECToR, the UK's previous Tier-1 facility, clearly demonstrating the continued high demand for HPC by the scientific community.

In addition, more than **1,800 academic user accounts and over 100 industrial user accounts were reported across the six Tier-2 centres established in 2016[3].** Almost 70% of HPC users responding to the survey undertaken for this study reported having used multiple HPC facilities.

Unsurprisingly, the majority of HPC users are based in the UK. Nevertheless, HPC's geographical reach, through international collaborations and initiatives such as PRACE, extends far beyond the UK.

Access to HPC is free for academics at the point of use, or provided at a small nominal cost.

[3] Note that data on the number of users was not available for all Tier-2 centres. As such, this number reflects a lower bound.

Benefits of HPC to the UK

HPC delivers benefits across a wide range of areas. These include:

Scientific research and discovery. HPC is a pillar of modern research and discovery, from materials modelling, and computational fluid dynamics to biomolecular simulations and beyond. HPC allows researchers to undertake complex simulations which enable them to answer more granular and deeper questions than is possible through experimentation and theory alone. 82% of researchers who responded to the study survey reported that HPC has enabled them to do more research, 88% reported that HPC has improved the efficiency of their research and 84% responded that HPC has improved the quality of their research (Section 3.1).

International competitiveness of UK science. High quality HPC facilities, and leading HPC research groups, play an important role in attracting high quality researchers and students to the UK. Of academics who responded to the survey, 93% agreed that HPC helps make UK science more competitive internationally; 76% agreed that HPC helps attract international students to study at UK universities, whilst 86% agreed that it attracts researchers to undertake their work at UK universities (Section 3.2).

Diffusion of research benefits to the private sector. Evidence from the academic literature suggests that investments in public sector research and development can have significant spillover benefits to the private sector. An analysis based on data obtained from EPSRC's research outcome system, Researchfish™, and estimates of productivity spillovers from the literature, indicates that HPC research by EPSRC grant holders results in spillover benefits of between £1.8 billion and £5.8 billion[4] (Section 4.1).

Software benefits to research. Alongside hardware and people, software forms a crucial part of the HPC ecosystem. Allowing existing hardware to be used more efficiently enables researchers to undertake more, or more complex, simulations on the same machine. The EPSRC has invested in a number of software development programmes that benefit the HPC community. Drawing upon estimated benefits of time savings from code developed on ARCHER[5], and assuming similar time savings benefits of software reported by academic users to the EPSRC, total value of time saved due to software development would be in the region of £53.4 million (Section 4.2).

Software benefits to industry. HPC software also brings significant benefits to industry. This is highlighted by CASTEP, an HPC simulation software used to calculate the properties of materials from first principle, which is used by around 900 industrial users. As the Goldbeck report indicates, companies using materials modelling software such as CASTEP, among others, achieved cost savings ranging from €100,000 to €50 million, with an average return on investment of 8:1 (Section 4.2).

[4] Note that these estimates are based only on the proportion of research that would not have been undertaken without HPC (70%, see Section 3.1.2). Spillover benefits of HPC research not dependent on HPC were excluded.

[5] Estimates provided by EPCC, based on 91 of 100 EPSRC/NERC eCSE projects completed as of December 2018, find that software code improvements enabled by eCSE are estimated to have generated £24.5 million in benefits through time savings allowing additional research on ARCHER. This represents a benefit cost ratio of more than 4:1.

Skills development and training. HPC training brings significant benefits to individuals. 85% of survey respondents agreed that HPC provided them with skills and knowledge that can be used across a range of jobs and industries; 75% agreed that the use of HPC has made them more attractive to potential employers; and, 77% agreed that it helped them advance their career (Section 4.3).

PhD training. Over 170 PhD or post-doctoral students were trained on ARCHER across the EPSRC scientific consortia in 2016/17 alone. Data on PhD and postdoctoral access is not recorded in all HPC centres. Nevertheless, the analysis in Section 4.3.3, suggests that a total of between 787 and 1,477 PhD and postdoctoral students could have been trained across the HPC centres. Taking account of the earnings premium associated with HPC doctoral and postdoctoral training the following benefit estimates can be made:

- the total present value of additional benefits accruing to the Exchequer is between approximately £50.7 million and £96.3 million; while,
- the total present value of benefits accruing to the graduates themselves is between approximately £48.3 million and £91.8 million. (Section 4.3.3).

Enabling collaborations. HPC plays a crucial role in enabling collaborations. EPSRC grant holders reported 367 collaborations with partners from the UK and at least 29 other countries. Stakeholders consulted for this study repeatedly emphasised the crucial role that HPC plays in enabling these collaborations. For example, amongst survey respondents, 76% agreed that HPC helped in facilitating collaborations and 74% that the collaborations they have been involved in would not have been possible without HPC (Section 4.4).[6]

Benefits to industry. HPC brings significant benefits to industry through a number of channels such as direct access or collaborations with academics. Key direct benefits of HPC are the size of computing resources, which are much more powerful than many industry users would be able to afford on their own, the flexibility and cost of the services, and the professionalism with which the services are operated. Assuming an effective increase in profits rate[7] of between 0.6% and 1.7%, industry use of HPCs is estimated to have contributed between £0.9 billion and £2.9 billion to the UK economy over the operational time of these services. (Section 5.1.1)

Outreach, diversity and inclusion. To ensure an efficient and productive research ecosystem, a number of diversity and inclusion programmes have been introduced so that no demographic characteristic is disadvantaged and HPC access benefits all members of the society (Section 4.4.1). These activities include:

- **Diversity in HPC** which supports the inclusion of under-represented groups working in HPC by encouraging participation through showcasing HPC as a career path.
- **Women in HPC** which encourages and promotes women in the HPC community through raising awareness and the support of collaborations and networking.
- **Public outreach** such as the use of Wee Archie, a miniature supercomputer, taken to science fairs across the country, a build-a-PC workshop, a bean bag

[6] Note that benefits of collaborations are indirectly evaluated in the benefits to industry and research spillovers.
[7] I.e. the proportion of firms seeing an increase in profit times the average increase in profit per firm.

algorithm sorting activity and a supercomputing app. These activities have been used over 58,000 times across 55 UK locations and 6 international venues.

- **ARCHER driving test** this gives new users basic access to ARCHER in order to encourage their future participation, conditional on passing a short test to ensure sufficient knowledge of HPC and ARCHER.

1 Introduction

In October 2018, the Engineering and Physical Sciences Research Council (EPSRC) commissioned London Economics to undertake an independent study to evaluate the impact of EPSRC's investment in High Performance Computing (HPC). The focus of the research is on the socio-economic and research benefits enabled by HPC.

The study comes at a time when increasing pressure from competing funding priorities across government, combined with significant investments in HPC infrastructure by other countries, especially the USA and China, mean that the UK's HPC capabilities are falling behind other nations.

In light of this, this study sought to address two key questions:

- What are the scientific, economic, societal and environmental impacts of EPSRC's investments in HPC provisions?
- What is the return on EPSRC's investment in HPC?

In scope of the study were EPSRC's HPC investments, over the last ten years, into national Tier-1[8] and regional Tier-2 HPC infrastructure (see Section 1.2 for further details), specifically:

- EPSRC and NERC funded ARCHER service (Tier-1);
- EPSRC funded five regional Computing Centres of excellence established in 2012 (Tier-2); and,
- EPSRC funded six Tier-2 centres that started in 2016.

The impact from the preceding Tier-1 service HECToR was already captured in 'The story of HECToR'[9]. Whilst this study is not specifically looking for impacts that continue from HECToR it would nonetheless capture any impacts that come to the fore during the course of this study.

In addition to evaluating EPSRC's HPC investments, the study also sought to benchmark these investments against similar HPC investments made by other EU countries.

The remainder of this section sets out the background to the study (Section 1.1) and provides an overview of EPSRC supported computing provision (Section 1.2) as well as EPSRC's investments in HPC infrastructure (Section 1.3). Section 1.4 discusses the representativeness of the data sources used in this study. Finally, Section 1.6 provides an overview of the structure of the remainder of the report.

1.1 Background

High Performance Computing (HPC) is critical to scientific research and innovation as well as economic leadership. Among HPC projects analysed by the IDC[10], average revenue generated was approximately £560[11] ($867) for each dollar invested in HPC, with a European average profit of $69

[8] Note that both EPSRC and NERC invested in national Tier-1 infrastructure.
[9] Available at http://www.storyofhector.org/
[10] IDC (2014). *EESI-2 Special Study To Measure And Model How Investments In HPC Can Create Financial ROI And Scientific Innovation In Europe*
[11] Original value reported in US $. Converted into Sterling using the 2014 Bank of England average annual spot exchange rate and adjusted for inflation using the ONS Consumer Price Index.

for each dollar invested, and 1,152 new jobs created across 52 industrial organisations; while industrial sectors that leverage HPC are estimated to add up to between 2% and 3% to Europe's GDP by 2020[12]. Many countries have recognised the importance of HPC investments. Overall performance of HPC infrastructure, in the USA, China, Japan, Germany, the UK and France, increased by a factor of 211 between November 2007[13] and November 2018.[14]

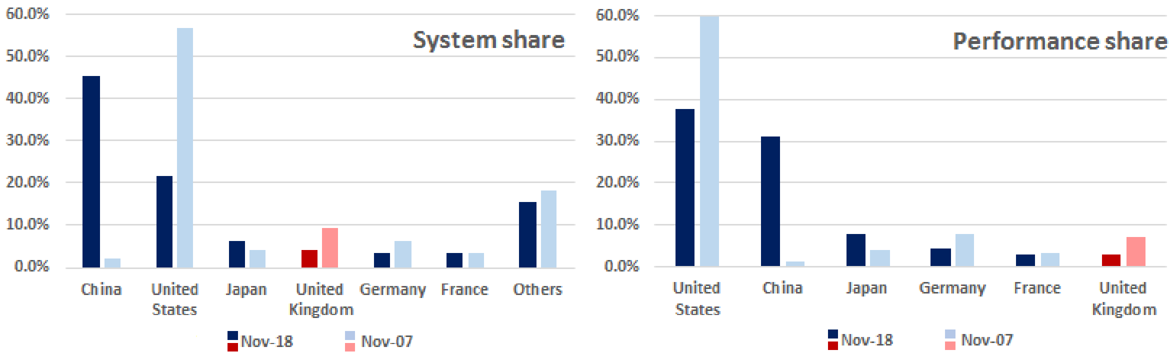
The European Union has also recognised the importance of HPC infrastructure and, in March 2017, started a joint undertaking, EuroHPC, that will pool European resources in order to develop the EU's HPC capacity, with an ambition to have EU supercomputers in the global top 3 by 2022/23[15].

In the UK, research council investments in HPC forms a crucial part of the scientific infrastructure, underpinning many of the UK's Eight Great Technologies[16], with more than 100,000 researchers reliant on the UK's e-infrastructure[17].

Whilst a number of the benefits identified in this report are related to an industrial or business context, the HPC activity that enables these benefits are a necessary pillar for research and scientific advancement as highlighted in the impact study of HECToR and ARCHER[18].

Investment in HPC infrastructure has allowed the UK to maintain its strong position in HPC since the introduction of HECToR in late 2007. In November 2018, the UK ranked fourth in the world in terms of HPC system share and sixth in terms of HPC performance share (Figure 1).

Figure 1 Proportion of HPC infrastructure, by country – Nov-18 and Nov-07



Note: Performance measure = maximal LINPACK performance achieved.

Source: London Economics analysis based on data obtained from Top500.org

However, other countries increased their public and/or private HPC investments at a faster rate over the same period[19]. China in particular has made significant investments in its HPC infrastructure. Since November 2007, China increased its share of HPC systems from 2.0% to 45.4% in November 2018, and now ranks first in the world in terms of HPC system share (Figure 1). As a result, the UK's share of HPC infrastructure has dropped from 9.4% in November 2007 to 4.0% in November 2018, while the UK's performance share has dropped from 7.3% to 2.9% over the same period (Figure 1).

12 European Commission (n.d.). *Pooling Resources for a European High Performance Computing Infrastructure*
13 When HECToR, the UK's former national HPC service, was introduced.
14 London Economics analysis based on *Top500.org data.* Performance measure = maximal LINPACK performance achieved.
15 European Commission (n.d.). *Pooling Resources for a European High Performance Computing Infrastructure*
16 EPSRC (2014). *The impact of HECToR*
17 Clarke, E., and Larmour, I. (2016). *The Impact of National High Performance Computing: An analysis of the impacts and outputs of investment in national HPC*
18 Ibid
19 EPSRC (2014). *The impact of HECToR*

This indicates that whilst other leading economies are increasing their HPC investment, the UK is falling behind and may, as a result, see a decline in its international competitiveness.

Looking to the future, the HPC industry is awaiting the next generation of computing systems, known as exascale systems, capable of performing a billion billion calculations per second. These systems, which reduce energy consumption 100-fold, will require substantial development of programming models to more efficiently use this new technology[20]. In order for the UK to remain competitive, significant continued investments in HPC infrastructure are needed. These investments are required across the scale to build capability at all levels which can efficiently and effectively exploit the benefits, identified in this report, from HPC investment.

1.2 Computing provision supported by EPSRC

Supercomputers allow researchers to solve a range of scientific and engineering problems, running simulations and calculations which require large numbers of processing cores working in parallel[21]. There are many examples of research areas which have benefitted from HPC including the development of new techniques to more accurately simulate the behaviour of materials, advancing our understanding at an atomic level and enabling the development of new materials, to complex simulations to understand the role of biomolecules within bacterial, plant and animal cells. The latter research area enabling the design of novel drugs.

For practical purposes, HPC provision for academic users in the UK can be classified into four tiers, with EPSRC investments at all levels. In reality, the ecosystem, in terms of use, is not as distinct, with academics using different or all provisions in their research. These four tiers are depicted in Figure 2 and summarised below:

- **Tier-3**, the lowest tier, comprises local ‘work-horse’ computers that enable basic computational research.
- **Tier-2** comprises regional HPC centres, offering higher capability and power for more complex simulations and bridging the gap between local systems and the national infrastructure. Tier-2 systems also provide a range of different architectures (such as GPUs), optimised for different computation tasks.
- **Tier-1** comprises the national supercomputing services, providing capabilities for most complex modelling and simulations. Tier-1 includes ARCHER (EPSRC and NERC investment) and DiRAC (STFC).
- **Tier-0** comprises pan-European HPC infrastructures, requiring collaboration between countries.

Launched in November 2013, ARCHER (Academic Research Computing High End Resource) is the UK’s national HPC facility, replacing its predecessor HECToR (2007-2013) funded by EPSRC, NERC and BBSRC. ARCHER was funded by the EPSRC and NERC and is hosted at the EPCC (Edinburgh Parallel Computing Centre) at the University of Edinburgh. When it first came into operation, ARCHER was among the top 20 fastest HPC systems in the world, however, at the time of writing this report, it had dropped to 186th but plans are underway to replace the infrastructure with a new, upgraded system[22].

[20] PRACE (2012). *The scientific case for High Performance Computing in Europe 2012-2020: From petascale to exascale*
[21] *About ARCHER*. http://www.archer.ac.uk/about-archer/ [accessed on 08/11/2018]
[22] Top 500 (2018). *TOP500 List - November 2018*

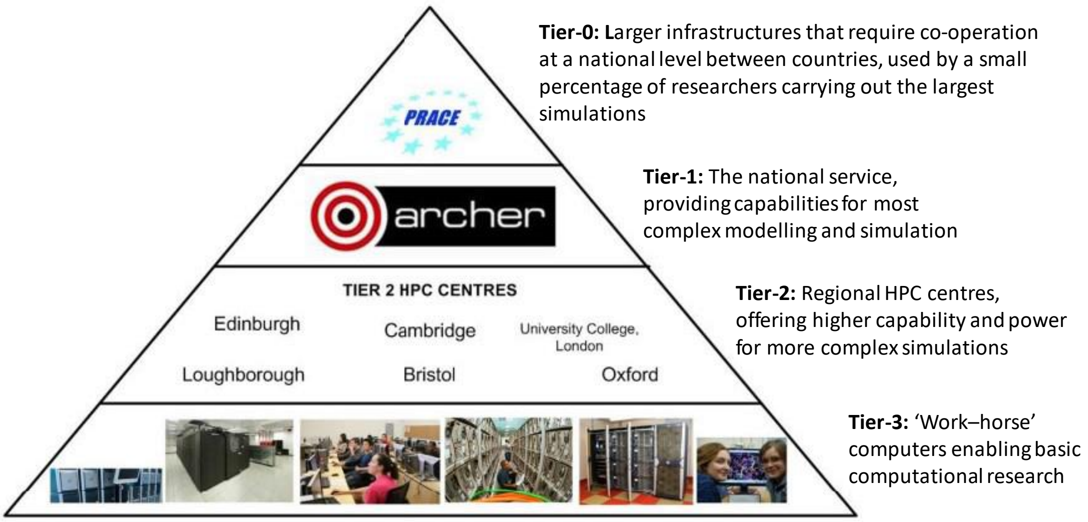
Despite these new investments into ARCHER 2, stakeholders were under the belief that the investments being made are not comparable to the investments made in the rest of Europe (e.g. Germany and Switzerland), let alone leaders such as the USA or China. As a result, the UK would be accepting that it is not the leader in this field and would lose researchers working at the cutting edge on state-of-the-art research projects. Instead, these users would work on HPC in China and the USA.

Tier-2 computers are fundamental for UK high performance computing as they provide a range of computing architectures which are driven by science needs not met by national or university-based facilities. Tier-2 permits training of computer scientists and technicians. It is also integrated with Tier-1 services and can be used as a test-bed to promote code development before being scaled up for Tier-1 use. There are currently six active Tier-2 facilities funded by EPSRC:

- Peta-5 / CSD3 (led by the University of Cambridge)
- Materials Modelling Hub (led by University College London)
- JADE (led by the University of Oxford)
- HPC Midlands Plus (led by Loughborough University)
- Isambard the GW4 Tier-2 HPC Service (led by the University of Bristol)[23]
- Cirrus (led by the University of Edinburgh)

These six Tier-2 facilities were established in 2016, replacing the previous regional Computing Centres of Excellence - Midplus, HPC Midlands, ARCHIE-WeSt, e-Infrastructure South and N8 Centre of Excellence - which were operational from 2012-13.

Figure 2 Computing provision supported by EPSRC



Source: EPSRC

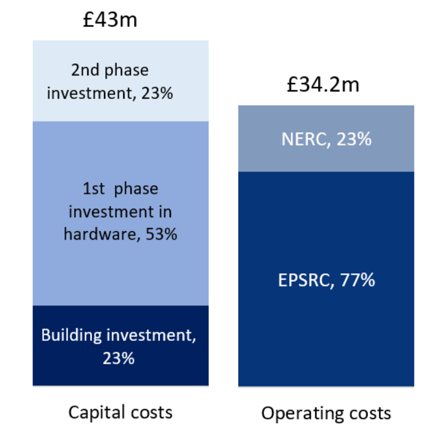
[23] Note that this service has only recently come online and as such has not yet generated statistics on usage.

1.3 EPSRC's investments in HPC infrastructure

To assess the **net socio-economic benefits** of EPSRC's investments in HPC infrastructure, it is important to understand the costs of **EPSRC's investments** into national and Tier-2 HPC infrastructure under the baseline scenario.

The EPSRC wishes to understand the impact of its national (ARCHER) and Tier-2 HPC investments over the last 10 years[24], as well as continuing benefits from ARCHER's predecessor, HECToR. These investments are shown in Figure 3 (ARCHER) and Figure 4 (Tier-2 centres), and briefly detailed below.

Figure 3 Costs of ARCHER



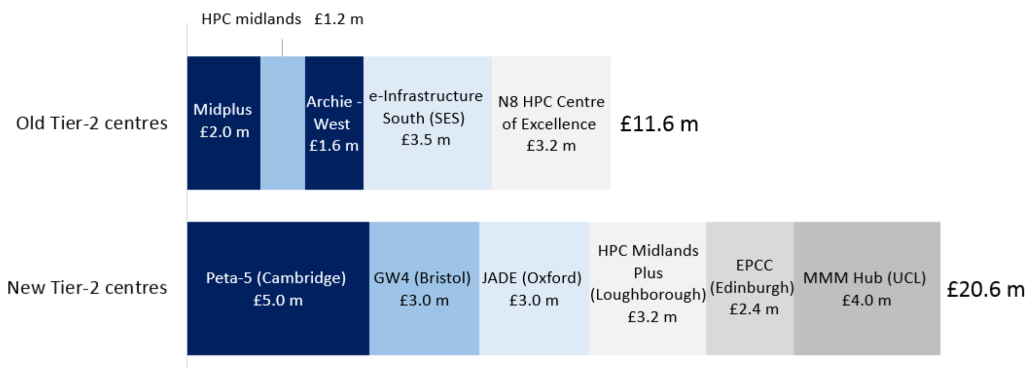
Note: Capital costs were funded as part of the spending review settlement provided by the government. ***Source: EPSRC***

ARCHER, the current national HPC service, cost a total of **£78.2 million** over the five years of its lifetime. This comprises **£43 million** capital investments by the Government, including a **£10 million** building investment. Operational costs are shared by the EPSRC and NERC and totalled **£34.2 million** over the five years of ARCHER's lifetime.

The total cost of ARCHER's predecessor, **HECToR**, was **£118 million** over the course of HECToR's lifetime. Capital expenditure, such as acquisition of the hardware, accounted for **£60 million** of this total cost and was covered by Research Councils UK (RCUK)[25]. Operational costs made up a total of **£58 million** over the course of HECToR's lifetime and split between the EPSRC (**73%**), NERC (**22%**), and BBSRC (**5%**).

In terms of the **Tier-2 centres** EPSRC invested **£11.6 million** in 2012-13 to support the establishment of five regional Tier-2 Centres of Excellence as well as a further **£20.6 million** in six new Tier-2 Centres in 2016. In addition to EPSRC's investments, part or all of EPSRC's funding was matched by host institutions and collaborators. Operational costs of the Tier-2 centres are covered by the centres or their host institutions.

Figure 4 EPSRC capital investments in Tier-2 centres



Note: Due to rounding approximations the sum of components may not equal the total. ***Source: EPSRC***

[24] Note that investment into national (ARCHER) HPC investment was shared with NERC.
[25] Research Councils UK was an umbrella body for the seven UK Research Councils. It was superseded by UK Research and Innovation (UKRI) in 2018.

In addition to investments in HPC infrastructure, EPSRC also allocates significant investments across the HPC ecosystems including grant funding for HPC research, support for software development efforts, and investments in outreach, diversity and inclusion activities. Further details on these investments can be found in the relevant chapters of each impact strand.

1.4 Caveats and limitations

The research has been conducted by a team of independent professional economists. Estimates of economic impacts are based on best practice and best judgement to calculate the most robust and fair estimates. Two detailed methodological annexes describe the overarching approach and counterfactual (Annex 1), and the methodology used and assumptions made as well as the caveats and limitations of the analysis, where appropriate (Annex 2).

Nonetheless, the reader should bear in mind the following high-level limitations and caveats of this study throughout:

- Estimating benefits of scientific R&D investments poses significant challenges and uncertainty. Therefore, where appropriate, benefits were estimated as a range to take these uncertainties into account. For this reason, a low and a high estimate are provided for the aggregate economic impact. In reality, the impact of HPC will lie somewhere between these estimates.
- While the study was focused on EPSRC's investments in HPC, the operational and funding structure of HPC means that it was not possible to clearly distinguish between benefits attributable to EPSRC and benefits attributable to other sources of funding. Therefore, the analysis of economic benefits includes benefits arising from funding by multiple sources including EPSRC, NERC, BBSRC, RCUK, the UK Government and the Tier-2 centres and partner institutions. Nevertheless, the study does not provide comprehensive coverage of benefits arising from HPC research funded by non-EPSRC sources. (See Box 2 for further discussion).
- Similarly, costs were assessed in terms of total costs to the public purse and include costs carried by non-EPSRC sources including NERC, BBSRC, RCUK, the UK Government and the Tier-2 centres and partner institutions. Costs assumed not to be accruing to the public purse were excluded from the overall cost estimate. (See Box 2 for further discussion).
- Parts of the analysis are based on stakeholder consultations including an online survey of HPC users. Therefore, all the usual caveats of stakeholder consultations and online surveys, such as non-representativeness and selection bias, apply wherever the analysis relies on these data sources. Where part of analysis is based on survey results, this is made clear in the relevant sections. The caveats of stakeholder consultations and online surveys are discussed in Section 1.5, which also examines the representativeness of the survey.
- Parts of the analysis are based on EPSRC's research outcome system, Researchfish™. This system only records data for EPSRC grant holders. As such, analysis based on Researchfish™ is specific to EPSRC and does not, unlike the stakeholder data, include NERC benefits. Moreover, Researchfish™ data relies on grant holders reporting back to EPSRC. As such, analysis based on Researchfish™ may significantly underestimate the true figures as a result of under-reporting. In particular, spillover benefits of HPC research are based on Researchfish™ and therefore were estimated only for research funded directly by the EPSRC, and further funding given to EPSRC grant holders by other bodies, if these were reported to the EPSRC (See Box 2 for further discussion).

- While every effort has been made to quantify benefits wherever possible, not all benefits were quantifiable. Unquantifiable benefits are highlighted through qualitative discussions and case studies, wherever the authors were aware of these benefits. In particular, benefits of software to industry were not monetised.

1.5 Representativeness of data sources used

1.5.1 Online survey and stakeholder consultations

To understand the importance of HPC for academic research, a wide range of stakeholders were consulted as part of this study on the impact of EPSRC and NERC funded ARCHER, and EPSRC funded Tier-2 centres. Consultations included five workshops and a number of telephone interviews with leading academics in the HPC field, as well as an online survey of a wide range of academics using HPC for their own research.

Whilst the stakeholder consultations and online survey were targeted at EPSRC researchers, a small proportion of respondents were NERC users (see Figure 5 for proportions of respondents undertaking research in NERC areas), or researchers benefitting from investments made by NERC, and as such are included in this analysis.

As is always the case when relying on surveys and stakeholder feedback, the results may not be fully representative of all users. It is possible that the experiences of some users may have been missed, and other experiences may have been over- or under-represented by the survey and consultations.

Sample selection bias may be present if certain users are more likely to complete the survey or engage in consultation compared with other users. This can happen because the benefits of engaging in stakeholder consultations is costly, in foregone time, and the benefits, in the form of greater HPC investment, may not be immediate. As a result, it is possible that users who benefit the most from HPC took the time and effort to engage in the consultation, thus leading to their over-representation. Alternatively, those who benefit the most from HPC might be occupied researching and not have time to respond to requests for consultation, in which case they would be under-represented.

To overcome the potential representation issues, primary evidence from the survey is combined with secondary evidence where available. Such secondary evidence includes the **EPSRC research outcomes system**, Researchfish™, along with direct data from EPCC on usage statistics.

It should be noted that any analysis based on data from EPSRC's research outcomes system, Researchfish™, is specific to EPSRC and does not, unlike the stakeholder data, include NERC benefits. Moreover, EPSRC's research outcomes system only includes research output and collaborative projects which were reported to EPSRC by EPSRC grant holders. As such, evidence based on EPSRC's research outcome system should be seen as a 'minimum achieved benefit', which may significantly under-represent the true figures as a result of under-reporting.

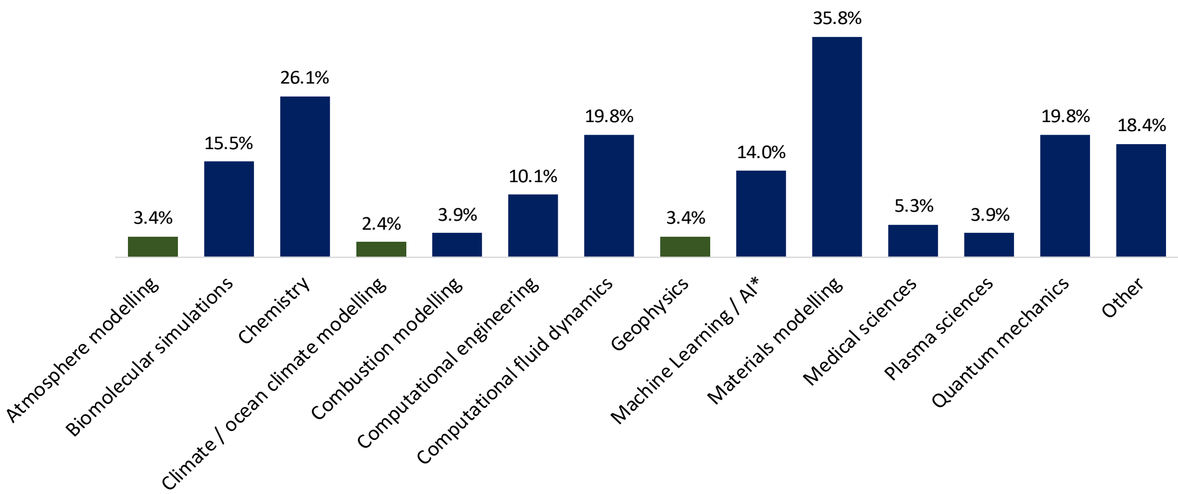
To gain a better understanding of the potential biases present in the analysis, Section 1.5.2 provides a comparison between respondents to the online survey and the population of registered users on ARCHER and the Tier-2 centres.

1.5.2 Online survey representativeness

Comparing Figure 5 and Figure 6, which show research areas of survey respondents and those of the population of HPC users, indicates that the survey under-represents NERC users undertaking research in atmosphere and climate / ocean climate modelling. This is unsurprising given that the focus of the study was on EPSRC's investments in HPC.

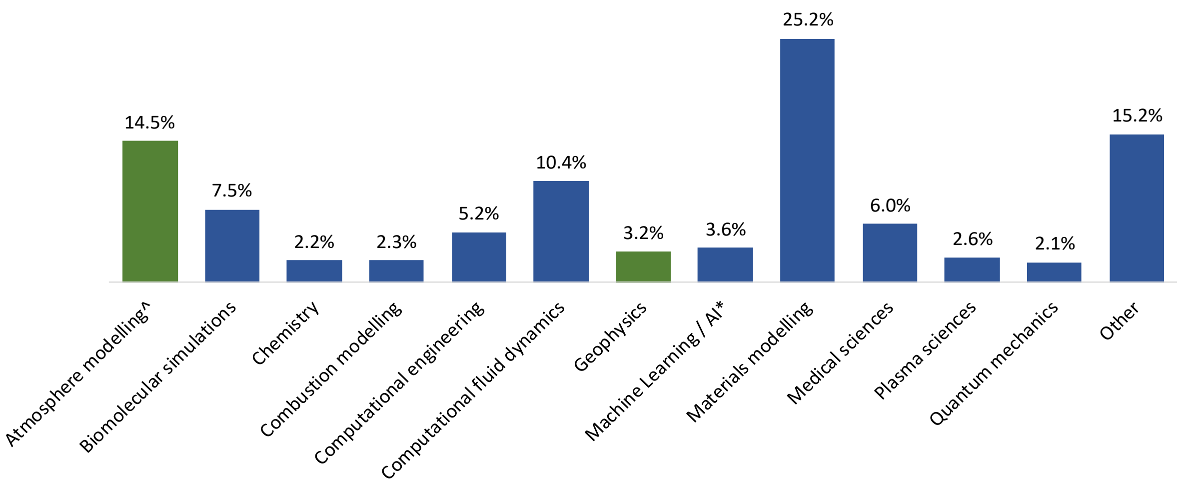
It should be noted that the survey allowed users to select multiple research areas, whereas data collected by HPCs user management system, shown in Figure 6, is restricted to one research area per user. Therefore, more general conclusions into the representativeness of the survey sample are not possible.

Figure 5 'In which broad area are you undertaking research?'



Note: Based on 207 respondents (including a small proportion of NERC users). Respondents could select multiple answers and hence percentages do not sum to 100%. (*) incl. natural language processing. Research areas marked in green are primarily NERC funded, though some science might be EPSRC or joint-funded. ***Source: London Economics survey of users of HPC capabilities***

Figure 6 Population of research areas by HPC user accounts

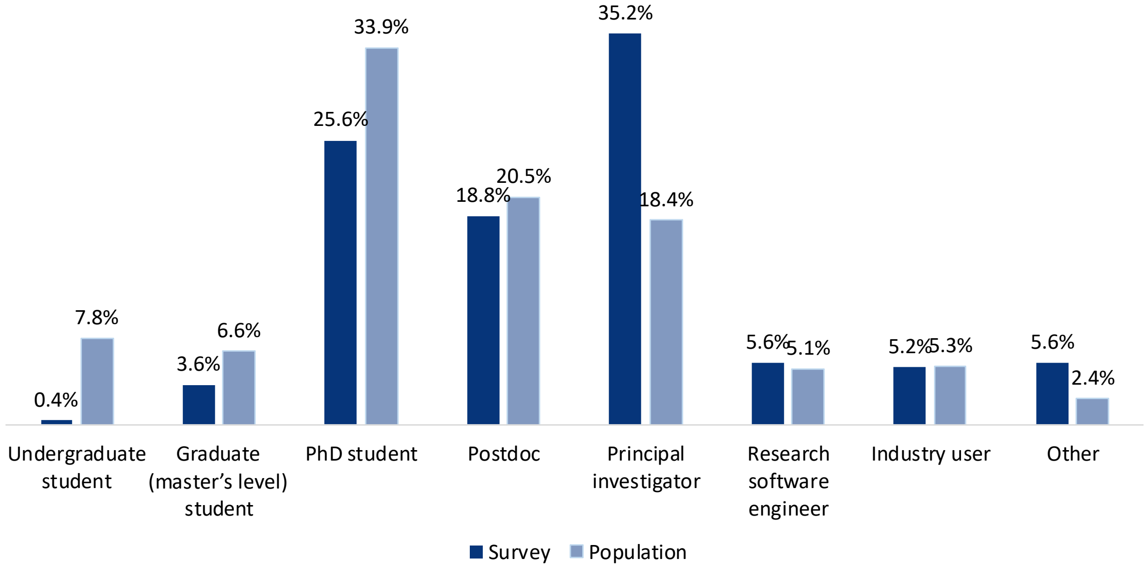


Note: Includes users from both ARCHER and Tier-2 machines. User accounts excludes those which were unknown or used for training and support. (^) includes climate/ocean modelling as this cannot be distinguished by the EPCC, (*) incl. natural language processing. Research areas marked in green are primarily NERC funded, though some science might be EPSRC or joint-funded.

Source: EPCC usage data

Figure 7 compares the demographic of survey respondents by career stage. The survey under-represents undergraduate students, graduate students, PhD students, postdocs and other users. Conversely, the survey over-represents principal investigators.

Figure 7 'What type of HPC user are you?'



Note: Survey based on 250 respondents (including NERC users). Population results includes users from both ARCHER and TIER-2 machines. User accounts excludes those which were unknown.

Source: London Economics survey of users of HPC capabilities; EPCC usage data

1.6 Structure of the report

The remainder of the report is structured as follows:

- **Section 2** provides an overview of the usage of HPC provision for academic research, or similar purposes.
- **Section 3** discusses the benefits of HPC to the academic research community as well as the benefits to the international competitiveness of UK science.
- **Section 4** highlights the contributions EPSRC's HPC investments have made to the UK economy including the impact of research, training and skills development, software development and support and collaborations.
- **Section 5** explores the impact of HPC on UK industry.
- **Section 6** provides an overview of the estimated aggregate economic impact of HPC and the corresponding return on investment.
- **Section 7** highlights EPSRC supported activities promoting outreach, diversity and inclusion, and the impact of these activities on the HPC community.
- **Section 8** examines the likely impact of the no HPC scenario on UK science.
- **Section 9** provides a comparison of the operational model and funding structure of EPSRC's HPC investments with those of other European HPC centres.
- **Section 10** provides concluding remarks.

Four annexes provide further details:

- **Annex 1** details the study approach, the baseline (EPSRC's investments in HPC) and counterfactual (No HPC) scenarios, and EPSRC's HPC investments under the baseline scenario. The implications of the *No HPC* scenario, and how realistic this scenario is, are also explored further in this section.
- **Annex 2** provides details of the methodology used to estimate economic benefits and costs and lists the main assumptions used in the modelling. Caveats and limitations of the methodology are discussed where appropriate.
- **Annex 3** provides additional material to supplement this report, including a logic map, commercial pricing of HPC, commercial pricing of HPC training and the groups comprising the EPSRC and NERC consortia.
- **Annex 4** provides a methodological summary of academic papers used in the analysis of the research spillover benefits.

2 Usage of HPC

2.1 Number of users

HPC facilities are used by a wide array of academics from different disciplines, both in the UK as well as abroad – for example via initiatives such as PRACE. **ARCHER had over 4,100 registered UK academic users from more than 65 institutions**. Details on the number of users of the Tier-2 centres are presented in Table 3, below.

In addition to supporting UK academic research, HPC facilities are also accessed directly by a number of industry users; **more than 150 industry users have accessed ARCHER directly**. Moreover, many more industrial users indirectly benefit from HPC capabilities via collaborations with academic users (see Section 5).

Table 3 No. of users by HPC facility

HPC facility type	HPC facility	Academic users	Industry users
National service	ARCHER	4,267[1]	158
	HECToR	2,405[2]	81
New Tier-2 centres	Cirrus	794	64
	Materials Modelling Hub	600	-
	Peta-5 / CSD3	102 (~18% industrial users)[3]	
	JADE	170	44
	GW4 Isambard[4]	-	-
	HPC Midlands Plus	222	-

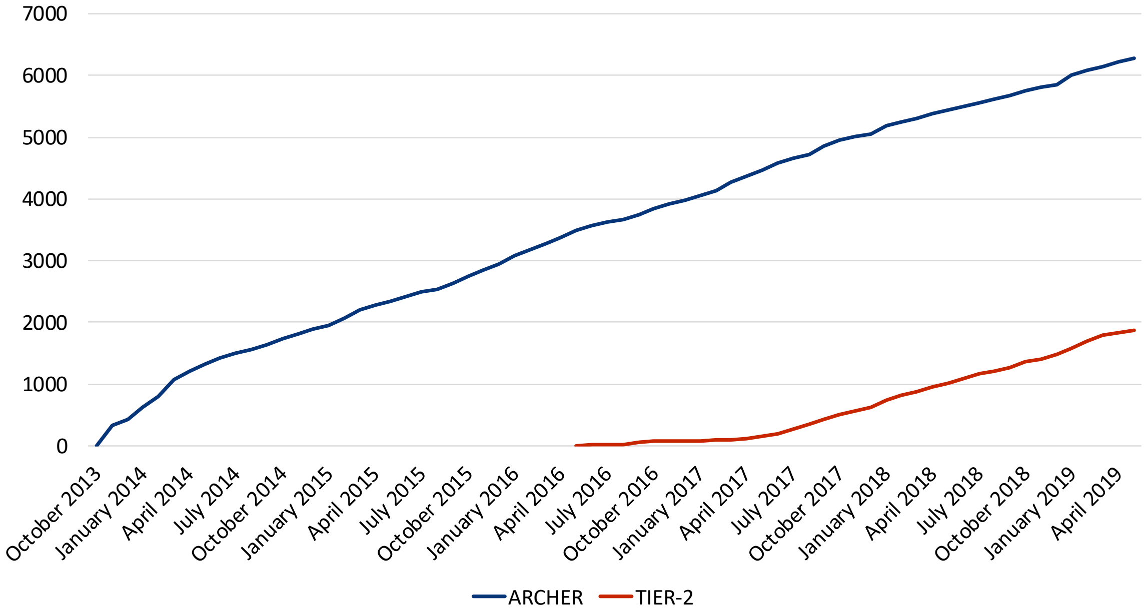
Note: Users on ARCHER include researchers from both EPSRC and NERC and as such NERC has contributed to the operational costs of ARCHER. Similarly, with Tier-2 centres, whilst the operational cost is covered by EPSRC, researchers represent disciplines outside of EPSRC's remit. (1) Includes 166 academic users at non-UK institutions. (2) Includes 1,994 user accounts by UK universities, 170 other UK academic research institutions, 112 accounts by European institutions via PRACE, and 129 accounts by other overseas academic and research organisations. (3) Refers to the Tier-2 system only. (4) GW4 Isambard only became operational recently. As such usage data is not provided at this stage.

Source: EPCC; Tier-2 centres; HECToR Report; Tier-2 annual reports

From the stakeholder consultations it was clear that a significant proportion of respondents felt that the UK was beginning to fall behind in its HPC capabilities, and that more investment was required to improve HPC provision, at all levels including Tier-0. As a result, several academics consulted for this study emphasised that demand for high-end HPC resources already exceeds supply. Given this, the numbers shown in Table 3 may not represent actual demand but the supply-constraint.

Figure 8 shows the number of users on ARCHER and Tier-2 facilities has increased over time. ARCHER overtook HECToR in terms of user numbers in June 2015 and currently has nearly double this number, thereby demonstrating the increased demand for HPC by the scientific community.

Figure 8 User registrations over time

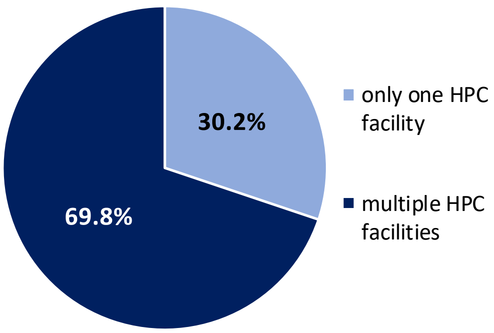


Note: Tier-2 centres came into operation starting from May 2016 (Cirrus). For Tier-2 centres users are only counted once they become "active", i.e. from the date they ran a job for the first time. As such the actual number of users for Tier-2 centres is likely higher. Tier-2 centres include Materials Modelling Hub, Jade, HPC-Midlands, Peta-5/CSD3 and Cirrus. Figures differ from Table 3 as these figures measure registrations not active users.

Source: EPCC

2.2 Facilities used by HPC users

Figure 9 Proportion of HPC users who have used multiple HPC facilities



Note: Based on 233 respondents. One respondent selected Don't know/Not applicable.

Source: London Economics survey of users of HPC capabilities

The user base of the national services is not completely distinct from that of the Tier-2 centres. Rather, each facility forms an important part of the wider HPC ecosystem.

On the one hand, Tier-2 centres have a role to play in preparing users and code for use on the national services. This includes ensuring code is scalable to exploit the extra computing power of the national services as well as gaining experience in using HPC for new users.

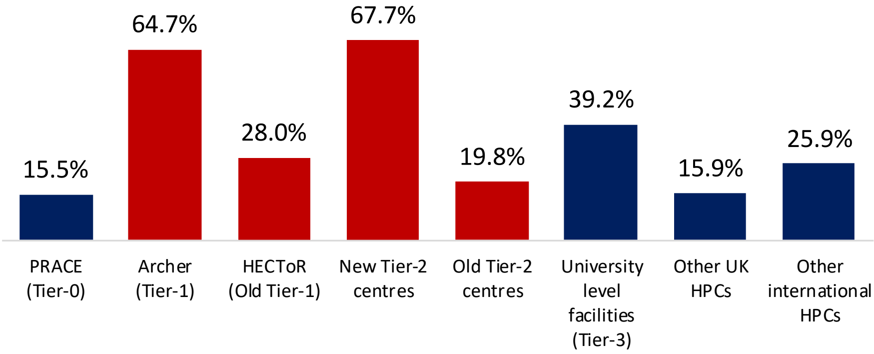
On the other hand, Tier-2 centres also offer a range of different computing architectures. This means different facilities are better suited to solving different types of problems and allows Tier-2 centres to meet different science needs.

Moreover, not all users need the extra computing power that ARCHER provides. Tier-2 centres, as well as local university facilities, thus provide a good way for these users to access the HPC capability they need while ensuring time on the national service is used more efficiently.

The importance of having a multi-facility ecosystem is further highlighted by the large proportion of HPC users using more than one facility. Specifically, of those users that responded to the survey of

HPC users undertaken for this study, **70%** have used more than one facility (Figure 9). In addition to the number of facilities used, Figure 10 further shows the types of facilities used by HPC users who responded to the survey.

Figure 10 Proportion of HPC users using specific HPC facilities, by type of facility



Note: Based on 233 respondents. Respondents could select multiple answers. One respondent selected Don't know/Not applicable. This respondent was excluded from the graph. Facilities highlighted in red are those facilities that are in scope for this study.

Source: London Economics survey of users of HPC capabilities

2.3 Geographical reach of ARCHER

Users of ARCHER are widely distributed across the United Kingdom, and across a wide range of institutions. This shows the diversity of ARCHER usage and ARCHER contribution to UK research and development across the four nations.

Figure 11 presents the number of ARCHER users across the main regions of the UK.

The University of Edinburgh, where ARCHER is based, accounts for the largest numbers of user accounts at 900. This figure also includes user accounts set-up for the students of the MSc in High Performance Computing run by the University of Edinburgh. **London** has the second-highest number of user accounts (**874**), driven by three universities – **University College London** (**372** accounts), **Imperial College** (**271**), and **King's College London** (**154**).

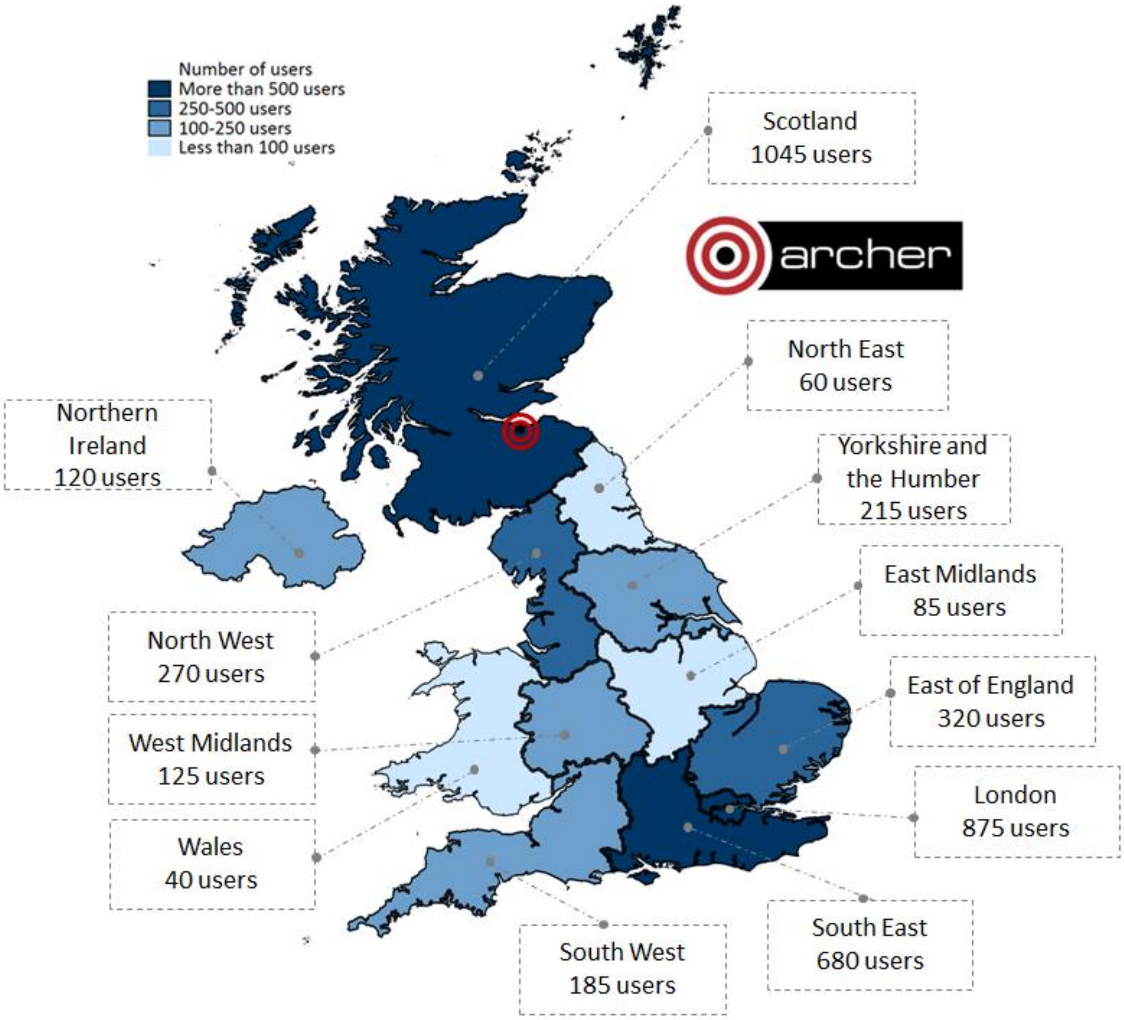
Table 4 presents the top 10 institutions in terms of ARCHER time allocation (Mega Allocation Units). University College London was the highest user in terms of time allocation.

Table 4 Usage of ARCHER: Top 10 institutions

Institution	Usage (MAU)	% of total usage
University College London	11,047	16.6%
University of Cambridge	6,015	9.0%
University of Reading	4,654	7.0%
Imperial College London	4,439	6.7%
University of Edinburgh	3,414	5.1%
University of Oxford	3,217	4.8%
University of Southampton	3,195	4.8%
University of Leeds	2,616	3.9%
King's College London	2,057	3.1%
University of Bath	1,956	2.9%

Note: 1 MAU (Mega Allocation Unit) = 1,000 kAU. A kAU (kilo Allocation Unit) is a standard allocation unit on ARCHER and previous services. On ARCHER: 1 core hour = 0.015 kAUs; 1 node hour = 0.015 x 24 = 0.36 kAUs. ***Source: London Economics***

Figure 11 Number of ARCHER users by UK region



Note: Based on home institution of users. All user numbers are rounded to the nearest 5. Excludes 89 users at other UK academic and research institutions, 17 UK government users, and 112 other commercial users for which no regional information was available.

Source: London Economics based on data provided by EPCC. Contains National Statistics data © Crown copyright and database right 2018.

2.3.1 International use of ARCHER

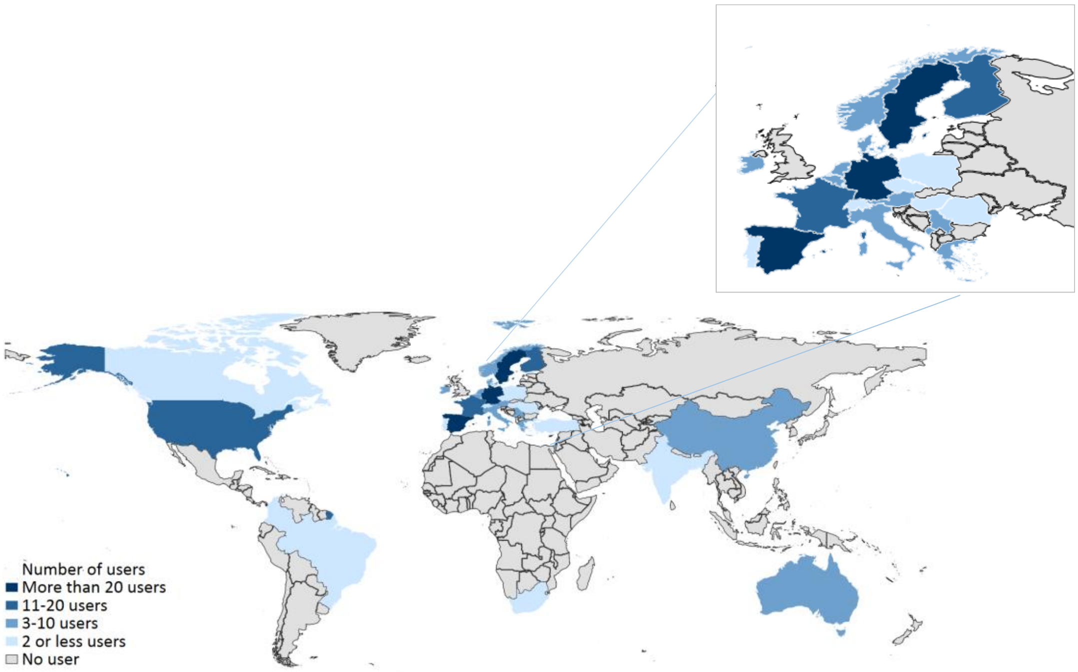
While the main purpose of ARCHER is to support UK scientific research and development, ARCHER's geographical reach is not limited to the UK.

Figure 12 presents the number of ARCHER users across the globe. Countries with the largest number of users are **Sweden** (**23 users**), **Spain** (**22 users**), and **Germany** (**21 users).** However, there are also a number of ARCHER users located at non-European institutions, including the **USA** (**14 users**), **China** (**8 users**), **Australia** (**4 users**), **Israel** (**3 users**) and **India** (**2 users**).

Figure 13 shows the geographical distribution of ARCHER in terms of international usage. While countries such as **China** and **Poland** account for relatively few ARCHER user accounts (**8** and **2** users, respectively), they account for significant usage of ARCHER (**15.8%** and **10.2%** of usage by non-UK institutions, respectively)[26]. Other countries with significant use of ARCHER include **France** (**8.1%**), **Finland** (**7.4%**), **Germany** (**7.2%**) and **Sweden** (**7.0%**).

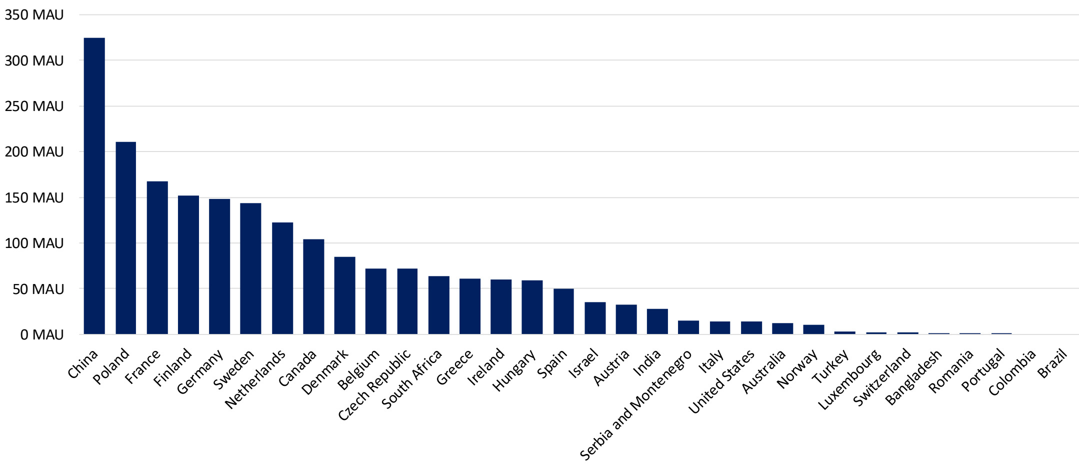
This shows the importance of ARCHER to international collaborations and its contribution to international research and development.

Figure 12 Number of ARCHER users at non-UK institutions



Source: London Economics based on data provided by EPCC. Made with Natural Earth.

Figure 13 Usage of ARCHER by users at non-UK institutions



Notes: 1 MAU (Mega Allocation Unit) = 1,000 kAU. A kAU (kilo Allocation Unit) is a standard allocation unit on ARCHER and previous services. On ARCHER: 1 core hour = 0.015 kAUs; 1 node hour = 0.015 x 24 = 0.36 kAUs. Note that the high usage by Chinese users appears to be a result of a single significant ongoing research collaboration between the UK and China.

Source: London Economics based on data provided by EPCC

[26] Note that the high usage by Chinese users appears to be a result of a single significant ongoing research collaboration between the UK and China.

2.4 Estimating benefits of HPC access

The main purpose of HPC is to support scientific research and discovery. To this end, access to the HPC capabilities is provided free of charge, or at a small nominal cost, at the point of access to academic users. In comparison, access to commercial computing capabilities such as Amazon Web Services or Microsoft Azure can cost anywhere from £ 0.04 to £ 0.10 per core hour (see Annex A3.1), depending on the number of cores required and whether computing time is purchased on-demand or in bulk. Similar, commercial access to GPU capabilities can cost anywhere in the region of £0.45 to £1.91 per GPU hour (see Annex A3.1).

The benefits of access to academic users are estimated as the cost avoided by academics that would otherwise have had to pay commercial rates:

Avoided cost of free access to researchers	=	*Number of core hours / GPU hours facilities were accessed by academics*	x	*Commercial price per core hour / GPU hour*

It should be noted that under the counterfactual, EPSRC would only have to pay access charges for research that could have been undertaken in the absence of HPC. Therefore, the total estimated costs were adjusted for the proportion of research that would not have been undertaken in the absence of EPSRC HPC, obtained from the survey of EPSRC HPC users (70%, see Section 3.1.2):

Adjusted avoided cost of free access to researchers	=	*Direct benefit of free access to researchers*	X	(	1	–	*Proportion of research dependent on HPC (70%)*	)

This is because we assumed that 70% of research would not take place without HPC. Leaving 30% of research that could still take place but would need to access commercial providers.

To estimate the direct benefits of access, avoided cost estimates at the top of the above-mentioned core and GPU hour price ranges were chosen. Specifically, avoided costs of £0.09 per core hour and of £1.66 per GPU hour were used in the estimation[27]. The reason for this is multi-fold:

- The above prices are for cloud services today. However, the assessment is backward looking. Prices for cloud services would likely have been much higher five to ten years ago than the high-end range (£0.09 and £1.66) for avoided cost estimates based on today's prices.
- The cloud prices discussed above are baseline prices for hourly use only. Additional costs for storage, software or other services may have applied. The avoided cost of these services is not estimated separately.
- Tier-1 and Tier-2 centres also offer additional benefits that cloud technologies may not have offered. These include faster network bandwidths, centralised data storage, and,

[27] These figures are based on the average of the highest three cost estimates provided in Annex A3.1.

perhaps most importantly, dedicated support to help users get up and running and ensure users get the most out of the machine.

For these reasons using cost estimates at the high end of today's prices will still yield an underestimate of the actual avoided cost[28].

Please note that the above assessment is backward looking and compares HPC to cloud services five to ten years ago. Technology is constantly evolving and as such the above assessment may not apply to cloud services today. Comparing HPC with cloud at present and weighing advantages and disadvantages is outside the scope of this study.

Combining the number of core / GPU hours accessed by academic users with the above-mentioned commercial rates indicates that **direct benefits of access are in the region of £213.3 million** (Table 5).

Table 5 Direct benefits of access – academic users

HPC facilities	Avoided cost	Adjusted avoided cost
ARCHER	£ 399.7 m	£ 119.9 m
HECToR	£ 151.4 m	£ 45.4 m
New Tier-2 centres*	£ 49.1 m	£ 14.7 m
Regional Tier-2 centres	£ 110.9 m	£ 33.3 m
Total	**£ 711.1 m**	**£ 213.3 m**

Note: Avoided cost represents cost avoided relative to paying commercial providers of all academic research undertaken. To derive the adjusted avoided cost, the avoided cost estimate was adjusted by the proportion of research that would not have been undertaken without HPC. Based on core-hours accessed by academic users supplied by the ARCHER team and the Tier-2 centres. Where data was not available the number of core-hours accessed by academics was approximated. The estimate uses a price per core hour of £0.09 and a price per GPU hour of £1.66. (*) Note that new Tier-2 centres have only been operational for about 1/3 of their lifetime. As such the lifetime benefit of these centres is expected to be significantly higher. Multiplying the estimated benefits by 3 suggests that the avoided cost of new Tier-2 centres will be in the region of £147.3 million over their lifetime, while the adjusted avoided cost will be approximately £44.1 million.

Source: London Economics

It is worth noting that access to commercial computing infrastructure for researchers would likely still be publicly funded, so the cost avoided is approximately equal to the cost avoided by the public purse.

Moreover, in the absence of EPSRC funded HPC systems, it is unlikely that commercial computing providers would provide an adequate substitute for scientific research and discovery (this is explored in further detail in Section 8 and Annex A1.2).

[28] Note that, in addition, more research in the past would likely not have been able to be undertaken without HPC as cloud alternatives were not available. While this would yield a lower direct benefit of access than the one presented here, no spillover benefits of this research would have taken place. As such the estimate of benefits of research (Section 4.1.3) would be lower in this scenario. As spillover benefits of research can reasonably be expected to be larger than avoided costs of access, the overall aggregate economic benefit would thus also be lower in this scenario.

3 Benefits of HPC to research

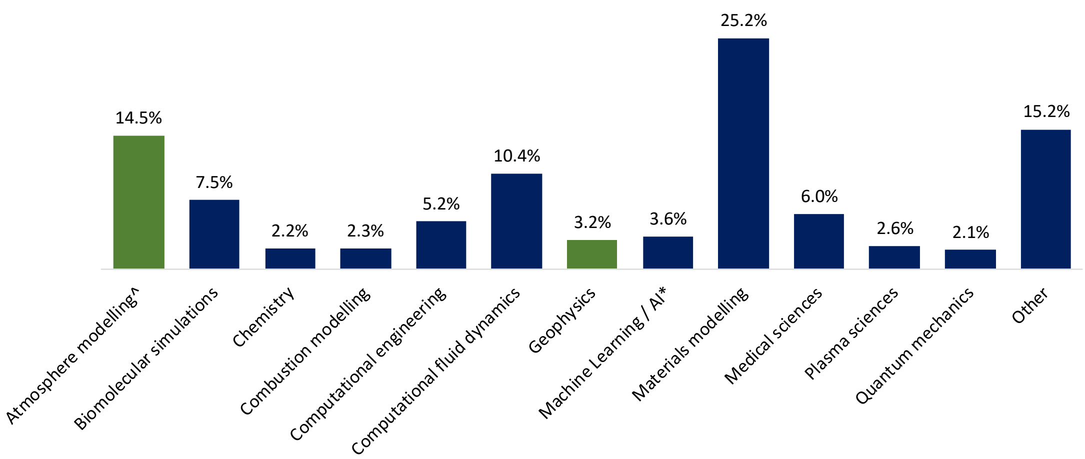
To understand the importance of HPC for modern academic research, a wide range of stakeholders were consulted as part of this study on the impact of ARCHER and Tier-2 centres on their own research (Section 3.1) as well as the benefits to the competitiveness of UK science internationally (Section 3.2).

It is important to emphasise that the benefits mentioned in this section are indicative areas which broadly reflect the HPC user community based on responses from the survey and stakeholder consultations. Specific benefits felt by user groups who were not reached through the consultations would thus not be captured here. Similarly, there might be certain areas which are under- or over-represented. An indicative analysis of the representativeness of the online survey is provided in Section 1.4.

3.1 Benefits to academic research

High Performance Computing is a crucial part of modern research in many fields. Figure 14 highlights the breadth of research areas in which academics undertake research on HPC, based on the primary research area reported by users when subscribing their account. The results show that access to HPC not only benefits research in core engineering and physical sciences but also benefits research within the wider research landscape.

Figure 14 Population of research areas by HPC user accounts



Note: Includes users from both ARCHER and Tier-2 machines. User accounts excludes those which were unknown or used for training and support. User accounts are attached to a project, with a specified research area, and a given researcher can have multiple user accounts for each project they work on. (^) includes climate/ocean modelling as this cannot be distinguished by the EPCC, (*) incl. natural language processing. Research areas marked in green are primarily NERC funded, though some science might be EPSRC or joint-funded. ***Source: EPCC usage data***

HPC allows researchers to simulate very complex problems, to ask more granular and deeper questions, and to obtain more accurate results; it helps improve existing theories; and makes it easier to avoid duplication and increases reproducibility.

"HPC allows me to do my research. Without access to internally competitive HPC very little of my research would be possible." – survey respondent

HPC has two key impacts on fundamental research: it enables research that would not otherwise have happened, and it improves existing research or research that would happen even without HPC. Each of these impacts are discussed below.

A bibliometric analysis of publications arising from EPSRC research is currently underway as a separate study.

3.1.1 Enabling research that would otherwise not have happened

"HPC allows me to tackle otherwise intractable problems in my field." – survey respondent

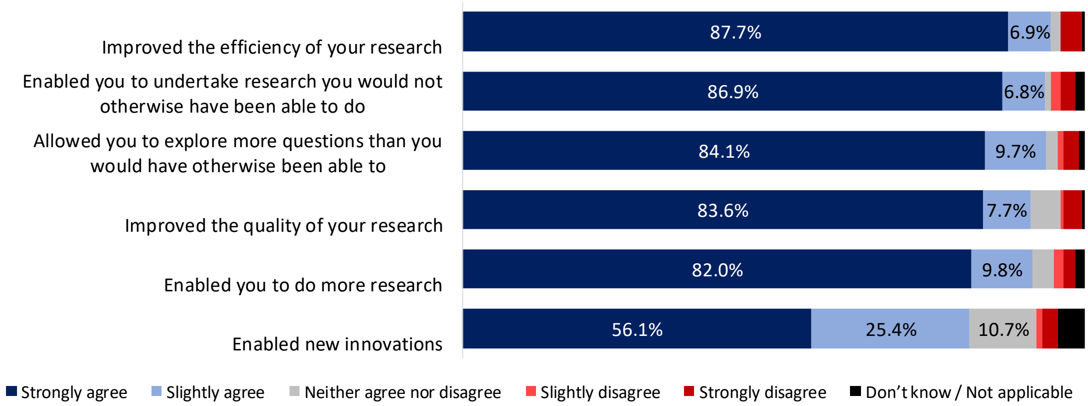
HPC allows researchers to conduct large-scale simulations of otherwise intractable problems, tackling questions that would be impossible to answer using conventional techniques. In a survey conducted for this study, **87%** of respondents strongly agreed that HPC enabled them to conduct research they would not otherwise have been able to do, **82%** that HPC enabled them to do more research and **84%** strongly agreed that HPC enabled them to explore questions they would not otherwise have been able to (Figure 15).

A few examples of the wide-ranging research conducted using HPC include:

	Running simulations on biomolecular systems that would be otherwise too large to simulate, thus allowing the study of the molecular basis of muscle and heart disease to ultimately find new pharmacological therapies and performing large scale detailed patient specific simulations of organs such as heart.
	Performing large scale simulations of computational fluid dynamics (CFD) models to improve the operation of gas-turbine engines.
	Investigating new physics essential to wind and tidal stream turbine design.
	Conducting simulations into wind flow and how this affects the blades of wind turbines to support the design of moveable wings which can harness greater energy yield.
	Investigating the micro-dynamics of ionisation-induced DNA damage.
	Studying the structure and processes of bacteria to examine how they develop resistance to antibiotics to enable the development of new drugs to prevent diseases.

Icon credit: © Microsoft word

Figure 15 'In terms of benefits to your research, to what extent do you agree or disagree that your use of HPC …?'



Note: Based on between 204 and 207 respondents*. **Source: London Economics survey of users of HPC capabilities***

3.1.2 Improving research

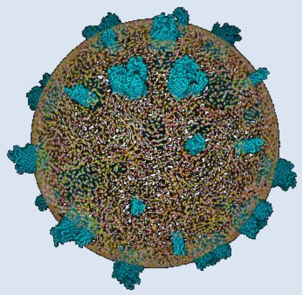
> **"The HPC allows me to conduct research on a much larger scale. The capacity to run hundreds of simulations rather than tens of simulations allows for better optimisation, a larger sample range and for quicker research." – survey respondent**

HPC also helps improve research that could have been undertaken without HPC. **84%** of HPC users surveyed for this study strongly agreed that HPC improved the quality of their research, and **88%** strongly agreed that HPC improved the efficiency of their research. HPC users indicated that HPC allowed them to conduct research faster, more economically, at a more granular level, on a larger scale, "do bigger better science faster", and generally conduct research exceeding "the limitations of desktops and workstations".

To gauge the importance of HPC to its users, survey respondents were further asked what proportion of their research they would be unable to do without HPC. Unsurprisingly, responses varied significantly from researcher to researcher. Nevertheless, on average, researchers said that they would be unable to undertake **70%**[29] of their research without HPC. This further highlights the importance of HPC to the research community.

> **"HPC allows me to complete my research much faster and in much more detail, adding a significant amount of value to the outcomes." – survey respondent**

Box 3 Developing models to study antimicrobial resistance



One of the ways in which antimicrobial resistance occurs is through bacteria evolving in response to the use of medicines designed to kill them. This is happening due, in part, to indiscriminate use of antibacterial treatment and excessive use in agriculture. As a result, infections, which these bacteria cause in humans, animals and plants are becoming more difficult to treat, leading to higher medical costs and ultimately higher mortality rates. At the current rate, by 2050,

[29] Based on the median. The mean was 64%.

infections and illnesses previously curable by antibiotics will kill more people worldwide than cancer[1].

Syma Khalid and her team of researchers at the University of Southampton are one among several groups worldwide, using molecular models and simulations to examine key structures and molecular processes within bacteria to gain insights into the development of resistance. This will eventually facilitate the development of new drugs with which to treat bacterial infections.

The first strand of research focuses on understanding how the bacterial cell envelope functions. Of particular relevance for antibiotic design is to understand how bacteria allow essential nutrients to enter the cell yet reject antibiotics. Furthermore, bacteria also contain efflux pumps, which pump foreign molecules out of the cell, if they do somehow get in. The Khalid group is studying the structures of the membranes and cell wall and also individual proteins such as those that make up the efflux pumps with the initial aim of understanding how they work.

Such understanding requires molecular level computational models which researchers can use to predict how a normal bacterial cell may function. Then these models can be used to predict how potential drug molecules will behave, and eventually such models are likely to inform the rational design of new antibiotics. To tackle this problem via the experimental route alone would be enormously expensive and more time-consuming. If the experiments are guided by the simulations, time and financial costs are reduced. However, running molecular dynamics simulations of bacterial systems requires significant computing power due to the complexity and size of the simulations. As such, ARCHER is vital to solving these problems.

Notes: Picture shows a Coarse-Grain simulation of a vesicle that mimics the *E.coli* outer membrane Khalid et al, Biochemsoc Trans, 2015. ***Source: Interview with Syma Khalid; University of Southampton (nd). Breaking down antibiotic resistance. Using supercomputers and simulations to find solutions and fight infections. Available at: https://www.southampton.ac.uk/news/2018/03/breaking-down-antibiotic-resistance.page [accessed 15/05/2109]; (1) AMR Review (2016). Available at: https://amr-review.org/ [accessed 31/05/2019]***

3.2 Benefits to UK science competitiveness

Computational modelling and simulations have become increasingly important to scientific research and discovery. Modelling and simulations are now fundamental in many areas of research, both in their own right and to support and supplement other methods of research. Indeed, computing has become so fundamental to the research process in many fields that it is often labelled as the third "pillar" of science, alongside theory and experiments.

"In my research area it is vital to have access to an international scale HPC facility, in order to carry out challenging and increasingly realistic simulations. Much interesting and worthwhile science can be done using small clusters, but it tends to be filling in gaps and solidifying understanding, rather than breaking new ground with first-of-a-kind simulations. Performing these highest impact simulations requires world-class HPC facilities." – survey respondent

"HPC is fundamental to scientific research now, many areas of research would be impossible without HPC, including my own (plasma turbulence simulation)." – survey respondent

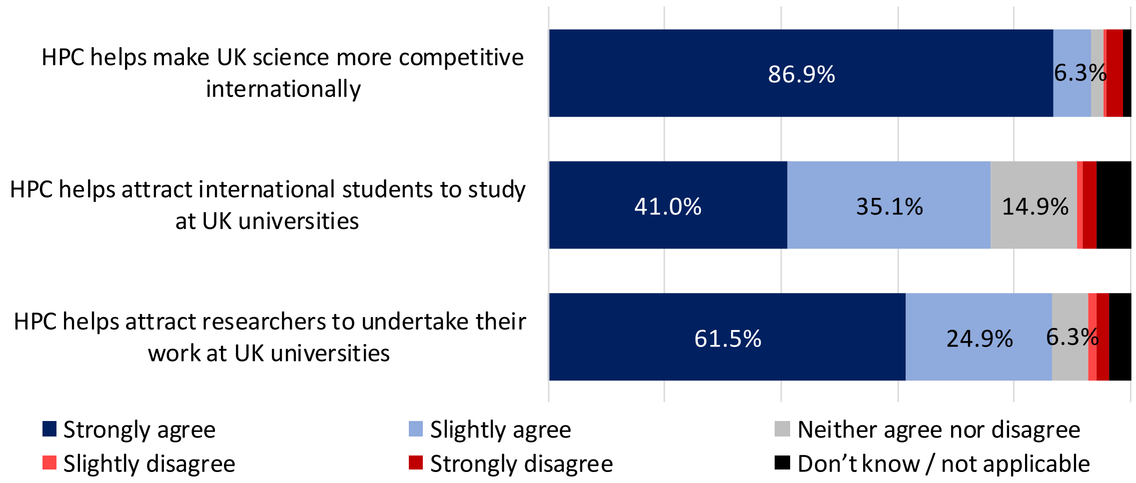
High Performance Computing is the capability that underpins modern computational modelling. HPC allows scientists to perform much more complex simulations than would be possible using a desktop computer; allowing researchers to ask, and find answers to, a much wider range of questions as well as to explore questions in much more depth than otherwise possible.

Given the interdependence between theory, experiments and computational modelling in many areas of modern research and development, investments in High Performance Computing play a crucial role in helping UK science remain competitive internationally. Whilst the focus of EPSRC's funding has been on targeting core Engineering and Physical Sciences research, there are additionally benefits to wider science areas.

The importance of HPC to the competitiveness of UK science in general, and the competitiveness of their own research in particular, was emphasised by a vast number of stakeholders consulted for this study. Of academics who responded to the survey undertaken for this study, **93%** agreed, either slightly or strongly, that HPC helps make UK science more competitive internationally (Figure 16). However, researchers also highlighted the fact that current capabilities are already falling behind other countries and the need for continued investment in order for the UK to remain competitive.

"There are many areas of International science that currently are almost entirely dependent on HPC access. If the UK did not give researchers access to such facilities, we would be completely non-competitive." – survey respondent

Figure 16 'To what extent do you agree or disagree that …?'



Note: Based on between 221 and 222 respondents*. Source: London Economics survey of users of HPC capabilities*

3.2.1 Attracting talent

Having an internationally competitive science base is also important for attracting high quality scientists to the UK, with approximately **86%** of survey respondents agreeing, either slightly or strongly, that HPC helps attract researchers to undertake their work at UK universities. (Figure 16)

Moreover, for approximately **42%** of principal investigators the availability of high quality HPC facilities in the UK played either a key or an important role in their decision to undertake their research in the UK. For a further **21%** HPC played some role. The presence of leading HPC research groups in the UK played either a key or an important role for **44%** of principal investigators, and some role for a further **17%**. (Figure 17)

"The availability of HPC Tier-1 resources (HECToR, then ARCHER) were major factors in my decision to join my current department in the UK. The resources are world-class and easy access to them provides great leverage to produce outstanding science. I would not be in the UK without these resources." – survey respondent

Box 4 Increasing Safety in Autonomous Vehicles

Autonomous driving presents a range of benefits for society including less traffic congestion, as a result of optimal routing, reduced emissions, which could reduce greenhouse gas emissions caused by vehicles by 60%, due to efficient fuel consumption and an increase in safety. Across the world, 1.2 million people are killed in traffic accidents each year. Autonomous vehicles have the potential to eliminate 90% of all traffic accidents, eliminating both drunk and distracted driving accidents.

However, the development of such technology is hindered by the complexity and variety in driving environments. For example, at a busy intersection, a car must interpret stationary elements like traffic lights and lanes as well as respond to moving objects such as pedestrians, cyclists and other vehicles.

One solution comes from utilising deep learning techniques, something being explored by the Torr Vision Group at the University of Oxford using the JADE HPC. They are constructing a framework to predict how stationary and moving elements interact. This framework makes a series of hypotheses about the most likely scenario for a moving object, expanding on current research by assuming a moving object could go anywhere rather than estimate how an object will move from one point to another. The framework is able to evaluate both the context of the scene in addition to the interactions between neighbouring objects. It then makes a strategic prediction about which of several hypothesis is most probable to occur. These predictions have proven highly accurate when compared with real world behaviours.

This research has great benefits on existing technology, as it does not limit the possibilities for what might happen but is instead capable of predicting various future outcomes.

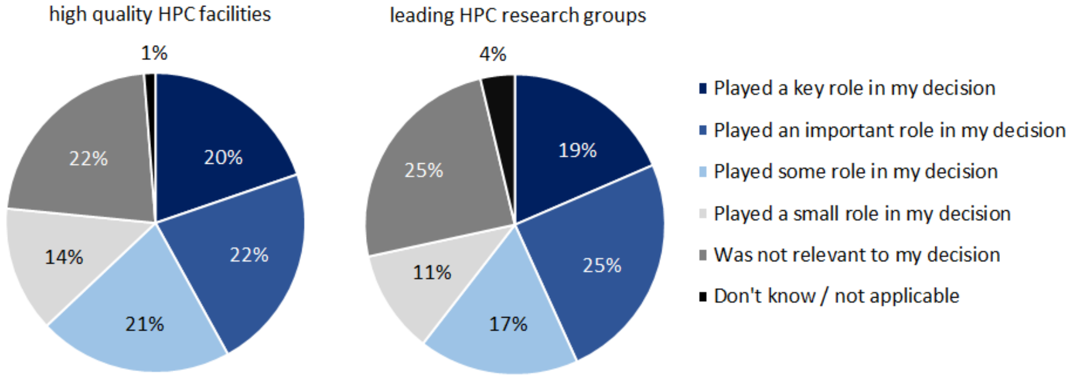
A further research strand, undertaken by the Torr Vision Group, concerns the use of adversarial examples which confuse even state-of-the-art computer vision models, called deep neural networks (DNNs), leading to miscalculation of objects and as a result reducing safety. Adversarial examples work by adding small amounts of white noise to an image which to the human eye is imperceptible, but for DNNs can result in catastrophic failure of perception. It is possible that agents with a malicious intent could utilise these techniques to cause harm – for example, by altering traffic signs with an adversarial example that could confuse DNNs leading to traffic accidents.

The existence of such techniques casts doubts over the safety of using DNNs in driverless vehicles or medical diagnosis due to the possibility of disastrous misclassification. Researchers operating on the JADE HPC have been experimenting with a method, known as semantic segmentation, to reduce the effect of adversarial examples, thus improving the safety of DNNs which can be used in autonomous vehicles.

Both research strands are targeted at increasing the flexibility and accuracy of computational vision, which is vital to improving autonomous vehicle road safety and improving public trust in the capabilities of autonomous vehicles.

Source: Arnab, A. et. al. (2018). On the Robustness of Semantic Segmentation Models to Adversarial Attacks. Available at: http://www.robots.ox.ac.uk/~aarnab/projects/cvpr_2018/Arnab_CVPR_2018_extended.pdf [accessed 07/06/2019]; Ohio University (n.d.). The Future of Driving. Available at: https://onlinemasters.ohio.edu/blog/the-future-of-driving/ [accessed 07/06/2019]; JADE Case Study

Figure 17 'To what extent did the availability of … in the UK play a role in your decision to study, work, or undertake your research in the UK?' - Principal investigators



Note: Based on 81 responses from **principal investigators**. Due to rounding approximations the sum of components may not equal the total. ***Source: London Economics survey of users of HPC capabilities***

> **"The UK's world class computing resources improves the ability for UK-based university and industry researchers to tackle important but compute-demanding problems. This attracts foreign researchers who seek the very best resources. […] My general research area (materials modelling) is thriving in the UK thanks to excellent and dedicated resources." – survey respondent**

HPC also plays a role in attracting international students to UK universities through a number of channels. Most directly, students interested in High Performance Computing may choose to come to the UK to study in a graduate or postgraduate program making direct use of HPC resources - for example the University of Edinburgh's MSc in High Performance Computing – or to undertake a program with a particular research group known to be strong in High Performance Computing. However, HPC also plays a more indirect role in attracting students, by allowing UK researchers to undertake cutting-edge research and thus contributing to the reputation of UK science overall.

Among academics who responded to the survey, approximately **76%** agreed, either slightly or strongly, that HPC helps attract international students to study at UK universities. (Figure 16) The proportion of PhD students and postdocs stating that high quality HPC facilities and leading HPC research groups played an important or a key role in their decision to study or undertake research in the UK was similar to that for principal investigators (not shown here).

4 Wider benefits of HPC

The main purpose of HPCs is to enable scientific research and discovery. As discussed in Section 3, HPC has led to an increase in the quality and productivity of UK research, enabled science not otherwise possible and contributed to the competitiveness of UK science internationally. However, science does not happen in isolation. Indeed, the wider academic literature indicates that investments in intangible assets such as research and development may induce positive externalities[30]. There are many ways in which academic research can induce such positive spillover effects to the UK economy. For example, spillovers are enabled through direct R&D collaborations between universities and other organisations, the publication and dissemination of research outputs, or through university graduates who enter into the labour market.

In light of this, this section explores the wider benefits of HPC. In particular, this section first reviews the existing academic literature on the impacts of research on the UK economy (Section 4.1). It then discusses the impact of HPC on, and the importance of, UK software (Section 4.2). This is followed by a discussion of the impact of HPC on skills and career progression, and the wider benefits of an improved skill base to the UK (Section 4.3). Finally, Section 4.4 discusses the role of HPC in enabling or facilitating collaborations. The impact on UK industry is discussed in Section 5.

4.1 Academic research

Extensive research suggests that academic research, not specifically related to HPC, can have considerable economic benefits, both direct and indirect. This section firstly discusses the economic benefits to broad academic research, not specifically related to HPC, before proceeding to discuss the limited literature which highlights the economic benefits to HPC-specific research.

Box 5 Bone development and response to disease

Researchers at the University of Hull have employed ARCHER to model the development of bones, thereby offering exciting insights into bone biomechanics, as well as improving the understanding of musculoskeletal conditions.

As is often the case, the results of this research have also found applications far beyond the realms of medical sciences. For instance, Professor Michael Fagan is investigating whether cancellous bone-like geometries could serve as the internal supporting structure of wind turbine blades. To that end, ARCHER, alongside local HPC at the University of Hull, will be used to simulate the growth of these structures inside blades. The software, which relies on HPC, will optimise the process such that the blades may withstand the complex loads they typically face while in operation.

Source: Clarke, E and Larmour, I. (2016). The impact of national High Performance Computing. Available at: https://epsrc.ukri.org/newsevents/pubs/impactofnationalhpc [accessed 07/06/2019]; Additional information provided by Prof. Fagan

[30] Economists refer to the term 'externality' to describe situations in which the activities of one 'agent' in the market induces external effects on other agents in that market (where these external effects can be either positive or negative and are not reflected in the price mechanism). In other words, 'an externality is present whenever the well-being of a consumer or the production possibilities of a firm are directly affected by the actions of another agent in the economy'. See Mas-Collell, A., Whinston, M. and Green, J. (1995). Microeconomic theory, New York: Oxford University Press.

4.1.1 Broad academic research

Direct benefits on the economy of broad research include:

- **Innovation:** Academic research can introduce new products, services, methodologies and concepts into markets and the public domain;
- **Higher productivity:** Improved methodologies can reduce the resources required to produce a product or service, improving total factor productivity[31];
- **Commercial spin-offs:** Academic research could be 'spun off' to commercial enterprises, bringing academic innovation to market. Research shows that university-based firms tend to have higher market values than otherwise similar independent firms[32], which suggests high commercial values to academic research and innovation[33].
- **Industry-science collaboration:** Evidence suggests that between two-fifths and one half of EPSRC-funded researchers are collaborative or contract research with industry[34], where commercial collaborators incorporate academic research into commercially-used technology. The literature suggests that academia-industry collaborations can stimulate further private research activity. For example, research shows that firms participating in industry-science collaborations invest more in research and development activity than in comparable situations without such collaborations[35].
- **Knowledge transfer:** Consulting transfers knowledge outside of academia, disseminating skills and methodologies long after specific consulting engagements end.

The impacts mentioned above are private benefits either to organisations conducting research or are felt more widely in the economy. However, research investment can also result in other impacts which can be less easily quantifiable e.g. impacts on health, well-being or security[36]. For example, the Wellcome Trust[37] estimates that cardiovascular research spending generated health gains equivalent to 9% of annual returns.

In addition to direct benefits, research can have **spillover economic impacts**, which arise when economic activities in one part of a market have effects elsewhere in the market. For example, **knowledge spillovers** can arise when skills and techniques arising from research activity are diffused more widely. Haskel and Wallis[38] find that research conducted in UK public research councils significantly filters through to the market sector, suggesting a strong role for knowledge spillovers. As a result, any study which looks only at direct effects of publicly funded research would fail to capture these benefits and investment in R&D may continue to be sub-optimally low. For example,

---

[31] Romer, P. 1990, Endogenous technological change. Journal of Political Economy 98(5), S71–S102

[32] Bonardo, D., Vismara, S., and Paleari, S. (2011), Valuing University-Based Firms: The Effects of Academic Affiliation on IPO Performance, Entrepreneurship Theory and Practice 35(4)

[33] However, the research does not suggest that university-based firms have greater longer-term success than other comparable independent firms.

[34] D'Este, P., and Perkmann, M. (2011). Why do academics engage with industry? The entrepreneurial university and individual motivations, The Journal of Technology Transfer, 36 (2011), pp. 316-339

[35] Czarnitzki, D. (2009). "The virtue of industry-science collaborations," EIB Papers 5/2009, European Investment Bank, Economics Department.

[36] Frontier Economics (2014), Rates of return to investment in science and innovation: a report prepared for the Department for Business, innovation and Skills (BIS)

[37] Wellcome Trust (2009), "Medical Research: What's it worth?", retrieved from: http://www.wellcome.ac.uk/stellent/groups/corporatesite/@sitestudioobjects/documents/web_document/wtx052110.pdf

[38] Haskel J, Wallis G, 2010, Public Support for Innovation, Intangible Investment and Productivity Growth in the UK Market Sector

Jones and Williams[39] find that after modelling social returns on R&D investment, optimal R&D should be at least two to four times actual investment.

Box 6 Using HPC to tackle Alzheimer's disease

Alzheimer's disease is widely accepted to be caused by the biological malfunctioning of a protein (beta-amyloid) disrupting cell membranes leading to the degeneration of brain cells and the onset of Alzheimer's. Yet, our understanding of the mechanisms in which this malfunctioning occurs are limited by experimental techniques. Consequently, there is a need for molecular modelling and simulations to better understand this process, supplementing existing experimental techniques.

Molecular Dynamic simulations are being used by the University of Strathclyde to research how beta-amyloid breaks down (a process known as protein misfold) and aggregates to allow researchers to investigate techniques to prevent the accumulation of this protein. Current projects are focused on comparing simulations with experimental data to ensure that simulated data is accurate before further exploration aims to provide insight which is not observable experimentally.

The number of people with dementia (including those with Alzheimer's) is around 850,000 in the UK and is forecast to increase to over 1 million by 2025 and 2 million by 2051. Any research developments which delay the onset of dementia by five years would halve the number of deaths from dementia, resulting in 30,000 fewer deaths a year. Furthermore, Alzheimer's in the UK is associated with a £26 billion cost per year with two-thirds of the cost of dementia being paid by the dementia-sufferer or their family. These figures demonstrate that this is developing into a growing health crisis which HPC can help to tackle, with obvious social and financial benefits.

Source: Alzheimer's Society: https://www.alzheimers.org.uk; ARCHIE WeSt Stopping Alzheimer's Case Study; Alzheimer's Research UK (n.d.): Statistics about dementia. Available at: https://www.dementiastatistics.org/statistics-about-dementia/ [accessed 07/06/2019]

However, it is difficult to quantify spillover benefits. In addition, when the impacts of an economic activity are felt outside of a transaction, there is a risk of either too little or too much of the activity being performed. For example, consider knowledge spillovers from academic research. When a laboratory or a collaboration conducts academic research, the technology or methods produced may benefit other organisations who did not perform or pay for the R&D. The institution conducting the research takes into account the costs they incur or their private benefit but does not incorporate the wider benefits on other organisations. Therefore, the research organisation may underinvest in research activity[40].

Therefore, there is an economic rationale for government to stimulate R&D activity beyond the levels that might occur if industry and research organisations were left to their own devices.

Knowledge spillovers can arise through a number of different channels:

- **Interactions between researchers and industry:** A key channel of knowledge transfer is through interactions between academic colleagues as well as with industry, through

[39] Jones, C., and Williams, J. (1998), Measuring the social return of R&D, The Quarterly Journal of Economics, 113(4), pp. 1119 – 1135.
[40] However, research organisations could engage in optimal R&D activity levels if there was a mechanism for research institutions to internalise the spillover benefits of their research activities (e.g. by benefitting organisations bargaining with research organisations to use/license the technologies they create).

seminars, conferences, collaborations or more informally through day-to-day interactions and discussions.

- **Labour mobility:** Researchers moving institutions or moving to industry, either changing where they work or via secondments, take with them knowledge and skills which can be used in new collaborations and dissemination in their new institutions. The same effect can happen with undergraduate or graduate students, who then change educational institution or apply their research in industry.
- **Applications of published research in academic or commercial research:** academic research builds on an existing body of knowledge, either empirical or theoretical. This means that stimulating research today enables current and future academic research, potentially in completely unthought-of areas. Commercial research can also benefit from applying academic research: Cohen et al.[41] conduct a survey of R&D managers at US firms and find that publications of academic research are the dominant method for transfer of information from public to private research. Frontier Economics[42] find that the most consistent benefits to private sector productivity stem from science-based, applied research council investment[43]. In fact, a 10% increase in university research spending has been found to have an approximately 1% increase in firm patents[44].

Academic research can also lead to market and network spillovers[45]. **Market spillovers** occur when buyers of a new product or a product made with new processes receive some of the benefit as the market may under-price the good relative to the benefits it produces (i.e. firms may not be able to fully capture the consumer surplus in the form of higher prices). **Network spillovers** occur when the value of participating in an activity depends on the number of other participants. For example, there is no benefit from being the only possessor of a telephone in an area. But having a telephone becomes more useful as more and more people are connected. However, without being able to predict the likely number of other participants in a network, people are not willing to make an up-front investment. This is known as a co-ordination problem[46] and can lead to under-investment. However, funding from sources such as the government can incentivise unilateral investments, helping to overcome the coordination problem. Academic research can then diffuse to related fields, increasing the economic value of not only funded research but innovations in other fields.

Public research and development investment also **stimulates private sector R&D**, thereby increasing total R&D spend and potentially further boosting the benefits of economic growth. For example, Haskel et al.[47] find that public sector funding of science is consistent with "crowding in" of private sector investment i.e. private sector and public sector R&D investment are frequently complements, not substitutes. Other studies (e.g. Jeaumotte and Pain[48]) confirm this finding.

[41] Cohen, W.M., Nelson, R.R., Walsh, J.P., 2002. Links and impacts: the influence of public research on industrial R&D. Management Science 48(1), 1-23.

[42] Frontier Economics (2014), Rates of return to investment in science and innovation: a report prepared for the Department for Business, innovation and Skills (BIS)

[43] Note, however, that it is very difficult to causally attribute benefits, and further that research and development impacts are not easily captured and may occur with a significant lag. For example, Adams (1990) finds that the lag between academic science investment and productivity growth is around 20 years. See Adams, J. (1990), Fundamental Stocks of Knowledge and Productivity Growth, Journal of Political Economy Vol. 98, No. 4 (Aug. 1990), pp. 673-702

[44] Jaffe, A.B., 1989. Real effects of academic research. American Economic Review 79(5), 957-970

[45] Ibid.

[46] Farrell, J., and Klemperer, P. (2007), Coordination and lock-in: competition with switching costs and network effects, handbook of Industrial Organisation, Vol(3) Ch.31, edited by M. Armstrong and R. Porter

[47] Haskel J, Haskel J, Hughes A, Bascavusoglu-Moreau Eet al., 2014, The economic significance of the UK science base: a report for the Campaign for Science and Engineering, Publisher: Imperial College Business School

[48] Jeaumotte, F. and Pain, N. (2005) 'Innovation in the Business Sector', OECD Economics Department Working Paper No.459, OECD, Paris

Note, however, that academic research can also lead to negative spillovers such as obsolescence. Research could create new ideas, technologies, products and services, rendering existing solutions obsolete. As a result, firms' investment or intellectual property could lose its value and firms could lose out from innovation. This phenomenon is known as creative destruction[49].

Box 7 Desalination Processes using Molecular Dynamics

According to the WHO/UNICEF, 844 million people don't have access to clean water and the United Nations finds that 40% of the world's population are affected by water scarcity. Therefore, developing large-scale technologies to convert seawater and contaminated water into potable water is urgently needed.

Tools to understand and model the non-equilibrium fluid flows that this problem presents are being developed at the University of Strathclyde using the ARCHIE-WeSt High Performance Computer. In particular, next generation semi-permeable membranes are being studied to allow the engineering of more efficient membrane designs which act as molecular sieves to deliver clean water, using a smaller energy footprint than is currently possible.

HPC facilities are required for this research as molecular dynamics techniques are necessary to study the systems of liquid flow through the nanotubes of the molecular sieves; such techniques demand greater processing power. Without such facilities, this research would not have been possible.

This research has the potential for enormous social benefits both in the UK and across the world, where up to 433 million school days are lost every year because of water-related illnesses and a new-born dies every minute from infection due to lack of clean water. Along with the social benefits, large economic gains would also be realised, with an expanded labour-force and greater labour productivity, permitting a higher standard of living for the world's citizens.

Source: WHO/UNICEF Joint Monitoring Programme (2017); United Nations, Sustainable Development Goals; ARCHIE-WeSt Desalination Case Study; Water Aid (n.d.). Facts and statistics. Available at: https://www.wateraid.org/facts-and-statistics [accessed 07/06/2019]

4.1.2 HPC-specific academic research

All of the benefits mentioned above are applicable to HPC-specific academic research. Particular channels through which HPC knowledge spillovers occur, and for which data is available through EPSRC's research outcome system are listed below:

- **Interactions between researchers:** Researchers can share knowledge of HPC (and particular benefits to their research through HPC) more widely, and colleagues can employ HPC in novel ways, generating new (or improving existing) techniques, products or insights. This might happen through the HPC consortia which exist, which allow researchers to collaborate and interact with other academics in their field. Furthermore, outreach and other HPC-related events provide an opportunity for researchers to interact with one another and share research.
- **Labour mobility:** 168 secondments for HPC researchers were reported in EPSRC's research outcome system, 86 within academia and 79 to the public or private sector[50]. Researchers

[49] Schumpeter, J. (1942), Capitalism, socialism and democracy, New York: Harper, 1975 (original publication 1942), pp. 82 – 85.
[50] Destinations for the remaining 3 secondments were not specified.

moving institutions take with them the knowledge of HPC and skills they have acquired (or acquire new knowledge when they visit other institutions), which can be used in new collaborations and disseminated in new institutions. This in turn can spark novel forms of research using HPC, new ideas or new techniques, in new institutions, new sectors and new countries.

- The effect is not restricted to HPC researchers themselves: graduate or undergraduate students under HPC researchers' guidance can benefit from newly-acquired knowledge, skills and techniques, which they can then disseminate to other academic, public or private organisations.

There is limited academic research analysing the benefits of HPC-specific research for the overall economy. However, one study, in the US, finds that investment into HPC is associated with a contemporaneous increase in the number of academic publications[51]. Another study, by the same lead author, find that, in the US, locally available HPC resources enhance the technical efficiency of research output in subject areas related to Chemistry, Civil Engineering and Physics, amongst other subjects[52].

4.1.3 Estimating productivity spillovers of academic research

Published estimates of productivity spillovers from higher education research

A small body of research literature provides estimates of productivity spillovers from academic R&D. A study by Haskel and Wallis[53] investigates evidence of **spillovers from public funding of Research & Development**. The authors analyse productivity spillovers to the private sector from public spending on R&D by the UK Research Councils[54].

Haskel and Wallis find strong evidence of the existence of market sector productivity spillovers from public R&D expenditure originating from UK Research Councils[55, 56]: the marginal spillover effect of public spending on research through the **Research Councils** stands at **12.7 (i.e. for every £1 spent on university research through the Research Councils results in an additional output of £12.70 in UK companies)**. The analysis also suggests that the spillover benefits of public spending on research in higher education are greater than those from other R&D areas supported by government.

A more recent study by Haskel et al.[57] provides additional insight into the size of potential productivity spillovers from university research. Rather than estimating effects on the UK economy as a whole, the authors analyse the size of spillover effects from public research across different UK industries[58]. The authors investigate the correlation between the combined research conducted by the Research Councils, the higher education sector, and central government (e.g. through public

[51] Apon et al. (2010), High Performance Computing Instrumentation and Research Productivity in US Universities, *JITI Journal of Information Technology Impact*

[52] Apon et al. (2015). Assessing the Effect of High Performance Computing Capabilities on Academic Research Output, *Empirical Economics*

[53] Haskel, J., & Wallis, G. (2010). *Public support for innovation, intangible investment and productivity growth in the UK market sector*.

[54] The authors use data on government expenditure published by the Department for Business, Innovation and Skills for the financial years between 1986-87 and 2005-06.

[55] Based on regression of total factor productivity growth in the UK on various measures of public sector R&D spending.

[56] Note that the authors' regressions only test for correlation, so that their results could be subject to the problem of reverse causation (i.e. it might be the case that increased market sector productivity induced the government to raise public sector spending on R&D). To address this issue, the authors not only test for 1-year lags, but for lags of 2 and 3 years respectively, and obtain similar estimates. The time lags imply that if there were a reverse causation issue, it would have to be the government's *anticipation* of increased total factor productivity growth in 2 or 3 years which would induce the government to raise its spending on research; as this seems an unlikely relationship, Haskel and Walls argue that their results appear robust in relation to reverse causation.

[57] Haskel, J., Hughes, A., & Bascavusoglu-Moreau, E. (2014). *The economic significance of the UK science base: a report for the Campaign for Science and Engineering.*

[58] Haskel et al. (2014) use data on 7 industries in the United Kingdom for the years 1995 to 2007.

research laboratories)[59], interacted with measures of industry research activity, and total factor productivity within the different market sectors[60]. Their findings imply a total rate of return on public sector research of **0.2 (i.e. every £1 spent on public R&D results in an additional output of £0.20 within the UK private sector).**

Further quantitative evidence on the return to public investment in scientific research is limited. An overview of the limited number of other studies that look at returns to public R&D investments can be found in a 2014 Frontier Economics report on returns to investment in science and innovation[61] as well as in in 2015 Economic Insights paper on the relationship between public and private investment in science, research and innovation[62]. In addition, the above-mentioned papers are, to the best knowledge of the authors, the only papers that specifically look at returns to science funding via the research councils.

It is important to note that both studies examine productivity spillovers of research across a wide range of research areas and are not specific to HPC. HPC research is likely to generate additional productivity spillovers not captured by these estimates. However, research specifically on productivity spillovers of HPC research is even rarer, though a recent study by the IDC[63] suggests that the return from investments into academic HPC projects may be significantly higher: among the European HPC projects analysed, **every $ invested in academic HPC projects generates a return on investment of $30**. Though it should be noted that the study is based on a small sample of only seven academic HPC projects. Moreover, the study only includes success stories, not projects that didn't generate economic or scientific results. This likely biases the results, resulting in inflated HPC estimates.

Given the limited evidence on productivity spillovers by research council and the observations discussed above, the analysis in this study uses the two papers by Haskel et al. for the low estimate, and the IDC paper for the high estimate.

Estimating productivity spillovers

Based on the findings in the literature, productivity spillover multipliers were applied to the different items of research-related funding to estimate the productivity spillovers associated with research undertaken on EPSRC funded HPC infrastructure (Table 6):

[59] A key difference to the multiplier estimate for Research Council spending in Haskel and Wallis (2010) lies in the distinction between *performed* and *funded* research, as outlined by Haskel et al. (2014). In particular, whereas Haskel and Wallis estimated the impact of research *funding* by the Research Councils on private sector productivity, Haskel et al. instead focus on the *performance* of R&D. Hence, they use measures of the research undertaken by the Research Councils and the government, rather than the research funding which they provide for external research, e.g. by higher education institutions. The distinction is less relevant in the higher education sector: to measure the research performed in higher education, the authors use Higher Education Funding Council funding (where research is both funded by and performed in higher education).

[60] The authors regress the three-year natural log difference of total factor productivity on the three-year and six-year lagged ratio of total research performed by the Research Councils, government and the Higher Education Funding Councils over real gross output per industry. To arrive at the relevant multiplier, this ratio is then interacted with a measure of co-operation of private sector firms with universities and public research institutes, capturing the fraction of firms in each industry co-operating with government or universities. The lagged independent variables are adjusted to ensure that the resulting coefficients can be interpreted as annual elasticities and rates of return.

[61] Frontier Economics (2014). *Rates of return to investment in science and innovation*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/333006/bis-14-990-rates-of-return-to-investment-in-science-and-innovation-revised-final-report.pdf [accessed 01/08/2019]

[62] Economic Insight (2015). *What is the relationship between public and private investment in science, research and innovation*? https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/438763/bis-15-340-relationship-between-public-and-private-investment-in-R-D.pdf [accessed 01/08/2019]

[63] IDC (2014). *EESI-2 Special Study To Measure And Model How Investments In HPC Can Create Financial ROI And Scientific Innovation In Europe*

- the multiplier of **12.7** was assigned to the research funding that researchers received directly from the EPSRC as well as further funding received from other UK Research Councils; and
- the multiplier of **0.2** was used for all other research funding received[64].

Table 6 EPSRC research funding, in 2018 prices

Funding type	Type / source	Value	Multiplier
EPSRC direct research funding	Consortia grant	£ 13.0 m	12.7
	Fellowship	£ 13.6 m	
	Research grant	£ 163.1 m	
Non-EPSRC further funding	Academic/University	£ 3.6 m	0.2
	Charity/Non-profit, and learned societies	£ 1.4 m	0.2
	Other public (UK)	£ 9.1 m	0.2
	Other public (non-UK)	£ 57.5 m	0.2
	Private	£ 4.2 m	0.2
	UK Research Councils (non-EPSRC)	£ 12.5 m	12.7

Note: Analysis based on EPSRC's research outcomes system. Direct funding only includes funding by EPSRC. Only further funding reported to EPSRC is included, and as such the true figure may be greater.
Source: Source: London Economics based on Researchfish™ data provided by EPSRC

A loss of HPC capabilities would have significant impact on the work undertaken by UK researchers. A large proportion of researchers using HPCs said they would be unable to undertake their work at all or undertake work of the same quality (see Section 8). However, HPC is not the only factor determining research output. To capture the proportion of benefit that is attributable to HPC investments, the average proportion of research which could not be undertaken without HPC, reported by users of HPCs (**70%**, see Section 3.1.2), is used to adjust the estimated productivity spillovers.

Applying these productivity spillover multipliers to the research-related funding, and adjusting for attribution, indicates that research relying on HPC capabilities results in **total market sector productivity spillovers of approximately £1.8 billion**. This represents a return on investment of approximately **9.5** associated with HPC research: **for every £1 invested in HPC research activities by the EPSRC, an additional economic output of £9.5 is generated across the UK economy**.

It should be noted that this is a conservative estimate of the productivity spillovers generated. Estimates of funding are based on EPSRC's research outcomes systems, which records quantitative information only for:

- research activities funded directly by the EPSRC, and
- further funding given to EPSRC grant holders by other bodies, *if* these are reported to the EPSRC.

[64] In terms of the large difference in magnitude between these multipliers, explaining the size of the 12.7 multiplier in particular, Haskel and Wallis (2010) argue that they would expect the productivity spillovers from Research Council funding to be large, 'given that the support provided by Research Councils is freely available and likely to be basic science'. To the best knowledge of the authors, there exists no further and recent empirical evidence to support this. As a result, we apply the separate multipliers to the different income strands.

Research activities using HPC systems funded by other UK Research Councils, other public bodies, or private organisations are not included in the estimate *unless* the funding was fed back into EPSRC's research outcomes systems.

Moreover, as discussed in Section 4.1, HPC research is likely to generate additional productivity spillovers not captured by the general literature on spillover benefits of academic research. Indeed, using the IDC's estimate of ROI discussed in Section 4.1.2, indicates that **returns of HPC research activities may be as high as £5.8 billion**, still counting only funding reported back to the EPSRC.

4.2 Software development and support

Alongside hardware and people, software forms a crucial component of the HPC ecosystem. Developments in software allow more efficient use of HPC hardware and thus allow researchers to perform more, or more complex, simulations on the same machine. As such, without optimised software that can exploit the thousands and thousands of cores of modern HPC systems, many of the benefits that HPC brings would not be realisable.

Throughout the stakeholder consultations undertaken for this study, the internationally leading position of UK HPC software was repeatedly stressed. Consistent with consultations undertaken for the previous HECToR impact assessment[65], the limited availability of HPC resources in the UK, forcing users to use existing resources as efficiently as possible, was repeatedly cited as a reason for this world-leading status. Another key characteristic cited is the **collaborative nature of UK scientific software development**.

> **"The funding provided by EPSRC covers my post and a post-doc over a five year period. The theme running through the proposal is to push the scalability of software on the very highest end hardware, and so is very much based in High Performance Computing […]. We will be developing a small set of codes to better utilise HPC, both in terms of scalability and in terms of solving scientifically exciting problems." – EPSRC RSE fellow**

Over the last few decades, High Performance Computing has contributed to the development of communities and support structures around software in the UK. One such community is the Research Software Engineering (RSE) community. Research software engineers are trained scientists that are also specialists in computing and work at the intersection of science and computing. Research software engineers play a vital role in creating efficient scientific software capable of exploiting modern HPC hardware. Accurate estimates on the number of research software engineers in the UK are not currently available. The 2017 Research Software Engineers: State of the Nation Report suggests that there are at least 1,000 research software engineers in the UK – though the number of software development roles in UK academia may be as high as 14,000 in 2015/16[66].

The importance of software has long been recognised by the EPSRC, with the EPSRC having set out a strategic framework for investment in software in their 2012 software as an infrastructure strategy[67]. Over the last five years, EPSRC has invested approximately **£9 million per annum** in software via a number of activities[68], many of which also benefit the HPC community.

[65] EPSRC (2014). *The impact of HECToR*
[66] The Research Software Engineer Network (2017). *Research Software Engineers: State of the Nation Report 2017*
[67] EPSRC (2012) Software as an infrastructure. Available at: https://epsrc.ukri.org/newsevents/pubs/software-as-an-infrastructure/ [accessed 11/04/2019]
[68] EPSRC website:https://epsrc.ukri.org/research/ourportfolio/themes/researchinfrastructure/subthemes/einfrastructure/software/ [accessed 11/04/2019]

Firstly, recognising the importance of research software engineers, EPSRC has invested in a number of RSE fellowships, providing funding for research software engineers for a period of up to five years.

Box 8 EPSRC Research Software Engineer Fellowships: Case study of Dr. Ian Bush

Dr. Ian Bush holds a 5-year post as an EPSRC Research Software Engineer (RSE) Fellow at the University of Oxford where he is researching methods to extend the scalability of software on HPC to fully exploit the current hardware capabilities. Fundamentally, this involves developing codes which are better able to utilise HPC, both in terms of scalability and in terms of solving scientifically exciting problems.

Ian has a background as a chemist and condensed matter physicist. He got involved in coding during his doctorate and thrived on the challenge of developing methods and algorithms to solve problems. This led him to move towards supporting and developing parallel computing for what is now the STFC Daresbury Laboratory. Since then Ian has been working on the development, optimisation and support of a number of software packages in the fields of materials science and chemistry.

One exciting project that Ian is working on is in trying to understand the dynamics of X-rays and their effects on materials. Further research on the properties of X-rays and their effects on materials are needed. One piece of software used to analyse these problems is CRYSTAL, which starts from first principles to define properties of different materials. Ian is working to improve the internal code of CRYSTAL so that it can more effectively utilise the large number of cores that HPC offers. This allows researchers using the software to improve the speed of their research and therefore the pace of writing academic papers. Moreover, it allows problems to be scaled up and makes possible research strands that would otherwise not be possible.

Another project that Ian is working on is a new collaboration with the Culham Centre for Fusion Energy who are examining the possibility of using hydrogen fusion power as a cheap and plentiful supply of energy. As part of this work simulation HPC is used to model how the plasma in a tokomak works. As plasma melts and evaporates very quickly, experimentation is not possible. Work over the past year by Ian and his group has resulted in a speed up of 20-30% in the code used to model plasma within the tokomak, at the core counts currently employed, and this improvement will increase with the number of cores used. Future work will involve speeding up the code further and allowing the software to be scaled up so that it can solve bigger problems using existing hardware capabilities.

HPC enables scientists to examine a wide range of issues they would not otherwise be able to investigate. However, it is important to recognise that HPC is not just about the hardware. Developments in software, which allow HPC hardware to be used more efficiently, are equally as important. EPSRC's investments in software, for example through RSE fellowships, computational software engineering (CSE) support, and collaborative computational projects (CCPs), are crucial to support Ian and others in maximising the value of the UK's HPC infrastructure.

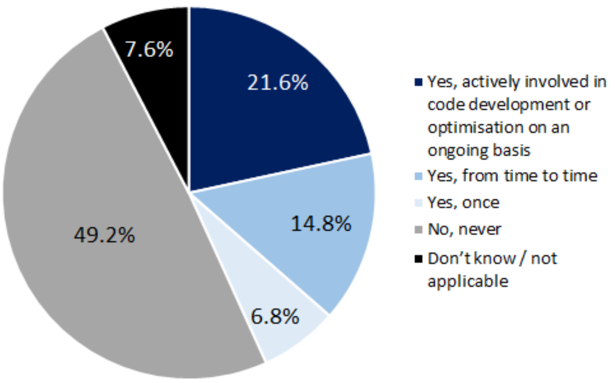
Source: London Economics based on interview with Ian Bush

Investments in software development, through Computational Software Engineering (CSE) support services, also form a key part of the EPSRC's HPC investments. In particular, funding to develop and optimise software is provided to the ARCHER community via the embedded CSE (eCSE) programme. Stakeholder consultations emphasised that this scheme, which is unique to the UK, is very important for the development of software which aids researchers writing scientific papers. As of December 2018, benefits of eCSE projects reached **£24.5 million** based on efficiency improvements of 91 of 100 EPSRC/NERC eCSE projects completed thus far[69]. Given an overall cost of £6 million, this represents a **benefit to cost ratio of over 4:1**. This figure is likely to increase over time as further benefits from improved software coding materialise, so this impact should be considered a conservative estimate.

It should be noted that these benefits represent only the direct benefits of optimised software, based on the time savings to users of the software. In addition, the additional research made possible through optimisation of the software will yield further, potentially much larger, benefits.

Moreover, these benefits only account for software optimisations made via the 91 evaluated eCSE projects. Software developed or optimised by HPC outside of the eCSE programme yield further benefits for the wider HPC community.

Figure 18 'Did you develop or optimise code that benefited the wider HPC community?'



Note: Based on 236 responses.

Source: London Economics survey of users of HPC capabilities

Overall, more than **40% of HPC users** responding to the survey undertaken for this study **developed or optimised code that benefitted the wider HPC community**; with over **20%** actively involved in code development or optimisation on an ongoing basis (Figure 18).

In addition to eCSE support, the EPSRC also supports Collaborative Computational Projects (CCPs). CCPs are flagship code development projects that build on a collaborative software development approach by bringing together leading expertise in research and computing to tackle innovative large-scale software development projects. There are currently 17 active CCPs in a wide range of areas from biomolecular simulations to computational plasma physics[70].

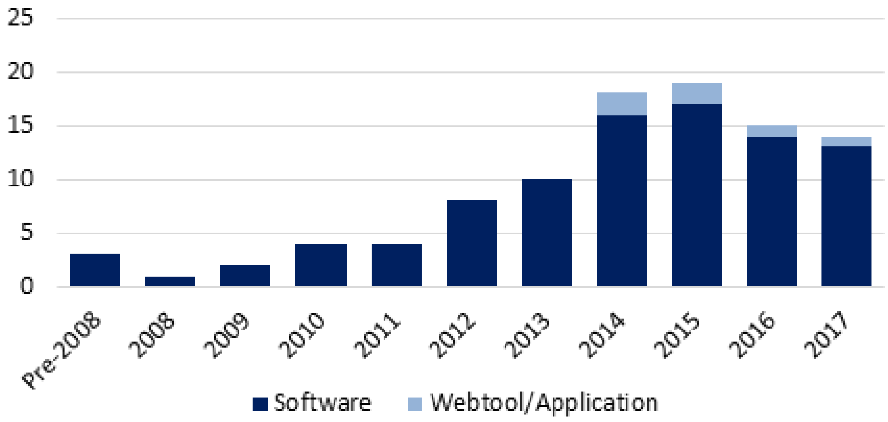
Analysis based on EPSRC's research outcomes system, indicates that **HPC supported development of 8 new software products per year on average** over the last decade[71]. In addition, HPC also underpinned continuous development of many of these software products since their release.

[69] It should be noted that around 22% of ARCHER users are funded by NERC and so NERC funding will be included in eCSE projects. Data is provided by the EPCC.

[70] See http://www.ccp.ac.uk/current.html [accessed 10/04/2019] for a list of current CCPs.

[71] Note that this does not include NERC. Only software products which were reported to EPSRC are included, and as such the true figure may be much greater.

Figure 19 Software development by HPC users reported to EPSRC



Note: Figure shows the number of software and technical products developed, linked to HPC, that have been made public. Analysis based on EPSRC's research outcomes system and does not include NERC. Only projects which were reported to EPSRC are included, and as such the true figure may be much greater. Where multiple versions of a software were funded or a software was funded in multiple years, only the first versions / year in which funding was provided is shown in the graph. Pre-2008 represents software development by HPC users reported to EPSRC prior to HECToR. (*) Only partial data for 2018 was available at the time of reporting, so this year is excluded. ***Source: London Economics based on Researchfish™ data provided by EPSRC***

Publication of new software products increased significantly from 2012 onwards, reaching more than 15 software products a year in 2014 and 2015. This rise coincides with the opening of the Tier-2 centres in 2012, though it may also be an artefact of better data collection[72].

The vast majority (70%) of software products developed were published under an open source license, allowing other researchers to access and modify these software products free of charge[73].

4.2.1 Valuing benefits to software

If software projects shown in Figure 19 generate a similar benefit to that of eCSE projects generated, the additional benefits, in terms of freeing up resources for additional science, could be in the region of **£28.9 million**[74]. However, this number significantly underestimates the true value of improved HPC software to the UK economy for several reasons:

1) Only software projects reported in EPSRC's research outcome system are counted. However, with more than 40% of respondents to the user survey suggesting that they have developed or optimised code that benefitted the wider HPC community, the number of software projects not captured could be very large.
2) The benefits only capture the benefit to science in terms of being able to do more science on existing HPC hardware. The true benefits of improved or new software is the contribution that this software makes to scientific research & discovery. Note that these benefits are implicitly monetised when estimating the spillover benefits of HPC research (Section 4.1.3), therefore software is excluded in the calculation of total aggregate benefits.
3) HPC software may also bring significant benefits to industry. This is highlighted by CASTEP (see Box 9), an HPC simulation software used to calculate the properties of materials from

[72] EPSRC started collecting outcomes data through Researchfish™ from 2014, prior to that outcomes for EPSRC were collected through a system called ROSS.
[73] It should be noted that the exact conditions of what users are allowed to do with the open source software depend on the details of the license, though most open source software is free and allows modification by users.
[74] 107 software projects between 2008 and 2018 * [(£24.5 million / 91 eCSE projects)]

first principle, which is used by around 900 industrial users. As the Goldbeck report[75] indicates, companies using materials modelling software such as CASTEP, among others, achieved cost savings ranging from €100,000 to €50 million, with an average return on investment of 8:1. Benefits of software to industry were not quantified in this study.

Box 9 CASTEP

CASTEP is a simulation software used to calculate the properties of materials from first principles. The software applies density functional theory to simulate the atomic-level structure as well as a wide range of other properties of materials, and thus enables its users to gain a deeper understanding of the properties of materials.

CASTEP is free for academic use and is used by a wide range of academics, from theoretical physicists to experimental scientists, in order to answer research questions in fields such as molecular dynamics, semiconductors and liquid crystal displays. In addition to academic use, CASTEP is also used by around 900 industrial customers around the world. Total revenue from sales of CASTEP is in excess of $40 million, while the last reported annual sales revenue was in the region of $3 million. CASTEP has also been cited in over 260 patents.

As simulations can be cheaper and more flexible than experiments, CASTEP can help reduce costs and speed up product development for its customers. Indeed, according to a recent EU study, companies using materials modelling, including CASTEP among others, achieved cost savings ranging from €100,000 to €50 million, with an average return on investment of 8:1. Commercial applications of CASTEP include integrating organic electronic materials for light-weight flexible displays in the case of Sony, new catalysts for Johnson-Matthey's hydrogen-powered fuel-cells and developing new battery materials and electrodes to improve performance of Toyota's electric cars.[1]

EPSRC's HPC investments have contributed to the development of CASTEP in a number of ways. Through its investments in HPC infrastructure as well as support services such as the eCSE service, the EPSRC has provided an ecosystem that allowed CASTEP to thrive. EPSRC also funds a number of scientific consortia, which in turn have chosen to provide funding for the continued development of CASTEP. In 2018, EPSRC has further funded a five-year research software engineering fellowship, part of which will be used to transform CASTEP's ease of use by non-computational scientists.

Note: (1) Goldbeck and Court (2016), The Economic Impact of Materials Modelling.
Source: Interview with CASTEP Developer; Castep website: http://www.castep.org; Goldbeck and Court (2016). The Economic Impact of Materials Modelling.

4.3 Training and skills

The importance of skills development and doctoral training were consistently emphasised in the stakeholder consultations undertaken for this study. Without people who possess the right skills and knowledge to exploit the HPC hardware, much of the benefits of HPC would not be realised.

[75] Goldbeck and Court (2016). *The Economic Impact of Materials Modelling.*

More widely, skills development and doctoral training are also of crucial importance to UK science as well as the UK economy. Firstly, doctoral training equips the next generation of researchers with key skills for their career in academia, thus helping to ensure that UK science continues to thrive. Secondly, graduates trained in High Performance Computing are sought after by industry and thus bring benefits to UK companies, the UK economy and the UK exchequer.

As such, enabling doctoral training in High Performance Computing is a key benefit. As reported in the HECToR impact assessment[76], at least **130** PhD students were trained on HECToR. For ARCHER, this number is even higher, with over **170** PhD or post-doctoral students trained on ARCHER across the EPSRC scientific consortia in 2016/17 alone.[77]

Tier-2 centres also play a key role in skills development and training. As gaining access to Tier-2 centres is easier than gaining access to the national service, Tier-2 centres allow training for a wider array of students. Data for the number of PhD and postdoc students accessing the Tier-2 centres is not recorded across all centres. However, of those centres that were able to provide data, two centres reported PhD and postdoc access in excess of **200** users. A further two centres reported that the majority of their users are PhD and postdoctoral students. Moreover, the Tier-2 ReICN report[78] indicates that, on average, **404.5** graduate and post doctorate users used each of the regional Tier-2 centres over their lifetime (2012-2015).

The remainder of this section is structured as follows. Section 4.3.1 explores the impact of HPC on skills and career progression; while Section 4.3.2 further explores the destinations of PhD and postdoctoral students trained on HPC and Section 4.3.4 looks at HPC training provision and the benefits of training.

4.3.1 The impact on skills and career progression

To understand the impact of HPC on skills development and career progression, academics responding to the online survey undertaken for this study were asked about their opinions on the impact of HPC on overall skills development, development of specific skills, and career development. Principal investigators who trained PhD or doctoral students in the use of HPC were further asked about their opinion on whether HPC had improved their students' prospects of moving into industry. The results of this consultation are presented in Figure 20 and Figure 21, and detailed below.

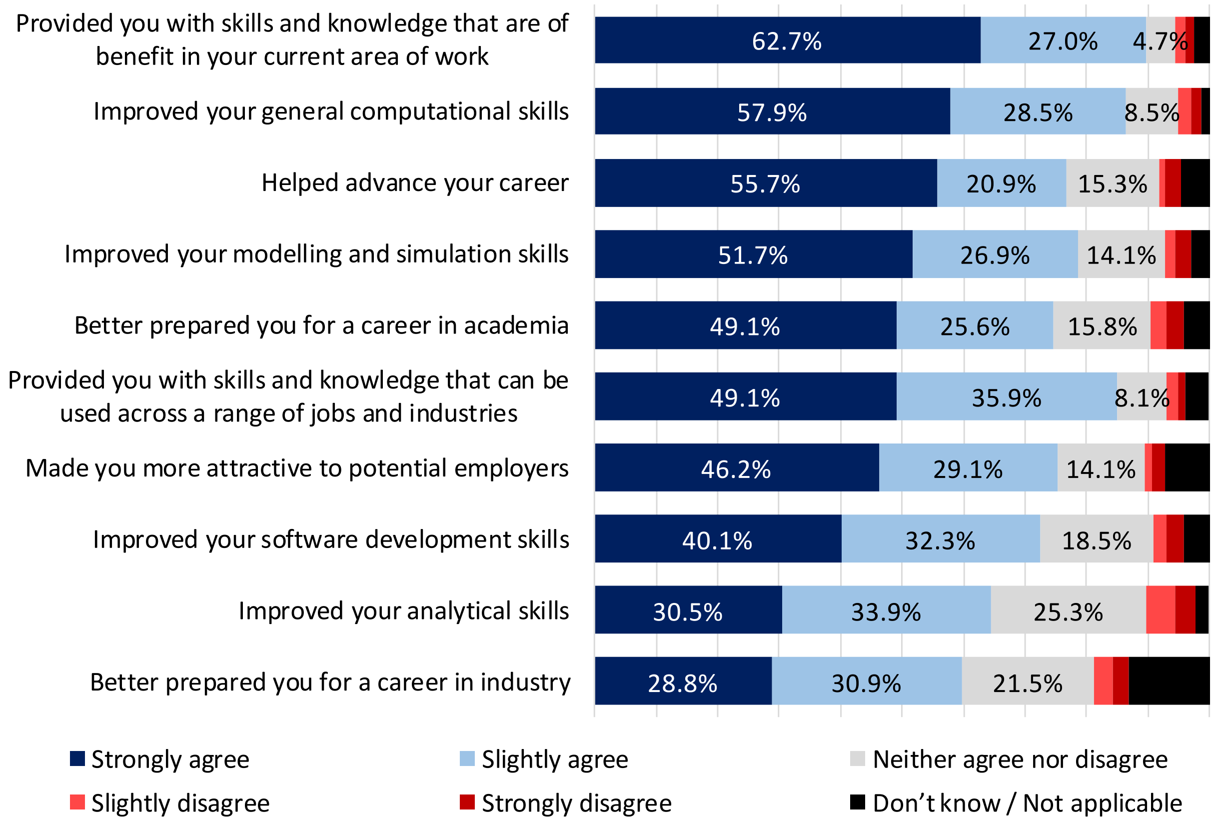
In terms of **overall skills development**: **90%** of respondents agreed, either strongly or slightly, that the use of HPC provided them with skills and knowledge that are of benefit in their **current area of work**; **85%** agreed that it provided them with skills and knowledge that can be used **across a range of jobs and industries**; **75%** agreed that it better prepared them for a **career in academia**; and **60%** agreed that it better prepared them for a **career in industry**.

[76] EPSRC (2014). *The impact of HECToR*

[77] Based on Annual Reports 2016/17 of UKCOMES, UKCTRF, HEC BioSim, UKPP, UKMC; data for UKCP, UKTC and UK AMOR was not available. The actual number of PhD students and postdocs trained on ARCHER is therefore likely even higher.

[78] The Importance of Regional e-Infrastructure Within the National Landscape: A submission compiled by the Regional e-Infrastructure Centres Network

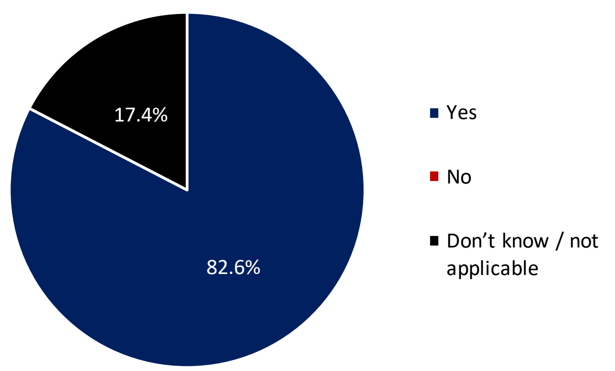
Figure 20 'In terms of skills and career progression, to what extent do you agree or disagree that your use of HPC …?'



Note: Based between 232 and 235 responses. ***Source: London Economics survey of users of HPC capabilities***

In terms of **career development**, **75%** agreed, either strongly or slightly, that the use of HPC has made them **more attractive to potential employers**, while **77%** agreed that it helped them **advanced their career**.

Figure 21 'In your opinion has training in HPC improved your PHD students' / postdocs' prospects of moving into industry?'



Note: Based on 69 responses. Source: London Economics survey of users of HPC capabilities

The use of HPC also helps with **specific skills development** with respondents agreeing that the use of HPC improved their **general computational** (**86%**); their **modelling and simulation** (**79%**); their **software development** (**72%**) and their **analytical** (**64%**) skills.

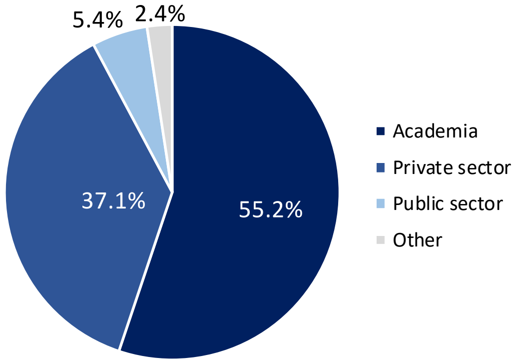
In terms of the benefit of doctoral training, **83%** of principal investigators responding to the survey **believed that HPC training improved their PhD / postdocs students' prospects of moving into industry**.

4.3.2 Destination of PhD and postdoc students trained on HPC

To understand the destinations of doctoral and post-doctoral students trained in HPC, principal investigators consulted via the online survey were asked how many doctoral students they had trained in HPC over the last ten years and what proportion of these stayed in academia, moved to the private or public sectors or to other destinations.

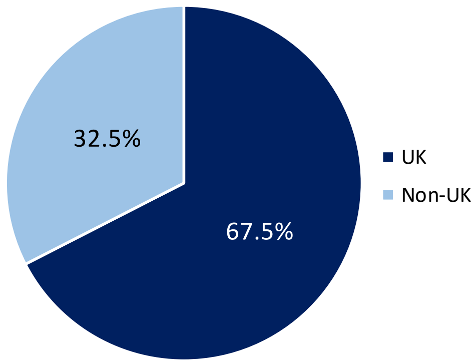
This consultation indicates that more than half (**55.2%**) of PhD and postdoctoral students trained on HPCs stayed in **academia**; over one-third (**37.1%**) moved to the **private sector**; and **5.4%** moved to the **public sector**. A small proportion (**2.4%**) of students moved to **other destinations**. (Figure 22)

Figure 22 Sectoral destination of PhD and postdoc students trained on HPCs



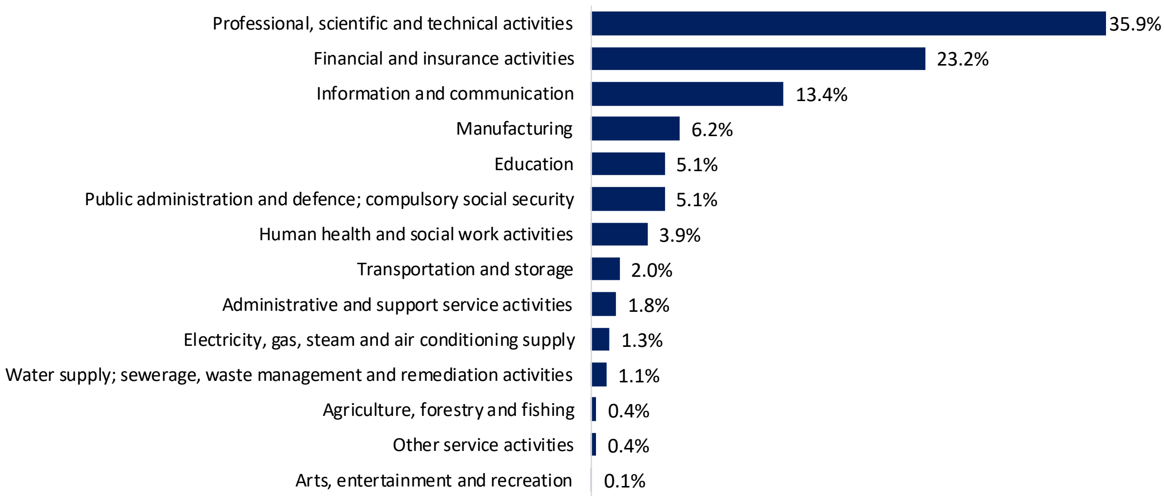
Note: Based on 65 responses from principal investigators. Due to rounding approximations the sum of components may not equal the total. Source: London Economics survey of users of HPC capabilities

Figure 23 Geographic destination of PhD and postdoc students trained on HPCs



Note: Based on 64 responses from **principal investigators**.
Source: London Economics survey of users of HPC capabilities

Figure 24 Detailed sectoral destination of PhD and postdoc students as a proportion of those moving to the private sector



Note: Based on 35 responses from principal investigators. Due to rounding approximations the sum of components may not equal the total. Source: London Economics survey of users of HPC capabilities

The top destinations of those PhD and postdoctoral students that moved into the private sector were **Professional, scientific and technical activities** (**35.9%**); **Financial and insurance activities** (**23.2%**); **and Information and communication** (**13.4%**). (Figure 24)

In terms of geographic destination, the consultation indicates that just over two-thirds (**67.5%**) of PhD and postdoctoral students stayed in the **UK**, while slightly fewer than one-third (**32.5%**) moved **abroad**. (Figure 23)

4.3.3 Valuing the economic contribution of doctoral training in HPC

Although there are many non-economic benefits associated with higher education, Atkinson's[79] report to the Office for National Statistics asserted that the economic value of education and training is essentially the **value placed on that qualification as determined by the labour market**. As such, to place a value on the **doctoral and postdoctoral HPC training** supported by HPC, the labour market benefits associated with enhanced qualification attainment and skills acquisition – to **both the individual and the public purse** - are considered.

To measure the economic benefits to doctoral and postdoctoral HPC training, the labour market value associated with HPC qualifications is assessed relative to the labour market value associated with a comparable group of workers without HPC training (the counterfactual group). The labour market value associated with HPC qualifications is estimated based on data on graduate destinations and earnings, collected via a survey of principal investigators who trained doctoral and postdoctoral students in HPC. This is compared to earnings data, obtained from the 2017 Labour Force Survey, for recent graduates (i.e. workers aged between 25 and 34 with any postgraduate degree).

Table 7 shows the results of this comparison for graduates entering technical fields where their HPC qualifications can be of direct benefit (such as professional, scientific and technical activities, financial services, information and communication, etc.).[80]

Table 7 HPC earnings premium

Benefit	Annual	Present value
Average salary of PhD or postdoctoral graduates skilled in HPC entering technical fields	£65,625	-
Average salary of recent graduates (aged 24-35) with any postgraduate qualification	£42,100	-
Gross HPC earnings premium	**£23,525**	**£507,460**
Additional income tax collected	£7,830	£168,900
Additional National Insurance contributions collected (employee contributions)	£1,260	£27,200
Additional VAT collected	£1,360	£29,410
Net HPC earnings premium	**£13,070**	**£281,950**
Additional National Insurance contributions collected (employer contributions)	£3,250	£70,030
Additional tax receipts and NI contributions accruing to Exchequer	**£13,700**	**£295,535**

Note: Numbers rounded to the nearest 5. Due to rounding numbers may not add up to the total. No data was available on the difference in earnings growth between graduates with and without HPC skills. Therefore, the earnings premium is assumed constant over the working life. Average age of competition assumed to be 27. Average age of retirement assumed to be 65. HM Treasury Green Book discount rate of 3.5% used to discount future earnings. ***Source: London Economics analysis based on survey of HPC users and Labour Force Survey 2017***

[79] Atkinson, B. (2005). *Atkinson Review: Final Report. Measurement of Government output and productivity for national accounts*. Palgrave Macmillan. Basingstoke, England. ISBN 9781403996466.

[80] The earnings premium of graduates entering non-technical fields such those who enter other fields such as education, admin and human health activities was found to be small or non-existent. As such graduates entering these fields are excluded from the analysis.

The benefits accruing to the Exchequer from the provision of higher education are derived from the increased taxation receipts from the increased earnings associated with more highly skilled and productive employees. Based on the analysis of the earnings premium associated with HPC training, combined with administrative information on the relevant taxation rates and bands (from HM Revenue and Customs), the **present value** of additional income tax, National Insurance and VAT associated with doctoral and postdoctoral HPC training is estimated.

Combining data on HPC earnings premium, the additional tax receipts and NI contributions collected, with data on the number of PhD and postdoctoral students trained on HPC, and their graduate destinations allows to calculate the total present value of additional benefits accruing to the Exchequer as well as the total present value accruing to the graduates themselves.

As shown at the beginning of this section, at least 170 PhD and postdocs were trained on ARCHER (≥ 170), while a further ≥ 130 were trained on HECToR. The number of PhD and postdoc students accessing the Tier-2 centres is not recorded across all centres. Furthermore, the estimates presented at the beginning of this section suggest that between 200 and 404.5 PhDs and postdocs, on average, accessed each of the Tier-2 centres. Adjusting for the proportion of users accessing multiple facilities (68.9%, Figure 9) suggests that a total of between 787 and 1,477 PhD and postdoctoral students could have been trained across the HPC centres.

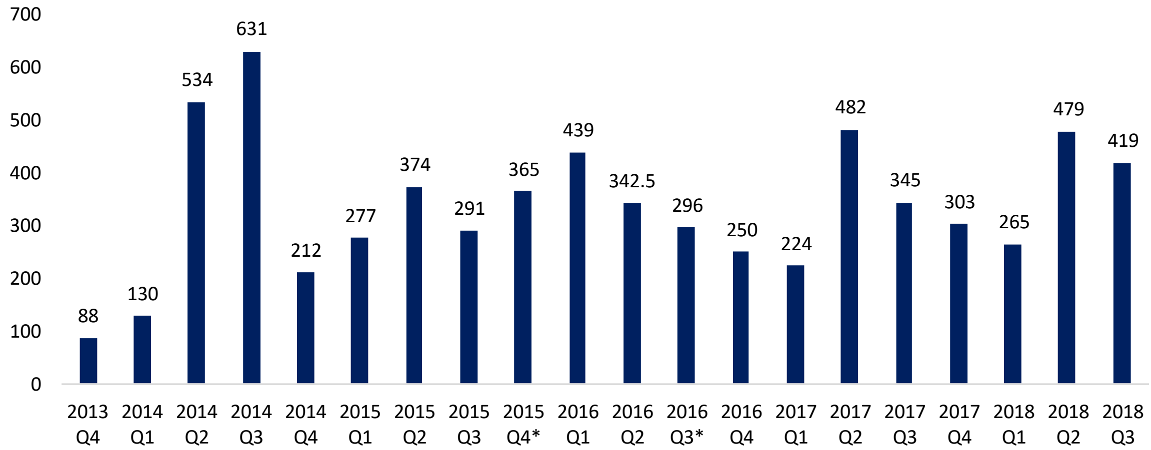
This indicates that **the total present value of additional benefits accruing to the Exchequer is between approximately £50.7 million and £96.3 million,** while **the total present value of benefits accruing to the graduates themselves is between approximately £48.3 million and £91.8 million.**

4.3.4 HPC training provision

Another key activity of HPC centres is the provision of specialised training. The focus of training is mostly on facilitating knowledge of advanced computational techniques and programming languages which can be applied on different HPC systems.

Figure 25 shows the quarterly distribution of ARCHER trainees-days for the period between Q4 2013 and Q3 2018. Each year an average of **57 face-to-face training** in **25 different locations** across the UK have been accommodated by the ARCHER team. Moreover, approximately **6 training** annually were provided through an online platform.

Figure 25 Quarterly number of ARCHER trainees–days in years from Q4 2013 to Q3 2018



Note: Online training are excluded from the graph as the numbers of online participants are unknown. (*) Calculated as average of preceding and subsequent quarter due to missing data. ***Source: ARCHER CSE Service Quarterly Reports Q1 2014 – Q3 2018***

At the Tier-2 level, the number of offered training varies across different facilities. For example:

- ARCHIE-WeSt held 90 free training sessions over its lifetime with an average number of 7 attendees.
- Cirrus organised 10 training with an average duration of 2 days and 22 attendees.
- CSD3 holds introductory course on the Linux command line and usage of HPC system for internal Cambridge users each term. More advanced courses, which are open to all CSD3 users (i.e. Cambridge internal, Tier-2 as well as DiRAC) are also delivered internally as well as in partnership with other UK HPC organisations.
- MMM Hub organised 3 big workshops attended by 30 researchers on average.
- HPC Midlands Plus coordinated 20 trainings within the first year since its launch. The average number of training participants was 30 people.

Benefits of training

Training provided by the ARCHER team and Tier-2 centres is provided free of charge to academics. Assuming a nominal cost of charged trainings for academics of around £250 per day[81], indicates that the value of these trainings to academics is in the region of **£1.7 million** for ARCHER, and an additional **£0.5 million** for trainings provided by the Tier-2 centres.

It should be noted that these valuations only capture the direct benefits of provision of training. To understand the wider benefits of training to the HPC user community, stakeholders were asked about the benefits they had received as a result of the training undertaken. Three main benefit areas emerged from these consultations; these are detailed below:

- **Efficient use of HPC systems:** Unsurprisingly, a large number of stakeholders who attended HPC training emphasised the contribution that the training made to using HPC systems more efficiently.
- **High demand for people with HPC skills:** Training are designed to equip users with, or improve upon users', advanced computational skills. These skills are in high demand both in academia and industry, as is evidenced by the high demand for people with advanced computational skills from industry as well as the high employability rates of students with computational skills.
- **Contribution to software base:** HPC training indirectly contribute to the advancement of HPC software, by equipping HPC users with the necessary skills.

"It has greatly improved my computational, programming, and modelling knowledge allowing me to [be] more efficient and productive." – survey respondent

"HPC skills are very transferrable. Numerical and computational skills make it much easier to get different types of jobs, not just post-docs in that area." - workshop participant

"Without the HPC training people would lose the ability to develop new software." - workshop participant

[81] This follows the recent Technopolis (2018). *Hartree Centre Phase 1&2 Baseline Evaluation Final Report.* Our own comparison of commercial software engineering training courses suggests that this is a reasonable value; see Annex A3.2.

Box 10 Making Musical Moods Metadata

The BBC broadcasts over 200,000 different musical tracks every week both on TV and on radio. With an ever growing musical library containing items to choose from, the BBC needed a way to easily navigate its collection of over a million songs, depending on the mood and emotion that needs to be conveyed.

To label all of the tracks in this library by hand would have taken many years and would have been subjective, depending on the interpretation of the employee labelling the music, so the BBC turned to software which could automatically label a track based on tempo, key, emotion and mood. The project involved three stages: (1) how to define the mood of a track, (2) converting the raw digital music into a format understandable by a computer, and (3) discover which properties of music produce different emotions. The first task of the project was conducted using a combination of mood category models whilst the second was completed using a set of algorithms to classify a track into musical properties such as tempo, key, loudness, rhythm and frequencies. Perhaps the most difficult challenge was the third stage. To overcome this, machine learning techniques were employed, whereby a computer learns how things are related through analysing lots of real world examples.

The N8 High Performance Computing Cluster was used to perform this machine learning analysis on 128,000 tracks – which had already been hand-labelled to describe genre, mood and instrumentation - to discover which musical features are most critical in determining the mood of music. The power of HPC reduced the analysis time from 1.5 years to only six hours, resulting in the cost saving of the labour time which would otherwise have been spent categorising the mood of music.

***Source:** BBC (2012, updated 2015). Pickin' up good vibrations. Available at: https://www.bbc.co.uk/rd/blog/2012-12-pickin-up-good-vibrations [accessed 07/06/2019]; BBC (n.d.). Making Musical Mood Metadata. Available at: https://www.bbc.co.uk/rd/projects/making-musical-mood-metadata [accessed 07/06/2019]*

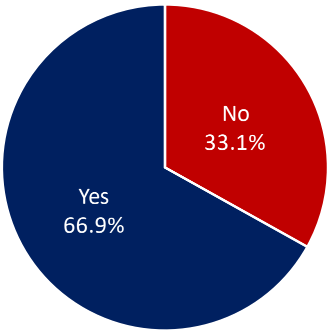
4.4 Collaborations

HPC also supports UK researchers in collaborating with other academics around the world as well as with industry, charities, the public sector, and other organisations. Among principal investigators who responded to the online survey of users undertaken for this study, **two-thirds were involved in collaborations over the last ten years.** (Figure 26)

Analysis based on EPSRC's research outcomes system, indicates that, between 2008 and 2018, **users of HPC were involved in at least 367 collaborations with partners from the UK and at least 29 other countries**.

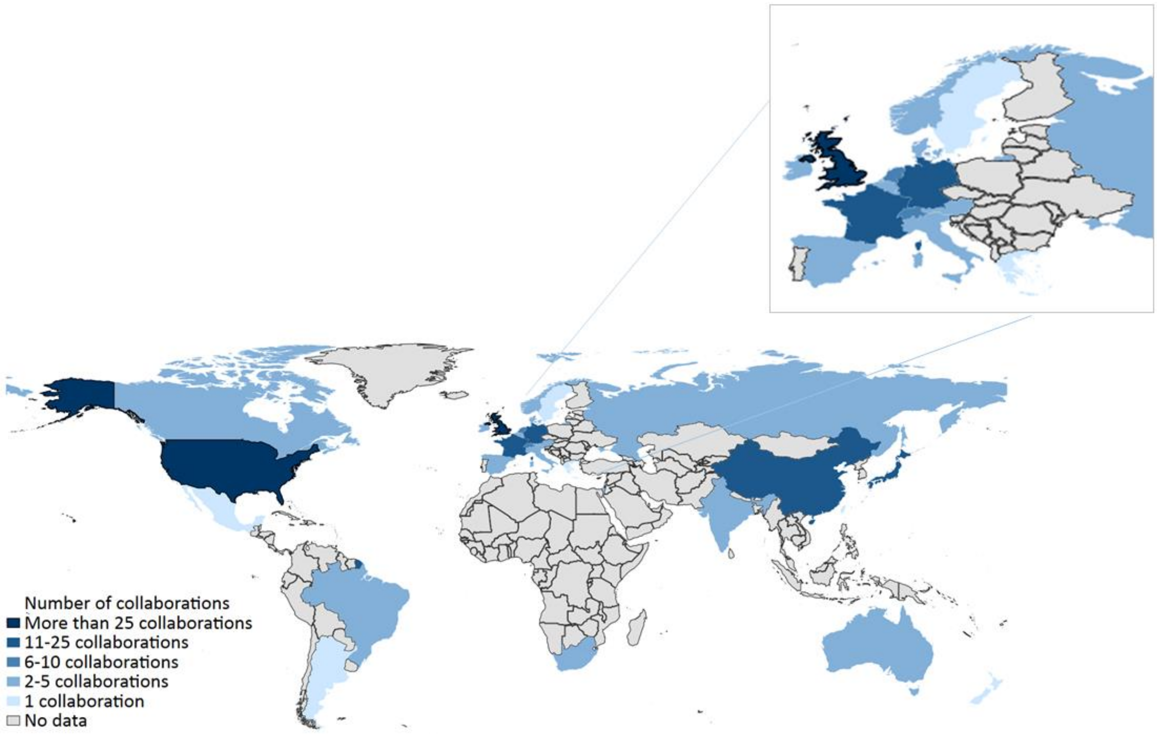
Of these, **192** were collaborations that included a **UK** based partner; **172** collaborations included a **non-UK** based partner. Among international collaborations, collaborations with partners from the **United States** were the most prominent, with **42** collaborations including partners from the USA. Other countries which featured in a significant number of collaborations include **Germany** (**19** collaborations), **France** (**14**), **China** (**13**), **Japan** (**11**) and the **Netherlands** (**9**). (Figure 27)

Figure 26 'As part of your HPC work, over the last ten years, were you part of any collaborations with industry, the public sector or other academics (in the UK or abroad)?'



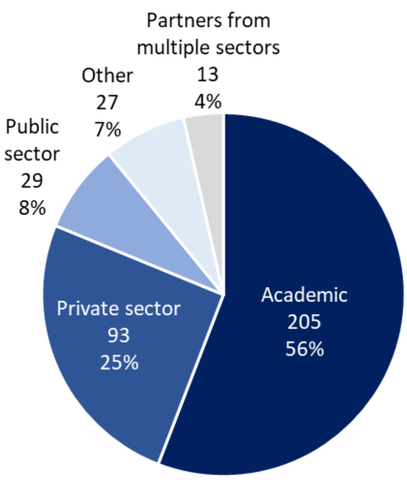
Note: Based on 139 responses from **principal investigators**. 7 principal investigators responded 'Not applicable', these were excluded from the graph.
Source: London Economics survey of users of HPC capabilities

Figure 27 Collaborations by country, 2008-2018



Note: The graph shows the number of collaborations that included a partner from a specific country. Collaborations including partners from multiple countries are included in the numbers for each country featuring in the collaboration. Analysis based on EPSRC's research outcomes system ***and does not include NERC. Only collaborations which were reported to EPSRC are included, and as such the true figure may be much greater.*** Only partial data for 2018 was available at the time of reporting. ***Source: London Economics based on Researchfish™ data provided by EPSRC. Made with Natural Earth.***

Figure 28 No. of collaborations by sector, 2008-2018



Note: Analysis based on EPSRC's research outcomes system and does not include NERC. Only collaborations which were reported to EPSRC are included, and as such the true figure may be much greater. Only partial data for 2018 was available at the time of reporting. ***Source: London Economics based on Researchfish™ data provided by EPSRC***

Unsurprisingly, the majority (**205** of **367**) of collaborations were academic collaborations. Collaborations with industry also accounted for a large share of collaborations (**93** of **367**).

In addition, EPSRC's research outcomes system indicates that users of HPC were involved in at least **29** collaborations with the public sector; **27** collaborations with other organisations such as charities and non-profits, hospital, or learned societies; and **13** collaboration with organisations from multiple sectors. (Figure 28)

As discussed in Section 4.1, the benefits of collaboration to the wider UK economy include the transfer of knowledge and skills from the public to private sector, along with the commercialisation of new products and services.

Box 11 Industrial Collaborations using Cirrus HPC

Collaborations with industry form an important part of the UK HPC community, bringing benefits to industrial partners, and thus the UK economy, that would otherwise not be possible or require large investments in HPC infrastructure by the company. Indeed, some of the Tier-2 centres actively collaborate with industry. For example, in any given year, the EPCC works with around 80 paying users who benefit from their Cirrus HPC. Three of these collaborations are highlighted here.

On demand CFD simulations

Helyx is a general purpose CFD software solution that can be used for engineering analysis and design optimisation. However, for users to fully engage with Helyx would require expensive investment in HPC hardware, which is out of the price range of many ENGYS customers. The EPCC worked with ENGYS to install Helyx on the Cirrus HPC cluster and built a client access platform to provide users with easy access to the CFD simulations. This allows users to access the Helyx software – combined with HPC operating power – from their local workstations, on demand. This allows cost effective access to Helyx on HPC cloud for quicker engineering analysis. It is estimated that this solution is 5 times cheaper than building an in-house computer cluster for those firms which have the capabilities to do so. A further upshot of this is an increased time to market for products and improved design capabilities. With simulations 10 times faster on HPC than standard systems there are potential savings of up to €2,000 per simulation per day. ENGYS itself saw sales rise by 20% as a result of the introduction of on-demand usage for HELYX software in the HPC cloud.

Modelling explosion, flammable and toxic releases

Another partner, Gexcon, uses Cirrus HPC capabilities more directly to run their own simulations (called FLACS) for modelling explosion, flammable and toxic releases in a technical safety context. The Cirrus cluster is ideally suited to run these multiple short-lasting simulations simultaneously, processing several hundred simulations in a matter of days. This allows Gexcon to deliver results to their customers much faster without the need to themselves invest in large-scale HPC infrastructure. Following on from this successful collaboration, EPCC and Gexcon developed a cloud service version of FLACS which is applicable across the petrochemical, nuclear and general process manufacturing industries. Such users will benefit from easy, cost-effective access to this software, reducing simulation times from weeks to a few hours.

Irish Marine Institute

Another example is the Irish Marine Institute, who are responsible for promoting Ireland's marine resources and run the Irish National Weather Buoy Network and the Tide Gauge Network alongside modelling applications which provide simulations of the ocean for the coast guard, government bodies and the fishing industries. Such modelling involves the forecasting of sea state and modelling oil-spills, harmful algal bloom predictions, marine habitat classifications and water quality modelling, to name just a few. Modelling these phenomena requires HPC capabilities, which are provided by the EPCC. This removes the need for expensive investments in their own HPC capabilities for the Irish Marine Institute.

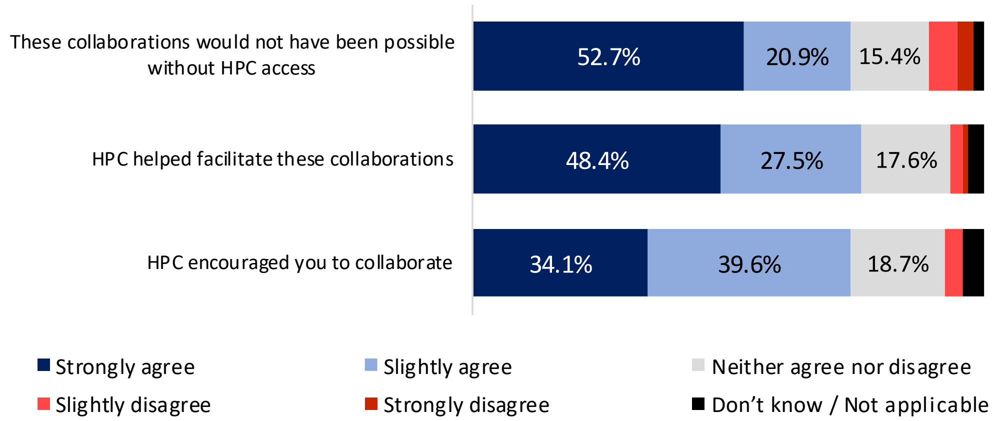
Source: Interview with EPCC and Case Studies provided by the EPCC; Cloud-based simulation of pipeline components for the oil & gas industry. Available at: https://www.fortissimo-project.eu/experiments/512 [accessed 21/05/2019]
Picture credit: Konstantin Chagin / Shutterstock

4.4.1 The role of HPC in supporting collaborations

To understand the contribution of HPC to these collaborations, users were further asked about the importance of HPC to their collaborations. Specifically, of principal investigators who said they had collaborated over the last ten years, approximately (Figure 29):

- **74%** agreed, either slightly or strongly, that HPC **encouraged them to collaborate**;
- **76%** agreed that HPC **helped in facilitating these collaborations**; and,
- **74%** agreed that these **collaborations would not have been possible without HPC**.

Figure 29 'To what extent do you agree or disagree that …?'



Note: Based on 91 responses from principal investigators. Source: London Economics survey of users of HPC capabilities

HPC users consulted for this study also highlighted the importance of HPC in enabling them to gain access to other HPC facilities internationally and to collaborate with other leading international research groups. For example, participation in PRACE has allowed UK researchers to access HPC facilities across Europe.

"I have a permanent post in France, and I was attracted by CSD3 to work with my collaborators in Cambridge. Without CSD3, I probably would not have started these very fruitful scientific collaborations." – survey respondent

Stakeholders also emphasised the role that access to and use of UK HPCs play in UK researchers gaining the necessary skills and HPC experience needed to apply for time on other international HPCs. This was seen as particularly important for young scientists, for whom it would be very difficult to gain access to international HPCs without first being able to gain experience in high performance computing via UK facilities.

"My entire research career is founded on collaborations, many of which are outside the UK. None of these would be possible without HPC facilities from EPSRC." – survey respondent

"[Collaborations have] allowed me to integrate experimental and computational work successfully and provided me with a community in which I can expand my research base and into which I can introduce novel teaching methodologies" – survey respondent

"Collaboration is essential in science. We can provide stronger conclusions when working with other groups. Problems can be approached with more methods from more directions. This provides stronger evidence to back claims." – survey respondent

5 Impact on industry

> **"We use the HPCs to run Computation Fluid Dynamics simulations. Having access to these HPC systems means we can run our simulations in hours rather than days and so increases the amount/scale of simulation we can do for ourselves or on behalf of our customers." – industry survey respondent**

While the main purpose of HPC is to support cutting-edge academic research, HPCs also benefitted UK industry through a number of channels.

As highlighted in Section 2, **over 150 industry users accessed ARCHER directly, with many more having accessed the Tier-2 centres or ARCHER's predecessor HECToR**. Accessing HPC provides industry with computing capabilities far larger than most companies would be able to afford in-house. This allows industrial users to test and scale their code, reduce the time needed to perform calculations compared to in-house capabilities or run large-scale simulations they would not have been able to run in-house, and support future product design, among others.

Key benefits of HPC highlighted by industrial users were, among others:

- The **size of the computing resources** that would otherwise be out of reach of most businesses.
- The **flexibility of the services**, allowing users to supplement their in-house computing resources as needed.
- The **cost of the services**, which was seen as very competitive compared to purely commercial solutions.
- The **professionalism with which the services are operated**, including the expert knowledge of centre staff.

In addition to direct users, UK industry also benefits from collaborations with academics using HPC. As discussed in Section 4.4, EPSRC's research outcomes system indicates that users of **HPC were involved in over 90 collaborations with the private sector**[82].

It should be noted, that EPSRC's research outcome system only captures collaborations with industry reported back to the EPSRC. The actual number of collaborations with industry is likely much larger. Moreover, some of the Tier-2 centres also collaborate directly with industry. For example, as part of the Fortissimo project, led by the EPCC, the EPCC collaborated with over 215 partners from industry.

Fortissimo is an online marketplace offering European businesses cost-effective, on-demand access to HPC facilities for computationally intensive simulations. In essence, Fortissimo acts as a broker: connecting firms with the need for HPC with HPC providers.

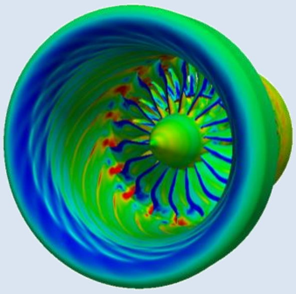
The benefits to industry come from on-demand access to advanced simulation and modelling software, without the need to invest in their own HPC infrastructure which would be infeasible for many SMEs. Access to HPC facilities both permits complex simulations, which may otherwise not be possible, and allows simulations to be run in a matter of hours rather than weeks.

[82] Note that this figure does not include NERC.

Fortissimo is coordinated by the EPCC and, across two projects, involved 215 partners, many of which are European SMEs, working in 92 experiments to demonstrate the value of HPC to the end-user community for the first time[83].

The EPCC and other Tier-2 centres also collaborate directly with industry. JADE provides an opportunity for industry to use its HPC facilities and expertise to develop their Artificial Intelligence, Deep Learning and Machine Learning products. JADE works with industrial customers to identify how HPC can benefit them and provides solutions for these problems. Some examples of these collaborations are shown in Box 11 for the EPCC and Box 4 for JADE alongside Table 8 which explores some success stories and the associated benefit, both in monetary and business terms.

Box 12 Rolls-Royce



Rolls-Royce used access to ARCHER for pre-competitive research to scale its simulation codes well beyond what was possible at the time. This allowed the company to perform software development that would have been impossible without access to the leadership capability of ARCHER. Highlights include scaling their in-house CFD codes to 100 times more cores than possible at the time and demonstrating the Company's first ever 2 billion cell simulation – still the largest that has been achieved. The impact of this work is that the codes developed on ARCHER were able to be deployed immediately when the company upgraded its HPC system. Some cases are running 3-4 times faster as a result of the ARCHER developments and others are expected to show a factor of 10 speed-up when they are deployed. Better HPC utilisation improves engineering productivity – over the lifetime of the HPC, the improvements developed on ARCHER are expected to generate millions of pounds of benefits.

Rolls-Royce's access to ARCHER was also a fundamental contributor to the Company leading the ASiMoV EPSRC Prosperity Partnership – the largest UK funded computational science research project in recent times. The project's ambition, to achieve the world's first high-fidelity simulation of a complete gas-turbine engine during operation, will require trillion cell simulations on millions of computing cores. Rolls-Royce is contributing several £m to UK academics in one of the UK's most aspirational computational and physical science projects that would not have been conceived if it weren't for ARCHER.

The figure shows a very high resolution solution for the noise generated (pressure contours) by the fan at the front of an engine. The ultra-high resolution was only possible by running this case on ARCHER.

Source: Rolls-Royce

[83] EPCC (2019). Fortissimo projects in numbers. Fortissimo 2 Final Review

Table 8 Examples of industry benefits

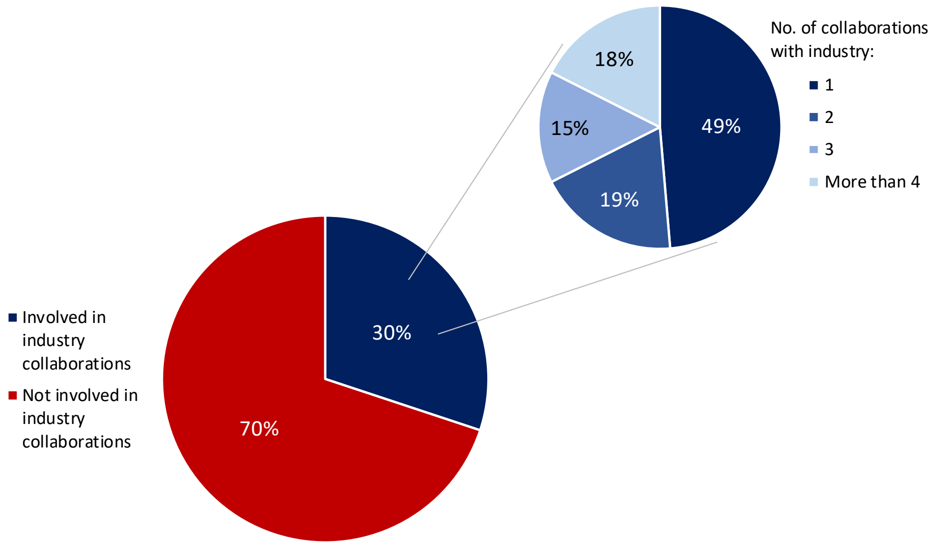
Project name	Problem and solution	Benefits
Optimisation of production in the dairy sector (Fortissimo)	- Dairy machines can break-down between 3 and 5 times a year leading to costs from penalty fines, lost product and clean-up. - HPC allows optimisation software to harness big data and permit efficient maintenance	- Savings for dairy manufacturers in excess of €50 million from avoided downtime as penalty fines are avoided - Increased output - Reduction in waste
Simulation for the oil and gas industry (Fortissimo)	- For many engineering/manufacturing SMEs there is a need for HPC capability to run occasional computational fluid dynamic simulations - Cloud-based HPC is a cost-effective way to run such simulations without the need to invest in own HPC infrastructure	- SMEs which would be unable to invest in their own HPC infrastructure gain access - Costs 5 times lower with a cloud-based HPC solution compared to own systems - Simulations can be up to 10x faster with potential savings of €2,000 per simulation per day
Simulation of the binding capacities of target drug compounds (Fortissimo)	- Repositioning existing drugs has huge potential to combat diseases but requires large computational resources to examine the properties of drugs and which diseases they might tackle - HPC cloud infrastructure can be used in combination with algorithms to evaluate drug compounds	-The firm providing this software foresees a potential increase in profits of around 3% per annum -More broadly this software is able to increase the productivity of the drug industry by increasing the number of uses each drug can tackle
Vehicle engineering simulation (Fortissimo)	- Physical testing of air flows around a vehicle is expensive but increasing airflow makes a vehicle faster and more efficient - Computer simulation allows millions of tests to be run at low-cost, enabling engineers to experiment with different designs - ICON developed browser accessible apps to analyse the necessary systems using HPC cloud systems	- No need for SME to purchase expensive equipment - Significant time saving with no need for physical testing - High end simulations cost €900 compared to annual costs of €300,000-€400,000 for in-house simulation capabilities which can only run around 50 simulations per year
Urban planning (Fortissimo)	- City planning tools (energy efficiency, qualify of living etc) require HPC due to the scale of the simulations being run - The use of cloud-based HPC allows the simulations to be run using HPC but displays data in a familiar workstation	- Simulations which would take days or weeks on a workstation could be run in a few hours using the cloud-based HPC system - An in-house system running continuously would cost £0.10 per core hour of simulation compared to £0.05 per core hour of cloud-based HPC
Earthquake resistant design (Fortissimo)	- The effect of earthquakes, explosions and flooding on civil engineering structures requires complex computational resources to solve	- Reduction in cost for an SME is estimated to be between 40% and 60% - Permits access to SMEs who can't invest in own infrastructure - Engineer productivity is doubled

Application	Challenge	Impact
Optimising gas and flame detector layouts in hazardous manufacturing and production plants (Fortissimo)	- The computing power required for a computational fluid dynamics-based approach to optimising gas detector layouts in manufacturing and production facilities is prohibitive in terms of expense and time - Installing the software on Cirrus allows customers to use HPC cloud resources at a much lower cost than accessing a local cluster	- Potential savings into the millions for a typical installation - Increased revenue for the software firm - Computational speed-up of a factor of 10 - Improved placement of gas detectors leading to improved safety
Detecting faults in aeroplane wings (JADE)	- Using Deep Learning on visual images, the wing manufacturer was able to model the inside of the wing during assembly and detect anomalies in components	- Reduction in QA inspection process of 1-2 days and thus greater manufacture time - Capability to reduce in-service inspection periods by several days, working in a hazardous environment
Identifying ice cream leakage for a global ice cream manufacturer (JADE)	- Ice cream manufacturer had customer complaints from ice cream leakage through the cone. This resulted in high customer return costs - Stemmed from faulty production where wafer cones were not coated correctly	- Deep Learning technique applied in the manufacturing process to quickly classify defects in product allowing proactive manufacturing adjustments before serious issues - Potential savings of 1-5$ of production costs and reduction in customer return cost
Improving welding process for a commercial digger manufacturer (JADE)	- Digger construction has significant amount of manual welding which leads to inconsistent quality and thus unpredictable maintenance causing significant down time	- Deep Learning modelling was used to improve welding quality by analysing images and detecting anomalies (imperfections) which can be addressed in real-time - Leading to reduced build time and lower maintenance costs and downtime
Increasing quality of submarine construction (JADE)	- Submarine has 4km of welding performed by thousands of different welders of various skills and competence - This leads to potential weld imperfections which constitutes a risk at great depths and can lead to huge expense of rework	- Images taken from x-ray, ultrasound, sonar and visual cameras are input into DL model to generate best practice - Ability to fix defects before launch, leading to large savings - Reduced overall build time
Virtual jury (JADE)	- Legal cases can take excessive time to complete which leads to high costs at tax-payer expense	- AI Deep Learning modelling supports the defendant and prosecutor to guide the proceeding to get to the truth quicker and more efficiently using knowledge from past legal cases - Leads to quicker closure of cases, lower court costs and tax-payer expense.

Source: Success Stories of The University of Edinburgh. Available at: https://www.fortissimo-project.eu/partners/university-edinburgh [accessed 21/05/2019]; JADE DLaaS references

Further evidence collected via the online survey of users undertaken for this study indicates that **approximately 30% of HPC users who responded to the survey were involved in collaborations with industry**. Of those, approximately half (**51%**) were involved in multiple collaborations with industry (Figure 30).

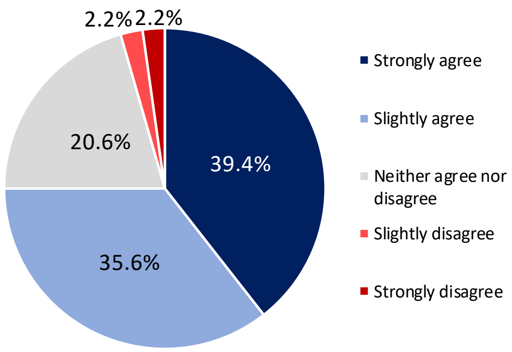
Figure 30 'As part of your HPC work, over the last ten years, how many projects involving collaborators with industry were you part of? How many collaborating partners have you worked with?'



Note: Based on 246 respondents, of which 74 respondents collaborated with industry. ***Source: London Economics survey of industry users of HPC capabilities***

UK industry further benefits from the highly skilled graduates that choose a career in industry following their HPC training, with **more than one-third** of PhD and postdoctoral students surveyed choosing to move to the private sector (see Section 4.3.2).

Figure 31 'To what extent do you agree or disagree that HPC helps attract industry to the UK?'



Note: Based on 180 respondents. Excludes 41 respondents who selected 'Don't know / Not applicable'. ***Source: London Economics survey of users of HPC capabilities***

HPC also contributes to the foundation of spin-outs. Data on spin-outs obtained from EPSRC's research outcomes system and the user survey undertaken for this study suggests that **at least 14 spin-outs benefitted, either directly or indirectly, from HPCs**[84]. Moreover, one user responding to the survey undertaken for this study reported that they were involved in a spin-out which was a direct result of access to the HPCs, while three further users said that access to the HPCs helped a lot in establishing the spin-outs they are involved in.

The contribution that HPC makes to attract industry to the UK - through its contribution to UK science competitiveness, the excellent UK

[84] Note that data from the research outcomes system does not include NERC. Only spin-outs which were reported to EPSRC are included in the research outcomes system, and as such the true figure may be much greater.

software base, and the availability of highly skilled people - was also emphasised in stakeholder consultations undertaken for this study. For example, **75% of respondents to the online survey agreed**, either slightly or strongly, **that HPC helps attract industry to the UK** (Figure 31).

Box 13 Understanding corrosion at the molecular level

The BP-International Centre for Advanced Materials (BP-ICAM) is a partnership between BP and the University of Manchester, the University of Cambridge, Imperial College London and the University of Illinois at Urbana Champaign. The BP-ICAM was set up by BP in Autumn 2012 with a $100m investment bringing together capabilities in a wide array of materials research including structural materials, corrosion, separations, surfaces, deposits, imaging, modelling and self-healing materials.

An understanding of how materials work is vital for the wide operations of BP, from exploring oil and gas, to producing it, transporting it and refining it. Understanding how the materials involved in these operations work at a fundamental level is vital to understand their behaviour in operating environments. A recent project focuses on research into surface degradation and corrosion of steel in demanding environments. The research forms part of a multi-million pound collaborative Prosperity Partnership, funded by the EPSRC, between BP and the BP-ICAM universities including two new partners – University of Edinburgh and the University of Leeds.

It is estimated that the annual global cost of corrosion is around $2.5 trillion a year, £55bn of which occurs in the UK, and leads to a variety of unexpected failures[1, 2]. The current method to prevent corrosion is through a barrier layer such as dynamic painting but how corrosion works on a molecular level is still largely unknown. As part of the partnership, researchers are trying to study how corrosion happens, why it happens in some places and not others, and predicting the longevity of materials. Researchers will also explore new techniques to combat corrosion such as new coatings and alloys which can extend the working life of materials. Over a billion tonnes of steel are produced every year, so any improvements in the lifetime of steel products would have a significant monetary benefit.

At Imperial College BP-ICAM researchers have developed a new model of corrosion that links molecular level mechanisms to observed growth and degradation of the corrosion film. This requires exceptional computational resources that require HPC for both CPU and memory. Whilst BP has access to its own HPC facility in Houston, access to ARCHER is essential due to wider accessibility, flexible interfaces and in some case the size of the calculations.

The gains from this avenue of research are potentially very large. Any new insights into the behaviour of steel and corrosion could save significant costs and may result in a switch towards a self-protecting system which acts more like stainless steel. One possible successful outcome would be to treat pipes with one chemical wash to significantly increase corrosion resistance. This could save BP around $120 million *per annum* through better control of corrosion.

***Source:** Interview with Nicholas Harrison (Imperial College London) and Sheetal Handa (BP); EPSRC (nd). Transcript for Prosperity Partnerships - The University of Manchester and BP. Available at: https://epsrc.ukri.org/newsevents/multimedia/ppmanchesterbp/transcript-for-prosperity-partnerships-the-university-of-manchester-and-bp/ [accessed 15/05/2019]; (1) Nace International (2016). International Measures of Prevention, Application, and Economics of Corrosion Technologies Study; (2) ICAM (nd). Lecture Series: Corrosion Studies. An overview from Dr. Brian Conolly. Available at: http://www.icam-online.org/research/lectureseries/corrosionstudies/ [accessed 15/05/2019]*

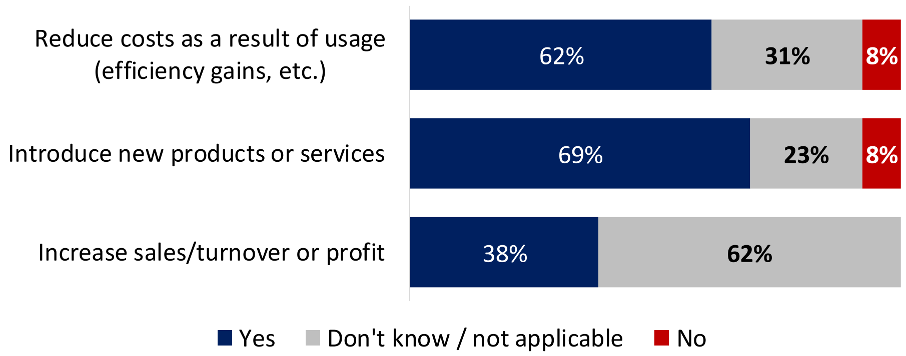
5.1 Valuing benefits to industry

Benefits to industry users accessing HPC accrue through a number of channels. First, in the short term, being able to access HPC brings a direct reduction in costs compared to accessing commercial HPC capabilities and avoids significant investment costs in companies own clusters. Second, work undertaken on HPC contributes to increased business profitability in the medium to long run, for example, through cost reductions as a result of HPC usage, efficiency gains, improvements to existing products or services or by contributing to the introduction of new products or services, or enhanced business growth.

Evidence collected from industry users via stakeholder consultations, the online survey and by comparing access rates of HPC capabilities to those of commercial computing providers suggests that cost-savings to industry users of accessing Tier-1 / Tier-2 HPC capabilities compared to using commercial providers or investing in their own cluster are significant.

Moreover, **62%** of industry respondents said that they had reduced costs as a result of usage (e.g. via improvements in efficiency, etc.), while **69%** said HPC helped them introduce new products or services. While direct cost savings have an immediate result on business profitability, other benefits such as the introduction of new products or services can have a significant time lag before impacting business profitability. Among industry survey respondents, **38%** said that HPC had already resulted in increased sales/turnover or profit. (Figure 32)

Figure 32 Figure 33 'Has access to / usage of EPSRC's HPCs helped your organisation to …?'



Note: Based on 13 responses from <u>industry</u>. ***Source: London Economics survey of users of HPC capabilities***

5.1.1 Valuing benefits to industry

To monetise benefits to business profitability among industrial users, and thus to the UK economy, data on the demographics of businesses, by size (SME, mid-cap, large), accessing ARCHER is combined with data on average profits of firms (within each size class), the proportion of firms seeing an impact on profit, and the average increase in profits:

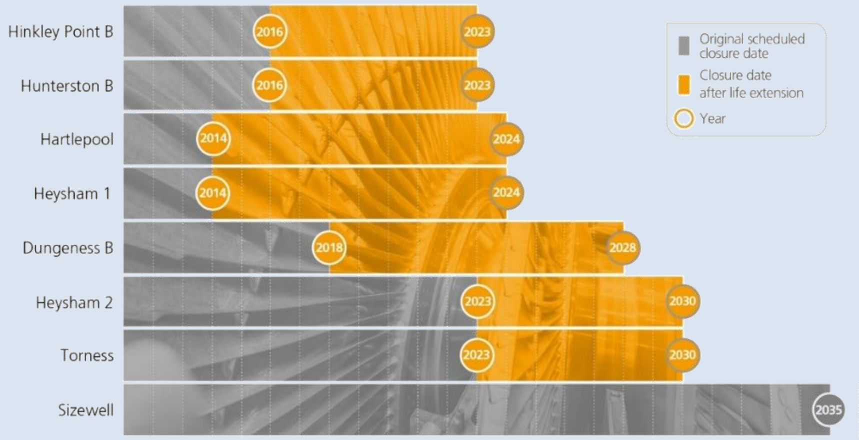
$$\textit{Benefits to UK industry} = \textit{Average profit of firms in size group} \times \textit{Proportion of firms seeing an increase in profit} \times \textit{Average proportional increase in profit per firm} \times \textit{No. of firms accessing HPC in size group}$$

Box 14 Safeguarding the UK's energy supply: The role of HPC in extending the lifetime of EDF Energy's nuclear power plants

The ever-increasing energy demand brings challenges for the efficient production of safe and clean energy. Nuclear energy plays an important role in meeting this demand, with nuclear energy currently accounting for approximately 20% of UK electricity generation. Nuclear energy, in the UK, is supplied by eight nuclear power plants, operated by EDF Energy.

However, many of the UK's nuclear plants are nearing the end of their lifetime, with the next generation of new nuclear plants not expected to arrive until at least the mid-2020s. In order to ensure the UK's energy supply, EDF Energy has launched a lifetime strategy seeking life extensions for all its nuclear stations, where it is safe and commercially viable to do so. Since 2008 EDF Energy has added an average of 8 years to the expected life of their nuclear plants. Four stations (eight reactors) would already be closed without these extensions.

Figure 33 Nuclear lifetime extensions



Source: EDF Energy: Nuclear lifetime management.

In order to ensure that strict operations safety regulations are adhered to, researchers are using ARCHER and other UK HPC systems in conjunction with computational fluid dynamics (CFD) to model the problems involved in nuclear reactor fluid flows as well as the possibility of fracture in the graphite core of nuclear reactors.

Performing these checks in other ways would be very costly and difficult to do. Without access to UK HPC systems, EDF would be unable to fund this critical research and may thus be unable to keep their plants opened for as long as they are opened now. Given the high proportion of energy generated by these plants, and a lack of immediate replacements, this would inevitably result in a significant shortfall in UK energy supplies. To make up the shortfall, the UK would either need to import energy or invest in additional capacity. Both options are expensive, and the latter is associated with long-lead times to plan and build new energy infrastructure.

In addition to extending the lifetime of plants, simulations are also used to reduce plant downtime which may otherwise occur due to safety concerns or maintenance issues. This brings considerable benefits, with downtime of a plant resulting in costs of approximately £1m a day for EDF energy.

Source: Interview with EDF Energy; EDF Energy: Nuclear lifetime management Available at: https://www.edfenergy.com/energy/nuclear-lifetime-management [accessed 09/05/2019]; Cray (2014) ARCHER Flexes Its Muscles on Code_Saturne. Available at: http://www.archer.ac.uk/casestudies/AB-Code_Saturne-201403.pdf [accessed 09/05/2019]; Department for Business, Energy & Industrial Strategy (2017). Digest of UK Energy Statistics 2018 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/729451/DUKES_PN.pdf [accessed 09/05/2019]

Industry stakeholders consulted for this study, emphasised the positive impact of HPC on their business. However, putting a value on the contribution that HPC makes to business profitability was felt to be difficult as calculations utilising HPC may not directly feed into new products or service. Rather, HPC is often used as a test bed to improve or scale up existing models or software or to develop new models that may ultimately feed into new products. For example, among respondents to the survey, only one company was able to provide an indicative value for the increase in profit linked to their work on HPC.

In the absence of this data, the analysis relies on the following observations:

- The only respondent providing a range for benefits in the online survey suggested an increase in profit of around 25% over the last five years (or a CAGR of around 4.6% per year). In addition, one SME working with the EPCC as part of the Fortissimo project reported an annual increase in profits of around 3% (Table 8). However, as mentioned previously, not all firms accessing HPCs see an increase in their sales. Combining these estimates with the 38% of respondents who said they had received increases in sales / turnover or profit, suggests a total effective increase in profits of between **1.1%** to **1.7%** per annum[85].
- In the absence of ARCHER / Tier-2 HPC, large companies may invest in their own HPC machines. While the price of HPC can vary significantly depending on the configuration, a cheap supercomputer can be acquired for between £0.3 million to £2.1 million[86]. However, even the more expensive machine is significantly less powerful than ARCHER / Tier-2. Assuming a five-year replacement cycle and allowing for operational costs of up to 20% of the purchase value[87], would mean total annual costs in the region of £0.8 million (based on £2.1 million machine), or approximately **0.4%** of average profit of FTSE100 enterprises[88]. In contrast, accessing publicly funded HPCs can be done at negligible costs[89].
- For Midcap firms it is unlikely that they would invest in the more expensive machine, given that this would mean costs of approximately 4% of profit. Assuming mid-sized firms instead opt for the lower cost machine (£0.3 million), would imply costs of approximately **0.7%** of profit per year.[90]

Given these observations, the analysis used an effective increase in profits rate of **0.6%** (average of cost reduction of mid-sized and large firms, assuming no increase in profit beyond cost savings) for the low estimate, and of **1.7%** (based on 4.6% profit increase per year accruing to 38% of companies) for the high estimate.

Note that small firms do not usually have the means or expertise to acquire and efficiently utilise their own supercomputers. Therefore, the above lower bound estimate does not directly apply to them. However, in the absence of publicly funded HPC, it is likely that a large proportion of HPC research at smaller firms would not have taken place. Moreover, diminishing returns to R&D[91]

[85] 3% X 38% = 1.14%; 4.6% X 38% = 1.7%

[86] In 2018 prices. Original prices $0.5 million to $3 million (in 2013 prices), based on: https://www.eweek.com/servers/cray-offers-low-cost-xc30-ac-supercomputer-for-500-000. Converted into £ using the 2013 average Bank of England annual spot exchange rate. Deflated using ONS Gross Fixed Capital Formation deflators.

[87] Based on average replacement time of EPSRC and comparators, and operational cost allowance of NSF.

[88] Based on average proportion in profit of FTSE100 companies, obtained from Top Track 100, between 2008 and 2018.

[89] E.g. Cirrus charges £0.037 per core hour for industry access; see https://www.epcc.ed.ac.uk/facilities/demand-computing/cirrus [accessed 26/07/2019]

[90] Again based on annual costs assuming a five year replacement cycle and operational costs of 20% of purchase value, as a proportion of average profit of FTSE250 companies, obtained from Top Track 250, between 2008 and 2018.

[91] See e.g. Kim, Konshik. (2018*). Diminishing returns to R&D investment on innovation in manufacturing SMEs: do the technological intensity of industry matter?* International Journal of Innovation Management. 1850056. 10.1142/S1363919618500561.

suggest that SMEs that do invest in R&D may see higher marginal returns than larger counterparts already investing large amounts into R&D. Therefore, the same lower bound was used for small firms.

Table 9 Estimated benefits of ARCHER to UK industry

Size of firm	Firms accessing ARCHER		Total effective annual increase in profits		Total effective increase in profits over ARCHER's operational time	
	%	No.	Low estimate (0.6%)	High estimate (1.7%)	Low estimate (0.6%)	High estimate (1.7%)
SME	55%	83	£ 0.1 m	£ 0.2 m	£ 0.3 m	£ 0.9 m
Mid-cap	20%	30	£ 3.4 m	£ 10.8 m	£ 17.7 m	£ 55.9 m
Large	25%	38	£ 42.9 m	£ 135.1 m	£ 221.3 m	£ 697.6 m
Total	**100%**	**150**	**£ 46.3 m**	**£ 146.1 m**	**£ 239.3 m**	**£ 754.4 m**

Note: For the purpose of this study, SMEs are defined as companies with up to 250 employees or annual turnover up € 50 million. Mid-cap are defined as companies with a company with a market capitalisation of between €2 billion and €10 billion. Large companies are defined as those with a market capitalisation greater than €10 billion. Proportion of ARCHER industry users by size provided by EPCC, average profit for SMEs, by year, calculated based on turnover of SMEs, and ratio of turnover and profit in the total business economy obtained from the EC Structural Business Statistics database. Where values were missing linear extrapolation was used to estimate the average profit. Average profit of mid-cap enterprises calculated based on profit of FTSE250 companies obtained from Top Track 100; average profit of large enterprises calculated based on profit of FTSE100 companies obtained from Top Track 100.

Source: London Economics

Combining this total effective annual increase in profits with data on the average ratio of profit to GVA in the UK business economy[92], suggests that **industry use of ARCHER could contribute between £95.5 million and £301.0 million to the UK economy each year, or between £0.5 billion and £1.6 billion over the operational time of ARCHER.** (Table 10)

Assuming that the distribution of firms accessing HECToR and the Tier-2 centres is similar to that of ARCHER indicates that **industry use of HPCs could have contributed between £0.9 billion and £2.9 billion to the UK economy so far**[93].

It should be noted that some industry users may be accessing multiple facilities. This means that the overall number of firms accessing HPC may be lower than the number used in the analysis. Due to the limited number of survey responses from industry, no reliable measure of the proportion of industry users accessing multiple facilities could be derived.

On the other hand, no data on industry access for the five regional Tier-2 centres was available. As such, benefits of industry access for these centres are excluded from the analysis, meaning that the overall number of firms used in the analysis may actually be an underestimate. Moreover, given that regional Tier-2 centres were operating for more than five years, whereas new Tier-2 centres have only been operating for a bit over a year, suggests that excluded benefits of industry users accessing Regional Tier-2 centres may outweigh benefits of firms accessing multiple services.

[92] 0.5 between 2008 and 2016, based on data obtained from the EC Structural Business Statistics database.

[93] Note that data on industry access for regional Tier-2 centres was not available. As such these benefits are not included in the calculation. Moreover, new Tier-2 centres have only been operational for about 1/3 of their lifetime. As such the lifetime benefit of these centres is expected to be significantly higher. Multiplying the estimated benefits by 3 suggests that the benefit of new Tier-2 centres could be between £0.4 billion and £1.2 billion over their lifetime.

Table 10 Contribution of benefits to industry to UK output

Service	Assumed no. of firms accessing HPCs	Operational time of services (years)	Total annual contribution to GVA		Total contribution to GVA over operational time of service	
			Low estimate (0.6%)	High estimate (1.7%)	Low estimate (0.6%)	High estimate (1.7%)
ARCHER	150	5.2	£ 95.5 m	£ 301.0 m	£ 493.1 m	£ 1,554.3 m
HECToR	81	6.2	£ 46.5 m	£ 146.6 m	£ 289.1 m	£ 911.4 m
New Tier-2 centres*	125	1.4	£ 89.9 m	£ 283.2 m	£ 124.5 m	£ 392.4 m
Regional Tier-2 centres**	-	5.2	-	-	-	-
Total	**356**	**-**	**£ 231.9 m**	**£ 730.8 m**	**£ 906.7 m**	**£ 2,858.0 m**

Note: Profit increases were converted to GVA contributions using data on the average ratio of profit to GVA in the UK business economy obtained from the EC Structural Business Statistics database. (*) Note that new Tier-2 centres have only been operational for about 1/3 of their lifetime. As such the lifetime benefit of these centres is expected to be significantly higher. Multiplying the estimated benefits by 3 suggests that the benefit of new Tier-2 centres could be between £0.4 billion and £1.2 billion over their lifetime. (**) Data for the regional Tier-2 centres was not available.

Source: London Economics

6 Aggregate economic impact of HPC

Aggregating across all strands of impact, **the total economic impact of HPC, over the operational time of the services, is estimated to between £3.0 billion, and £9.1 billion.** (Table 11)

In comparison the total cost of ARCHER stood at approximately **£82.4 million** (EPSRC and NERC) in 2018 prices, the total cost of HECToR stood at approximately, **£137.8 million** (EPSRC, NERC and BBSRC) in 2018 prices, and EPSRC's contribution to the six new and the five regional Tier-2 centres stood at approximately **£21.5 million** and **£12.8 million**, respectively. (Table 12)

In addition, EPSRC invested approximately **£163.1 million** in HPC related research grants, **£13.6 million** in HPC related fellowships, and **£13.0 million** in research grants to the scientific consortia between 2008 and 2018. Further public funding, e.g. from other UK research councils or UK public bodies, for EPSRC grant holders reported back to EPSRC amounted to **£21.6 million**.

The return on these investments is thus between 6.5:1 and 19.5:1. The actual impact of HPC will lie somewhere between the low and high estimates presented here. However, due to the significant challenges and uncertainty inherent to assessments of scientific R&D investments, a point estimate is not provided.

Box 15 Note on estimated benefits and costs

While the study was focused on EPSRC's investments in HPC, the estimated economic impacts include benefits arising from funding by multiple sources including EPSRC, NERC, BBSRC, RCUK, the UK Government and the Tier-2 centres and partner institutions (see Section 1.3).

Nevertheless, the study does not provide comprehensive coverage of benefits arising from HPC research funded by non-EPSRC sources. In particular, spillover benefits of HPC research were estimated only for research funded directly by the EPSRC, and further funding given to EPSRC grant holders by other bodies, if these were reported to the EPSRC. Research activities using HPC systems funded by other UK Research Councils, other public bodies, or private organisations are not included in the estimate *unless* the funding was fed back into EPSRC's research outcomes systems.

In line with benefits, cost estimates represent the cost to the public purse, in 2018 prices, and include the total cost of ARCHER (funded by the UK Government, EPSRC and NERC); the total costs of HECToR (funded by RCUK, EPSRC, NERC, and BBSRC); EPSRC's funding of HPC research and other public purse costs of funding of EPSRC grant holders reported back to the EPSRC; and EPSRC's contribution to the Tier-2 centres.

Tier-2 centres and the hosting institutions further carry part of the capital expenditure and all of the operational expenditure themselves. This cost is assumed not to be carried by the public purse and is thus not included in the cost calculation. In reality, some costs may ultimately be covered by the public purse. Costs to the public purse may thus be higher than those reported here.

On the other hand, as discussed above, only benefits of research by EPSRC grant holders were included in the calculation of benefits. Moreover, benefits of software to industry were also not monetised (see discussion below).

Source: London Economics

Table 11 Aggregate economic impact of EPSRC* HPC investments in the UK (£m, 2018 prices)

Type of impact (£m in 2018 terms, over operational time of services)	Low estimate (£m)	High estimate (£m)
Impact of UK scientific research and discovery	**£ 2,021.1 m**	**£ 6,051.1 m**
Avoided cost of free HPC access for academics	£ 213.3 m	
Spillover impact of HPC research on UK output (EPSRC grant holders only)	£ 1,807.7 m	£ 5,837.7 m
Impact of direct industry access	**£ 906.7 m**	**£ 2,858.0 m**
Contribution of industry impacts to UK output	£ 906.7 m	£ 2,858.0 m
Impact of training and skills development	**£ 101.2 m**	**£ 190.3 m**
Benefits of PhD and postdoc training of students entering industry to students and the UK exchequer	£ 99.0 m	£ 188.1 m
Benefits of provision of free HPC training courses	£2.2 m	
Total economic impact	**£ 3,029.0 m**	**£ 9,099.4 m**

Note: The true benefits of improved or new software is the contribution that this software makes to scientific research & discovery. However, these benefits are already implicitly monetised when estimating the spillover benefits of EPSRC funded HPC research. Therefore, to avoid double counting, software was excluded from the calculation of total aggregate benefits. Similarly, the true benefits of skills development in HPC is the value brought to UK companies and UK science. However, these benefits are already implicitly monetised in the benefits to industry and the spillover impacts of HPC research. In addition, software has benefits on industry as well, these benefits were not quantified. All estimates are rounded to the nearest £0.1 m. * See Box 15, some of the impacts captured could be from NERC funded research.
Source: London Economics' analysis. Icon credits: bioraven – Shutterstock

Table 12 HPC investments (£m, 2018 prices)

Type of investment	EPSRC	NERC	BBSRC	RCUK / UK Government	Total
Investments in centres	**£ 109.6 m**	**£ 22.8 m**	**£ 3.3 m**	**£ 119.0 m**	**£ 254.6 m**
ARCHER	£ 27.3 m	£ 8.3 m		£ 46.8 m	£ 82.4 m
HECToR	£ 47.9 m	£ 14.4 m	£ 3.3 m	£ 72.1 m	£ 137.8 m
New Tier-2 centres[1]	£ 21.5 m				£ 21.5 m
Regional Tier-2 centres[1]	£ 12.8 m				£ 12.8 m
Public purse costs of EPSRC research funding	**£ 189.7 m**			**£ 21.6 m**	**£ 211.3 m**
EPSRC HPC research funding[2]	£ 189.7 m				£ 189.7 m
Further public funding of EPSRC researchers[2]				£ 21.6 m[3]	£ 21.6 m
Total	**£ 299.2 m**	**£ 22.8 m**	**£ 3.3 m**	**£ 140.5 m**	**£ 465.9 m**

Note: (1) Investment in Tier-2 centres only includes EPSRC investments; i.e. excluding CAPEX and OPEX costs to the centres themselves and hosting institutions; (2) Estimates of funding are based on EPSRC's research outcomes systems, which records quantitative information only for research activities funded directly by the EPSRC, and further funding given to EPSRC grant holders by other bodies, if these are reported to the EPSRC. Research activities using HPC systems funded by other UK Research Councils, other public bodies, or private organisations are not included in the estimate *unless* the funding was fed back into EPSRC's research outcomes systems. Only costs of further funding accruing to the public purse were counted in the calculation. This includes funding provided by UK research councils and other UK public bodies. (3) May include funding by NERC or BBSRC.

Source: London Economics' analysis based on data provided by EPSRC

In terms of the components of this economic impact, the impact of UK scientific research and discovery was estimated to be between £2.0 billion and £6.0 billion, while the impact of direct industry access was estimated to be in the region of £0.9 billion and £2.9 billion.

Estimated benefits of software (Section 4.2.1) as well as of training and skills development (Section 4.3.3 and 4.3.4) are relatively small. However, this does not mean that these areas are not important. Indeed, the impacts of software and of training and skills development were identified among the most important benefit areas.

Better software aids researchers in scientific research and discovery, and ultimately allows researchers to undertake better or more science, for example by allowing researchers to simulate at a more granular level. The real impact of software is thus the contribution it makes to scientific research and discovery. The benefits of better software are thus already implicitly captured in the assessment of benefits of scientific research and discovery. Therefore, the benefits of software estimated in Section 4.2.1 were excluded from the aggregate figures presented in this section.

Moreover, HPC software may also bring significant benefits to industry. This is highlighted by CASTEP (see Box 9), an HPC simulation software used to calculate the properties of materials from first principle, which is used by around 900 industrial users. As the Goldbeck report[94] indicates, companies using materials modelling software such as CASTEP, among others, achieved cost savings ranging from €100,000 to €50 million, with an average return on investment of 8:1. Benefits of software to industry were not quantified in this study.

Similarly, the nominal value of training courses as well benefits to students and exchequer from higher earnings are important benefits. However, the real value of skills development comes from the benefit skilled people bring to UK science, as the next generation of HPC trained researchers, as well as UK industry. Again, these benefits are, at least to some degree, already captured in the estimations of the benefits of scientific research and discovery and the impact on industry. Therefore, these benefits were not assessed separately.

[94] Goldbeck and Court (2016). *The Economic Impact of Materials Modelling.*

7 Outreach, diversity and inclusion

"HPC suffers from the general diversity issues of STEM but even more so by being seen as a geeky computer science subject. To me HPC is a tool to do science, no different from doing experiments. We have to show that you don't have to be a stereotypical computer geek to use HPC successfully." – survey respondent

Several of the subject areas within EPSRC's remit have historically been male dominated. Moreover, computing itself has also historically been a male-dominated area. As HPC sit at the cross-over of these disciplines it is perhaps no surprise that HPC itself suffers from diversity issues.

Recognising the importance of outreach and diversity, the EPSRC is committed to supporting a diverse and inclusive research environment with equal access to opportunities. In 2018, EPSRC funded eleven projects addressing equality, diversity and inclusion within engineering and the physical sciences with a total of £5.5 million in funding via its Inclusion Matters call.[95]

Within the HPC community, outreach and diversity is also an important area for the EPSRC, with a number of outreach and diversity programmes and initiatives supported by the EPSRC as well as by Tier-2 centres. In particular, EPSRC has supported outreach activities at the EPCC, who host ARCHER, with a £450,000 grant. Diversity, outreach and inclusion activities at the EPCC, revolve around four key themes: diversity, public outreach, training and user engagement activities.

The **Diversity in HPC** programme supports the inclusion of under-represented groups working in the HPC community by encouraging participation and showcasing that HPC is a career path available to everyone. One particular example of the programme's work is **Faces of HPC,**[96] an initiative celebrating diversity in the UK HPC community by highlighting the profiles of individuals with a wide range of backgrounds involved in the HPC field.

"Keep up the good work! I'm a woman and it does make an impact - it's never nice being the only women in the room and it still happens far too often to me!" – survey respondent

Women in HPC encourages and promotes a more diverse and inclusive HPC community by encouraging women to engage in the HPC community. Women in HPC brings together women through a number of outreach activities including raising awareness, collaborations and networking, and providing fellowship, education, and support to women, as well as the organisations employing them.

Recognising the efforts of Women in HPC the programme has won the HPCwire *Readers' Choice for Workforce Diversity Leadership* award in 2015, an international award for highlighting issues of diversity within HPC[97]. In 2016, Women in HPC was again recognised with the *Readers' Choice: Workforce Diversity Leadership Award,* the *Editors' Choice: Workforce Diversity Leadership Award.* Moreover, as well as the *Readers' Choice: Outstanding Leadership in HPC* for Toni Collis, the chair and co-founder of Women in HPC[98].

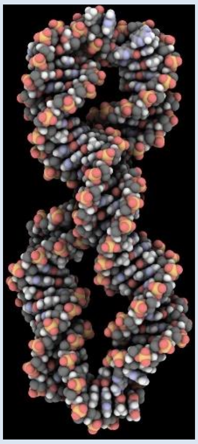
[95] EPSRC (2018). *Inclusion Matters - EPSRC announces £5.5 million for projects to address Equality, Diversity and Inclusion*. Available at: https://epsrc.ukri.org/newsevents/news/inclusionmatters/ [accessed 28/03/2019]
[96] The initiative can be found here: https://www.hpc-diversity.ac.uk/faces-of-hpc [accessed 28/03/2019]
[97] EPSRC (2015). *Diversity Award for Women in High Performance Computing.* Available at: https://epsrc.ukri.org/newsevents/news/diversityawardforwomeninhpc/ [accessed 28/03/2019]
[98] *Women in High Performance Computing (WHPC) Receives Honours in 2016 HPCwire Readers' and Editors' Choice Awards.* Available at: http://www.archer.ac.uk/about-archer/news-events/news/2016-11-15-WHPC-HPCwire-Honours.php [accessed 28/03/2019]

Box 16 Computational models to generate innovative drug designs

Dr Sarah Harris is a researcher at the University of Leeds working in the Theoretical Physics Group. She uses HPC to model the physical behaviour of biological macromolecules with *biomolecular simulation.* This research area is important to answering key questions in drug design because it involves simulating the atomic structure of biomolecules to understand their interactions.

Medicines work by interacting with the biomolecules which make up living matter, such as proteins and DNA. As part of a vibrant biomolecular simulation community, Dr Harris designs better models to understand the physical interactions of biomolecules to provide more powerful design tools for drugs to combat disease. The most promising solution is to build more accurate computational models, which requires expanding the limits of current supercomputing capabilities to handle more advanced computation. This would lead to more efficient methods to design medicines and could reduce the need for animal testing.

Furthermore, in separate research, she is designing computer models which are being used to predict whether an observed genetic variation is likely to cause disease. By studying an atomic model of how the protein is affected, Sarah and her group try to predict whether its function will be disrupted. This will increase our understanding of inherited diseases.



Both research strands have the potential to transform the way medicine works and provide radical solutions to tackling disease. Biomolecular simulation and HPC is so valuable to the pharmaceutical industry as many possible molecular designs for drugs can be tested rapidly and efficiently using computer modelling. Furthermore, it reduces the need for costly experiments to be run on every single compound, an unfeasible task, instead allowing experiments to be prioritised based on the research from the modelling.

In addition to her research, Dr Harris has worked on making her simulations available in film-format which can be used in outreach programmes to educate the general public. One such film was featured at the Science Museum and this medium has the potential to be used in counselling by informing patients about their diseases and treatments in the future.

HPC has given Dr Harris the platform to disseminate her own research and has provided opportunities to collaborate with other academics. This has resulted in the formation of strong international contacts which has further progressed her research.

Source: Interview with Sarah Harris. Picture credits: Computational Biophysics Group University of Leeds

EPCC also runs a wider **Public Outreach Programme** comprising of a range of public engagement activities. A highlight of this programme is the **Wee Archie** initiative aimed at engaging with younger audiences. Wee Archie is a suitcase-sized supercomputer modelled after its bigger counterparts. The mini supercomputer was designed to explain what a supercomputer is and is showcased during events such as science festivals and conferences, and other outreach activities such as visits to schools. Wee Archie was so successful, that EPCC in 2018 had to build a second miniature

Figure 34 Wee Archie

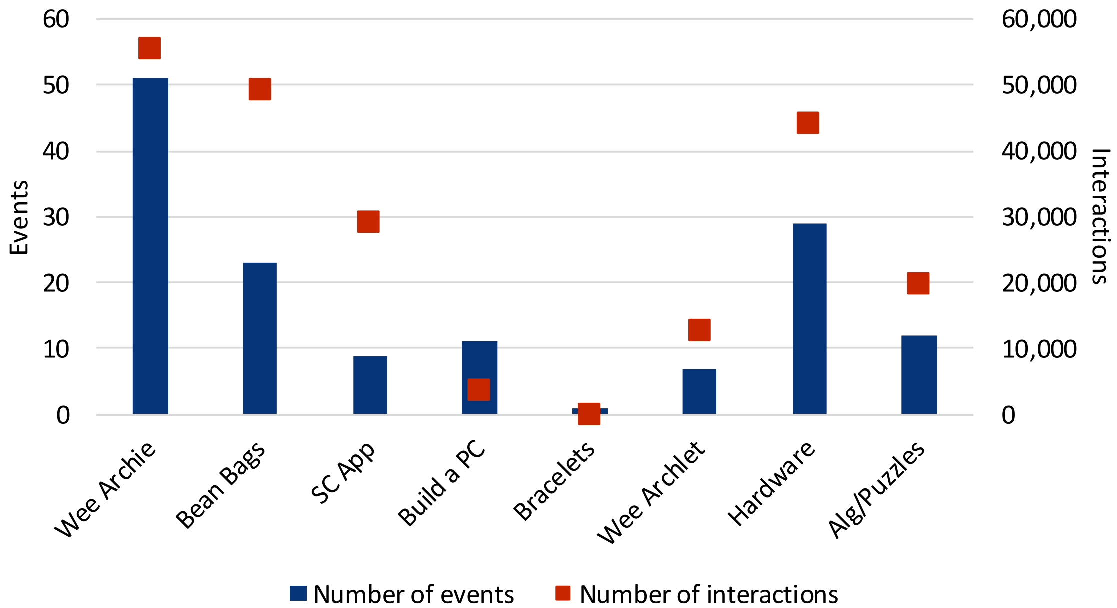
Source: EPCC; Picture credit: Maverick

supercomputer[99]. A series of demonstrations have been developed to showcase the science being carried out on ARCHER. These include a demonstration that allows you to design your own aircraft wing; one that allows you to model waves and design your own coastal flood defences; the Mouse Wee demonstration which showcases molecular modelling; and the dinosaur racer.

Other activities within the programme include: a build-a-PC workshop designed to explain how technology is built; a bean bag sort activity to showcase parallelism and algorithms; a series of interactive puzzles to help understand programming concepts; Wee Archlet, a miniature build your own supercomputer and a supercomputing game app that allows you to run your own HPC Centre.

To gain a better understanding of the impact of the public outreach programme, the number of people who have actively participated in an event (and activity) has been recorded (Figure 35). This shows that since the start of the Outreach funding, in April 2015, the activities have been used over 58,000 times across 55 different locations in the UK and 6 international venues. This includes over 75 different events such as the Big Bang Fair in Birmingham, New Scientist Live in London and the Edinburgh International Science Festival.

Figure 35 Public interactions with different activities, April 2015 – May 2019



Note: Different activities were present at different events. Interactions refers to the number of people engaging with a given activity.

Source: EPCC

[99] Gibb, G. (2018). *Wee Archie earns its wings!* Available at: http://www.epcc.ed.ac.uk/blog/2018/07/24/wee-archie-earns-its-wings [accessed 28/03/2019]

Box 17 EPSRC Early Career Fellows

Dr. Agnes Noy is an EPSRC Early Career Fellow in the department of Physics at the University of York investigating the physical and chemical properties of different biomolecules like DNA and proteins through the use of molecular modelling techniques. Because her degree was in biochemistry, Dr. Noy became familiar with the basics of simulations during her PhD with Prof. Modesto Orozco at the Institute for Research in Biomedicine Barcelona (IRBB). Towards the middle of her PhD, she gained access to the brand-new supercomputer *MareNostrum* established in Barcelona, which was the most important machine across Europe at the time of its inauguration. That allowed her to raise the level of her research, leading to the publication of several papers in *Nucleic Acids Research* and *Journal of the American Chemical Society*. She was also invited in a review at the journal *Current Opinion in Structural Biology*.

For her post-doctoral stage, Dr. Noy moved to the UK, where she continued developing her research career and her skills on HPC. She was the main person responsible in managing allocation to ARCHER in a BBSRC grant awarded to the theoretical physicist Dr. Sarah Harris (University of Leeds) and to the biological chemist Prof. Tony Maxwell (John Innes Centre, Norwich) for performing cutting-edge simulations and comparing them with experiments.

In the course of her career, Dr. Noy has published a total of 19 articles, 11 of which made use of a supercomputer, including Tier-1 machines like ARCHER or MareNostrum, Tier-2 computers like N8 Polaris, JADE and CSD3, and the local clusters of Oxford, Leeds and York universities. Dr. Noy's papers, containing original research using HPC, have been published in journals like *Nature Methods* and *Physical Review Letters* and have collected more than 500 citations to date.

Source: Dr Agnes Noy

To ensure that ARCHER users get the most out of HPC, **training programs** have been developed to support HPC users at all stages of their HPC careers (see Section 4.3.4 for additional information).

The core component of this training is face-to-face classroom based teaching in academic institutions throughout the UK. Around 65 days of this training are provided each year with 2-3 day courses led by members of EPCC, supported by specialist tutors. Additionally, online training is provided, often from recordings of the face-to-face courses and a five-week MOOC called "Supercomputing" has been run since 2017 on the FutureLearn platform.

To measure the short and long term training impact, online questionnaires are sent to attendees both immediately after the event and 6 months later. Feedback from the 2018 questionnaires are generally very positive with strong evidence that attendees find the courses of immediate use to their work and that they spread their newfound knowledge amongst colleagues[100] (see Section 9.2.2 for further evaluation on the quality of trainings provided).

[100] EPCC, ARCHER CSE Service Report of Fifth Training Impact Survey, June 2018

User engagement activities include the ARCHER Champions, image and calendar competition, ARCHER Driving Test, easy-access training and the Supercomputing Massive Open Online Course (MOOC).

The **ARCHER Champions** initiative provides a self-sustaining user community where an individual Champion in a given community (e.g. university or consortia) can offer practical, informal, advice to their peers on accessing and effectively exploiting computational resources. Meetings occur between these Champions to discuss ways at improving access and best practices for engagement. Since the program was set up, three meetings have been held across the country with a total of 82 attendees.

Another popular engagement activity is the **image and calendar competition** where the ARCHER user community is invited to submit images which demonstrate their research. This provides a wealth of material for outreach and engagement activities and the images are collated into an annual calendar which is distributed to key HPC stakeholders.

Accessing HPC for the first time can be daunting for new users, particularly when they have no significant previous computing experience. To overcome this the **ARCHER Driving Test** gives new users basic access to ARCHER for a year, to do exploratory work and become comfortable with the use of ARCHER, conditional on completing a short test to ensure users have sufficient knowledge of HPC and ARCHER. To help users pass the test introductory materials are readily available online. The Driving Test has been very successful, with over 350 active users using over 200,000 kAUs of ARCHER. The scheme has been most successful with PhD students and those in the early stages of their career and is therefore likely to be additional rather than displacing.

While these findings are encouraging, diversity in the HPC field remains a challenge. As such, continuing efforts in improving outreach, diversity and inclusion within the HPC community remain highly important.

8 The impact of the no HPC scenario on UK science

To understand the impact that a loss of HPC capability would have on the UK science community, the No HPC counterfactual was explored with stakeholders via workshops, interviews as well as via the online survey of users. Specifically, users were asked to imagine a scenario where UK research councils would not fund HPC. In this scenario[101]:

- National or Tier-2 HPC capabilities do NOT exist in the UK.
- Research councils also do not fund potential substitutes for these capabilities (e.g. the research councils / UKRI would make no investments in cloud computing).
- Research councils do not provide funds to access commercial / international cloud capabilities or HPC capabilities.

The impact of this scenario on UK science is expected to be significant. High Performance Computing has become a very important part of scientific research, complementing or indeed underpinning the other parts of the research process.

Most research undertaken on HPC facilities is scientific in nature and produces benefits which are not immediately commercialised or cannot be captured by a single entity; e.g. fundamental research which disperses across the economy. Meanwhile, HPC investment requires large upfront capital investments and continuing operational expenses. As a result, it is unlikely to be commercially viable for private firms to provide commercial HPC infrastructure for scientific research and discovery. This is explored in further detail in Annex A1.2.1.

Another possibility, in the no HPC scenario, is for HPC users to substitute to other HPC facilities, including other public HPCs in the UK or overseas. This would result in oversubscription and hence rationing for HPC time. Only a select number of users would be able to gain access to overseas facilities. Therefore, it is unlikely that researchers would be able to substitute publicly funded HPCs. For more information on these counterfactuals refer to Annex A1.2.2.

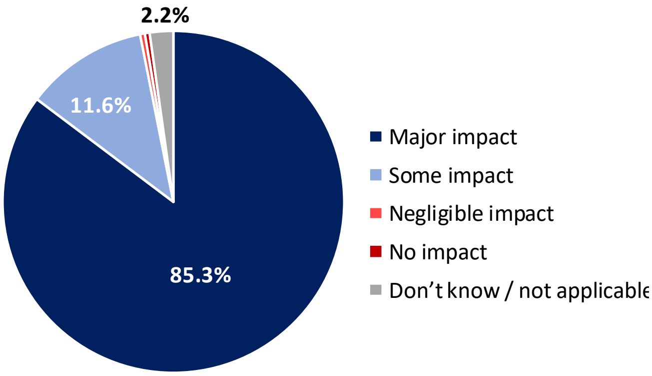
A loss of publicly funded HPC facilities would have a direct negative impact on the research of users of these facilities. Indeed, when asked about the impact that the counterfactual scenario would have on their research or work, **85% of users** of HPC facilities, that responded to a survey undertaken for this study, **said that this scenario would have a major impact**. A further **12%** said that this scenario would have some impact on their research or work. (Figure 36)

"If ARCHER and the Tier-2 centres were turned off it would have a major detriment for everyone. We'd limp on for a year or two but after that the whole thing would collapse. Not immediately but it would come crashing down in a year or two." – workshop participant

"We would lose our leadership in several key scientific areas, and a lot of users who are dependent on HPC availability would likely leave the UK to find employment in places that do provide HPC access. A lot of these researchers would be unable to do so and would thus leave academia." – survey respondent

[101] The counterfactual is explored in more detail in Annex A1.2. It should be noted that the precise counterfactual is that there are no pubicly funded HPC. However, for simplicity, and as funding of similar HPC facilities by other research councils would have incurred similar costs to the public purse, users were asked about the more general scenario in which none of the research councils fund HPC facilities.

Figure 36 'What impact would this scenario have on your research / work?'



Note: No. of respondents = 224

Source: London Economics survey of users of HPC capabilities

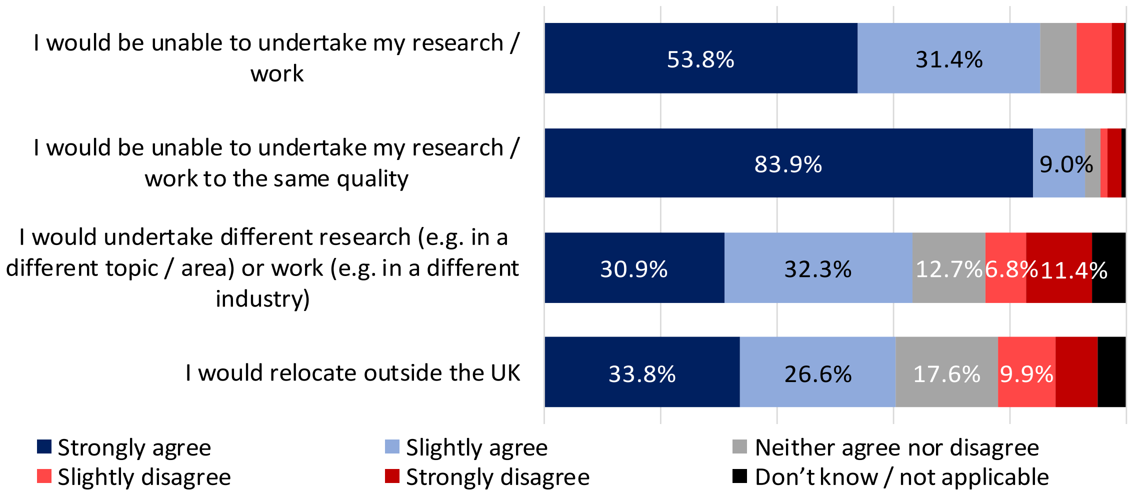
> **"We would take a lot less risk. The product of what we do might eventually go commercial and be run on a large under-the-desk server, but that's the result. We would have to go down blind alleys and if each of them took a month we wouldn't even start that enquiry. Having access to HPCs means we can do this in a few days which is viable. We can take risks and fail quickly even if the result isn't the code that will run." – workshop participant**

Of those users that agreed that there would be at least some impact on their work, **93% agreed**, either slightly or strongly, **that they would be unable to undertake their research or work to the same quality**; and **85%** agreed, either slightly or strongly, that they would be unable to undertake their research or work at all. (Figure 37)

The direct impact on researchers would in turn also lead to knock-on impacts on the quality and competitiveness of UK science. Given the impact of the counterfactual scenario on their work, **63%** of respondents agreed, either slightly or strongly, that **they would undertake different research or work**; and **60%** agreed, either slightly or strongly, that **they would relocate outside the UK** under this scenario. (Figure 37)

> **"It will make the UK less competitive in a wide range of things, we would need to make things less complex, less accurate. It will make UK industry non-competitive." – workshop participant**

Figure 37 'In this scenario, to what extent do you agree or disagree that …?'



Note: No. of respondents = 223. ***Source: London Economics survey of users of HPC capabilities***

"The scope of scientific research in the UK would be massively restricted. Simulations needed to produce publishable interpretation or prediction of experiments or observations would not be possible." – survey respondent

This loss of talent was also highlighted by workshop participants who emphasised the importance of HPC to the competitiveness of UK science. Without significant supercomputing resources, many researchers would be unable to explore the kind of questions they are able to explore now. This would limit the ambition of UK researchers, resulting in less blue-sky thinking, and limiting the capability of UK science to be world-leading. As a result, the UK's reputation for the quality of its science base would suffer and the best researchers and students may consider whether to pursue their careers elsewhere.

"It's straightforward, anyone young would leave. The UK has conned itself into this idea we get the best academics in the world. That's changed now, people go where the resources are." – workshop participant

Without adequate HPC facilities, the UK would also find it much more difficult to take part in international HPC collaborations such as PRACE. A loss of HPC facilities would therefore also limit UK researcher's ability to access international HPC facilities and collaborate with other researchers internationally.

"Companies who need to do advanced simulations would move somewhere else. They can do it anywhere, but they are doing it in the UK because we have the HPC facilities." – workshop participant

However, the impact would not only be limited to UK science, but affect other areas too. In particular, the UK has a very strong HPC software and HPC skill base developed over decades. Without adequate supercomputing facilities, these capabilities would also suffer and over time lose their international competitiveness.

Moreover, a degradation of the UK's HPC skills base will also have impacts on industry and other organisations which require those skills. Without appropriate HPC facilities and access to skilled staff, firms may ultimately decide to invest, or even locate, elsewhere. It was highlighted that without HPC, there will be less training of domestic talent and the entire eco-system would breakdown such that HPC experience could be lost permanently.

"It makes you think whether there are areas in the UK where the big computer firms have invested because they are local to HPC. You wonder if that level of investment would have occurred without it." – workshop participant

9 The EPSRC model: Comparison to other models of HPC provision

Section 9.1 provides a comparison of the model of HPC provision in the UK compared with **selected international examples** including **IT Centre for Science (CSC)** in Finland, the **Juelich Supercomputing Centre (JSC)** in Germany, **SURFSara** in the Netherlands and the XSEDE project coordinated by the **National Science Foundation (NSF)** in the United States. The comparison focuses on five key features of each model: service provision, financing, operating model, access and usage, and industrial access. Further details of the model of HPC provision are provided in Section 9.2, for the UK, and Section 9.3 for other international models of HPC provision.

9.1 UK EPSRC model vs. international models

HPC service provision

For the purpose of this study, only HPC centres focused on the scientific research community were considered. Therefore, their services are, as a rule, publicly available for researchers. Nevertheless, service provision of HPC varies across the selected countries.

In the UK, academic HPC provision is largely handled by the research councils. ARCHER is part funded by EPSRC and NERC and hosted by the EPCC at the University of Edinburgh. The Tier-2 centres were established in response to EPSRC grant calls for the establishment of Tier-2 HPC centres. Tier-2 centres are hosted by one, or a partnership of several, academic institution(s).

In the US, provision of academic HPC services is coordinated by an independent federal agency, the National Science Foundation (NSF). The NSF funds major new machines through direct grants with institutions and offers them to the community via the XSEDE project.

In contrast to the UK and the US, Finland only has one sole centralised national HPC services provider, the IT Centre for Science (CSC). CSC is registered as a private non-profit company with the state being its partial shareholder.

In the Netherlands, national HPC services are provided by SURFsara. SURFsara, is managed by a co-operative association, SURF, of Dutch educational and research institutions who are the owners of SURF.

Supercomputing provision in Germany follows a similar Tier-structure to the UK, with HPC provision at a national, regional and university level. The national supercomputing centre in Germany is the Gauss Centre for Supercomputing, consisting of the Juelich Supercomputing Centre (JSC) and two other HPC hubs in Garching and Stuttgart.

Financing and operating model

As detailed in Section 1.3 the total capital investment in ARCHER was £43.0 million by the UK Government. The operational costs of ARCHER were £34.2 million over the five-year lifetime. Operational costs for ARCHER are shared between EPSRC and NERC in a 77 : 23% partnership respectively.

In terms of the Tier-2 centres EPSRC invested £11.6 million in five regional Tier-2 Centres of Excellence in 2012-13 and a further £20.6 million in six new Tier-2 Centres in 2016. This funding covered part or all of the hardware costs. Operational costs of the Tier-2 centres are covered by the centres or their host institutions themselves.

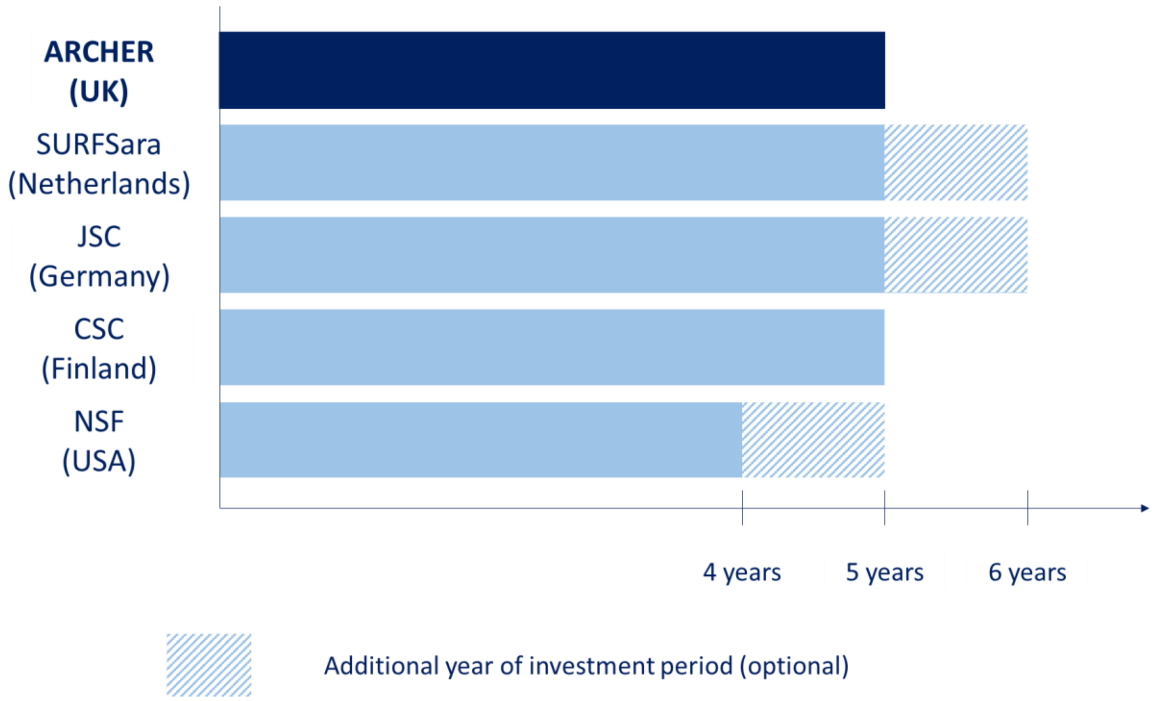
The most recent capital investments by the Finnish CSC and Dutch SURFSara were lower than the investments in ARCHER, with total capital expenditures of £25.4 million (€29 million) and £14 million (€16 million), respectively.

In Germany, funding received by the Juelich Supercomputing Centres for the national supercomputer is close to £128.4 million (€150 million) over the period between 2017 and 2025.

In the US, the NSF would historically invest a sum between $50-$60 million in new HPC hardware and operational budgets, plus additional funding of approximately $20-$30 million dedicated to the XSEDE project. In the upcoming years the XSEDE anticipates the NSF to fund approximately four new HPC machines with the total anticipated cost of around $30 million that will be integrated into the XSEDE distributed environment.

Figure 38 illustrates the average replacement cycle of the HPCs; i.e. the average operational time before new hardware is acquired. It should be noted that in some cases midlife upgrades may be made.

Figure 38 Replacement cycle of HPCs



Source: London Economics' analysis based on data provided by EPSRC and selected HPC international centres

Similar to the UK, almost 100% of the funding in the examined centres comes from state- or EU-funded grants. In some cases, comparatively small or occasional investments are made by third parties.

All the examined international HPC centres have to provide a scientific case to justify further investment. In the US, NSF grants are awarded to HPC centres who won publicly announced competitions for funding. None of the centres are required to provide evidence for HPC centres' economic impact to support their application for funding. For example, the German JSC uses the number of research papers published as their main measure of success.

In the UK, ARCHER operational costs are jointly covered by EPSRC and NERC, while operational costs for Tier-2 centres are covered by the centres themselves or their partner institutions. In the Dutch, Finish and German centres evaluated, the day-to-day operations are funded using dedicated resources from their approved annual budgets. In the US, operational costs for NSF-funded HPC centres are funded as part of an additional grant beyond the hardware acquisition grant, with an annual budget of up to 20% of the hardware acquisition costs (and in some cases 25%).

Access and usage

Access to HPC public systems, both in the UK and for the international centres used as benchmarks, is free of charge for academic researchers.

In order to receive access, researchers need to submit a research proposal directly to the HPC centre or through a research council. The proposal review procedure is similar across the examined centres; similar to the UK, research requests are assessed through a peer-review process (JSC in Germany and NSF in the US) or managed centrally by the organisation (CSC in Finland).

Unlike ARCHER, which is predominantly used for research in EPSRC and NERC funded disciplines, the Finnish CSC and Dutch SURFSara grant access to supercomputers to researchers working across all scientific domains based on merit, with no preference over their field of specialisation. Similarly, the German JSC and US NSF do not limit their services to researchers from specific scientific disciplines.

Industrial access

Private businesses in the UK can access ARCHER via a number of routes such as research collaborations, working within a scientific consortia or via assistance on a consultancy basis. Some Tier-2 centres further offer industry access via a fixed fee.

A similar approach can be found in the selected Finnish, Dutch and German models, were private users are charged a fee for using the HPC service. In the Finnish and Dutch centres, the income achieved through this channel is re-invested. Private users in these centres constitute between 5% to 10% of all HPC users.

In the US, the NSF excludes private access from its scope. Though, US education institutions accessing NSF HPC resources may support industry work, under the condition that the research outcomes are published.

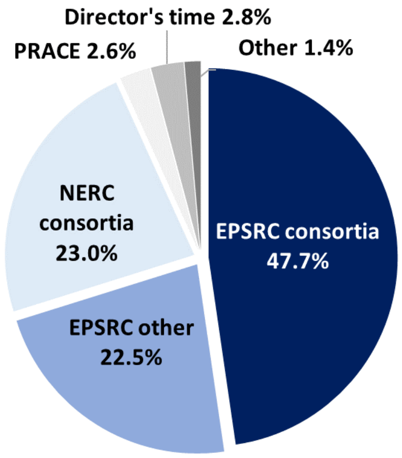
9.2 UK model

Access and usage

The classification of past and existing HPC centres has been introduced in Section 1.2. Access to the Tier-1 facility, ARCHER, can be gained through one of the eight EPSRC-funded and three NERC-funded High-End Computing consortia (HEC). Each of the consortia focuses on a different area of computational science (Annex A3.3). The HEC consortia act as the central point of HPC access and constitute the leading networking hubs for researchers. Consortia do not only enable new science but also facilitate collaborations with international and non-academic users. This allows for a high degree of flexibility in the allocation of machine time slots and aims to maximise efficient usage of the HPC resources.

As an alternative to accessing ARCHER through consortia, HPC researchers can apply for access to ARCHER through the Resource Allocation Panel (RAP). The RAP comprises of selected members of the Engineering and Physical Sciences community who collect and evaluate all research proposals sent in response to the published calls.

Figure 39 Proportion of time used on ARCHER, 2014-2017



Note: Time used is calculated in terms of kAU (kilo Allocation Units). 'Other' includes BBSRC, eCSE, CSE, industrial and internal usage. Part of industrial use is also captured via Director's time. Almost all of the time used by NERC was allocated to and used by the consortia. ***Source: London Economics based on ARCHER Quarterly SAFE reports Q1 2014 – Q4 2017***

Between 2014 and 2017 approximately **47.7%** of the usage of ARCHER was through EPSRC scientific consortia, with a further **23.0%** of usage through NERC scientific consortia.

Other EPSRC allocations such as allocations for Leadership Awards, direct access via the Resource Allocation Panel, instant access, training and other activities accounted for a further **22.5%** of ARCHER use. The remaining time was split between international usage via the PRACE initiative (**2.6%**), Director's time (**2.8%**), and Other activities (**1.4%**). (Figure 39)

In contrast, Tier-2 centres provide a somewhat less formal method of access, which is particularly beneficial for newer HPC users. Access to the Tier-2 centre is open to all academic users and can be gained by making an application via the twice-yearly EPSRC Tier-2 Resource Allocation Panel.

Industrial usage

Since the beginning of its operations in 2013, ARCHER has been used by more than 150 industrial users. Commercial organisation can gain access to ARCHER via a number of routes, for example through research collaborations, working within the scientific consortia or by asking for assistance on a consultancy basis. Industry access for Tier-2 centres varies by centre, with some centres engaging heavily with industry users while other centres focus solely on academic researchers.

Financing

Financing of Tier-1 and Tier-2 centres was detailed in Section 1.3. Therefore, only a brief overview is provided here. Capital costs of ARCHER were funded by the UK Government, with the operational costs split between EPSRC (77%) and NERC (23%). Tier-1 centres are replaced, roughly, on a five-year basis. For Tier-2 centres, EPSRC provided significant parts of the capital investments, with the remainder covered by the centres or their partner institutions themselves. Operational costs are carried by the in Tier-2 centres.

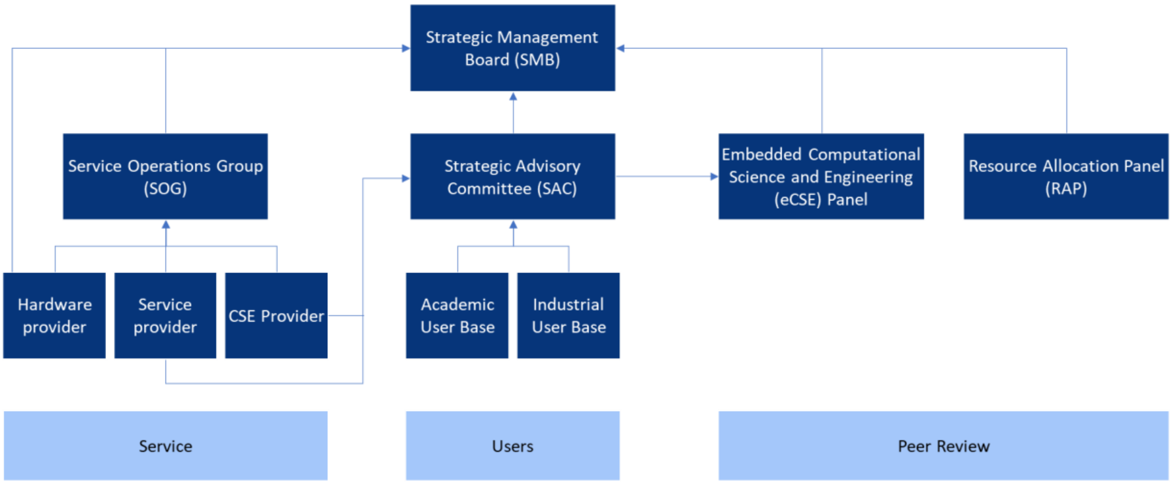
ARCHER's management structure

This subsection highlights the key components of ARCHER's management structure which facilitates its day-to-day operations as well as helps develop new strategies for the future.

The main bodies of the ARCHER organisational structure are:

1. The **Strategic Management Board (SMB)** responsible for the management of the ARCHER service. It is highly involved in all the operational aspects of the ARCHER service from the monitoring of operational risks to the introduction of changes in operational policy recommended by the Scientific Advisory Committee (SAC).
2. The **Scientific Advisory Committee (SAC)** is in charge of the ARCHER supervision process. The team meets every four to six months to discuss current and potential needs of the user base.
3. The **Service Operations Group (SOG)** ensures that all technical requirements for ARCHER are met through its collaboration with the hardware and service providers.

Figure 40 ARCHER management structure



Source: EPSRC (n.d.). ARCHER management structure. Available at: https://epsrc.ukri.org/research/ourportfolio/themes/researchinfrastructure/subthemes/einfrastructure/highperformance/archer/structure/ [accessed 05/06/2019]

9.2.1 User support

User support is another important service provided by HPC facilities. The helpdesks of the respective centres are the first ports of call when users experience problems, dealing with a wide range of issues users experience. This ranges from user registration, change and admin requests to technical queries and resolving in-depth technical problems.

ARCHER's helpdesk was continuously very busy, recording more than **7,000** queries each year over the operational time of the ARCHER service (2014-2017). Self-service admin queries were by far the most prevalent, accounting for more than 5,000 requests each year. (Table 13)

The first year of services, also saw a significant number of technical queries, accounting for nearly one-third of all queries in that year. However, this is unsurprising given that users had to transition from HECToR. The number of both technical queries and in-depth queries passed on to CSE and Cray has substantially declined over time, suggesting the service has been working well for users.

Indeed, more than **98%** of all queries were resolved within two days across all years and users were generally pleased with the services offered by the helpdesk. This is evidenced by ARCHER's annual user surveys, in which users rated the helpdesk a **4.5 out of 5**, on average, in 2014, 2015, and 2016 and **4.6 out of 5** in 2017 (with 1 representing unsatisfactory and 5 representing excellent service)[102]. Helpdesk calls were also all answered within 2 minutes.

"I always get a quick response from ARCHER Helpdesk. I am grateful for the Helpdesk team for their support." – ARCHER User Survey respondent

Table 13 Number of technical and administrative queries received (ARCHER)

Type of query	2014	2015	2016	2017
Self-Service Admin	5,749	5,233	5,288	8,048
Admin	1,873	2,278	1,869	1,757
Technical	539	363	269	277
Total Queries	**8,161**	**7,874**	**7,426**	**10,082**
In-depth queries passed on to CSE and Cray	**> 300**	**296**	**222**	**136**

Source: ARCHER Service Annual Reports for 2014, 2015, 2016, and 2017.

In addition to technical and admin queries, the helpdesk also deals with more general queries, such as change requests and user registrations; an overview of these is provided in Table 14.

Table 14 Number of other queries received (ARCHER)

Query type	2014	2015	2016	2017
Phone Calls Received	599 (117)	431 (97)	299 (79)	367 (66)
Change Requests	38	25	23	12
User Registration Requests	851	1,049	1,084	1,043

Note: The numbers shown in brackets for the phone calls received represent actual ARCHER user calls resulting in new or updated queries.
Source: ARCHER Service Annual Reports for 2014, 2015, 2016, and 2017.

Tier-2 centres, or their partner institutions, also have their own help desks, ensuring adequate support is available for users of the relevant facilities.

9.2.2 Quality of training

To assess the quality of training, ARCHER users are asked to fill-in subsequent follow-up surveys. The attendees of all courses are surveyed over an extended period after finishing the training (from 3 to 15 months). The threshold of 3 months ensures that the researcher had enough time to implement the newly acquired HPC knowledge in their research. The main findings emerging from the user and training surveys are then presented in the quarterly reports as well as the detailed Training Impact reports received by EPSRC. In addition, the most recently developed approach attempts to put more emphasis on examining the long-term impact of ARCHER training.

The ARCHER Training Impact survey first took place in February 2015 and has been organised approximately every year since then. Each year respondents were asked 8 questions about the

[102] EPCC ARCHER User Surveys 2014, 2015, 2016, and 2017

trainings' impact on their work productivity and performance, skills improvement and HPC expertise sharing (see Table 15).

Table 15 ARCHER Impact Survey questions

Category	Survey question
Productivity and performance improvements	It has increased my overall understanding of HPC
	It has resulted in productivity improvements in my work
	It has resulted in performance improvements in my codes
Research work	I have applied the training in my research work
	It has helped in submission of grant applications
	I have passed on aspects of the training to colleagues
New knowledge	It has enabled me to port new codes to HPC systems
	It has enabled me to make use of new HPC packages

Note: Each question could be rated from 1 to 6, where 1 corresponds to "Strongly Disagree" and 6 corresponds to "Strongly Agree".

Source: LE analysis using ARCHER CSE Service Report on First Training Impact Survey

On average, across the last five Training Surveys for each question category, survey respondents agreed that participation in the ARCHER training improved their productivity and performance (average score above 4 each year). Moreover, the average training participant found the training moderately beneficial to its research work (approximate average score of 4). Acquiring new knowledge about HPC received an average score of 3.2.

9.3 International models

9.3.1 Finland

CSC – IT Centre for Science (CSC) is the sole, centralised HPC services national provider in Finland. CSC is a private non-profit company, with the public sector owning shares in the company. The main goal of the organisation is to support IT-related research and science in Finland. The Centre offers access to various computing solutions including HPC and cloud technologies.

Financing:

The HPC supercomputers are fully financed by the Finnish Ministry of Education, Science and Culture. The capital investment in hardware is made approximately every 5 years. At present, CSC owns two types of HPC infrastructure: a high-end HPC dedicated to large-scale parallel computation and a cluster environment for jobs requiring lower capacity. Both machines are not replaced simultaneously; the latter is replaced one year after the replacement of the high-end supercomputer.

The approximate value of the most recently purchased machines[103] was £25.4 million[104] (€29 million) (excluding VAT). The net cost of the hardware acquired five years ago reached around £15.8 million (€18 million).

[103] This figure includes both HPC systems and other IT systems available at CSC e.g. cloud computing technologies.
[104] Original value reported in EUR €. Converted into Sterling using the 2017 Bank of England average annual spot exchange rate and adjusted for inflation using the ONS Consumer Price Index.

Operating model:

The HPC maintenance costs are mostly covered by the CSC's annual budget contractually agreed with the Ministry of Education, Science and Culture. In addition, maintenance services can be purchased through a separate procurement process. Additional contributors to the annual budget are the European Commission and various National Funds.

Access & Usage:

Researchers can access CSC services free of charge. One of the advantages of the current operational model of CSC is its flexibility in time allocation. All applications are processed centrally and independently from projects' research areas. The overall machine capacity available to HPC users can be split into 3 categories, depending on the size of the project:

- Small-size projects: the decision about giving access to the system is fully automated; the researcher receives an email notification about the assigned time slot soon after sending a request;
- Medium-size projects: time allocated every three weeks; the decision made by an in-house committee consisting of post-docs and professors;
- Large-size projects: time allocated twice per year by the above-mentioned panel of researchers.

Nevertheless, the two available systems (high-end HPC and additional cluster for smaller jobs) are often overbooked due to the high demand.

Industrial access:

CSC HPC can be accessed by industrial users for a fee. Private users contribute to a small percentage of the total HPC usage (around 5%). The fees-generated income is used to finance further research or is being re-invested in the new HPC systems.

University HPC centres:

Some Finnish universities use their own, independent local HPC facilities. The university clusters are partly supported by CSC through the process of procurement supervision. This additional layer of HPC systems facilitates access for researchers and relieves the burden on the national computer.

9.3.2 The Netherlands

SURFsara is the Dutch National HPC Centre, previously known as SARA before undergoing merger with SURF. SURFSara is part of the SURF co-operative association of Dutch educational and research institutions, who are the owners of SURF. Established in 1971, as a joint computing centre of the two Amsterdam universities and the Mathematical Centre, SURFSara has expanded its range of services since. Currently, the centre offers not only access to HPC infrastructure but also runs grid-based infrastructure, cloud computing and a large archive for storage. The mission of the organisation is to facilitate public research in the Netherlands.

Financing:

Due to the adopted legal form of a co-operative, SURF is managed by its 108 members among others: universities, applied universities and knowledge institutes. Most of the funding comes from the government, namely the Ministry of Education and Research, through the research council. A small part is also funded by the Ministry of Economic Affairs. SURF members pay for services offered by SURF.

An extension to the HPC facilities is made approximately every year or two years. The capital investment in hardware in the past five years was equivalent to approximately £14 million[105] (€16 million). A total replacement of the system is undertaken approximately every 5 to 6 years.

Operating model:

The day-to-day operations are covered by the annual budget, mainly provided by the research council. The budget is running long-term running budget; subject to approval each year. SURF reports directly to the research council for the activities aimed at facilitating public research.

Access & Usage:

Access to SURFsara facilities is free on point of access for academics and researchers. In other words, each user in academia is charged a fee, however, access granted is prepaid by government. Nonetheless, in special cases when a researcher wants to expand the range of awarded services beyond the limits of the central funding coverage (e.g. increase the storage), it has to be compensated for by their research community.

Industrial access:

Access to HPC equipment is available for private users under certain constraints. It must be market-conforming and cannot be in competition with public providers. Private users contribute to around 5 to 10% of HPC usage.

9.3.3 United States

The USA offers multiple channels of access to HPC for the research community:

- **The National Science Foundation (NSF)** has set up supercomputing facilities in numerous states. The Foundation invested in HPC hubs in order to facilitate research requiring top-quality advance research computing resources and services. The NSF is fully funded by the US Federal Government and its annual budgeting must be approved by the US Congress. In particular, the NSF-funded XSEDE project plays a key role in academic HPC provision in the US.
- **The US Department of Energy (DoE)** operates six major laboratories around the US, all of which enable access to the state-of-the-art high-performance computing equipment. Three of the labs are used for classified work, mostly in the security and defence areas. DoE laboratories support the laboratories' own research, though some centres also offer access to researchers not affiliated with the laboratories. Currently, the DoE has access to the fastest supercomputer in the world, located at its Oak Ridge National Laboratory[106].

[105] Original value reported in EUR €. Converted into Sterling using the 2017 Bank of England average annual spot exchange rate and adjusted for inflation using the ONS Consumer Price Index.
[106] Top 500 HPCG List for June 2018. Available at: www.top500.org

- **Other organisations of the US Federal Government** e.g. NASA or the US Department of Defence invest in their own HPC machines as well. Similar to the DoE, their main focus is put on their own internal projects, however, some of the organisations offer access to external researchers through dedicated programmes.

The remainder of this section discussed the model of HPC provision of NSF HPC capabilities.

Financing:

The NSF funds major new machines through direct grants with institutions and offers them to the community via one of its largest projects – the XSEDE. XSEDE federates a number of major HPC centres based in different states including Texas, Illinois, California and Pennsylvania. The organisation has a mandate to support any research commissioned by federal agencies. Over the last 10 years, the XSEDE initiative has provided integrative and support services with a total worth of approximately $250 million to compliment a series of capital investments in new HPC machines.

In the upcoming years the XSEDE anticipates the NSF to fund approximately four new HPC machines with the total anticipated cost of around $30 million that will be integrated into the XSEDE distributed environment. Historically, the NSF would invest a sum between $50-$60 million in new HPC hardware and operational budgets. On top of that, the Foundation would dedicate approximately $20-$30 million to the leading-edge XSEDE project. Moreover, the NSF has recently acquired a supercomputer with significantly higher capacity. However, this investment expands the usual scope due to the associated high capital investment costs reaching $60 million for the Phase I machine.

In order to obtain funding from the NSF, aspiring HPC centres must compete for a series of individual awards announced by NSF. To be awarded a grant from the NSF, a centre must submit a proposal in response to one of NSF's public announcements. To ensure that wealthier universities are not favoured during the reviewing process, partner universities are not allowed to provide direct funding support to winning projects. As a result, in the majority of the cases, 100% of the awarded grant comes from the NSF. The organisation would allow additional discretionary investments from external stakeholders on an occasional basis.

NSF funded machines are typically funded for an operational period of four- to five-years. After this time, HPC centres have to submit a new proposal, and be awarded a new grant, in response to one of NSF's funding calls. Though, NSF-grant awards can be renewed at the end of a cycle, on the condition of meeting certain quality performance targets. The hardware investment contracts do not normally foresee any mid-cycle investment opportunities, though the possibility for mid-cycle replacements exist if foreseen in the initial proposal submitted.

Operational model:

Under the NSF funding model, operational expenses are funded as part of an additional grant beyond the hardware acquisition grant. An annual budget of up to 20% (and in some cases 25%) of the hardware acquisition cost is usually foreseen for user support and operational expenses.

Access & Usage:

The application process for access to NSF's capabilities resembles the model found in the UK. Researchers first submit a detailed research proposal, elaborating on the project's goal and number of machine hours needed to carry it out. Next, a dedicated committee reviews the proposal based on its merit and allocates the available computing time. The reviewing process is agile and ensures that the researchers are granted the requested access within a short period of time.

The XSEDE grants access to the NSF-funded HPC environments with no differentiation between researchers' scientific domains. Moreover, access to NSF HPC resources is open to research funded by any federal investment. However, due to the limited capacity of available machines, the research proposal award process does have some bias towards researchers funded by the NSF.

Industrial access

Industrial access to NSF HPC centres is not part of their scope and is not funded by the NSF. US education institutions accessing NSF HPC resources may support industry work, under the condition that the research outcomes are published.

9.3.4 Germany

Supercomputing provision on a national level follows a similar Tier-structure to the UK:

- **National HPC centres** are provided by the Gauss Centre for Supercomputing and include three hubs in Garching, Juelich and Stuttgart. The centres grant access to supercomputers at the EU Tier 0 level (part of the EU PRACE project) as well as the national Tier 1 level. The centres are available for researchers working across all scientific domains.
- **Regional HPC centres** are available to researchers at a local level, offering additional capacity for mid-sized research projects. Germany offers around 10-15 centres serving particular regions.
- **University level HPC centres** based in the universities allow students to use advanced computing methods throughout their studies.

The Juelich Supercomputing Centre (JSC), part of the Juelich Research Centre, operates a range of supercomputers. Importantly, as part of the Gauss Centre for Supercomputing, the national supercomputing centre in Germany, Juelich hosts a Tier-1 machine. In addition, the centre also operates a second HPC system, providing capacity for the region.

In addition to HPC provision JSC offers two main additional services: users support and training. Training is provided both at the centre level, as well as through the Gauss centre. Also through the Gauss centre, JSC Juelich is also one of the PRACE training centres, offering training at the European level. In terms of user support, JSC offers a multi-layered user support service, ranging from the standard helpdesk to implementing measures leading to improved performance. In addition, JSC has established "simulation labs" which are small groups, composed of researchers and computer scientists working on projects for specific communities, such as improving community code.

The remainder of this section discusses HPC provision by JSC.

Financing

The Juelich Supercomputing Centre itself is funded via institutional funding. The lion share of this funding (90%) comes from the federal government of Germany. The remaining 10% of the funding comes from the region of North Rhine-Westphalia. In addition, the centre also has third party funding.

Institutional funding is used to cover most of the centre's staff costs as well as part of their supercomputing investments. JSC's institutional funding covers around €3 to €5 million for acquiring and operating supercomputers per year.

As part of the Gauss centre, JSC also receives funding for the national supercomputers, provided 50% by the federal government and 50% by the region of North Rhine-Westphalia. Funding for the national supercomputer received is close to €150 million over the period between 2017 and 2025. Almost all of this investment is used to acquire and operate JSC's Tier 0/1 HPC machine.

The institutional funding through the Helmholtz Association is provided on the basis of 5-7 year scientific programmes, which are scientifically assessed (this so called Programme oriented funding applies to all Helmholtz centres). The funding for the Gauss centre is provided as third-party funding. Gauss' project proposal is justified also by a scientific case. Evidence of economic impact does not have to be systematically gathered to support it. Rather, the main measure of success of previous developments are the number and quality of publications using JSC's HPC systems.

Operating model

Typically, JSC replaces their supercomputer every 5 to 6 years. Intermediate system upgrades are done on an occasional basis.

Access and usage

Access to JSC's HPCs is open to researchers from all disciplines, with users representing a wide range of different disciplines.

Access to the national supercomputer is granted through a joint peer-review process with the other two Gauss centres. Typically, there are two calls per year, where researchers from academia or government labs can apply. Successful projects receive a grant for machine time for one year. In addition to that, fixed shares of the machine time are contributed to PRACE and the Earth Science community. The latter grants access to their dedicated time slots using their own peer review system.

Cycles on JSC's regional machine are pooled with the Technical University in Aachen. Both collaborating parties announce joint calls for researchers from Juelich and Aachen to apply. In addition, a separate share of machine time is reserved for researchers from the Juelich Supercomputing Centre itself.

Industrial usage

Providing private, paid access does not constitute part of JSC's core mission, though industrial users can access JSC's HPCs on a case-to-case basis. A number of smaller industrial projects are ongoing. Typically, firms are not only interested in the machine time slot, but also in additional consultancy and support services provided by JSC. The results of the study do not have to be published unless users received access via PRACE.

10 Conclusion

High Performance Computing is a fundamental pillar of modern scientific research and discovery. EPSRC's investments in High Performance Computing have underpinned world-leading research in a wide range of fields, from biomolecular simulations and medical sciences, machine learning and AI, to advances in materials research, plasma science, physics and many more disciplines. HPC has enabled researchers to answer deeper and more granular questions they would not otherwise have been able to explore and helped contribute to the overall competitiveness of UK science.

Even though the fundamental goal of HPC was enabling research, these capabilities have also delivered a wide range of further impacts as highlighted throughout this report. Doctoral and skills training utilising HPC has helped equip the next generation of UK scientists with the skills they need to succeed in academia, as well as giving students the right skills to make significant contributions in industry. HPC has enabled collaborations both with industry and academia across the globe, has contributed to the development of a strong UK software base, and delivered significant benefits to UK industry, both directly through industry usage of HPC and collaborations with academics as well as indirectly through spillovers from academic research.

The case for continued UK High Performance Computing capabilities is clear. As highlighted throughout this report, and in particular through exploration of the no HPC counterfactual (Section 8), the impact of a loss of these capabilities on UK science would be devastating; making UK science less competitive, leading to a loss of skills and crucial software capabilities, and a potential brain-drain of top talent out of the UK.

While putting a precise monetary value on these benefits is difficult, the analysis undertaken for this study indicates that, over the operational time of the services, **the aggregate economic impact of HPC is between £3.0 billion and £9.1 billion.** This represents a significant **return on investment of between 6.5:1 and 19.5:1,** when compared to the costs of HPC to the public purse. (see Section 6)

The case for continued investments in HPC is recognised by other countries, both in Europe and around the globe, as highlighted in the introduction to this study (Section 1). Without continued investment, the UK's strong position in this area is in jeopardy. The UK's national service, ARCHER, already fell to 186th place in the ranking of the world's fastest supercomputers[107], and, without continued investments, the UK will continue to lose ground in the HPC field.

However, provision of HPC hardware is only part of the puzzle. Without skilled people, the right software, training, user support, maintenance and other support activities, the best hardware cannot deliver the full potential benefits. This is rightly recognised by the EPSRC who have invested in a range of crucial activities from training provision, to software development, to doctoral training, user support, to outreach and diversity activities and more. Continued investment is therefore needed not only in state-of-the-art HPC hardware, but right across the HPC ecosystem.

The conclusions of this report echo the findings of the Tildesley report[108] and the previous HECToR impact assessment[109], highlighting both the benefits of a strong UK e-infrastructure, as well as the need for continued investment in order to maintain this strong position.

[107] Top 500 ranking November 2018. https://www.top500.org/list/2018/11/?page=1 [accessed 11/04/2019]
[108] Tildesley, D. (2012). *e-infrastructure strategy: roadmap for developing advanced computing, data and networks*
[109] EPSRC (2014). *The impact of HECToR*

INDEX OF TABLES, FIGURES AND BOXES

Index of Tables, Figures and Boxes

Tables

Figures

Boxes

ANNEXES

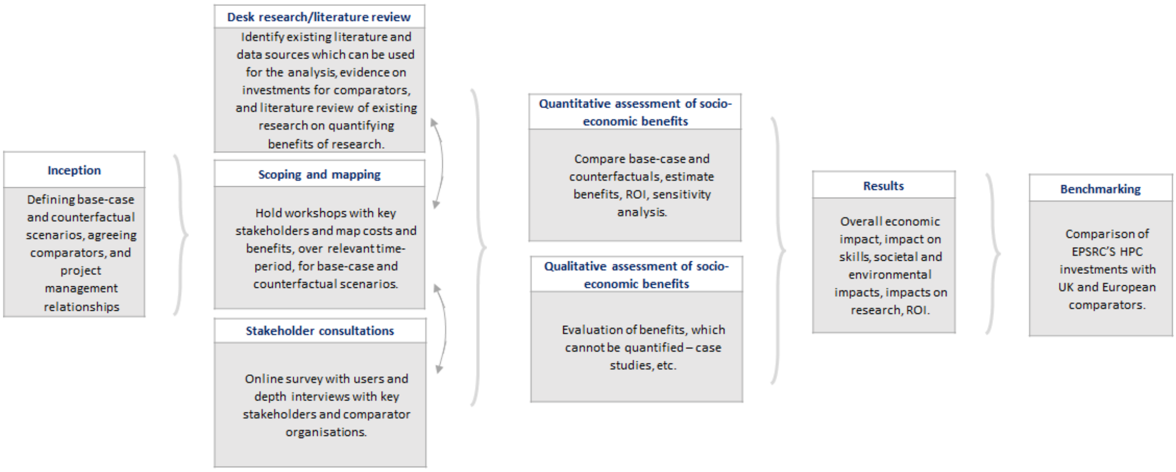
Annex 1 Approach and counterfactual

A1.1 Approach

The study approach was based on five strands of research, summarised in Figure 41 and detailed below.

First, a systematic **scoping and mapping** of the flow of socio-economic benefits of EPSRC's HPC investments and the research enabled by these investments was carried out. As part of this stage, five half-day workshops, focused on key areas of research undertaken on ARCHER and the Tier-2 centres, were held. These workshops were attended by key representatives of ARCHER's research consortia and the Tier-2 centres, as well as key users of the HPC facilities.

Figure 41 Overview of study approach



Source: London Economics

In parallel to the scoping and mapping of benefits a **review of existing literature and data** was undertaken. This stage also included a literature review on the multitude of ways in which academic research translates into economic benefits, focused on the following four key areas:

- the more general literature on the benefits of fundamental and other relevant types of research;
- literature on the benefits of particular types of research undertaken on the HPC facilities (e.g. literature on the socio-economic benefits of research in fluid dynamics);
- literature on the particular benefits of HPC supported/enabled research (compared to other types of research); and,
- literature on quantifying specific benefits identified in the workshops (e.g. literature on the benefits of improvements in skills).

In this step, desk-based research was also carried out to identify any existing literature or data sources to aid benchmarking of EPSRC's investments with other HPC investments in the UK and in Europe.

Third, **consultations** with a wide range of stakeholders were undertaken. These consultations were used to supplement, follow up on, and refine the information gathered in the workshops. The stakeholder consultations consisted of two parts:

- in-depth interviews with stakeholders from academia and industry; and,
- an online survey of all users of ARCHER (including NERC users) and the Tier-2 centres.

The fourth strand of the study synthesised the collected information and data to **assess the costs and benefits of EPSRC's HPC investments** on**:**

- **skills** (e.g. impact of HPC provision on training and skills, career choices, and destinations, etc.);
- **research and innovation** (e.g. HPC enabling research, quality of research benefiting from HPC, national / international collaborations, software and datasets generated by HPC, impact on doctoral training, generation of ideas for simulation and testing, etc.);
- **the economy** (overall economic impact, e.g. in terms of supply chain effects, effects on jobs, effects on growth (regional/local), leverage and further funding, supporting innovations, cost/efficiency, etc.);
- **the environment** (e.g. impact on climate change, climate forecasting, environmental conservation, etc. – mainly NERC funded); and,
- **society** (e.g. impact on societal challenges (housing, transport, etc.), policy impacts, impacts on science communication and outreach, etc.).

This assessment included a quantitative assessment of EPSRC's impact as well as a qualitative assessment of benefits. Case studies of specific examples of research undertaken on the HPC capabilities were also developed in this strand.

Finally, a **benchmarking** exercise of EPSRC's HPC investments was undertaken. This strand sought to compare EPSRC's HPC investments to UK and EU comparators and gain an understanding of the strengths and weaknesses and investment competitiveness of EPSRC's investments relative to comparators.

A1.2 Baseline and counterfactual

The EPSRC wishes to understand the impact of its national (ARCHER) and regional HPC investments over the last 10 years, as well as any continuing benefits from ARCHER's predecessor, HECToR. These investments form the **baseline scenario.** The investments under the baseline scenario are set out in further detail in Section 1.3.

Benefits derived under the baseline scenario constitute the gross economic impact of EPSRC's HPC investments. To arrive at the net economic impact of EPSRC's HPC investments, the socio-economic benefits and costs of the baseline scenario are assessed relative to the counterfactual.

There are a range of potential alternatives that could be used as a counterfactual, however the Treasury's 'Green Book' guidance suggests that a 'do nothing' option or a 'do minimum' option is used as a basis for judging other options.

The **counterfactual** scenario adopted in this study is a **'do-nothing'** scenario. Under this scenario, the EPSRC does not invest in HPC at all:

- EPSRC investments in HPC under the baseline scenario would not have been made; and,
- the EPSRC would not have made other HPC related investments which may substitute for its HPC capabilities (e.g. investments in cloud computing, etc.).

Under this scenario there would therefore be no EPSRC supported Tier-1 and Tier-2 capabilities, and no other EPSRC supported HPC related capabilities.

Even though this 'no HPC' scenario may not be realistic in practice (the implications of this counterfactual are explored in more detail in the following sections), it allows valuation of all benefits of EPSRC's HPC investments. The choice of the second option, where the EPSRC provides a minimal set of core HPC services, would mean that that minimal set of core services would not have a value associated with them – only the services that are additional to that core. Moreover, given the high capital and operational costs of HPC, even providing such a minimal set of core services would likely be associated with relatively high costs.

To understand the implications of the chosen counterfactual scenario, one needs to consider to what extent HPC capabilities may have been provided by other actors such as private companies (Section A1.2.1) and to what extent users of EPSRC's HPCs may have been able to access these or other capabilities in the absence of EPSRC's HPC capabilities (Section A1.2.2). The impact of the no HPC scenario on UK science is explored in Section 8.

As the analyses in these sections highlights, private actors would have been unlikely to substitute for public HPC capabilities, while other existing public capabilities would have been unlikely to cope with increased demand in the medium term. As such the chosen counterfactual is the practical baseline against which benefits of HPC investments should be compared against.

Please note that the assessment in these sections, and the above conclusion, is backward looking. In particular, comparison with other technologies are made with the state of these technologies five to ten years ago in mind. Technology is constantly evolving and as such the assessments made may not apply today or in the future. Comparing HPC with other technologies at present and weighing advantages and disadvantages is outside the scope of this study.

A1.2.1 To what extent would private investment have substituted for EPSRC's HPC capabilities?

The work undertaken on EPSRC supported HPCs is scientific in nature with the vast majority of users being academic researchers. While academic research benefits society in general, benefits can often take a long time to materialise. Moreover, it is not clear at the outset to whom benefits will accrue or what the size of these benefits may be. This is particularly true for fundamental research, that may have no immediate commercial applications.

At the same time HPC provision requires large upfront capital investments as well as continuing operational expenses and capital investments to keep machines up to date. Private companies are unlikely to have made these investments in the absence of EPSRC HPC capabilities given the uncertain returns.

This is highlighted by the current HPC landscape in Europe, which is primarily driven by the public sector: Over 90% of Europe's HPC capacity is allocated to universities or academic research centres, with only 10% installed for commercial use. Moreover, most HPC centres in Europe are publicly funded by national budgets, university funds or grants, or via support from European Union funds. While private investors are already engaged in funding HPC infrastructure, this funding is concentrated in commercial HPC infrastructure. Private funding for public HPC infrastructure is limited due to the limited 'bankability' prospects.[110]

Universities may also have decided to invest in their own HPC facilities under this scenario. However, due to the large investments needed, many universities would likely have decided not to invest in their own facility. Moreover, if funding for university level facilities would have come, even partly, from public sources such as research council funding, the cost to the public purse would have remained the same.

Similarly, other research councils may have decided to invest in HPC. For example, NERC funded part of ARCHER and has shared HPC facilities with the Met Office. In the absence of EPSRC investment, NERC may have considered to invest in their own machine or collaborate with another research council or the Met Office. However, this would again have incurred costs to the public purse.

A1.2.2 To what extent would users of EPSRC's HPC capabilities have been able to use other HPC capabilities?

To understand whether users of EPSRC's HPC capabilities would have been able to substitute to other HPCs, one needs to consider which options were available under this scenario. Consultations with users of EPSRC's capabilities indicate that the following options would have been available under this scenario:

- Users could have used university level (Tier-3) or 'homebrew'[111] capabilities
- Users could have used other public HPC facilities in the UK (e.g. DiRAC)
- Users could have used HPC facilities overseas
- Users could have accessed commercial facilities, including cloud solutions

In the short term, these alternatives would likely have allowed users to continue their HPC work. However, over time, without government investment in new HPC facilities, the loss of EPSRC's HPC capabilities would likely have meant that other public or university level facilities would have become oversubscribed. Moreover, with increased demand these facilities would likely have had to introduce stricter access criteria.

"You would end up with a market of universities with their own HPC and then second-class universities who couldn't do the science." - workshop participant

As mentioned in Section A1.2.1, as a result of increased demand and a lack of sufficient national HPC facilities, some universities may have decided to invest, or further invest, in their own HPC facilities. This would have further mitigated the impact on some users. However, it should be noted that these local facilities would likely have only provided access to users associated with the university (or users who have gained access via a collaboration, etc.). This would have put researchers at other universities, which would not have been able to or who would have decided not to invest in their own HPC, at a disadvantage.

[110] European Commission (2018). *Financing the future of supercomputing: How to increase investment in high performance computing in Europe*

[111] Homebrew capabilities in this context refers to HPC capabilities set up by a particular researcher or research group for internal use by themselves or their group.

Similarly, HPC users who already had, or would have been able to set-up, their own homebrew facilities would also have been able to continue to use these. Again, this would have put users who would not have been able to invest in their facilities at a disadvantage.

Another point to note is that university level or homebrew facilities would not have provided the same computing power and range of architecture that the current Tier-1 and Tier-2 infrastructure provides. This means that, while users may have been able to access HPC to some degree, the type of problems that could have been addressed would have been more limited.

Accessing HPC facilities overseas may also have been a short-term alternative. However, gaining access to these facilities is more difficult. As a result, only a select number of users would have been able to substitute to overseas facilities. Moreover, if more and more users applied for access to these facilities, gaining access would have become even more difficult.

"[Accessing HPC facilities abroad] is not easy because it's highly competitive. You apply 10 times and get access once. It will become even more competitive. It will be almost impossible." - workshop participant

Finally, users could have accessed commercial facilities or bought computing time in the cloud. However, this would again only have been an option for a certain user group – i.e. those with the budget to pay for these services. Researchers who would have been unable to pay for these services would again have been put at a disadvantage. Moreover, Tier-1 and Tier-2 centres also offer additional benefits that cloud technologies may not have offered. These include faster network bandwidths, centralised data storage, and dedicated support to help users get up and running and ensure users get the most out of the machine.[112]

[112] Please note that this assessment is backward looking and compares HPC to cloud services five to ten years ago. Technology is constantly evolving and as such this assessment may not apply to cloud services today. Comparing HPC with cloud at present and weighing advantages and disadvantages is outside the scope of this study.

Annex 2 Estimation of benefits and costs

Benefits were estimated across a number of strands, detailed below. This section outlines the methodology used to calculate benefits, as well as the main assumptions used. Caveats and limitations of the methodology are discussed where appropriate.

Benefits were estimated in two ways:

- The low estimate represents the lower bound estimate using conservative assumptions to derive benefits.
- The high estimate represents the upper bound estimate using less conservative assumptions to derive benefits.

In reality, the impact of HPC will lie somewhere between the low and high estimates presented in this study. However, due to the significant challenges and uncertainty inherent to assessments of scientific R&D investments, a point estimate is not provided.

A2.1 Impact of UK scientific research and discovery

Avoided cost of free HPC access for academics

Direct benefits of academic usage represent the benefits of access to EPSRC funded HPCs by academics. The calculation uses an avoided cost approach. That is, the direct benefit of access to academics is calculated as the cost avoided if academics had to access commercial HPCs. Where usage data was not available, it was approximated based on the operation time of the centres, the number of cores and an assumed utilisation of the theoretically available core hours.

The main drivers of this strand are the price of access that commercial HPC centres charge. The calculation uses the access costs in the table below. These costs were calculated as the average of the highest three prices charged by commercial providers listed in Annex A3.1.

Assumption	
Price per core hour	£ 0.09
Price per GPU hour	£ 1.66

The reason for using the upper range of the price estimates is multi-fold:

- The above prices are for cloud services today. However, the assessment is backward looking. Prices for cloud services would likely have been much higher five to ten years ago than the high-end range (£0.09 and £1.66) for avoided cost estimates based on today's prices.
- The cloud prices discussed above are baseline prices for hourly use only. Additional costs for storage, software or other services may have applied. The avoided cost of these services is not estimated separately.
- Tier-1 and Tier-2 centres also offer additional benefits that cloud technologies may not have offered. These include faster network bandwidths, centralised data storage, and, perhaps most importantly, dedicated support to help users get up and running and ensure users get the most out of the machine.

For these reasons using cost estimates at the high end of today's prices will still yield an underestimate of the actual avoided cost[113].

Please note that the above assessment is backward looking and compares HPC to cloud services five to ten years ago. Technology is constantly evolving and as such the above assessment may not apply to cloud services today. Comparing HPC with cloud at present and weighing advantages and disadvantages is outside the scope of this study.

It should be noted that under the counterfactual, EPSRC would only have to pay access charges for research that could have been undertaken in the absence of HPC. Therefore, the total estimated costs are adjusted for the proportion of research that would not have been undertaken in the absence of EPSRC HPC, obtained from the survey of HPC users:

Proportion of research that could not be undertaken without HPC (median)	70%
Adjustment used in calculation (1-HPC dependent research)	30%

Note that these estimates are based on a survey of HPC users. Therefore, all the usual caveats of online surveys apply to these estimates. These caveats are discussed in Section 1.5, which also examines the representativeness of the survey.

HPC centres included in estimation: all

A2.2 Spillover impact of HPC research on UK output (EPSRC only)

Academic research generates spillover impacts through a number of channels. To calculate productivity spillovers from EPSRC funded research, data on EPSRC's investments in HPC research, obtained from Researchfish™, are combined with literature estimates of productivity spillovers.

The main drivers of this strand are the spillover estimates. Spillover estimates are based on multipliers from the literature. Specifically, the analysis uses productivity spillover estimates from Haskell and Wallis's 2010[114], which analyses productivity spillovers to the private sector from public spending on R&D by the UK research councils, as well as productivity estimates from Haskell et. al.'s 2014 paper[115], which provides further evidence on the size of potential productivity spillovers from public sector R&D. The main results of these studies are summarised below, for a methodological summary of see Annex 4.

Haskel and Wallis[116] investigates evidence of spillovers from public funding of Research & Development. The authors analyse productivity spillovers to the private sector from public spending on R&D by the UK Research Councils[117]. Haskel and Wallis find strong evidence of the existence of

[113] Note that, in addition, more research in the past would likely not have been able to be undertaken without HPC as cloud alternatives were not available. While this would yield a lower direct benefit of access than the one presented here, no spillover benefits of this research would have taken place. As such the estimate of benefits of research (Section 4.1.3) would be lower in this scenario. As spillover benefits of research can reasonably be expected to be larger than avoided costs of access, the overall aggregate economic benefit would thus also be lower in this scenario.
[114] Haskel, J., & Wallis, G. (2010). *Public support for innovation, intangible investment and productivity growth in the UK market sector*.
[115] Haskel, J., Hughes, A., & Bascavusoglu-Moreau, E. (2014). *The economic significance of the UK science base: a report for the Campaign for Science and Engineering.*
[116] Haskel, J., & Wallis, G. (2010). *Public support for innovation, intangible investment and productivity growth in the UK market sector*.
[117] The authors use data on government expenditure published by the Department for Business, Innovation and Skills for the financial years between 1986-87 and 2005-06.

market sector productivity spillovers from public R&D expenditure originating from UK Research Councils[118, 119]: the marginal spillover effect of public spending on research through the Research Councils stands at 12.7 (i.e. for every £1 spent on university research through the Research Councils results in an additional output of £12.70 in UK companies). The analysis also suggests that the spillover benefits of public spending on research in higher education are greater than those from other R&D areas supported by government.

A more recent study by Haskel et al.[120] provides additional insight into the size of potential productivity spillovers from university research. Rather than estimating effects on the UK economy as a whole, the authors analyse the size of spillover effects from public research across different UK industries[121]. The authors investigate the correlation between the combined research conducted by the Research Councils, the higher education sector, and central government (e.g. through public research laboratories)[122], interacted with measures of industry research activity, and total factor productivity within the different market sectors[123]. Their findings imply a total rate of return on public sector research of 0.2 (i.e. every £1 spent on public R&D results in an additional output of £0.20 within the UK private sector).

Note that further quantitative evidence on the return to public investment in scientific research is limited. An overview of the limited number of other studies that look at returns to public R&D investments can be found in a 2014 Frontier Economics report on returns to investment in science and innovation[124] as well as in in 2015 Economic Insights paper on the relationship between public and private investment in science, research and innovation[125]. In addition, the above-mentioned papers are, to the best knowledge of the authors, the only papers that specifically look at returns to science funding via the research councils.

It is important to note that both studies examine productivity spillovers of research across a wide range of research areas and are not specific to HPC. HPC research is likely to generate additional productivity spillovers not captured by these estimates. However, research specifically on

[118] Based on regression of total factor productivity growth in the UK on various measures of public sector R&D spending.

[119] Note that the authors' regressions only test for correlation, so that their results could be subject to the problem of reverse causation (i.e. it might be the case that increased market sector productivity induced the government to raise public sector spending on R&D). To address this issue, the authors not only test for 1-year lags, but for lags of 2 and 3 years respectively, and obtain similar estimates. The time lags imply that if there were a reverse causation issue, it would have to be the government's *anticipation* of increased total factor productivity growth in 2 or 3 years which would induce the government to raise its spending on research; as this seems an unlikely relationship, Haskel and Walls argue that their results appear robust in relation to reverse causation.

[120] Haskel, J., Hughes, A., & Bascavusoglu-Moreau, E. (2014). *The economic significance of the UK science base: a report for the Campaign for Science and Engineering.*

[121] Haskel et al. (2014) use data on 7 industries in the United Kingdom for the years 1995 to 2007.

[122] A key difference to the multiplier estimate for Research Council spending in Haskel and Wallis (2010) lies in the distinction between *performed* and *funded* research, as outlined by Haskel et al. (2014). In particular, whereas Haskel and Wallis estimated the impact of research *funding* by the Research Councils on private sector productivity, Haskel et al. instead focus on the *performance* of R&D. Hence, they use measures of the research undertaken by the Research Councils and the government, rather than the research funding which they provide for external research, e.g. by higher education institutions. The distinction is less relevant in the higher education sector: to measure the research performed in higher education, the authors use Higher Education Funding Council funding (where research is both funded by and performed in higher education).

[123] The authors regress the three-year natural log difference of total factor productivity on the three-year and six-year lagged ratio of total research performed by the Research Councils, government and the Higher Education Funding Councils over real gross output per industry. To arrive at the relevant multiplier, this ratio is then interacted with a measure of co-operation of private sector firms with universities and public research institutes, capturing the fraction of firms in each industry co-operating with government or universities. The lagged independent variables are adjusted to ensure that the resulting coefficients can be interpreted as annual elasticities and rates of return.

[124] Frontier Economics (2014). *Rates of return to investment in science and innovation*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/333006/bis-14-990-rates-of-return-to-investment-in-science-and-innovation-revised-final-report.pdf [accessed 01/08/2019]

[125] Economic Insight (2015). *What is the relationship between public and private investment in science, research and innovation?* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/438763/bis-15-340-relationship-between-public-and-private-investment-in-R-D.pdf [accessed 01/08/2019]

productivity spillovers of HPC research is even rarer, though a recent study by the IDC[126] suggests that the return from investments into academic HPC projects may be significantly higher: among the European HPC projects analysed, every $ invested in academic HPC projects generates a return on investment of $30. Though it should be noted that the study is based on a small sample of only seven academic HPC projects. Moreover, the study only includes success stories, not projects that didn't generate economic or scientific results. This likely biases the results, resulting in inflated HPC estimates.

Given the limited evidence on productivity spillovers by research councils, and the observations discussed above, the analysis in this study uses the two papers by Haskel et al. for the low estimate, and the IDC paper for the high estimate:

Research multipliers (low estimate):

Research Council Funding	12.7
Other R&D performed by universities	0.2

Research multipliers (high estimate):

IDC ROI academic HPC projects	30

In the absence of Tier-1 and Tier-2 centres, researchers using HPC centres would still undertake research. To capture the benefit of research that depends on HPC, benefits are adjusted using the average proportion of research depending on HPC obtained from the LE survey of HPC users:

Proportion of research that could not be undertaken without HPC	70%

Note that, as this estimate is based on a survey of HPC users, all the usual caveats of online surveys apply to these estimates. These caveats are discussed in Section 1.5, which also examines the representativeness of the survey.

HPC centres included in estimation:	All, but only impacts of EPSRC grant receivers are captured. Investments from other research councils are only captured if this investment was reported back to the EPSRC.

A2.3 Impact of direct industry access

A2.3.1 Contribution of industry impacts to UK output

Industry receives benefits in terms of reduced costs of access compared to commercial services, as well as potential long term benefits in terms of cost reductions as a result of HPC usage, efficiency gains, improvements to existing products or services, or by contributing to the introduction of new products or services, or enhanced business growth. Benefits to industry are calculated based on the contribution that HPC makes to firms using the Tier-1 or Tier-2 centres, in terms of increased profits. The contribution of this increased profit to UK Gross Value Added is then calculated using the average ratio of Profit to GVA in the UK business economy based on Eurostat SBS data.

[126] IDC (2014). *EESI-2 Special Study To Measure And Model How Investments In HPC Can Create Financial ROI And Scientific Innovation In Europe*

To calculate benefits to industry, a distribution of firms, by size, accessing Tier-1 / Tier-2 centres has to be assumed. For the calculation it is assumed that firms accessing HECToR and Tier-2 centres follow a similar distribution as those firms accessing ARCHER:

Proportion of firms on ARCHER:

SME	55%
Midcap	20%
Large	25%

Note that this assumption may not hold in practice. In particular, the distribution of firms accessing Tier-2 centres may vary from those accessing Tier-1 centres. Moreover, the distribution may also have varied over time. Unfortunately, more granular data could not be obtained during the course of this study. However, given that ARCHER and its predecessor, HECToR, both were national machines with the same target audience and many similar features it may perhaps reasonably be assumed that the distribution of firms accessing ARCHER would be similar to those which accessed HECToR.

The distribution of firms accessing Tier-2 centres may be subject to greater uncertainty. In particular, given the less stringent access restrictions and initiatives such as the Fortissimo project (discussed in Section 5) it may be expected that Tier-2 centres have a higher proportion of SMEs using their capabilities. Given that benefits of industry access of new Tier-2 centres only account for a relatively small proportion of overall industry benefits (see Table 10 in Section 5), accounting for these considerations would likely only have a marginal impact on the estimated benefit overall.

It should also be noted that some industry users may be accessing multiple facilities. This means that the overall number of firms accessing HPC may be lower than the number used in the analysis.[127] On the other hand, no data on industry access for the five regional Tier-2 centres was available. As such, benefits of industry access for these centres are excluded from the analysis, meaning that the overall number of firms used in the analysis may actually be an underestimate. Moreover, given that regional Tier-2 centres were operating for more than five years, whereas new Tier-2 centres have only been operating for a bit over a year, suggests that excluded benefits of industry users accessing Regional Tier-2 centres may outweigh benefits of firms accessing multiple services.

The distribution of firms is then combined with annual profit data for SMEs (using turnover data for SMEs from Eurostat SBS converted into Profits using the average profit-turnover ratio in the UK business economy), Midcaps (using data for FTSE250 companies from TopTrack250 as a proxy), and Large (using data for FTSE100 companies from TopTrack100 as a proxy) firms. To calculate the annual profit increase for firms, the low estimate in the table below is used in the calculation of the low estimate, while the high estimate is used for the calculation of the High estimate:

Assumption:	**Low**	**High**
Effective increase in profit	0.55%	1.75%

These estimates are based on the following observations:

[127] Note that the survey only received limited responses from industry. As such, unlike for academic users, no reliable measure of the proportion of industry users accessing multiple facilities could be derived.

- The only respondent providing a range for benefits in the online survey suggested an increase in profit of around 25% over the last five years (or a CAGR of around 4.6% per year). In addition, one SME working with the EPCC as part of the Fortissimo project reported an annual increase in profits of around 3% (Table 8). However, as mentioned previously, not all firms accessing HPCs see an increase in their sales. Combining these estimates with the 38% of respondents who said they had received increases in sales / turnover or profit, suggests a total effective increase in profits of between **1.1%** to **1.7%** per annum[128].
- In the absence of ARCHER / Tier-2 HPC, large companies may invest in their own HPC machines. While the price of HPC can vary significantly depending on the configuration, a cheap supercomputer can be acquired for between £0.3 million to £2.1 million[129]. However, even the more expensive machine is significantly less powerful than ARCHER / Tier-2. Assuming a five-year replacement cycle and allowing for operational costs of up to 20% of the purchase value[130], would mean total annual costs in the region of £0.8 million (based on £2.1 million machine), or approximately **0.4%** of average profit of FTSE100 enterprises[131]. In contrast, accessing publicly funded HPCs can be done at negligible costs[132].
- For Midcap firms it is unlikely that they would invest in the more expensive machine, given that this would mean costs of approximately 4% of profit. Assuming mid-sized firms instead opt for the lower cost machine (£0.3 million), would imply costs of approximately **0.7%** of profit per year.[133]

Given these observations, the analysis used an effective increase in profits rate of **0.6%** (average of cost reduction of mid-sized and large firms, assuming no increase in profit beyond cost savings) for the low estimate, and of **1.7%** (based on 4.6% profit increase per year accruing to 38% of companies) for the high estimate.

Note that small firms do not usually have the means or expertise to acquire and efficiently utilise their own supercomputers. Therefore, the above lower bound estimate does not directly apply to them. However, in the absence of publicly funded HPC, it is likely that a large proportion of HPC research at smaller firms would not have taken place. Moreover, diminishing returns to R&D[134] suggest that SMEs that do invest in R&D may see higher marginal returns than larger counterparts already investing large amounts into R&D. Therefore, the same lower bound was used for small firms.

Finally, to derive the impact on UK GVA, the average ratio of profit to GVA in the UK business economy (excluding financial and insurance activities) between 2008 and 2016 (the latest available year) was applied to calculated profit increases of firms:

[128] 3% X 38% = 1.14%; 4.6% X 38% = 1.7%

[129] In 2018 prices. Original prices $0.5 million to $3 million (in 2013 prices), based on: https://www.eweek.com/servers/cray-offers-low-cost-xc30-ac-supercomputer-for-500-000. Converted into £ using the 2013 average Bank of England annual spot exchange rate. Deflated using ONS Gross Fixed Capital Formation deflators.

[130] Based on average replacement time of EPSRC and comparators, and operational cost allowance of NSF.

[131] Based on average proportion in profit of FTSE100 companies, obtained from Top Track 100, between 2008 and 2018.

[132] E.g. Cirrus charges £0.037 per core hour for industry access; see https://www.epcc.ed.ac.uk/facilities/demand-computing/cirrus [accessed 26/07/2019]

[133] Again based on annual costs assuming a five year replacement cycle and operational costs of 20% of purchase value, as a proportion of average profit of FTSE250 companies, obtained from Top Track 250, between 2008 and 2018.

[134] See e.g. Kim, Konshik. (2018*). Diminishing returns to R&D investment on innovation in manufacturing SMEs: do the technological intensity of industry matter?* International Journal of Innovation Management. 1850056. 10.1142/S1363919618500561.

Assumption:

Ratio profit to GVA	0.49

For this exercise profit and GVA data were obtained from the Eurostat Structural Business Statistics.

HPC centres included in estimation: ARCHER, HECToR and new Tier-2 centres

A2.4 Impact of training and skills development

A2.4.1 Benefits of PhD and postdoc training of students entering industry to students and the UK exchequer

Benefits of PhD and postdoctoral students with HPC training are calculated as the lifetime earnings premium HPC graduates obtain compared to a counterfactual group. The counterfactual used in the analysis are students with any postgraduate degree aged between 25 and 34 (i.e. recent graduates) in the same professions. Benefits are split into two parts: benefits accruing to graduates themselves (the HPC earnings premium) and benefits accruing to the UK exchequer, calculated as the additional income tax, national insurance contributions, and VAT accruing to the exchequer.

The number of PhD and postdoc students are estimated based on the following data for Tier-1:

- The ARCHER consortia reports indicate that at least 170 PhD or postdoc students were trained on archer in 2016/17 alone.[135]
- The Story of HECToR[136] reports that at least 130 PhD students were trained on HECToR.

Data for the number of PhD and postdoc students accessing the Tier-2 centres is not recorded across all centres; however:

- Of those centres that were able to provide data, two centres reported PhD and postdoc access in excess of 200 users.
- The Tier-2 ReICN report[137] indicates that, on average, 404.5 graduate and post doctorate users used each of the regional Tier-2 centres over their lifetime

Therefore, a low estimate (using 200 PhD/postdoc users per centre on average) and a high estimate (using 404.5 PhD/postdoc users per centre on average) is used in the analysis.

Note that there is some degree of uncertainty surrounding these figures. In particular, figures for ARCHER were based on only on PhD / postdoc students trained by consortia in 2016/17. This likely underestimates the overall number of students trained. Similarly, the HECToR report reported “at least” 130 students being trained, suggesting that this number is also a conservative estimate.

Data on students trained at Tier-2 centres was not recorded across all centres. Through, in addition to the data used in the estimation (outlined above), two centres suggested that the majority of their academic users would have been PhD / postdoc students. Therefore, it appears reasonably that the

[135] Based on Annual Reports 2016/17 of UKCOMES, UKCTRF, HEC BioSim, UKPP, UKMC; data for UKCP, UKTC and UK AMOR was not available. The actual number of PhD students and postdocs trained on ARCHER is therefore likely even higher.
[136] EPSRC (2014). *The impact of HECToR*
[137] The Importance of Regional e--Infrastructure Within the National Landscape: A submission compiled by the Regional e--Infrastructure Centres Network

200 users per centre would be an underestimate of the actual number of PhDs and postdocs trained at Tier-2 facilities.

The upper bound estimate is based on a report for regional Tier-2 centres. Since the overall number of users accessing HPC has grown significantly over the last decade, it could reasonably be expected that the number of PhDs and postdocs users accessing Tier-2 centres would also have grown. On the other hand, the number includes some graduate students in addition to PhDs and postdocs.

To account for PhDs and postdocs accessing multiple facilities, the number of users is adjusted by the proportion of HPC users indicating that they use multiple facilities, obtained from the LE user survey:

Proportion of users accessing multiple facilities	68.9%

Note that, as this estimate is based on a survey of HPC users, all the usual caveats of online surveys apply to these estimates. These caveats are discussed in Section 1.5, which also examines the representativeness of the survey.

Moreover, benefits were calculated for students entering 'Technical fields' in industry only[138]. Students staying in academia, entering the public sector, moving abroad, or those with other destinations are excluded from the analysis. To do this, data on the number of PhD and postdoc students trained were combined with survey data on graduate destinations (see Section 4.3.2):

Graduate destinations:

Proportion staying in the UK	67.5%
Proportion entering private sector	37.1%
Proportion entering technical fields	88.2%
Proportion of students considered for analysis:	22.1%

The key drivers of the analysis are the earnings of HPC graduates (obtained via the LE survey of HPC users) and the earnings of the counterfactual group (obtained from the 2017 Labour Force Survey):

Salary data:

Average salary of HPC graduates	£ 65,625
Average salary of other postgrads	£ 42,100

It should be noted that salary data for HPC graduates was based on a small sample (nine data points for students entering technical fields) of salary data reported by academics who have trained PhD students. Therefore, the estimate used in this analysis is dependent on academics' knowledge of their former students' career development. Moreover, as the estimate is based on a survey, all the usual caveats of online surveys (discussed in Section 1.5) also apply to these estimates.

[138] The analysis was split into those entering 'technical fields', where their HPC skills could reasonably be expected to be of use, and those entering 'Education, and human health activities'. This included Professional, scientific and technical activities; Financial and insurance activities; Information and communication; Manufacturing; Public administration and defence; compulsory social security; Transportation and storage; Electricity, gas, steam and air conditioning supply; and Water supply; sewerage, waste management and remediation activities. The graduate premium for students entering education; human health and social work activities; and administrative and support service activities was found to be negligible or non-existent. Only a very small proportion of students entered the agriculture, forestry and fishing; other service activities; and arts, entertainment and recreation sectors, as such these sectors were not included in the analysis.

A degree of uncertainty is also present in the counterfactual, which is based on workers with any postgraduate degree between 25 and 34 in the same professions. Therefore, the counterfactual salary estimate may be biased downwards as it did not account for degree subject (i.e. by only including STEM fields in the counterfactual we may expect the average salary to be higher). On the other hand, the salary estimate may be biased upward as it did not only include recent graduates, but also individuals who may have been in the job for several years. Due to the relatively small sample size (66 observations for technical fields), these effects could not be explored further.

To derive lifetime benefit estimates, salary data is combined with assumptions on the average age of completion of a PhD degree and the average retirement age. Future benefits are discounted using the HM Treasury Green Book discount rate:

Assumptions:

Average age on completion of PhD	27
Average retirement age	65
Years in labour force	38
Discount rate	3.5%

To derive benefits to students and the exchequer, the 2019 income tax rates and national insurance contributions, were applied to the gross HPC premium. Note that in practice tax and national insurance rates, rates are likely to vary over the lifetime of graduates. To derive estimates of additional VAT receipts accruing to the exchequer a household savings rate of 6% based on OBR economic and fiscal outlook for March 2018[139] was assumed. In addition, it was assumed that 50% of consumed expenditure is subject to VAT[140]. Finally, for simplicity, it was assumed that the same 20% VAT rate applies across all goods and services consumed by the individuals.

Salary data:

Household savings rate	6%
% of consumed expenditure subject to VAT	50%
VAT rate	20%

HPC centres included in estimation: all

A2.4.2 Benefits of provision of free HPC training courses

Benefits of training are estimated using an avoided cost approach. That is, the benefit per attendee, per training day is calculated as the cost of attending a similar training from a commercial provider.

The key driver of the analysis is the cost of accessing training from a commercial driver, this is assumed to be £250 per attendee per training day, following the 2018 Technopolis evaluation of the Hartree Centre[141]:

Total number of attendee days of trainings	8,676
Cost of training per attendee per day	£ 250

[139] OBR (2018). *Economic and fiscal outlook - March 2018.* http://obr.uk/efo/economic-fiscal-outlook-march-2018/
[140] Based on OBR: *Tax by tax, spend by spend: VAT.* http://obr.uk/forecasts-in-depth/tax-by-tax-spend-by-spend/vat/
[141] Technopolis (2018). *Hartree Centre Phase 1&2 Baseline Evaluation Final Report.*

Total benefit of training	£	2.2 m

Our own comparison of commercial software engineering training courses (see Annex A3.2) suggests that this is a reasonable estimate, though prices of training vary considerably across providers.

It should also be noted that training data was only available for five of the eleven Tier-2 centres. Training data was also not available for HECToR. Therefore, the overall benefits of trainings are likely underestimated.

HPC centres included in estimation:	ARCHER and Tier-2 centres (where data was available)

A2.5 Benefits of software

Benefits of software are based on the evaluation of benefits of eCSE software project by the EPCC. That is, the benefits to software are evaluated as the benefit of additional science enabled by improvements to software. The potential benefits of additional, non-eCSE software projects, was calculated assuming that these projects bring similar benefits to those of the eCSE projects:

Assumption:

eCSE benefits	£ 24.5 m
No. of eCSE projects assessed	91
Benefits per eCSE project	£ 0.27 m
No. of software products reported back	107
Benefit of software products reported back	£ 28.9 m
Total benefit	£ 53.4 m

Note that these estimates only include benefits in terms of freeing up resources for additional science. However, this number significantly underestimates the true value of improved HPC software to the UK economy for several reasons:

- Only software projects reported in EPSRC's research outcome system are counted. However, with more than 40% of respondents to the user survey suggesting that they have developed or optimised code that benefitted the wider HPC community, the number of software projects not captured could be very large.
- The benefits only capture the benefit to science in terms of being able to do more science on existing HPC hardware. The true benefits of improved or new software is the contribution that this software makes to scientific research & discovery. Note that these benefits are implicitly monetised when estimating the spillover benefits of HPC research (Section 4.1.3), therefore software is excluded in the calculation of total aggregate benefits.
- HPC software may also bring significant benefits to industry. This is highlighted by CASTEP (see Box 9), an HPC simulation software used to calculate the properties of materials from

first principle, which is used by around 900 industrial users. As the Goldbeck report[142] indicates, companies using materials modelling software such as CASTEP, among others, achieved cost savings ranging from €100,000 to €50 million, with an average return on investment of 8:1. Benefits of software to industry were not quantified in this study.

HPC centres included in estimation: ARCHER

A2.6 Estimation of costs

Estimated benefits are compared to costs of investments in HPC. Where costs spanned several years, it was assumed that these costs accrued equally in each period of the analysis. Costs were then converted into 2018 money terms using ONS Gross Fixed Capital Formation deflators.

The following costs were considered in the analysis:

- Total CAPEX and OPEX costs in ARCHER and HECToR (across research councils and UK Government)
- EPSRC investments in Tier-2 centres (EPSRC only)
- EPSRC investments in HPC research (including consortia grants, EPSRC grant holders only)

It should be noted that the costs to the public purse used in the calculation exclude CAPEX and OPEX investments made by the Tier-2 centres and partner institutions themselves. The assumption is that these costs do not accrue to the public purse. In reality, some of these costs may ultimately be covered by the public purse. Therefore, the costs to the public purse may be higher than the costs reported here.

It should further be noted that estimates of research funding are based on EPSRC's research outcomes systems, which records quantitative information only for research activities funded directly by the EPSRC, and further funding given to EPSRC grant holders by other bodies, if these are reported to the EPSRC. Research activities using HPC systems funded by other UK Research Councils, other public bodies, or private organisations were therefore not included in the estimates *unless* the funding was fed back into EPSRC's research outcomes systems. In terms of cost calculations, only costs of further funding accruing to the public purse were counted in the calculation. This includes funding provided by UK research councils and other UK public bodies.

Annex 3 Additional material

A3.1 Costs of accessing commercial HPC

Table 16 CPU pricing for industrial usage

Provider	Cost per hour	Cores	Cost per core hour
Amazon (AWS)	£0.15 – £3.10	2 - 72	£0.04
Microsoft Azure	£0.81 - £1.61	8 - 16	£0.10
Google Cloud	£0.07 - £3.28	2 - 96	£0.04
Penguin Computing	NA	12 – 40	£0.06 - £0.08
Sabalcore	NA	NA	£0.07 – £0.09

[142] Goldbeck and Court (2016). *The Economic Impact of Materials Modelling.*

R Systems	NA	128	£0.04 - £0.07

Note: Figures reported are baseline prices for hourly use, additional support, storage, software or other costs may apply. Prices rounded to 2 decimal places. Prices are converted, where necessary, to Sterling using the Bank of England average exchange rate in 2018.

Source: LE analysis of industry pricing; https://aws.amazon.com/emr/pricing/; https://azure.microsoft.com/en-gb/pricing/details/cloud-services/; https://cloud.google.com/compute/pricing; https://www.penguincomputing.com/wp-content/uploads/2019/03/penguin-computing-hpc-cloud-pod-pricing.pdf; https://www.sabalcore.com/services/pricing/; http://rsystemsinc.com/pricing/

Table 17 GPU pricing for industrial usage

Provider	Cost per hour	GPU	Cost per GPU hour
Amazon (AWS)	£2.69 – £21.51	2 - 16	£1.34
Microsoft Azure	£0.87 - £3.81	1 – 4	£0.87 – £0.95
Google Cloud	NA	NA	£0.49 - £1.91
Penguin Computing	NA	2	£1.74
R Systems	NA	16	£0.45 - £0.97

Note: Figures reported are baseline prices for hourly use, additional support, storage, software or other costs may apply. Prices rounded to 2 decimal places. Prices are converted, where necessary, to Sterling using the Bank of England average exchange rate in 2018. (*) GoogleCloud GPU is not available in the UK so prices from the Netherlands are used.

Source: LE analysis of industry pricing; https://aws.amazon.com/emr/pricing/; https://azure.microsoft.com/en-us/pricing/details/virtual-machines/windows/; https://cloud.google.com/compute/pricing; https://www.penguincomputing.com/wp-content/uploads/2019/03/penguin-computing-hpc-cloud-pod-pricing.pdf; http://rsystemsinc.com/pricing/

A3.2 Costs of accessing commercial training

Table 18 Training pricing for industry

Course Name	Cost	Days	Cost/day
HPC specific courses			
Uncertainty Quantification Training (Hartree)	£250	3	£83
Summer School in HPC	£1,400	10	£140
OpenFOAM Training	£1,050	2	£525
CASTEP	£350	5	£70
Generalised programming courses			
Java Programming Intermediate Course	£1,200	5	£240
Advanced Python: Best Practices and Design Patterns	£1,695	4	£424
Advanced Python	£199	2	£100
Advanced C# Programming Training Course	£250	2	£125

Note: Course prices are indicative and may include accommodation.
Source: LE analysis of industry pricing; https://cossan.co.uk/training/training_UQ2018.php; https://www.csc.fi/web/training/-/csc_summerschool_2019; https://www.openfoam.com/training/schedule.php; http://www.castep.org/CASTEP/Workshop2018; https://www.hotcourses.com/courses-by-subject/Java-Programming-Intermediate-Course-5-days-London-courses/56385210/; https://www.learningtree.co.uk/courses/1906/advanced-python-training-best-practices-and-design-patterns/; https://www.citylit.ac.uk/courses/advanced-python; http://www.londonacademyofit.co.uk/training-course-advanced-c-sharp

A3.3 EPSRC and NERC Consortia

Table 19 EPSRC consortia

EPSRC consortia	Description
UK Turbulence Consortium (UKTC)	Consortium facilitating non-reacting turbulence flows research in the UK using High-End computing.
Materials Chemistry Consortium (MCC)	Consortium working on modelling, predicting structures, properties and reactivities of materials.
UK Car-Parrinello Consortium (UKCP)	Consortium applying quantum mechanics to understand properties of materials.
Plasma HEC Consortium	Consortium supporting research in applied plasma simulations including magnetic fusion and laser-plasma physics.
HEC Biomolecular Simulation Consortium (HECBioSim)	Consortium using HPC to complement experiments with molecular simulations.
UK Consortium on Mesoscale Engineering Science (UKCOMES)	Consortium bringing together expertise from various disciplines contributing to world-class research on mesoscale modelling.
UK Turbulent Reacting Flows Consortium (UKCTRF)	Consortium performing energy efficiency simulations through the modelling of turbulent reacting flows.
UK Atomic, Molecular and Optical physics R-matrix consortium (UK AMOR)	Consortium working in the area of atomic, molecular and optical physics responsible for the development of the internationally recognised R-matrix methodology.

Source: ARCHER website https://www.archer.ac.uk/community/consortia/

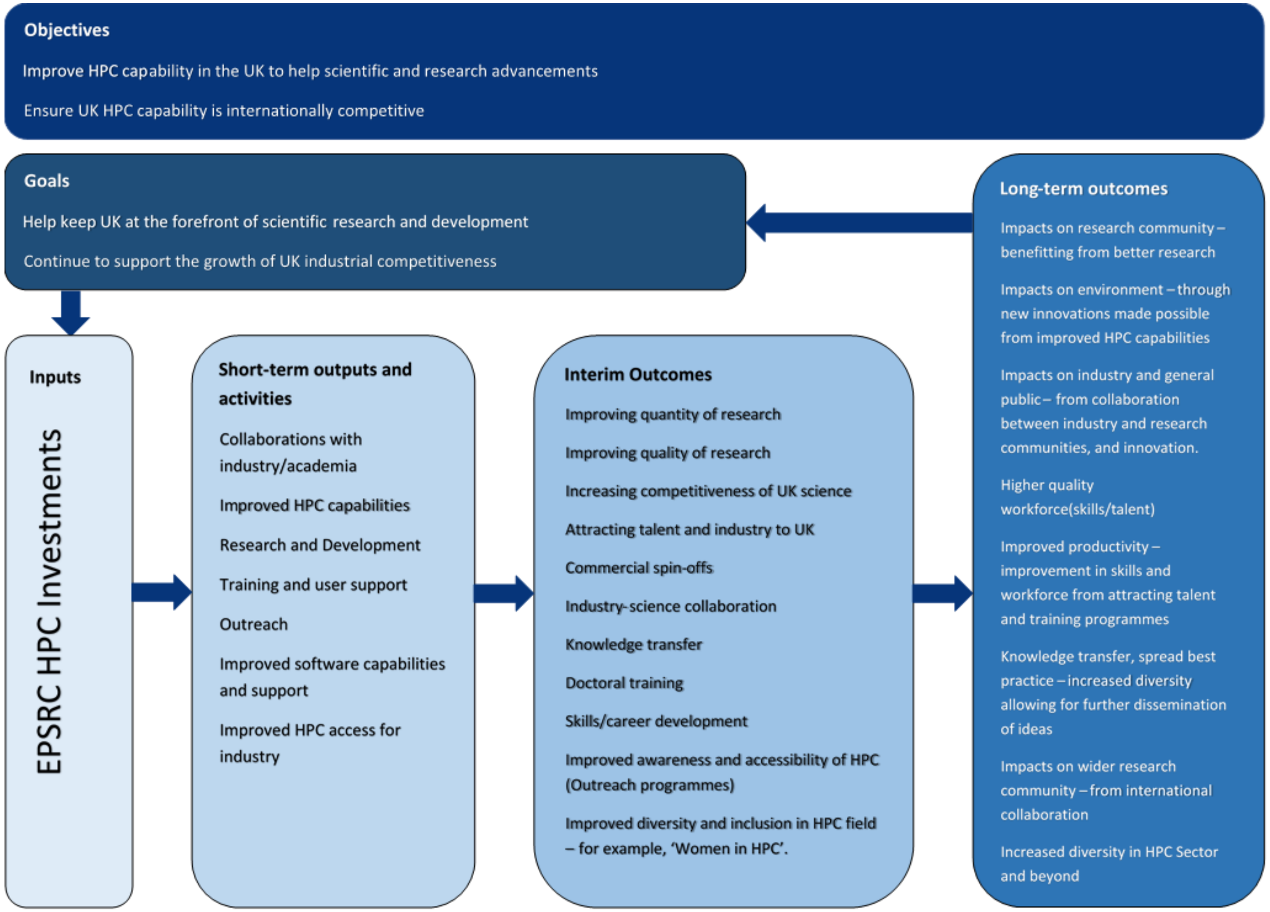
Table 20 NERC consortia

NERC consortia	Description
Oceanography	Consortium undertaking integrated ocean research and technology development.
Atmospheric and Polar Science	Consortium carrying out research on atmospheric science including climate change and hazardous weather.
Mineral and Geo Physics	Consortium using HPC to understand the properties of minerals and melts in Earth’s interior.

Source: ARCHER website https://www.archer.ac.uk/community/consortia/

A3.4 Logic map

Figure 42 Logic map



Source: London Economics

Annex 4 Methodological summary of research spillover papers

This section provides a methodological summary of the academic papers used in the analysis of the research spillover benefits; these are:

- Haskel, J., & Wallis, G. (2010). *Public support for innovation, intangible investment and productivity growth in the UK market sector.*
- Haskel, J., Hughes, A., & Bascavusoglu-Moreau, E. (2014). *The economic significance of the UK science base: a report for the Campaign for Science and Engineering.*

A4.1 The Economic Significance of the UK Science Base

A4.1.1 Broad form model

Haskel, Hughes and Bascavusoglu-Moreau calculate the relationship between total factor productivity and UK public sector science funding between 1995-2007. Their findings suggest a rate of return to public sector R&D of 0.2. They use a model that assumes the change in the industry specific TFP depends on the change in knowledge stock in:

(i) the private sector within the industry,
(ii) the private sector outside the industry and
(iii) the public sector.

Therefore, their model takes the broad form:

$$\Delta \ln TFP_{it} = \alpha_{it}\left(\Delta \ln R_{it}^{PRIV}\right) + \beta_{it}\left(M\Delta \ln R_{\_i,t}^{PRIV}\right) + \gamma_{it}(P\Delta \ln R_{t}^{PUB}) + v_{it}$$

Where $\Delta \ln R$ denotes change in the knowledge stock and M and P reflect the extent to which outside knowledge stock is usable by industry i.

A4.1.2 Data used

Haskel et al. largely constructed R&D data based on the sector that performed the R&D according to data from the official national accounts. For private sector R&D this includes some government funded R&D performed by business, but it is allocated to business because they will learn from it and an ONS survey indicated that a large proportion will actually be owned by business. For public R&D, performance by research councils, higher education and government are summed. Then, the amount of this funded by business is subtracted to reach public R&D.

A4.1.3 Econometric work

Public knowledge stock

To transform their broad form model into a workable equation, Haskel et al. start by writing out γ and $\Delta \ln R_t^{PUB}$ explicitly. To calculate the change in public knowledge stock, the authors use the perpetual inventory model. This assumes that the knowledge stock in a given period to be a function of

(i) the existing knowledge stock, less knowledge depreciation and
(ii) the annual expenditure on research.

Haskel et al.'s model looks at the change in knowledge stock so, assuming the depreciation of public knowledge to be zero, the existing knowledge stock does not need to be included. This results in:

$$\Delta \ln TFP_{it} = \alpha_{it}\left(\Delta \ln R_{it}^{PRIV}\right) + \beta_{it}\left(M\Delta \ln R_{\_i,t}^{PRIV}\right) + \rho_{it} P\left(\frac{N_t^{PUB}}{G_{it}}\right)$$

where $\frac{N_t^{PUB}}{G_{it}}$ is a sum of research council spending, higher education funding council and government performed R&D as a proportion of industry gross output.

When modelling ρ and P, (the elasticity of TFP growth with respect to public knowledge stock and the degree that public sector knowledge is useful to industry i, respectively), Haskel et al. found that they depend on similar variables. Therefore, to avoid collinearity, they used:

$$\Delta \ln TFP_{it} = \alpha_{it}\left(\Delta \ln R_{it}^{PRIV}\right) + \beta_{it}\left(M\Delta \ln R_{\_i,t}^{PRIV}\right) + (\rho_0 + \rho_1 X_{it})\left(\frac{N_t^{PUB}}{G_{it}}\right)$$

where for X_{it} the authors attempt to measure the absorptive capacity of the industry. This is to model the industry's ability to access, understand and apply the results of outside research. For this, Haskel et al. use a measure of co-operation of industry with universities and public research institutes. This data is from the UK Wave of Community Innovation Survey (UKIS), which asks firms if they have formal co-operation agreements with universities or government research centres.

The authors found that short differencing their data resulted in noisy measurements due to measurement error in TFP. Therefore, they long differenced the data, using a lag of three years.

Next, X_{it} is entered as a fraction of its total value in that year. Thus,

$$\frac{X_{it-3}}{\Sigma X_{it-3}}$$

represents industry co-operation with universities and public research divided by its annual sum over all industries. Modelling this, and including time lags, results in:

$$\Delta_3 \ln TFP_{it} = \alpha_{it}\left(\Delta_3 \ln R_{it}^{PRIV}\right) + \beta_{it}\left(\mathrm{M} \ln R_{\_i,t}^{PRIV}\right) + \left(\rho_0 + \rho_1\left(\frac{X_{it-3}}{\Sigma X_{it-3}}\right)\right)\left(\frac{N_t^{PUB}}{G_{it}}\right)_{t-3} + \lambda_i + \lambda_t + \varepsilon_{it}$$

This allows ρ_1 to be a direct reading of the average rate of return to public sector R&D spend. Using a composite measure of research council spend, higher education funding council and government R&D for N_t^{PUB}, and lagging by 6 years, results in a rate of return of 0.2. Note that outside private knowledge stock was dropped from the equation because it was statistically insignificant.

Outside private knowledge stock

The M matrix (the extent to which outside private knowledge stock is useful to the industry) is calculated through attaching a weight based on labour transition. For example, if many workers transfer from industry j to industry i, the weight for industry j will be high. Denoting the labour transition weights by ω and including the three-year time difference, Δ_3, results in:

$$\Delta_3 \ln TFP_{it} = \alpha_{it}(\Delta_3 \ln R_{it}^{PRIV}) + \beta_{it}(\Sigma\omega_{i,t}\,\Delta_3 \ln R_{\_i,t}^{PRIV}) + \left(\rho_0 + \rho_1\left(\frac{X_{it-3}}{\Sigma X_{it-3}}\right)\right)\left(\frac{N_t^{PUB}}{G_{it}}\right)_{t-3} + \lambda_i + \lambda_t + \varepsilon_{it}$$

A4.2 Public support for innovation, intangible investment and productivity growth in the UK market sector

A4.2.1 Broad form model

Haskel and Wallis estimate the relationship between public R&D spend on research councils and market sector spillovers. Their findings suggest that the marginal spillover effect of research council spending is 12.7. To reach this figure, they use growth accounting methods to calculate TFP growth as it is not directly observable. This is then regressed on measures of direct public sector R&D spend, including research councils, to examine spillovers.

The authors begin with the function:

$$Y_t = A_t F(L_t, K_t, N_t^{PRIV}, N_t^{PUB})$$

This implies that output (Y_t) depends on labour input (L_t), tangible capital input (K_t), intangible capital (N_t^{PRIV}), and the stock of freely available public R&D (N_t^{PUB}). A_t captures increase in output not accounted for by increases in the factors of production.

A4.2.2 Data used

The data used for government R&D spend is from the annually published information on science engineering and technology statistics by the Department for Business Innovation and Skills. This includes data on the break-down of spending into its primary purposes: research councils, defence, civil and Higher Education Funding Council (HEFC). For data on TFP and capital stocks, the authors use data from Giorgio Marrano, Haskel, and Wallis (2009)[143].

A4.2.3 Econometric model

Taking the logarithmic form of the function and denoting ε as output elasticity:

$$\Delta \ln Y_t = \Delta \ln A_t + \sum_{X=L,K,N^{PRIV}} \varepsilon_X \Delta \ln X + \varepsilon_{N^{PUB}} \Delta \ln N_t^{PUB}$$

where X is the first three inputs.

The authors rely on a series of assumptions to be able to estimate this.

Firstly,

$$\Delta \ln A_t = a_0 + v_t$$

[143] Giorgio Marrano, M., Haskel, J., and Wallis G., (2009), "What Happened to the Knowledge Economy? ICT, Intangible Investment and Britain's Productivity Record Revisited", Review of Income and Wealth 55, 3, 686-716, August 2019

where v_t is an independent identically distributed error term.

Growth accounting assumes that the output elasticities to the factors of production are equal to that factor's share of income (s_X). Haskel and Wallis follow this practise, but also include a term to account for spillovers from that factor/deviations from perfect competition (d_X). Hence, their second assumption is:

$$\varepsilon_X = s_X + d_X \forall X$$

Looking at the $\varepsilon_{N^{PUB}}$ term, it cannot be measured in terms of factor share as the private factors are because it is freely available. Instead, Haskel and Wallis' third assumption is that the change in the stock of freely available public R&D ($\Delta \ln N_t^{PUB}$) depends on public sector spending on R&D, denoted R^{PUB}. In order to capture other freely available public knowledge, for example from the internet, the authors also include Z. Therefore, the effect of public knowledge is expressed as:

$$\varepsilon_{N^{PUB}} \Delta \ln N_t^{PUB} = \alpha_1 \left(\frac{R^{PUB}}{Y}\right)_{t-1} + \alpha_2 Z_t$$

where α_1 is the rate of return on public sector R&D spend and the ratio is lagged to account for the time taken for R&D spend to result in knowledge gain. Note that this assumes a zero-depreciation rate to public knowledge.

The fourth and final assumption defines TFP growth as the change in total output minus the change in private factors of production:

$$\Delta \ln TFP_t \equiv \Delta \ln Y_t - \sum_{X=L,K,N^{PRIV}} s_x \Delta \ln X$$

Using these assumptions, the authors reach:

$$\Delta \ln TFP_t = \alpha_1 \left(\frac{R^{PUB}}{Y}\right)_{t-1} + \alpha_2 Z_t + \sum_{X=L,K,N^{PRIV}} d_X \Delta \ln X + v_t$$

In words, this means that the change in TFP depends on the freely available public knowledge and spillovers/deviations from perfect competition of factors of production.

To examine the effect of research council spending, the authors use a measure of this R&D spend for R^{PUB}. Therefore, α_1 represents the spillover effect from research council spending. When using TFP data from 1988-2007, Haskel and Wallis find the marginal effect is 12.7. That is to say, every £1 spent by research councils on R&D results in an additional output of £12.70 for UK companies.

LE
London
Economics

Somerset House, New Wing, Strand,
London, WC2R 1LA, United Kingdom
info@londoneconomics.co.uk
londoneconomics.co.uk
@LondonEconomics
+44 (0)20 3701 7700