European Commission

Policy Support Facility Mutual Learning Exercise: Evaluation of Business R&D Grants Schemes

Combining Mixed Approaches to Evaluation

Thematic Paper No 3

March – 2018

Research and Innovation

Policy Support Facility Mutual Learning Exercise: Evaluation of Business R&D Grants Schemes - Combining Mixed Approaches to Evaluation - Thematic Paper No 3

European Commission
Directorate-General for Research and Innovation
Directorate A: Policy development and Coordination
Unit A4: Analysis and monitoring of national research and innovation policies

Contact (H2020 PSF MLE EVALUATION OF BUSINESS R&D GRANT SCHEMES):
Eva Rueckert, Coordinator of the MLE, Unit A.4 - eva.rueckert@*ec.europa.eu*
Contact (H2020 PSF coordination team):
Román ARJONA, Chief Economist and Head of Unit A4 - Roman.ARJONA-GRACIA@ec.europa.eu
Stéphane VANKALCK, PSF Head of Sector, Unit A4 - Stéphane.VANKALCK@ec.europa.eu
Diana SENCZYSZYN, PSF Team Leader, Unit A4 - Diana.SENCZYSZYN@ec.europa.eu

European Commission
B-1049 Brussels

Manuscript completed in March 2018.

This document has been prepared for the European Commission however it reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

© European Union, 2018.
Reuse is authorised provided the source is acknowledged. The reuse policy of European Commission documents is regulated by Decision 2011/833/EU (OJ L 330, 14.12.2011, p. 39).

Cover Image © Eurotop.be 2017

EUROPEAN COMMISSION

Policy Support Facility Mutual Learning Exercise: Evaluation of Business R&D Grants Schemes

Combining Mixed Approaches to Evaluation

Thematic Paper No 3

Prepared by the independent expert:

Paul Cunningham

Directorate-General for Research and Innovation

2018 EN

Table of Contents

List of Boxes

1 INTRODUCTION

This paper has been prepared for the Mutual Learning Exercise (MLE) on the evaluation of business research and development (R&D) grant schemes in European countries. It addresses the topic of applying mixed-method approaches to the evaluation of public schemes to support R&D and innovation in firms, specifically business R&D grants and associated innovation schemes. The first section outlines the objectives and structure of the report.

The report then sets out the broad context and rationale for the use of mixed-method approaches in the evaluation of innovation support schemes, with a focus on business R&D grants. Clearly, with such a very wide ranging topic, in order to be concise, the report has restricted its scope to a relatively narrow set of issues.

The subsequent sections are then structured around the major issues addressed by the component sessions of the site visit. These are:

The bigger picture – an overview of the support and evaluation of innovation in the UK. This section examines the role of UK Research and Innovation (the UK's new overall coordinating body for research and innovation policy and support which will become operational in 2018). It also looks at Innovate UK's approach to programme planning and evaluation.

Building evaluation capability – examining the major evaluation challenges perceived by Innovate UK and the activities of the What Works Centre in experimental approaches to evaluation and in disseminating the lessons learned.

New methodological approaches: A series of introductions to recent studies that have either used novel and innovative approaches to evaluation or have used more traditional approaches in innovative ways. These include:

- A comparative study which attempts to better understand business characteristics and the impact of R&D grants (including a focus on the Economic Research Centre's work on advanced econometrics and big data sources). This also introduces the idea of 'growth hero' companies.

- The combination of advanced econometric and other approaches (with a specific focus on the use of Regression Discontinuity Design in an evaluation of the UK's Biomedical Catalyst initiative).

- Using mixed evaluation methods to overcome the problems of time lag affects, with examples from some new evaluations of a long-standing UK support scheme for SMEs – Smart.

- Approaches aimed at addressing issues of target heterogeneity, complex anticipated outcomes and behavioural additionality; looking at the use of case studies, contribution analysis and Agent Based Modelling in the evaluation of the UK's Catapult Centres.

- Approaches involving the design of policy instrument delivery for improved evaluation, focusing on the use of Randomised Control Trials (RCTs) and other forms of experimentation in evaluation design.

Each section offers a number of concise examples of the approaches discussed and the main lessons emerging (presented as boxes).

The final section presents a number of the key findings and conclusions arising from the site visit.

1.1 Objective and background of this report

Business R&D grant schemes represent a specific type of public policy instrument intended to support R&D and/or innovation by companies. In their simplest form, such schemes are provided to individual companies (sometimes according to size or sector) although there can be subsidiary (and complementary) objectives, such as stimulating collaboration with other companies, universities or research institutes, etc. Consequently, the effects of these schemes may extend beyond the immediate beneficiaries (recipients) and their internal R&D and related innovation activities, to collaborating partners, value chain and supply-side partners to the wider local, regional or national innovation system.

The *ex-post* evaluation of business R&D grant schemes also formed the topic of the first cycle of this Mutual Learning Exercise (MLE) which took place in 2016 (see Cunningham et al., 2017). MLE participants include R&D and innovation policy practitioners and experts in European countries.

One of the main conclusions to emerge from this first cycle was that, as regards evaluation methods, there has been and continues to be an increasing trend towards the use of more sophisticated econometric analyses in measuring the impact of these types of policy instruments. However, there is an accompanying need both to better understand the behavioural effects (i.e. 'how' and 'why' such effects are engendered rather than simply measuring 'what' effects have emerged) and to examine the 'innovation journey of firms', particularly when firms use R&D and innovations grants in a simultaneous or sequential manner or combine them with alternative forms of support.

This Thematic Report is the third in the present (second) MLE cycle (2017-18) and explores recent policy developments in the use of mixed-method approaches to evaluation. The first site visit (Oslo) addressed the opportunities and challenges of big data, whilst the second (Stockholm) explored new approaches to understanding and measuring behavioural change and behavioural additionality (as defined as a third type of additionality alongside those associated with input and output additionalities). The final site visit (London) aimed to bring together elements and topics addressed by the previous two site visits and placed particular attention on recent advances in econometrics, the use of control groups, case study approaches and evaluation and programme design. As such, the London site visit attempted to build on the findings of the previous two site visits, and to further explore the use of advanced evaluation techniques (such as econometric approaches and RCTs) both in order to better identify the impact of innovation support interventions and to better understand the underlying mechanisms that contribute to such impacts.

This Thematic Report will, where relevant, use examples that derive from presentations made at the London site visit and which are seen as providing lessons to the broader audience.

2 THE RATIONALE FOR MIXED-METHOD APPROACHES IN EVALUATION

2.1 The context of mixed approaches to evaluation

2.1.1 The objectives and purpose of evaluation

As noted in the preceding Challenge Paper for the London site visit, evaluation serves a number of purposes, under a number of rationales. In the past, evaluation was often used with the primary objective of justification, i.e. as a means of ensuring (or checking) that the resources invested in a specific programme or scheme were being used appropriately and were generating the anticipated results. These early evaluations, often performed for ministries of finance and similar agencies, were constructed around the three pillars of 'economy, efficiency and effectiveness' – i.e. ensuring that the programme received the appropriate level of resources (costs), that it was operated in a cost-efficient way and that it achieved the best outcomes and impacts given the inputs used.

However, such purposes neglected a major benefit of evaluations – that of the opportunity for policy learning – an omission that has been largely rectified. Thus, a major purpose of evaluation now is to inform policy learning at a variety of levels. Typically, three levels of policy learning have been defined:

- Operational learning: In this case, evaluation is used as a management feedback tool to improve the effectiveness, efficiency and quality of policy intervention by the organisation responsible for the implementation of the programme. Essentially, evaluation provides lessons on how organisations (ministries, agencies, etc.) can do things better, in terms of designing, managing and implementing programmes. Lessons may also be learned from the evaluation itself in order to improve the evaluation of future programmes.

- Policy feedback: Here evaluation is used in its 'traditional' sense to determine the outcome and impacts of policy measures and programmes. This type of use is aimed at checking whether, and the extent to which, programmes have achieved their objectives. In addition to contributing to the design of future programmes, evaluation in this context also offers a method for policy makers to assess whether the assumptions they made about the identified bottlenecks and market or system failures which prompted the policy intervention in the first place were, in fact, accurate.

- System impact: At this level, evaluations serve to improve the efficiency and effectiveness of national innovation systems through guiding the design and formulation of intervention policies and programmes. Evaluations are used to provide answers to broader level questions concerning the innovation system, such as when certain interventions are appropriate, which complementary programmes should be used and when, what is the appropriate policy mix needed to achieve the desired effects, and so on" (adapted from VINNOVA, 2004, cited in Miles and Cunningham, 2006).

Another way of categorising the purpose of the evaluation is whether it is intended to be summative or formative.

Summative evaluation typically looks at the impact of an intervention on the target group to find out what the project achieved. It is often associated with more objective, quantitative methods of data collection and tends to be linked to the evaluation drivers of accountability. Summative evaluation is generally outcome-focused more than process-focused and tends to be undertaken *ex post*.

Formative evaluation generally takes place *ex ante* or, more typically, during the implementation of a project with the aim of improving its design and performance. Formative evaluation complements summative evaluation and provides insights into understanding why a programme does or does not work. It also takes account of other

factors (internal and external) that may be influencing the project. Typically, formative evaluation is more resource intensive than summative evaluation although it represents a better investment since it contributes to better policy learning and improved programme design.[1]

Evaluations can also be conducted for a variety of audiences: for instance, programme managers will seek information on the implementation of their policy instruments in order to better improve aspects of design and delivery. At the same time, they will want to assess both the immediate results and longer term outcomes and effects, again to learn lessons on effectiveness and efficiency. Auditors and ministries of finance will still be interested in value-for-money and efficiency – not to mention any significant leveraging effects (the 'bang for the buck') generated. Added to these, a broader range of policy makers will seek to determine the comparative effectiveness of different policy interventions and any synergistic or contradictory impacts they may exhibit.

2.1.2 Types of policy intervention

Typically, at least at the level of the individual policy instrument, policy interventions will vary in their modality (i.e. how they interact with and influence the behaviour of the targets of their support). These modalities, and the targets they address, are of course dependent upon the purpose (whether an aspect of market failure or systems failure) for which the instrument has been designed.

Policy instruments exhibit a large range of modalities, varying according to the specific purpose for which they have been designed and for the context in which they are implemented. In designing a typology for the characterisation of these instruments, the SIPER project[2] identified eight main categories of modality, although each one can be subjected to a range of smaller scale adjustments in terms of the targets it addresses, the eligibility criteria, etc:

1. Direct financial support: grants, loans, guarantees, contracts, etc.
2. Direct financial support: scholarships, fellowships, etc.
3. Direct financial support: (non-project-specific) institutional block grants
4. Indirect financial support: tax and fiscal incentives (e.g. R&D credits)
5. Indirect financial support: norms, standards, regulations
6. Infrastructure support (e.g. provision of access to and construction/upgrading of research infrastructure)
7. Non-financial support (e.g. training, coordination and advisory/information support/provision)
8. Prizes and awards (ex-ante inducement, ex-post performance recognition, etc.).

Although this MLE is concerned primarily with the evaluation of policies to support business R&D, policy makers have recognised the complexity of the innovation process (and of the actors and infrastructures which impact on it) which has led to a shift in policies away from relatively simple grants and loans to more sophisticated packages of support (Cunningham, Gök and Laredo, 2015; Edler et al., 2015). Thus, for example, direct grant schemes are

[1] See: http://evaluationtoolbox.net.au/

[2] See: http://www.si-per.eu/

now likely to include elements of training or inducements for intra- or inter-sectoral collaboration.

Allied to these developments, and again in recognition of the pervasive and complex nature of the outcomes and impacts of the innovation process, the purposes of the support instrument (i.e. what policymakers are seeking to achieve) are also likely to encompass a broader range of goals. An illustration of these policy objectives or goals is also provided by the SIPER typology:

1. Enhancement of education and initial/further training
2. Facilitating personnel mobility (including career enhancement)
3. Internationalisation of RTDI activities
4. Awareness raising and promotion of public acceptance
5. Strengthening/improving research excellence, relevance and management practices
6. Improving absorptive capabilities and capacity
7. Supporting collaborative interactions for the production of new knowledge (including project-focused approaches, innovation vouchers, etc.)
8. Supporting broader (multiple) interactions (e.g. through clusters or networks)
9. Supporting the protection of IP
10. Mobilising additional (non-public) financing for innovation (e.g. support of business angels, Venture Capital Trusts (VCTs), equity schemes, etc.)
11. Stimulation of additional RTDI activity (e.g. increasing R&D expenditure)
12. Strengthening the quality of RTDI activities (promotion of excellence)
13. Creating new RTDI capacity (e.g. new organisations, start-ups, technology-based companies, etc.)
14. Diffusion of innovation (including creation or exploitation of new markets, public procurement of innovation).

In this respect, it is highly likely that the decision to implement schemes to support business R&D will be prompted by the wish to achieve a number of policy goals beyond that of simply stimulating additional RDTI activity in order to increase R&D expenditures (Category no. 11 above) and could easily encompass three or more of the above-mentioned policy objectives. Further context-driven objectives may also be applied, such as the wish to regenerate industries in disadvantaged regions, or to develop technological leadership in emerging industrial sectors, or to help encourage the growth and development of small firms or start-up companies, etc.

2.1.3 Challenges: attribution and policy mixes

Clearly, given the above considerations, any efforts to determine the outputs, outcomes and impacts of a given policy measure will necessitate the construction of a number of measurement indicators, each of which will require significantly different evaluative methodologies. As noted by Miles and Cunningham (2006), evaluation will generally only tell us about certain aspects of the issue in question. "For instance, it may seem a simple matter to determine whether a car or computer is value for money. But consider how many aspects there are to the performance of a car or computer (let alone aspects such as the

design and 'feel' of the product." Likewise, policymakers may ask a simple-sounding question such as, "What are the economic benefits of this investment?", the answer to which can be far from simple. Thus, the evaluators will have to specify the question in terms of operational aspects that are open to measurement, such as, "How did this programme influence the performance of X [a particular set of economic actors] in terms of Y [a particular set of parameters of activity and outputs]?".

In evaluation studies, the impacts and outcomes of interventions are usually highly complicated (particularly so if the interest lies in impacts and outcomes that are further 'downstream' and which extend beyond the merely research- or technically oriented outputs of a scheme). Moreover, different sorts of target will experience impacts on several dimensions. Some of the outcomes and impacts may be captured in official statistics, company reports and other documents; some may require direct investigation. Some of them may be a matter of counting observable phenomena, while others may be more a matter of subjective judgement. Some may be a matter of individual experience, some of organisational behaviour, and some may even cross organisational boundaries (e.g. networks) (Miles and Cunningham, 2006).

As a consequence, the diversity of methods available for performing an evaluation is an acknowledgement of the multiple dimensions in which the impacts of policy intervention might manifest themselves. For this reason, no single best evaluation methodology exists for all purposes and studies. Each methodology will be more suitable for analysing particular dimensions of impacts. In general, an evaluation study will require a combination of various evaluation methods. Thus, for instance, different methods may be used at different levels of data aggregation, or to capture immediate and longer-term impacts. Despite the increased resource-cost implications, the use of more than one method has advantages in that it allows for cross-checking the robustness of conclusions about the observed effects of the intervention – i.e. it permits triangulation (Miles and Cunningham, 2006).

As seen above, even a relatively straightforward evaluation of a single policy measure may entail the use of a mix of evaluation methods. However, there is yet a further level of complexity which can be added to the evaluator's set of challenges. As indicated by the typologies outlined above, it is evident that all policy instruments operate as part of a policy mix. Given the breadth of actors and activities that can be included under the notion of an innovation system, the boundaries of such a policy mix are extremely difficult to define. Guy et al. (2009) found that, even when such a policy mix is limited to the primary policy objective of increasing R&D expenditure, any attempts to investigate the interactions between elements of the policy mix is highly problematic. This finding was further confirmed in work by Cunningham et al. (2015) who, in an extensive review of the evaluation literature, were able to find little substantive evidence of evaluations of entire or partial policy mixes.

Perhaps as a consequence (or cause) of this absence of evaluation activity, Guy et al. (2009) concluded that policy mixes are overwhelmingly the result of an organic process of growth in the policy portfolio, whereby new measures are added in response to policy demands either at the expense of older measures or alongside existing ones. Very rarely is a policy mix constructed in a rational process ab initio. Indeed, given the range of motivations (political, policy, industrial, research, etc.) of the various stakeholders involved in the generation of policy mixes, Flanagan et al. (2011) caution against the assumption of a high degree of rationality underpinning their evolution.

Thus, it is becoming increasingly imperative that policy makers, programme designers and managers and evaluation practitioners engage in greater dialogue in order to develop more sophisticated, yet robust, approaches to evaluation. At the same time, these approaches must contribute to a greater understanding of the innovation processes that are being stimulated by the instruments under consideration and ensure that their results are transparent and comprehensible to a wide audience. Subsequent sections of this report present a number of examples of this form of evaluation policy dialogue.

2.2 Examples of mixed approaches to evaluation

As noted above, no single evaluation approach is able to provide a fully comprehensive picture of the performance of a policy instrument. Even if the policy instrument in question has a relatively simple modality for its delivery and a restricted set of policy goals, it is still often desirable to employ combinations of several methodologies in order to triangulate the aspects of performance to better understand if, how and why a particular instrument is successfully addressing the rationale for which it was designed and implemented.

It should also be noted that, in the past, there have been numerous efforts to develop so-called evaluation 'tool-boxes'[3] – manuals from which policy makers and programme managers may select appropriate evaluation approaches in order to assess the performance of their policy instruments. Examples of such manuals include:

- The ePUB RTD Evaluation Toolbox (2002)
- DG Enterprise and Industry: Smart Innovation: a practical guide to evaluating innovation programmes (2006)
- HM Treasury (UK): The Magenta Book: Guidance for evaluation (2011)
- European Commission, DG Regio: Evaluation of Innovation Activities: Guidance on methods and practices (2012)

However, based on the foregoing discussion, it is clear that 'out-of-the-box' methodologies are not sufficient to answer the questions posed by policy makers and evaluators. Generally, they do not deliver a full understanding of the range of potential impacts on the innovation process policy instruments may have. This is mainly because policy makers now seek to answer more complex questions which require increasingly sophisticated approaches to evaluation.

It is beyond the remit and resources of this report to present a comprehensive account of the methodologies available to the evaluator of innovation policy support instruments. However, a number of the most commonly encountered are worthy of mention. To this end, Miles and Cunningham (2006) provide an overview of available evaluation methodologies, grouping them according to the functions they play and the specific information they contribute:

- Methods for accessing and generating data – techniques of data production that are used in primary and secondary research, such as surveys, interviews or document review. Here, the types of indicator available is a major issue.
- Methods for structuring and exploring interventions – ranging from 'traditional' approaches using experiments and control groups to more action-based and participatory approaches. Examples could include the construction of counter-factual sampling approaches or RCTs.
- Methods for analysis of data – ways of processing and drawing conclusions from statistics and qualitative material (such as descriptive statistics), and more elaborate modelling and simulation approaches (such as econometric modelling).
- Methods for drawing conclusions, including impact assessments and similar studies.

[3] An unfortunate term which implies that innovation policy instruments operate in a mechanical or mechanistic way when, in reality, they seek to influence a complex multi-actor, sociological process.

The ePub RTD Evaluation Toolbox (2002) gives some more specific examples of evaluation methodologies, but distinguishes between quantitative and qualitative approaches:

Examples of quantitative methodologies:

- Statistical data analysis: Innovation Surveys and Benchmarking
- Modelling methodologies: Macroeconomic modelling and simulation approaches; Micro-econometric modelling; Productivity analysis; and Control group approaches

Examples of qualitative and semi-quantitative methodologies:

- Interviews and case studies
- Cost-benefit analysis
- Expert Panels/Peer Review
- Network Analysis
- Foresight/ Technology

When selecting the appropriate methodologies to use in an evaluation, it can also be useful to distinguish between data *generation* and data *analysis* techniques, since the requirements (in terms of access, resources, skills, etc.) for each approach may influence the final choice. Miles and Cunningham (2006) offer a comparison of data-generation methodologies, listing the major advantages and challenges associated with their use (see Box 1), while Polt and Rojo, in the ePub RTD Evaluation Toolbox, give a number of approaches that focus more closely on data analysis options (see Box 2). In the second box, a distinction is again made between qualitative and semi-qualitative or quantitative approaches. In each case, the selection of methods is entirely dependent on the context and purpose of the evaluation.

A detailed investigation of each methodological approach exceeds the space available in this report. However, Miles and Cunningham (2006) provide an extensive in-depth discussion of the pros and cons of each specific methodological approach and the appropriate circumstances in which each may be applied. Similar guidance for a range of methodologies is also provided by Reid et al. (2012). Once again, the choice of which method to use will be highly influenced by the purpose and timing of the evaluation, the objectives of the policy intervention, the nature of the specific policy questions, the availability of data and information, and other associated issues.

Box 1: Summary comparison of data generation methods

Method	Overall Purpose	Advantages	Challenges
Questionnaires Surveys *Checklists*	When there is a need to quickly and easily gather lots of information from people in a non-threatening way	Can complete anonymously, inexpensive to administer, easy to compare and analyse, administer to many people, can gather lots of data, many sample questionnaires already exist	Might not get careful feedback, wording can bias client's responses, are impersonal in surveys, may need sampling, expert does not get full story
Interviews	When there is a need to fully understand someone's impressions or experiences, or learn more about their answers to questionnaires	Accessing a full range and depth of information develops a relationship with the client; can be more flexible with client	Can take a lot of time, can be hard to analyse and compare, can be costly, interviewer can bias client's responses
Documentation Review	When there is a need to an impression of how programme operates without interrupting the programme; comes from a review of applications, finances, memos, minutes	Get comprehensive and historical information, does not interrupt programme or client's routine in programme, information already exists, few biases about information	Often takes a lot of time, info may be incomplete, need to be clear about what you are looking for, not a flexible way to get data, data restricted to what already exists
Observation	Gather accurate information about how a programme actually operates, particularly about processes	View programme operations as they actually occur, can adapt to events as they happen	Can be difficult to interpret, seen behaviour can be complex to categorise, observations can influence behaviour of programme participants, can be expensive
Focus groups	Explore a topic in depth through group discussion, e.g. about reactions to an experience or suggestion, understanding common complaints, etc., useful in evaluation and marketing	Quick and reliable, way to gather common impressions, can be efficient way to get a wide range and depth of information in a short time, can convey key information about programmes	Can be hard to analyse responses, need good facilitator for safety and closure, difficult to schedule six to eight people together
Case studies	To fully understand or depict client's experiences in a programme, and conduct comprehensive examination through cross comparison of cases	Fully depicts client's experience in programme input, process and results, powerful means of portraying programme to outsiders	Usually quite time consuming to collect, organise and describe; represents depth of information, rather than breadth

Box 2: Analytical approaches applied in ex-post evaluation

Quantitative methods

- Statistical data analysis

– *Innovation surveys*: These provide basic data to describe the innovation process, which are summarised using descriptive statistics.

– *Benchmarking studies:* These permit comparisons between sets of entities, based on a relevant set of indicators and accompanied by a reasoned explanation of their values.

- Modelling approaches

– *Macroeconomic modelling and simulation approaches*: These are used to provide an assessment of the broader socio-economic impact of policy interventions.

– *Microeconometric modelling*: This range of approaches permits the study of the effect of policy intervention at various levels (individuals, institutes, firms, etc.). Mechanisms allow for the control of counterfactual events, by specifying models that enable an estimation of the effects on the outcome for participants had the programme not taken place.

– *Productivity analysis approaches*: These allow an assessment of the impact of R&D on productivity growth at different levels of data aggregation. This set of approaches is particularly relevant for the analysis of the broader effects of R&D on the economy.

– *Control group approaches*: These permit evaluators to capture the effect of programmes on different participants (and, in certain cases, non-participants) using statistically sophisticated techniques.

Qualitative and semi-quantitative methods:

- *Interviews and case studies*: These approaches use direct observation of naturally occurring events, generally from the perspective of programme participants and stakeholders, to investigate behaviour, stimulated by the innovation programme, in their indigenous social setting.

- *Cost-benefit analysis*: By appraising all the economic and social effects of a programme, these approaches allow evaluators to ascertain whether a programme (or project) is economically efficient

- *Expert panels/peer review*: This type of approach is frequently used to measure scientific output, where they rely on the perception scientists have of the scientific contributions made by other peers. As such, peer review is the method most widely used to evaluate the output of scientific research. However, the approach has also been used for the evaluation of R&D collaboration programmes, such as EUREKA.

- *Network analysis*: Approaches of this type enable the analysis of the structure of co-operation relationships and the consequences of participants' decisions on actions. They provide explanations for the behaviour observed by analysing their social connections in networks.

- *Foresight/technology assessment*: Although more typically used for *ex-ante* evaluation, these approaches may be employed to identify potential mismatches in the strategic efficiency of projects and programmes.

Source: Polt and Rojo, in ePub RTD Evaluation Toolbox (2002)

As already noted, it is rare that a single evaluation approach will provide a full and comprehensive picture of the scheme being evaluated, so some form of methodological triangulation, i.e. a 'mixed-method' approach, will be required. An early and 'classical', example of an evaluation employing such an approach was that of the EUREKA programme (see Box 3).

Box 3: Example of a mixed-method evaluation: EUREKA

The first evaluation of the **EUREKA** initiative (Ormala et al., 1993) provides a helpful example of combining both qualitative and quantitative methodological approaches. Created as an intergovernmental initiative in 1985, EUREKA aims to enhance European competitiveness through its support to businesses, research centres and universities which carry out pan-European projects to develop innovative products, processes and services. EUREKA offers project partners rapid access to a wealth of knowledge, skills and expertise across Europe. Through a EUREKA project, partners develop new technologies for which they agree the Intellectual Property Rights and build partnerships to penetrate new markets (see: http://www.eurekanetwork.org/).

The evaluation analysis was based on three levels of information utilising different methodological approaches, both qualitative and quantitative.

- Level 1 concerned compilation of data on the basic characteristics of projects and participants, carried out using data held by national administrations and other documentary sources.
- Level-2 analysis was based on the identification of direct and indirect economic effects of projects. This involved sending separate questionnaires to industrial and non-industrial participants within in the sample. The response rates seem very good, averaging around 57 %.
- Level-3 information was collected via an in-depth analysis of a representative sample of projects: A representative sample of 73 projects (17.5 % of the total) was selected to be interviewed in depth of which 70 were eventually completed. These issues could only be addressed through the combination of both qualitative and quantitative approaches, as outlined above.

Source: Miles and Cunningham (2006)

Whilst the selection of appropriate methodologies can be drawn from the examples provided in the above boxes, Reid et al. (2012) offer a specific selection of methodological approaches that are frequently encountered in the evaluation of innovation funding schemes for companies. Several of these form the basis of the examples provided later in this report (see Box 4).

Box 4: Suggested methods for business innovation support programmes

Relevant evaluation methods include one or more of the following:

- Use of structural business statistics or data from tax authorities (company accounts, etc.) to compare beneficiary performance over time with those of a comparison group of non-assisted enterprises. Such data can also be used to assess multiplier effects of the public funding (gross-value added, etc.).
- Bespoke surveys of beneficiaries (either a sample or the entire population) extended to a comparison group (or double comparison group, i.e. non-applicant and unsuccessful applicants) to allow for a counterfactual analysis.
- Counterfactual econometric analysis of micro-data (e.g. the Community Innovation Survey data from Eurostat) or national/regional panel data (e.g. the Mannheim Innovation Panel in Germany). An econometric analysis is generally only appropriate when an evaluation covers a large enough number of beneficiary firms for which a reliable and complete economic time series can be obtained.
- An in-depth analysis of a subset of beneficiary firms applying interview/case study methods to understand the synergies between innovation measures. This can be done by tracking over time companies that have received a 'package' of support (e.g. innovation voucher, grant for R&D, prototyping and follow-on investment, training and export grants). By covering the full 'project life cycle' from the firm's viewpoint, the evaluation avoids the risk of project fallacy (i.e. the assumption that a grant, which may only cover part of the product development phase, leads to a direct and verifiable outcome for the firm).

Source: Reid et al. (2012)

It is clear from the above that the evaluation of business R&D support has both a long history and that a range of methodologies have been in place for a considerable period of time. Some of these approaches have undergone considerable evolution and development (for example, econometric modelling has benefited from improved data monitoring and the growth of 'big data' sources). The following section gives a number of insights into recent developments in some of these approaches, as presented during the recent London site visit.

3 POLICY CHALLENGES AND MIXED METHODOLOGIES

3.1 Introduction

The London site visit provided an opportunity to examine how at least some of the challenges outlined in the preceding sections might be overcome. It gave an insight into systems already in place and which might be considered good practice, together with examples of recent methodological developments. These encompassed the overall organisation of evaluation and how it could be supported, and more detailed descriptions of programme and evaluation design, including some experimental evaluation approaches.

3.2 The bigger picture – support and evaluation of innovation

One of the main challenges facing policy makers in science, technology and innovation is the fact that the process of innovation has become virtually all-encompassing, engaging a variety of actors who provide support, infrastructure and knowledge at a range of levels. The policy approach to supporting innovation must now adopt a holistic viewpoint.

This throws up a number of specific challenges and issues, such as how to best coordinate the delivery of a range of support schemes (from university block funding and industry R&D support loans, on the one hand, to cross-sectoral collaborative initiatives on a range of scales). In the UK, the latter range from Catapult Centres, Knowledge Transfer Networks, Knowledge Transfer Partnerships, the (demand-driven) Small Business Research Initiative, and Innovation Vouchers, to name but a few examples.

An additional challenge for policy makers is how best to arrange the spectrum of funding and support institutions across the research and innovation landscape and the pressing need to better understand how their different forms of support combine to realise impact.

As regards the coordination of the UK's support system for R&D, a recent reorganisation has created UK Research and Innovation (UKRI) (see Box 5).

Box 5: UKRI – a coordinating structure for RDI

UK Research and Innovation (UKRI) is a new body that incorporates the relevant agencies in the UK (the six Research Councils and Innovate UK) as well as Research England (which covers regional policy and funding issues in England). The three pillars of UKRI are: knowledge, economy and society. The new organisation will be launched by the end of 2018. One of its overall objectives is to lead on the development and delivery of a coherent national research and innovation strategy.[4] Sitting within this new structure, Innovate UK will have the strategic freedom and autonomy to develop a UK innovation strategy.

UKRI also offers a coordinated approach to data collection regarding the outcomes of its various funding instruments through the Gateway to Research (GtR). This website[5] has been developed by Research Councils UK (RCUK) and Innovate UK to enable users to search and analyse information about publicly funded research, including details of current research projects and outcomes of past projects. Two Application Programming Interfaces (APIs) have been provided as a way of accessing GtR's information directly from other information systems, allowing third parties to link GtR data to other data sets and analyse the information for their own purposes.

[4] https://www.ukri.org/

[5] http://gtr.rcuk.ac.uk/resources/about.html

A clear message to emerge from the discussions during the Oslo and Stockholm site visits and from the preceding MLE (in particular, the Copenhagen site visit) was the importance of careful programme design in the subsequent evaluation process. Policy design tools, such as logic frameworks and logic charts, are extremely useful in deriving a comprehensive view of the programme rationale and how its activities, their outcomes, effects and impacts all link to the achievement of objectives. In turn, the achievement of objectives may be compared to the initial programme rationale, feeding back lessons for subsequent programme design and testing the underlying assumptions for the implementation of the programme itself, thereby closing the policy learning cycle.

In addition, logic frameworks may be used to review the programme rationale in a broader sense, allowing comparative assessment of the role of the programme under consideration within the broader policy mix and thus providing a more holistic view of the innovation policy environment. This view can help to identify the wider set of innovation stakeholders who may be impacted by a policy instrument (rather than the narrower set of direct targets it addresses). This, in turn, contributes to the formulation of an evaluation design which can capture a more comprehensive set of outcomes and impacts, including those that are unintended or unforeseen.

Some of these issues are dealt with in following sections, but one major contribution to an effective evaluation system seems to be the UK's use of standard guidance and approaches to the overall process of evaluation (see Box 6).

Box 6: Innovate UK's approach to programme planning and evaluation

The approach to evaluation starts with the policy cycle approach, part of HM Treasury's ROAME-F framework (from rationale and objectives to appraisal, monitoring and evaluation, and linked via feedback to the design/modification of programmes). However, in practice, things may happen in a slightly different order or in a more iterative way.

- Rationale: Justifies the programme, setting out the policy goal(s) to which it is intended to contribute, and explaining why the programme is necessary (in terms of market or system failures).
- Objectives: Specific (ideally measurable) objectives that should be achieved – an 'operationalised' statement of contributions to the overall policy goal. The statement of objectives should include specification of performance indicators that provide feedback on the success of the programme.
- Appraisal: Sets out the choice of activities that are selected to achieve the programme objectives, explaining why these particular activities were chosen and justifying their selection.
- Monitoring: Explains the procedures to be used routinely during the course of the programme to check that its progress is proceeding as planned. Monitoring requires flows of information related to management (how far activities are taking place as planned) and results (what progress is being made in achieving the objectives, e.g. as specified by the performance indicators).
- Evaluation: Involves examining the efficiency with which the programme has been conducted and its effectiveness in terms of achieving the intended objectives. An attempt is also made to account for the (positive and negative) wider effects of the policy or side effects that were not envisaged initially. It may also examine issues related to programme rationale and objectives, and with the unexpected consequences of, and lessons learned during, the programme.
- Feedback: A statement of how the evaluation results are intended to inform the strategy and design of subsequent programmes: for instance, procedures for publishing, disseminating, and gathering responses to the evaluation.

In evaluations, Innovate UK tends to use the logic model approach; from input and activities to output, outcomes and impact. For each of these elements, indicators are developed or selected and data collected. Two of the methodological priorities are data linking and creating control groups (e.g. companies that apply for support but are not successful). Innovate UK also examines the evaluation stakeholders and audience. In many cases, the evaluation study is commissioned from an external, independent organisation (using a tender procedure).

3.3 Building evaluation capability

As noted in the first section of this report, the range of targets, modalities and purposes served by the various innovation support schemes offered by agencies generate particular challenges for their evaluation, which are further complicated by the potential interplay (ideally, synergistic) between them and their outcomes over varying timescales. For example, how is it possible to disentangle the direct impact of schemes on firms' R&D activities from the persistent behavioural impacts on, for instance, their propensity to collaborate with other firms or with researchers from the higher education sector or from public research organisations?

Box 7: Innovate UK's evaluation challenges

Innovate UK's Economics and Performance team was formed in 2013, with a key initial focus of implementing a robust evaluation framework across all of its activities. This was developed into a new evaluation framework, setting out IUK's guiding principles for designing and implementing evaluation. The strategy highlights a number of challenges and present solutions that can be used to address them:

Challenge - paucity of data

- Innovation grant programmes support a relatively small number of participants.
- Where programmes have different strands or segments, sample size issues can be significant.

Solutions:

- *Track cohorts in real time to enhance data quality and encourage participation.*
- *Keep in touch with sample between survey waves.*
- *Use cohorts from longer time periods to increase sample.*

Challenge – low observability

- Many outcomes and impacts of innovation support are not well documented.
- Primary output – knowledge – can be embedded in innovation project outputs (e.g. products).
- It also moves with people, to different companies, industries and applications, creating benefits elsewhere.
- These spillover impacts are impossible to predict and difficult to track, observe and measure.

Solutions:

- *Evaluation of Smart asked direct beneficiaries whether any customers, suppliers or competitors would have benefited from the project.*
- *Asked what form those benefits took, to build a typology of different spill-overs.*
- *Asked for contact details to interview the indirect beneficiaries, to follow up.*
- *Experienced difficulties in contacting indirect beneficiaries, limited analysis to non-representative qualitative view.*

Challenge - fluidity

- Companies are fluid: they change frequently and unpredictably.
- Introduction of new products or processes.
- Entry to new markets.
- Changes in strategy or leadership.
- Mergers and acquisitions.

Solutions:

- *Use external data for a better understanding of changes in company ownership and exit strategies.*
- *Provides a clearer picture as to how supported companies change over time; ties in with evaluation activities to understand whether grant support impacts survival or company structure.*
- *Looking at advanced analytical techniques to dynamically analyse internet data to look for product launch activity.*

Challenge – skewed and lagged effects

- Statistical models often assume a 'normal' distribution of observations around a mean.
- Impacts of innovation tend to be highly skewed towards a small number of very successful projects with a long tail of low- or no-impact projects.
- Many evaluation techniques seek to estimate the average treatment effect - the mean impact of an intervention on a participant, but a profile of impacts can be difficult to capture in sample-based analysis.
- Impacts also occur over many years, generally way beyond the duration of support.
- In the initial years following support, returns can appear to be low or even negative.

Solutions:

- *Innovate UK evaluations span long time periods, from the start of projects to at least three years after they end.*
- *Recent evaluation of support for agri-tech conducted fieldwork six years after the programme started and intending to go back again after eight years. Initial survey*

found only one quarter of the projects had been completed, so the second wave was put back a further two years.

Challenge – attribution

- Innovation support acts as part of a complex science and innovation system
- Multiple actors and programmes at national and sub-national levels
- Companies may receive support from several programmes across multiple organisations
- Attribution of any observed impact to any single intervention can be very difficult, with each programme being necessary but not sufficient to achieve outcomes.

Solutions:

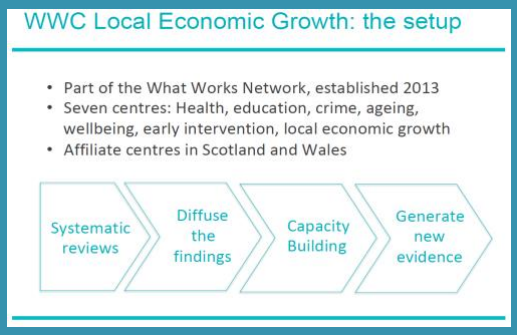
- *Surveys can ask about other forms of support received, but self-reported information is likely to be flawed and incomplete.*
- *Greater linking of administrative data would allow for a more detailed analysis.*
- *Innovate UK has linked data to other programmes delivered by BEIS or the British Business Bank to identify where the support overlaps.*
- *Nesta have been compiling data from local support schemes to provide the sub-national picture.*

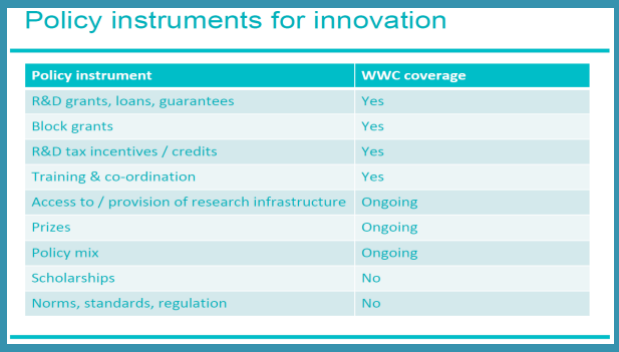
As noted above, Innovate UK now sits within UKRI, alongside the UK Research Councils. Thus, it is able to take a more holistic view of the UK's innovation landscape, an aspect that is reflected in its current strategy and programme of work to improve evidence. Innovate UK's new evaluation framework, which will be published in 2018, focuses on a number of the challenges being faced and the solutions being implemented to overcome them (see Box 7, above). The challenges listed are ubiquitous to policy evaluations across the world and, whilst the solutions are those adopted by Innovate UK, they may be relevant for use in a much wider context.

In addition to the challenge of determining the appropriate methodologies to measure the impact of innovation policy instruments, it is essential that policy-makers and programme managers gain a better understanding of the process of evaluation and its requirements to better understand the results of their evaluations. The activities of the What Works Centre for Local Economic Growth provide a good example of how lessons about the process of evaluation may be learned and disseminated and may be relevant to other countries, provided national contexts are taken into consideration.

Box 8: Understanding and dissemination: the What Works Centre for Local Economic Growth

The Centre was set up in October 2013 as part of the What Works network to analyse which policies are most effective in supporting and increasing local economic growth.





This MLE's preceding site visits have identified a further challenge: it is not only important to determine the effects of a particular policy instrument, in order to maximise policy learning, but it is also essential to understand 'how' and 'why' policies achieve their effects and outcomes. The Centre's evidence reviews and policy design toolkits currently cover 10 policy areas affecting local economic growth – from business support to employment training and innovation. These reviews analysed thousands of existing policy evaluations in order to elicit those which were most useful for understanding what really works in the different policy areas. Toolkits focus on more detailed aspects of how to design policies. As well as highlighting what the evidence revealed about impact, the Centre team selected a range of case studies representing particularly useful evaluation techniques, which can be copied by practitioners as they plan their own policy evaluations. In addition, as noted by Edler et al. (2015), reviews indicate there is less robust evidence available than is desirable for clearly identifying how to improve policy effectiveness. Consequently, the Centre is supporting local practitioners to contribute to the evidence base themselves by piloting and testing new approaches (see Box 8).

3.4 New methodological approaches

3.4.1 Understanding business characteristics and the impact of R&D grants

Although the application of econometric methods to look at how the performance of recipients of funding compares to a counterfactual situation is a long-standing part of the evaluator's tool-box (see Section 2, Box 4, in particular), it remains a technique that is still used correctly in only a handful of evaluations. The advantage of a counterfactual approach is that, potentially, it can offer results that are both rigorous and accurate. The disadvantage is that such methods remain strongly dependent on the availability of sufficiently robust and complete (in terms of time frame) data on the use of different sources of funding by a large enough sample or panel of firms. Consequently, evaluators frequently adopt an approach based on beneficiaries' survey data and comparison of beneficiary performance with the average performance of firms in their sector, etc., using standard economic statistics (Reid et al. 2012).

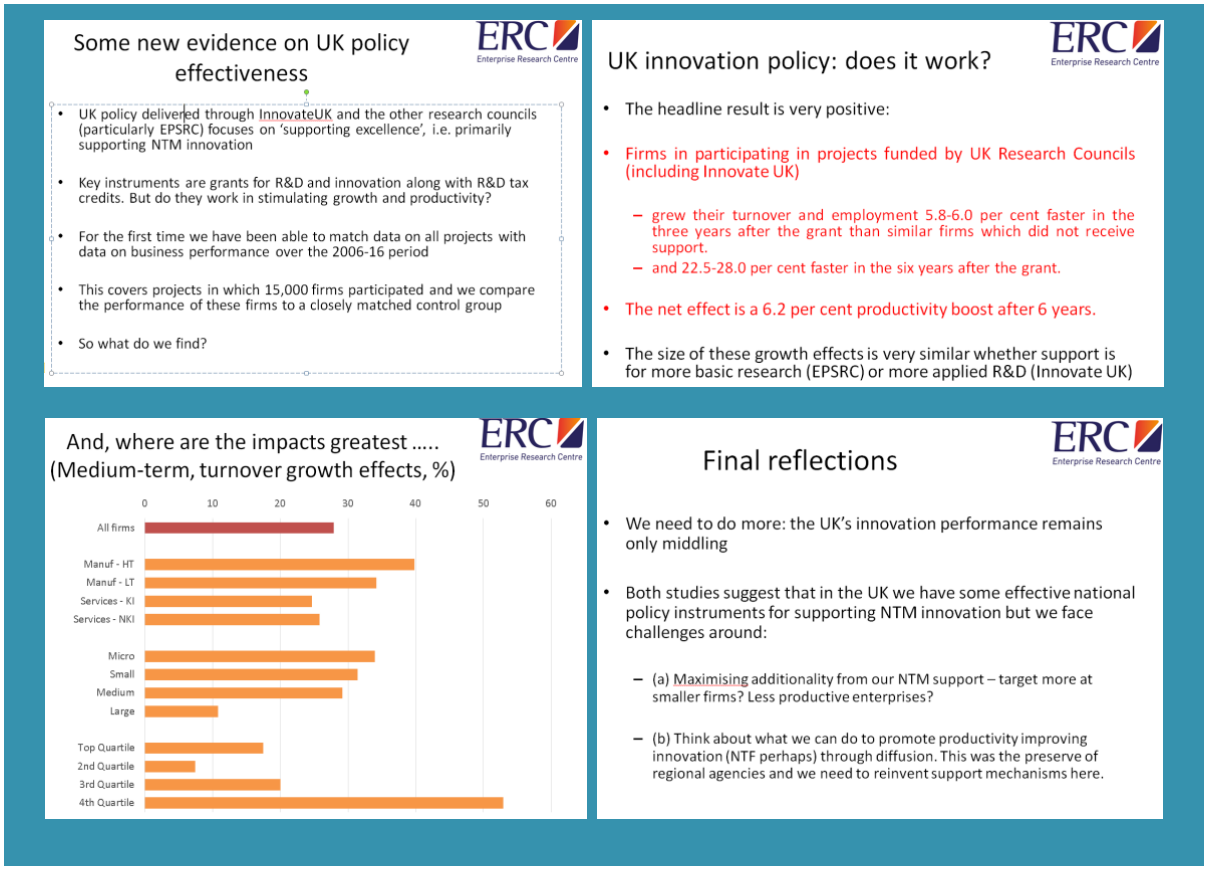
The Economics Research Centre (ERC) at the University of Warwick has recently undertaken studies of so-called 'growth heroes' – small companies which have grown both their employment and productivity and make a disproportionate contribution to national productivity growth. Previously, ERC research has found that a small proportion of UK firms generate revenue growth faster than jobs growth, enhance productive efficiency and demonstrate greater resilience. Subsequent work has explored the characteristics of these growth heroes and examines their wider impact on the local economy. This work involved micro-econometric analysis based on data on the ERC's 'Millennial 2000' firms linked to the ONS Business Structure Database,[6] as well as regional data. Hence, the topic is particularly relevant to discussions in the previous MLE site visit concerning the use of big data sets and their combination (see **Error! Reference source not found.**).

ERC research in this area aims to provide a better understanding of the channels through which growth heroes have an impact at the local scale (e.g. through local demonstration and exemplar effects, as well as potential competition and crowding out) and what types of growth heroes (e.g. knowledge-based fast growing firms) generate positive effects on a region with particular characteristics (e.g. entrepreneurial regions, agglomeration regions, etc.).

In analyses such as these, it is important to consider carefully the criterion to be used when making performance assessments: should this be based on the employment criterion or the productivity criterion? Often, these criteria can apply to different groups of firms, which implies that policy makers should be very clear about their goals and target groups, for example, helping a larger group of firms to become more productive and/or to generate higher employment. It is also necessary to differentiate between high-growth firms and firms with high-growth episodes – growth parameters frequently do not exhibit a continuous upwards trajectory but rather a more punctuated series of stepped increases. This can be a political issue, e.g. whether to support high-flying companies or companies that may have more long-term, sustainable improvements in productivity and/or employment. One view put forward was that High Growth Firms, per se, do not really exist – rather, some firms tend to experience episodes of high growth more frequently.

[6] See: https://discover.ukdataservice.ac.uk/catalogue?sn=6697

Box 9: Advanced econometrics and big data sources



3.4.2 Combining advanced econometric and other approaches

Having identified a number of the challenges facing evaluation methodologies (see Section 2.3), it is important that evaluations continue to incorporate new and experimental approaches, or to use older approaches in new configurations, so as to continually improve their value in policy-learning. One example of a mixed method evaluation incorporating a long established technique, Regression Discontinuity Design,[7] with the combination of multiple data sources, has been utilised in the evaluation of the Biomedical Catalyst.

The Biomedical Catalyst programme is funded by Innovate UK and the Medical Research Council and is intended to support innovative solutions to healthcare challenges. The scheme offers funding to life sciences projects at varying stages of technical and commercial development: Confidence-In-Concept (CiC) awards for portfolios of small projects at the earliest stages of technical development by academic institutions, Feasibility awards (comparable in focus to the CiC awards, but awarded on a firm-by-firm basis by Innovate UK), and more substantial funding for pre-clinical and clinical work through the early and late-stage awards (funding is available up to a Phase II clinical trial or equivalent).

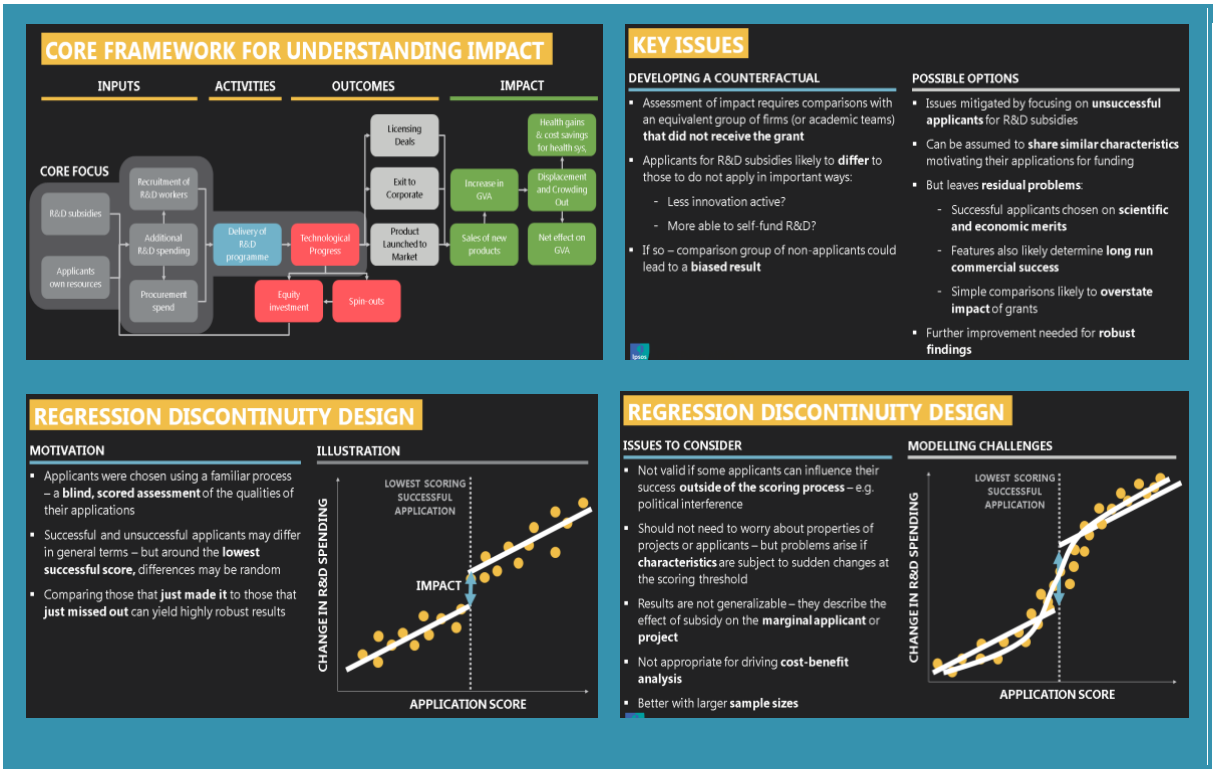
[7] The regression-discontinuity (RD) research design is a quasi-experimental method that can be used to assess the effects of a treatment or intervention. Unique to the RD design is that participants are assigned to groups solely on the basis of a pretreatment cut-off score. The name 'regression-discontinuity' comes from the fact that a treatment effect appears as a 'jump' or discontinuity at the cut-off point in the regression function linking the assignment variable to the outcome. In its simplest form, the design has a pre-test or pre-treatment (the assignment variable) measure, two groups (those scoring above and below the cut-off), and a post-test or posttreatment (the outcome) measure. More complex variations are also possible (Trochim, 1984).

An initial evaluation was conducted by Ipsos MORI in 2015.[8] This focused on: determining a baseline for the assessment of impact; the intermediate and final impacts which could be measured within the time frame of the initial study (together with the proposed approach for doing this); and the intermediate and final impacts which may be realised beyond the time frame of the initial study.

Part of the study methodology involved a data-linking exercise which linked records of successful and unsuccessful applicants to sources of longitudinal secondary data, such as the Business Structure Database and Business Expenditure on Research and Development survey (via the Office for National Statistics Secure Data Service), and computerised patent records available through the European Patent Office (PATSTAT). Further fieldwork was conducted in 2017 to capture more fully and evaluate the impact of the Biomedical Catalyst, including the use of RDD to analyse this data.

Of particular relevance to this MLE was the design of an intervention logic (core framework) and the use of a mixed-methods study to examine process and impact evaluation objectives, with the evidence on results gathered through a survey of successful and unsuccessful applicants. As can be seen from the core framework, the range of anticipated outcomes and impacts clearly necessitated a mixed methods approach in this example (see Box 10).

Box 10: An example of the use of RDD



Specific challenges arising from the evaluation include the poor relevance of traditional measures of assessing the economic impacts of R&D subsidies in this industry (such as productivity, output, or jobs created). These problems are driven by the long timescales for product development and regulatory approval, business models targeting a medium term exit to large corporations, and the growing presence of 'virtual' companies that execute programmes of R&D exclusively through subcontracting arrangements.

8 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/600023/Biomedical_Catalyst_Baseline_Evaluation_Report.pdf

More specifically, trying to apply any form of cost-benefit analysis to the companies involved in the Biomedical Catalyst can create problems since conventional measures of economic benefit (i.e. productivity gains) are not useful because the firms are not expected to generate sales or outputs. Therefore, an alternative way of understanding economic benefit is required. A possible solution might be found in examining changes in firm valuations, which represent expectations of future abnormal profits. However, the issue of the illiquidity of VC investments could mean that such values may not always be observed. Alternative approaches are therefore required.

3.4.3 Mixed methods to overcome time lag effects

As noted in Box 7 above, a further challenge in evaluation is to account for the fact that some effects take longer to emerge than others and sometimes may not become apparent until several years following the intervention itself. In such cases, not only can a two-phased approach be useful, but the application of quantitative in combination with more qualitative approaches may uncover time dependent effects.

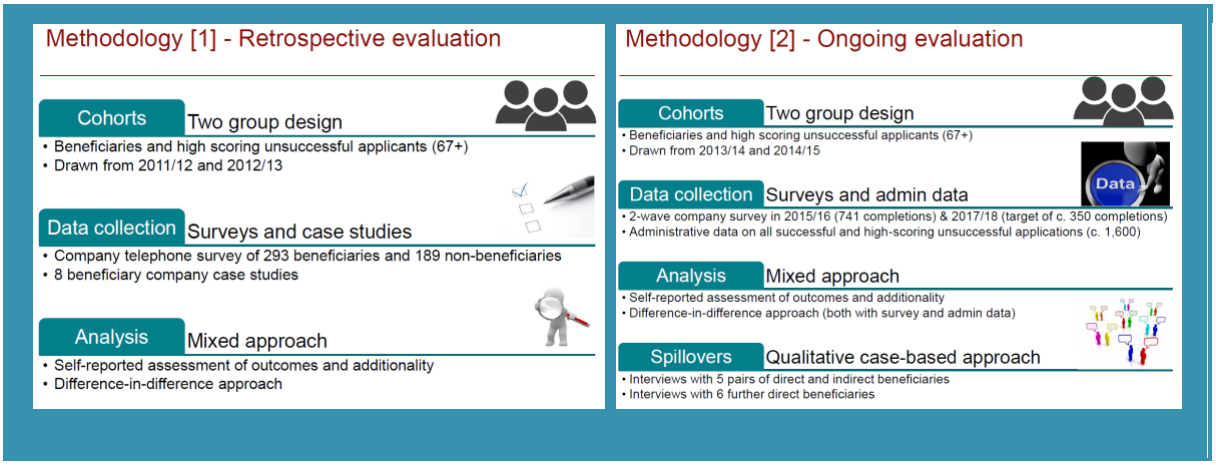
Such an approach was used in a recent (and ongoing) evaluation of the UK's Smart scheme. Operated by Innovate UK, Smart was one of the UK's longest running publicly funded innovation support instruments, providing funding of up to £250,000 to support SMEs across the UK working on innovative R&D projects that could lead to new products or services.

The study, conducted by SQW, was carried out in two phases. The first included large-scale surveys of businesses that had been awarded Smart grants and a comparison group of businesses that applied for but not received one (note the typical use of a counterfactual again). In addition, it also featured a set of in-depth case studies. As such, the evaluation provides a valuable example of the mixed methods approach to evaluate the full range of impacts of Smart on its target population.

A second longitudinal phase of the evaluation is ongoing. This has adopted a similar mixed-methods approach, although with two rounds of large-scale surveys and the addition of a data-linking component. A final impact evaluation will be reported in 2018.

A particularly interesting element of the evaluation is how the businesses involved develop and take forward their innovations – a question that is highly resonant with preceding discussions in the MLE concerning aspects of a firm's 'innovation journey'. Box 11 gives an overview of the evaluation design which illustrates the use of a combination of evaluation approaches.

Box 11: SQW's evaluation of Smart



The study has provided evidence of project additionality (when comparing the beneficiary and non-beneficiary groups), and that projects were being taken forward through further R&D after the end of the project funding. There is also a high degree of variation between project areas and companies (illustrating the 'journey of the project' as part of the firm's innovation journey). The latter finding raises the question of whether projects should be selected as the unit of analysis rather than the entire firm when evaluating such schemes.

Many business outcomes were still anticipated and dependent on further R&D. Those achieved already included not only tangible outcomes such as patents and licences, but intangibles such as enhanced skills, capabilities and collaborations. An interesting finding was that the econometric evidence was less clear-cut in terms of the difference made by Smart participation, thereby underscoring the usefulness of the accompanying, more qualitative, methodologies in determining impact.

3.4.4 Target heterogeneity, complex anticipated outcomes and behavioural additionality

In cases where the targets of a particular intervention exhibit a significant degree of variability in their composition, activities and expected outputs and effects, once again it is clear that a mixed method approach will offer the best opportunity to capture what can be a highly complex set of impacts – one which is likely to be, at best, only partially susceptible to identification by a single methodology.

However, at the same time, it is essential that these mixed approaches are applied within a relatively rigorous framework which encompasses all the programme's targets to ensure that some degree of comparability is maintained.

Typically, when considering heterogeneous targets with complex activities, evaluators are often interested in the issue of behavioural additionality. This received detailed attention during the MLE site visit in Sweden, as is described in the second Thematic Report (den Hertog and Bongers, 2018). Two specific challenges that emerged from this discussion were related to determining how projects influence the behaviour of companies, and how to account for unintended behavioural consequences of innovation support. A further challenge is how to extend the identification of behavioural effects to the system level, since this is ultimately what policy makers are trying to influence.

The use of qualitative, in-depth approaches such as case studies provides for an in-depth investigation of a particular aspect of the innovation programme (generally, its processes and outcomes) and its reasons for success or failure. However, they extend well beyond the collection and analysis of information from the cases. Resources must be allocated to analyse data and report it in ways that are useful to programme managers, staff, sponsors and any external stakeholders. Further, the results need to be explored to understand what caused what, and why, and how much was attributable to the programme. Case study work is also often a matter of participatory evaluation.

In some contrast to the econometric approaches that have been examined in earlier phases of this MLE, the use of case studies places much more emphasis on the 'learning' aspect of evaluating innovation programmes (i.e. lessons drawn, training and knowledge exchange issues). Case studies are typically detailed examinations of how a programme has affected one particular firm or organisation, and often several case studies can be used to explore several different types of programme participant or target. Also, if a programme has many specific projects or lines of activity, the case study may focus on one of these.

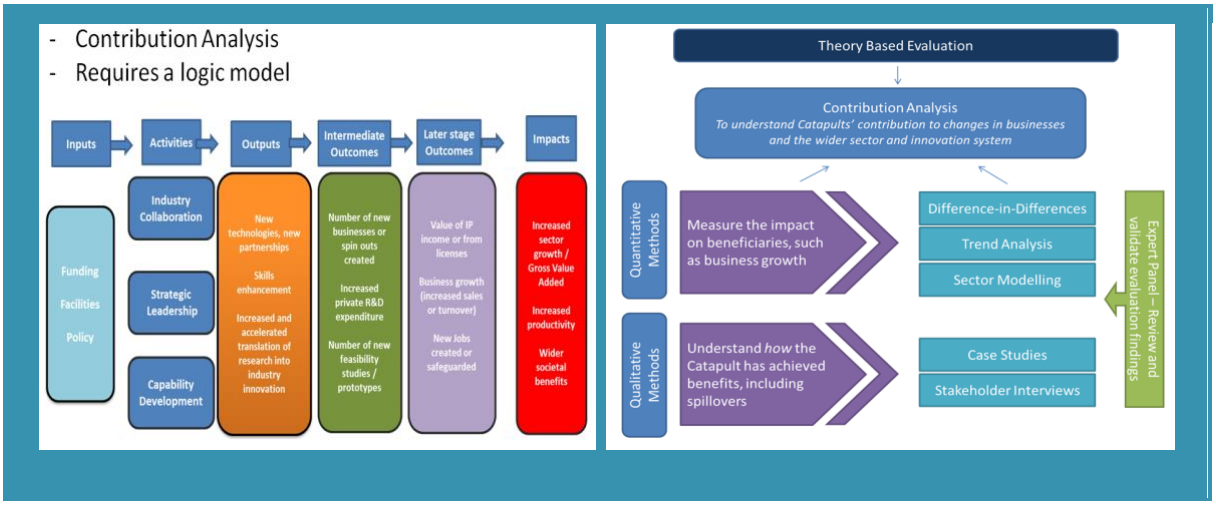
Case studies may also be used to track companies that have received a 'package' of support over time (e.g. innovation voucher, grant for R&D, prototyping and follow-on investment, training and export grants). Hence, by covering the full 'project life cycle' from the target's viewpoint, the evaluation avoids the risk of project fallacy (assuming that a grant, which may only cover part of the product development phase, leads to a direct and verifiable outcome). Likewise, a case study approach can also pinpoint behavioural changes within beneficiary firms and consolidate conclusions on sustainability (for example, the ability of firms to maintain the innovation activity).

The issues discussed above are well illustrated in Innovate UK's evaluation of the performance of the Catapult Centres. These are a series of centres, with physical locations accommodating advanced facilities, which bring together researchers from business and academia to collaborate on late-stage, transformative R&D across a range of cross- and interdisciplinary areas, namely: Cell and Gene Therapy; Compound Semiconductor Applications; Digital; Energy Systems; Future Cities; High Value Manufacturing; Medicines Discovery; Offshore Renewable Energy; Satellite Applications and Transport Systems.

The evaluation framework designed by Innovate UK addresses challenges such as time scale and identifying counterfactuals, and includes a logic model, from input to impacts. One of the main goals of the evaluation framework is to urge and help evaluation studies to capture (behavioural) additionality. Individual frameworks were developed for each of the catapults, e.g. for Catapults aiming at short-term or long-term effects (e.g. Digital vs. Biomedical). Development of the framework for evaluation had to confront a number of challenges – for example, the lack of an agreed overall approach, the complexity of the activities undertaken, the timescale of impacts, the problem of identifying a counterfactual, data collection, establishing a baseline, capturing additionality, etc.

The framework provides an example of a comprehensive mixed methods approach for evaluation, comprising: a theory-based approach; contribution analysis; mixed methods approach to evidence collection and analysis; quantitative evaluation methods (difference-in-differences (DiD), trend analysis (TA), before-after analysis); and qualitative evaluation methods (case studies, in-depth interviews with key stakeholders, expert consultations) (see Box 12). Although the framework is, of necessity, quite generic, through the use of contribution analysis it implicitly addresses the complexity of the anticipated (and unanticipated) outcomes and effects and the possibility of induced spill-overs (such as those arising from behavioural change). Seven Catapults have completed the first phase evaluation, which has focused on the intermediate outputs and outcomes.

Box 12: Catapults evaluation framework approach



At a more focused level, a retrospective evaluation of the High Value Manufacturing Catapult (HVMC), the largest and most mature of the Catapult Centres network across the UK has been conducted by SQW. The HVMC evaluation illustrates well the complexity of anticipated outputs since it covers 27 technology areas. Moreover, not only does it offer a broad, often bespoke service to industry across a range of target types, but it acts strategically, seeking to provide leadership and influence on policy, and is active in skills and competence development to those within high value manufacturing.

The retrospective evaluation was intended to assess the impact of the activities of the HVMC from its inception in 2011 through to mid-2016. In keeping with the Catapult evaluation framework described above, the evaluation adopted a mixed-methods approach, reflecting the diversity, complexity and scale of the HVMC (see Box 13).

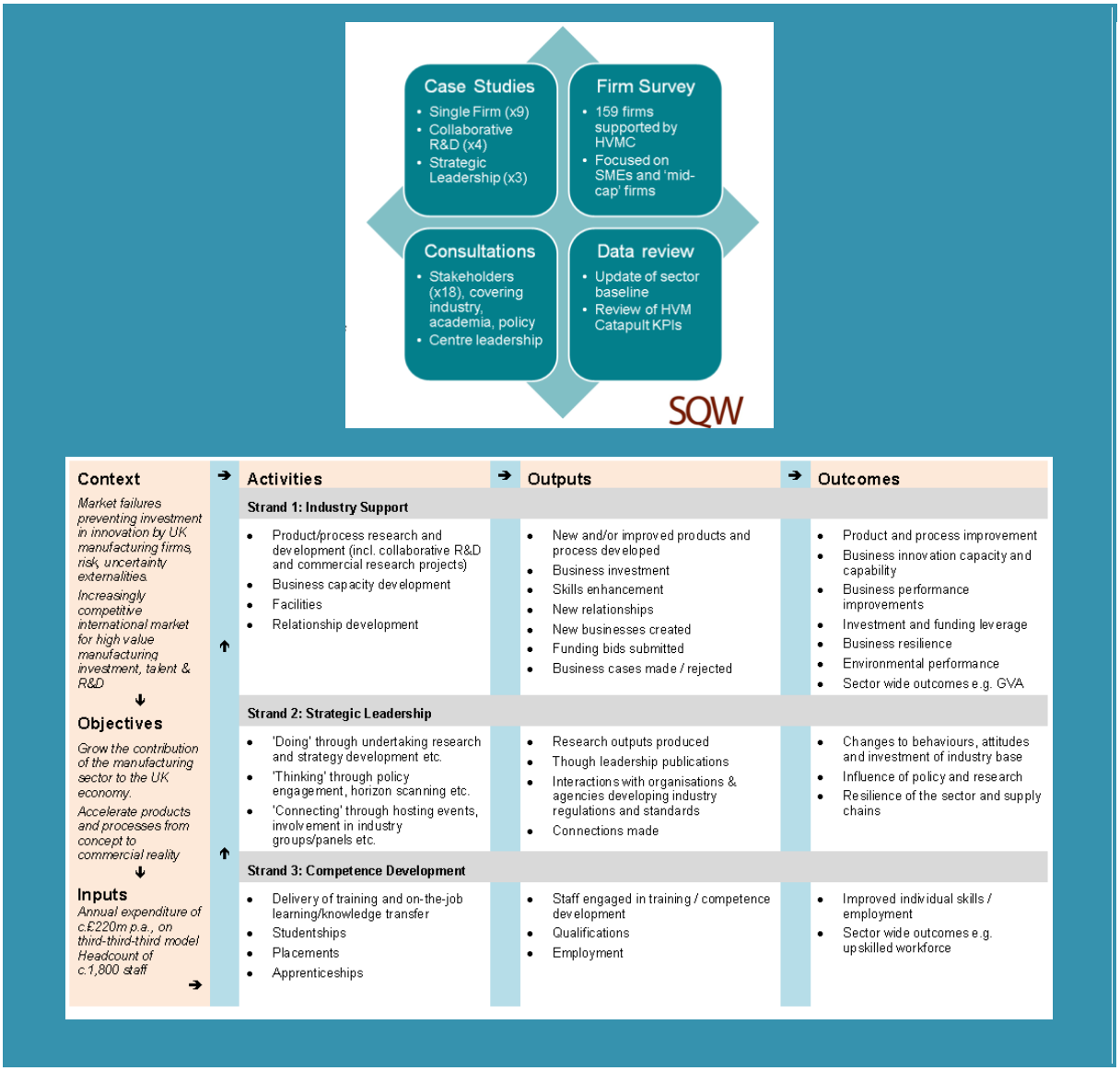
Contribution analysis was selected as an overarching approach in order to reflect the breadth of service offer and non-standard customer journeys, the non-linear and varied routes and time paths to impact, outcomes that are difficult to trace or quantify, and the importance of external factors impacting performance.

The evaluation covered the three major 'Strands' of HVMC activity and the links between them, namely:

- Industry Support, i.e. developing the performance of firms.
- Strategic Leadership, i.e. developing the sector context to support innovation and growth.
- Competence Development, i.e. developing people and knowledge.

As can also be seen in Box 13, the overarching logic framework described above was used to form the basis of one more tailored to the activities of the HVMC.

Box 13: Focus on the High Value Manufacturing Catapult



Most of the evaluation results were quite positive: the results on industry support and impact on innovation were mainly based on the survey, while those concerning strategic leadership and competences were mainly based on qualitative approaches, although they also benefit from the survey. In addition, the findings serve as progress indicators for external audiences (e.g. ministers).

It became evident, however, that even with qualitative approaches such as those used in this example, there are particular limitations with defining and capturing diffusion aspects and spill-over effects. It can be very difficult to track the technologies being adopted. One conclusion is that for this type of policy measure it is difficult to fully capture the complexity and to map the full range of actors engaged. While existing firm taxonomies can provide some simplification of the longer-term beneficiaries in terms of, for example, differential rates of adoption, in reality, the picture is much more complex.

A possible way forward may be suggested from approaches used in a study of the activities of another of the Catapult Centres – the Digital Catapult. This investigated the approaches that might be used to understand the impact of these activities and any behavioural changes induced through the Catapult's collaboration with business.

The Digital Catapult team are seeking to understand the potential impacts of interventions in a complex and dynamic environment. There is a growing recognition that existing models of economic change fail to effectively reflect the potential for large disruptive changes to economic systems – this area of digital technologies is perhaps one of the leading disruptive technology sectors. Traditional statistical models rely on past relationships and trends continuing into the future.

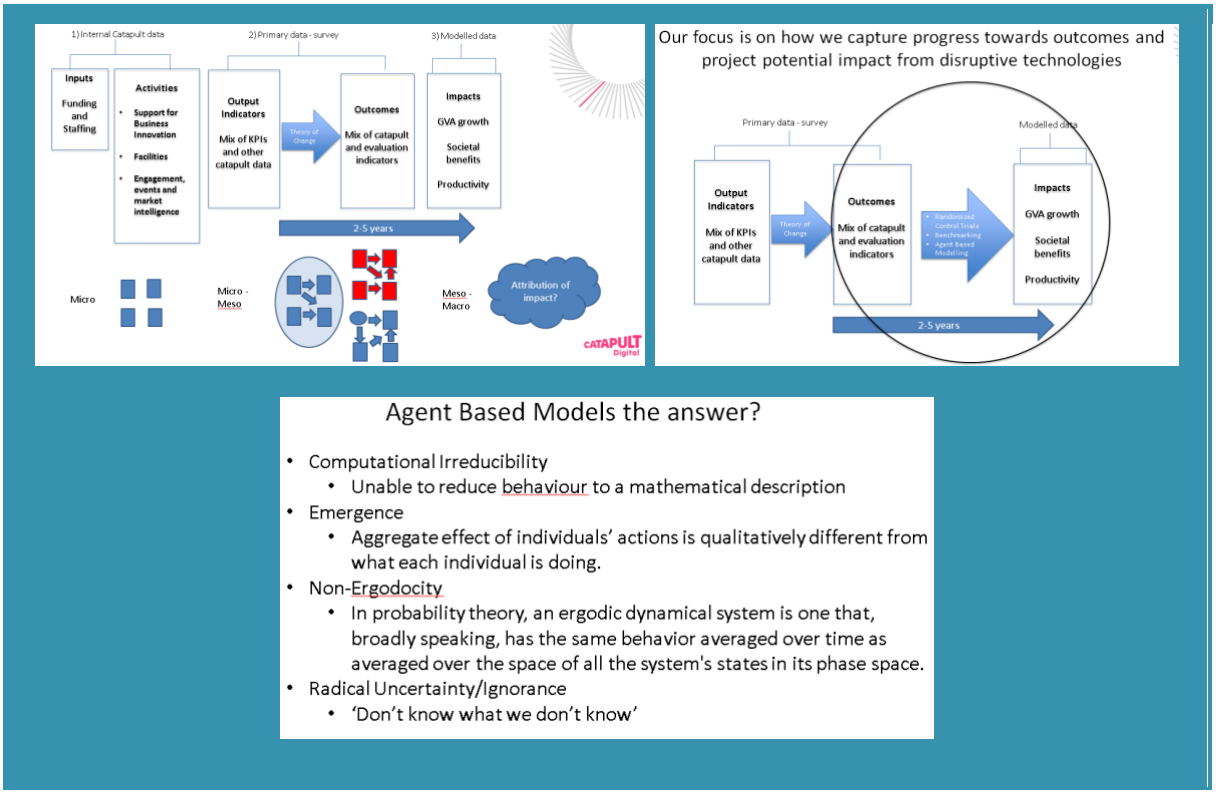
Agent Based Modelling (ABM) is increasingly seen as a solution to the challenges of modelling emergent and disruptive change. ABMs enable the exploration of new markets and opportunities by basing the system on the behaviour of actors within the system. While widely accepted among economists who understand evolutionary dynamics, ABMs have yet to experience wide-spread adoption as the economic community seem to be more locked into existing statistical models.

The Digital Catapult commissioned research on the potential of an ABM to explore the potential impacts of new digital technologies. By building on the internal data collected as part of the Catapult's CRM system, and by collaborating with experts from modelling, data and policy, the Catapult team has explored the development of an advanced ABM to deliver robust exploratory assessments of Catapult support and intervention projects.

ABM is one way of modelling complex systems. Steps include defining the boundaries of the system and defining rules to classify the objects/actors within it. One assumption is that agents are identical (or at least, that there is a small set of agents – with specific rules – interacting with each other). The team used an evolutionary modelling approach: one feature of this is that the future need not be a replication or extension of the past. This aspect matches the uncertainties involved with (radical) innovation (see Bookstaber, 2017).

Much of the required data was internally sourced: the Digital Catapult routinely collects data aligned with its KPIs as well as for management and strategy development in general. Internally, data is collected about inputs and activities in addition to basic company data (using existing datasets to avoid placing a data collection burden on their associated companies). Company data is collected from participants of the Catapult, in a Customer Relationship Management system. Survey data was also used. The team are exploring the use of the Office of National Statistics' Virtual Micro Lab. This would enable them to use data relating to more companies and entire sectors, including traditionally, formally defined sectors and emerging sectors (with the use of text mining to identify relevant companies). For data regarding output and outcomes, the team used surveys and, to assess impact, they used modelling, looking at Value Added, productivity and societal benefits (see Box 14).

Box 14: Agent-based modelling of the Digital Catapult



The model has a forward-looking perspective and includes the use of large (un)structured datasets, data linking, text mining of business registries (and avoids the need to ask companies while not being hindered by industry classifications). This new approach complements existing evaluation approaches, while representing a new paradigm of doing evaluations. In addition, the conceptualisation of the model, carried out in consultation with stakeholders, is a very important first step (i.e. regarding questions on which technologies, sectors, public domains, actors, relations, etc. should be included). One view is that the perspective could be seen to challenge the entire way that evaluations can be carried out and, depending on the availability of the large-scale data and the associated analytical techniques, could be used to derive empirical data. Thus, the study also provides a clear link with the topic of Big Data as discussed in the first Thematic Report (Poel, 2017).

3.4.5 Designing instrument delivery for improved evaluation

The final section of this report picks up and brings together two of the issues already raised: the first concerns the need for evaluators to construct a robust counterfactual to enable a valid assessment of all the facets of additionality (input, output and behavioural), whilst the second answers the need for greater experimentation in the use of novel evaluation techniques. The need for more experimentation has been a long-stated ambition for both programme managers and evaluators. It is seen as a means of exploring further solutions to the range of challenges still encountered in evaluation while overcoming the limitations of the current set of methodologies, particularly in response to the more complex ones (see Box 7). Typically, although programme managers and policy makers have encouraged evaluators to utilise more innovative evaluation approaches, they are generally constrained to the traditional or more tried and tested approaches due to resource constrictions and the need to guarantee that the methods they employ will generate useful and meaningful results. Therefore, room for experimentation is considered a luxury in most cases. However, it is now apparent that policy makers are beginning to make more room for experimentation in the expectation that improved methodologies will deliver a higher level of policy learning.

This section concerns the experimental use of Randomised Control Trials (RCTs) to facilitate programme evaluation. RCTs are designed to directly address the problem of identifying robust comparative samples and are therefore heralded as the 'gold standard' for demonstrating causality. But whilst widely used in other fields, there has been minimal adoption within the design and delivery of innovation policy support instruments.

The Innovation and Growth Lab (IGL) was launched three years ago to address this. Based at Nesta[9], IGL is a global partnership (with partners including Nesta, BEIS (UK), Innovation Norway and others) bringing together innovation agencies, governments, researchers and foundations to foster a more experimental approach to policy. The European Commission has also recognised the potential benefits of such a policy approach, launching a new call to fund innovation pilot projects which will use RCTs to assess impacts.[10]

The implementation and development of RCTs is one of IGL's key ambitions. Typically, an RCT is designed by first screening possible participants for eligibility and other key characteristics, and then using a lottery (a random selection process) to create a 'treatment' and a 'control group'. A major challenge is to address non-observable characteristics such as the quality of the management team, although this is less of an issue when using RCT approaches compared to creating other types of control groups, such as the need to find a group that resembles or closely matches the group of companies that received support (see Box 15).

IGL has developed a guide[11] and an online toolkit[12] to help those undertaking trials on innovation, entrepreneurship and business growth. It maintains a database[13] of RCTs from around the world and has itself supported around 30 trials.

One of the key lessons from IGL is that the collection and sharing of experiences on evaluation, together with experimenting and learning, are key processes. Together, they help with the design of better policies, experimentation, learning and scaling up.

Several examples of innovation programmes in which the use of RCTs had been trialled were provided, including:

- The evaluation of the UK innovation vouchers programme (RCT applied to receipt of voucher or not)
- A matched grant programme for Mexican SMEs to promote high impact entrepreneurship (RCT applied to use of selection process)
- A Chilean start-up and incubator scheme (RCT linked to the choice to share feedback to applications)
- Harvard Medical School scientific collaborations (RCT linked to the provision of information-sharing sessions with possible collaborators)
- An Italian start-up programme (RCT linked to provision of two types of start-up training courses).

---

[9] http://www.nesta.org.uk/about-us

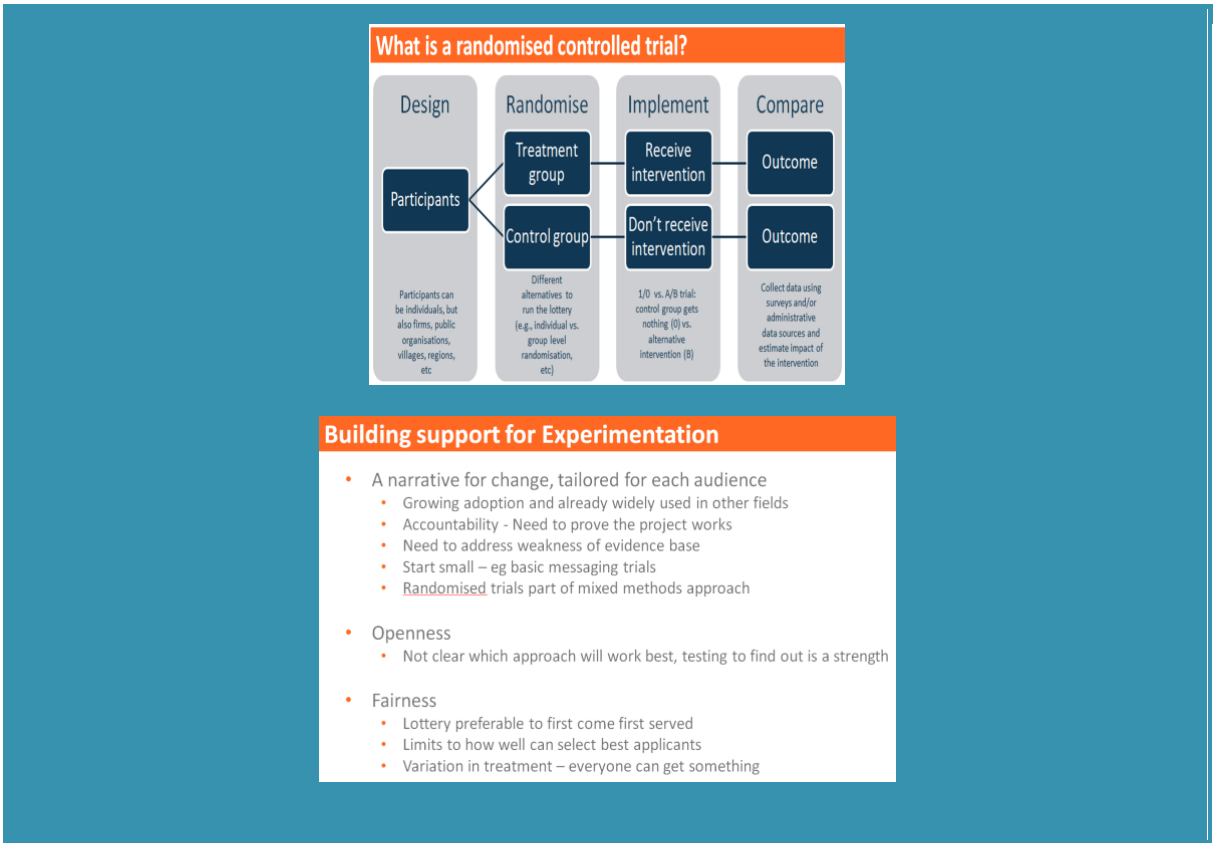
[10] http://www.innovationgrowthlab.org/blog/eu-launches-new-funding-call-experimental-innovation-support-programmes

[11] http://www.nesta.org.uk/publications/running-randomised-controlled-trials-innovation-entrepreneurship-and-growth

[12] http://toolkit.innovationgrowthlab.org/home

[13] http://www.innovationgrowthlab.org/igl-database-map

Box 15: RCTs and experimentation



Key benefits of the approach include the fact that RCTs can provide a concise and clear-cut conclusion of intervention effectiveness. If the randomised groups are large enough, it is possible to be confident that differences in outcomes are due to the intervention and not to other factors. In addition, the evaluator can decide when and how to randomise into groups and thereby gains a clear understanding of the nature of the counterfactual(s). This is not the case in alternative approaches.

For this reason, finding a 'counterfactual' of similar firms is often not a problem: innovation programme users are often atypical SMEs, and their observable factors (for example, age, size, sector and past growth) explain a lot about those who are unlikely to use support.

One question concerning the use of RCTs is how confident can we be that matching firms or individuals on the basis of observable characteristics removes any selection bias? It can be argued that significant effort goes into the collection of detailed information and expert assessments are used to select the best applicants – and it has been noted that observable characteristics have very little explanatory power in business growth models.

A further challenge is to be able to identify the point at which the performances of the two samples diverge. There may be several underlying characteristics that are not measurable but which may influence performance, such as the reasons which started the business on the journey (e.g. undertaking innovation) or which helped it be successful (e.g. the presence of a management team that is aware and able to successfully apply for support).

There is evidence of the increasing use of RCTs in other policy fields (not just medicine) – Innovation Norway has also reported that it is experimenting with the approach. One downside, however, is the need for evaluators to be engaged from the design stage of the policy instrument.

4 LESSONS LEARNED

4.1 Introduction

Participants at the London site visit all agreed that there were a number of valuable lessons to be learned from the event and the discussions that accompanied the presentations. Many of the approaches were familiar to some participants, but there were also significant new developments that could be taken back to their national contexts. In addition, the discussion had highlighted several challenges that still needed to be addressed.

The use of mixed method approaches was widespread and typically the preferred way of carrying out evaluations. Many participants were able to provide examples that had been used at their national or regional level (see Box 16). However, there was agreement that the investment of time in planning policy instruments and their evaluations could pay major dividends and that the use of new methodologies might provide valuable insights about outcomes, effects and impacts. As described by one participant, there is a need to thoroughly reflect on the type of results and lessons you are interested in, instead of accepting that "the [available] data decided which questions should be asked". There is also room to build ex-ante hypotheses, such as 'why and what should we research?'.

Box 16: Examples from the breakout sessions

Typical methodologies employed include:

- Surveys
- Administrative data (national statistical offices)
- Econometrics
- Case studies (including multiple methods and using data from other methods)
- Workshops

Examples:

- Three 2018 French Tax Credit scheme evaluations: 1. Modelling, econometrics and CIS data – looking at input additionality; 2. Combination of the TC scheme and more direct support schemes; 3. Examining human capital, R&D employment, focusing on a scheme connected to the TC scheme.
- Evaluation of state-aid in Spain, including qualitative and quantitative approaches for triangulation of additionality effects.
- Programme evaluations in Sweden using 200+ case studies.
- Support schemes for young tech firms in Brussels, combining in-depth interviews, survey and Qualitative Comparative Analysis
- Evaluation of the Norwegian cluster programme (using innovation database with network analysis, difference-in-difference, interviews, case studies); also examining how cluster support/participation complements other public schemes, with regional workshops to discuss the future.
- Sweden has used a lot of case studies (100-500): some in-depth (using data from various methods), some at a higher level of abstraction; two rounds is useful but can be resource-intensive.

- Austrian Evaluation Platform, although limited to RTD evaluation, does disseminate information on evaluation activities and holds a repository of evaluation reports.
- A 2013 two-phase evaluation of the Austrian Kompetenz Centres may have lessons for the UK Catapult evaluations.
- A series of R&D impact studies on major support instruments conducted in Turkey since 2015 use similar mixed method approaches. Quantitative approaches are assisted by unique entrepreneurs' information system which enables the construction of control groups.
- Germany had two examples: 1. the ZIM-programme evaluation (by ZEW) with a typical thorough mixed methods approach combining panel data, econometrics, survey and case studies; 2. evaluation of the KMU Innovation initiative (by DIW) – included a horizontal perspective, i.e. looking at other similar schemes and their evaluation (benchmarking element).

The need to move towards more systemic approaches and to investigate the potential of big data were also seen as challenges that would have to be met in the near future in order to improve knowledge about the performance of the recipients of innovation support.

4.2 Main takeaway messages

The presentations and their following discussions covered an extensive range of issues.

The participants referred to almost all the presentations during the London site visit as a source of learning and inspiration, both from the perspective of the evaluations and their design as well as the content of the instruments themselves. For example, the presentation on Tax Credits was highly relevant as various member countries are engaged in starting or carrying out evaluations of their own tax credit schemes. Particularly useful aspects were the three perspective views (including its design using the cost of capital approach and the issue of the importance of diffusion, i.e. what firms are being targeted). Similarly, the work on growth heroes made participants aware that high growth can mean very different things. Thus, defining what should be measured (growth in employment, productivity gains) is a key preliminary step. Likewise, the Catapults evaluation was regarded as highly informative, both on the evaluation design and its customisation to different Catapults.

Useful practical elements included the use of borderline rejected proposals/applications as a means of constructing control groups, the drawbacks of some methods, the use of steering groups for evaluations, the various ways in which case studies can be used, when to use the regression discontinuity approach, and the more intensive use of logical framework analysis.

Participants recognised the typical approach in a mixed method evaluation involves a combination of administrative data, surveys, cases and econometric analysis. An attractive, yet challenging development (in addition to further improving the standard of mixed methods and their triangulation) is the use of Big Data (including more or less 'off the peg' datasets) and the agent-based modelling perspective or approach). More countries also seem interested in developing the use of RCTs more extensively.

However, there were several key messages, many of which were neatly encapsulated in the Innovate UK presentation.

1. In planning an evaluation, the choice of methods is typically informed by:

   a. The purpose of the evaluation, which may, in very broad terms, be summative or formative (or a mix of the two), justificatory or learning-oriented.

   b. The rationales underlying the policy instrument and the main expected outcomes (also remembering to be sensitive to unanticipated outcomes).

   c. The timing of the evaluation (what outcomes and objectives are likely to have been already achieved, what are the needs for policy learning).

   d. The resources available, including data availability, personnel resources (including the likely burden imposed on respondents, programme managers and stakeholders), time constraints, etc.

2. Logic charts can be extremely important in programme design and in formulating the key issues and questions to frame the evaluation design.

3. Evaluation should be designed into programmes from the beginning (in combination with the construction of a logic chart). It is never too soon to start planning an evaluation. Moreover, the use of evaluation frameworks (as used in the Catapult Centres evaluation) might provide a more systemic approach to evaluating a group of related schemes or instruments and could offer an insight into methods to evaluate policy mixes, either in part or in their entirety.

4. It has to be accepted that the evaluation of innovation support is difficult, with several, sometimes intractable, challenges. This does not mean that evaluation should not be attempted. The best strategy is to ensure that the most practical (given the constraints) and most robust (i.e. likely to produce a convincing answer) methods are applied to each element of the programme. A mixed methods approach is usually the most suitable, provided those selected reflect the key evaluation issues and questions.

5. A key to successful evaluations is the availability of good quality data, either pre-existing or derived from data-gathering approaches. Ideally, by identifying what data will be required for the evaluation, steps can be taken to ensure sufficient (and appropriate) data collection processes are in place. The use of carefully considered monitoring processes can provide a useful tool here.

6. Where the use of data is concerned, the available sample size is fundamental. Evaluations should be designed so that a sufficient sample size can be obtained in both the treatment and the control group. This may mean using a cohort over a longer period of time. If sufficiently large samples cannot be obtained, alternative, more qualitative approaches should be considered providing they can be used to address similar evaluation questions.

7. Linked to the previous point, evaluations will generally require survey data. However, it is important to recognise that this is typically likely to be imperfect. Therefore, evaluators should attempt to complement and verify the available data by linking it to third-party data sources, where possible. The combination of several data sets can serve the same triangulation purpose as offered by mixed method approaches.

8. Avoid over preoccupation with a single number – evaluation findings will always come with some gaps and uncertainties. It is also imperative to clearly state any caveats associated with the findings. While the headline return on investment is important (particularly to politicians and auditors) it will always be accompanied by a significant margin of error. In most cases, the narrative and lessons around it (many of which may arise from the process of undertaking the actual evaluation) will be just as useful in informing decision-making, if not more so.

9. Where resources permit, try to be innovative and experimental when carrying out evaluations. Although the primary imperative is to ensure that the evaluation methodologies being used are indeed the most appropriate and robust, it can be useful to examine opportunities for using novel techniques. For example, if a number of challenges and uncertainties are associated with the traditional evaluation approaches being employed, it may be an opportunity to push the boundaries of the evaluation a bit further to investigate whether alternative methods can overcome the existing shortcomings.

10. Finally, when experimenting with new approaches, either in the programme or evaluation design, start small or build on any previous successes.

5 REFERENCES

Bookstaber, R. (2017). Agent-based models to help economics do a better job, OECD. Available at:

http://oecdinsights.org/2017/01/23/agent-based-models-to-help-economics-do-a-better-job/

Cunningham, P., den Hertog, P. and Peter, V. (2017). *Ex-post* evaluation of business R&I grant schemes. Expert panel report prepared for the Horizon 2020 Policy Support Facility, Mutual Learning Exercise, Brussels, European Commission.

Cunningham, P., Edler, J., Flanagan, K. and Larédo, P. (2015). The innovation policy mix. In: Handbook of Innovation Policy Impact, Edward Elgar Publishing Limited, 2015, ISBN1784711845, 9781784711849.

Cunningham, P.N., Gök, A. and Larédo, P. (2015). The impact of direct support to R&D and innovation in firms. In: Handbook of Innovation Policy Impact, Edward Elgar Publishing Limited, 2015, ISBN1784711845, 9781784711849.

den Hertog, P. and Bongers, F. (2018). Capturing Behavioural Change. Thematic Report 2 prepared for the PSF-MLE on Evaluation of Business R&D Grant Schemes, European Commission, Brussels.

Edler, J., Cunningham, P., Gök, A. and Shapira, P. (2015). Handbook of Innovation Policy Impact, Edward Elgar Publishing Limited, 2015, ISBN1784711845, 9781784711849.

European Commission (2002), RTD Evaluation Toolbox - Assessing the Socio-Economic Impact of RTD-Policies, IPTS Technical Report Series, EUR 20382 EN, published August 2002, Seville, IPTS and Vienna, Joanneum Research Institute of Technology and Regional Policy. Available at:
https://ec.europa.eu/research/evaluations/pdf/archive/other_reports_studies_and_documents/assessing_the_socio_economic_impact_of_rtd_policies_2002.pdf

Flanagan, K., Uyarra, E. and Laranja, M. (2011). Reconceptualising the 'policy mix' for innovation. Research Policy, vol. 40, no. 5, pp. 702-713. DOI: 10.1016/j.respol.2011.02.005

Guy, K., Boekholt, P., Cunningham, P., Hofer, R., Nauwelaers, C. and Rammer, C. (2009). Designing Policy Mixes: Enhancing Innovation System Performance and R&D Investment Level. European Commission – DG Research, February 2009. Available at: https://cordis.europa.eu/result/rcn/153516_en.html

HM Treasury, The Green Book: Appraisal and Evaluation in Central Government (2003). [Note: refers to the broader context of evaluation practice across government]. Available at:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220541/green_book_complete.pdf

HM Treasury, The Magenta Book, Guidance for evaluation (2011). Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220542/magenta_book_combined.pdf

Miles, I. and Cunningham, P.N., Smart Innovation: A practical guide to evaluating innovation programmes, ISBN: 92-79-01697-0, Commission of the European Communities, January 2006. Available at: https://cordis.europa.eu/publication/rcn/7244_en.html

Ormala et al. (1993). Evaluation of EUREKA Industrial and Economic Effects, Brussels, EUREKA Secretariat.

Poel, M. (2017). Big data: data linking, new data sources and new data analytics methods. Thematic Report 1 prepared for the PSF-MLE on Evaluation of Business R&D Grant Schemes, European Commission, Brussels.

Technopolis Group and Manchester Institute of Innovation Research (2012) [cited in the text as Reid et al. (2012)] Evaluation of Innovation Activities: Guidance on methods and practices - A publication for the Directorate-General for Regional Policy, European Commission, March 2012. Available at: http://ec.europa.eu/regional_policy/sources/docgener/evaluation/pdf/eval2007/innovation_activities/inno_activities_guidance_en.pdf

Trochim, W.M.K., 1984. Research Design for Program Evaluation: The Regression Discontinuity Approach, Sage, Beverly, CA.

Getting in touch with the EU

IN PERSON
All over the European Union there are hundreds of Europe Direct Information Centres.
You can find the address of the centre nearest you at: http://europa.eu/contact

ON THE PHONE OR BY E-MAIL
Europe Direct is a service that answers your questions about the European Union.
You can contact this service
– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
– at the following standard number: +32 22999696 or
– by electronic mail via: http://europa.eu/contact

Finding information about the EU

ONLINE
Information about the European Union in all the official languages of the EU is available on the Europa website at: http://europa.eu

EU PUBLICATIONS
You can download or order free and priced EU publications from EU Bookshop at: http://bookshop.europa.eu. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see http://europa.eu/contact)

EU LAW AND RELATED DOCUMENTS
For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

OPEN DATA FROM THE EU
The EU Open Data Portal (http://data.europa.eu/euodp/en/data) provides access to datasets from the EU. Data can be downloaded and reused for free, both for commercial and non-commercial purposes.

This thematic report has been prepared for the Mutual Learning Exercise (MLE) on the evaluation of business research and development (R&D) grant schemes in European countries. It addresses the topic of applying mixed method approaches to the evaluation of public schemes to support R&D and innovation in firms, specifically business R&D grants and associated innovation schemes. It sets out the broad context for the use of mixed-method approaches in the evaluation of innovation support schemes, with a focus on business R&D grants. The report focuses on a number of major issues which were addressed in the third site visit to London and illustrates these with selected examples. These include; an overview of the UK's approach to the support and evaluation of innovation; some of the major evaluation challenges encountered; and a number of new methodological approaches to evaluation which rely on combined evaluation methodologies. Finally, the report presents a number of the major lessons learned during the site visit concerning programme and evaluation design

Studies and reports