technopolis group

WHAT METHODS WORK FOR EVALUATING THE IMPACT OF PUBLIC INVESTMENTS IN RD&I

A report prepared for the Department of Science, Innovation and Technology (DSIT)

May 2024

Annex

May 2024

Contents

1 The Influence of Patents in Twenty R&D Portfolios Funded by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy 13
1.1 Overview 13
1.2 Introduction 13
1.3 Background description of RD&I intervention 14
1.4 Challenges measuring outcomes and impact 14
1.5 Methodology and data sources 15
1.5.1 Methodology 15
1.5.2 Data sources 16
1.6 Main findings 16
1.7 Analysis of methodological suitability and effectiveness 17
1.8 Transferability and context 18
1.9 Conclusions 19
2 Location-dependent Public-private Interaction in Catalysing Solar Technology Commercialization 20
2.1 Overview 20
2.2 Introduction 20
2.3 Background description of RD&I intervention 21
2.4 Challenges measuring outcomes and impact 21
2.5 Methodology and data sources 22
2.5.1 Methodology 22
2.5.2 Data sources 22
2.6 Main findings 23
2.7 Analysis of methodological suitability and effectiveness 24
2.8 Transferability and context 24
2.9 Conclusions 25
3 Ex-post impact evaluation of Partnerships for Enhanced Engagement in Research PEER programme 26
3.1 Overview 26
3.2 Introduction 26
3.3 Background description of RD&I intervention 27

3.4 Challenges measuring impact 27
3.5 Methodology and data sources 28
3.5.1 Methodology 28
3.5.2 Data sources 29
3.6 Main findings 30
3.7 Analysis of methodological suitability and effectiveness 30
3.8 Transferability and context 31
3.9 Conclusions 32
4 An Assessment of ARPA-E 33
4.1 Overview 33
4.2 Introduction 33
4.3 Background description of RD&I intervention 33
4.4 Challenges measuring impact 34
4.5 Methodology and data sources 35
4.5.1 Methodology 35
4.5.2 Data sources 35
4.6 Main findings 36
4.7 Analysis of methodological suitability and effectiveness 36
4.8 Transferability and context 37
4.9 Conclusions 37
5 Ex-post evaluation of the Support to large enterprises – Work package 4 38
5.1 Overview 38
5.2 Introduction 38
5.3 Background description of RD&I intervention 38
5.4 Challenges measuring impact 40
5.5 Methodology and data sources 40
5.5.1 Methodology 40
5.5.2 Data sources 42
5.6 Main findings 42
5.7 Analysis of methodological suitability and effectiveness 43
5.8 Transferability and context 44
5.9 Conclusions 44

6 Realisation of a Final Impact Assessment Study for Horizon 2020 for the COST Association 46
6.1 Overview 46
6.2 Introduction 46
6.3 Background description of RD&I intervention 47
6.4 Challenges measuring impact 47
6.5 Methodology and data sources 48
6.5.1 Methodology 48
6.5.2 Data sources 49
6.6 Main findings 49
6.7 Analysis of methodological suitability and effectiveness 50
6.8 Transferability and context 51
6.9 Conclusions 51
7 Evaluation of investments in Research and Technological Development (RTD) infrastructures and activities supported by the European Regional Development Funds (ERDF) in the period 2007-2013 53
7.1 Overview 53
7.2 Introduction 53
7.3 Background description of RD&I intervention 54
7.4 Challenges measuring outcomes and impact 54
7.5 Methodology and data sources 55
7.5.1 Methodology 55
7.5.2 Data sources 56
7.6 Main findings 56
7.7 Analysis of methodological suitability and effectiveness 57
7.8 Transferability and context 58
7.9 Conclusions 59
8 A microeconomic perspective on the impact of the Fraunhofer-Gesellschaft 60
8.1 Overview 60
8.2 Introduction 60
8.3 Background description of RD&I intervention 60
8.4 Challenges measuring outcomes and impact 61
8.5 Methodology and data sources 62
8.5.1 Methodology 62

8.5.2 Data sources 62
8.6 Main findings 63
8.7 Analysis of methodological suitability and effectiveness 63
8.8 Transferability and context 64
8.9 Conclusions 65
9 The macroeconomic effects of the Fraunhofer-Gesellschaft 66
9.1 Overview 66
9.2 Introduction 66
9.3 Background description of RD&I intervention 66
9.4 Challenges measuring outcomes and impact 67
9.5 Methodology and data sources 67
9.5.1 Methodology 67
9.5.2 Data sources 68
9.6 Main findings 69
9.7 Analysis of methodological suitability and effectiveness 69
9.8 Transferability and context 70
9.9 Conclusions 71
10 Accompanying evaluation of the pilot project "ZIM Cooperation Networks International" and investigating the promotion of international cooperation in research and development in the Central Innovation Program SMEs (ZIM) 72
10.1 Overview 72
10.2 Introduction 72
10.3 Background description of RD&I intervention 73
10.4 Challenges measuring impact 73
10.5 Methodology and data sources 74
10.5.1 Methodology 74
10.5.2 Data sources 76
10.6 Main findings 77
10.7 Analysis of methodological suitability and effectiveness 78
10.8 Transferability and context 79
10.9 Conclusions 79
11 Evaluation of the Nano 2017 Programme 80
11.1 Overview 80

11.2 Introduction 80
11.3 Background description of RD&I intervention 81
11.4 Challenges measuring outcomes and impact 81
11.5 Methodology and data sources 82
11.5.1 Methodology 82
11.5.2 Data sources 84
11.6 Main findings 85
11.7 Analysis of methodological suitability and effectiveness 85
11.8 Transferability and context 86
11.9 Conclusions 87
12 Econometric evaluation of aid to collaborative R&D projects (2005 -2019) 88
12.1 Overview 88
12.2 Introduction 88
12.3 Background description of RD&I intervention 89
12.4 Challenges measuring outcomes and impact 89
12.5 Methodology and data sources 90
12.5.1 Methodology 90
12.5.2 Data sources 92
12.6 Main findings 92
12.7 Analysis of methodological suitability and effectiveness 93
12.8 Transferability and context 93
12.9 Conclusions 94
13 Evaluation of the Impact of Individual Innovation Grants Distributed by Bpifrance 95
13.1 Overview 95
13.2 Introduction 95
13.3 Background description of RD&I intervention 96
13.4 Challenges measuring impact 96
13.5 Methodology and data sources 97
13.5.1 Methodology 97
13.5.2 Data sources 98
13.6 Main findings 98
13.7 Analysis of methodological suitability and effectiveness 99
13.8 Transferability and context 99

13.9 Conclusions 100
14 Assessing the collaboration and network additionality of innovation policies: a counterfactual approach to the French cluster policy 101
14.1 Overview 101
14.2 Introduction 101
14.3 Background description of RD&I intervention 102
14.4 Challenges measuring outcomes and impact 102
14.5 Methodology and data sources 103
14.5.1 Methodology 103
14.5.2 Data sources 104
14.6 Main findings 105
14.7 Analysis of methodological suitability and effectiveness 105
14.8 Transferability and context 106
14.9 Conclusions 107
15 World Class Ecosystems in the Finnish Economy 108
15.1 Overview 108
15.2 Introduction 108
15.3 Background description of RD&I intervention 109
Instrument 110
15.4 Challenges measuring impact 110
15.5 Methodology and data sources 111
15.5.1 Methodology 111
15.5.2 Data sources 112
15.6 Main findings of the evaluation 113
15.7 Analysis of methodological suitability and effectiveness 114
15.8 Transferability and context 115
15.9 Conclusions 116
16 The Value of CSIRO: The broader impact of CSIRO's portfolio of activities 117
16.1 Overview 117
16.2 Introduction 117
16.3 Background description of RD&I intervention 118
16.4 Challenges measuring impact 119
16.5 Methodology and data sources 119

16.5.1 Methodology 119
16.5.2 Data and use of monitoring and evaluation inputs 121
16.6 Main findings 121
16.7 Analysis of methodological suitability and effectiveness 122
16.8 Transferability and context 122
16.9 Conclusions 123
17 Business Research and Innovation Initiative Impact Evaluation 124
17.1 Overview 124
17.2 Introduction 124
17.3 Background description of RD&I intervention 125
17.4 Challenges measuring impact 125
17.5 Methodology and data sources 126
17.5.1 Methodology 126
17.5.2 Data sources 128
17.6 Main findings 128
17.7 Analysis of methodological suitability and effectiveness 128
17.8 Transferability and context 129
17.9 Conclusions 129
18 Cooperative Research Centres Programme Impact Evaluation 131
18.1 Overview 131
18.2 Introduction 131
18.3 Background description of RD&I intervention 132
18.4 Challenges measuring outcomes and impact 132
18.5 Methodology and data sources 132
18.5.1 Methodology 132
18.5.2 Data sources 133
18.6 Main findings 134
18.7 Analysis of methodological suitability and effectiveness 135
18.8 Transferability and context 135
18.9 Conclusions 136
19 Evaluation of the Energy Entrepreneurs Fund EEF 137
19.1 Overview 137
19.2 Introduction 137

19.3 Background description of RD&I intervention 138
19.4 Challenges measuring impact 139
19.5 Methodology and data sources 139
19.5.1 Methodology 139
19.5.2 Data sources 142
19.6 Main findings 143
19.7 Analysis of methodological suitability and effectiveness 144
19.8 Transferability and context 145
19.9 Conclusions 145
20 Evaluation of the Enterprise Ireland Research, Development and Innovation Programme 146
20.1 Overview 146
20.2 Introduction 146
20.3 Background description of RD&I intervention 147
20.4 Challenges measuring outcomes and impact 147
20.5 Methodology and data sources 148
20.5.1 Methodology 148
20.5.2 Data sources 149
20.6 Main findings 149
20.7 Analysis of methodological suitability and effectiveness 149
20.8 Transferability and context 150
20.9 Conclusions 150
21 Value of Engineering and Physical Sciences Research Council (EPSRC) Fellowships 152
21.1 Overview 152
21.2 Introduction 152
21.3 Background description of RD&I intervention 153
21.4 Challenges measuring outcomes and impact 153
21.5 Methodology and datasources 154
21.5.1 Methodology 154
21.5.2 Data sources 156
21.6 Main findings 156
21.7 Analysis of methodological suitability and effectiveness 157
21.8 Transferability and context 158
21.9 Conclusions 158

22 Returns on Research Funded under the NIHR Health Technology Assessment 160
22.1 Overview 160
22.2 Introduction 160
22.3 Background description of RD&I intervention 160
22.4 Challenges measuring outcomes and impact 161
22.5 Methodology and data sources 161
22.5.1 Methodology 161
22.5.2 Data sources 163
22.6 Main findings 163
22.7 Analysis of methodological suitability and effectiveness 164
22.8 Transferability and context 165
22.9 Conclusions 165
23 The ISIS Neutron and Muon Source at the Rutherford Appleton Laboratory: Lifetime Impact Study 166
23.1 Overview 166
23.2 Introduction 166
23.3 Background description of RD&I intervention 167
23.4 Challenges measuring impact 167
23.5 Methodology and data sources 168
23.5.1 Methodology 168
23.5.2 Data sources 169
23.6 Main findings 169
23.7 Transferability and context 169
23.8 Conclusions 170
24 Strategic Priority Fund. Interim impact evaluation 171
24.1 Overview 171
24.2 Introduction 171
24.3 Background description of RD&I intervention 172
24.4 Challenges measuring outcomes and impact 173
24.5 Methodology and data sources 173
24.5.1 Methodology 173
24.5.2 Data sources 175
24.6 Main findings 175

24.7 Analysis of methodological suitability and effectiveness 176
24.8 Transferability and context 177
24.9 Conclusions 177

1 The Influence of Patents in Twenty R&D Portfolios Funded by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy

1.1 Overview

Key characteristics	
Country	United States
Institution	U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy
Type of RD&I	Basic and applied research and innovation
Type of Intervention	Research and innovation grants
Evaluation challenges	Contribution; Intangible benefits The key evaluation challenge is observing and identifying contributions and tracing the impact of basic research
Evaluation methodology and methods	• Patent analysis of patent portfolios funded by the programme • Patent citation tracing (backwards and forward)

1.2 Introduction

This case analyses the influence of 20 research portfolios funded across 9 units/offices in the Office of Energy Efficiency and Renewable Energy in the U.S. Department of Energy (DoE) between 1976 and 2018 using a patent tracing method. Using forward and backward tracing makes it possible to measure links between basic science being used in an applied technology over a long period and identify how funded basic research has formed the foundations of innovations and applied technologies.

1.3 Background description of RD&I intervention

This report provides a synthesis of a series of recent patent analyses conducted for the Office of Energy Efficiency & Renewable Energy (EERE) in the U.S. Department of Energy (DoE). The 20 EERE-funded portfolios correspond to approximately 60% of total R&D funding by the nine offices over the period 1976-2018. The nine offices have broad thematic coverage funding projects in areas such as advanced manufacturing, bioenergy, building technologies, geothermal technologies, hydrogen and fuel cell technologies, solar energy, vehicle technologies, water power and wind energy. Accordingly, interventions funded by these offices support research in these areas.

The portfolios have very different profiles concerning research risks, funding levels and time periods covered and there are wide variations in the propensity to patent across technologies.

1.4 Challenges measuring outcomes and impact

Measuring links between basic science being used in an applied technology over a long period presents a difficulty in observing and identifying contributions and tracing the impact. Using citation patent tracing made it possible to link EERE-funded patents to subsequent technology innovations.

The patent tracing analysis requires a specialised skill set to set up the correct filters to identify the patents to be traced. Without them, it is likely to miss out on important data, which can generate misleading evaluation results.

According to the Evaluation Lead in the EERE office[1], once the patents to be traced were identified, verifying if they had been funded by the EERE office presented significant challenges. Around 50% of the patents were easily matched to existing monitoring data of EERE and EERE funding was confirmed. This is because awardees must report the effects of their projects for 5 years after completion. However, the remaining 50% of the EERE funding is, on average, awarded to national labs, who do not distinguish sources of funding, track their data, and therefore do not report to EERE the final project outputs.

In multiple cases, the EERE evaluation team had to contact inventors based in labs and referenced in the patent application to verify the source of funding. In the interview with the Evaluation lead at EERE, it was highlighted that this process took approximately three months and was the only mechanism to confirm EERE funding, with a response rate in almost 80% of the cases. It is worth noting this was a labour-intensive task to ensure the data set was properly built.

Another challenge is to observe the associated benefits of a patent and its links in the absence of prior art references. This is because the responsibility for adding prior art references differs

1 Interview conducted on the 28th of November 2023 to Jeff Dowd, Evaluation Lead at EERE, Department of Energy, US.

across patent systems. According to the evaluation authors, in the U.S. patent system, patent applicants must reference (or cite) all prior art of which they are aware that may affect the patentability of their invention. Patent examiners may then reference additional prior art. In contrast, in patents filed at the EPO and WIPO, the examiner adds prior art references. The number of prior art references on EPO and WIPO patents thus tends to be much lower than the number on U.S. patents.

1.5 Methodology and data sources

1.5.1 Methodology

The rationale for the method choice is a very close link that patent analysis provides to arriving at conclusions about the impact of programme funding on subsequent innovations.

The evaluation questions were:

- To what extent the funded research forms a foundation for subsequent innovations?
- How frequently the portfolio of programmes' funded patents has been cited as prior art by subsequent patents?

In many patent systems, patent documents contain a list of references to prior art. The purpose of these prior art references is to detail the state of the art at the time of the patent application and to demonstrate how the new invention is original over and above this prior art. Prior art references may include many different types of public documents. A large number of the references are to earlier patents, and these references form the basis for this study.

Patent citation analysis focuses on the links between generations of patents that are made by prior art references in patent applications. In simple terms, this type of analysis is based on the idea that the prior art referenced by patents has had some influence, however slight, upon the development of these patents. The prior art is thus regarded as part of the foundation for the later inventions. In assessing the influence of individual patents, citation analysis centres on the idea that highly cited patents (i.e., those cited by many later patents) tend to contain technological information of particular interest or importance. As such, they form the basis for many innovations and research efforts, and so are cited frequently by later patents. While it is not true to say that every highly cited patent is important, or that every infrequently cited patent is necessarily trivial, many research studies have shown a correlation between patent citations and measures of technological and economic importance.

The evaluation used two approaches to citation tracing – forward tracing and backward tracing. Forward tracing takes a given body of research and traces the influence of this research upon subsequent technological developments. Forward tracing involved identifying all patents in each EERE-funded portfolio. The influence of these patents on later generations of technology was then evaluated. This tracing is not restricted to subsequent patents from the technology associated with each portfolio, since the influence of a body of research may extend beyond its immediate technology. Thus, the forward tracing element evaluates the influence of EERE-funded patents upon developments both inside and outside their associated technology.

Backward tracing took a particular technology, product, or industry, and traced it back to identify the earlier technologies upon which it was built. The leading organizations in a given technology (in terms of patent portfolio size) were identified, and tracing was carried out backwards in time from the patents owned by these organizations. This made it possible to determine the extent to which innovations associated with these leading organizations build on earlier EERE-funded research.

The analysis begins with a search undertaken for other U.S., EPO and WIPO patents that are members of the same patent families as these initial patents. These family members were added to the original patent lists.

1.5.2 Data sources

The evaluation constructed a database containing all DoE-funded patents. These include patents assigned to DoE itself, and also patents assigned to individual labs, their contractors, and other organizations and companies funded by DoE. This patent database was constructed using several sources:

- DoE Patents Database. The first source is a database of DoE-funded patents put together by DoE's Office of Scientific & Technical Information, and available on the web. This database contains information on research grants provided by DoE. It also links these grants to the organizations or DoE labs that carried out the research, the sponsor organization within DoE, and the patents that resulted from these DoE grants
- iEdison Database. EERE staff supplied an output from the iEdison database, which is used by government grantees and contractors to report government-funded subject inventions, patents, and utilization data to the government agency that issued the funding award
- Visual Patent Finder Database. EERE also supplied an output from its Visual Patent Finder tool. This tool takes DoE-funded patents and clusters them based on word occurrence patterns. In this case, the output was a file containing DoE-funded patents
- Patents Assigned to DoE in the USPTO database. There are a small number of U.S. patents assigned to DoE itself that were not in any of the sources above. These patents were added to the list of DoE patents
- Patents with DoE Government Interest. A U.S. patent has on its front page a section entitled 'Government Interest', which details the rights that the government has in a particular invention. For example, if a government agency funds research at a company, the government may have certain rights to patents granted based on this research. All patents that refer to 'Department of Energy' or 'DoE' in their Government Interest field, (including different variants of these strings) were therefore identified

1.6 Main findings

The study identified 5,988 patents resulting from EERE funding across the 20 portfolios. The main finding of this evaluation is that patents in the 20 EERE R&D portfolios have had a strong influence on subsequent technological developments. Although EERE-funded patents only represent a small percentage of the total patent universe in their respective technologies (0.6% overall), the analysis reveals that they have been cited 67% more frequently than expected. Patents in the 20 EERE-funded portfolios are also linked via citations to an average of almost 10% of subsequent patent families owned by the leading patenting organizations (98% of which are companies) in their respective technologies.

A significant time lag exists between project start dates and patent application/issue dates. The mean time span from the project start date to the patent application date is 4.5 years, and the mean time span from the project start date to the patent issue date is 7.6 years.

Several companies have large portfolios of EERE-funded patent families. Examples include General Motors (114 EERE-funded patent families), General Electric (105 families), Novozymes (91 families), and Caterpillar (63 families). The 20 EERE-funded portfolios contain many highly cited patents. EERE-funded research is also linked to subsequent spillover technology innovations beyond EERE’s primary research focus areas. Tracing forwards in time from the 20 portfolios through two generations of citations also reveals spillovers into technologies outside EERE’s targeted technologies. These spillovers from EERE-funded research were located in a wide range of industries, including chemicals, electronics, semiconductors, waste management, optics and advanced materials.

1.7 Analysis of methodological suitability and effectiveness

The evaluation had specific evaluation questions focusing on the extent to which the funded research forms a foundation for subsequent innovations. The approach used in this study is a way to address the challenge of tracing the contribution of publicly funded knowledge to the development of key innovations and measuring their intangible benefits. The methodology was successful in providing an answer to the evaluation question and addressing these challenges of tracing contribution. Patent analysis provides a strong link between the funding and subsequent innovation; however, the method does not explain how the impact arises/impact pathways and the role of the funding mechanism design, etc. This type of analysis was not feasible in this study as it covered a large portfolio over a long time. Qualitative insights might be feasible in studies focusing on a single intervention.

The analysis covered a portfolio of instruments with different risk appetite levels, funding amounts, duration, etc. Therefore, the results presented in the evaluation are not intended to be used to compare the EERE-funded portfolios. Potentially, the instruments could be arranged in smaller groups of similar instruments and the analysis could then have a comparative element.

From the Evaluation Lead perspective, the evaluation provides evidence about the intangible benefits of funding basic research and how this research has contributed to innovation in specific research fields. The evaluation allows the EERE to determine in which research fields the EERE was leading based on citation rates, and award rates. The methodology shows how important is the patent, how often it is cited and how many other organisations’ patents are built on EERE-supported patents. It also revealed that EERE patents were critical and more central / very important in the research field, based on the citation index.

It helped confirm the EERE's leadership role in contributing to research and innovation in the field, and therefore the importance of providing funding.

1.8 Transferability and context

The evaluation covered a large portfolio over a long period and the approach of this evaluation is a robust attempt to trace how the basic research materialises in commercial applications over time. The methodological approach is clearly explained and articulated to meet the study objectives and assess the contribution of the funding portfolio to subsequent innovations.

The approach is very specific and relies only on patent data and does not consider other options for assessing the impact on commercialisation (e.g. qualitative inquiry about impact mechanisms), however, because it covers a long period and a large portfolio, it provides robust conclusions about the link between funding and innovation.

The methodology has a good potential to be transferred to the UK context with some limitations regarding prior art references in WIPO and the EPO. The overall approach can be used similarly to assess a portfolio of programmes or for single interventions. case, the approach would likely be accompanied by additional method components.

The evaluation relied on several secondary data sources and any replication of the study would require availability of similar sources. This should not be a problem in the UK context. iEdison database used in this evaluation to identify the patents requires recipients of federal research funding to report inventions and patents to the federal funding agency that issued the funding award. This is similar to many UK research funders using ResearchFish also requiring reports on IP. National, EPO, and WIPO databases also utilised in the evaluation, can be used by any other evaluator. Thus, overall, in technical terms, the study is easy to replicate.

However, the present evaluation heavily relies on prior art references in the patent systems. The responsibility for adding prior art references differs across patent systems. According to the evaluation study authors, in the U.S. patent system, patent applicants must reference (or cite) all prior art of which they are aware that may affect the patentability of their invention. Patent examiners may then reference additional prior art. In contrast, in patents filed at the EPO and WIPO, prior art references are added solely by the examiner, rather than by both the applicant and examiner. The number of prior art references on EPO and WIPO patents tends to be much lower than the number on U.S. patents. This is a limitation or nuance that has to be considered when replicating the approach in European and UK contexts.

A final limitation in using patent data is that companies may choose not to patent, and different industries may exhibit different patenting tendencies or strategies. Therefore, the approach will not be feasible, for example, if the public intervention focuses on disciplines or industry sectors with low patenting activity. Other methods have to be used to uncover other uses of knowledge.

1.9 Conclusions

The approach used in this evaluation is one of the robust ways to establish the socio-economic impact of publicly funded research – an increasing need in the UK context. The approach is a way to address the challenge of measuring intangible benefits of funding basic research and tracing its contribution to the development of key innovations. The approach relies heavily on secondary data and requires well-functioning research information systems. Data sources used in this evaluation should be easily available also in the UK RD&I system.

The reliance on secondary data and methodological approach provides high-level quantitative figures to showcase the value of public investment to the decision makers and general public and thus make the case for further investment. However, it cannot explain pathways to impact and will not yield insights to improve the implementation mechanisms and processes of the funding instruments.

2 Location-dependent Public-private Interaction in Catalysing Solar Technology Commercialization

2.1 Overview

Key characteristics	
Country	United States
Institution	US Department of Energy (DOE), Office of Energy Efficiency and Renewable Energy (EERE)
Type of RD&I	Applied research and innovation
Type of Intervention	Research and innovation grants
Evaluation challenges	Additionality; attribution The key evaluation challenge is determining the causal relationship between receipt of Solar Energy Technologies Office funding and the ability of small US solar tech businesses to secure follow-on private investment.
Evaluation methodology and methods	• Quasi-experimental (regression discontinuity design)

2.2 Introduction

This evaluation investigates the impact of US EERE Solar Energy Technologies Office (SETO) funding on the ability of small US solar tech businesses to solicit follow-on private investment, with a focus on private equity deals.

To address the evaluation challenges of additionality and attribution, this evaluation employs a high-quality regression discontinuity design (RDD) to establish a causal link between SETO funding and the success of a small business in soliciting follow-on private investment.

2.3 Background description of RD&I intervention

US EERE's SETO aims to accelerate the development and deployment of solar technology to support an equitable transition to a decarbonised electricity system in the United States by 2035 and a decarbonised energy sector in the country by 2050.

SETO's activities have included multiple initiatives that provided matching grants to small US businesses working to develop new solar hardware or software technologies (henceforth, solar tech businesses). These initiatives have included the Solar America Initiative, the SunShot Initiative and the Technology to Market Initiative. Over the period 2007-2018, SETO awarded matching grants to 129 small businesses with a total value of 240.68 million USD.[2]

2.4 Challenges measuring outcomes and impact

The evaluation addresses two key evaluation challenges to determining the effect of SETO funding on the ability of small US solar tech businesses to solicit follow-on private investment: additionality and attribution.

Regarding the challenge of additionality, the impact evaluation must address the question of whether follow-on investment received by SETO beneficiaries is likely to have occurred in the absence of SETO funding. This is achieved by employing a high-quality quasi-experimental research design.

Regarding the challenge of attribution, the impact evaluation must address the question of whether SETO beneficiaries share characteristics aside from SETO funding (for example, their geographic location or the type of technology they are developing) which underlie their ability to solicit follow-on private investment. Put differently, it must address the "third variable" problem, in which an unobserved variable correlates highly with SETO funding and muddies any causal explanation. While not all quasi-experimental research methods are able to establish programme attribution, a high-quality 'sharp' RDD[3] such as this one can make credible claims about attribution because assignment to treatment or control groups is as good as random at the cut-off point. This similarity to randomised controlled trials, the gold standard in establishing causality, enhances the credibility of the causal inferences drawn from high-quality RDD studies compared to other quasi-experimental methods such as matching or difference-in-differences.

2 This includes grants awarded under previous iterations of SETO's programmes at DOE.

3 'Sharp' RDD strictly assigns treatment based on a cut-off point, whereas 'fuzzy' RDD sees some crossover in treatment assignment around this threshold.

2.5 Methodology and data sources

2.5.1 Methodology

This impact evaluation investigates the effect of SETO funding on the ability of small US solar tech businesses to solicit follow-on private investment, with a focus on private equity deals. Follow-on funding is defined as that obtained in the first three years after the time of application to SETO. The impact evaluation investigates the impact of SETO funding awarded over the period 2007-2018. After including the three-year follow up period, the overall time period covered by the evaluation is 2007-2021.

The evaluation employs a high-quality RDD to establish a causal link between SETO funding and the success of a small business in soliciting follow-on private investment. This approach leverages companies' SETO application rankings for causal inference. It compares follow-on investment among solar tech companies with similar SETO application rankings, some of which were either narrowly included from receiving SETO funding and others which were narrowly excluded from receiving SETO funding. Because the funding cut-off line is arbitrary enough that most company characteristics near the cut-off are randomised, comparing companies on either side of the cut-off line allows the researchers to make credible causal claims about the effect of SETO funding.

Importantly, the quantitative analysis conducted for this evaluation is of high quality. The researchers employ several methods to ensure that the RDD analysis is sufficiently rigorous to address the challenges of additionality and attribution. First, the evaluation explicitly tests the assumptions of the RDD using a comprehensive set of empirical analyses. This provides a compelling justification for the validity of this research design in addressing the additionality challenge. Second, the researchers conduct robustness tests to ensure that their findings are not artefactual to the RDD specification used in the main models. Third, the analysis includes control variables for a comprehensive set of company and environmental factors to help address the challenge of attribution/ contribution. Notably, the process for selecting these variables is clearly described and justified by the researchers.

Finally, the study conducts exploratory analyses to investigate the relationship between SETO funding and two types of follow-on investment that are of secondary interest to the evaluation: 1) private debt financing, and 2) public or private grants. This involves performing general linear regression models that control for a comprehensive set of company and environmental factors. RDD analyses are not carried out for these two types of follow-on investment.

2.5.2 Data sources

The impact evaluation draws on administrative and secondary data about the companies under study. This data is available over a substantial time period (2007-2021), allowing the researchers to capture data for the three-year follow-up period for 584 solar tech small businesses. In terms of administrative data, the evaluation uses DOE administrative data on unsuccessful SETO applicants and SETO beneficiaries, as well as SETO application rankings for both groups of companies. Data on funding events were collected from three different data sources:

- Pitchbook, a specialized firm focused on research and data analysis on companies, deals, funds, investors and service providers across the entire private investment lifecycle including venture capital, private equity and M&A transactions;
- Crunchbase, a crowd-sourced database of the start-up ecosystem, consisting of investors, incubators, start-ups, key people, funds, funding rounds and events;
- Bloomberg New Energy Finance, data company focused on energy investment and carbon markets research, tracking investment trends and deal flow.

Not only do these data sources allow for retrospective analysis over a substantial period, the cost of collecting such data is low in comparison to that of collecting primary data. Unfortunately, the evaluation report does not detail the data sources used by the researchers to construct variables related to company characteristics used in the analysis.

2.6 Main findings

The evaluation finds that SETO funding alone does not have a statistically significant impact on the ability of small solar tech companies to secure follow-on private equity deals and that the impact of SETO funding is dependent on the location of small solar tech businesses. For those located in a “solar hub”, companies just above the SETO funding cut-off line secure three times as many private equity deals as those just below the line. The evaluation defines a “solar hub” as a geographic location in the United States where there is a significant solar start-up ecosystem. It identifies three such hubs: Silicon Valley, Los Angeles/ San Diego, and Boston. In contrast, the reverse trend is found for companies located outside a “solar hub”. Those just above the SETO funding cut-off line secure a smaller number of private equity deals compared to those just below the line.

With regard to private debt financing and public or private grants, the exploratory analysis finds no relationship between SETO funding and the ability to secure these types of follow-on investment. This finding holds both for the overall sample of small solar tech companies as well as for various subcategories of these companies (for example, those located inside/outside a “solar hub” or software versus hardware companies). Similarly, SETO funding is not found to be related to the number of follow-on public or private grants received from other sources. Again, this finding holds both for the overall sample of small solar tech companies as well as for various subcategories of these companies. Because SETO funding aims to assist independent small businesses to become commercially sustainable, the latter finding is interpreted to suggest that SETO beneficiaries did not require continued government support to operate and is interpreted by the evaluation report to be an ex-post validation of programme success on this dimension.

2.7 Analysis of methodological suitability and effectiveness

The methodology employed in this evaluation has a number of strengths.

- The high-quality RDD analysis allows the research team to draw conclusions about the questions of additionality and attribution surrounding SETO matching grant programs.
- The reliance on administrative and secondary data is cost-effective compared to the collection of primary data.
- The investigation of the heterogenous (i.e. differential) effects that SETO funding may have on companies with different characteristics provides nuanced, policy-relevant insights on the programme's effectiveness. Such nuanced findings provide important insights into the internal validity of evaluation findings (that is, the extent to which the research findings accurately represent the population under study). They can also provide suggestive insights into the external validity of evaluation findings (that is, the extent to which one might expect the same intervention to have similar effects in a different context).

This methodology also has several limitations.

- As the evaluation report notes, SETO funding may be having a displacement effect at the company level, whereby private investors looking to invest in "solar hubs" may be funnelling funding into SETO beneficiaries because the ability to secure a SETO grant is seen as a signal of a company's quality. In this scenario, SETO would be playing the role of a third-party, independent validator of nascent solar technologies. The methodology used in this evaluation does not investigate the causal mechanisms that lead investors to invest in SETO-funded companies, but rather only the outcome of whether this funding is secured. Thus, it cannot determine whether such a displacement effect is taking place.
- It cannot determine the reasons for the finding that SETO funding has a heterogenous (i.e. differential) effect on companies located inside/outside "solar hubs".
- The reliance on administrative and secondary data in this methodology means that this evaluation can only investigate the impact related to outcomes and variables where there is existing off-the-shelf data. Because the off-the-shelf data available does not necessarily map onto a programme's theory of change or policy goals, this can limit the scope of the evaluation in ways that are not optimal for answering evaluation questions.

2.8 Transferability and context

The methodology employed in this evaluation is transferable to the evaluation of some RD&I programmes in the UK context, given the following conditions are met:

- The programme must have an inclusion cut-off line which is arbitrary enough that most company characteristics near the cut-off are randomized.

- There must be a sufficient number (sample size) of programme applicants and beneficiaries near the programme inclusion cut-off line in order to conduct the RDD analysis.
- There must be administrative data available on key characteristics of the programme applicants and beneficiaries that are relevant to the programme objectives (for example, company size, age and/or location) so that these may be included as control variables in the RDD model.
- There must be administrative or secondary data available on the programme impact being studied. The secondary data used for this particular evaluation (Pitchbook, Crunchbase and Bloomberg New Energy Finance) is international in scope and could be used to measure investment in UK-based renewable energy companies. In the case of Pitchbook and Crunchbase, this also extends to investment in other economic sectors.

2.9 Conclusions

This evaluation provides a methodologically rigorous analysis of the extent to which SETO funding improves the ability of small US solar tech businesses to solicit follow-on private investment, with a focus on private equity deals. It employs a relatively cost-effective research design to provide strong evidence that SETO funding improves the ability of small US solar tech businesses located in “solar hubs” to secure private equity deals, but this does not hold for those located outside of these hubs. However, the methodology used in this evaluation cannot determine whether a funding displacement effect is occurring at the company level. Nor can it explain the reasons for the heterogenous effects observed between companies located inside/outside “solar hubs”. This methodology is transferable to the evaluation of some RD&I programmes in the UK context, given that certain conditions related to programme design and data availability issues are met.

3 Ex-post impact evaluation of Partnerships for Enhanced Engagement in Research PEER programme

3.1 Overview

Key characteristics	
Country	United States
Institution	United States Agency for International Development (USAID)
Type of RD&I	Collaborative research and development
Type of Intervention	Fellowships
Evaluation challenges	Additionality The key challenge is to establish the programme's additionality and determine what would have been the outputs of PEER researchers in the absence of PEER funding in terms of research production and improvements to researcher capacity
Evaluation methodology and methods	• Quasi-experimental (matching) • Bibliometric analysis • Generic methods (survey; focus groups)

3.2 Introduction

This case presents the evaluation of the Partnerships for Enhanced Engagement in Research PEER programme funded by USAID over the first five implementation cycles since 2012. The evaluation assessed the contributions and additionality that PEER made to research production, researcher capacity improvements, and the production of evidenced-based policy.

The evaluation adopted a mixed-method approach to answer the evaluation questions. A quasi-experimental method is used to address the programme additionality and a qualitative approach to examine the factors shaping the pathways to impacts.

3.3 Background description of RD&I intervention

The Partnerships for Enhanced Engagement in Research PEER programme directly supports scientists in USAID partnered countries through 2–3-year institutional research awards ranging from $40,000 to $80,000 per year. PEER provides resources to researchers for generating evidence which might have policy implications. Funding builds their capacity to develop and implement high-quality research in development priority areas. They also receive capacity-building support with project management and reporting. As an output of these activities, research projects are designed and conducted. Research activities funded through PEER lead to new evidence to improve researchers' country's development objectives. This research will have long-lasting development impacts if applied successfully in policy-making processes.

PEER has funded over 300 scientists and engineers across geographies and thematic research areas for a nine-year cooperative agreement. Producers of research comprise the Principal Investigators and Co-PIs in universities and research institutions who receive support to conduct research. These actors also directly collaborate with the users of the research.

During its first five years, PEER emphasized increasing the number of projects, improving the relationships with missions/buy-in, and streamlining the solicitation and awards processes. In the ongoing second half of the programme, PEER emphasizes project impact, translation of results to key stakeholders, and capacity-building activities in partner institutions. From PEER Cycle 5 onwards, a section on Government & NGO partner collaboration was added, and the development impact section was expanded in the application. Grants awarded from 2016 to 2018 included opportunities to apply for evidence to action or travel grant funding for U.S. partners.

3.4 Challenges measuring impact

The main challenge faced by the evaluation team was to assess the programme contribution that PEER made to research production, research capacity improvements, and the production of evidenced-based policy.

The mixed-method approach adopted helped tackle the breadth of the research questions, while the quasi-experimental method sought to address mainly additionality. Nevertheless, the lack of a control group led the evaluation team to find an alternative to approximate a counterfactual, using non-successful applicant data held during the selection process. The most pressing challenge lay in the balance in the composition of the comparator group, which may have influenced the evaluation results. Similarly, other sources of funding available to unsuccessful applicants represented another challenge to address by the evaluation.

The evaluation examined the factors increasing the likelihood of achievement of the programme objectives and its impacts and understood under what conditions research results may influence decisions in policymaking. Looking at these factors enabled the evaluation to confirm pathways to impact and more detail about the areas in which PEER's additionality was more salient.

3.5 Methodology and data sources

3.5.1 Methodology

The evaluation seeks to answer the following questions:

- What factors (at the levels of PI, institution, and country) increase the likelihood of achieving programme and policy impacts? How has the PEER grant helped to generate new knowledge and be used for policymaking?
- To what extent and how has PEER contributed to programme and policy impacts?
- What kind of evidence is most likely to influence policymakers?
- To what extent and how has PEER contributed to capacity-related outcomes? (career progress, students with degrees, data production)
- What factors (at the levels of PI, institution, and country) increase the likelihood of achieving capacity-related outcomes? (publications, students with degrees, data production)
- Are there regional differences in motivational factors for involvement in PEER-funded activities?

The evaluation adopted a mixed-method approach and utilised secondary and primary data. The quantitative analysis employed a quasi-experimental matching methodology. The evaluation team travelled to three countries (Uzbekistan, Colombia, and Morocco) to conduct in-person interviews and site visits.

The analytical approach sought to **establish the programme's additionality** and answer the question: "What would have been the outputs of PEER researchers in the absence of PEER funding?". However, the evaluation lacked an established control group to analyse the survey and bibliometric data against it.

Despite this, the good availability of secondary data made possible the adoption of a quasi-experimental method. The programme maintained a database of all applicants to the PEER programme — both those applications which were funded and those which were not. The programme also held meaningful documentation on the selection process of the applicants, including scoring and recommendation procedures by review panels. Thus, evaluators were able to employ quasi-experimental matching methods using these points of information.

To approximate the counterfactual, the evaluators observed the outputs and outcomes of PEER applicants who ultimately did not receive funding from the PEER programme. The evaluation team reviewed the data from the PEER applications provided by the National Academy of Sciences. The application data set includes the title of the project, the sector, the gender of and highest degree earned by the principal investigator, the country of the research, the budget of the project, and reviewers' scores of the application. The evaluation applied the following approach:

- First ran balance tests to confirm that the raw data demonstrates that balance is not achieved between the treatment group (PEER-funded researchers) and the comparison group (non-PEER funded researchers) on some of these indicators
- Explored three matching methods:
- Employed coarsened exact matching, which improves balance by an amount specified by the analyst through temporarily coarsening the variables used for matching and matching on the coarsened variables
- Entropy balancing employed a reweighting method to achieve balance on the specified matching variables while preserving as many of the observations as possible
- Mahalanobis distance matching calculated a "distance" between each treatment and comparison observations, which then pares down the data to matched pairs among the nearest matches

3.5.2 Data sources

The evaluators reviewed 1,937 PEER rejected and accepted applicants across 54 countries to compare research output from 2012 onwards. The evaluation included:

- Document reviews
- An online survey of all PEER applicants: In addition to secondary data analysis, the evaluation also collected primary data. An online survey was drafted by the evaluation team in consultation with key stakeholders. Evaluators aimed to ensure that survey questions were consistent with those asked in the midterm evaluation and that they were relevant to both PEER and non-PEER funded researchers so they could be asked to both groups. Due to concerns about recall, some questions were asked about two time periods— since applying for the grant and in the last two years. The survey was opened to respondents twice, in October 2019 and January 2020, to increase the response rate.
- Interviews: The qualitative methodology utilised semi-structured interviews with predefined evaluation questions and preliminary results from the online survey. The semi-structured interview guides were reviewed and revised by key stakeholders at USAID and translated into local languages when necessary. The evaluation team sought to interview all PEER-funded PIs within the selected country that were funded in Cycles 1-5, as well as corresponding research users identified for each PI. No qualitative interviews were conducted with non-PEER researchers (the comparison group).
- Web-scraping, bibliometric analysis: To gain more information about research output, the evaluators also conducted a web-scraping activity that searched for all research outputs of each researcher in the dataset. The evaluation created a triangulated list of digital object identifiers (DOIs) of all publications from researchers in the dataset. The first source of information for this was the list of PEER researcher outputs provided by the PEER program itself. Next, the evaluators requested this information through the

online survey, asking researchers to upload a CV or provide a link to a research page on Google Scholar or ResearchGate. Publication information was pulled from these sources. Finally, if the researcher did not respond to the survey, the evaluators manually searched for their Google Scholar or ResearchGate pages. Once the researchers' Google Scholar IDs or ResearchGate websites were determined, their publications were scraped from their profiles. The R package scholar was used for Google Scholar profiles, and the R package rvest was used to help scrape the ResearchGate profiles. DOIs were scraped from these sites using Python or R.

- Focus group discussions

3.6 Main findings

The findings of the evaluation are mixed. The qualitative research team found effects of the programme on policy and programming, academic output, production of academic output, and career advancement of both researchers and students. Analysis of the quantitative data did not produce such results. With the notable exception of career advancement, the evaluation did not detect a statistically significant impact on any outcomes analysed.

PEER grantees are 10% more likely to have received a promotion, an increase in salary, or an additional paid or advisory role than PEER applicants who did not receive funding. This result could reveal an element of prestige and respect gained by PEER researchers. PEER researchers can purchase high-quality equipment, bring in funding from the United States, and collaborate with U.S. government-funded researchers to produce high-quality output.

There is, however, a negative relationship between participation in the PEER programme and the production of technology or patents. This could be due to the focus of the research selection process on development and policy-related issues. In addition, the focus of the PEER programme activities may have prioritised policy-related outputs over technology or patent production.

The quantitative analysis did not detect an effect on academic output, policy impact, receiving additional funding, or other capacity building outcomes.

3.7 Analysis of methodological suitability and effectiveness

The use of a mixed method approach allowed to assess the additionality of the programme and its results. The evaluation presents several shortcomings:

- The evaluation did not have sufficient observations to quantitatively detect the programme's effect. Low response rates to the survey likely introduced shortcomings to the data and the null results emerging from the quantitative data
- There is a possibility that comparison researchers were acquiring funding elsewhere and achieving the same outcomes as the PEER researchers. This could explain why the evaluation did not observe a difference in the number of research outcomes. If that is

the case, it raises questions about the validity of using un-funded researchers as a control group. In the future, a market analysis could help provide an understanding of the dynamics at play to establish if there is a surplus of grant funding in PEER research countries or sectors

- Recall bias because of the long period covered

3.8 Transferability and context

In principle, the mixed methods approach and use of quasi-experimental matching methodology to establish the counterfactual were appropriate to provide answers to the evaluation questions and establish the additionality of the programme. However, mixed findings point to several challenges with the un-funded applicants as a control group.

To apply the methodology in other contexts, first, the evaluators should have access to programme data on un-funded applicants to use the data in matching methodology. If the programme funder does not systematically collect such data, the replication of the approach is not possible. Second, this evaluation provides a lesson about the need to consider the appropriateness of the control group and the appropriateness of using the approach for the specific funding instrument. Evaluators conclude that in the future, a market analysis could help provide an understanding of the dynamics at play to establish if there is a surplus of grant funding. The final beneficiaries of the funding are individuals. Perhaps there is a higher likelihood that un-funded individuals will find other funding and achieve the same or similar results as the funded ones and thus cannot properly function as a control group.

The evaluation was not successful in dealing with recall bias. The evaluation faced a low survey response rate, which might be partly explained by and linked to recall bias because of the long period covered. This confirms the need for cautiousness in using surveys to collect data on awards funded long ago.

The methodology employed in this evaluation is transferable to the UK if certain conditions are met. To apply the quasi-experimental matching methodology, the funder of the RD&I programme needs to collect systematic monitoring data on un-funded applicants, and this practice differs for different funders and programmes in the UK. As the lessons from the evaluation show, the evaluation needs to collect sufficient data points for analysis to avoid inconclusive findings. This might be challenging if the evaluation has to cover a programme with awards that started long ago. Overall, because of the specific topic of the programme and evaluation, it may be especially interesting to evaluate the impacts of the RD&I programmes from an international development angle. This includes formulating expected outputs and outcomes, data collection (including in target countries) and an overall mixed-methods approach.

3.9 Conclusions

The evaluation is a robust attempt to establish the research outcomes, policy impact and capacity building of researchers and students and overall additionality of the programme. The evaluation applies a mixed-method approach and quasi-experimental matching methodology to assess the counterfactual. The evaluation delivered answers to the evaluation questions, although the findings are mixed. The quasi-experimental matching methodology, in principle, is appropriate for assessing the programme additionality if good quality monitoring data on non-funded applicants exist.

However, the experience from this evaluation highlights two lessons. First, the quantitative methodologies require rich data/number of observations to avoid concerns about the validity of the findings, and this can be challenging if data are acquired from primary sources, e.g. surveys, as in this evaluation. Second, the choice of counterfactual has to be carefully considered. Lessons from this evaluation highlight that there might be various reasons for the control group achieving results comparable to the treatment group (e.g. they benefit from other funding). Thus, a wider market analysis is necessary at the study's onset to establish the control group's appropriateness.

This methodology is transferable to the evaluation of RD&I programmes in the UK context, especially for the programmes with development objectives where the approach to qualitative analysis can be relevant. The programme's funder needs to hold monitoring data on non-funded applicants for the evaluation to replicate the quasi-experimental matching methodology.

4 An Assessment of ARPA-E

4.1 Overview

Key characteristics	
Country	United States
Institution	Advanced Research Projects Agency-Energy
Type of RD&I	Basic and applied research and innovation
Type of Intervention	Research and innovation grants
Evaluation challenges	Additionality; Contribution; Lagged effects Two key challenges are to assess the contribution made by the funding to the formation of new communities of research and the emergence of significant technologies from basic research and APA-E's relevance (additionality)
Evaluation methodology and methods	• Generic research methods (outputs analysis)

4.2 Introduction

This case analyses the preliminary impact of the funding awarded by the Advanced Research Projects Agency-Energy ARPA-E after the first six years of its creation (2009 and 2015). ARPA-E supports high-risk and early-stage research in Energy, and therefore, the evaluation faces significant challenges in assessing its impact at a very early stage, as well as the contribution made by the funding to outcomes generation and its complementary to other energy research funding offered by complementary agencies.

The evaluation adopts a generic research method approach using patent data, bibliometrics, monitoring and evaluation data and primary data from interviews to quantify preliminary outcomes, determine their relevance and impact on the field, and benchmark them against other complementary funding offered by the Department of Energy in the US.

4.3 Background description of RD&I intervention

The Advanced Research Projects Agency-Energy ARPA-E was established in 2009 for the stated purpose of funding energy technology projects by “identifying and promoting revolutionary advances in fundamental and applied sciences; translating scientific discoveries and cutting-edge inventions into technological innovations; and accelerating transformational technological advances in areas that industry by itself is not likely to undertake because of technical and financial uncertainty." ARPA-E was intended to function as an adaptation of the Defence Advanced Research Projects Agency (DARPA) model which is widely considered a successful example of funding high-risk high-reward research. ARPA-E was envisioned as a means of tackling energy challenges in a way that could translate basic research into technological breakthroughs while also addressing economic, environmental, and security issues. However, it was unclear whether the DARPA model would work also in energy technologies and therefore the government mandated a review of the agency's performance in six years of its operations.

To enable the agency to be well positioned to identify and support this kind of research, it is exempted from many federal rules and regulations. A key element of the ARPA-E operations is reliance on programme directors who are hired for limited terms and are highly empowered to act outside of the box from the design of the programmes up to involvement in the management of the research and technical aspects of funded projects. The organisational structure is flat, and it can easily initiate and terminate projects based on performance. Funding decisions are made faster compared to other government research funders.

ARPA-E funds projects very early in the TRL scale, typically around the first translation from scientific discovery to engineering - focusing on ideas at technical readiness levels 2-4 with the possibility of leading to a marketable product.

4.4 Challenges measuring impact

This evaluation faces two main challenges, one is to determine the **contribution** made by the funding to the formation of new communities of research and the emergence of significant technologies from basic research. Answers to these evaluation questions provide evidence of the role of ARPA-E, its performance and early impact. Alongside contribution lies the challenge of assessing ARPA-E's funding **additionality**. Benchmarking ARPA-E intermediate results against those generated by comparator agencies, such as the EERE, allowed the evaluation to state their relevance and the difference made by this funding.

The rationale for the chosen methodology and heavy reliance on secondary sources is linked to the type of funding ARPA-E provides. ARPA-E's focus is on early-stage, high-risk, high-reward energy research. Development of transformative energy technologies from initial discovery to broad market deployment typically takes several decades. Most ARPA-E awards last for about three years, much shorter than the decades required to commercialize energy technologies. Few data were available for this study regarding ARPA-E's impact on energy technologies or the sector as a whole. Therefore, the evaluation utilised the secondary data resulting from six years of operation to demonstrate the intermediate impacts of ARPA-E's activities and address the lagged effects of these investments.

4.5 Methodology and data sources

4.5.1 Methodology

The evaluation used generic research methods. Most methods and analyses focus on the organisational evaluation of the ARPA-E and are not covered in this case study. This case study focuses on the evaluation task asking to assess the success of the focused technology programmes in spurring the formation of new communities of researchers in specific fields.

Specific evaluation questions were:

- Have the focused technology programs spurred the formation of new communities of research? If so, how sustainable have these communities been?
- What, if any, significant products, and technologies have emerged from these communities of researchers?
- How have the products and technologies from these communities of researchers contributed to ARPA-E's mission or stated goals?

The evaluation relied heavily on secondary data sources and completed a benchmark with other programmes funding energy research to establish whether the ARPA-E fulfils its unique mission in the funding mix. The main secondary data sources were general and ARPA-E publications, patents, commercialized outcomes, revenues, novel collaborations, and communities built. The primary data source was interviews with stakeholders about outcomes.

The quantitative analysis of publicly available secondary data was used to identify where ARPA-E projects have made an important difference and to help assess the performance of the portfolio of awards and evaluate whether and to what extent the portfolio has helped ARPA-E achieve its goals. The impact of ARPA-E was compared with that of Department of Energy (DOE) Office of Science and Office of Energy Efficiency and Renewable Energy (EERE).

4.5.2 Data sources

A dataset of all awards offered by the three agencies using publicly available data was created. This dataset linked award-level specifications for recipient type, amount of funding, and project length to publicly available outcomes credited to each award. The sample included data on patents, patent quality, publications, and publication quality.

The dataset was created using award data from the Data Download page of USAspending.gov run by the Department of the Treasury. USAspending.gov provides publicly accessible data on all federal awards. The resulting dataset contained 5,896 awards, 263 of which were from ARPA-E. Publication outputs for DOE awards were obtained from Web of Science. To determine whether a publication was highly cited, additional information was downloaded from Thomson Reuters to supplement the publication analysis.

A patent-level dataset also was created based on distinct patents that acknowledge an award. The evaluation measured also patent quality by patent and the number of claims made on a specific patent.

4.6 Main findings

The evaluation found that ARPA-E projects were more likely to publish and to do so with high frequency relative to other DOE offices. 44 % of ARPA-E awards recipients published at least once, compared with the Office of Science and EERE at 27 % and 18 % of awards, respectively. ARPA-E awards resulted in more publishing in top journals relative to EERE awards and in more energy journals relative to Office of Science awards.

ARPA-E's portfolio of projects resulted in more patents per project than the portfolios of either the Office of Science or EERE. The 13 % of ARPA-E awards have resulted in at least one patent while only 2% of Office of Science projects and 5% EERE projects have resulted in patents. Moreover, the odds of an ARPA-E awardee being granted at least one patent are three times higher than those for an EERE awardee.

ARPA-E projects had a greater likelihood than EERE projects of producing patents that would be cited. ARPA-E projects produced patents with more claims relative to the projects of both the Office of Science and EERE. These measures suggest that on average, the patent portfolio of ARPA-E projects is of higher quality than those of the other two offices.

Among the three comparator organizations, ARPA-E is serving a unique role in terms of funding projects that are focused on the energy system, hold both scientific and technological potential, and are associated with early-stage indicators of practical impact. The metrics used in this evaluation represent only intermediate outcomes but overall, these results demonstrate ARPA-E's productivity and the contribution the agency makes to the DOE mission.

4.7 Analysis of methodological suitability and effectiveness

The methodology was successful in establishing what are the intermediate outcomes of the ARPA-E programmes and assessing whether the funding is complementary to other energy research funding. Given the focus on low TRL funding, investment in high-risk high-reward research, short period elapsed since the start of the ARPA-E programmes and the overall challenge of lagged effects of research investment, the evaluation efficiently achieved the evaluation objectives and demonstrated the possibility of secondary data use.

The intermediate impact evaluation was part of the larger assessment of the operations and processes of ARPA-E and could not have a more comprehensive approach to impact assessment. Therefore, the use of secondary data sources to establish the intermediate outcomes and detect early positive effects of the funding was the most suitable approach in these circumstances. Future methodological improvement might be introducing attempts not only to establish intermediate outcomes but also estimate future outcomes by, for example, surveying award recipients and asking for estimates based on progress achieved so far.

4.8 Transferability and context

The main lesson of this evaluation is that secondary data sources are useful to arrive at intermediate outcomes of funding that support low TRL, high-risk, high-reward research with long timescales to impact. In this case, the approach was used in a high-level evaluation of organisational performance and was deemed sufficient to judge the organisation's and its programmes' success.

The approach is transferable to any other context where secondary data on project outcomes can be linked to funding. Two key measures used in this evaluation – publications and patents – are usually tracible in scientometric databases if funding acknowledgements are included. The evaluation of ARPA-E as such and how the impact of the funding is assessed in this evaluation is a relevant and interesting example for all governments that similarly try to replicate the DARPA model and fund blue skies research.

4.9 Conclusions

The evaluation of the ARPA-E programme's impact is a methodologically rigorous, although simple, analysis of the emerging impact of high-risk, high-reward research with long impact timeframes. The evaluation relies heavily on secondary data and demonstrates intermediate outcomes, acknowledging the difficulties of evaluating the impact that might only emerge in 10 or more years after the funding.

This approach is transferable to evaluating the RD&I programmes in the UK, as similar secondary data should also be available in the UK funding landscape.

5 Ex-post evaluation of the Support to large enterprises – Work package 4

5.1 Overview

Key characteristics	
Country	European Union
Institution	European Commission
Type of RD&I	R&D grants and investments
Type of Intervention	Advanced support services for firms or groups of firms, R&I grants and other investments on firms
Evaluation challenges	Additionality The challenge is to assess the effect on large enterprises' investment decisions (behavioural additionality) and how that compares to prior impact evaluation evidence of this type of intervention
Evaluation methodology and methods	• Theory-based evaluation (contribution analysis; comparative case studies)

5.2 Introduction

This case analyses the evaluation of the programme supporting R&I activities in large enterprises with funding from the Cohesion Fund CF and the European Regional Development Fund (ERDF) between 2007-2013. The evaluation assesses the behavioural additionality of the programme and establishes the socio-economic impacts of the intervention.

To address the challenge of measuring the behavioural additionality of the funding in large companies, the evaluation adopts a Theory-Based method underpinned by a contribution analysis.

5.3 Background description of RD&I intervention

The work package 4 is the programme supporting R&I activities in large enterprises with funding from the Cohesion Fund CF and the European Regional Development Fund (ERDF).

Between 2007-2013 this support reached EUR 6 billion at the EU-28 level. This evaluation covers eight countries, Poland, Germany, Portugal, Spain, Hungary, Italy, the Czech Republic, and Austria, who received EUR 4.6 billion from the total funding. Most of this funding is provided as non-refundable grants, although for four members refundable funds were also available to large companies as loans (Spain, Portugal, Italy and Austria).

The main goal of the support to large firms is to foster regional employment, increasing GDP and fostering economic growth, alongside increased R&D and innovation[4], as well as exports growth. The underlying assumption was that the support would lead to positive indirect effect and wide benefits and spillovers, such as improved practices in large companies that could be replicated by SMEs later, as well as improved infrastructure.

On average, each project is worth EUR 1 million, and each funded company – 3,700 funded approximately – executed on average 1.6 projects[5]. The support provided to companies fall under three main categories, Advanced support services for firms or groups of firms (Code 05), Investment in firms directly linked to research and innovation (Code 07), and other investment in firms (Code 08) of the European Regional Development Fund ERDF.

Funding is channelled through operational programmes, which respond to the local conditions and needs being tackled by the intervention. For some, competitiveness is more relevant, while for others the support is to improve their embeddedness in the region (Austria), technological upgrades, more environmentally friendly production and innovation (Czech Republic). For each country, one operational programme was selected for evaluation.

Country	Operational Programme
Austria	Operational Programme Styria 2007-2013
Czech Republic	Operational Programme Enterprise and Innovation
Germany	Operational Programme Thuringia 2007-2013
Hungary	Economic Development Operational Programme
Italy	National Operational Programme for Research and Competitiveness
Poland	Operational Programme Innovative Economy
Portugal	Operational Programme Thematic Factors of Competitiveness

4 According to the report, this objective was set up for some of the countries analysed but was not shared generally.

5 Companies may have received funding more than once.

Spain	Comunidad Valenciana Operational Programme

Source: Taken from KPMG/Prognos (2016)

In terms of participating firms, 43% have less than 250 employees – but are treated as large companies as they are subsidiaries of large corporations, 42% have between 250-999 employees, and 13% employ more than 1,000.

5.4 Challenges measuring impact

The main challenge being addressed by the evaluation is the **behavioural additionality** of the intervention, considering the assumption that the support to large enterprises may have a positive direct or indirect effect apart from wider impacts at the regional level.

Another challenge identified is the time lag of impacts as not all of them were realised by the time the evaluation or because projects failed to generate the expected outcomes.

The fluidity of enterprises and their investments is also mentioned, although not entirely and explicitly addressed in the evaluation.

5.5 Methodology and data sources

5.5.1 Methodology

This evaluation assesses the rationale for the intervention, its implementation and the effectiveness of the Cohesion Policy support to large enterprises between 2007–2013. To this end, it seeks to understand the circumstances and conditions under which EU Cohesion Policy support to large enterprises is desirable and justified. It evaluates the socio-economic impacts of the intervention, and the factors contributing to its success or failure.

This evaluation seeks to answer the following questions[6]:

- What direct effects and indirect/wider benefits can the support achieve and on what time scale?
- In view of these benefits, what kind of large enterprises or projects should be supported, and how can they be influenced?
- What kind of large enterprises or projects should not be supported?
- What can be done to keep large enterprises in the region in which they were supported?

6 Report, p. 17

The evaluation adopts a theory-based approach focusing on the causal relationship between the support and observed outcomes, the assumptions and alternative explanations for change. A contribution analysis is chosen to unpack the causality and contribution of the intervention to the expected outcomes, underpinned by comparative case study analysis.

A central objective of this evaluation is to "de-compose complex causal chains into micro-steps"[7]. Theories of Change are used to analyse expected outcomes, assumptions and external factors influencing causal chains.

Under the assumption that financial support leads to a change in the large firms' behaviour, the evaluation assesses the **behavioural additionality** of the programme. In addition, it seeks to offer alternative views to previous impact evaluations that have found no impact or low impact on this type of intervention. Therefore, to determine what difference has the intervention made and how that compares to prior impact evaluation evidence.

A step-by-step explanation of the evaluation process is provided in the report. It begins with data collection of this support in the eight EU member states, to aggregate information about enterprise support under Cohesion Policy between 2007-2013. Most national interventions do not show a logic behind the intervention, for which the evaluation develops specific theories of change, using policy documents, interviews with Managing Authorities, and complementing it with a literature review and existing evaluations and research studies on the intervention.

A total of 27 programme-level Theories of Change are reconstructed covering eight countries with operational programmes. To systematically assess and refine cause-effect linkages within the Theories of Change, the evaluation carried out 45 company case studies to analyse the outputs and outcomes emerging at the project level and to confirm or reject causality claims.

Company case studies adopt a multi-respondent approach and entail about 130 semi-structured interviews with corporate officials and regional stakeholders. Evidence is collected around relationships (micro-steps), outcomes, assumptions and influencing factors. Further analysis is carried out to establish whether outputs were achieved and the reliability of evidence. Alongside this information, monitoring data and literature review support triangulation, which was then balanced with academic expert interviews.

These 27 Theories of Change were aggregated and generalised into four Theories of Change, following the explanation of the causal chains. The comparative analysis of case studies builds on these Theories of Change, which also substantiates the conclusions of the final report.

An operational programme is selected for each country to develop the country-level case studies, based on three main factors. The significance of the support provided by the programme to large enterprises, its representativeness of large firms' spending at the national level and the availability of counterfactual impact evaluations for comparison.

7 Report, p. 20

Findings from the case studies and the aggregated analysis of the Theories of Change were enriched with a stakeholder seminar bringing together more than 40 participants from the eight case study countries. Throughout the evaluation, four steering group meetings strengthened the methodological approach and helped refine the causal chain of the theories of change and the information of the operational programmes.

5.5.2 Data sources

- Monitoring and information data of operational programmes.
- Programme and project documents
- Programme of interviews with Managing Authorities, company officials (45 companies), academic experts, intermediate bodies
- Stakeholder workshop
- Business case studies (literature)
- Steering group meetings with other work packages of enterprise support to refine concepts and discuss early findings

5.6 Main findings

The evaluation found that the investment decisions of large companies are influenced by a set of factors, and ERDF funding acts as a pre-condition for such investments. Some of these factors are tax incentives, availability of infrastructure, corporate strategy. Only in 20% of the cases, the intervention accounts for a behavioural change in large firms leading to project implementation. Larger effects, in 50% of the cases, were found in terms of behavioural changes around timing and scope of investment decisions. For 30% of the firms the intervention showed little of no influence and additionality.

Outcomes generated by projects showed wider positive effects, but these were difficult to attribute solely to the ERDF support. The intervention showed the highest effects in job creation and exports, generating over 60,000 new jobs and increasing the export based of those regions where firms are established.

Innovation capacity is also favoured by the intervention, enhancing R&D infrastructure and the involvement of local universities, and SMEs. Particularly in cases where large enterprise projects were aligned with the local industrial base and well-articulated to existing research and innovation ecosystem. Thus, another finding was that outcomes and impacts thought during the design of the operational programme, the ERDF support was likely to generate wider and stronger positive effects. This holds true for cases in which funding attracted new big corporations to set up in the regions studied, and those companies were a good matching for the current regional capabilities and industrial base.

In a similar vein, the effectiveness of the ERDF support to generate impacts is higher in mid-caps companies – those having between 250- 3,000 employees – as those face more frequently limited access to finance and therefore, the intervention can determine the decision of implementing an innovation project.

Across the case studies, the lack of planning around wider effects came out as a significant barrier for the realisation of wider impacts, especially for SMEs. Finally, the evaluation drew recommendations in three main domains to improve the effectiveness and additionality of the ERDF support. First improving the selectiveness of the intervention, focus on attracting funding from third countries under conditions that can support longer-term effects (sustainability). Likewise, inducing effects on SMEs via programme design and treating them as desirable beneficiaries of wider impacts of intervention.

5.7 Analysis of methodological suitability and effectiveness

This evaluation assessed the effectiveness of the ERDF support to large firms in eight operational programmes rolled out in Poland, Germany, Portugal, Spain, Hungary, Italy, the Czech Republic, and Austria. These countries account for the highest spending[8] of the total EU investment in supporting large enterprises R&I activities. The evaluation sought to provide a more granular understanding of the additionality of the ERDF support, to contrast and complement previous counterfactual analysis of the intervention showing no or little effect of the support in large enterprises.

To this end, the evaluation adopted a theory-based approach and used a contribution analysis method to establish causality chains of the intervention and determine to what extent this influences R&I investment decisions of large firms. The contribution analysis builds on comparative case studies developed for each of the countries based on eight operational programmes and 45 cases firms' projects that helped reconstruct theories of change of the intervention.

Overall, the evaluation explained causality using multiple sources, including programme information for each country, literature, experts views, and important primary information from 130 interviews. Country-level case studies played an instrumental role in this, answering most of the evaluation questions and substantiating causality claims. Evidence was triangulated and complemented with a stakeholder seminar held with participants from the eight nations and several steering group meetings with officials and Managing authorities from complementary work packages 3 and 4 to refine scope and objectives of the work package 4.

Through this methodological approach, the evaluation assessed the behavioural additionality of the intervention, and unpacked the conditions under which wider effects take place. These findings offered recommendations to improve the effectiveness of the instrument and future policy decisions. Considering the complexity of the Cohesion policy, and the various policy instruments supporting R&D at different levels (EU, national and regional level), additionality appeared as the main impact evaluation challenge. To address this, the evaluation narrowed

8 75% of the total EU spending in the codes analysed

down the analysis to investigate and assess behavioural additionality. Other impact evaluation challenges identified were the lagged effects of investments and the fluidity of enterprises.

Company cases studies (45) were used to addressed lagged effects of the intervention, distinguishing successful from failed projects and categorising short-term, medium-term and longer-term effects in those case studies. Still, the evaluation acknowledged that some wider effects were not possible to be captured especially in ongoing projects where wider socio-economic effects were not clearly identified from the initial stage of the intervention.

Fluidity of enterprises is addressed through the comparative case study analysis, which brought up the corporate strategy as one key factor influencing investment decisions and the conditions making more likely for a company supported to remain in their location.

Overall, this encompassing methodology answered all the evaluation questions and provided insights as to future improvements of the effectiveness of the instrument and to complement what has been found in other counterfactual impact evaluations. The step-by-step process followed to ascertain contribution and establish causality is thorough and carefully implemented.

Part of the limitations of the evaluation relate to the quality of programme data, monitoring data and the lack of counterfactual in several country-level cases.

5.8 Transferability and context

The methods used in this evaluation are underpinned by a significant amount of primary data collected by interviews, steering group meetings and a stakeholder seminar, which can be used seamlessly in different contexts. Programme data and monitoring data are central to the evaluation, but due to its patchy and incomplete nature constituted a limitation for the evaluation. The UK’s long evaluation tradition has led to systematic monitoring and evaluation data collection process, which could facilitate this type of evaluation approaches.

The wide-spread use of counterfactual analysis and econometrics as part of evaluations in the UK could improve the conditions for this type of evaluations.

5.9 Conclusions

This evaluation assesses the effectiveness and behavioural additionality of the support offered by the ERDF to large enterprises (work package 4) under the Cohesion Policy between 2007-2013 in the highest recipients of this funding. Three main activities are funded through this intervention: Advanced support services for firms or groups of firms (Code 05), Investment in firms directly linked to research and innovation (Code 07), and other investment in firms (Code 08) of ERDF. The evaluation analysed eight countries and selected the most relevant operational programme considering the funding allocated to large enterprises under these three codes.

A theory-based approach was selected to explain causality, using a contribution analysis and a comparative case study analysis method. This approach answered the evaluation questions and provided further insights into how to improve the effectiveness of the instrument. This type of evaluations can be transferable to the UK and could complement well econometric analysis of similar types of interventions.

6 Realisation of a Final Impact Assessment Study for Horizon 2020 for the COST Association

6.1 Overview

Key characteristics	
Country	European Union
Institution	COST EU
Type of RD&I	Collaborative R&D
Type of Intervention	Research and innovation grants
Evaluation challenges	Low observability of research impacts; Additionality The challenge is capturing scientific collaboration, knowledge generation and flow beyond publications, and establishing the added value of COST
Evaluation methodology and methods	• Theory-based (Social Network Analysis; bibliometric analysis; semantic analysis; contribution analysis; comparative case studies) • Econometric analysis

6.2 Introduction

This case examines the evaluation of the COST collaborative R&D activities. The evaluation seeks to capture scientific collaboration, knowledge generation and flow beyond publications, to establish the added value of COST. The evaluation addresses two main challenges in assessing the impact of COST, the low observability of the research impacts and the additionality of the funding compared to the Horizon 2020 network.

To address these challenges and assess COST impacts, the evaluation undertakes a Social Network Analysis to capture the effects of COST activities in the evolution of the research networks and a bibliometric and textual analysis of the outputs associated with COST.

6.3 Background description of RD&I intervention

Established in 1971, COST is an intergovernmental programme dedicated to promoting research networks among researchers from COST member countries as well as from partner countries known as Near Neighbour Countries (NNC) and International Partner Countries (IPC). It aims to promote and spread excellence, foster interdisciplinary research for breakthrough science and empower and retain Young Researchers and innovators. COST helps researchers to establish and extend their networks for scientific exchange. The programme's impact reflects on both researchers' individual careers and networking, as well as on research and innovation at-large.

The long-term goal of the programme is to narrow the gap between science, politics, and society in Europe. The COST Association has three strategic priorities in this regard: 1) Promoting and spreading excellence, 2) Foster interdisciplinary research for breakthrough science, and 3) Empowering and retaining Young Researchers.

The main instrument of COST is the funding of COST Actions, which are networks of researchers and innovators. On average, a COST Action has 50 participants in the Management Committee. In general, COST Actions are funded for a period of four years. During the funding period, COST funding can be allocated to organise conferences and workshops and to cover related costs (travelling, accommodation, etc.). Moreover, COST promotes the careers of researchers through the funding of training schools or Short-Term Scientific Missions (STSM). Additionally, COST Action funds can be used to finance targeted activities which serve to disseminate the outcomes and results of the Action.

The programme is thematically open and structured in a bottom-up way. Apart from scientific excellence, applicants have to prove inclusion of those member countries designated as Inclusiveness Target Countries (ITCs), female researchers, and Young Researchers. COST is thus designed to bridge the innovation and participation gaps between well-established member countries on the one hand, and ITCs, IPCs and NNCs, on the other hand, as well as promoting gender equality and career development.

In 2013, COST was re-organised as an international non-profit organisation (the COST Association) under Belgian law (Association internationale sans but lucratif, AISBL), taking over from the European Science Foundation (ESF). It is administered by its Brussels-based bureau (the COST Administration) and is funded from the EU Framework Programme budget of roughly EUR 300 million for the 2014-20 period.

6.4 Challenges measuring impact

The main challenge addressed by this evaluation is to capture scientific collaboration, knowledge generation and flow beyond publications, and to establish the added value of COST. Some outcomes emerging from publications can be easily observed and measured. However, the emergence of new linkages, the evolution of research networks and their patterns of collaboration are more difficult to observe.

To address this challenge this evaluation uses data from networking activities promoted by the programme, such as meetings, training schools, short-term scientific missions (STSMs), and conference grants to carry out a Social Network Analysis of these networks and assess their evolution as well as the change in the outputs being generated.

6.5 Methodology and data sources

6.5.1 Methodology

The evaluation aimed to estimate the added value of the COST activities. It provides an analysis of how the research networks built because of COST activities are structured: considering the geographical scope and professional backgrounds of the researchers that participate in these activities. Additionally, it evaluates the complementarity of the COST network with the Horizon 2020 network. Regarding the scientific impacts of the COST activities, an overview of knowledge creation – in terms of the number of outputs (conferences, workshops, scientific publications) generated by the COST Actions – and of knowledge diffusion – in terms of the size of networks, spin-off Horizon 2020 projects and the number of publications' citations – is performed and reported upon.

The methodological approach includes:

- A **social network analysis** at the participant and regional level, aiming at, first, providing a descriptive analysis of the networks underlying the COST Actions and characterise the structure of these networks, and second, comparing networks constituted by the COST Actions with 'default' networks in science and technology. The analysis focused on two aspects:
- The links between participating researchers to COST activities and the factors that characterise those links: interdisciplinarity, early vs. advanced career stage (professional background), gender, geographical localisation, etc.
- The extent to which the networks generated as a result of the COST activities compare to (public-funded) scientific collaboration networks and what the value-added of COST is regarding the original features of those COST network.
- A **bibliometric and textual analysis** to create an overview of structured and unstructured outputs of the COST Actions, focusing on the scientific and societal breakthroughs supported by the programme. As economic/societal breakthroughs are wider than the impact of COST Actions on patents, only 1 societal/economic breakthrough was selected using the bibliometric data. The other societal/economic breakthrough was selected using natural language processing.
- **Case studies** are presented alongside the bibliometric analysis to help identifying significant breakthroughs among COST Actions.

6.5.2 Data sources

Different types of primary and secondary data sources were collected for this analysis:

- For the social network analysis, data was gathered for each COST Action, instrument, and participant to create a consolidated database. This data englobed the following aspects: characteristics of the Actions (date of start, date of end (if any), scientific domain); characteristics of the instruments (date, type of instrument); and characteristics of the participants (gender, age, title, country of affiliation, NUTS-2 region, type of organisational affiliation).
- Output analysis was performed using COST Action final reports, as well as bibliometric data sourced from the bibliometric database Scopus. Publications were identified using funding information and DOIs. Patent citation data provided by Lens was also used.
- Secondary data sources and documents from the programme were also reviewed and analysed
- In-depth interviews (56) with internal and third-party stakeholders: 18 interviews for the strategic position, 9 interviews for COST Connect, 9 interviews for COST Global Networking, and 20 interviews for the Stewardship approach (including the COST Academy, scientific stewardship, communication stewardship and the Innovators Grant)

6.6 Main findings

The evaluation found that within the COST network, knowledge and ideas spread efficiently and quickly thanks to a flat network structure resembling a ‘small-world’, where connectivity between participants is high. Additionally, COST offers low entry barriers to Young Researchers and researchers from less research-intensive areas, which is seen as a defining feature. Furthermore, one-quarter of COST publications’ authors can be considered as Young Researchers.

Thus, COST is perceived to be the primary networking tool in the European research and innovation landscape, spanning disciplines, countries, career stages and different types of actors. Equally important is that participation in COST Actions increases the chance for success of applications to other European programmes, making COST a pre-portal to other European funding instruments.

Career stage still drives a significant part of the connections made on Actions, but less so for connections made during individual meetings. About one out of two connections in the COST network is between a man and a woman, a higher share than in both FP7 and FP6. Also, when taking titles into account, male participants are not better connected to any other participant or more central to a network.

The COST programme enables interactions between Inclusiveness Target Countries (ITC) and other COST countries, as seen by the high share of connections between participants from ITC and non-ITC. Within the EU, geographic separation and language barrier do not significantly impede cross-regional connections.

More than half of the COST Actions are interdisciplinary (on average 5.8 different disciplines), and the COST programme enables more interdisciplinary Actions for Humanities and Social Sciences fields than Horizon 2020. In addition, connections in COST activities tend to be slightly hierarchical (between participants bearing the same title) whereas the Horizon 2020 network is governed by a negative hierarchy (disassortative) effect.

On average, Actions contributed to publishing a little bit more than 30 publications, based on self-reported data, and around 53 publications on average based on Scopus. Those spin-off publications generated over 200,000 citations and 89% of them have been cited at least once. COST publications have a collaborative nature (on average, 6.7 authors are listed on a COST publication).

6.7 Analysis of methodological suitability and effectiveness

As the programme evaluated is intended to foster collaboration and network creation among researchers, social network analysis is a suitable method to assess the extent to which the programme has contributed to that goal. Furthermore, the evaluation performs an output analysis to identify the main scientific breakthroughs and societal impacts of the programme, by using bibliometric analysis, patent analysis and natural language modelling. These methods are appropriate as they provide an overview of how the programme has delivered against its original intentions.

When performing social network analysis to assess RDI programmes, a common challenge is building a correlational data set that enables to visualize and quantify the structure of nodes and links resulting from a funding programme. This is usually performed, for example, using bibliometric data focusing on co-authorships, when available. However, scientific collaboration is not limited to publications but also encompasses a broad range of activities during the research process. This challenge is addressed here by using data from networking activities promoted by the programme such as meetings, training schools, short-term scientific missions (STSMs), and conference grants.

The methodological approach is also robust regarding:

- The combination of methods, using bibliometric data, interviews and case study analyses to triangulate the impact of COST Action program
- The comparison of the results to findings and indicators collected from desk and literature research
- The use of social network analysis indicators to perform econometric analysis
- The fact that the revision of self-reported output from the reports was complemented with bibliometric data analysis

An additional challenge is the complexity of the COST programme, which includes a wide range of activities and instruments. The evaluation addresses this by combining different qualitative and quantitative methods. While the approach and methods employed to assess this large programme are comprehensive, they are also resource-intensive and time-consuming given the multitude of data collection and analysis techniques required. Gathering, processing, and interpreting data from various sources, especially when dealing with programmes that involve a large number of instruments and participants such as COST, can be logistically challenging and costly, potentially limiting the scalability of the evaluation. To address this, the evaluation uses existing secondary data sources, especially monitoring data of the programme to perform the social network analysis.

Also, in the analysis of the selected scientific breakthroughs, there are several considerations that are to be taken into account for a nuanced interpretation of the programme's effects. First, in some cases the breakthroughs are assumed as societal impacts of the programme. However, the data analysed in each case is mostly related to innovation and scientific indicators (citation, patents, publications) that do not necessarily translate into societal impact. Furthermore, the uptake of the breakthroughs by industry is supported by citation data, which only provides a partial as to how innovations are diffused among firms. Finally, the study (and readers) would have benefit from a better understanding as to how the cases were selected.

6.8 Transferability and context

The methods employed in this study can be applied to the context of the UK. Like in the case of COST, multiple R&D programmes and instruments could have supported the pre-existing relationships among the networks. However, the focus on the characteristics of the network supported by a specific programme, and the extent to which those characteristics are different from 'business as usual' could offer valuable insights into the effectiveness of programmes aimed at widening participation. To isolate the effects of the programme under evaluation it is useful to employ counterfactual analysis or performing case studies that delve deeper into the causal mechanisms that explain the effects of the programme.

Furthermore, for the UK context there are several data sources that can be exploited to test the characteristics and composition of the network. (e.g. Researchfish, Gateway to Research data).

6.9 Conclusions

This evaluation showcases a relevant combination of social network analysis and econometric analysis to assess the collaboration, scientific and societal impacts of R&D programmes. This methods can be relevant when determining the effects of collaborative R&D programmes that intend to build interdisciplinary networks of researchers that enable knowledge production and diffusion, while providing a networking platform that integrates young and remote researchers from less research-intensive fields and/or regions.

A similar study can be performed in the context of the UK as long as it evaluated collaborative R&D programmes for which there is network data available, i.e. data that relates one node with another, whether researchers or research institutions. This data can be collected via the revision of secondary sources, such as programme monitoring reports or primary sources such as self-reported data on collaboration, collected via surveys and interviews with beneficiaries.

In this case, the unit of analysis is the whole network, so conclusions must be limited to the structural conditions and composition features of these. Furthermore, the instruments, projects or grants that relate each node must provide an insight into the scientific outputs, so that further bibliometric analysis can be performed.

7 Evaluation of investments in Research and Technological Development (RTD) infrastructures and activities supported by the European Regional Development Funds (ERDF) in the period 2007-2013

7.1 Overview

Key characteristics	
Country	European Union
Institution	European Commission, Directorate-General Regional and Urban Policy (DG REGIO)
Type of RD&I	Applied research and innovation
Type of Intervention	Research and innovation grants; Research infrastructure
Evaluation challenges	Attribution; Contribution The challenge is the identify the impacts that can be attributed to the investments in RTD infrastructure and associated R&D activities
Evaluation methodology and methods	• Theory-based (contribution analysis; cases studies; literature review; stakeholder interviews) • Econometrics (multivariate regression)

7.2 Introduction

The case analyses the Evaluation of investments in Research and Technological Development (RTD) infrastructures and activities supported by the European Regional Development Funds (ERDF) between 2007-2013.

To address the evaluation challenge of attribution of impacts to the investments on RTD infrastructure and activities, this evaluation adopted theory-based mixed methods approach. Contribution Analysis (CA) was employed to determine the causal relationships between specific interventions and observed changes. The CA draws evidence from case studies and quantitative methods employed in the CA. The quantitative component of the CA involved conducting multivariate regression analysis to test hypotheses regarding how RD&I support contributed to regional economic and innovation outcomes (e.g. patenting activity, private sector R&D expenditure).

7.3 Background description of RD&I intervention

Over EUR 17 billion of European Regional Development Funds (ERDF) was invested through 215 operational programmes (OPs) in projects to support Research and Technological Development (RTD) activities in research centres[9] and to develop RTD infrastructure and centres of competence in specific technology areas. The rationale and need for the RTD investments were to overcome two main barriers in the research and innovation ecosystem. Firstly, the lack of a critical mass of infrastructure capabilities and research capacities to enable the production of top-class research. Secondly, the need to increase the industrial relevance of the regional science base by linking existing or emerging research centres of scientific excellence to areas of industrial strength.

The RDI investments were closely linked to the objective of fostering regional research and innovation ecosystems and creating a larger pool of high-quality research talent in Europe. The role of research infrastructures was seen as key to enhancing national and regional RTD capacities and fostering regional ecosystems to drive economic growth.

The ex-post evaluation of investment in RTD infrastructures and activities supported by the EDRF covers the programming period 2007-2013. It focussed on 53 OPs selected by the European Commission out of the total of 215 OPs funded by the ERDF. Selected OPs covered 18 of the 28 Member States involved and represented substantial RTD investment of EUR 14.64 billion (EUR 11 billion for research infrastructure support and EUR 5.8 billion for research activities support.

7.4 Challenges measuring outcomes and impact

The main methodological challenge faced by the evaluation was the attribution of impacts to the investments on RTD infrastructure and activities. According to the Evaluation Officer of the Department for Regional and Urban Policy DG REGIO in the European Commission, there are two intertwined challenges when investing in infrastructure and related activities. The lagged effects and the attribution of impacts to those investments. During the first years after the initial investment R&D activity spikes and so do short term outcomes, which tend to fade over time. However, the main purpose of the evaluation is to see longer term effects and to assess to

[9]RTD projects included support for scientific R&D activities; collaborative research activities; support for the internationalisation of research activities; development of researchers and other personnel involved in R&D activities; support for technology-transfer activities; and the valorisation of research results.

what extent infrastructure investments translate into increasing possibilities for regions, building on their strengths.

What is expected by this intervention is to facilitate convergence of lagged regions, and this is complex and long-term process that can partly be assessed in the short and midterm. This the main reason to adopt a Contribution Analysis to discern pathways to impact and causality chains. Nonetheless, identifying causality and attribution of impacts pose significant challenges due to the heterogeneity of the regions and their Operational Programmes.

A salient aspect of this evaluation highlighted by the Evaluation Officer, was its mixed method approach underpinned by the Contribution Analysis. This Theory-based method guided the quantitative analysis, providing insights as to what effects and relationships to test in the multivariate analysis.

7.5 Methodology and data sources

7.5.1 Methodology

The evaluation was guided by a set of evaluation questions concerning the relevance, coherence, effectiveness, efficiency, added value and sustainability of ERDF support to RTD. A theory-based impact evaluation approach, combining qualitative and quantitative methods was applied at different levels of analysis (at country, Operational Programme, instrument and project level) to provide comprehensive answers to the evaluation questions.

For the assessment of the degree of effectiveness of selected policy instruments, the study followed the approach of Contribution Analysis (CA), a specific form of theory-based evaluation that focuses on causal relationships and explanatory conclusions between observed changes and specific interventions. This approach enabled both the outputs, outcomes and impacts to be assessed to understand what happened (i.e., the direct effects of the ERDF support for RTD) and also the supporting factors, pre-conditions and underlying assumptions to the achievement of causal packages to explain why and how the observed effects had occurred. The principles of CA were used to guide the collection and processing of evidence into seven case studies at Member-State level addressing nine OPs (including in-depth analyses of three selected policy instruments per case study). A cross-case analysis was conducted to understand how policy instruments work collectively as part of a broader ‘causal package’. The case studies were informed by a literature review to develop theories of change regarding ERDF support for RTD investments and their expected results.

Due to the lack of available data, it was not possible to design a control group of non-treated European (EU) regions, thus a pure counterfactual method, such as propensity score matching, regression discontinuity design or difference in difference, could not be carried out to provide an economic assessment of the size of the observed effects of ERDF support to RTD. An economic analysis employing multivariate regression analysis was selected as it allowed the testing of a set of hypotheses to understand the contribution of the various types of instruments to a set of regional outcomes. A strength of this method is that it permits one to isolate the contribution of the ERDF instruments to specific regional outcomes from the other potential factors (e.g., regional socio-economic conditions, other R&D policies beyond the ERDF instruments, etc.) influencing those outcomes. The economic analysis was not meant to provide an estimation of the size of the observed effects but intended to offer a complement to the qualitative analysis of the case studies.

7.5.2 Data sources

The analysis utilised monitoring data collected on results and outputs[10] collected by Managing Authorities. It also draws on a database aggregated at the regional NUTS2 level from several sources, including CORDIS, Eurostat, Patstat, and Web of Science to support testing of the hypotheses. In addition to statistical indicators and expenditure data, sources of evidence included more than 200 direct interviews with stakeholders and project beneficiaries collected within the framework of the case studies.

7.6 Main findings

The evaluation reported these main findings related to the effectiveness of the ERDF support for RTD investment:

R&D personnel and researchers. ERDF support for RTD investment was found to have a positive and significant contribution to the observed improvement of R&D capacities in targeted EU regions. The multivariate regression analysis showed ERDF investments in research infrastructures and individual R&D projects in higher education institutes contributed to an average 40% increase in the number of R&D personnel and researchers at the regional level. Case studies showed that ERDF investments enabled the modernisation of education facilities and research infrastructures, which improved the conditions and teaching environments to attract new R&D personnel, students and researchers. Individual R&D projects in HEIs were shown to positively increase the number of R&D personnel and researchers at the regional level.

Scientific output. ERDF support for RTD investment was also found to have a positive and statistically significant relationship between ERDF support and the growth rate in the number of scientific publications. The multivariate regression analysis indicated scientific publications almost doubled in volume across the EU as a whole in the period 2007-2017. Further, the higher the ERDF expenditure, the higher the growth rate in the number of scientific publications. However, the analysis found no statistically significant relationship between ERDF investments and the quality of scientific production (as evidenced by the growth rate in the number of regional scientific publications in the top 25% of most cited publications).

[10] Monitoring data included data on number of: RTD projects, cooperation projects enterprises, research jobs created, new researchers in supported entities, researchers working in improved research infrastructure facilities, enterprises cooperating with research institutions and value of private investments matching public support in innovation or R&D projects.

Findings from the cases studies highlighted individual R&D projects played a significant role in contributing to the increase of scientific output.

Technological development and innovation. More limited, however, was the capacity of ERDF funded projects to generate economic benefits from the commercial exploitation of R&D results:

- No statistically significant relationships are found in the econometric analysis between ERDF support and the growth rate of technological outputs (as measured by patenting and public-private co-publications). Evidence from case studies revealed reasons for the limited uptake of research results, which included sustainability issues where science-industry collaborations did not lead to systematic follow-up projects, or the results were simply not relevant for the industry. Furthermore, the synergy and complementarity with existing funding sources were not always well exploited. Another explanation put forward for the lack of relationship between ERDF support and technological output is that of time lag, where more time is needed for research activities to generate a technological output. Moreover, ERDF support alone may not have been enough as other supporting factors were probably necessary for a noticeable shift in technological capacity.
- The multivariate regression analysis also tested whether ERDF policy instruments showed an indirect link with the growth rate in the number of patents through the R&D expenditure in the business sector. No significant correlation was found between ERDF policy instruments and increases in business R&D expenditure. However, it was noted that other factors can play a more direct and significant role in triggering business R&D expenditure. Further, the role of the economic downturn of 2008 which resulted in severe cuts in public funding for education and research should not be overlooked.
- The multivariate regression analysis also tested correlations between the ERDF and ‘softer’ innovation outcomes. The analysis suggests that ERDF support positively and significantly correlates with the growth rate of EU trademark applications. Interestingly, in addition to ERDF support for RTD investments, additional drivers were the level of ERDF expenditure on business support and R&D expenditure in the business sector (as a percentage of GDP). Although not conclusive evidence in itself, this finding seems to confirm the impression that the role of the ERDF is more related to behavioural changes and less to technologically intensive innovation. The ERDF role in triggering behavioural changes among its beneficiaries is supported by evidence from the case studies.

7.7 Analysis of methodological suitability and effectiveness

The combination of different methods was instrumental in collecting a comprehensive set of evidence to attribute impact to ERDF investment in RTD activities in different EU regions. The multivariate regression analysis in combination with the understanding of the causal mechanisms and contextual factors of ERDF interventions explored in the case studies was a key strength. Evidence from the case studies complemented and corroborated the findings from the multivariate regression analysis.

Despite the huge amount of data and evidence collected during the evaluation activities, some limitations remain:

- The monitoring indicators collected by Managing Authorities had limitations both in terms of coverage of the actions supported and in terms of quality (i.e., inconsistencies, gaps and focus on outputs and results and not impacts).
- Evidence on outcome and impact at the level of individual policy instruments is mostly qualitative and based on stakeholder perception.
- The poor quality of some beneficiary data (lack of partner beneficiaries for collaborative projects, duplication or missing data) prevented a more in-depth analysis, and it also prevented systematic matching with external databases (e.g., CORDIS database).

According to the Evaluation Officer during interview, at the time of the evaluation, this methodological approach was used as a pilot. Using the contribution analysis represented a challenge and learning opportunity for the Evaluation team at DG REGIO. Thus, the expertise of the evaluation team was central to support the development of such skills in the Evaluation Office at DG REGIO. As a result of this evaluation and the learnings of this pilot, methodological improvements were introduced to the data collection at Operational Programme level, which are expected to enable the use of counterfactuals in evaluation. Considering the 7-year cycle of the policy, these improvements did not feed the second cycle of the programme but has been incorporated for the third cycle. Likewise, the officer confirmed that the contribution analysis method has become a pervasive evaluation method used at DG REGIO.

The evaluation also highlighted recommendations for future methodological improvements:

- A longer-term analysis or a back-to-back approach of evaluations of different programming periods would better account for longer-term effects, including the link between infrastructure investment and innovation. It would also help to assess whether the stand-alone approach to investments has effectively been overcome, absorbing research infrastructure into the ecosystem approach.
- Incorporating agent-based modelling approaches or combining a system-dynamics approach with theory-based evaluation techniques. This is particularly important in the field of RTD support, as the unit of analysis typically changes throughout the evaluation exercise, starting from individual operations and ending in innovation ecosystems towards the end of the evaluation.

7.8 Transferability and context

The evaluation relied on several secondary data sources and any replication of the study would require availability of similar sources. This should not be a problem in the UK context providing equivalent programme monitoring data is available. The secondary data sources used for this particular evaluation (Patstat and Web of Science) is international in scope. However, it should be noted that the evaluation does not provide methodological details on how regional data matching (e.g. patent and publication bibliometric data) was performed. This can be complex and time consuming, which could in part be automated by quantitative text analysis.

7.9 Conclusions

This evaluation provides a methodologically rigorous analysis of the extent to which EDRF funding is contributing to strengthening existing regional RTD capacities. In the absence of a counterfactual, it employs a relatively simple multivariate model to test a set of hypotheses that permitted being able to isolate the contribution of ERDF instruments from other potential contributing factors. The use of case studies was key to expand and corroborate the findings from the multivariate regression analysis. Data limitations posed a number of evaluation challenges; however, these challenges were effectively mitigated by employing the triangulation of data sources to draw robust conclusions.

8 A microeconomic perspective on the impact of the Fraunhofer-Gesellschaft

8.1 Overview

Key characteristics	
Country	Germany
Institution	Fraunhofer-Gesellschaft
Type of RD&I	Applied research and innovation
Type of Intervention	Research and innovation grants
Evaluation challenges	Additionality; contribution The key evaluation challenges are 1) Determining the causal relationship between a company's collaboration with the Fraunhofer-Gesellschaft research organisation and its economic performance. 2) Obtaining insights into causal mechanisms driving the observed causal relationships.
Evaluation methodology and methods	• Quasi-experimental (matching)

8.2 Introduction

This evaluation investigates the impact of collaboration with the Fraunhofer-Gesellschaft (FG) research organisation on the organisation's industry partners. This includes the company-level effects on various economic and innovation outcomes.

The evaluation utilises a quasi-experimental research design involving multivariate regression models combined with matching techniques. This approach allows evaluators to address the challenges of determining programme additionality, as well as providing some insights into programme contribution.

8.3 Background description of RD&I intervention

FG is a German research organisation with 76 research institutes located across Germany, each focused on a specific field of applied science. It is the largest organization for applied research and development services in Europe, employing around 30,800 employees (predominantly scientists and engineers) working with an annual research budget of around €3.0 billion, €2.6 billion of which is contract research. Around two thirds of Fraunhofer's contract research revenue is generated from industry contracts and publicly funded research projects. The remaining third is largely comprised of core funding from the German federal and state governments, which is used by the Fraunhofer institutes to develop solutions to social and industrial challenges. Much of Fraunhofer's contract research is conducted for small and medium sized enterprises (SMEs) and German "Mittelstand" (mid-tier-business) companies, the latter being defined as those with less than 5,000 employees.

In addition to contract research, FG also participates in publicly funded joint projects with industry. In most cases, multiple industry partners participate in each project and other public research organisations (PROs) may also participate. These projects usually involve pre-competitive research, rather than research oriented toward product development. Thus, from an industry perspective, publicly funded joint projects tend to address mid- to long-term challenges. Collaborating industry partners are largely SMEs.

8.4 Challenges measuring outcomes and impact

This evaluation addresses the following key evaluation challenges:

Additionality: To estimate what effect the programme has had in excess of what would have occurred in its absence, the evaluation applied a quasi-experimental research design that utilises matching procedures. As noted by Professor Torben Schubert, a member of the study team and the Deputy Director of the FG Competence Centre for Innovation and Knowledge Economy[11] which carried out this evaluation, it is vital to employ a quasi-experimental approach to evaluate the impact of FG collaboration on industry partners because companies become FG industry partners through a dual selection process. First, innovative, hi-tech companies are more likely to select into approaching FG for a collaboration. Second, Fraunhofer institutes agree to enter into industry partnerships which are deemed to involve high quality research.[12]

Contribution: By considering the impact of FG collaboration on a range of company economic and innovation outcomes, the evaluation findings can provide nuanced insights into the specific effects that FG collaboration has on industry partners. This provides some insights into the causal mechanisms underlying the programme's impact (e.g. increased revenue, increased productivity, increased patentable innovations). An understanding of causal mechanisms is key to addressing the evaluation challenge of demonstrating contribution, as it allows evaluators to link a programme's activities to its impact.

11 The FG Competence Centre for Innovation and Knowledge Economy is the organisation's in-house innovation economics research centre.

12 Interview with Professor Torben Schubert, 16/11/2023.

8.5 Methodology and data sources

8.5.1 Methodology

This evaluation aims to measure the economic impact of FG collaboration on industry partners. It investigates the relationship between FG collaboration and a range of company-level economic outcomes: turnover, earnings before interest and taxes (EBIT), earnings before interest, taxes, depreciation, and amortization (EBITDA), operating revenue, return on equity, employment and value-added per employee. It also investigates the relationship between FG collaboration and several company-level innovation outcomes: product innovation, new product turnover, service innovation, new service turnover and patenting activity.

The evaluation employs a quasi-experimental research design, using multivariate regression models combined with matching techniques. The analysis involved a matched-pair approach in which each FG industry partner was assigned a statistical “twin”. These twins were determined by their similarity to the cooperating firm in terms of number of employees, industrial sub-sector, R&D status and product complexity. This matching technique helps correct for observable differences between companies that have collaborated with FG and comparator companies which have not, thus mimicking a randomised experiment and allowing for a credible comparison between the two groups.

8.5.2 Data sources

The evaluation uses data from two original company-level datasets constructed for the purpose of this study. FG mostly cooperates with industrial firms, and this is reflected in the companies included in these datasets.

- One dataset is based on the ORBIS database published by Bureau van Dijk, a publisher of business information specialised in private company data. The ORBIS database contains basic company data (e.g. economic sector and ownership structure) on industrial German companies for the most recent available year. To construct the dataset used for the evaluation, German industrial companies in the ORBIS database were matched to four other datasets: 1) Data from the German Public Funding Catalogue was used to provide company-level information on participation in collaborative research with FG and other PROs; 2) Data from FG’s administrative records was used to provide company-level information on FG contracted research projects. This information is only available for FG, and not for other PROs, but it allowed the researchers to separate contract research from joint research projects in the analyses; 3) Data from Bureau van Dijk's Amadeus database was used to provide company-level information on a range of financial indicators; 4) Data from the European Patent Office's PATSTAT database was used to provide company-level information on companies’ transnational and German patent applications. The resulting dataset contains ~34,000 company level observations.
- The second dataset is based on the representative German Manufacturing Survey (GMS), which is the German component of the European Manufacturing Survey.

Although it only covers manufacturing firms and thus has a smaller sample than the dataset described above, it contains information on a wider range of variables related to firm behaviour. For this evaluation, GMS data from the years 2012, 2015 and 2018 was pooled to construct a dataset containing ~4,000 company level observations. This dataset was matched to the same four datasets described above.

- The Orbis and GMS data was matched to that from the four other datasets using a string-matching algorithm based on the Levenshtein distance at the level of company names. For the Amadeus data, companies' VAT number was additionally used for matching. This algorithmic matching approach greatly reduces the effort and cost required to construct such datasets.

8.6 Main findings

Collaboration with FG (as well as universities and other PROs) was found to have positive impacts on several aspects of economic performance among partner companies (labor productivity, EBIT, turnover and revenue per employee). Positive economic impacts were found for both samples (that based on the GMS and that based on the Orbis database). This relationship also holds true for SMES and large firms, but SMEs were found to benefit relatively more from collaboration with FG. It was also found that FG contract research collaboration had more direct impacts on companies' economic performance than joint research projects (which are more long-term oriented).

8.7 Analysis of methodological suitability and effectiveness

The methodology employed in this evaluation has a number of strengths.

- By combining six sources of data, the evaluation team was able to construct rich datasets that can provide nuanced insights into the economic effects of FG collaboration on industry partners. In contrast to relying on a small number of generic economic and innovation indicators, this provides some insight into how FG collaboration leads to improved economic performance among companies. Put differently, it allows for greater understanding of the causal mechanisms underlying the economic benefits of the programme (e.g. increased revenue, increased productivity, increased patentable innovations).
- In the absence of a comprehensive dataset on all relevant German companies that includes data on their economic and innovation characteristics, this evaluation conducted complementary analyses on two different datasets and compared the research results for triangulation. This is a powerful approach to demonstrating the robustness and internal validity of evaluation findings. As noted by Professor Schubert, it is well-established that the application of quasi-experimental econometric techniques to complex contexts yields imprecise causal estimates. For this reason, the evaluation team conducted multiple analyses in order to validate the research findings. In doing so, they sought to establish whether the various related analyses all point in the same

direction, yielding positive and sizeable results. Where this was found to be the case, it was interpreted as giving high confidence in the findings.[13]

This methodology also has several limitations.

- The report does not detail the rationale for the choices made in the matching procedure. It is therefore unclear whether the four variables used to match FG collaborators with their statistical twins (number of employees, industrial sub-sector, R&D status and product complexity) are sufficient to create a credible counterfactual.
- The matched-pair approach mitigates the evaluation challenge of establishing additionality by matching FG collaborators to non-collaborators on several observable characteristics. However, it does not account for unobservable differences between these two groups of companies. In contrast, quasi-experimental approaches such as difference-in-differences and regression discontinuity design can help to account for unobservable differences. However, it is likely that these approaches were not possible in this evaluation due to data limitations and the nature of the programme design.
- Due to the unavailability of medium-to-long term time series data, it was not possible to determine whether the finding that joint research projects (as compared to contract research collaborations) have fewer economic benefits is conclusive or whether it is due to lagged effects that cannot be measured with the data available (given that these projects are more long-term oriented).

8.8 Transferability and context

The key advantages of this evaluation approach can be transferred to the evaluation of RD&I programmes in the UK. The use of matching algorithms to merge multiple sources of data is an efficient and cost-effective way to construct rich datasets for impact evaluation. It is worth noting here that linking multiple data sources for individual companies in the UK context is broadly less challenging than it is in the German context, due to the more widespread use of unique company identifying numbers.[14] Such datasets enable evaluators to more easily and rigorously test for the causal mechanisms hypothesised to drive programme impacts. The ability to identify causal mechanisms helps address the evaluation challenge of contribution and makes evaluation results more actionable for policymakers.

On a related point, conducting complementary analyses on multiple datasets is an effective approach to demonstrating the robustness and internal validity of evaluation findings, particularly in the absence of a comprehensive dataset on programme beneficiaries. While this is not always feasible, it may be more practical when off-the-shelf data is available and/ or

13 Interview with Professor Torben Schubert, 16/11/2023.

14 Interview with Professor Torben Schubert, 16/11/2023.

where matching algorithms can be used to construct datasets in an efficient and cost-effective manner.

8.9 Conclusions

This evaluation constructs two rich company-level datasets which provide nuanced insights into the economic impact of FG collaboration on industry partners. It finds that collaboration with FG (as well as universities and other PROs) has positive impacts on several aspects of economic performance among partner companies. The key advantages of this evaluation approach can be transferred to the evaluation of RD&I programmes in the UK. Most notably, the use of matching algorithms to merge multiple sources of data is an efficient and cost-effective way to construct rich datasets for impact evaluation. Such datasets can help to address the evaluation challenge of contribution, making evaluation results more actionable for policymakers.

9 The macroeconomic effects of the Fraunhofer-Gesellschaft

9.1 Overview

Key characteristics	
Country	Germany
Institution	Fraunhofer-Gesellschaft
Type of RD&I	Applied research and innovation
Type of Intervention	Research and innovation grants
Evaluation challenges	Additionality The key evaluation challenge is determining the causal relationship between the presence of a Fraunhofer-Gesellschaft research institute in a German region and the region's economic performance
Evaluation methodology and methods	• Quasi-experimental (instrumental variables; matching) • Regression analysis

9.2 Introduction

This evaluation investigates the impact of the Fraunhofer-Gesellschaft (FG) research organisation on the German economy over the period 2003-2017, focusing specifically on the impact of Fraunhofer institutes on regional GDP and regional patenting activity.

The evaluation utilises static panel data (SPD) regression analysis alongside two quasi-experimental research designs: instrumental variables and matching. This approach allows evaluators to address the evaluation challenge of determining programme additionality.

9.3 Background description of RD&I intervention

Fraunhofer-Gesellschaft (Fraunhofer Society) is a German research organisation with 76 research institutes located across Germany, each focused on a specific field of applied science. It is the largest organization for applied research and development services in Europe, employing around 30,800 employees (predominantly scientists and engineers) working with an annual research budget of around €3.0 billion, €2.6 billion of which is contract research.

Around two thirds of Fraunhofer's contract research revenue is generated from industry contracts and publicly funded research projects. The remaining third is largely comprised of core funding from the German federal and state governments, which is used by the Fraunhofer institutes to develop solutions to social and industrial challenges.

9.4 Challenges measuring outcomes and impact

This evaluation investigates the macroeconomic impact of the FG research organisation by comparing the economic performance of regions that contain Fraunhofer institutes with that of regions that do not. Given the systematic differences between Fraunhofer and non-Fraunhofer regions, establishing programme additionality is a significant challenge in this evaluation. As the evaluation report points out, regions containing Fraunhofer institutes have substantially higher GDP per capita and number of patents than the national average. They also have a slightly higher than average number of high-tech workers. To address this challenge, the report employs two different quasi-experimental research designs and compares the results for triangulation.

9.5 Methodology and data sources

9.5.1 Methodology

This evaluation investigates the impact of Fraunhofer-Gesellschaft on the German economy over the period 2003-2017, focusing specifically on the impact of Fraunhofer institutes on regional GDP and regional patenting activity.

To estimate the local economic impacts of Fraunhofer institutes, the evaluation employs SPD regression analysis alongside two quasi-experimental research designs: instrumental variables and matching. These analyses compare the economic trajectories of German regions that have Fraunhofer-Gesellschaft institutes to those which do not. This geographical variation in the location of Fraunhofer-Gesellschaft's institutes underpins the causal inference strategy used in the study. If it is the case that these institutes boost local GDP and patenting activity, the regions where they are located should – on average and over time – outperform similar regions where they are not located, after accounting for confounding factors.

The challenge of establishing additionality looms large in this evaluation. This is because the choice of where to locate Fraunhofer institutes is highly likely to be strategic. For example, if Fraunhofer institutes choose to locate in regions that are a priori economically stronger, any observed associations between Fraunhofer presence and economic outcomes may be partly or even completely spurious. As mentioned above, regions containing Fraunhofer institutes have substantially higher GDP per capita and number of patents than the national average. They also have a slightly higher than average number of high-tech workers.

The use of panel data regression analysis with regional fixed effects in this evaluation helps to overcome these evaluation challenges. The inclusion of regional fixed effects in the regression models allows researchers to control for unobserved time-invariant characteristics that are specific to each region, such as historical development trajectories or geographical features. If one region is inherently more industrious for some historical reason, this will not affect the model estimates because that inherent industriousness is consistent over time.

To further address the evaluation challenge of additionality, the analysis also estimates SPD models with entropy balancing. This matching technique helps to correct for systematic differences between Fraunhofer and non-Fraunhofer regions. Entropy balancing calculates regression weights to balance the samples of Fraunhofer and non-Fraunhofer regions on observable variables that may muddy causal inference, such as the share of agricultural employment or number of high-tech workers in a region.

In addition, the evaluation estimates models that employ dynamic panel data (DPD) modelling techniques using the Arellano-Bond estimator. If locational choice depends on unobserved time-variant variables, entropy balancing will not eliminate all the bias resulting from locational choice. DPD modelling using the Arellano-Bond estimator incorporates lagged dependent variables as instrumental variables into the models to address this issue. The evaluation uses the Hansen test to test for the validity of the instruments used.

The evaluation rigorously tests the robustness of its findings by implementing several statistical models with various specifications as well as employing several econometric techniques. Where findings are consistent across the different model specifications and techniques, the evaluation can make highly plausible causal claims. As noted by Professor Torben Schubert, the author of the study and Deputy Director of the FG Competence Centre for Innovation and Knowledge Economy,[15] it is well-established that the application of quasi-experimental econometric techniques to complex contexts yields imprecise causal estimates. For this reason, the evaluation involved multiple parallel analyses in order to validate the research findings. In doing so, it sought to establish whether the various related analyses all point in the same direction, yielding positive and sizeable results. Where this was found to be the case, it was interpreted as giving high confidence in the findings.[16]

9.5.2 Data sources

The analysis draws on an original panel dataset constructed by merging Fraunhofer-Gesellschaft administrative data with regional economic data provided by the German Federal Statistical Office (DESTATIS). Fraunhofer-Gesellschaft administrative data on the location, budget and employee headcount for each of its institutes was aggregated to the “NUTS 3” regional level and matched with DESTATIS data provided at the same geographical level. The DESTATIS statistics provide “NUTS 3” level data on population, GDP per capita, share of agricultural employment and number of high-tech workers. It appears that the DESTATIS data is also the source of regional data on patenting activity, although this is not explicitly stated in the published report. In total, the dataset covers 400 “NUTS 3” regions over a fifteen-year

15 The FG Competence Centre for Innovation and Knowledge Economy is the organisation’s in-house innovation economics research centre.

16 Interview with Professor Torben Schubert, 16/11/2023.

period, resulting in a total of 6,000 time-year observations. This dataset allows for retrospective analysis over a substantial period of time, thus addressing the evaluation challenge of lagged effects.

9.6 Main findings

The evaluation finds that Fraunhofer institutes have a substantial impact on regional GDP. In the baseline model,[17] a €1 increase in the budget of a Fraunhofer institute is associated with a €21 increase in regional GDP per capita. In the other models, the estimated increase in regional GDP per capita for a €1 increase in the budget of a Fraunhofer institute ranges from €15 - €29. The finding that Fraunhofer institutes have a large positive impact on regional GDP is robust to a number of model specifications and to various econometric techniques.

The evaluation also finds evidence that the effect of Fraunhofer institutes on regional GDP increases substantially over time, with the GDP multiplier in 2015-2017 around 23% larger than in the period up to 2014. However, robustness tests for this finding are not presented in the evaluation report.

For patenting activity, the baseline model[18] finds that a €10 million increase in funding to Fraunhofer institutes is associated with approximately 12 regional patent applications (after accounting for the average value of patents in regions hosting Fraunhofer activities). However, this finding is not statistically significant in the DPD model using the Arellano-Bond estimator. In addition, it was only possible to implement restricted versions of the dynamic ordinary least squares analysis for patenting activity, due to a high level of correlation between some of the variables included in these models. Thus, the positive finding related to patenting activity should be considered suggestive.

9.7 Analysis of methodological suitability and effectiveness

The methodology employed in this evaluation has a number of strengths.

- The evaluation can make highly plausible causal claims about the macroeconomic impact of Fraunhofer institutes due to its use of panel data analysis and quasi-experimental methods, combined with rigorous robustness tests. The findings provide strong evidence of additionality: that these economic benefits would not have materialised anyway in the absence of the Fraunhofer institutes.
- The findings also provide some suggestive evidence for attribution/contribution: that these economic benefits can be attributed to the activity of the Fraunhofer institutes, rather than other elements of the innovation system in the institutes' local regions. To

17 SPD model with year and regional fixed effects.

18 SPD model with year and regional fixed effects.

further strengthen the analysis, one possibility is to include additional control variables in the models that measure different aspects of the innovation ecosystem in German regions. The analyses presented in the evaluation report already control for the number of high-tech workers in a region, but this could be supplemented with other key measures – for example, data on regional levels of private sector R&D investment (subject to data availability).

- As noted by Professor Schubert, the approach to economic impact analysis employed here offers an important advantage over the use of input-output or Keynesian multipliers to measure macroeconomic impacts. This is because the aforementioned multiplier approaches treat organisations as generic economic actors merely injecting demand into the economy. They cannot account for how RD&I can have an outsized economic impact through knowledge production and innovation.

This methodology also has several limitations.

- The analysis does not establish the causal mechanism underlying the macroeconomic impact of the Fraunhofer institutes, that is – it does not provide insights into how the institutes increase regional GDP. The ambiguous findings related to patenting activity mean that we do not have compelling evidence that the observed impact is driven by increased innovation in these regions. To strengthen this aspect of the analysis, the evaluation could attempt to explore the effect of Fraunhofer institutes on innovation outcomes other than patenting activity. This would allow the evaluation to test whether the institutes' macroeconomic impact is indeed related to their research outputs.
- A key limitation of the approach employed by this evaluation is the focus on regional economic impacts, given that the economic benefits of many RD&I initiatives are not limited to a programme's local region. In the case of the Fraunhofer institutes, it may be that particular institutes have been located in economic hubs relevant to the specific field of applied science that they specialise. In such a case, the focus on local economic impacts would be appropriate.

9.8 Transferability and context

The methodology employed in this evaluation has limited transferability. It is transferable to the evaluation of RD&I programmes in the UK where the following conditions are met:

- Such an approach can be used to evaluate the impact of programmes where there is significant variation in the programme's geographic footprint. Ideally, this variation would include no programme presence in some geographic areas. In addition, the socioeconomic composition of geographic areas where there is significant programme presence and those where there is no-to-low programme presence should not be very large. This is because matching techniques (whether entropy balancing or other matching techniques) may not be successful where these differences are very large. A related point is that data on relevant socioeconomic variables must be available at a level of geographical disaggregation that is appropriate for studying the programme under evaluation.

- In addition, as discussed above, such an approach is only suitable for evaluating the impact of programmes that have impacts in their local region. The benefits of many RD&I initiatives are not limited to their local region and indeed, may be primarily outside this region. The extent to which a particular programme is expected to have local impact should be determined on a case-by-case basis, ideally with reference to the programme's logic model or theory of change.

It is worth emphasising that this methodology can be applied to the evaluation of a large range of programme outcomes, more specifically, any regional outcome for which geographically disaggregated data is available (for example, employment, productivity or private sector RD&I spending).

Finally, Professor Schubert noted the importance of recognising the diminishing returns of repeatedly conducting such analyses for a single programme. Given the imprecise nature of such econometric estimates, there is little to be gained by comparing the macro-level effects estimated using such models on an annual basis.[19] If such analyses are to be repeated for a single programme, it is more suitable to do so at extended time intervals.

9.9 Conclusions

This evaluation provides a methodologically rigorous analysis of the local economic impacts of Europe's largest organization for applied research and development services. It employs a relatively cost-effective research design using secondary data to provide strong evidence that Fraunhofer institutes have a substantial impact on GDP in the regions where they are located. However, the research findings do not provide clear insight into the causal mechanism underlying this economic impact. This methodology may be transferable to the evaluation of a limited number of RD&I programmes in the UK context, given that certain conditions related to the programme's design and geographical footprint are met.

19 Interview with Professor Torben Schubert, 16/11/2023.

10 Accompanying evaluation of the pilot project "ZIM Cooperation Networks International" and investigating the promotion of international cooperation in research and development in the Central Innovation Program SMEs (ZIM)

10.1 Overview

Key characteristics	
Country	Germany
Institution	Federal Ministry for Economic Affairs and Energy (BMWi)
Type of RD&I	Collaborative R&D
Type of Intervention	Research and innovation grants
Evaluation challenges	Additionality; Spillovers The key challenge is assessing the immediate effects of ZIM, in terms of behavioural additionality in both firms and research institutions and capture any spillovers
Evaluation methodology and methods	• Theory-based (Logic model; literature-based analysis; Social Network Analysis; Case studies) • Quasi-experimental (propensity score matching PSM)

10.2 Introduction

This case presents the evaluation of the international component of the Central Innovation Programme for SMEs (ZIM) created in 2008 as a result from the merger of several previous programmes becoming one of the pillars of German innovation policy to support SMEs efforts around research and innovation. The evaluation seeks to determine the direct and indirect effects of the programme, including its effectiveness, success, behavioural additionality of ZIM in the companies funded, and spillovers.

The evaluation adopts a comprehensive evaluation framework underpinned by mixed methods, guided by a theory-based analysis to determine the causal linkages of the programme and using Propensity Score Matching to assess additionality and indirect effects. Social Network Analysis is used to assess changes in the network structures, alongside case studies to provide further details on the impact of the programme.

10.3 Background description of RD&I intervention

The Central Innovation Programme (ZIM) aims to foster market-driven technology-based R&D activities within German companies. The programme is designed to enhance companies' capacity to innovate and to strengthen their long-term competitiveness. ZIM has amongst its main goals to contribute to the national target of achieving a 3.5% of GDP R&D expenditure and supporting firms to achieve an innovator rate of 50%. ZIM seeks to stimulate start-ups activity and SMEs digitization, cooperation between research and business, and strengthen technology transfer. The programme was launched in 2008 seeking to consolidate other instruments into one and more effective intervention.

Under ZIM, there are three main funding instruments:

- **Individual projects:** the programme provides funding for individual companies doing their in-house R&D work, with a maximum of € 380,000 budget per project, which can be spent on subcontracts, and activities with research institutions (to a maximum of 25% of the total project).
- **Cooperation projects**: the programme provides funding for R&D work carried out jointly by two or more companies, or by one company and one or more research institutes. The cooperation projects are possible in two variants (at least two companies or at least one company and one research institution). Compared to individual projects, higher indirect effects (spillovers) are expected from this funding mechanism. Projects are funded up to two million euros, with sub-projects totalling €380,000 per company, and research institutions €190,000.
- **Cooperation networks**: the programme provides funding for external network management work carried out by innovative networks that comprise at least six SMEs which jointly develop a common innovation. The maximum grant for national network management is €380,000 with phase 1 not exceeding €160,000. For international networks

Since 2008, ZIM has funded 470 international R&D cooperation projects, as well as 18 international cooperation networks since 2019.

10.4 Challenges measuring impact

The evaluation seeks to assess the immediate effects of ZIM, in terms of behavioural additionality in both firms and research institutions. For companies, it concerns cooperation behaviour in brainstorming, sustainability of the cooperation, and innovation and cooperation management. The programme additionality on R&D investments made by companies and their propensity to innovate is also part of the impacts analysed. Similarly, for research institutes, additionality is explored around their attractiveness as cooperation partners, expansion of their research networks, emergence of spin-offs and expansion of their expertise.

A second challenge relates to identifying and measuring the spillovers generated by the different programme pillars to arrive at conclusions of direct and indirect effects of the programme.

10.5 Methodology and data sources

10.5.1 Methodology

The evaluation seeks to assess the commercial impacts of the intervention on sales, employment, and exports as well as expenditure on R&D, internationalisation effects, and changes in competence and behaviour in companies.

A control group approach was applied to enable an in-depth impact analysis of the international cooperation of ZIM supported projects compared to comparable instruments and companies. Two control group approaches were used:

- Control group 1: nationally cooperating companies in the ZIM
- Control group 2: companies that are not ZIM-funded but cooperate internationally

The matching procedure was based on a systematic analysis of the differences between treatment and control groups described above. This analysis drew on over 10,000 responses from the RKW Competence Center surveys and around 700 data sets from internationally cooperating companies in the ZIM. This allowed the assessment to determine to what extent there were significant differences between the treatment and control groups described above. This approach allowed to draw conclusions about differences in the target groups reached, particularly about their number of employees, R&D activity and export behaviour, in order to inform the choice of matching variables used in the PSM procedure.

PSM was used to examine in depth the effects that result from the international cooperation in the ZIM. This PSM procedure compared the international programme beneficiaries with comparable non-beneficiary companies that are as similar as possible (control group consisting of “statistical twins”). This approach compensated for distortions that arise from comparing companies that were already different before the funding (known as selection bias).

The PSM method is based on a propensity score P(X), which is defined as the conditional probability of participating in a funding program based on observed characteristics. The participation probability of the funded person is estimated individually based on multidimensional statistical characteristics of the funded person (e.g. turnover, industry affiliation). Subsequently, participant/non-participant pairs or treatment and control groups are formed with (approximately) the same probability of participation. The pairs are not only brought together based on the agreement in the individual characteristics. Rather, the combination of the characteristics for the derivation of a participation probability P(X) is decisive. In a subsequent step - with the matching - a non-participant is identified for individual funded persons and their individual propensity score value, whose P(X) comes closest to that of the funded person. This results in a participant/non-participant pair and the control group of non-participants formed in this way can now be used for the analysis to compare the participant groups.

To create the different models, a core data set was first developed in which the project management data and information from company databases on company characteristics, the results of the online surveys by Fraunhofer IMW, the RKW Competence Center and KMU-Forschung Austria are brought together to inform 17 variables (see below).

Models and indicators

Target group analysis and indicators for matching similar companies: treatment and control groups 1, 2		Impact analysis: Indicators for recording different development dynamics (1.5-2 years after the end of the project or 4 years after T0)	
1	Founding year (T0)	10	Employees (t4-t0)
2	Turnover (€ million) (T0)	11	Conversion (t4-t0)
3	Number of employees (T0)	12	Exports in millions (t4-t0)
4	Exports (million €)(T0)	13	Exports (t4-t0)
5	Export quota (T0)	14	Export quota (t4-t0)
6	R&D expenditure (million €)(T0)	15	Export rate (t4-t0)
7	R&D rate (T0)	16	R&D expenditure in millions (t4-t0)
8	Sector/ WT (T0)	17	R&D expenditure (t4-t0), %
9	Knowledge intensity of Industry (T0)		

Variables 1 to 9 were used to carry out the target group analysis in order to examine central structural characteristics of the funded and non-funded companies and to identify possible significant differences between the groups. Variables 1 through 9 were also used to match similar companies to form control groups of companies that mirrored the treatment group.

Parametric and non-parametric tests were used to examine the treatment effects, depending on the nature of the data, and the direction, strength and significance of the effect of the treatment (the ZIM promotion of international cooperation) were calculated (indicators 10 to 17)

To complement the quantitative analysis, the evaluation developed detail case studies to examine the individual direct and indirect effect on companies and research institutions and to validate the results of the analytical quantitative analyses. Case studies were based on individual projects, cooperation projects with all cooperation partners and ZIM innovation networks, each with the network manager and selected network partners to be able to distinguish between the three different funding mechanisms.

To select the case studies, the evaluation used a stratification of selected projects and networks based on a cluster analysis to reflect the heterogeneity of the project participants. Cases and reports previously prepared by RKW and available to the evaluation team were integrated to select a total of 23 reports (8 individual projects, ten cooperation projects and five networks). The evaluation developed 25 cases studies in total, five of which belong to individual projects, 12 to cooperation projects and eight to networks.

Focus groups were organised in September 2018 at several points with different stakeholders as follows:

- A Network Focus Group with five network managers and three network partners (companies)
- A research institution focus group with four of them were acting as network managers as well.
- A companies focus group some of which are or were active in networks

10.5.2 Data sources

- Online surveys of stakeholders involved in international R&D projects in the ZIM and for international cooperation networks.
- Interviews with 52 stakeholders: programme managers (2), German managers of networks (15), foreign coordinators of networks (10), representatives with German companies (20) and representatives from partner countries (6). Insights from the interviews informed the bases of case studies on R&D cooperations and cooperation networks and were used for triangulating the insights gained from the surveys and secondary data.
- A focus group with 17 representatives from network management institutions involved in international cooperation networks to validate results.
- Programme documents including application and reporting documents.
- Data from surveys conducted by:
    - RKW Competence Center for ZIM projects that had completed in 2010 and 2016
    - KMU-Forschung Austria on ZIM projects in 2019 on ZIM projects starting in 2015
    - Fraunhofer IMW on national and international cooperation in the ZIM from 2015

For the impact analysis (PSM):

- Data provided by the Center for European Economic Research (ZEW) as part of a project running parallel to the main evaluation of the ZIM.
- Mannheim Innovation Panel (MIP) CORDIS funding data that allowed in-depth insights into the target group, their structural characteristics and the effects of international cooperation. In addition, the MIP data contains information on the cooperation behaviour of the companies and the locations of the cooperation partners. This data source was used to create a control group for companies that have experience in international cooperation in R&D but are not funded in the ZIM.
- Corporate databases: MARKUS and AMADEUS were used to determine company-related variables.

10.6 Main findings

Overall, the evaluation found that ZIM has had positive effects on companies' R&D funding. In 67% of the cases, additional investments have been made by the companies to further develop the project results up to market entry, with nuanced differences across firm sizes. For micro-enterprises represent 70% of the cases, while small enterprises are 68% and medium-sized enterprises stand at 64%.

The location of firms appears to have significant effects on additionality. For example, companies from western Germany made additional investments in 71% of the cases, this was only the case in 58% of the projects by companies from eastern Germany. This is explained by the fact that only 53% of the small eastern German companies made further investments.

The international cooperation ZIM programme has similar positive effects on beneficiaries' turnover and employment as those observed in the overall evaluation of the ZIM programme. In comparison to non-subsidised, internationally active enterprises, the enterprises funded within the framework of the international component of ZIM show significantly higher growth in employment, turnover and exports two years after the end of the funding period.

In terms of cooperative behaviour, ZIM showed a positive effect on its beneficiaries, even on individual projects, with 10% of their project ideas being sourced from outside the company, 50% for cooperation projects and in the case of network projects 75%.

The evaluation found that ZIM has had effects beyond the funded companies, particularly in collaborative projects and individual projects. These effects refer to higher demand, and contributions to the state of the art in the technology field, generating a distributed impact across the value chain. An important finding is the presence of technology spillovers from cooperation projects, enabling knowledge transference and industrial application of this knowledge.

10.7 Analysis of methodological suitability and effectiveness

The evaluation uses a mixed methods approach, which combines the use of existing programme data and corporate data sets with primary data collection via surveys, interviews and focus groups. Analytical approaches used:

- A control group analysis where comparisons were carried out between ZIM-funded companies in the international and national variants as well as between ZIM-internationally funded companies and internationally cooperating but not funded companies. On the one hand, differences in the company groups were analysed (target group analysis), and on the other hand, the additional effects of international funding were isolated and determined using a PSM process. Importantly, the PSM analysis conducted for this evaluation is of high quality. The evaluation team employs several methods to ensure that the analysis sufficiently rigorous to address the challenge of programme additionality. First, the matching procedure is based on a systematic analysis of the differences between programme beneficiaries and non-beneficiaries. Second, the evaluation conducts two sets of analyses, each using a different counterfactual group. Implementing both sets of analyses allows the evaluation team to rigorously test the robustness of the econometric results.
- A network analysis of international cooperation networks was carried out for the three international cooperation networks. To collect the network data, a question module on the actor relationships in the relevant network was added as part of the online survey. The network partners were evaluated by the survey participants on various dimensions (closeness of cooperation, distribution of the added value of the cooperation and contribution to the success of the network). The network analysis was carried out with the software package R.
- Although, the response rate for the survey was low (~25%) it was deemed satisfactory response level from a methodological point of view. The low response rate was likely in part due to the international R&D projects being several years old and the contact person not being reachable.
- For the control group: Mannheim Innovation Panel (MIP) CORDIS funding data that allowed in-depth insights into the target group, their structural characteristics and the effects of international cooperation. In addition, the MIP data contains information on the cooperation behaviour of the companies and the locations of the cooperation partners. This data source was used to create a control group for companies that have experience in international cooperation in R&D but are not funded in the ZIM.

Overall, the vast majority of evaluation questions were the subject of at least two empirical instruments. The test results could thus be checked for their resilience by means of method triangulation. Two control groups were used to capture the effects of collaboration (one for ZIM beneficiaries cooperating at the national level and companies not ZIM-funded but cooperating internationally). In addition, an in-depth impact analysis of the target groups was applied to account for differences among groups. The design of methods provided a process assessment (to capture the evolution of networks) and impact at the micro level

10.8 Transferability and context

This evaluation is underpinned by a theory-based approach, using an encompassing set of methods to assess the direct and indirect impact of the programme and triangulate evidence from the quantitative analysis via qualitative analysis.

The methodology assembled is robust and suited for the evaluation, answering the evaluation questions and capturing the additionality of ZIM across its three main funding pillars. Moreover, the methods provide means to corroborate results and effects and details on causing mechanisms.

Transferability of the methodology is possible provided monitoring and evaluation data is available and outputs and emerging outcomes are recorded. Primary data collection methods were instrumental to the evaluation, via surveys, interviews and focus groups, which do not represent a limitation in the UK context.

10.9 Conclusions

This evaluation provides a comprehensive overview of the impacts of the international cooperation ZIM programme. It seeks to determine the direct and indirect effects of the programme, including its effectiveness, success, spill overs and behavioural additionality. The evaluation employs a theory-based evaluation approach using mixed methods, including interviews, case studies, quasi-experimental research and social network analysis.

The international cooperation ZIM programme is found to have had similar positive effects on beneficiaries' turnover and employment as those observed in the national ZIM programme. When compared to non-subsidised, internationally active enterprises, the enterprises funded by the international cooperation ZIM programme show significantly better economic performance. The international cooperation programme was also found to have had a positive effect on its beneficiaries in terms of cooperative behaviour and technology spillovers.

The evaluation methodology is robust and suited for the evaluation objectives. Notably, the mixed methods approach provides means to corroborate results and detail the causal mechanisms. Transferability of this methodology is possible provided that the relevant monitoring and evaluation data is available, in addition to data on the comparator companies chosen as a control group.

11 Evaluation of the Nano 2017 Programme

11.1 Overview

Key characteristics	
Country	France
Institution	Applied research and innovation
Type of RD&I	Research and innovation grants
Type of Intervention	Ministry of Economics and Finance; local authorities in the Grenoble region; European ENIAC/ECSEL initiative
Evaluation challenges	Additionality The key evaluation challenge is determining the causal relationship between NANO 2017 funding and economic outcomes including job creation as well as beneficiary companies' R&D expenditure and economic performance
Evaluation methodology and methods	• Econometric analysis • Quasi-experimental (difference-in-differences; matching)

11.2 Introduction

This evaluation investigates the impact of the French government's NANO 2017 Programme on the economic performance of programme beneficiaries as well as on the Rhône-Alpes region more generally.

While the programme had a significant number of beneficiaries, the application of a single, unified quasi-experimental methodology was not possible in this case. This is because it was not possible to identify suitable counterfactuals for the two large organisations (STMicroelectronics and CEA-Leti) which received a large majority of the programme's funding. The study therefore adopted a bifurcated approach involving two parallel sets of analyses. The first involved focusing exclusively on estimating the direct, indirect and induced economic outcomes of the programme for the large programme beneficiaries. The second involved a quasi-experimental approach focused on the remaining beneficiaries.

The quasi-experimental approach applied to smaller beneficiaries allowed evaluators to obtain some insights into programme additionality. However, the econometric approach used for the large beneficiaries cannot provide any such insights. Nevertheless, the latter approach may be a valuable complement to other evaluation methods, especially in cases where theory-based and quasi-experimental methods are inappropriate, not feasible or insufficient.

11.3 Background description of RD&I intervention

Launched in 2013, the French government's NANO 2017 Programme aimed to position the Grenoble-Crolles regional cluster as a global leader in the field of Complementary Metal Oxide Semiconductor (CMOS) technologies by supporting an R&D programme for design and production technologies for integrated circuits. It constituted the third generation of such support, following the NANO 2008 and NANO 2012 Programmes.

NANO 2017's R&D programme was based on a technological roadmap developed jointly by the programme's academic and industrial partners, focused on advancing nanolithography techniques and diversifying CMOS production capabilities. The programme was implemented by 81 beneficiary organisations including 23 SMEs, 13 Entreprises de Taille Intermédiaire (ETIs)[20], 12 large companies and 29 public research laboratories. Over €1.9 billion in R&D was carried out by public and private actors on all NANO 2017 projects, €775 million of which was public funding.

The programme largely focused on two key players in the sector. Companies belonging to the STMicroelectronics group represented 64 percent of total programme expenditure and carried out €1.2 billion in R&D as part of the programme over a period of five years. CEA-Leti, a public research organisation specialised in micro and nanotechnologies, represented 24 percent of total programme expenditure.

11.4 Challenges measuring outcomes and impact

The key methodological challenge faced in this impact evaluation was the diverse range of NANO 2017 beneficiaries. While the programme had a significant number of beneficiaries, it was not possible to apply a single, unified quasi-experimental methodology for all beneficiaries. Two large organisations (STMicroelectronics and CEA-Leti) received a large majority of the programme's funding. However, it was not possible to identify suitable counterfactuals for these organisations in order to include them in a quasi-experimental analysis with the remaining beneficiaries. Thus, the evaluation adopted a bifurcated approach involving two parallel sets of analyses: 1) An econometric analysis of economic impact focused exclusively

[20] Intermediate-sized companies, defined as those with 250 to 4,999 employees and a turnover that does not exceed €1.5 billion, or a balance sheet total which does not exceed €2 billion.

on STMicroelectronics/ CEA-Leti and, 2) A quasi-experimental impact evaluation approach focused on the remaining beneficiaries.

The quasi-experimental analysis enabled the evaluation team to obtain insights into the issue of programme **additionality** for some programme beneficiaries. However, the econometric approach used to evaluate programme outcome related to the large beneficiaries does not provide insights into programme additionality. Nevertheless, this approach to econometric analysis can be a valuable complement to other evaluation methods, especially in cases where theory-based and quasi-experimental methods are inappropriate, not feasible or insufficient.

11.5 Methodology and data sources

11.5.1 Methodology

This impact evaluation aimed to determine whether the programme achieved its objectives of strengthening the economic development of beneficiaries as well as the Rhône-Alpes region more generally. It adopted a bifurcated approach involving two parallel sets of analyses: one focused exclusively on STMicroelectronics/ CEA-Leti and a quasi-experimental approach focused on the remaining beneficiaries.

The analysis focusing on STMicroelectronics and CEA-Leti involved estimating the direct, indirect and induced economic outcomes of the programme. SOITEC, a third large organisation which received a significant amount of NANO 2017 funding, was also included in this analysis. The financial and administrative records of the three organisations were used to obtain data on the direct outcomes of the programme, as follows:

- R&D employment related to the programme;
- Turnover generated by products attributable to the programme;
- The workforce employed in the manufacturing of products attributable to the programme.
- The indirect economic outcomes of the programme were estimated by analysing the linkages between these three companies and their suppliers, including the value of local and national purchases attributable to the programme and the number of jobs associated with these purchases.

The induced economic outcomes of the programme were estimated at both the local and national level. At the local level, induced economic outcomes were calculated by estimating the additional employment supported by the consumption and fiscal contributions of direct and indirect employees linked to the programme. The socioeconomic profile of employees associated with the NANO 2017 Programme was taken into consideration in estimating their consumption patterns. Data on the number of salaried positions in local service sectors was used to conduct a sector-by-sector analysis of the contribution of employees associated with the NANO 2017 Programme to each sector.

Induced economic outcomes at the national level were estimated based on three companies' salary costs, national purchases, and social contributions that were attributable to the NANO 2017 Programme. The totality of this expenditure reflects the additional demand injected into the economy by the programme. Using an input-output analysis approach, the evaluation modelled the effect of this additional demand on the economy and labour market. The national induced effects were distinguished from the local induced effects by subtracting the local share from the overall impact, thereby isolating the induced effects outside the Rhône-Alpes region.

The evaluation also employed a difference-in-differences (DiD) research design with propensity score matching in order to establish a causal link between NANO 2017 funding and a beneficiary's R&D/ economic performance. This approach compares the trajectory of NANO 2017 beneficiaries with that of non-beneficiaries that share similar characteristics. STMicroelectronics and CEA-Leti were excluded from this analysis.

The evaluation finds that NANO 2017 beneficiaries differ from non-beneficiaries in several ways, most notably in terms of their R&D expenditure, export performance and value-added. Thus, to construct a credible counterfactual group, the study employed a propensity score matching technique to balance the distribution of NANO 2017 beneficiaries and non-beneficiaries on these observed characteristics, thus mimicking a randomised experiment.

Matching techniques, however, do not address the fact that unobserved beneficiary characteristics may bias the comparison between the two groups. For this reason, the study combines propensity score matching with a DiD research design. The DiD approach considers the difference in outcomes between NANO 2017 beneficiaries and non-beneficiaries before the establishment of the programme (the pre-intervention period), and the difference between the two groups after the programme's establishment (the post-intervention period). It then takes the difference between these two differences. DiD estimation is commonly used for evaluating nonrandomized interventions. This is because it removes biases in post-intervention comparisons between the treatment (NANO 2017 beneficiary) and control (comparison) groups that may result from unobserved differences between these groups, in addition to biases from comparisons over time in the treatment group which result from trends due to other causes of the outcome.

Three sets of DiD analyses with propensity score matching were undertaken, as follows:

- The first compared direct and indirect beneficiaries of the programme to a control group of similar R&D active companies in the same economic sector in order to examine the impact of the programme at the company level.
- The second examined territorial spillover effects by comparing R&D companies in beneficiary municipalities to those in other parts of France operating in similar economic sectors.
- The third examined the regional effect on the computer, electronic and optical industries, compared to other industrial subsectors. These analyses compared companies in these selected subsectors that are located in Rhône-Alpes to companies in other industrial subsectors that are also located within the region.

The above-described analyses were used to estimate the impact of the programme on the following outcomes: R&D expenditure, R&D employees, total number of employees, turnover and export value.

Finally, the above figures were used to conduct an analysis of the programme's value for money which estimated:

- The average public cost of each direct NANO 2017 job created or maintained per year, compared to the average cost per year of each unemployed person in France.
- The tax revenue raised as a result of the programme, including sales tax and social contributions.

11.5.2 Data sources

The following data sources were used for the analysis of NANO 2017 Programme economic impact achieved by STMicroelectronics, CEA-Leti and SOITEC.

- Data on the direct outcomes of the programme was obtained from the financial and administrative records of the three organisations.
- For the estimation of indirect economic outcomes of the programme companies' purchasing files were used to ascertain the value of local and national purchases attributable to the programme. Indirect employment estimates were based on either: 1) contractual agreements with suppliers, where these companies had contractually committed to providing a specific number of employees to the client, or 2) employment to turnover ratios for specific supplier companies, or 3) annual sectoral employment to turnover ratios published by INSEE, the French national statistical agency.
- French National Institute of Statistics and Economic Studies (INSEE) data on the number of salaried positions in local service sectors was used to conduct a sector-by-sector analysis of the contribution of employees associated with the NANO 2017 programme to each sector.
- The input-output analysis employed data from input-output tables published by INSEE.

The evaluation does not provide a detailed description of the data sources used for the DiD with propensity score matching analysis. However, it lists the following as key data sources for this analysis:

- The CORDIS database on EU-funded research and innovation projects.
- The Diane database on French companies published by Bureau van Dijk, a publisher of business information specialised in private company data.

11.6 Main findings

The analysis conducted on STMicroelectronics, CEA-Leti and SOITEC suggests that the NANO 2017 programme directly employed 2,700-4,500 workers. In addition, it created 500-800 indirect jobs and 2,500-4,400 induced jobs across France. Focusing specifically on employment outcomes for the Rhône-Alpes region, the evaluation found that all of the direct employment opportunities created were in located in the region, in addition to 300-500 of the indirect jobs and 1,800-3,100 of the induced jobs. The evaluation report does not detail the findings from this analysis in terms of company expenditure attributable to the programme.

The DiD with propensity score matching analyses conducted for all NANO 2017 beneficiaries excluding STMicroelectronics and CEA-Leti findings suggest that:

- NANO 2017 increased levels of R&D expenditure, R&D employment and export turnover among participating companies. However, the programme was not found to have had an impact on overall company turnover.
- R&D companies in the Crolles-Grenoble cluster showed a greater increase in R&D employment and export turnover than matched companies in the control group. However, no such local effect was observed for R&D expenditure. For the remaining indicators, the results were difficult to interpret because the DiD with propensity score matching analyses were not successful in eliminating pre-treatment differences between the treatment and control groups.
- Finally, R&D companies in the computer, electronic and optical industries that are located in Rhône-Alpes were found to have outperformed companies in other industrial subsectors located within the same region in terms of: R&D expenditure, R&D salaries and export turnover. No significant effect was found for the other indicators under study.

In terms of value for money, the evaluation estimated that the average cost per job created or preserved is €15,000 – €24,000 per year (including direct, indirect and induced jobs). This compares to an average cost of between €20,000 – €28,000 per year per unemployed person in France. The evaluation report notes that these estimates should be taken with caution, as salaries within STMicroelectronics are above the French average. A full year of unemployment therefore costs more than the French average. However, STMicroelectronics employees would also likely suffer a shorter period of unemployment than the national average due to their high levels of qualifications. The evaluation also estimated that for every €1 of public support for the NANO 2017 programme, €1.1 of social contributions and sales were generated.

11.7 Analysis of methodological suitability and effectiveness

The methodology employed in this evaluation has a number of strengths.

- The estimation of the economic outcomes resulting from STMicroelectronics/ CEA-Leti participation in the programme provides valuable insights into the programme's economic outcomes. Although it does not address the key question of additionality, it is important to recognise that the quasi-experimental methods often used to demonstrate additionality in impact evaluations are difficult to implement in the evaluation of programmes with a small number of beneficiaries. As highlighted in the evaluation report, it was not possible to identify suitable counterfactuals for these organisations so that they could be included in the DiD with propensity score matching analyses. It is notable that the synthetic control method is an alternative quasi-experimental approach often employed for impact evaluation when the number of programme beneficiaries is small. However, it is likely that this method would also be challenging to implement for programmes such as this one, with large and highly specialised beneficiaries such as STMicroelectronics and CEA-Leti. This is because constructing a synthetic control also requires the identification of counterfactual organisations.

- The DiD analysis with propensity score matching allows the evaluation team to draw some conclusions about the programme's additionality.
- In general, the use of matching methods can be contentious due to the fact that the use of different matching procedures can sometimes yield entirely different research results. This study provides a transparent description of the systematic matching procedure employed. This helps address concerns about the matching methodology and is a good practice to be emulated.
- The study presents formal tests of the parallel trends assumption, a key assumption that underpins the validity of DiD research designs. This provides strong evidence for the validity of the research design and is a good practice to be emulated.

This methodology also has several limitations.

- The beneficiaries eligible for inclusion in the DiD with propensity score matching analysis account for a mere 12 percent of programme expenditure. The results of this analysis therefore provide limited insights into overall programme effectiveness.
- The NANO 2017 Programme was implemented over the period 2013-2017, with the evaluation's analyses covering the same period. However, the impacts of R&D support programmes can take many years to materialise. Such lagged effects are not captured by this evaluation.
- The NANO 2017 Programme is the third generation of supports for CMOS technologies in the Grenoble-Crolles regional cluster, closely following the NANO 2008 and NANO 2012 Programmes. It is challenging to disentangle the effects of the NANO 2017 programme from previous generations of the programme, particularly because the beneficiaries of the three programmes are highly overlapping.

11.8 Transferability and context

Some aspects of the methodology employed in this evaluation are transferable to evaluating programmes designed to support RD&I among UK enterprises. The DiD with propensity score matching approach can be applied for evaluating programmes with:

- A number of programme beneficiaries that is sufficiently large to conduct a DiD analysis.
- Available company-level data for indicators of the programme's intended impacts, both for programme beneficiaries and a sufficient number of comparator beneficiaries, as well as for both the pre-intervention and post-intervention periods.

In addition, the approach used in this evaluation to estimate the direct, indirect and induced economic outcomes of the programme can be a valuable complement to other evaluation methods, especially in cases where theory-based and quasi-experimental methods are inappropriate, not feasible or insufficient. Implementing this approach requires access to detailed beneficiary administrative and financial records. Ideally, it also requires access to employment and turnover data for beneficiaries' suppliers.

11.9 Conclusions

This impact evaluation aimed to determine whether the programme achieved its objectives of strengthening the economic development of beneficiaries and that of the Rhône-Alpes region.

It adopted an approach involving two parallel sets of analyses. The first involved estimating the direct, indirect and induced economic outcomes of the programme for its largest three beneficiaries. The second employed a DiD research design with propensity score matching to investigate the impact of programme participation on R&D expenditure, R&D employees, total number of employees, turnover and export value. The findings suggest that:

- The programme generated a significant number of jobs, and that the public revenue generated due to programme expenditure likely exceeds the public cost of the programme.
- Participation in the programme led to higher levels of R&D expenditure, R&D employment and export turnover among programme beneficiaries.
- The programme improved some aspects of the economic performance of R&D companies in the Crolles-Grenoble cluster as well as companies in the computer, electronic and optical industries that are located in the Rhône-Alpes region.

The approach used in this evaluation to estimate the direct, indirect and induced economic outcomes of the programme can be a valuable complement to other evaluation methods, especially in cases where theory-based and quasi-experimental methods are inappropriate, not feasible or insufficient. However, this approach does not allow evaluators to address the key question of programme additionality. The DiD with propensity score matching analysis methodology is transferable to the evaluation of large-scale programmes to support RD&I among UK enterprises, subject to availability of the data required for such an analysis.

12 Econometric evaluation of aid to collaborative R&D projects (2005 -2019)

12.1 Overview

Key characteristics	
Country	France
Institution	Government of France
Type of RD&I	Collaborative RD&I
Type of Intervention	Direct assistance
Evaluation challenges	Additionality; attribution The key evaluation challenges are 1) Determining the causal relationship between programme funding and beneficiary companies' R&D expenditure, patenting activity and economic performance. 2) Untangling the effects of this programme from those of other major RD&I support programmes that beneficiaries may have accessed simultaneously.
Evaluation methodology and methods	• Quasi-experimental (difference-in-differences; matching)

12.2 Introduction

This evaluation investigates the impact of two collaborative RD&I interventions implemented in France: the European Commission's State Aid for RD&I Programme (SARDIP) and the Future Investment Programme (FIP). It focuses on the programmes' impact on companies' R&D spending, patenting activity, employment and economic performance.

The evaluation utilises a difference-in-differences (DiD) research design with propensity score matching to determine the impact of programme funding on businesses' R&D expenditure, patenting activity and economic performance. This approach allows evaluators to address the evaluation challenge of determining programme additionality. The econometric analyses also account for whether a company received funding for collaborative projects from other major RD&I programmes. This means that the econometric models can isolate the impacts of the RD&I programmes under study from those of key similar programmes, allowing evaluators to provide some tentative insights regarding programme attribution.

12.3 Background description of RD&I intervention

Two collaborative RD&I interventions implemented in France are studied in this evaluation: the European Commission's State Aid for RD&I Programme (SARDIP) and the Future Investment Programme (FIP). More specifically, the evaluation investigates the impact of:

- Five components of SARDIP: the Single Inter-Ministerial Fund, Structuring Projects for Competitiveness, Industrial Projects for the Future, B Missions of Competitiveness Clusters and the Dual Innovation Support Scheme. These mainly consist of support for R&D collaborations, in particular between businesses and research organizations.
- The environmental component of FIP, a French state investment program which aims to stimulate RD&I activities in growth-promoting sectors as diverse as the digital economy, health, and higher education and research. The environmental component of FIP finances innovations to support the energy and environmental transition and is implemented by the French Agency for Ecological Transition (ADEME). Over the period 2010-2019, ADEME provided over €2 billion in support to around 700 projects under this programme. Because this funding was largely allocated to collaborative R&D projects, the environmental component of FIP has been grouped with the aforementioned SARDIP initiatives for the purposes of this evaluation.

Over the period under evaluation, these programmes (2005-2019) provided total funding of around €7 billion to support 2,972 projects.

12.4 Challenges measuring outcomes and impact

This evaluation addresses three key evaluation challenges: additionality, attribution and data limitations.

- Regarding the challenge of additionality, the impact evaluation must address the question of whether any observed increases in R&D spending, patenting activity, employment and economic performance are likely to have occurred in the absence of SARDIP/FIP funding. This is achieved by employing a high-quality quasi-experimental research design.
- Regarding the challenge of attribution, the impact evaluation must address the question of whether SARDIP/FIP beneficiaries share characteristics aside from programme funding (for example, participation in other RD&I programmes) which underlie any observed changes in R&D spending, patenting activity, employment and economic performance. Put differently, it must address the "third variable" problem, in which an unobserved variable correlates highly with SARDIP/FIP funding and muddies any causal explanation. A common challenge in attributing impact to a specific RD&I intervention is the fact that many programme beneficiaries participate in multiple similar programmes that aim to promote RD&I. By accounting for whether a company received funding for collaborative projects from other major overlapping RD&I programmes, the

analysis used in the evaluation provides some tentative insights regarding programme attribution.

- It employs an innovative econometric approach ("chained DiD") to deal with data limitations, specifically the fact that a complete longitudinal (panel) dataset for programme beneficiaries and comparator businesses was not available.

12.5 Methodology and data sources

12.5.1 Methodology

This impact evaluation aims to determine the effect of the SARDIP and FIP programmes described above on businesses' R&D spending,[21] patenting activity, employment and economic performance.

It utilises a high-quality DiD research design with propensity score matching to investigate whether there is a causal link between programme funding and a company's R&D expenditure, patenting activity and economic performance. This research design compares the trajectory of programme beneficiaries with that of non-beneficiaries that share similar characteristics. An analysis of these two groups of companies shows that programme beneficiaries differ from non-beneficiaries in several ways, notably that they are larger, have more experience in using public R&D support systems and are more often specialized in high technology sectors. Thus, to construct a credible counterfactual group, the study employed a propensity score matching technique to balance the distribution of programme beneficiaries and non-beneficiaries on these observed characteristics. In this way, the analysis mimics a randomised experiment. It was not possible to find counterfactuals for large companies and so they were excluded from the analysis. The analysis therefore only contains estimates for SMEs and Entreprises de Taille Intermédiaire (ETI).[22]

Matching techniques, however, do not address the fact that unobserved company characteristics may bias the comparison between the two groups. For this reason, the study combines propensity score matching with a DiD research design. The DiD approach considers the difference in outcomes between programme beneficiaries and non-beneficiaries before the establishment of the programme (the pre-intervention period), and the difference between the two groups three years after the programme's establishment (the post-intervention period). It then takes the difference between these two differences. DiD estimation is commonly used for evaluating nonrandomized interventions. This is because it removes biases in post-intervention comparisons between the treatment (programme beneficiary) and control (comparison) groups that may result from unobserved differences between these groups, in addition to biases from

21 Data on RD&I expenditure more broadly was not available.

22 Intermediate-sized companies, defined as those with 250 to 4,999 employees and a turnover that does not exceed €1.5 billion, or a balance sheet total which does not exceed €2 billion.

comparisons over time in the treatment group which result from trends due to other causes of the outcome.

The evaluation uses a novel DiD model designed specifically for DiD analyses where: 1) programme beneficiaries are not treated at the same time but rather participate in programmes gradually over time, and 2) data for treated companies is not consistently available over a period that covers several years before and after the start of their participation. The "chained DiD" model used in the analysis is an extension of well-established staggered DiD approaches. It aggregates a collection of short-term treatment effects to estimate a long-term treatment effect, and was recently developed by the authors of this evaluation. At the time of writing, a detailed description of this model and its properties has yet to be published in a peer-reviewed publication.

In addition to estimating the overall effects of the programme, the analysis explores the heterogeneity of programme effects among various subgroups of beneficiaries. A large number of these disaggregated analyses are conducted, including analyses of:

- Projects in different economic sectors.
- Projects in which public research organisations have a significant (versus minor) presence.
- Projects with a large (versus small) number of partners.
- Project in which at least one large company is participating (versus those with no large company participation).

Importantly, the quantitative analysis conducted for this evaluation is of high quality. The evaluation team employs several methods to ensure that the analysis sufficiently rigorous to address the challenge of programme additionality. First, the matching procedure is based on a systematic analysis of the differences between programme beneficiaries and non-beneficiaries.

Second, the evaluation conducts two sets of analyses, each using a different counterfactual group. The first counterfactual group is drawn directly from the Ministry of Education, Higher Education, Research and Innovation (MEHRI) survey of firms' R&D spending. The second counterfactual group is drawn from a version of the MEHRI dataset that is restricted only to companies that have either unsuccessfully applied for funding to one of the programmes under evaluation or are members of France's competitiveness cluster policy.23 The latter group is likely to better reflect the (unobservable) characteristics of companies interested in RD&I collaboration. However, the sample size of the latter dataset is smaller, making the DiD estimates obtained from these analyses less precise. Implementing both sets of analyses allows the evaluation team to rigorously test the robustness of the econometric results.

The study also accounts for whether a company received funding for collaborative projects from the French National Research Agency and/or funding for collaborative projects awarded by the research framework programs FP7 and Horizon 2020. This means that the econometric

23 A major French government collaborative RD&I programme.

models can isolate the impacts of the SARDIP/FIP components under study from the impacts of these other major overlapping programmes, allowing evaluators to provide some tentative insights regarding programme attribution.

12.5.2 Data sources

The evaluation employs French company data from multiple sources matched to administrative data on participation in the programmes under evaluation. The external data sources used are:

- The MEHRI survey of firms' R&D spending, which provides company-level data on R&D funding, R&D spending and R&D employment.
- French governmental tax records, which provide company-level financial data.
- The Patent Atlas, which provides company-level data on the number of patent applications made to the French and European Patents Offices.
- The French statistical agency's Annual Social Data Declarations, which provides company-level data on total employment and employment of engineers.

12.6 Main findings

The evaluation finds that, on average, programme participation has a positive effect on the R&D expenditure directly carried out by companies (GERD) but no significant effect on outsourced R&D expenditure. The positive effect on R&D expenditure is largely directed to increased R&D employment and, to a lesser extent, increased remuneration for R&D employees. Importantly, the evaluation finds a significant positive effect of programme participation on the private R&D spending of programme beneficiaries (total R&D spending minus total public grants). This suggests that programme participation caused beneficiaries to increase their private investment in R&D.

With regard to the programmes' effect on economic performance, the evaluation finds this to be limited. Even five years after programme participation begins, the analysis does not find any significant effect on company turnover or valued added. The analyses conducted also suggest a significant effect on the number of patents filed, although the size of this effect varied greatly depending on the data source used.

With regard to the heterogeneity of programme effects among subgroups of beneficiaries, the evaluation conducted a large number of analyses. The key findings that emerged from these analyses are:

- Later cohorts of projects financed by the Single Inter-Ministerial Fund produced effects significantly different from those observed for early cohorts. In later cohorts, the increase in GERD is much larger and unlike early cohorts, this increase is almost exclusively dedicated to applied research.
- The participation of research organizations in projects is associated with a stronger effect on GERD and the number of patents filed.

- SME and ETI beneficiaries participating in projects with large companies file a lower number of patents than those participating in projects with no large company partners.

12.7 Analysis of methodological suitability and effectiveness

The methodology employed in this evaluation has a number of strengths.

- The high-quality DiD analysis with propensity score matching allows the evaluation team to draw credible conclusions about the question of additionality surrounding programme participation. The use of two different counterfactual comparison groups in the analyses demonstrates the robustness of the econometric findings in a compelling manner.
- The use of the chained DiD model is an innovative approach to implementing a DiD research design with incomplete panel data. However, as noted above, a detailed description of this model and its properties has yet to be published in a peer-reviewed publication.
- Investigating the heterogeneity of programme effects provides valuable policy-relevant insights. For example, it points to the important role that public research organisations can play as RD&I collaborators. It also points to changes in the effectiveness of the Single Inter-Ministerial Fund over time, which programme funders and administrators may choose to investigate further.

The methodology also has several limitations.

- Most notably, the evaluation was not able find a counterfactual for large companies which receive nearly 20 percent of the funding disbursed by the programmes under study. Thus, the evaluation findings only relate to programme impact on SMEs and ETIs. The dynamics of RD&I investment and activity among large companies may differ substantially from those of other companies, and it cannot be assumed that the results of the evaluation are generalisable to large companies.
- The evaluation only examines the programmes' impact on R&D spending, patenting activity, employment and economic performance. It does not provide insights into whether the programmes increased RD&I collaboration, a key objective of French collaborative RD&I programmes.

12.8 Transferability and context

The methodology employed in this evaluation is transferable to evaluating programmes designed to support RD&I among UK enterprises, given the following conditions are met:

- The number of programme beneficiaries is sufficiently large to conduct a DiD analysis.
- Company-level data on RD&I expenditure and economic performance is available for programme beneficiaries and a sufficient number of comparator companies for both the pre-intervention and post-intervention periods.

- Data is available on the participation of programme beneficiaries and comparators in key RD&I programmes similar to that being evaluated.

It is worth noting, however, that the chained DiD approach may be challenging to implement given that it has yet to be integrated into common statistical analysis software packages.

12.9 Conclusions

This evaluation provides a methodologically rigorous analysis of whether certain components of SARDIP and FIP funding have an impact on beneficiaries' R&D expenditure, patenting activity and economic performance. It employs a DiD research design with propensity score matching to provide strong evidence that these programmes have a positive effect on companies' GERD and the number of patents they file. However, the econometric analysis does not find any significant effect on beneficiaries' financial performance. The most significant limitation of the methodology used in this evaluation is that it was not possible to find a counterfactual for large companies, which receive 20 percent of the funding disbursed by the programmes under study. Thus, the evaluation findings only relate to programme impact on SMEs and ETIs. This methodology is transferable to the evaluation of large-scale programmes to support RD&I among UK enterprises, subject to availability of the data required for a chained DiD with propensity score matching analysis.

13 Evaluation of the Impact of Individual Innovation Grants Distributed by Bpifrance

13.1 Overview

Key characteristics	
Country	France
Institution	Bpifrance
Type of RD&I	Experimental development
Type of Intervention	Direct assistance
Evaluation challenges	Additionality; attribution The key evaluation challenges are 1) Determining the causal relationship between Bpifrance Individual Support for Innovation Programme funding and beneficiary companies' RD&I investment and their economic performance. 2) Untangling the effects of this programme from those of other major RD&I support programmes that beneficiaries may have simultaneously accessed.
Evaluation methodology and methods	• Quasi-experimental (difference-in-differences; matching)

13.2 Introduction

This evaluation investigates the impact of Bpifrance's Individual Support for Innovation Programme (ISIP). It focuses on the programmes' impact on companies' level of RD&I investment and their economic performance.

The evaluation utilises a difference-in-differences (DiD) research design with propensity score matching to determine programme impact. This approach allows evaluators to address the evaluation challenge of determining programme additionality. The econometric analyses conducted also account for whether a company received funding from three other key RD&I support programmes deemed comparable to ISIP. This means that the econometric models can isolate the impacts of ISIP from those of key similar programmes, allowing evaluators to provide some tentative insights regarding programme attribution.

13.3 Background description of RD&I intervention

Bpifrance is a French public sector investment bank that aims to promote the growth of the French economy by supporting entrepreneurship. Among the bank's programmes is ISIP, which promotes innovation among French enterprises and is comprised of seven individual initiatives. The largest ISIP initiative, both in terms of expenditure and number of beneficiaries, is a system of direct financial supports provided to companies in the form of matching grants, repayable advances or zero-interest loans, which is administered through a decentralized application process. ISIP also runs the Bourse French Tech initiative, which provides grants to cover eligible business expenses for innovative, early-stage start-ups with high growth potential. In addition, ISIP runs three innovation competitions and two regional innovation initiatives. Over the period 2014-2018 (the period covered by this evaluation), ISIP committed around €600 million annually to around 3,500 distinct beneficiaries each year.

13.4 Challenges measuring impact

The evaluation addresses two key evaluation challenges: additionality and attribution.

- Regarding the challenge of additionality, the impact evaluation must address the question of whether any changes in RD&I expenditure and/or economic performance that are observed among ISIP beneficiaries would have been likely to occur even if they had not participated in the programme. Given the systematic differences between ISIP beneficiary companies and non-beneficiaries, establishing programme additionality is a significant challenge. As the evaluation report points out, ISIP beneficiaries are more likely to have received public support for innovation in recent years, to have a high share of R&D expenditure[24] and to be in economic sectors that employ high numbers of engineers and technicians. To address this challenge, the study employs a high-quality quasi-experimental research design.
- Regarding the challenge of attribution, the impact evaluation must address the question of whether ISIP beneficiaries share characteristics aside from ISIP funding (for example, participation in other RD&I programmes or their economic sector) which underlie any observed changes in their RD&I expenditure and/or economic performance. Put differently, it must address the "third variable" problem, in which a certain variable correlates highly with ISIP funding and muddies causal explanation. A common challenge in attributing impact to a specific RD&I intervention is the fact that many programme beneficiaries participate in multiple similar programmes that aim to promote RD&I. Attributing observed impacts to participation in a particular programme can therefore be difficult. To address this aspect of the attribution challenge, the analyses account for whether a company received funding from key RD&I support programmes deemed similar to ISIP. This allows evaluation to provide some tentative insights regarding programme attribution.

24 Data on RD&I expenditure more broadly was not available.

13.5 Methodology and data sources

13.5.1 Methodology

This impact evaluation aims to determine ISIP's impact on:

- Companies' level of RD&I investment. This includes both private RD&I expenditure and total RD&I expenditure (i.e. expenditure including public grants).
- Company economic performance, as measured by a variety of indicators. These are: total turnover, export turnover, value added, total employment, RD&I employment, physical investments, use of bank/private equity financing.

The evaluation employs a high-quality difference-in-differences (DiD) research design with propensity score matching in order to establish a causal link between ISIP funding and a company's RD&I expenditure/economic performance. This approach compares the trajectory of ISIP beneficiaries with that of non-beneficiaries that share similar characteristics. An analysis of these two groups of companies shows that ISIP beneficiaries differ from non-beneficiaries in several ways, notably that they are more likely to have received public support for innovation in recent years, to have a high share of R&D expenditure[25] and to be in economic sectors that employ high numbers of engineers and technicians. Thus, to construct a credible counterfactual group, the study employed a propensity score matching technique to balance the distribution of ISIP beneficiaries and non-beneficiaries on these observed characteristics, thus mimicking a randomised experiment.

Matching techniques, however, do not address the fact that unobserved company characteristics may bias the comparison between the two groups. For this reason, the study combines propensity score matching with a DiD research design. The DiD approach considers the difference in outcomes between ISIP beneficiaries and non-beneficiaries before the establishment of the programme (the pre-intervention period), and the difference between the two groups three years after the programme's establishment (the post-intervention period). It then takes the difference between these two differences. DiD estimation is commonly used for evaluating nonrandomized interventions. This is because it removes biases in post-intervention comparisons between the treatment (ISIP beneficiary) and control (comparison) groups that may result from unobserved differences between these groups, in addition to biases from comparisons over time in the treatment group which result from trends due to other causes of the outcome.

The evaluation also matches data on companies' R&D expenditure[26] and economic performance with data on whether these companies received other forms of Bpifrance individual support, the Research Tax Credit or support from France's Young Innovative Companies Programme over the eight years preceding the evaluation's endline (2018). These

25 Data on RD&I expenditure more broadly was not available.

26 Data on RD&I expenditure more broadly was not available.

three RD&I support programmes are those deemed most comparable to ISIP. The analysis matched ISIP beneficiaries to non-beneficiaries with the same status in regard to participation in these programmes.

Importantly, the quantitative analysis conducted for this evaluation is of high quality. The evaluation team employs several methods to ensure that the analysis is sufficiently rigorous. First, the evaluation employs empirical tests to demonstrate that the analysis meets the key assumption of DiD research designs (the parallel trends assumption). This provides a compelling justification for the validity of this research design in addressing the additionality challenge. Second, the matching procedure is based on a systematic analysis of the differences between ISIP beneficiaries and non-beneficiaries. Third, the evaluation investigates the programme's impact on a range of interrelated economic outcomes. The consistency of evaluation results across these various measures (or lack thereof) tests the robustness of the econometric findings.

13.5.2 Data sources

The evaluation employs French company data from multiple sources which is matched to Bpifrance administrative data on participation in ISIP and other Bpifrance programmes. The two primary external data sources used are the governmental GECIR database which contains administrative data on company declarations related to expenses eligible for the Research Tax Credit and data from the Ministry of Education, Higher Education, Research and Innovation (MEHRI) survey of firms' R&D spending. However, the GECIR does not cover all RD&I activities while the MEHRI survey data has poor coverage of microenterprises. Thus, the evaluation also conducts supplementary analyses using the French statistical agency's Annual Social Data Declarations (DADS) in order to investigate ISIP's impact on proxy measures of RD&I employment (employment of engineers and technicians, as well as highly qualified employment more broadly).

13.6 Main findings

The evaluation finds that, on average, receipt of ISIP funding is associated with a €250,000 increase in a company's overall RD&I expenditure over the three-year period after the funding was received. This includes private RD&I expenditure as well as funding from public grants. Importantly, receipt of ISIP funding was not found to be associated with any change in a company's private investment in RD&I over the same period. Put differently, ISIP funding does not change companies' level of private RD&I investment but rather supplements it.

With regard to economic performance, the evaluation finds that receipt of ISIP funding is associated with an increase in companies':

- Turnover, export turnover and value added.
- Total employment, employment of engineers and technicians, and employment of highly qualified workers.
- Physical investments.

- Use of bank financing and private equity financing.

13.7 Analysis of methodological suitability and effectiveness

The methodology employed in this evaluation has a number of strengths.

- The high-quality DiD analysis with propensity score matching allows the evaluation team to draw conclusions about the questions of additionality and, to a lesser extent, attribution surrounding ISIP.
- Investigating ISIP's effect on a range on interrelated indicators of company economic performance demonstrates the robustness of the econometric results.

This methodology also has several limitations.

- While the evaluation provides general insights into ISIP's effectiveness, it does not provide insights into the causal mechanisms underlying the programme's impact. Put differently, it does not provide insights into the specifics of how ISIP funding translates into improved company-level economic performance. In particular, the evaluation does not differentiate between different ISIP instruments and cannot provide insights into the differential effects they might have.
- Matching programme beneficiaries to non-beneficiaries is not always computationally possible. In the case of this evaluation, the report states matching was successful for VSEs and SMEs. However, it was not possible for Entreprise de Taille Intermédiaire[27] due to the fact that this is a small population. This category of companies was therefore excluded from the analysis.

13.8 Transferability and context

The methodology employed in this evaluation is transferable to evaluating programmes designed to support RD&I among UK enterprises, given the following conditions are met:

- The number of programme beneficiaries is sufficiently large to conduct a DiD analysis.
- Company-level data on RD&I expenditure and economic performance is available for programme beneficiaries and a sufficient number of comparator companies for both the pre-intervention and post-intervention periods.
- Data is available on the participation of programme beneficiaries and comparators in key RD&I programmes similar to that being evaluated.

27 Intermediate-sized companies, defined as those with 250 to 4,999 employees and a turnover that does not exceed €1.5 billion, or a balance sheet total which does not exceed €2 billion.

13.9 Conclusions

This evaluation provides a methodologically rigorous analysis of whether ISIP funding impacts companies' RD&I expenditure and economic performance. It employs a DiD research design with propensity score matching to provide strong evidence that ISIP funding neither substitutes nor increases firms' private RD&I expenditure, and that ISIP funding is associated with improved economic performance at the company level. A key limitation of the methodology used in this evaluation are that it cannot provide insights into how ISIP funding translates into improved company-level economic performance. Relatedly, it analyses all seven ISIP initiatives in an aggregate manner and therefore cannot provide insights into whether different ISIP instruments have differential effects. This methodology is transferable to the evaluation of large-scale programmes to support RD&I among UK enterprises, subject to availability of the data required for a DiD with propensity score matching analysis.

14 Assessing the collaboration and network additionality of innovation policies: a counterfactual approach to the French cluster policy

14.1 Overview

Key characteristics	
Country	France
Institution	Centre National de la Recherche Scientifique; Jean Monnet University
Type of RD&I	Collaborative R&D
Type of Intervention	Collaborative research and innovation grants Direct assistance
Evaluation challenges	Additionality; intangible benefits The key evaluation challenges are 1) Determining the causal relationship between businesses' membership in the Competitiveness Cluster Programme and their involvement in collaborative R&D projects and/or integration into R&D networks. 2) Measuring integration into R&D networks.
Evaluation methodology and methods	• Quasi-experimental (difference-in-differences; matching) • Social network analysis

14.2 Introduction

This study investigates the impact of the French government's Competitiveness Cluster Programme (CCP) on beneficiaries' involvement in collaborative R&D projects and their integration into R&D networks.

To determine the programme's impact, the study utilises a difference-in-differences (DiD) research design with propensity score matching. This quasi-experimental approach allows evaluators to address the challenge of determining programme additionality. As part of this analysis, the study quantifies R&D collaboration by combining patent analysis with concepts from the field of social network analysis. This enables the study to overcome the challenge of measuring participation in collaborative innovation and in innovation networks, two intangible collaboration outcomes that many RD&I programmes and policies aim to foster.

14.3 Background description of RD&I intervention

Established in 2005, CCP aims to boost the competitiveness of the French economy by supporting collaborative innovation, with a focus on networking, knowledge exchange and collaborative R&D projects between cluster members.

At the time the programme was established, it comprised 67 sectoral competitiveness clusters. Following the creation of new clusters and the merger of others over time, the number of clusters stood at 71 in 2014, and 54 in 2023. Each competitiveness cluster is an accredited non-profit organization financed through a mix of public funding and annual member subscriptions. The clusters operate in a range of economic sectors including energy, mechanics, aerospace, transport, information and communication technologies, health, environment, and ecotechnology. Government support for these clusters is largely comprised of:

- Partial funding of cluster governance structures.
- Direct financial supports to collaborative R&D projects emerging from clusters.
- Non-financial support mainly targeted at SMEs cluster members, including training and technical assistance.

Each cluster is regionally anchored, with its field of specialisation having been matched to the economic sectors located in its region and the research themes addressed by public research institutions in that region. Regional authorities have played a key role in the programme since its inception, particularly in terms of providing joint funding alongside national public funding.

According to the latest publicly available data, around 14,000 private establishments and 2,000 research institutions are members of the CCP. Over its lifetime, the programme has invested €7.5 billion Euros in a total of 12,000 innovation projects.

14.4 Challenges measuring outcomes and impact

The study addresses two key evaluation challenges: additionality and measuring the intangible benefits of RD&I programmes.

- Regarding the challenge of additionality, the impact evaluation must address the question of whether observed changes in CCP members’ involvement in collaborative R&D would have been likely to occur even if they had not participated in the programme. Given the systematic differences between CCP members and non-members, establishing programme additionality is challenging. As the evaluation report points out, CCP member businesses tend to be larger, have a greater share of highly qualified employees and are more likely to be in high-tech industries than non-members. The use of a high-quality quasi-experimental research design enables the study to overcome the challenge of determining additionality by enabling the construction of a credible counterfactual (control group).

- The evaluation quantifies network effects by adopting measures from the social network analysis literature as outcome measures in the quasi-experimental analysis. This is an innovative approach to capturing some of the intangible networking benefits of collaborative RD&I programmes that often go unmeasured.

14.5 Methodology and data sources

14.5.1 Methodology

This research paper aims to answer three questions:

- Does a business' participation in the CCP increase its collaborative innovations, as a proportion of its overall number of innovations?
- Does a business' participation in the CCP increase its regional collaborative innovations, as a proportion of its overall number of innovations?
- Does a business' participation in the CCP increase its centrality in innovation networks?

By investigating these questions, this study aims to improve understanding of how CCP membership shapes business' innovation activity.

While the research article frames its conclusions in terms of the CCP's general impact, it is important to note that only a specific subset of cluster members were included in this analysis: those which applied for patents to the French Patent Office over the period 2005-2010. This subset of CCP members comprises only 116 enterprises out of over 10,000 private enterprises which have participated in the programme. Importantly, it is highly unlikely that firms engaged in patenting activity are a representative sample of CCP members. Thus, the scope of these findings should be considered as limited to cluster members that have been successful in producing patentable innovations. Notably, it is well established that patenting activity is more common in certain economic sectors, such as the energy sector and industrial manufacturing. Research also suggests that businesses are less likely to apply for patents related to their process innovations as compared to their product innovations.

To construct a counterfactual for CCP members engaged in patenting activity, the researchers compiled an original dataset of all French enterprises that applied to the French Patent Office for patents over the period 2005-2010. Firms were identified either as CCP members or non-members, with non-members used as a comparison group to investigate the effect of CCP membership on collaborative innovation. However, the study finds that CCP member businesses differ from non-members in several ways, including that they tend to be larger, have a greater share of highly qualified employees and are more likely to be in high-tech industries. Thus, to construct a credible counterfactual group, the study employed a propensity score matching technique to balance the distribution of CCP members and non-members on these observed characteristics, thus mimicking a randomised experiment.

Matching techniques, however, do not address the fact that unobserved firm characteristics may bias the comparison between the two groups. For this reason, the study combines propensity score matching with a DiD research design. The DiD approach considers the difference in patenting activity between CCP members and non-members before the establishment of the programme (the pre-intervention period), and the difference between the two groups three years after the programme's establishment (the post-intervention period). It then takes the difference between these two differences. DiD estimation is commonly used for evaluating nonrandomized interventions. This is because it removes biases in post-intervention comparisons between the treatment (CCP member) and control (comparison) groups that may result from unobserved differences between these groups, in addition to biases from comparisons over time in the treatment group which result from trends due to other causes of the outcome.

The study estimates the effect of CCP membership on four outcomes measures (dependent variables):

- The "co-invention rate", which represents the collaborative share of a business' total innovations. This is defined as the number of co-invented patent applications associated with an individual business. Total innovations are defined as the overall number of patent applications associated with the business.
- The regional co-invention rate. This is defined as the number of co-invented patent applications associated an individual business, where at least one of the co-inventors is located in the same region as the business.
- The size of a business' innovation network. This is defined as the number of direct collaboration partners that an individual business has.
- The centrality of a business in the overall innovation network. This employs a concept from social network analysis called "betweenness centrality" which measures an entity's centrality in an overall network.

Finally, the analysis includes control variables for whether a business: 1) has a history of collaborative patenting activity in the pre-intervention period, and 2) is a member of the EU's Framework Programme, a key policy instrument to support medium- to large-sized collaborative research projects in Europe.

14.5.2 Data sources

The researchers compiled an original dataset of all French enterprises which applied to the French Patent Office for patents over the period 2005-2010. For each enterprise, the dataset included information regarding:

- **Collaboration networks**. Information on patent applicants and inventors listed in the primary patent data was used to construct variables about collaboration networks.

- **Company location**. The addresses of patent applicants and inventors listed in the primary patent data was used to geolocate patents.
- **Company characteristics**. The number of employees as well as their qualifications was obtained from the French statistical agency's Annual Social Data Declarations and linked to the company patent data.
- **CCP membership**. CCP membership was identified using CCP administrative data and linked to the company patent data.
- The study also employs data on EU Framework Programme membership and company sector. However, the paper does not provide information on the data sources used to obtain this information.

14.6 Main findings

The study finds that CCP membership has a positive effective on a business' likelihood of collaborating on its R&D initiatives. It estimates that, on average, CCP membership leads to a 4.1 percent increase in a business' co-invention rate. This increase was much higher (6.3 percentage points) for CCP members who had little involvement in collaborative R&D projects prior to joining the programme.

The study further finds that CCP membership has no significant effect on regional collaboration. This finding is consistent across both those CCP members who had little involvement in collaborative R&D projects prior to joining the programme, and those who had greater involvement.

Regarding the effect of the programme on business' RD&I collaboration network, the study also finds no significant effect. Again, this finding is consistent across both CCP members who had little involvement in collaborative R&D projects prior to joining the programme, and those who had greater involvement. Given the positive effect of the programme on business' co-invention rate, it may be that CCP members collaborate with a fixed group of partners across their different projects.

Finally, a sensitivity analysis reveals that the research findings are not highly sensitive to the presence of unmeasured confounding variables.

14.7 Analysis of methodological suitability and effectiveness

The methodology employed in this evaluation has a number of strengths.

- The construction of a rich dataset on patenting activity among CCP members and comparator businesses. As discussed in the methodology section, this dataset draws on several data sources, including patenting activity data from the French Patent Office which is matched with: 1) administrative data for the CCP and the EU Framework Programme; 2) data on individual establishments' characteristics collected by the

French statistical agency. In addition, each patent application in the dataset was geolocated to enable the analysis described in the methodology section.

- The use of the “betweenness centrality” concept from the social network analysis literature, which allowed the study to quantify the programme’s effects on the RD&I network of participating businesses.
- In general, the use of matching methods can be problematic due to the fact that the use of different matching procedures can sometimes yield entirely different research results. This study provides a transparent description of the matching procedure employed. This helps address concerns about the matching methodology and is a good practice to be emulated. Just as important is the fact that the study demonstrated consistency between the primary matching procedure used in the study and a second matching procedure. Such robustness tests are also a good practice to be emulated.

This methodology also has several limitations.

- As briefly discussed in the methodology section above, there are well-established limitations to the use of patenting activity as a proxy measure for innovation. Most notably: 1) All innovations do not lead to a patent application; 2) Not all patents applications refer to an innovation; 3) The value of patents can differ greatly; 4) The use of patents varies widely according to economic sector.
- Related to the above point, and as discussed in the methodology section, only a specific subset of cluster members were included in this analysis: patent applicants. This subset of CCP members comprises only a small proportion of CCP member firms. Importantly, it is unlikely that CCP member firms engaged in patenting activity are a representative sample of CCP members.
- The DiD estimator used in the study is imprecise. For simplicity, it divides the sample into a single pre-treatment and post-treatment period, when in reality businesses joined the programme on an ongoing basis. While this approach simplifies the statistical analysis, it also biases the statistical findings. A preferable approach would be to use a staggered DiD estimator, which accounts for the fact that businesses are exposed to the treatment at different time periods.
- The study does not present a formal test of the parallel trends assumption, a key assumption that underpins the validity of DiD research designs. This is perhaps due to the fact that the use of matching techniques helps address violations of the parallel trends assumption. Nonetheless, it is good practice to present a formal test of the parallel trends assumption in order to allow readers to formulate a judgement on the strength of the evidence presented in the study.

14.8 Transferability and context

The methodology employed in this evaluation is transferable to the evaluation of some RD&I programmes in the UK context.

- The methodology would be suitable for measuring the impact of large-scale collaborative RD&I programmes which are expected to result in a significant increase in collaboration on patentable innovations. It is worth noting, however, that the construction of the required dataset is a labour-intensive endeavour.
- In principle, it would be possible to apply a similar approach to measure the impact of collaborative RD&I programmes that are not focused on patentable innovations. However, this would require fielding an original survey of UK businesses using a sophisticated survey instrument designed to measure non-patentable RD&I collaboration. Such an exercise would require considerable financial resources.
- The “betweenness centrality” measure could be applied to bibliometric data on the authorship of academic research papers to evaluate the effect of collaborative research programmes on the network positionality of participating academic researchers.

14.9 Conclusions

This study provides a methodologically rigorous analysis of the impact of CCP participation on patentable RD&I collaborations, employing high-quality data to provide strong evidence that the programme increases such collaboration. However, the study also finds that participation in the CCP does not increase patentable RD&I collaboration at the regional level, nor does it have an effect on CCP members’ RD&I network. The key limitation of the methodology used in this evaluation is that it cannot capture CCP effects on non-patentable innovations.

This methodology is transferable to the evaluation of large-scale collaborative RD&I programmes in the UK context which are expected to result in a significant increase in collaboration on patentable innovations. Elements of the methodology could also be used to evaluate the effect of collaborative research programmes on the network positionality of participating academic researchers.

15 World Class Ecosystems in the Finnish Economy

15.1 Overview

Key characteristics	
Country	Finland
Institution	Business Finland
Type of RD&I	Applied research and innovation; Mission funding (secondary)
Type of Intervention	Research and innovation grants; Networking activities: support to innovation ecosystems
Evaluation challenges	Lagged effects; Additionality The two key challenges are measuring the additionality (across four areas: input, behavioural, output, and wider impacts) and the long-term effect of the policy promoting ecosystem development
Evaluation methodology and methods	• Theory-based (Contribution analysis; web scraping; case studies) • Economic modelling

15.2 Introduction

This case study looks at the evaluation of the World Class Ecosystems in The Finnish Economy between 2019-2021. The programme provides funding for orchestration activities and capital loans, seeking to support collaboration and network dynamics, and creating synergies with other policies offering direct support to R&D activities to companies.

This evaluation deals with two main challenges in measuring impact. The programme additionality and lagged effects or this public investment. Additionality is assessed across four main areas (input, behavioural, output, and wider impacts) to disentangle the added value of the ecosystems policy. The lagged effects are addressed via economic modelling to overcome the short timeframe of the policy implementation.

15.3 Background description of RD&I intervention

Business Finland is the public agency for innovation funding and trade, travel and investment promotion, supporting companies to grow at a global level.[28] Ecosystems policy is key to achieving this. Its primary objective is to: “Create new 20 billion-euro ecosystems and strengthen existing ecosystems in Finland to drive economic growth.

The programme was launched in 2019 with the initial purpose of funding five new business ecosystems per year. Through its roll out, Ecosystems have gone through adjustments responding to emerging challenges, such as the adoption of a funnel approach to track progress and provide differentiated support to ecosystems according to their development stage and facilitate progress evaluation. Over time, the number of supported ecosystems also surpassed the initial target.

The rationale for intervention is to overcome system failures, to address higher-level coordination issues, create new sources for economic growth and revitalise traditional industries[29]. The underlying reasons for adopting an ecosystem policy perspective are:

- The challenges of renewal, internationalisation and SDGs can be better addressed by companies working collectively, as these challenges ask for more dynamic knowledge generation and exchange, increased public-private collaboration, stronger involvement of end-users as well as the use of co-creation methods.
- Together, companies can take on larger challenges and risks in terms of innovation, entrepreneurship and internationalisation.
- By departing from an analysis of shared challenges, a shared strategy can be defined to support effective collaboration within these ecosystems
- Collectively, they can build their capacity, meet new partners, work towards industrial transformation and address relevant societal challenges.
- Ecosystems can support both formal activities, often organised by a central player or neutral body within the ecosystem, as well as informal activities between members through networking in the ecosystem.

The programme’s value-added lies in the support to collaboration and network dynamics, its potential to create **synergies with other policies** aimed at supporting companies directly (e.g. instruments promoting innovation and internationalisation, funding and services and the Business Finland Programmes[30]) and those seeking more transformational change (e.g. promoting sustainability and industrial transformation).

28 https://www.businessfinland.fi/en/for-finnish-customers/strategy

29 (Technpolis Group and 4Front, 2021, p.30)

30 Ibid, p. 31)

Instrument

The ecosystems policy has used two **Growth Engines instruments** to support ecosystem development and functioning. One focuses on support for coordination and orchestration and the other relates to capital loans for platform companies.[31] The support for ecosystem development and functioning is aimed at improving aspects like governance and coordination within ecosystems.

- Orchestration funding: fosters innovation cooperation and facilitates joint activities of ecosystem members, with a 50% cap for funding the overall project activities. The instrument targets private companies, associations, and foundations, with exceptional cases made for research organisations and public bodies.
- Capital loan: provides loans between €2 million and €10 million with a specified repayment schedule and interest rates, targeting platform companies, seen as anchors for innovation and competitiveness around new innovations that can trigger the development of new ecosystems or their transformation.

At the time of the evaluation (2021), 13 ecosystems had been supported by the Growth Engines instruments, of which three were at an explorative stage, five at the experimental (birth/start-up) stage and another five at expansion and growth level. The period of analysis of this evaluation is 2018-2020, two years after the launch of the programme. Business Finland envisaged a periodic evaluation, on a bi-annual basis.

15.4 Challenges measuring impact

Two main impact-related challenges are identified in this evaluation, additionality and lagged effects of investments.

Additionality of the Ecosystem Policy in four main domains:

- Input additionality: the effects of resources put into the ecosystem policy. This concerns both the available financial and non-financial resources and capabilities of Business Finland as well as those available within the ecosystems (human resources, orchestration, etc.).
- Behavioural additionality: to capture the change in the processes of companies and ecosystems as a whole as a result of policy stimulus.
- Output additionality: results that are realised due to the ecosystem policy.
- Impacts on the Finnish economy and society

Lagged effects. In many cases the ecosystems are still in development, therefore company-level results have not been realised yet for all companies.

31 Ibid, p. 41

15.5 Methodology and data sources

15.5.1 Methodology

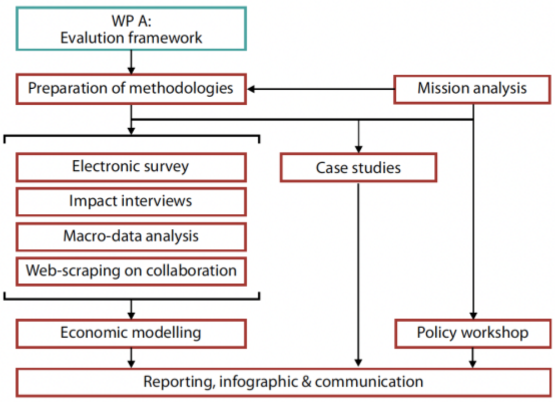
The analysis began with a definition of the ecosystem concept that sets up the actors and type of functions and resources facilitated/mobilised by ecosystems. A horizon of 10 years was used to frame the four stages of development of ecosystems named, emergence, start-up, growth/expansion and maturity/renewal. The evaluation points out the complexity of assessing ecosystems when using company results (turnover, export, employment, etc.), since ecosystem functions alongside direct and collaboration interventions can influence companies. This is difficult to disentangle and makes it more difficult to assess additionality.

The main evaluation questions are:

- How can the public sector in Finland improve its ability to build ecosystems to attract global actors to Finland?
- What has been the main value added of the funding and services of Business Finland for promoting business ecosystems in the Finnish economy?
- What kind of critical obstacles and bottlenecks have affected, or could affect, the ability to achieve these goals?
- What kind of societal impacts (renewal of economy, environment, well-being, capabilities, company growth, ecosystems) have been achieved and how they could be measured?

A representation of the methodologies used to answer these questions is presented in Figure 1. The evaluation was split into two work packages. Work package A carried out a mapping of high-potential ecosystems, identified potential thematic areas and developed an evaluation plan. Work package 2 undertook an impact study.

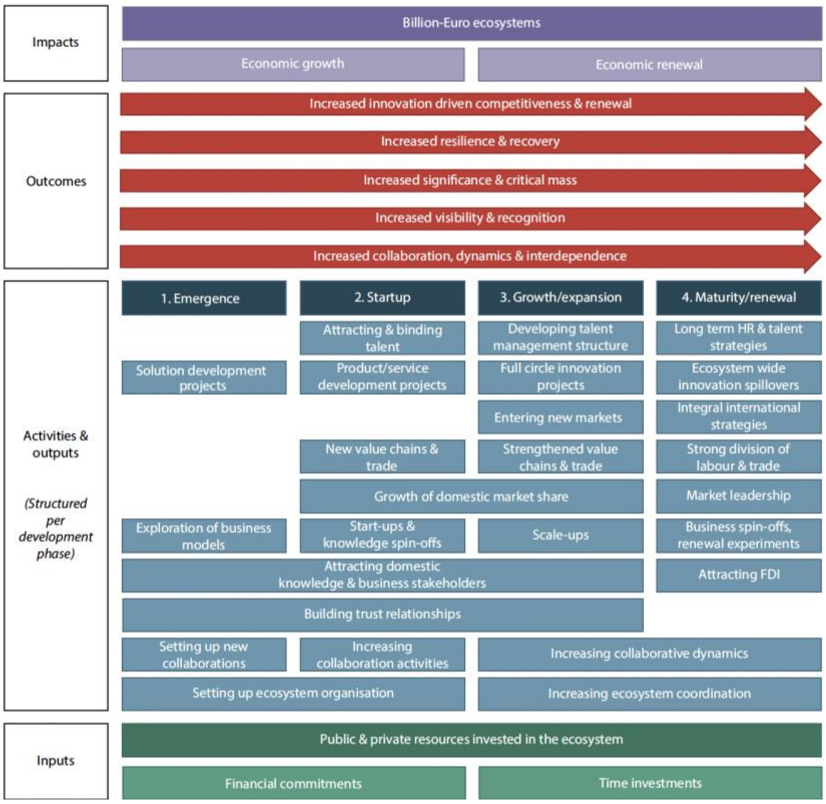
Figure 1: Methods used to evaluate the impact of Ecosystems



Source: Technopolis and 4Front, 2021

The evaluation framework developed in the work package A is summarised in Figure 2. The economic modelling uses regression analysis to estimate the effect of Growth Engine Support on turnover, employment and exports, with two scenarios one based on past turnover growth, and another based on past turnover growth plus the contribution of public policy.[32]

Figure 2: Evaluation framework of ecosystem policy



Source: Technopolis and 4Front, 2021

This impact assessment was carried out in three steps measuring the effects of ecosystems on companies, the results along the impact pathways, and the economic impact.

The evaluation used three other methods to measure additionality:

- Self-reflection of companies, based on guided estimations
- Differences between members within ecosystems, based on the level of engagement of ecosystem members
- Differences between ecosystems, based on the level of maturity of ecosystems

To address the lagged effects the evaluation carries out an economic modelling, with assumptions to set up the conditions under which the models would apply.

15.5.2 Data sources

- Insights from the WPA report and policy landscape analysis were inputs used to develop the evaluation framework of WPB

32 Report, p. 160

- Questionnaires developed for the electronic survey launched in October 2020 with 75 respondents
- Impact interviews to inform the development of four detailed case studies
- Statistical data at the company level regarding employment, turnover and export to carry out the economic modelling
- Survey data were used to enrich this data for a sample of the companies. This sample was used to extrapolate some of these findings for the entire data set.
- For the network analysis: the URLs of the domains (websites) of the ecosystem participants were collected as a starting point for the webscraping algorithm. For the purpose of the study, “html” and “pdf” data extractions were considered (searched) as these are common and relevant data types for the network analysis.
- Policy workshop was held to discuss the findings with representatives of Business Finland and the Ministry of Economic Affairs and Employment (TEM).

The main limitations of the methodology and process of the study process are:

- Contact details of company ecosystem members were not readily available, as a result, a lot of effort had to be put into retrieving responses, increasing the likelihood of response bias: respondents with a positive attitude will be more likely to participate in surveys, interviews.
- The available data had no overlap with the public intervention. Many interventions took place around 2018 and statistical data often lags a few years
- No control group was available, therefore isolating effect was more challenging
- No data was available on real collaborations, therefore webscraping was used to gather data on collaboration based on referrals between companies on their website

15.6 Main findings of the evaluation

Despite a clear general objective, its operationalisation into actions and results appeared more challenging. This is found to be related to the absence of indicators showing ecosystems’ journey across development pathways, against the use of more general KPIs, mainly focused on individual actions.

The evaluation found mixed results due to the heterogeneity of ecosystems, disparity in development levels, and different degrees of engagement across ecosystems’ members, which made it too early to determine the economic and social impact of the ecosystem policy.

On the other hand, the evaluation found that ecosystem members show a very strong innovation and growth profile, and for some, it is possible to see well-organised governance models and strong network ties. An interesting finding was statistical signs that the length and depth of ecosystem participation of companies is associated with stronger company growth.

Findings related to additionality:

- The policy framework fits the system failure being tackled by the policy but not completely its implementation, as the instruments consist mainly of financial support, and an important number of ecosystems are still in the early stages of development that require non-financial ecosystem services.
- In terms of behavioural additionality, results are also mixed. Some ecosystems show dense networks and strong operations, but about half of the members have a low attachment and engagement/commitment to the ecosystem.
- Output additionality is found although to a limited extent. Participants have shown strong employment and turnover growth. The overall turnover of ecosystems is very large, but it was not deemed to be attributed to the ecosystems but rather to other activities in which companies are involved. Still, a positive correlation was identified between economic performance and the participation of companies in the ecosystems. This appears to hold true for those companies being involved in the ecosystem for longer periods, who can reap the benefit of ecosystem functions.
- Economic impact. Although setting up an ambitious target can mobilise organisations around a vision, there is a need to break this down into measurable objectives. Given the strong performance of firms in some ecosystems, the €20 billion target appears attainable.
- Innovation is the key activity of ecosystems driving business growth rather than industrial transformation. This was identified via turnover sources, coming from the expansion of business activities rather than the replacement of existing ones.

15.7 Analysis of methodological suitability and effectiveness

The methodological approach is clearly explained and articulated to respond to the evaluation questions outlined. These are thought to determine to what extent the ecosystem policy has achieved its main objective. The timeframe of the evaluation, two years after the launch of the programme, posed significant challenges in terms of data availability and evaluation impact, as the development of ecosystems is a long-term process.

Organised in two work packages, this evaluation is well articulated and complete. With a thorough evaluation framework used in the initial phase, which embraced the complexity of ecosystems, their different stages of development and therefore, a variety of outputs and outcomes. The impact evaluation builds on the first phase and utilises valuable inputs.

Main challenges in measuring impact are also explained and framed around the programme additionality and lagged effects or this public investment. The core of the analysis around additionality and its segmentation into four main areas (input, behavioural, output and wider impacts) is relevant to scrutinise effects and disentangle the added value of ecosystems policy. This is also complemented by the landscape review that provides further insights into overlapping policy efforts and opportunities for improving the instrument.

Regarding lagged effects, the evaluation develops an economic modelling to estimate the economic impact of the programme, to overcome the short timeframe of the policy implementation - over two years. Assumptions and further elements considered in the modelling are well explained and relevant. Nevertheless, the instable context[33] that dominated the period of analysis reduces its strength.

A less discussed challenge in the evaluation deals with the skewness of impact, which may have a play in those ecosystems exhibiting poor performance, hence favouring stronger companies and reinforcing the effects on a small number of successful and well-connected companies.

15.8 Transferability and context

The methodological approach could be adopted to evaluate similar policy interventions seeking to support R&I ecosystems in a wide range of contexts given that it relies significantly on primary data, such as surveys, interviews and a policy workshop. Monitoring and evaluation data (not always available in all contexts) was less relevant for this evaluation – although pointed out by the evaluators as a desirable input to strengthen the evaluation results. On this, capturing interactions within ecosystems would be desirable to carry out more meaningful Social Network Analysis.

Supporting R&I ecosystems may take different forms and emphasise different aspects of collaborative work, from networking to collaborative R&D. Thus, this methodology could be reproduced for different types of interventions aimed at strengthening innovation ecosystems in the UK, considering two main factors. Access to company data before and after the period of evaluation and availability of primary sources of information to support case studies.

Case studies build on survey results and interviews, which can be flexible to capture particularities of ecosystems and policy intervention. Regarding the economic modelling, access to, for example, the UK Business Enterprise Research and Development (BERD) dataset from the Office for National Statistics would provide the data to carry out the economic analysis.

The landscape review and the work entailed in the work package A, are normally seen as part of standard evaluations in the UK[34]. This is to design an evaluation framework and develop the logic model of the intervention to determine its effectiveness in delivering the expected outputs and outcomes using the inputs and activities entailed by the programme.

33 The conditions under which the ecosystem policy were implemented were quite turbulent given the Tekes and Finpro merger in 2018 and the COVID19-crisis from Spring 2020 onwards.

34 Technopolis has conducted recently evaluations for the Royal Society and its International Collaboration Award, the Fund for International Collaboration FIC for the UKRI and the Socio-Economic impact of Mathematical Science Research for ESPRC including baseline evaluations and landscapes reviews.

15.9 Conclusions

This case study analysed the World Class Ecosystems in the Finish Economy programme. It reviewed the methodological approach and evaluation results and found the methods well-suited to meet the objective of the evaluation and to respond to the evaluation questions. The methods selected are appropriate to ensure triangulation of results and were relevant for addressing the prime impact evaluation challenges, establishing the additionality of the policy and dealing with the time lag of the results derived from the investments.

In this line, the methodology is robust and provides valuable insights to strengthen the policy intervention, offering recommendations about the segmentation of beneficiaries, improving data collecting to facilitate monitoring and evaluation and clarifying objectives and potential indicators for measuring ecosystem functioning and development, which were not clear at the beginning of the evaluation.

The methodology assembled appears highly transferable to the UK context considering that it is based on primary data that could be collected also in the UK by evaluators.

16 The Value of CSIRO: The broader impact of CSIRO's portfolio of activities

16.1 Overview

Key characteristics	
Country	Australia
Institution	CSIRO
Type of RD&I	Research and Innovation Grants
Type of Intervention	Challenge / Mission-oriented funding
Evaluation challenges	Lagged effects and low observability Two key challenges associated to the lagged effects of basic research funding and investment in national facilities, and the low observability of research impacts in the six challenge areas that entail behavioural change and adoption of new practices
Evaluation methodology and methods	• Benefit-cost analysis • Generic methods (Case studies)

16.2 Introduction

This case study analyses the impact evaluation of the Commonwealth Scientific and Industrial Research Organisation (CSIRO) portfolio of activities between 2010-2022. The portfolio encompasses R&D activities around challenge areas and CSIRO's three main business units, impact science, national facilities and collections and CSIRO services. This evaluation addresses two main challenges, the lagged effects of basic research funding and investment in national facilities, and the low observability of impacts is especially noted around the six challenge areas, as some may involve changing behaviours and practices, as a result of research.

To address these challenges, the evaluation adopts a cost-benefit analysis of the portfolio of investments over a 25-year span and analyses 112 case studies to capture pathways to impact and low observable outcomes.

16.3 Background description of RD&I intervention

The Commonwealth Scientific and Industrial Research Organisation (CSIRO) is Australia's national science agency, which coordinates and works collaboratively across the innovation ecosystem to solve the country's greatest challenges through science and technology.

Challenge areas are food security and quality, sustainable energy and resources, health and wellbeing, resilient and valuable environments, future industries, and a secure Australia and region. For each area, CSIRO offers support to basic research, applied research, technological development and production.[35] Thus, CSIRO funds and maintains science infrastructure and collections for public use, promotes science and technology adoption and patenting, funds entrepreneurial projects, spinouts and start-ups, and supports SME programmes. Equally important is the formation of scientific talent. These constitutes CSIRO's portfolio.

CSIRO operation is organised in three main business units[36]:

- Impact science: with nine national research business units focused on the biggest challenges facing the nation
- National facilities and collections: managing infrastructure and biological collections
- CSIRO services: commercial customer-centred products and services for government, community and industry

CSIRO's annual report for 2021-2022 states an annual budget of AUD $10.2 billion. These resources are channelled through the Business Units, responsible for the refinement and execution of programmes. Each programme has between three to five impact statements and are responsible for managing the impact pathways of their portfolio of projects. Research outcomes are captured via case studies alongside benefit-cost analysis and assessed bi-annually as an increasing annual portfolio.

This evaluation includes 112 case studies of published case studies between 2010-2022 on research and infrastructure outcomes. It involves assessments of the benefits and costs of research beginning mostly from 2000 and breaks down into 63 coming from the 2020 value report and 49 new studies. For the synthesis, only cases with complete costs and benefits data for at least 25 years (with projections capped at 10 years after completion) were include in the synthesis, for a compilation of 68 cases to assess their impact. For example, the High Pressure Processing programme started in 1998 and was completed in 2018 when its case study was published. The period of analysis for the benefit-cost evaluation was therefore 1998-2028 and appeared in both the 2020 and 2022 CSIRO's reports.[37]

35 p. 8

36 https://www.csiro.au/en/about/We-are-CSIRO

37 https://www.csiro.au/-/media/About/Files/Impact-assessment/2018/2018-High-Pressure-Processing--Case-Study---PPE---pdf.pdf

16.4 Challenges measuring impact

The evaluation faces two main challenges in assessing impact. Lagged effects of the investments, particularly those related to infrastructure investments and more basic research. Secondly, the low observability of impacts is especially noted around the six challenge areas, as some may involve changing behaviours and practices, or for example, the preservation of the Great Barrier Reef as a result of research and improved control measurements on the Crown of Thorns Starfish that preys on coral.

Another impact-related challenge is the attribution of the benefits, which is addressed at an earlier stage when case studies are elaborated. CSIRO attributes efforts based on cost sharing, proportional to the participation of organisations in the generation of outputs[38]. In this way, only CSIRO's attributed costs are considered for the calculation of the cost time series being aggregated at the portfolio level, ensuring consistency of the measure.

16.5 Methodology and data sources

16.5.1 Methodology

CSIRO uses primarily case studies and benefit-costs analysis to assess the impact of their research outcomes and investments, which can be at the project, programme, business unit level or as inn this report at the portfolio level.[39] The value of CSIRO is assessed bi-annually with an increasing annual portfolio of externally performed and/or validated impact case studies.

This evaluation seeks to assess the value delivered by SCIRO to Australia. Specific objectives of the report are to:

- Review case studies describing the impacts of CSIRO's technology development and innovation programs,
- Synthesise the monetised economic impacts described therein,
- Compare monetised benefits with costs to estimate Australia's return on investment in CSIRO's technology development and innovation portfolio
- Compile qualitative and quantitative metrics of non-monetised elements of CSIRO's impact.

38 According to the impact evaluation guide this method is useful for those projects where costs distribution is well proportionally distributed amongst participants but requires adjustments when that's not the case.

39 CSIRO has designed an impact evaluation guide to assist business units, managers, and innovation ecosystems users in their evaluations https://www.csiro.au/en/about/Corporate-governance/Ensuring-our-impact/A-CSIRO-wide-approach-to-impact

The report aggregates case study values over years to calculate the Present Value of costs and benefits, the Net Present Value NPV and the Benefit-Cost Ratio BCR.[40] Impact case studies are developed by CSIRO's internal impact evaluation team or commissioned by CSIRO and undertaken by ACIL Allen, ACIL Tasman, the Centre for International Economics CIE, Deloitte Acess Economics DAE, RTI, or Tractuum.

To ensure consistency across case studies Benefit-Costs calculations, the evaluation starts verifying the accuracy of data, and methods used for each case and standardising them, doing inflation and discounting adjustments for each year of data, using Australia's Consumer Price Index CPI. Following the Impact Evaluation Guide (2020) the benefit and costs time series are discounted using the 7% real social discount rate. Both discounts and adjustments are made using 2022 as the base year.

The main source of costs is the funding provided by CSIRO and the estimated benefits attributable to CSIRO. The evaluation does not review individual assumptions of cases. Nevertheless, CSIRO carried out a review of three cases for the 2017 Value of CSIRO report and found their assumptions robust and conservative. [41]

For the portfolio, cases were limited to those started within the past 25 years. As part of the analysis, the evaluation compares project versus realised values at the time of publication for each case and to improve accurateness, the evaluation capped benefits and cost projections beyond 10 years. Although, realised values from costs and benefits are not limited. The resulting dataset is used to estimate the return on investment of the funded research.

The present value (PV) of CSIRO's research benefits is aggregated across case studies and the same process is done to calculate the PV of their costs. The Net Present Value NPV is obtained by subtracting the PV of costs from the PV of benefits. The Benefit-Cost Ratio BCR is calculated by dividing the PV of benefits by the PV of costs. These estimates provide a moving average of the value of CSIRO research. Thus, adding new cases to the portfolio is expected to increase confidence in the estimates.

The report acknowledges that wider impacts can't be captured in monetary terms. Particularly, around knowledge spillovers, environment, and wellbeing. To demonstrate the impact that cannot be monetised, the evaluation includes short cases across the six challenge areas of CSIRO: health and wellbeing, food security and quality, secure Australia and region, resilient and valuable environments, sustainable energy and resources and future industries. For this purpose, the report collates excerpts of case studies produced in the last 2 years and generates one-page vignettes for each challenge area.

40 By dividing the PV of benefits by the PV of costs.

41 Report p. 9

Alongside this compilation, for those investments, whose impact is not readily monetised, the evaluation offers additional impact metrics around research translation, use of science infrastructure and collections and stimulating innovation.

16.5.2 Data and use of monitoring and evaluation inputs

- 112 case studies published or commissioned by CSIRO from 2010 – 2022, covering research initiated between 1965 and 2022, with most cases starting in 2000
- Time series of costs and benefits of projected and
- Australian Consumer Price Index CPI for1998-2022

16.6 Main findings

The evaluation estimated a research portfolio Benefit-Cost Ratio of 8.4 to 1, with an increase from the previous evaluation estimated at 7.6 in 2020. Case studies in the six challenge areas showed higher BCRs, which confirms their priority and relevance. A breakdown review of BCR per case evidenced that those exhibiting the highest BCR were the ones with the lowest costs, mostly related to high-impact software.

Case studies with negative returns (after 10 years of projections) are assumed to reflect the lagged effects of impact and long-term investment necessary to develop new technologies or the underlying high risk of conducting more cutting-edge research.

The new cases added an NPV of AUD$7.9 billion to the AUD$12.0 billion NPV of the portfolio in 2020. Together with the increased BCR of 8.4, the research funded between 2021-2022 shows higher returns. For the evaluator, BCR is lower bound as wider environmental and social effects are difficult to monetise and therefore, are not captured by this method.

To address this issue and surface those impacts difficult to observe and measure (intangibles) via benefit-cost analysis, the evaluation selects excerpts of case studies in the six challenge areas to showcase their social, economic and environmental impact. A total of 12 mini cases are included in the evaluation, two per challenge area demonstrating further social, environmental and wellbeing impacts. For example, in changing behavioural practices, allowing access to difficult-reaching vulnerable populations. Each of these cases had been developed individually to establish their performance and impact on Australia’s society.

Against its objectives, this evaluation meets its primary purpose of demonstrating the benefit of CSIRO’s portfolio of research to the Australian society. The report’s objectives are met using a standard methodology to assess the economic benefits and costs of multiple interventions in aggregate over time, providing further insights at the portfolio level via moving averages.

16.7 Analysis of methodological suitability and effectiveness

CSIRO evaluates the economic impact of its research mainly through benefit-cost analysis. Following the evaluation guide developed by CSIRO, outcomes of projects, programmes and business units are synthesised in case studies. This evaluation reviews and aggregates the economic benefits derived from the portfolio of research funded by CSIRO since 1998. It consists of 112 case studies from which 68 are aggregated to calculate the benefit-cost ratio of these investments. The wider effects and non-monetary impacts of this research is showcased by 12 case studies excerpts in the six challenge areas.

The evaluation provides evidence of the impact being generated across the three main business units and their portfolio of R&D investments. The selection of the benefit-cost analysis as a standard way to capture the economic impact of interventions allows to compare different types of interventions over time.

Lagged effects are addressed by using projected costs and benefits for case studies capped at 10 years to reduce uncertainty of future costs and returns[42]. The portfolio includes examples of big investments (Pawsey Supercomputing Centre) whose BCR is expected to be positive in around 30 years but is negative by the time of the report.

To address low observability of impacts, the report compiles excerpts of the original case studies to highlight social, environmental and economic impacts that cannot be monetised, but have contributed to addressing the challenges of the six priority areas. These benefits have been captured following the impact evaluation guide provided by CSIRO, using a logic model.

16.8 Transferability and context

Two evaluation practices outstand in this case and define the transferability of the methods used herein. The availability of an Impact Evaluation Guide used across the whole R&D system to assess the outputs and impact of this type of investments. Secondly, the standard use of Benefit-Cost Analysis and Benefit Costs Ratios as indicators of impact at individual level that can be aggregated to determine overall impact of CSIRO's portfolio.

This has implications in terms of monitoring data collection as this is done from the start of each project and under the specific process to facilitate aggregation and standardisation in future portfolio assessments.

Centralised evaluation, as done by CSIRO, may present challenges to the more decentralised R&I system in the UK and less standardised practices around monitoring and evaluation, for example, in terms of tracking the associated costs of initiatives.

42 Usage and adoption costs from benefits are subtracted from the benefits but are not included in the cost basis Guide to evaluate impact p. 30

16.9 Conclusions

Analyse the relevance of the methodology to the UK RD&I system and conditions under which this can be used for RD&I interventions going forward.

This evaluation assesses the value and impact of CSIRO research between 1998-2022 using 112 case studies and benefit-cost analysis at the portfolio level. It builds on individual evaluations of programmes previously undertaken in the form of case studies, for which benefit-cost analyses have also been carried out.

CSIRO's portfolio evaluation adopts a Benefit-Cost analysis as a way of capturing the economic impact of its research and facilitating comparability between different programmes and investments. The evaluation deals with the lagged effects of these programmes using projected costs and benefits – in some cases available for more than 30 years – capped at 10 years. These figures are updated annually with actual values adjusting the portfolio estimates. To showcase non-monetary and non-readily monetised impacts a collection of case study extracts from original case studies are presented.

This evaluation adopts a more general approach, using Benefit-Cost analysis, case studies and other output indicators that will lead to further impacts in the future. It is underpinned by a consistent process for capturing the costs and benefits of the research, as stated in the impact evaluation guide. Its transferability is dependent on monitoring and evaluation of data availability in the UK and the possibilities to standardise and aggregate them at a portfolio level.

17 Business Research and Innovation Initiative Impact Evaluation

17.1 Overview

Key characteristics	
Country	Australia
Institution	Department of Industry, Science, Energy and Resources
Type of RD&I	Industrial research
Type of Intervention	Research and innovation grants
Evaluation challenges	Additionality; intangible benefits; lagged effects of research investments The key evaluation challenges are 1) Determining whether there is a causal relationship between programme funding and company turnover among BRII beneficiaries before the programme impacts have fully materialised. 2) Measuring the intangible benefits of the commercial solutions developed with programme funding (e.g. time savings made by product users).
Evaluation methodology and methods	• Cost-benefit analysis • Econometric modelling • Quasi-experimental methods (synthetic control method)

17.2 Introduction

This evaluation investigates the impact of the Business Research and Innovation Initiative (BRII), an Australian government challenge-based innovation programme. It focuses on quantifying the economic impacts of the programme.

The evaluation presents a cost-benefit analysis (CBA) comparing the total costs to deliver the first round of BRII against its economic benefits. This CBA combines quasi-experimental methods and econometric modelling to: 1) Quantify the direct benefit of the BRII on the business performance of first round SME beneficiaries; 2) Quantify the benefit to the Australian government and society from successful BRII government/SME partnerships. The quasi-experimental component of the CBA allowed the evaluators to address the challenge of determining programme additionality. The econometric modelling conducted for this evaluation employed revenue projections and scenario modelling. The former component of the econometric modelling allowed the evaluation to address the challenge of capturing the lagged effects of research investments, while the latter component provides a systematic approach to measuring the less tangible benefits of RD&I programmes.

17.3 Background description of RD&I intervention

Established in 2016, BRII is an Australian government challenge-based innovation programme. The programme was designed to achieve two goals: foster innovation among small and medium sized enterprises (SMEs) and help Australian government agencies find novel solutions to challenging public policy and service delivery problems.

The BRII has partnered with numerous government agencies on a range of difficult problems, attracting agencies with challenges across energy, human services, information management, agriculture and environment, data and digital, and tourism. Overall, fourteen government agencies have participated in the BRII, only one of which has been involved in multiple BRII rounds. The first three rounds of the programme’s operations were launched over the period 2016-2020 and granted 20.4 million AUD to 58 SMEs across fifteen challenges. This funding was distributed across a diverse group of SMEs, with very few repeat SMEs in the programme.

17.4 Challenges measuring impact

The evaluation addresses three key methodological challenges to determining the effect of BRII funding: additionality, measuring the intangible benefits of RD&I investments and capturing the lagged effects of research investments.

- **Additionality**: The evaluation must address the question of whether observed changes in company turnover among BRII beneficiaries are likely to have occurred in the absence of BRII funding. This is achieved by employing a quasi-experimental approach.
- **Measuring the intangible benefits of RD&I investments**: RD&I programmes often generate intangible benefits and assets that are difficult to measure or express in quantitative terms. In the case of this programme, commercialised BRII solutions brought a range of benefits to their users which are difficult to quantify, such as reduced time and labour required to conduct a certain task. The evaluation addressed this challenge through a scenario modelling exercise, which mapped the changes introduced by BRII solutions and estimated the costs/ savings associated with each of these changes.
- **Capturing the lagged effects of research investments**: Research and innovation support often takes place over several years, and the subsequent impact can take a decade or more to fully materialise. To address this challenge, the current evaluation employs an economic forecasting approach to capture estimates of anticipated economic impact before they have fully materialised.

17.5 Methodology and data sources

17.5.1 Methodology

This impact evaluation assessed the first round of the BRII, launched in 2016. It aimed to answer the following evaluation questions:

- What longer-term impacts has the programme achieved? To what extent has there been an increase in:
  - The commercialisation of new to market products/services among participating SMEs?
  - Collaboration among SMEs, agencies and industry partners?
  - Agencies' and SMEs' confidence working with each other?
  - Innovation and collaboration activities among participating SMEs and agencies?
- To what extent has the programme generated value to participating SMEs, challenge agencies, and the Australian government? Have the benefits outweighed the costs in light of expected outcomes and programme impacts?
- Has the program exhibited any spill overs or other unexpected consequences, positive or negative?
- How do the outcomes of the programme compare with similar programs elsewhere (such as SBRI in the UK), or with alternative programme designs for achieving the same objectives?

The evaluation did not provide comprehensive answers to all of these questions but rather focused on quantifying the economic impacts of the programme. The evaluation provided only tentative insights into the programme's effect on collaboration between participating SMEs and agencies, based on limited qualitative research findings and descriptive statistics. A comparative analysis was also conducted to benchmark the findings of this impact evaluation against those of impact evaluations conducted for similar programmes in the US and UK.

The economic impacts of the programme were quantified using a cost-benefit analysis (CBA) which compared the total costs to deliver the first round of BRII against its economic benefits. The CBA combined quasi-experimental methods and econometric modelling to:

- Quantify the benefit from the direct impact of the BRII on the business performance of first round SME beneficiaries, as measured by changes in company turnover.
- Quantify the benefit to the Australian government and society from successful government/SME partnerships, as measured by savings made as a result of commercialised BRII solutions.

To quantify the BRII's benefit on beneficiaries' business performance, the CBA employed the following methodology:

- Constructing a quasi-experimental control group to isolate the impact of BRII on beneficiary turnover. This involved use of a synthetic control method (SCM). SCM is a family of methods that can be used to construct a weighted combination of control companies to create "synthetic" controls that closely match the trends and characteristics of companies participating in the BRII during the pre-treatment period (i.e. prior to their joining the programme). The post-treatment outcomes of a BRII beneficiary can then be compared to the projected outcomes for this synthetic control. SCMs are particularly suitable for counterfactual analysis when the number of programme participants is small and a clear control group is not available.
- Forecasting the impact of the BRII for the period FY19/20 to FY24/25 by combining BRII beneficiaries' self-assessed revenue projections with econometric analysis. It should be noted here when the evaluation was conducted, insufficient time had elapsed since the launch of the first round to conduct a robust ex-post assessment.
- Using the above estimates of increased SME revenue to quantify "benefits to society" through gross value added (GVA) estimates. Increased company revenue attributed to the BRII for the period FY16/17 and FY24/25 was converted into GVA using a GVA-to-revenue ratio, based on a weighted sectoral average. Second order economic effects were then estimated by adjusting GVA for both displacement (negative second order economic effects) and spill overs (positive second order economic effects).

To quantify savings made by the Australian government and society as a result of commercialised BRII solutions, the evaluation conducted a scenario modelling exercise. Three commercialised BRII solutions were identified as having sufficient evidence of implementation to enable quantification. This exercise involved mapping the changes introduced by BRII solutions and estimating the costs/ savings associated with each of these changes. Costings were largely based on broad estimates provided by government and SME interviewees with demonstrated experience in implementing the BRII solutions in question. Demonstrating the benefit of the BRII in this way requires a set of key assumptions about how solutions have changed user outcomes because of the new technology or process innovation. Changes in inputs and outputs for solution users typically came in the form of costs avoided through reduced labour, time or other resources needed to achieve an outcome that is as good (or better) than could be achieved before the BRII solution was implemented. In some cases, unverified best estimates were used to round out a scenario model. While this scenario modelling exercise does not produce precise estimates, it does provide a systematic approach to measuring the less tangible benefits of the programme.

The CBA methodology described above makes several strong assumptions. Perhaps the most salient is the use of BRII beneficiaries' self-assessed revenue projections for forecasting the programme's economic impact over the period FY19/20 to FY 24/25. It is also notable that the study's estimates for displacement and spill over effects are based on strong, broad-brush assumptions. Thus, the report includes a sensitivity analysis which evaluates how variations in key assumptions would influence the overall CBA results. This allows for better understanding of the uncertainty around these results and the robustness of the findings.

17.5.2 Data sources

- **SCM analysis:** The SCM analysis combined modelled sectoral market growth data obtained from the Australian Bureau of Statistics (ABS) with a nearest neighbour matching method which used anonymised company financial records from the Department of Industry, Science, Energy and Resources' (DISER) data warehouse.
- **Economic forecasting**: BRII beneficiaries' self-assessed revenue projections were collected through an original survey of these companies fielded for the purposes of this evaluation. These projections were combined with data on BRII beneficiaries' annual turnover which were also collected through the survey, as well as the findings of the SCM analysis to produce forecasted revenue estimates.
- **GVA analysis**: The GVA analysis combined the economic forecasting results with sectoral GVA data from ABS Input-Output tables.
- **Scenario modelling**: The scenario modelling exercise was principally based on data collected from interviews with government and SME interviewees who had demonstrated experience in implementing the BRII solutions under study.

17.6 Main findings

This evaluation found that the first round of the BRII achieved an estimated net benefit of 10.4 million AUD and returned 1.64 AUD for every 1 AUD that the government invested in the program. This included total benefits of 26.8 million AUD, which were made up of 25.7 million AUD of benefits created through improved SME performance and 1.1 million AUD in benefits created through the implementation of three round one solutions. The total costs amounted to 16.4 million AUD, which were predominantly made up of grant funding costs alongside some administrative costs. These results reflect a benefit to cost ratio (BCR) of 1.6 to 1.

The sensitivity analysis estimates a lower bound net benefit of 3.8 million AUD and an upper bound net benefit of 17.3 million AUD for the programme.

17.7 Analysis of methodological suitability and effectiveness

The methodology employed in this evaluation has a number of strengths.

- The use of a SCM to conduct counterfactual analysis allows the research team to draw conclusions about the additionality of the BRII programme, specifically the extent to which participation in the programme increased turnover among beneficiary companies over and above what would have occurred in the absence of the programme.
- The incorporation of revenue forecasting techniques in the CBA allows the research team to address the challenge of measuring lagged effects during earlier rounds of evaluation, a common challenge when attempting to capture the emerging impacts of RD&I programmes.

- The use of scenario modelling makes it possible for the evaluation to provide credible estimates of less tangible impacts of the BRII programme and (relatedly) those with low observability. In a similar vein, the GVA analysis conducted for the CBA also serves to quantify indirect economic benefits of the BRII programme. However, these indirect economic benefits are “observable” in the sense that they are easily accounted for using input-output tables.

It is important to recognise, however, that this methodology also has important limitations.

- As discussed in section 17.4, the CBA methodology described above makes several strong assumptions. This is difficult to avoid given the complexity of the programme and the timeframe of the evaluation. However, the inclusion of a sensitivity analysis is a good practice which mitigates the issues of potential over/underestimation of programme impacts due to modelling assumptions.
- As also discussed in section 17.4, the scenario modelling exercise conducted for the CBA does not produce precise estimates. However, given that the programme under evaluation is relatively small in scale, it may be that a more precise approach was not possible due to a lack of available monitoring data and resources. For larger programmes, it may be feasible to implement a more rigorous and data-driven approach to scenario modelling for the purposes of quantifying the monetary value of governmental/ societal benefits accruing from innovations that are developed as a result of RD&I programmes.

17.8 Transferability and context

The methodology employed in this evaluation is transferable to the evaluation of a wide range of RD&I programmes in the UK context, subject to some conditions. It is particularly appropriate for early-stage evaluations. This methodology can be used where:

- The RD&I programme being evaluated is expected to result in market-ready innovations.
- Company-level data is available on the organisational characteristics and financial performance of programme beneficiaries and comparator companies. For beneficiary companies, this must include company-level forecasted turnover for upcoming years.

17.9 Conclusions

This evaluation provides a comprehensive CBA of the challenge based BRII programme, evaluating the programme’s emerging impacts in its early stages. It combines the econometric techniques of counterfactual analysis, forecasting, GVA analysis and scenario modelling to compare the total costs of delivering the first round of BRII against the programme’s direct and indirect economic benefits. The analysis suggests that the first round of BRII returned 1.64 AUD for every 1 AUD that the government invested in the program.

The evaluation methodology combines several techniques in order to overcome the significant challenges of evaluating the impact of small-scale RD&I programmes in their early stages. In doing so, however, it makes several strong assumptions. Thus, the results should be interpreted with caution, and with reference to the sensitivity analysis conducted by the research team. The methodology employed in this evaluation is transferable to the evaluation of a wide range of RD&I programmes in the UK context that are expected to result in market-ready innovations. It is particularly appropriate for early-stage evaluations.

18 Cooperative Research Centres Programme Impact Evaluation

18.1 Overview

Key characteristics	
Country	Australia
Institution	Department of Industry, Science, Energy and Resources
Type of RD&I	Collaborative R&D
Type of Intervention	Research and innovation grants
Evaluation challenges	Intangible benefits; skewedness of impacts The key evaluation challenges are 1) Capturing social and environmental impacts which could not be measured using secondary data or fully captured through beneficiary surveys. 2) Capturing the outsized impacts of a small number of very successful projects.
Evaluation methodology and methods	• Case studies • Quasi-experimental (computable general equilibrium modelling)

18.2 Introduction

This evaluation investigates the impact of the Australian Cooperative Research Centres (CRC) programme. It aims to assess the programme's overall impacts and value for money by combining econometric modelling with a case study approach. To quantify the impact of the programme on Australia's economy, the evaluation employs a dynamic, global computable general equilibrium (CGE) model. This approach allows the evaluation to address the challenge of capturing the skewed programme impacts. In addition, the evaluation documents the social and environmental impacts of the programme through a combination of survey responses and case studies. The case studies conducted allow evaluators to address the evaluation challenge of measuring intangible benefits.

18.3 Background description of RD&I intervention

The CRC programme was established in 1990 to promote collaboration between industry and the research sector in Australia. A flagship initiative of the Australian government, it is of high strategic importance to the government, involves significant funding and has a high public profile. In 2015, a second component of the programme named CRC Projects (CRC-P) was introduced to encourage greater SME involvement in collaborative research.

Over the period 1990-2020, the Australian government invested around 5.1 billion AUD in the programme to support 230 CRCs and 154 CRC-Ps, with the annual value of grants averaging around 150 million AUD. Over the life of the programme, CRC(-P) partners have contributed 3 billion AUD in cash and an estimated 12 billion AUD through in-kind contributions.

18.4 Challenges measuring outcomes and impact

The evaluation addresses two key challenges to determining the effect of CRC(-P) funding: measuring the intangible benefits of RD&I investments and capturing skewed impacts.

- **Measuring the intangible benefits of RD&I investments**: RD&I programmes often generate intangible benefits and assets that are difficult to measure or express in quantitative terms. In the case of this programme, the evaluation used a case study approach to documenting the social and environmental impacts of the programme which could not be measured using secondary data or fully captured through CRC(-P) beneficiary surveys.
- **Capturing skewed impacts**: The impacts of RD&I programme are often highly skewed towards a small number of very successful projects with a long tail of low or no-impact projects. This presents a significant challenge for evaluating public RD&I investments because traditional evaluation approaches focus on analysing average outcomes across all beneficiaries and are therefore ill-suited to capturing skewed impacts. This evaluation captures skewed impacts by collecting data on economic outcomes across the entire population of programme participants, with particular attention to documenting and verifying high-value impacts. These outcomes were then aggregated to arrive at a lower-bound estimate of the overall economic benefits of the programme.

18.5 Methodology and data sources

18.5.1 Methodology

This evaluation of the CRC(-P) programme aims to assess the programme’s overall impacts and value for money. It focuses on addressing two key methodological challenges related to assessing this impact: capturing skewed impacts and measuring intangible benefits.

The evaluation employs a dynamic, global computable general equilibrium (CGE) model to quantify the impact of the programme on Australia’s economy. A CGE model is an economic framework that simulates the interaction of various economic agents (such as households, firms and governments) in markets for goods, services, and factors of production to find a simultaneous equilibrium in all markets. A global CGE model extends this simulation across world regions, while the dynamic component simulates interactions across multiple time periods. The evaluation uses this modelling approach to compare the actual performance of the Australian economy (i.e. a scenario in which the CRC was implemented) to a counterfactual scenario in which CRC programme funding was allocated across other government expenditures, potentially having positive impacts elsewhere. Programme impacts were categorised under four tiers:

- Tier 1: CRC outputs, fully delivered by and attributable to CRCs.
- Tier 2: Collaborative outputs, partly attributable to CRCs and partly to other parties.
- Tier 3: Imminent outputs, expected to occur over the five years following the evaluation (2021-25).
- Tier 4: Preparedness outputs, which address potential risks.

Tier 1, 2 and 3 outputs were included in the CGE analysis while tier 4 was excluded. The tier 3 category serves to capture lagged effects of CRC investments which have yet to materialise, although these self-reported estimates are likely to be imprecise.

It is important to note the CRC-Ps were not included in the CGE analysis, ostensibly due to the lack of data on these projects. Instead, the evaluation conducted a cost-benefit analysis for this component of the programme.

To overcome the challenge of measuring intangible benefits, the evaluation documents selected social and environmental impacts of CRC and CRC-Ps through a combination of beneficiary survey responses and case studies. Importantly, the evaluation documents these impacts but does not attempt to quantify their value or demonstrate additionality. Beneficiary survey responses were used provide a high-level overview of CRCs/CRC-Ps self-reported social and environmental impacts. This was supplemented with 17 case studies that provide detailed documentation of selected social and environmental impacts. The evaluation report does not describe the rationale behind the selection of case studies.

18.5.2 Data sources

- **CGE modelling**: The data collection process for identifying economic impacts arising from CRC research and commercialisation activities was central to conducting the evaluation. Data was collected for CRCs in receipt of funding during the period 2012-2020, with 74 CRCs active in this period. These CRCs were surveyed to identify the economic impacts of their projects. This included information on cost savings, contract income, increased sales/ revenue, value of patents sold and value of spin-off companies.
- Only twenty of the 74 CRCs funded during the period 2012-2020 responded to the evaluation survey, although a number of others provided material, reports and other information. Where no survey response was provided, the evaluation team relied on

programme Exit Reports to identify impacts. To obtain updated information on impacts identified in Exit Reports, the evaluation team contacted senior managers and CRC partners as necessary. Particular attention was given to verifying claimed high-value impacts. As a result of these investigations, the evaluation team obtained information from 77 per cent of the CRCs active during 2012-2020 and catalogued 191 economic impacts. Where significant projected impacts could not be verified, they were excluded from the analysis. Impacts projected beyond 2024 were also excluded. Some estimated projected impacts were scaled back where the evaluation team deemed that they were unlikely to be fully realised.

- **Cost-benefit analysis:** The economic benefits of CRC-Ps were estimated from a variety of sources, including end of project reports, applications, beneficiary survey responses and discussions with stakeholders. The evaluation report notes that in principle, one method to evaluate the additionality of CRC-Ps would be to survey unsuccessful applicants to the programme and compare their outcomes to those of programme beneficiaries. However, the report also notes that unsuccessful CRC-P applicants were expected to be a difficult group to engage with, especially given the difficulty faced by the evaluation team in soliciting survey responses from successful applicants to the CRC-P programme.
- **Social and environmental impacts:** The evaluation documents selected social and environmental impacts of CRC and CRC-Ps through a combination of beneficiary survey responses and case studies. The case studies are based on document review and key informant interviews.

18.6 Main findings

The evaluation finds that CRCs active in the period 2012-2020 generated economic impacts exceeding 12.1 billion in 2021 AUDs compared to the counterfactual scenario in which CRC funding had instead been used for general government expenditure (which would have contributed to GDP, but at a lower level). This figure includes impacts anticipated to occur over the period 2021-2024, as well as some CRC funding prior to 2012 which contributed to outcomes during 2012-2020. Because the CRCs included in the analysis comprise around 77 per cent of programme participants over the period under study, this finding should be considered a conservative estimate of the CRC programme’s macroeconomic impact.

In addition to the positive impact on GDP, the evaluation finds that CRCs achieved a number of environmental and social impacts during this period. While the evaluation did not attempt quantify the value of these impacts, key examples include:

- Health: improvements in health and well-being from outcomes such as improved cancer therapeutics and asthma diagnostic products.
- Education and training: around 2,600 doctorate and masters’ degrees awards and research careers started in applied research.
- Social costs avoided: improved schooling in remote areas.

- International collaboration: CRCs report international collaboration including with the EU Framework Programme and NASA.
- Environmental: environmental impacts such as reduced greenhouse gas emissions, reduced water consumption and protection of endangered species.

With regard to CRC-Ps, the evaluation valued their economic benefits at 514 million AUD in net present value terms. This results in a benefit-cost ratio of 7.7. Considering all project costs gives a benefit-cost ratio of 2.5.

18.7 Analysis of methodological suitability and effectiveness

The methodology employed in this evaluation has a number of strengths.

- The CGE analysis allows the evaluation team to capture the skewed impacts of the CRC programme and aggregate these impacts to the level of the Australian economy.
- The use of survey data and case studies provides suggestive evidence regarding some of the less tangible outcomes of the programme.

This methodology also has important limitations.

- The CGE analysis has several limitations. First, CGE models are generally based on a number of assumptions regarding market behaviours, agent decision-making, and economic structures. This particular model is based on further assumptions, most notably the accuracy of projected programme impacts and the extent to which some CRC impacts can be partially attributed to the programme. Second, data for a significant number of CRC programmes is not available, while data for some others is incomplete. As a result, the findings of this analysis should be interpreted with some caution.
- The evaluation does not provide insights into the additionality of the CRC-P component of the programme or of the social and environmental impacts of the CRC.

18.8 Transferability and context

The methodology employed in this evaluation is transferable to the evaluation of some RD&I programmes in the UK context. The principal challenge to implementing this approach is the construction of a comprehensive dataset on project-level impacts. As discussed above, this data collection process is a labour-intensive endeavour that involves surveying programme beneficiaries, compiling data from programme reports and, in some cases, correspondence with beneficiaries to collect supplementary data. The feasibility of compiling such a dataset for a particular programme should be carefully considered before employing this evaluation approach. In particular, this approach is unlikely to be well suited for the evaluation of programmes with a large number of beneficiaries.

18.9 Conclusions

This evaluation provides a comprehensive analysis of the macroeconomic impact of the CRC programme on the Australian economy for CRC activities which took place over the period 2012-2020. It combines original data collected for the purposes of the evaluation with CGE modelling to provide evidence that the programme has made a substantial contribution to Australian GDP over and above what would have been achieved by general government expenditure. This was supplemented with qualitative data on the social and environmental impacts of the programme. Importantly, however, the methodology used in this evaluation cannot provide insights into the additionality of the CRC programme. This methodology is transferable to the evaluation of some RD&I programmes in the UK context, given that the required data collection process is feasible for the programme in question.

19 Evaluation of the Energy Entrepreneurs Fund EEF

19.1 Overview

Key characteristics	
Country	United Kingdom
Institution	Department of Energy, Security and Net Zero DESNZ
Type of RD&I	Applied research and innovation
Type of Intervention	Industrial research
Evaluation challenges	Lagged effects; Contribution The two key challenges in this evaluation are the lagged effects of the intervention and the early stage of most technologies and assessing the causality and contribution of the EEG given the duplicity of programmes providing support to clean technologies
Evaluation methodology and methods	• Theory-based (contribution analysis, Qualitative Comparative Analysis QCA) • Value for money evaluation (cost-effectiveness and cost-benefit analysis • Econometrics (Logic Regression Analysis)

19.2 Introduction

This case analyses the evaluation of the Energy Entrepreneurship Fund EEF over seven rounds of funding awarded since its establishment in 2012. The economic and environmental impacts of the programme pose significant challenges due to the lagged effects of the intervention and the early stage of most technologies. To address these challenges, the impact evaluation is done concerning the medium-term outcomes and impacts expected from the programme post-completion of projects.

A mixed method is adopted with Qualitative Comparative Analysis QCA of nine case studies, quasi-experimental methods and logistic regression analysis to establish causality and identify the mechanisms/pathways to achieving outcomes in the six main impact domains (regulation and availability of follow-on funding, R&D spending, technological development, commercial readiness, follow-on investment, commercialisation and adoption, economic benefits, environmental impacts and spillovers).

19.3 Background description of RD&I intervention

The Energy Entrepreneurship Fund EEF Programme provides funding and incubation support to SMEs to help them develop and commercialise low-carbon technologies, products and processes. Established in 2012, its prime objective is to support the development of start-ups and SMEs in the clean technology sectors by closing the gap with private equity funding, under the assumption that firms were holding back their development due to this. Through seven rounds, the EFF has funded 156 projects totalling £72M[43]. Firms are required to match DESNZ funding with at least 10% of the overall cost of the project. Awards fall into three main technological areas: buildings, energy networks and storage and clean power, all found at mid-way of the TRL journey.

The purpose of the EEF is to support firms to take their technologies to higher levels of the TRL and increase their commercial readiness through funding and incubation support. The programme objectives are:

- Produce disruptive/lower-cost technologies that improve energy security, lower carbon emissions, or improve energy efficiency.
- Support SMEs and early-stage innovators to develop innovative technologies and processes.
- Leverage private sector funding into pre-commercial technologies and processes.
- Produce technologies that are market-ready with businesses capable of achieving sales in the five years after the grant.
- Produce projects and technologies that are ready for a large-scale demonstration or pilot.
- Support SMEs with technologies or products that are suitable for follow on private investment.

The EEF's main expected outcomes are to generate increased R&D activity and technical development, improvements in commercial readiness, follow-on funding, and increased economic benefits (higher revenues, employment, turnover and Gross Value Added (GVA). In terms of environmental impacts, the programme is expected to lead to a reduction in both carbon emissions and reliance on fossil-fuel energy supply. Likewise, to lower the costs of low-carbon technology, increasing energy system flexibility and improving energy efficiency. Finally, some knowledge spillovers are expected to influence government regulation or other

43 The EEF is part of the Energy Innovation Programme (£505M) running from 2015 to 2021 and from round 8th (out of scope) has become part of the Net Zero Innovation Portfolio running from 2021 to 2025.

policies, via demonstration of technical and commercial feasibility or by revealing barriers to commercialisation of new technologies.

19.4 Challenges measuring impact

A first challenge faced by this evaluation is to estimate the environmental and economic benefits of projects funded by the programme as they will take time to materialise and require other changes in the energy system, such as new regulations, standards, and user and consumer practices to fully realise the benefits of some technologies. To overcome the lagged effects of these R&D activities, the evaluation adopts a forward-looking approach using the underlying value of participating firms as an indicator of the expected future profits of the companies.

Another challenge is to assess the causality and contribution of the EEG given the duplicity of programmes providing support to clean technologies. In 2012, an increase in public support for industrial innovation (e.g. Energy Catalyst, having similar objectives to the EFF, and other complementary ones, such as the Low Carbon vehicles and Low Impact Building Innovation Platforms. Which was overcome using qualitative and quantitative methods bridged by the QCA to provide a systematic approach to causality.

An additional challenge relates to the skewness of impact as the programme selected participants most likely to achieve positive outcomes. To address this, a case-by-case analysis provided a more accurate overview of the programme and its impact. Econometric methods were also used to counteract further potential effects of the selection, by interviewing and using data from declined applicants that scored highly in the technical assessment, who exhibit closer characteristics to the evaluation group.

19.5 Methodology and data sources

19.5.1 Methodology

The purpose of the evaluation is to assess the delivery of the EEF with three main objectives, identify the overall benefits and impacts of the scheme, and establish the achievement of the programme objectives. Alongside this, assess the cost-effectiveness of the programme and determine whether it has delivered value for money. Thus, the evaluation entails a process, impact and economic evaluation. As part of the process evaluation, identify the opportunities for adjustments to the EEF to optimise its impacts, benefits and efficiency. This case study focuses on the impact evaluation.

The impact evaluation adopts a mixed method approach, using econometric analysis, logistic regression analysis and Qualitative Comparative Analysis QCA, alongside in-depth case studies. Initially, the evaluators developed two sets of evaluation research questions, one investigates observed impacts by answering to five evaluation questions and the other around the contribution and additionality of the programme. These questions were framed as hypotheses to test throughout the evaluation in various impact domains: regulation and availability of follow-on funding, R&D spending, technological development, commercial readiness, follow-on investment, commercialisation and adoption, economic benefits, environmental impacts and spillovers.

Impact evaluation research questions

- Did EEF achieve its objectives? If so, by which mechanisms? If not, why not? (IE1)
- To what extent has the EEF programme achieved impacts in the following areas of priority: TRL progression, commercial readiness, external investment, follow-on funding?
- Do these impacts differ for different companies, stage of innovation, starting TRL, grant size, technologies and sectors? (IE3)
- What factors within the lifetime of the project influence impacts (e.g. collaborations, inputs from matched funding providers)? (IE4)
- What is the role of incubation support in achieving these outcomes? Do companies value incubation support? By which mechanisms have incubation support helped to achieve these outcomes? (IE5)
- Have there been any unintended consequences? (IE6)
- Contribution of EEF to the achievement of the observed impacts
- To what extent are the EEF's observed impacts additional to what would have happened otherwise? (IE7)
- How does the nature and level of additionality vary across different types of grant recipients, technologies, and sectors? (IE8)
- What explains any differences in the level of additionality observed for the different grant recipients, technologies, and sectors? (IE9)
- If the intended outcomes are not observed for certain recipients, why was this? (IE10)

The impact evaluation is done concerning the medium-term outcomes and impacts expected from the programme post-completion of projects. Post-completion information was collected via in-depth interviews and underpinned by secondary data and monitoring and management information.

Logistic regression analysis is used alongside the QCA method to strengthen its findings and establish which factors considered in the QCA have a statistically significant influence on the likelihood of an applicant achieving a given outcome. The regression model to estimate the effect of potential causal mechanisms and contexts on achieving stated EEF outcomes and impacts uses two main independent variables, the exposure to the EEF support and the level of satisfaction with the incubation support. The model controls for the starting TRL of the technology being developed, maturity and size of the business, technology area and level of funding secured by the beneficiary prior to the EEF grant.

Data completeness and the tautologous nature of some outcome variables, such as employment (whether the firm increased or not employment) are limitations faced by the method. A quasi-experimental method is used to overcome these methodological issues.

A **quasi-experimental analysis** is used to test the main hypotheses identified in the Theory of Change of the EEF to inform the economic evaluation of R&D activity and spending, leverage of follow-on funding, employment, turnover and GVA, valuations and the effectiveness of incubation services (comparing to grants awarded without it, e.g. Innovate UK).

The design of the counterfactual considers the potential selection bias of the programme, which could lead to overstating the effects of the intervention. For this, the counterfactual is designed from the declined applicants to ensure a credible comparator group. Further data adjustments and checks to avoid systematic differences between successful and declined applicants are undertaken (assessment scores – applicants scoring over 60 – and baseline characteristics – removing outliers e.g. companies with > 250 employees from secondary datasets). Additional refinements are done by matching EEF beneficiaries and decline applicants showing similar TRL levels at the point of application, technology area associated with the project, and scale of operations (employment and turnover).

Case studies are designed to enrich the findings – with EEF applicants' interviews – and provide evidence for the impact evaluation, with a focus on the environmental and commercial impacts of the intervention to explain how these have emerged. The sample was purposive, targeting those projects showcasing the highest commercial outcomes, distributed across the seven rounds and technological areas. Nine cases were fully developed from a shortlist of 15. Cases were developed around four main topics, project delivery, commercial outcomes, environmental outcomes and knowledge-sharing outcomes.

Qualitative Comparative Analysis is used to link qualitative and quantitative analysis with the sole purpose of establishing causality using case study data (qualitative) due to the high heterogeneity of the cases and the presence of impact attribution challenges (causation). The method is used to identify mechanisms enabling EEF applicants to achieve outcomes in the six main impact domains mentioned above.

Data from in-depth interviews is coded based on a coding frame developed from the topic guides. The analysis departs from a theory-based definition of variables (conditions) contributing to the emergence of outcomes. Context-Mechanism-Outcome CMO statements are elaborated to exemplify various pathways to achieving these outcomes. Each CMO presents the underlying conditions for project, business and wider contextual factors, and the main conditions under which an outcome occurs.

Economic analysis compares the costs of the EEF programme against the outcomes achieved to determine whether the intervention has generated value for money, through a Cost-Benefit analysis. The biggest challenge facing this analysis is that most associated environmental and economic benefits are yet to be realised and cannot be included.[44]

Costs of the programme are estimated using multiple sources, including DESNZ records on monitoring costs, incubation costs, and grants (assuming beneficiaries take it all by the end of the project).

Estimation of the environmental benefits is done using data collected during the case study interviews and the Management Information (CPRs). Estimates are defined within a low and high range[45] considering that they are self-reported by project teams and could lead to overstating the effect of the programme. Likewise, a social welfare analysis is done to determine the effects of the EEF on R&D investment to its economic benefits capitalised into the valuations of firms.

Some limitations of the economic evaluation include:

- The nature of the sample, which is not representative of the firms.
- Using company valuations as a proxy of well-functioning markets, when the logic of intervention assumes otherwise

The environmental analysis is made based on 11 projects that have trialled technologies in labs and have provided an ex-ante assessment of their CO_2 reductions. These estimates were elaborated by the project team (beneficiaries), which is the reason for selecting a high-low estimate value.

19.5.2 Data sources

Three main data sources are used in this evaluation:

- Monitoring and management information data from rounds 1-7 including application forms, interim reports, Commercial Progress Reports CPR from beneficiaries at the end of their projects, and incubation support activities[46]
- An extensive programme interview, consisting of 167 depth interviews split between successful applicants – 101 and 66 non-successful applicants, stakeholders: 19 interviewees including DESNZ officials, contractors – incubation support and monitoring services – and members of the commercial assessment panel. Additional 7 in-depth

44 Valuation data uses data from the PitchBook. A total of 44 – out of 133 – EEF beneficiaries had valuations in the database (Technical Annex p. 96)

45 The value of potential carbon savings was taken from the DESNZ traded value of carbon and the estimate uses the CO_2 savings derived from the EEF technologies and multiplies it by the commercial outcomes achieved to date by firms (low estimate) and the value using company sales projections over the next five years (high estimate).

46 The report states that the quality of the Management data was good, and coverage was high, particularly for rounds 5-7. Previous rounds didn't have CPR as mandatory, resulting in some data missing from rounds 1-4.

interviews for case study development were conducted with firms, their customers and partner organisations

- Third, secondary datasets, encompassing data from the Business Structure Database (BSD) and the Business Enterprise Research and Development (BERD) dataset from the Office for National Statistics (2010-2018), the Beauhurst database and PitchBook platforms, which provide data on capital investments, and finally some data obtained from Innovate UK Transparency Data with data on EEF being awarded funding from Innovate UK.

19.6 Main findings

The evaluation finds that EEF has achieved most of its shorter-term impacts regarding R&D spending, estimated around £328m to £580m by 2020 compared to grant spending of £67m. In terms of employment, it is found that between 140 and 320 R&D jobs are created by the firms benefited by EEF.

According to the evaluation, the EEF's additionality on R&D activity is high, as beneficiaries accelerated the development of their technologies (2.4 TRL levels against 1.5 for unsuccessful applicants). Moreover, declined applicants struggled to raise alternative funding for their projects, slowing their development.

EEF beneficiaries showed rapid advance on their business models and commercial progress, as a results of the incubation support.

In terms of the longer-term, the programme is found to have a positive impact on follow-on funding enabling participants to raise around £4.3m in equity funding by 2020 compared to £0.6m by declined participants in the same time frame. Considering a £462K average grant value, the EEF funding enables an increase in equity investment raised by £0.7m to £2.1m on average, by 2020.

The evaluation found that firms commercially advanced tend to benefit more of the programme compared to those at earlier stages of their market validation. From this, a recommendation is to award funding according to the validation of the business models and the consideration of a precursor programme to help build readiness.

Effects on commercialisation and adoption of technologies were difficult to identify due to time lags. Isolated policy spillovers were identified during the evaluation. To benefit of learning and capturing barriers to commercialisation originating in policy, the evaluation suggests a knowledge management function to reveal and trace this.

Regarding the value for money evaluation, the EEF is found a cost-effective instrument for leveraging private R&D investment for highly innovative clean technologies and more moderately in helping leverage follow-on investment and increasing the underlying value of participating firms.

The environmental impact of the projects – based on ex-ante modelling of the environmental benefits made by beneficiaries – showed that about one-third of participating firms would not be expected to deliver emissions reductions exceeding the value of grants awarded over a 10-year period. For this, the evaluation recommends incorporating a further verification of the potential environmental benefits of the programme.

19.7 Analysis of methodological suitability and effectiveness

The evaluation answers the impact and contribution evaluation questions and estimates the economic value of the EEF via a cost-benefit analysis. Priority impact areas of the EEF, such as TRL progression, commercial readiness, follow-on investment and follow-on R&D funding are addressed, quantifying the effects and describing the mechanisms enabling short-term and longer-term outcomes.

Challenges in assessing impact are approached from different angles to strengthen the robustness of findings, for example, using QCA methods to reveal pathways to impact and complementing this with a quasi-experimental approach to investigate probabilities of causal relationships.

Amongst the main challenges faced to measure impact are the lagged effect of R&D investments, given the early stage of certain technologies. To deal with this, the evaluation differentiates between short-term and longer-term outcomes to offer an immediate economic effect analysing the most advanced projects and investigating in detail the factors contributing to their successful, to finally quantify the effect generated by the EEG funding on the value of those companies.

Although the above provides an estimate of impact, the most salient methodological limitation remain as to the appropriateness of using firm valuation as an indication of economic impact. This assumes well-functioning financial markets and considers that in absence of revenue and commercial readiness, valuations given by equity investments provide a good indicator. This is problematic for two reasons clearly highlighted in the evaluation. First, firms who do not receive follow-on investment are unobserved, and may result in underestimation of the EEF's effect. Moreover, a core assumption for the creation of the EEF programme is that financial markets do not value effectively clean technologies, questioning the reliability of estimated economic benefits.

Opportunities for improvement are outlined in the report in different areas. First, the programme can benefit from a closer and more detailed assessment of the applicants' technological progress and the maturity of their projects, offering support for early market validation and providing appropriate levels of funding. This is suggested in the form of stage-gate processes for beneficiaries who can get access to additional funding subject to commercial and technological progress. According to the evaluation, this could increase the economic benefit of the programme. In a similar vein, this could also increase the environmental impact of the programme, ensuring all projects target environmental issues for the design of their projects.

Finally, to capture and capitalise the benefits of individual learnings, the evaluation recommends adopting a knowledge management mechanism to facilitate this and enhance knowledge sharing.

19.8 Transferability and context

This evaluation is done in the UK context, using monitoring and programme management data, alongside primary sources of information, such as in-depth interviews and surveys and secondary datasets for venture capital investment in the UK. Monitoring and management information are instrumental for the evaluation, particularly the CPR provided by grant holders by the end of their project.

In this sense, this methodology can be replicated for evaluating similar programmes supporting experimental development and industrial research if monitoring data as well as CPRs are available from the start of programme roll-out. Effective measurement of economic impact will also be dependent on the quality of initial technological assessments of projects and detailed analysis of their maturity, which was limited to some extent in this evaluation.

Causality and contribution analysis explored in case studies and through a qualitative comparative case study analysis can be replicable in any context, as they mainly rely on primary data from in-depth interviews and complemented with projects reports. Economic and environmental impact are assessed using secondary data, and management data, making it feasible to replicate the approach in the national context.

19.9 Conclusions

This evaluation offers a comprehensive methodology to assess the impact of the Energy Entrepreneurship Fund in supporting the development of clean technologies at early stage and the leverage of equity funding by SMEs and start-ups in the sector. The methodology provides robust answers to the evaluation questions through the triangulation of evidence to examine causality of impacts and offers a clear view of where the value of the programme is and the opportunities for improving its effectiveness. With this, the methodology used can be implemented in other sectors and similar types of programmes providing support to experimental and industrial development.

20 Evaluation of the Enterprise Ireland Research, Development and Innovation Programme

20.1 Overview

Key characteristics	
Country	Ireland
Institution	Department of Business, Enterprise & Innovation (DBEI)
Type of RD&I	Applied research and innovation
Type of Intervention	Research and innovation grants
Evaluation challenges	Additionality; attribution The key evaluation challenges are 1) Determining the causal relationship between Enterprise Ireland RD&I Programme funding and firm-level R&D performance. 2) Attributing observed changes in companies' R&D performance to the programme, independent of confounding factors and external influences in Ireland's innovation system.
Evaluation methodology and methods	• Generic methods (interviews, survey) • Quasi-experimental (matching)

20.2 Introduction

This evaluation investigates the impact of the Enterprise Ireland (EI) RD&I Programme, with a focus on the programme's impact on firm-level R&D performance indicators for programme beneficiaries including R&D expenditure, R&D intensity and number of R&D employees.

To address the evaluation challenges of additionality and attribution, this study employs a quasi-experimental research design, alongside a survey and interviews with programme beneficiaries. Due to data limitations and the complex nature of the programme, econometric analyses were insufficient to address all of the evaluation themes. Thus, the survey and interviews served to fill data gaps and provide suggestive evidence on the causal mechanisms underlying the quantitative findings. This allowed the evaluation to provide methodologically rigorous answers to the questions of additionality and attribution where data was available, in addition to more suggestive answers where there were data limitations.

20.3 Background description of RD&I intervention

The EI RD&I Programme aims to support Irish companies of all sizes to engage in innovation and research activities. It is a comprehensive programme offering a range of direct financial supports as well as various indirect supports.

During the period covered by this evaluation, the programme offered the following direct financial support schemes: the Exploring Innovation Grant, the RD&I Fund, the Intellectual Property Strategy Offer, the Agile Innovation Fund, and the Business Innovation Offer. These may be delivered as standalone grants or as part of a tailored package of supports. Over the period 2007-2018, the programme awarded 2,005 grants with a total value of 464.7 million Euros to 1,562 individual firms.

Sitting around these direct financial supports are a range of indirect supports. These include two procurement-based schemes administered by EI: European Space Agency funding and Small Business Innovation Research funding. They also include two tax-based R&D supports, in the form of the R&D Tax Credit and the Knowledge Development Box.

The evaluation focused on the suite of direct financial supports to firms, with the exception of those provided to firms classified as high potential start-ups, which were outside the scope of the evaluation.

20.4 Challenges measuring outcomes and impact

This methodology addresses the evaluation challenges of additionality and attribution.

- **Additionality**: The impact evaluation must address the question of whether observed changes in R&D performance by EI R&D Programme beneficiaries is likely to have occurred in the absence of the programme. This poses a challenge because programme beneficiaries are systematically different to Irish companies that do not participate in the programme. Compared to non-beneficiaries, programme beneficiaries have a larger number of employees, higher levels of labour productivity and higher wages. They are more likely to be exporters and patent holders and more likely to have a link to a university. This evaluation overcomes the challenge of determining additionality by employing a quasi-experimental research design that enables the construction of a credible counterfactual (control group) to compare programme beneficiaries against.
- Importantly, however, rather than relying exclusively on econometric analysis, the evaluation complements this analysis with survey and qualitative methods in order to fill data gaps, as well as to provide suggestive evidence on the causal mechanisms underlying the quantitative findings. This allowed the evaluation to provide

methodologically rigorous answers to the question of additionality where data was available, in addition to providing more suggestive answers where data limitations precluded counterfactual analysis.

- **Attribution**: The online survey of EI RD&I Programme beneficiaries included questions to measure firms' self-reported attribution of improved R&D performance to the programme's support. This approach provides suggestive insights into programme attribution but is prone to reporting bias and misreporting.

20.5 Methodology and data sources

20.5.1 Methodology

This evaluation had four objectives:

- Determining the effectiveness of the programme in achieving the desired or any impact.
- Determining the effectiveness of the programme at increasing the number of companies investing in RD&I.
- Comparing the role and performance of the direct-to-firm financial supports of the EI RD&I Programme in terms of achieving increased business expenditure on R&D (BERD) to that of indirect supports.
- Exploring the synergies between direct and indirect supports.

To this end, the evaluation employed a three-pronged methodological approach:

- **Econometric analysis**: Propensity score matching was used to compare 654 beneficiaries of EI RD&I Programme funding to a comparator group of firms with similar characteristics that had not accessed this funding. The propensity score was designed to match beneficiaries of the EI RD&I Programme with firms that would have the same probability of being awarded RD&I direct financial supports. The comparator firms thus acted as a control group, allowing the evaluation team to assess the programme's impact on firm-level R&D performance indicators for a period of five years after the first approved award. These indicators include R&D expenditure, R&D intensity and number of R&D employees. Importantly, only the R&D Fund was included in the econometric analysis because the newer components of the programme (Agile Innovation Fund and Business Innovation Offer) had been launched too recently for their effects to have materialised.
- **Survey data**: The online survey of EI RD&I Programme beneficiaries included questions to measure the self-reported impacts and benefits of programme funding, as well as firms' attribution of improved R&D performance to the programme's support. Triangulating the results of the econometric analyses with this survey data serves as a validity check for the econometric results. Moreover, the survey data provides information on firms for which econometric analysis was not possible due to data limitations. This includes beneficiaries of newer components of the programme (as discussed above), as well as micro-sized enterprises which are not included in the DBEI

Annual Business Survey of Economic Impact (ABSEI) used in for the econometric analysis.

- **Interview data**: Interviews with EI RD&I Programme beneficiaries and programme staff were used to: 1) Provide insights into the causal mechanisms linking EI RD&I Programme funding to the various impacts observed; 2) Provide insights into the nuances of the high-level quantitative findings; 3) Explore impacts which could not be studied quantitatively due to data limitations. In particular, the impact of the R&D Tax Credit could not be studied quantitively due to restrictions on access to tax credit data.

20.5.2 Data sources

- The evaluation team constructed a dataset of Irish enterprise development agency client firms containing information on their characteristics, economic performance, and receipt of EI RD&I Programme funding (or lack thereof). The dataset was constructed by linking data from the ABSEI with EI administrative data.
- The online survey of EI RD&I Programme beneficiaries was an original survey conducted for the purposes of this evaluation. The survey of awardees was launched to 1,058 firms and remained open for three weeks. The survey received 220 responses, representing a final response rate of 21%.
- The study team conducted 34 interviews with EI RD&I Programme beneficiaries, identified via the online surveys, and eight interviews with EI Development Advisors, who were identified and recruited with the support of EI.

20.6 Main findings

The econometric analyses found that the programme had a positive impact on a number of innovation and economic performance indicators among beneficiary firms during the five years after their first approved award. This includes increased R&D expenditure, R&D employment, turnover, total value added, employment, export sales and export intensity. However, the econometric analyses found no impact of the programme on productivity among beneficiary firms.

The online survey and interview findings validate these results and provide suggestive evidence that the lack of productivity impact may be related to a long time lag between firms' implementation of RD&I initiatives and their release of new products and/or entry into new markets. In addition, the online survey and interview findings provided evidence of synergy between the programme's direct financial supports and its indirect supports.

20.7 Analysis of methodological suitability and effectiveness

The methodology employed in this evaluation has a number of strengths.

- The mixed-methods approach was carefully designed to ensure that the evaluation could answer all four evaluation themes in an effective manner. No single research methodology would have been sufficient for this purpose due to the complex nature of the programme and the limitations of the data available on participant firms and their comparators.
- The econometric analysis benefits from the availability of rich, longitudinal firm-level data on RD&I outcomes and economic performance captured through the ABSEI. This allows the evaluation to adopt a five-year follow-up period for investigating the impact of direct financial supports, thus capturing programme effects over a considerable time period.

The methodology also has several limitations.

- As is typical in RD&I programme evaluations, the original online surveys of programme beneficiaries and (especially) comparator firms have low response rates. This may have implications for the representativeness of these surveys. Moreover, it precludes analysis of the survey data by firm type (for example, findings disaggregated by enterprise size or sector).
- The lack of quantitative data available on indirect programme supports means that findings related to their impact can only be considered suggestive. These findings cannot be considered to provide robust evidence on the question of additionality.

20.8 Transferability and context

The methodology employed in this evaluation is transferable to the evaluation of some RD&I programmes in the UK context, given the following conditions are met:

- There must be rich, longitudinal firm-level data available on the RD&I outcomes and economic performance of both programme beneficiaries and plausible comparator firms (the latter could include unsuccessful applicants to the programme, clients of the enterprise development agency that did not access the programme or companies with similar characteristics more generally).
- When adopting a propensity score matching methodology, it is strongly advisable that there is a large number (sample size) of programme beneficiaries and comparators. This helps in ensuring that it is possible to match programme beneficiaries with suitable comparators and in ensuring that the statistical analysis yields reliable estimates.
- The evaluation must have sufficient resources to enable the collection of original data. Both original surveys and interviews are costly data collection methods.

20.9 Conclusions

This evaluation provides a comprehensive analysis of the EI RD&I Programme’s impact. Data limitations and the complexity of the programme’s design posed a number of evaluation challenges; however, these challenges were effectively mitigated by employing a mixed-methods approach drawing on a diverse range of data sources. This approach is transferable to the evaluation of some RD&I programmes in the UK context, given that certain conditions related to data availability, the number of programme beneficiaries and the evaluation budget are met.

21 Value of Engineering and Physical Sciences Research Council (EPSRC) Fellowships

21.1 Overview

Key characteristics	
Country	United Kingdom
Institution	UK Research and Innovation, EPSRC
Type of RD&I	Basic research
Type of Intervention	Fellowships
Evaluation challenges	Additionality; displacement The key evaluation challenges are 1) Determining the causal relationship between participation in ESPRC fellowships and academic performance/impact. 2) Accounting for the unintended negative externalities of the programme (i.e. opportunity costs).
Evaluation methodology and methods	• Theory-based (survey; interviews) • Bibliometric analysis • Patent analysis • Value for money

21.2 Introduction

This evaluation presents a comprehensive overview of EPSRC fellowship impacts, focusing on the programme's impact on knowledge production, fellows' career advancement and economic impact. It utilises a combination of theory-based evaluation principles and economic modelling.

The approach to economic modelling employed here allows evaluators to address the challenges of determining programme additionality and displacement effects within the framework of a value for money analysis. This is achieved by quantifying and monetising estimates of displacement, substitution and deadweight effects and incorporating these estimates into a value for money analysis.

21.3 Background description of RD&I intervention

EPSRC fellowships are aimed at nurturing aspiring and established research leaders by providing them the support, flexibility and freedom to develop their research ideas. Grants are awarded to individuals with the greatest potential at early and established career stages.

In addition, targeted fellowship schemes are launched periodically to build critical mass and/or capabilities in areas of strategic priority. The expectation is that fellows will deliver the highest quality research, make significant contributions to their field, improve their visibility (in their field and internationally) and progress in their career.

Over 2006-2020, the period under study in this evaluation, 603 EPSRC fellowships were funded. These fall into three groups:

- Traditional early career fellowships, consisting of the Postdoctoral, Career Acceleration and Early Career Fellowship schemes.
- Traditional fellowships for established researchers, consisting of the Senior, Leadership and Established Career Fellowship schemes.
- Targeted fellowship programmes, including Challenging Engineering, Manufacturing and Engineering for Growth Fellowship schemes.

The various fellowships are structured differently, but broadly speaking they offer fellows a bursary to conduct their research, set up a research group and/or visit potential international partners, for a period of 3-5 years. EPSRC Fellowships also facilitate access to training, mentoring, and research facilities.

During the period under study in this evaluation, the University of Oxford, Imperial College London and University of Cambridge accounted jointly for over 37 percent of all awarded fellowships, while the remaining fellowships are distributed among 44 different organisations.

21.4 Challenges measuring outcomes and impact

This methodology addresses the following evaluation challenges:

- **Additionality.** The impact evaluation must address the question of whether observed changes in fellows' knowledge production, career advancement and economic impact are likely to have occurred in the absence of the fellowship programme. This poses a challenge because fellows are selected on the basis of exceptional past performance and it is plausible that even in the absence of the fellowship programme, their academic performance and impact would improve as their career progresses. To address this challenge, the evaluation collected self-reported data from an original EPSRC fellowship alumni survey about the extent to which the programme provided scholars with time for research that they would not have otherwise had.
- **Displacement**. Displacement is a potential unintended effect where the positive outcomes of the programme being evaluated come at the cost of negative outcomes

elsewhere. In the context of this programme, the two key displacement effects are: 1) The opportunities lost by EPSRC fellowship applicants rejected after the interview round (i.e. highly competitive candidates who did not receive the funding); 2) The time and effort of unsuccessful applicants in applying for the EPSRC fellowship programme. To account for the impact of these unintended effects, the evaluation calculated a quantitative estimate of their size and adjusted the value for money analysis accordingly.

21.5 Methodology and data sources

21.5.1 Methodology

This evaluation aimed to determine the impact of EPSRC fellowships funded over the period 2006-2020, focusing on the programme's impact on knowledge production, fellows' career advancement and economic impact.

The evaluation adopts a mixed methods approach, combing theory-based evaluation principles with economic modelling. The table below presents how various quantitative and qualitative approaches were employed to determine various types of impacts.

- **Theory-based evaluation principles** were used to inform a qualitative analysis. This analysis drew on several data sources about the project: Researchfish programme monitoring data, a survey of EPSRC fellowship alumni, interviews with EPSRC fellowship alumni and bibliometric / citation analysis. These data sources were used to trace the outcomes anticipated by the programme.
- **Economic modelling** was used to determine the economic impact of the project. This approach combined patent analysis and value for money analysis, following the methodological guidelines of Sartori et al. (2014)[47] in respect to valuing investments in research, development and innovation. Four channels of economic impact were measured:
  - **Career progression: Wage premium**. The estimated value of the wage premium earned by EPSRC fellowship alumni, over their career, due to their participation in the programme. This is compared to a scenario where the researchers/fellows would not have been granted the EPSRC fellowship. The estimated wage premium was based on self-reported values from the alumni survey conducted in the context of this study.
  - **Knowledge and innovation (direct effects): Additional value of spinouts**. This is the estimated additional value of spinouts (turnover) created by EPSRC fellows/alumni. The value of EPSRC fellows/alumni spinouts is estimated using Researchfish, ONS and Eurostat data.

47 Sartori, D., Catalano, G., Genco, M., Pancotti, C., Sirtori, E., Vignetti, S., & Del Bo, C. (2014). Guide to cost-benefit analysis of investment projects. Economic appraisal tool for Cohesion Policy, 2020.

- **Knowledge and innovation (direct effects): Additional value of granted patents.** This is the estimated additional value of granted patents linked to EPSRC fellows/alumni. Researchfish data was used to identify these patents.
- **Knowledge and innovation (spillover effects): Additional value of granted patents to other stakeholders.** This is the estimated additional value of granted patents to other stakeholders (beyond EPSRC fellows/ alumni) that draw on knowledge produced by EPSRC fellows/alumni. Patent citation data was used to identify patents that reference research publications produced by EPSRC fellows/alumni.

The economic modelling also incorporates estimates for displacement, substitution and deadweight effects. In terms of displacement, two effects were considered.

- EPSRC fellows are given an opportunity to boost their careers. This makes them more competitive candidates for other/future grants, enabling them to source additional funding and possibly displacing other researchers competing for the same opportunities. To reflect this, the value for money analysis was adjusted to account for opportunities lost by EPSRC fellowship applicants that were rejected after the interview round (i.e. highly competitive candidates who did not receive the funding). These applicants comprised 49 percent of all interviewed candidates and were assumed to be displaced.
- The time and effort of unsuccessful applicants in applying for the EPSRC fellowship programme is also a displacement effect. This was estimated at five working days and monetised based on figures for average UK R&D salaries by age bracket.

Substitution and deadweight were estimated based on data collected from the alumni survey. Data on substitution was based on self-reported values regarding the number of teaching hours transferred from EPSRC fellows to other researcher/s for the duration of their fellowship. Data on deadweight was obtained from self-reported values regarding the extent to which the programme provided fellows with time for research that they would not have otherwise had.

The total assumed benefit of the programme is compared with the cost of investment in a return-on-investment analysis.

Finally, sensitivity analyses are conducted to estimate the impact of the key modelling assumptions on the total estimated economic benefits of the programme. The report presents a sensitivity analysis for each of these key assumptions. According to an interview with the economist that led the quantitative component of the evaluation, this suite of sensitivity analyses was valuable in providing the research funder with a concrete understanding of the modelling assumptions and aided the funder in interpreting the evaluation results.[48]

48 Interview with Diogo Machado, 10/01/2024.

21.5.2 Data sources

- **Researchfish data.** Researchfish programme monitoring data on outcomes such as research publications, patents, fellows' career trajectory and awards were analysed by fellowship type to assess and compare the strength of the different types of EPSRC fellowships. This data was also used in the bibliometric analysis and economic modelling (see below).
- **Bibiliometric analysis**. Researchfish was used to obtain data on research publications attributable to ESPRC fellowships and EPSRC fellows. 529 out of 603 EPSRC fellows were matched to Scopus to conduct the bibliometric analysis. The production of bibliometric indicators is limited to articles, reviews and conference papers (and some book chapters). Data for an additional 18 fellows was identified by tracing funding acknowledgements. In total, data for 546 of 603 (90. percent) fellows was identified. In total, 11,449 papers were matched to Scopus and an additional 910 papers are identified in Scopus funding acknowledgements, yielding a total of 12,359 papers included in the bibliometric analysis.
- **EPSRC fellow alumni survey**. The survey was used to collect data on the impact of the EPSRC fellowship on alumni research outcomes, the use of these outcomes by others, whether there had been any unexpected research outcomes and any wider impacts. In total 223 responses were collected.
- **EPSRC fellow alumni interviews**. A semi-structured interview programme with 23 alumni was conducted to explore personal career impact, including leadership capability, research outputs, economic impact, wider impact and unexpected impacts. The sample of interviewees was drawn from data provided by EPSRC and informed by the responses to the survey. The interview programme was specifically designed to include participants from all the programmes, across the disciplinary themes and covering a range of characteristics, as appropriate (e.g. gender, career stage, time since fellowship, etc.).
- **Economic modelling**. The economic modelling employs Researchfish programme monitoring data on spinouts, intellectual property protection (including patents granted) and EPSRC fellowship publications referenced by patents. In addition, it employs data from the EPSRC fellow alumni survey on wage increases obtained immediately following the fellowship.

21.6 Main findings

Knowledge production. The evaluation found that the EPSRC fellowship programme supported highly cited, high impact research. In particular, EPSRC fellowships have led to the publication of 11,775 research papers in total. The Average Relative Citation score of papers published by EPSRC fellows was found to be 2.06, which means that these are cited two times more frequently compared to the world level (i.e. 1.0). This is in line with similar impact measures for the EPSRC as a whole (e.g. 2.06 in Field-Weighted Citation Impact). However, 25 percent of the top 10 percent Highly Cited Publications, in their respective fields, come from EPSRC fellows.

Career progression. The evaluation findings also suggest that EPSRC fellowships have a positive impact on career development. 93 percent of EPSRC fellowship alumni surveyed agreed that the fellowship had made a significant difference to their career path. 86 percent agreed that it had impacted the level of seniority they had reached, with a similar proportion agreeing they had experienced faster career progression than they would have done without a fellowship.

Economic impact. The evaluation finds that the EPSRC Fellowships contributed to a range of positive economic impacts, including:

- 50 EPSRC fellowships from 2006-2016 contributed to the creation/development of spinouts, which have also received further contribution from other EPSRC grants.
- 106 patent applications are linked to the EPSRC Fellowships, with 35 patents granted at the time the evaluation was conducted.
- 414 publications associated to EPSRC fellows are referenced in 1,012 patents filed by others.

The economic assessment of monetisable benefits emerging from EPSRC Fellowships reveal a net benefit of £615.2 million under conservative assumptions after adjusting for displacement, substitutions and deadweight. Consequently, the estimated return on investment (ROI) of the EPSRC Fellowship programme is 1.39. For every £1 invested in the programme, there is an additional benefit of £0.39. This is most likely an underestimation of the impact of the fellowships as it only captures the impact that is monetised through the four channels of impact described in the methodology section above.

Wider impact. The evaluation report documents a number of case studies illustrating how EPSRC fellowship research has made significant contributions towards health, social and environmental impacts.

21.7 Analysis of methodological suitability and effectiveness

The methodology employed in this evaluation has a number of strengths.

- It draws on self-reported measures of programme impact from an original beneficiary survey to address the question of additionality. This presents a cost-effective alternative to addressing the issue of additionality by comparing beneficiaries with a control group of similar researchers. According to an interview with the economist that led the quantitative component of the evaluation, this approach was adopted because of budget constraints that did not allow for the implementation of a quasi-experimental approach to measuring impact.[49]

- It addresses the evaluation challenge of capturing displacement effects by mapping these effects and incorporating them into the value for money analysis.
- It utilises a combination of qualitative and quantitative data sources to assess the programme's impact. This triangulation between different data types is an effective approach to validating evaluation findings.

It is also important to recognise that this methodology has several limitations:

- The use of self-reported measures of programme impact is an efficient method to address the question of programme additionality, however, it may be less reliable and/ or accurate than the use of objective measures.
- The evaluation documents examples of how EPSRC fellowship research has made significant contributions towards health, social and environmental impacts. However, it does not quantify the aggregate value of these impacts or incorporate them into the economic modelling component of the evaluation.

21.8 Transferability and context

The methodology employed in this evaluation is particularly appropriate for evaluating UK RD&I programmes in which:

- Determining value for money is a primary objective of the evaluation.
- Data or resource constraints prevent the evaluators from measuring impact by comparing beneficiaries with a control group of similar individuals or organisations.
- The key outcomes of the programme include research outputs, patentable innovations and/or economic outcomes.

To implement this methodology, it is necessary for evaluators to have sufficient access to programme beneficiaries in order to conduct a comprehensive beneficiary survey. The evaluation also relied heavily on the availability of rich monitoring data for the EPSRC fellowship programme through Researchfish. Although it may be possible to implement this methodology in the absence of such monitoring data, this is likely to involve significant data collection efforts.

21.9 Conclusions

This evaluation presents a comprehensive overview of EPSRC Fellowship impacts, including a thorough analysis of the programme’s economic impact. It combines theory-based evaluation

49 Interview with Diogo Machado, 10/01/2024.

principles and economic modelling, utilising a combination of qualitative and quantitative data sources.

The evaluation finds that the EPSRC Fellowship programme supported highly cited, high impact research and had a positive impact on fellows' career development. The evaluation also finds that the EPSRC Fellowships contributed to a range of positive economic impacts, including the development of patentable innovations, the creation/development of spinouts and a wage premium for EPSRC Fellows. The estimated return on investment (ROI) of the EPSRC Fellowship programme is estimated at 1.39.

The methodology employed in this evaluation is particularly appropriate for evaluating UK RD&I programmes in which determining value for money is a primary objective of the evaluation and in which there are constraints preventing evaluators from measuring impact by comparing beneficiaries with a control group of similar individuals or organisations.

22 Returns on Research Funded under the NIHR Health Technology Assessment

22.1 Overview

Key characteristics	
Country	United Kingdom
Institution	Department of Health
Type of RD&I	Applied research and innovation
Type of Intervention	Research and innovation grants
Evaluation challenges	Skewness of impacts The key evaluation challenge is identifying the small number of highly successful projects in the programme's research portfolio and evaluating the economic and policy impacts of these projects.
Evaluation methodology and methods	• Case studies • Value for money analysis

22.2 Introduction

This study presents an analysis of the potential economic returns on research funded under the National Institute for Health and Care Research's (NIHR) Health Technology Assessment (HTA) Programme, as well as providing insights into wider impacts of the programme on policy and practice in the UK's National Health Service (NHS).

The study's approach to value for money analysis and case study methods allows evaluators to address the challenge of determining impact in instances where programme impacts are expected to be highly skewed toward a small number of very successful projects. This is achieved by a purposive sampling strategy that identifies the most high-impact programme outputs and restricts analysis to these cases.

22.3 Background description of RD&I intervention

The NIHR's HTA Programme was established in 1993 and is the largest research programme dedicated to the NHS. It funds research on the clinical effectiveness, cost-effectiveness and broader impact of healthcare treatments. The purpose of the programme is to ensure that high quality research evidence is made available to policymakers, practitioners and patients in a timely and efficient manner. Funded research is particularly intended to inform the key decision-making bodies within the UK health system, especially the National Institute for Health and Clinical Excellence (NICE). The work of the HTA Programme covers both primary research and evidence synthesis, with HTA studies typically conducted to provide clarity in situations where there is a balance of evidence. The research can be either commissioned or researcher led.

22.4 Challenges measuring outcomes and impact

This methodology addresses the challenge of evaluating RD&I programmes that have **impacts which are highly skewed** towards a small number of very successful projects. In this case, the NIHR's HTA Programme funds a substantial number of studies, however, only a small subset of these are expected to result in a substantial economic return to the NHS and the UK population for several reasons. First, only HTA research projects which demonstrate that a new intervention offers advantages over the existing standard of care will result in the implementation of new practices in the NHS. Second, not all new interventions found effective by HTA research will be adopted by the NHS. Third, not all new interventions introduced by the NHS as a result of HTA research will necessarily be high impact.

To address the challenge of evaluating a programme whose impacts are expected to be highly skewed toward a small number of very successful projects, this evaluation employs a purposive sampling approach designed specifically to identify high-impact projects and conduct an in-depth analysis of them.

22.5 Methodology and data sources

22.5.1 Methodology

This study on the HTA Programme's impact aimed to determine its potential economic benefits relative to its cost. More specifically, it aims to estimate the benefits that would have accrued to the NHS and the UK population if the findings of HTA studies had been implemented. It focuses on benefits that could be delivered by HTA research projects which demonstrate that a new intervention offers advantages over the existing standard of care. The extent to which the findings of the HTA studies were actually adopted is not within the scope of this study.

The study involves two components: an economic analysis and a case study approach. The economic analysis quantifies and monetises the benefits of a sample of potentially high impact HTA studies, comparing these benefits to the cost of the entire HTA Programme. The case study component aims to document HTA-funded research impacts which cannot captured by the economic analysis, highlighting the extent to which the selected cases have had an impact on policy and practice.

The economic analysis is based on an assumption that the NHS is an evidence-based institution. Thus, it assumes that an intervention which had not been in place before a study had been conducted would be introduced in the NHS once the study had established its benefits over the existing standard of care. Based on this assumption, the analysis focuses on HTA studies which demonstrate that a new intervention offers benefits over existing care, since studies which show no benefit to a new intervention do not offer direct benefit to the NHS. In addition, the analysis only considers the direct benefits that would have accrued to the NHS. It does not consider wider outcomes from the studies, such as further externalities on research or funding.

The economic analysis consisted of two phases. The first involved identifying a sample of HTA studies that meet the study's inclusion criteria. The second involved calculating the potential net benefit that would have accrued if the interventions were implemented and comparing this benefit with total spending on the HTA Programme since 1993.

It was expected that a small number of studies provided most of the potential benefit from the HTA Programme. Hence, the evaluation team purposively selected HTA studies that were judged likely to have the largest potential benefits. Studies were screened to ensure they met the following criteria:

- The study found that the intervention trialled is either cost-effective or cost-saving compared to the standard of care.
- The study reported health outcomes in conventional quality-adjusted life years (QALYs).
- The study showed that the probability that the intervention is either cost-effective or cost-saving is greater than 60 per cent at the £20,000 threshold for QALY value.

The identification of studies was an iterative process involving two stages. First, the NIHR's Evaluation, Trials and Studies Coordinating Centre (NETSCC) identified an initial list of ten HTA studies which would likely demonstrate a high potential benefit in terms of either cost savings or QALYs gained. In the initial sample of ten studies, only three met the inclusion criteria. Thus, in the second stage, the NETSCC provided a further list of studies with likely high potential benefit and the evaluation team additionally scanned the most recent two volumes of the NIHR's HTA Journal to identify studies with high potential benefit. The two most recent volumes of the journal were chosen to ensure that the studies were up to date with the most recent guidelines with regards to discounting costs and health effects and conducting a comprehensive cost-utility analysis.

The second phase of the economic analysis calculated the potential net benefit that would have accrued if the interventions in the selected HTA studies were implemented. The total net benefit of the studies was defined as the monetary benefit that would have accrued to the healthcare system if the recommendations of these ten studies had been implemented for one year, net of the costs of the intervention. According to the study's authors, it is likely that where interventions are implemented, they would be in place for longer than a year before they are superseded. Thus, the total net benefit calculated by the study is a conservative estimate. Under the conventional assumption that the value of a QALY in the UK is £20,000 to £30,000, the net benefit per case can be estimated as: (Incremental Health Gain(QALY) x Pay

Threshold) – Incremental Cost. The size of the affected population was either obtained from the studies themselves or from high quality academic sources or systematic reviews. The potential net benefit was compared to total spending on the HTA Programme since 1993

In addition to the economic analysis, the evaluation team conducted ten case studies. These were separate from the studies identified in the economic analysis, although there was some overlap. The aim of the case studies was to determine the wider impacts of a set of projects with a high potential for high impact. In particular, the evaluation team sought to determine impact on policy and practice. The case studies were selected in consultation with NETSCC, drawing on their knowledge of the HTA portfolio to identify studies likely to have an impact on policy and practice, including some of the studies included in the economic analysis. For each case study, data was collected based on document review and interviews with a member of the team involved in the HTA study. The case studies were drafted using a standard reporting template to aid comparability and ensure all case studies covered key main points.

22.5.2 Data sources

- **Economic benefits associated with HTA research**. The majority of data for the economic analysis was obtained directly from the selected HTA studies. This included data on the prevalence of the health condition in question, the incremental health effect, incremental costs relevant to the NHS, the estimated longevity of the health gain, whether treatment was seen as cost-effective by the study authors (compared to the monetised value for a QALY) and the overall conclusions on the implications of the study. Data from each study was extracted and entered into a spreadsheet.
- **Size of affected population**. Where the population size affected by a health condition was not available from the relevant HTA study, this figure was estimated based on data from high quality academic sources or systematic reviews.
- **HTA Programme costs.** The costs of the HTA Programme were obtained from the programme’s administrative data.
- **Case studies**. The case studies were developed by reviewing the relevant HTA publication and related journal articles. Where relevant and available, the evaluation team also reviewed systematic reviews, NICE guidance on the topic, other policy guidance on the topic, the project website and publications regarding other closely related studies. For all but one of the case studies, the evaluation team interviewed the chief investigator for that study.

22.6 Main findings

The economic analysis estimates that the ten HTA studies analysed provided a potential net-benefit of £3.0 billion based on a value of £20,000 per QALY, and £5.0bn based on a value of £30,000 per QALY. According to NETSCC, the total research cost of the HTA Programme since 1993 was £317m, with the estimated overall cost at £367m. The estimated overall cost of the HTA Programme includes the cost of NHS support for HTA research. The study therefore concludes that 12 percent of the calculated potential net benefit would cover the total cost of the HTA Programme from 1993 to 2012.

With regard to the ten case studies, four had a clear impact on either policy or practice or both. A further five showed some evidence of impact on either policy, practice or both. In these cases, either the impact had occurred but it was less clearly linked to the research, or the impact was forthcoming at the time of the report's publication. More specifically:

- Three cases had a clear impact on policy through citation in guidance, with a fourth expected to be included in guidance that was forthcoming after publication of the evaluation report.
- Three cases showed a clear impact on NHS practice. A further three cases showed some evidence of changes in practice. It was not possible to attribute these changes to the relevant HTA study with a high level of confidence, although the evaluation team reports it is likely the studies played a role.
- A number of the case studies investigated involved research which suggested that a procedure should be removed from practice, as it costs money and does not provide any significant benefit. In such cases, putting the findings into practice should be straightforward since there is no 'new' treatment to introduce. In the cases documented, however, it proved challenging to overcome the existing views and habits of practitioners.

22.7 Analysis of methodological suitability and effectiveness

The methodology employed in this evaluation has a number of strengths.

- It employs an iterative purposive sampling strategy that is well suited to identifying highly successful projects in an RD&I programme, combining the expertise of programme staff with objective case selection criteria. This allows evaluators to address the challenge of evaluating programmes with highly skewed impacts by conducting in-depth analysis of highly successful cases. Importantly, the resulting sample of cases can be analysed using quantitative and/or qualitative methods, as appropriate for the evaluation objectives.
- It employs a credible approach to monetising the potential net benefit of HTA research by aggregating the findings of ten key HTA studies and modelling their economic implications.

It is also important to recognise that this methodology has several limitations:

- It does not address the key issue of programme additionality/ deadweight. While the evaluation findings clearly establish that research supported by the HTA Programme provided a very high potential net benefit compared to programme costs, the methodology employed cannot provide insights into whether these studies were likely to have been undertaken in the absence of the programme.

- The studies included in the economic analysis were partly selected based on whether they reported health outcomes in conventional QALYs, rather than based entirely on their potential impact. Thus, studies with high potential impact may have been excluded from the economic analysis due to measurement issues rather than for substantive reasons.

22.8 Transferability and context

The iterative purposive sampling strategy employed in this evaluation is appropriate for evaluating UK RD&I programmes that are expected to have impacts which are highly skewed towards a small number of very successful initiatives/projects. Importantly, the resulting sample of cases can be analysed using quantitative and/or qualitative methods, as appropriate for the specific objectives of a given RD&I evaluation.

This approach is particularly effective for conducting a value for money analysis when programme impacts are skewed toward a small number of highly successful cases and the evaluation faces resource and/or data limitations. Restricting the estimate of monetary benefits to high impact cases allows evaluators to arrive at a credible conservative estimate of a programme’s overall value for money with fewer considerably fewer resources than required to collect data for all of the initiatives/projects implemented by a programme.

Finally, the study’s use of findings extracted from research papers in order to estimate the returns on research funded by the programme is valuable in the context of evaluating this particular programme. It is important to note, however, that this approach has limited transferability. To implement such an approach, it is necessary that a significant proportion of the research publications funded by an RD&I programme report readily monetisable estimates of the research’s applied benefits.

22.9 Conclusions

This study presents a credible conservative estimate of the potential economic returns on research funded under the NIHR’s HTA Programme, as well as providing insights into wider impacts of the programme on policy and practice in the NHS. It combines economic modelling with a case study approach and focuses on HTA-funded studies that have the highest potential for high impact. The evaluation finds that the HTA Programme provided a potential net-benefit of at least £3.0 billion, compared to an estimated overall programme cost of £367m. It also documents clear impacts that a number of HTA-funded studies have had on either NHS policy, practice or both. Elements of the methodology employed in this evaluation are particularly appropriate for evaluating UK RD&I programmes expected to have impacts that are highly skewed towards a small number of very successful initiatives/projects. More specifically, this study’s approach to can be effective for conducting a value for money analysis in the context of resource and/or data limitations, where programme impacts are skewed toward a small number of highly successful cases.

23 The ISIS Neutron and Muon Source at the Rutherford Appleton Laboratory: Lifetime Impact Study

23.1 Overview

Key characteristics	
Country	United Kingdom
Institution	Science and Technology Facilities Council (STFC)
Type of RD&I	Basic and applied research
Type of Intervention	R&D infrastructure
Evaluation challenges	Lagged effects, intangible benefits, attribution The key challenges are the lagged effects of infrastructure investments, the intangible nature of many of the benefits derived from the research and discoverings enabled by the facility, as well as their attribution to this investment
Evaluation methodology and methods	• Theory-based (Desk research; historical analysis; survey; case studies) • Economic analysis of 10 representative case studies • Bibliometric analysis

23.2 Introduction

This case analyses the lifetime impact study of the ISIS Neutron and Muon Source at the Rutherford Appleton Laboratory, funded by the UK Government, over a 30-year span, since its foundation in 1985. The evaluation assesses the immediate direct economic impact of this facility and characterises the wider economic impacts among industry users and suppliers.

The main challenges faced by this impact study are lagged effects, the intangible nature of many of the benefits derived from the research and discoverings enabled by the facility, as well as their attribution to this investment. To address these challenges, the evaluation utilised case studies, alongside bibliometric analysis to capture direct and indirect impacts of the intervention and ten extended case studies to determine the immediate economic impact of ISIS.

23.3 Background description of RD&I intervention

The ISIS pulsed neutron and muon source is part of the Rutherford Appleton Laboratory (RAL), a world-leading international facility for research in the physical and life sciences, in operation since 1985. Its operational life has been expanded from 20 to 30 years, with a refurbishment and updating of the facilities after 20 years of operation, given its positive and broad impact.

This unique facility allows scientists to study materials at the atomic level using a suite of custom-designed instruments and analysis tools. The research undertaken at ISIS facility aims to generate advances in understanding, as well as scientific knowledge and technical development in instrumentation derived from the multiple experiments carried out in the facilities. Likewise, the facility provides specialised opportunities for industrial users who can access to services and benefit from use-oriented research. These R&D activities also allow the development of experimental skills by students, post-docs, and industrial users.

Access to the facilities is mainly through the main user programme, which accounts for more than 85% of the beam time per year. Proposals are evaluated on their scientific merit by Facilities Access Panels. Another access route is the paid proprietary access, which provides third parties with. The option to pay a commercial rate to gain access to beam time and to use particular instruments, meaning private use of ISIS. This represents 1% of ISIS funding stream. In 2011, a new access mechanism was introduced, the Industry Collaborative Research and Development Programme (ICRD). This flexible access form allows industry users to choose whether or not to publish their results publicly and is a fast-track means to access ISIS during the first two years of the programme. The programme has become the industry user's preferred route of access to the facility.

In terms of investments, the evaluation team estimated that rebuilding the facility would cost around £0.5 billion. ISIS initial infrastructure and equipment were estimated at ~£130 million, while capital investments (building, instruments, equipment) were around £245 million between 2004-2013. The refurbishment costs totalled ~£260 million. By the time of the evaluation, and considering depreciation of capital of 5%, in 2013, these assets had a value of £176 million.

23.4 Challenges measuring impact

The main challenges faced by the evaluation are the attribution of impacts and their quantification, alongside the lagged effects of the realisation of impacts due to the type or research being undertaken at the facilities. Equally important is the low observability of some of the benefits.

To address this, the evaluation focuses on the immediate impacts, counting direct employment and expenditure over time and using standard economic multipliers to arrive at an estimate of the direct and indirect effects. To this end, the evaluation team used ISIS's financial and HR figures to estimate the immediate economic impact as well as the inputs provided in the survey by supplier and academics. Estimates of future impacts prepared by ICRD users were also included to provide a more comprehensive basis for estimating likely future benefits for users.

Tracing and assessing advances in understanding is difficult because they permeate new technologies and methods resulting in new products and processes. The selected way to tackle this low observability challenge, was by developing in-depth longitudinal case studies.

The evaluation had to assume a pragmatic approach to capture the (many) intangible benefits derived from ISIS and considering that when quantifiable they cannot be monetisable and therefore, impossible to aggregate. This led the evaluation team to stick to the immediate economic impact of the intervention and capture intangibles through case studies.

23.5 Methodology and data sources

23.5.1 Methodology

This evaluation adopts a mixed-method approach to assess the impact of ISIS over 30 years and characterise the wider economic impact of this investment amongst industry users and suppliers. The analysis utilises primary and secondary data collected from several sources. A noteworthy aspect of this evaluation is the collaborative work between the evaluator and the team at STFC/ISIS during this study.

The methodology consisted of:

- A historical analysis: A programme of interviews with ISIS directors (previous and current), supported by targeted desk research to arrive at an overview of the facility, its development over time (new investments, new beamlines, new targets, new instruments, etc.) and funding levels.
- Desk research: to develop an analytical framework that entails all the types of research undertaken in ISIS, skills developed and direct and indirect economic benefits likely to be realised. Additionally, desk research to gather data on the anticipated financial impact of the industry using ISIS instruments, through the ISIS Collaborative R&D (ICRD) programme
- A series of online surveys of university academics, companies and suppliers who use or work with ISIS to profile usage and to invite people to relate the role of ISIS and neutron science in their own work. With this, it was possible to draw conclusions about ISI's relevance for their research and to gather brief insights into academic achievements and subsequent benefits associated with the facility. The survey covered the impact of ISIS on human capital, career progression and intersectoral mobility. A total of 200 responses were obtained from academics. The industry survey was applied via academics but did not provide significant input, with only 10 responses. Finally, the supplier survey was run with large and high-value instrumentation providers from which 50 responses were obtained and provided valuable insights on the benefits derived from their interaction with ISIS.
- Case study development highlighting key examples from a substantial archive of information from ISIS. From a sample of 50 good examples, 30 case studies were

developed. Ten of them were developed in more detail to estimate the economic impact of ISIS.

- A bibliometric analysis: to profile journal use and citations for ISIS-related publications from searchers in the Scopus database

23.5.2 Data sources

- ISIS archive of annual reports and case studies were used to inform this rapid review
- Monitoring and evaluation data of the ISIS Collaborative R&D (ICRD) programme
- ISIS case studies archive
- Surveys of academic, industry and supplier stakeholders
- ISIS's financial and HR figures

23.6 Main findings

The evaluation finds that ISIS has created substantial long-term impact. Since its creation, ISIS has delivered major social and economic benefit for the UK and global economies, generating a wide range of benefits, and playing an essential role in global research agendas that can have a worldwide impact.

ISIS research has delivered a net economic impact of £400 million based on its operation and the employment of highly qualified personnel. Overall, the evaluation found that ISIS had generated at least £1,049 millions over its lifetime (1985-2013). The net present value of recently complemented research was estimated at £340 millions up to 2025.

The evaluation also estimated a Return on Investment (ROI) figure for ISIS around 214% that derived from total ISIS expenditure in the order of £700 million (£650 million being provided by the UK government) and the overall economic benefit of £1,389 millions. It is important to note that the evaluation used income data as a proxy for expenditure data.

In terms of publications, ISIS outperforms the UK average with selected publications achieving several 100 citations within a three-year window. Similarly, ISIS has enabled innovations to instrumentation that could have gone unexplored without this facility. At the time of evaluation, ISIS had 26 ongoing international agreements across 12 countries and had received at least £56 million in international contributions over its lifetime.

23.7 Transferability and context

This is an example of a UK evaluation. Hence, context-related elements do not represent an issue for the transferability of the methodology. Some considerations when using the method to assess other types of STI infrastructure funding refer to data access to calculate future impacts derived from the use of the facility.

For example, ISIS archives and case studies prepared by the STFC evaluation units helped significantly to gather evidence for case studies and to facilitate longitudinal tracking of outputs, outcomes, and impacts. Likewise, monitoring and evaluation data of the ICDR programme was instrumental for the economic analysis considering lagged effects of the research undertaken in ISIS. Data collected over long periods, in particular, at the initial stages of the facility's establishment is crucial for the economic analysis. In this case, the evaluation had to extrapolate employment and expenditure data to fill data gaps of initial periods of ISIS operation.

An administrative limitation as to the data held by ISIS from industry user proved to be an important operational challenge, which may be relevant to consider when assessing similar interventions. The evaluation team circumvented the issue, asking academic partners to share with their industry partners the survey. However, running the survey indirectly resulted in low response rates and an incomplete picture of the benefits from this type of user based on the few responses (10).

23.8 Conclusions

This case study has looked at the Lifetime impact evaluation of the ISIS Neutron and Muon Source over its 30 years of operation. Measuring the impact of this investment faced three main challenges associated with the lagged effects of thousands of experiments and research being undertaken in this facility, the attribution of impacts to ISIS as well as capturing the intangible benefits enabled by such infrastructure.

The evaluation adopted a mixed method approach underpinned by in-depth research of ISIS activities, historical data, monitoring and evaluation data of the Industry Collaborative Research and Development Programme (ICRD) and primary information from a programme interview.

Impact is assessed through 30 case studies, ten of them from the ICRD programme featuring an extended analysis and providing input to calculate the future economic benefits of the programme. Overall, the evaluation found that ISIS has delivered wide social and economic impact, generating £400 million of net economic impact and a ROI of 214%.

24 Strategic Priority Fund. Interim impact evaluation

24.1 Overview

Key characteristics	
Country	United Kingdom
Institution	UKRI
Type of RD&I	Basic and applied research
Type of Intervention	Research Grants
Evaluation challenges	Low observability of impact research Measuring the degree of multidisciplinarity of research outputs and the extent to which this exceeds what is already produced via other sources of funding, tracing knowledge flows and the uptake of results beyond academia (in particular, among policymakers)
Evaluation methodology and methods	• Theory-based (Desk review; case studies; interviews; focus group) • Bibliometric analysis • Case studies

24.2 Introduction

This case study looks at the evaluation of the Strategic Priorities Fund (SPF) launched in 2018 with three objectives, drive an increase in multidisciplinary and transdisciplinary research and innovation (MIDRI), address cross-departmental R&I priorities and ensure that UKRI's investments are well linked up across departments and respond to strategic priorities or opportunities. Two main challenges are found in assessing the impact of SPF, the low observability of research impact and the additionality of the funding. To address these, the evaluation adopts a comprehensive methodology, consisting of a bibliometric analysis and case studies that cut across three levels, research outputs, proposals and projects and programmes. Through this analysis is possible to quantify and characterise the aspects of MIDRI and identify pathways to impact.

24.3 Background description of RD&I intervention

The Strategic Priorities Fund is part of a wider package of UKRI measures to support the delivery of the Industrial Strategy launched in 2017. Its purpose is to strengthen the UK's research capacity as a world leader in R&I and address gaps in UK research funding highlighted in the Nurse Review50 in three main areas, the UK research system's awareness and coordination of strategic research efforts across the research councils and government, support of multi- and inter-disciplinary research (MIDRI), and the ability to respond quickly and materially to emerging challenges or opportunities.

SPF's objectives are:

- Drive an increase in multidisciplinary and transdisciplinary research and innovation: by de-risking the process of preparing and submitting MIDRI proposals and improving the efficacy of the funding system in assessing MIDRI proposals
- Address cross-departmental R&I priorities ensuring UKRI's investments and well link up with cross-departmental research and innovation priorities and opportunities.
- Respond to strategic priorities or opportunities, improving the agility of the funding system to respond to emerging opportunities and providing a funding route for medium-scale programmes

SPF's budget of £831m has supported 34 programmes in two waves, one launched in 2018, which allocated £334m to 15 programmes. In 2019, Wave 2 awarded £497m to 19 programmes. Programmes encompass a wide range of R&D activities at different levels of maturity, each seeking to tackle at least one of the three fund's objectives described above.

Another important feature of SPF is that all UKRI councils are leading at least one programme and partnering on others, alongside devolved administrations, government departments and executive agencies that involve significant budgets.

SPF portfolio covers a wide range of themes, from research on productivity, biology and medicine, health, wellbeing and human rights, environment, AI and digital and infrastructure. The health theme was incorporated in Wave 2.

SPF's programmes tend to be medium-scale, around £10m+, addressing a gap for larger, more complex projects. In addition to research funding, the funding has also covered spending reviews and allowed programmes to address emerging opportunities and priorities at scale. To the point of the interim evaluation (17 June 2022), SPF has supported 767 individual projects through open competition, and this number is expected to increase as only 3 out of the 34 programmes funded have been finalised as of December 2022.

---

50 Ensuring a Successful UK Research Endeavour, BIS/15/625, Nov 2015

24.4 Challenges measuring outcomes and impact

One of the three SPF's objectives is to increase the degree of MIDRI in the research funded by UKRI and across research councils. It is expected that more multidisciplinary and interdisciplinary research generate a wider impact and help address more complex challenges envisaged in the Industrial Strategy 2017.

Two primary challenges are faced by this evaluation. One relates to low observability of the research outcomes, particularly, assessing multidisciplinarity and transdisciplinarity in research and capturing their impacts. The first issue is that involving more diverse stakeholders and disciplines may increase the complexity of the research undertaken and reduce or delay the realisation of impact. Likewise, MIDRI can also generate other types of outcomes intended for other audiences beyond the academic community. For example, seeking to inform policy and substantiating policy changes. Also, part of this challenge is to assess how knowledge flows across a wide range of actors in different domains and the uptake of results beyond academia.

The second challenge is to determine the additionality of the SPF in driving an increase in MIDRI compared to what is being produced under other funding mechanisms from the UKRI and research councils. SPF is expected to de-risk the process of preparing and submitting MIDRI proposals and strengthen the efficiency of the funding system in assessing them. Capturing this requires an analysis at multiple levels and involves a variety of actors intervening in different processes, who generate outputs at various levels.

24.5 Methodology and data sources

24.5.1 Methodology

The purpose of the evaluation is to demonstrate what the fund has delivered and build the evidence base on what works in supporting MIDRI and ensuring R&D responds to strategic opportunities and priorities of the UK's government. This evaluation is expected to inform ongoing and future improvements to the Fund.

Guided by a theory-based approach, the methodology combines qualitative and quantitative methods to capture the main outputs emerging from individual programmes and determine to what extent the Fund objectives have been achieved and how. Each objective is analysed at three different levels, outputs, proposals/projects and programmes.

The evaluation methods used to assess the SPF are:

- **Desk review of programme documentation and secondary data (M&E data):** to assess the programme governance and its composition and to provide inputs for the bibliometric analysis carried out at the outputs level.
- **Bibliometrics**: to analyse publications emerging from the programmes funded and measure MI (diversity of co-authors' disciplinary backgrounds) and diversity of

knowledge and contrast SPF against benchmarks. The bibliometric analysis of research across disciplinary boundaries is undertaken in two dimensions:

- A human dimension, capturing collaborations of researchers with different disciplinary backgrounds (an input to cross-disciplinary work, and a proxy of multidisciplinarity). The evaluation captures disciplinary diversity amongst a publication's author list (multidisciplinarity) by adapting the Rao-Stirling diversity index[51] to the disciplinary profiles of co-authors, while disciplinary diversity in a publication's integrated knowledge (interdisciplinarity) is inferred by applying the Rao-Stirling index to the disciplines represented in the papers' reference list.
- An epistemic dimension, capturing publications that draw on knowledge from different disciplines (an output from cross-disciplinary work, and a proxy of interdisciplinarity).

To provide a point of comparison, these diversity indicators are computed for different groups of papers, for SPF and UKRI overall for two periods (2006-2017 and 2018-2021). Comparator groups were selected from four publication sets as adequate benchmarking references against which to measure the performance of SPF papers:

- UK papers - all publications with at least one UK-based author
- UKRI papers - all UK publications with funding from a UKRI council, identified in GtR and Scopus acknowledgements
- Prior publications from SPF researchers. These are papers authored by SPF researchers and published before the first year of any of the SPF projects in which the researcher has participated and published between 2006 and 2019
- Parallel publications from SPF researchers. These are papers authored by SPF researchers after their first year in any SPF project, excluding those that have been identified as an SPF paper in GtR or Scopus. These are presumed to be publications associated with concurrent projects by SPF-funded researchers. They include papers published between 2018 and 2021. A remark is made in the evaluation as to parallel papers group may include SPF papers not correctly identified as such in GtR or Scopus acknowledgements.

From the 767 projects awarded until 2022, only 25% of them have associated publications reported in GtR, which have been used as input for the bibliometric analysis. Considering this, an important remark of the evaluation is to use and interpretate these interim results with caution as outputs may change as the SPF portfolio evolves and grows. A drawback highlighted by the evaluation team is that GtR does not provide a comprehensive list of all researchers involved in each SPF project. Therefore, the analyses of the multidisciplinarity of research teams are focused on the remaining research grants that listed at least 3 researchers in GtR.

51 https://doi.org/10.1098/rsif.2007.0213

- **Programme lead consultation:** to understand in detail aspects of the programme design aimed at increasing the diversity of disciplines and analyse the actions adopted to attract MIDRI projects and assess them. These consultations consisted of four steps, a first initial request for basic information on the stakeholders involved in programme governance. This was followed by a programme information template, providing information on their programme and its implementation, as well as high-level views and perceptions on several key areas relating to Fund objectives. Semi-structured Interviews were conducted with the programme mainly to inform case study development. In both cases using self-reported data from programme leads. Finally, two workshops were held with programme leads and representatives from UKRI to discuss how SPF has driven an increase in high quality how SPF has driven an increase in high quality MIDRI.
- **Wider stakeholder consultation** (survey of programme advisory board members and stakeholders' interviews) to obtain further insights into their role and activities within the advisory group, their experiences of the SPF programme and their views on the added value of SPF objectives.
- **Longitudinal case studies of SPF programmes**: aimed to assess eight SPF programmes and developed throughout the evaluation. In the first stage, the cases describe the programmes and their origins, as well as any early learnings and preliminary outputs. In the second stage (interim) additional evidence about outputs, outcomes, and direction of travel in terms of impact achievement.

Each starts by exploring how the programme has been designed and its operational and day-to-day management. Each then discusses outputs, outcomes, and impact, based on the 3 objectives of the fund, and finishes by exploring what is expected to be seen in the final report and by the final evaluation.

24.5.2 Data sources

- Programme documentation and secondary data from GtR and Scopus
- Questionnaires and interviews for programme lead consultation at each phase of the evaluation (baseline, interim, and final). For the second phase, primary data was obtained from an initial request for basic information, a programme information template (questionnaire), interviews and two workshops.
- Surveys of programme advisory board members (wider consultation), consisting of 250 members, and excluding those to be interviewed as part of the case study development, a total of 183 individuals of which 90 provided responses (49% response rate).
- Stakeholder interviews (wider consultation) for a total of 42 interviews that were tied to the development of case studies.

24.6 Main findings

The evaluation found that SPF has established a portfolio of programmes that align with its high-level objectives about MIDRI, government priorities and system agility, with a centralised process for funding allocation and a decentralised process for programme design and implementation. SPF has supported 34 programmes in two waves. Each programme exhibits a diverse range of R&I activities and projects at various stages of maturity across the spectrum of R&I areas. Each programme addresses at least one of the main objectives of the Fund.

SPF programmes sit around £10m+ each closing a gap for larger and more complex projects and allowing flexibility to address emerging opportunities and priorities at scale. A total of 767 projects have been funded under these programs, expecting to fund more in the future. As of December 2022, three of the 34 programmes have been finalised.

SPF's programmes outcomes are making progress, although performance around publications per £m invested remains lower than for the wider UKRI portfolio. In contrast, citations within policy-related literature are three times higher for SPF outputs compared to other UKRI grants. These findings are associated with the complexity of MIDRI and the delay this can cause in the realisation of outputs.

The evaluation confirmed that SPF is helping to drive an increase in high-quality multi-disciplinary and interdisciplinary research and innovation from applications to research teams and publications, through to synthesis and dissemination. The focus on supporting MIDRI through funding programmes aimed at MIDRI, challenge-led programmes addressing cross-departmental government priorities and embedding wide participation throughout the programme cycle are found the main mechanism enabling this result. Additionally, MIDRI applications have been attracted via call text and criteria, and most programmes adopted processes specifically designed for the assessment of MIDRI, of which 57% are considered enhanced or new.

Regarding government R&I priorities, the Fund has helped address this via additional spend, and increased link-up of department priorities and strategies represented in cross-collaboration in programmes and projects. SPF has also facilitated the dissemination, uptake and use of outputs developed under the SPF's funded research. Research outputs have also informed policy decisions in priority areas, and government departments' participation in the production of publications, and the involvement of Public Sector Research Establishments PSREs.

24.7 Analysis of methodological suitability and effectiveness

This evaluation adopted an encompassing methodology, with mixed methods to look at the outputs and outcomes of SPF programmes at three different levels. This allows the evaluators to address the various aspects of MIDRI while assessing output performance from various processes, including programme design, calls administration, and research.

The bibliometric analysis is well suited to establish the multidisciplinarity and transdisciplinarity of outputs from human and cognitive dimensions, computing the Rao-Stirling diversity index. By doing so, it is possible to assess the diversity of research teams and disciplines brought into their research, making it easier to observe and measure outputs and outcomes, hence, addressing the challenge of low observability of impact.

Further characterisation of the outputs is done by case studies and consultations at the programme and proposal levels, helping trace the flow of knowledge among diverse audiences. Such triangulation of evidence also contributes to tackling low observability of impact that would be missed otherwise. For example, a key finding from triangulation is that emerging publications are aimed at policymakers three times more frequently than those funded by UKRI and that MIDRI fosters broader participation in research outputs, even from policymakers and public research institutions.

Despite the results of the interim evaluation do not account for most projects (25% of the total funded), the use of comparison points with four different sets of publications provides a good counterfactual and an indication of SPF's delivery of its MIDRI objective. This helps measure the additionality of the funding and provides insights into the timing and complexity of MIDRI.

Limitations regarding the completeness of project data recorded in GtR and how that may affect the bibliometric analysis are handled well, reducing the sample of projects to those including at least three authors. Although this may result in fewer projects providing evidence on aspects of MIDRI, it ensures consistency in the analysis.

Case studies provided a deeper analysis of how SPF has de-risked proposal development and led to adopting new ways to improve grant evaluation. This evaluation component is expected to be strengthened in the final stage of SPF evaluation.

24.8 Transferability and context

This evaluation is undertaken in the UK context. It utilises primary and secondary sources of information that would require tracking publications linked to the programme via monitoring systems such as Researchfish and Gateway to Research GtR. Similarly, identifying SPF (and non-SPF) researchers can be done from monitoring systems and data, but requires a great deal of data cleaning to match with bibliometric data, which may be a costly and time-consuming task. Another factor to consider is that the analysis requires access to proprietary data sources that link grant, publication, and policy document data.

Apart from the bibliometric analysis, the case studies build on primary data, focusing on processes around programme design and call evaluation, which can be easily adapted in other evaluations of similar characteristics.

24.9 Conclusions

This case analysed the evaluation (interim) of the Strategic Priorities Fund commissioned by the UKRI and delivered in December 2022. Using a mixed methodology, the evaluation assesses the extent to which the Fund has delivered on its three main objectives: drive an increase in multidisciplinary and transdisciplinary research and innovation, address cross-departmental R&I priorities and ensure that UKRI's investments are well linked up across departments and respond to strategic priorities or opportunities. A three-level approach looking at outputs, proposals/projects and programmes is used to assess the aspects of MIDRI.

Two main methods are used to assess progress on SPF objectives, a bibliometric analysis and case studies to quantify and characterise aspects of MIDRI at the three main levels. Despite outputs data providing an early indication of impact (covering 25% of total projects), triangulation of data offers valuable insights into the complexity of MIDRI, and what has worked to drive an increase in MIDRI and the pathways to impact. The evaluation provides a good overview of the progress made on each of the Fund's objectives and its methodology helped address the low observability of research impact and determine the additionality of the funding.

This publication is available from: www.gov.uk/dsit **[replace with direct URL if known]**

If you need a version of this document in a more accessible format, please email alt.formats@dsit.gov.uk. Please tell us what format you need. It will help us if you say what assistive technology you use.