Mutual Learning Exercise on Evaluation of Business R&D Grant Schemes: behavioural change, mixed-method approaches and big data

Horizon 2020 Policy Support Facility

Final report of the Mutual Learning Exercise under the Horizon 2020 Policy Support Facility (PSF): Evaluation of Business R&D Grant Schemes: behavioural change, mixed-method approaches and big data

Mutual Learning Exercise under the Horizon 2020 Policy Support Facility (PSF)

Evaluation of Business R&D Grant Schemes: behavioural change, mixed-method approaches and big data

Prepared by the independent experts:

Marzenna Weresa (Chair)

Martijn Poel (Rapporteur)

Paul Cunningham

Pim den Hertog

Table of Contents

List of Tables

List of Figures

THE PSF MLE PANEL

Marzenna Anna Weresa - Chair

Marzenna Weresa is a full Professor of Economics. Since 2005, she has been working as a director at the World Economy Research Institute; in 2016, she was elected Dean of the Collegium of World Economy at the Warsaw School of Economics for the period 2016-2020. She has been carrying out many advisory projects for enterprises and governmental organisations in the field of internationalisation strategies, R&D and innovation. From 2012-2015, she was a member of the High Level Economic Policy Expert Group (I4G and RISE), providing advice to the European Commission on policies for research and innovation.

The focus of her research and academic teaching is on international economics, and the economics of innovation in particular, issues related to foreign direct investment (FDI), technology transfer, innovation systems as well as the effects of FDI and foreign trade on competitiveness. She has authored and co-authored over 100 books and scientific articles.

Martijn Poel - Rapporteur

Martijn Poel is specialised in ICT-driven innovation, regional clusters and innovation policy. He manages prospective studies about technology trends, big data applications and skills needs. He conducts evaluations of research and innovation programmes. Martijn was deputy director at the Amsterdam office of Technopolis Group and is now working as a senior policy official at the Dutch Ministry of Education, Culture and Science. He wrote his PhD thesis at Delft University of Technology, exploring the impact of the policy mix on service innovation.

Paul Cunningham - Expert

Until recently, Paul Cunningham was Senior Research Fellow and Director of the Manchester Institute of Innovation Research (MIOIR) at Manchester Business School, University of Manchester. With over 30 years in the field of science, technology and innovation policy studies, his research interests now encompass a wide range of science and innovation policy-related fields, but with a particular focus on the broad issues surrounding the evaluation of the impact of policy initiatives. On the more applied side, he has undertaken numerous evaluations, reviews and studies and has acted in an expert role for a wide range of bodies, including UK Government departments and agencies, UK Research Councils, the British Council, the European Commission, the European Parliament, OECD, European Space Agency, and a range of foreign government departments and other agencies. His work has been influential in the

development and formulation of STI policy at a variety of levels, within and outside of the UK.

Pim den Hertog - Expert

Pim den Hertog is one of the founding partners and senior researchers at Dialogic, a research-based consultancy in Utrecht (NL) where he is one of the directors coordinating the science, technology and innovation (policy) domain. Before founding Dialogic, he was a senior researcher at TNO Strategy, Technology and Policy. Pim originally graduated as an economic geographer (Utrecht University, cum laude, 1990) and obtained his PhD in business studies (University of Amsterdam, 2010). He has participated in and led numerous national and international studies on innovation policies, innovation governance, service innovation (policies) as well as monitoring and evaluation studies of individual innovation instruments and organisations. Most recently, he was involved in the evaluation of the Innovation Box (2015), the Topsector approach, and the six TO2 institutes in the Netherlands. He coordinated Dialogic's contribution to the European Service Innovation Centre initiative. In the period 2015-2016, he was responsible for the RIO country reports for the Institute of Prospective Technological Studies. He further contributed to ERAnet projects such as IPPS and EPISIS. He acted as an expert for the first phase of the MLE on business R&D grants and continued to do so for the second phase.

The expert team was supported by the PSF Team comprising the PSF contractor represented by **Nikos Maroulis**, project manager at Technopolis Group, and the Commission services (DG Research and Innovation, Unit A4 – 'Analysis and monitoring of national research policies') with **Eva Rückert** as the contact point from DG Research and Innovation, who coordinated the exercise. **Kimmo Halme**, 4Front, acted as the quality reviewer.

EXECUTIVE SUMMARY

This report has been prepared for a Mutual Learning Exercise (MLE) on the evaluation of business R&D grant schemes. The MLE is part of the European Commission's Policy Support Facility (PSF). It has engaged policymakers and public agencies from 12 countries: Austria, Belgium, Croatia, France, Germany, Lithuania, Norway, Slovenia, Spain, Sweden, Turkey and the United Kingdom.

To a large extent, the methodological challenges of evaluating business R&D grant schemes resemble the challenges of evaluating other types of support schemes targeting businesses. Moreover, in many cases, businesses can benefit from multiple support schemes. For these reasons, the MLE also addressed the evaluation of tax incentives, voucher schemes, collaborative R&I programmes, cluster policies, etc.

The evaluation challenges faced when designing and conducting evaluation studies include:

1. skewed effects;

2. lagged effects;

3. paucity of data;

4. low observability (including spill-overs);

5. fluidity of companies; and

6. attribution.

In order to address these (persistent) challenges, evaluation communities continuously improve their perspectives, methods and data-collection approaches. This process started decades ago and will continue to evolve, based on advances in innovation theory, data-collection tools, data analytics, etc.

This report focuses on three incremental innovations in the evaluation of support schemes for business R&D and innovation:

- the added value of taking a behavioural change perspective and measuring and understanding how the R&D and innovation behaviour of companies changes in response to policy measures;

- recent advances in mixed-method approaches, including econometrics, the use of control groups and qualitative methods, in the evaluation of the impact of business R&D support measures; and

- the opportunities and challenges of big data in policy evaluations, including data linking.

These three innovations are complementary and can help to address evaluation challenges. Behavioural change is mainly a perspective, a conceptual framework

with an emphasis on understanding and measuring how support schemes change companies' R&D and innovation behaviour. A behavioural change perspective requires policymakers, public agencies and evaluators to make explicit *how* a support scheme should, and does, influence what behaviour in which companies, temporarily or persistently. In other words: what changes in behaviour can we attribute to a policy intervention? For example, business R&D grants, large or small, may emphasise the level of business R&D, the level of risk-taking and/or the level of R&D collaboration, by small, medium-sized or large companies with either little or vast experience in R&D.

Mixed-method approaches are effective for implementing a behavioural change perspective, using different types of relevant technics and developing the insights needed to assess and adapt R&D and innovation support schemes. For example, econometric tools using control groups and time-series data can be used to attribute behavioural changes to support schemes. In so doing, econometric tools can address the time lag between policy interventions and changes in behaviour (or identify effects that fade away when a policy intervention is stopped). Qualitative methods, such as case studies, are very complementary to quantitative methods, especially for explaining why, how and when companies change their R&D and innovation behaviour, as a result of one policy intervention (or a mix of interventions).

Big data increases the types and volumes of data that can be used in quantitative analyses. For instance, a dataset created for one support scheme (e.g. data about beneficiaries) can be linked to datasets concerning other support schemes as well as to private datasets (e.g. commercial company databases). As such, and taking into account confidentiality, more characteristics of individual companies can be monitored. This increases the possibilities to assess the effects of support schemes on certain types of companies (size, age, sector, region, technology-intensity, etc.). Another example of big data is text mining the final reports and websites of *all* beneficiaries, looking for the economic and social impact. This approach reduces the chances that any impacts are overlooked (cf. skewed impact).

MLE participants provided examples of recent evaluations with a behavioural change perspective, a mixed-method approach and/or data linking. Moreover, they stressed the following current/planned activities in terms of improving their evaluations of R&D and innovation support schemes:

- data linking, e.g. combining data from agencies and statistical offices;

- measuring (behavioural) effects at the level of innovation systems;

- evaluation of the policy mix, e.g. how certain support schemes are relevant for certain phases of a company, during their 'innovation journey';

- acknowledging the heterogeneity of companies, when designing and evaluating support schemes and policy mixes;

- building an evaluation community, including continuous skills development.

The table below summarises 14 recent evaluations by the MLE participants:

Country	Evaluated scheme	Examples of a behavioural change perspective, mixed-method approach and big data
Austria	Innovationscheck programmes (vouchers)	Using data linking and a mixed-method approach, the evaluation assessed whether the vouchers led to persistent changes in the R&D&I behaviour of SMEs
Brussels Capital-Region	Doctiris: collaborative R&D projects between doctoral researchers and enterprises	The first evaluation of the Doctiris programme used different qualitative methods and explored how the programme leads to which types of effect
Croatia	Evaluation of the Croatian innovation system	Collaborative behaviour was promoted by means of agenda setting, using a stakeholder engagement strategy (Entrepreneurial Discovery Process)
France	Competitiveness clusters policy	Data linking allowed the creation of a control group and the assessment of input additionality
Germany	Innovative Regional Growth Cores	A mixed-method approach was taken, 14 years after the introduction of the programme. This enabled an assessment of persistent changes in company behaviour
Lithuania	Inno-vouchers LT scheme	Web scraping was used to collect additional data about beneficiaries and other companies
Norway	Multiple schemes	Linked data platform for monitoring and evaluating support schemes
Norway	Technical Industrial Institutes in Norway	Three datasets were linked, and a quasi-experimental approach was taken to attribute effects to the intervention
Poland	Operational Programme (OP) on improving enterprise competitiveness	A mixed-method approach was used to fully understand how the programme led to effects on companies
Spain	Business R&D grants and other schemes	Comparing the effect of public support, before and during the recent economic crisis; linking internal and external datasets and using a control-group
Sweden	Entrepreneurial experimentation and collaboration	Detailed study of the complex and long-term behavioural effects which contribute to enhanced firm performance
Turkey	Technoparks	Using the linked data in the Entrepreneur Innovation System and taking a mixed-method approach
UK	Smart Scheme	A two-phase approach addressed the issue of time-lagged outcomes and effects
UK	Digital Catapult UK	The evaluation experiments with Agent Based Modelling, including the use of large (un)structured datasets, data linking and text mining of business registries

1   INTRODUCTION

1.1 *Mutual Learning Exercise*

This report has been prepared for the MLE on the evaluation of business R&D grant schemes in European countries, which was carried out from April 2017 to May 2018.

The MLE is one of three instruments available under the overarching Policy Support Facility (PSF), which was set up by the European Commission within Horizon 2020 (H2020). The aim of the PSF is to give EU Member States (and countries associated to H2020) practical support to design, implement and evaluate reforms that enhance the quality of their R&D and innovation investments, policies and systems.

1.2 *Background and participating countries*

The present MLE is a follow-up to that on '*Ex-post* evaluation of business R&I grant schemes' which ran throughout 2016. The MLE gave policymakers from Spain, Denmark, Sweden, Norway, Germany Hungary, the Czech Republic, Turkey and Romania an opportunity to exchange information and share views and experiences about evaluation methodologies, data management and dissemination techniques.

One of the main conclusions of the first MLE was the greater use of econometric analyses, including the use of control groups. Another finding was that additional methods are essential to overcome the limitations of econometric analyses, i.e. to better understand the behavioural effects of using R&D and innovation grants, e.g. the effects on the "innovation journey of firms" (Cunningham et al., 2017).

On the basis of these findings, the current MLE focuses on three topics:

- the added-value of taking a behavioural change perspective and measuring and understanding how the R&D and innovation behaviour of companies changes in response to policy measures;

- recent advances in mixed-method approaches, including econometrics, the use of control groups and qualitative methods, in the evaluation of the impact of business R&D support measures; and

- the opportunities and challenges of big data in policy evaluations, including data linking.

Innovations in each of these three topics, individually and in combination, have the potential to improve the quality of evaluation studies and to address persistent evaluation challenges, such as attribution and lagged effects (see Section 2).

To be able to explore the three topics as thoroughly as possible, the current MLE continued to address the evaluation of business R&D grant schemes but also

looked at evaluations of other types of support schemes that are targeted at businesses (tax incentives, collaborative R&I programmes, cluster policies, etc.).

The current MLE attracted the interest of 12 countries: Austria, Belgium, Croatia, France, Germany, Lithuania, Norway, Slovenia, Spain, Sweden, Turkey and the United Kingdom.

The hosts of the three workshops/site visits played an important role. The first workshop (on the use of big data) took place on 29-30 August 2017 in Oslo and was hosted by Innovation Norway. The second (focusing on behavioural change) took place on 26-27 October 2017 in Stockholm and was hosted by VINNOVA. The third workshop (focusing on mixed-method approaches) took place on 15-16 January 2018 and was hosted by Innovate UK, Warwick Business School and Nesta.

Section 1.3 introduces the three main topics of the MLE and, as such, presents the scope of the report. Section 1.4 presents the structure of the report.

1.3 Innovation in evaluation

1.3.1 Introduction

The perspectives and methods for the evaluation of R&D and innovation policy have been developed over more than four decades. New perspectives, such as systems of innovation, clusters and knowledge transfer, have been added; methods, such as surveys and econometric modelling, have been improved; and analytical tools, such as mapping the intervention logic of policy schemes, have been fine-tuned. These and other innovations in conducting evaluations continue to unfold (European Commission, 2002; Miles and Cunningham, 2006; Technopolis Group and Manchester Institute of Innovation Research, 2012; What Works Centre for Local Economic Growth, 2015).

This report, in response to the priorities identified by policymakers participating in the first round of this MLE, considers three innovations in the evaluation of R&D and innovation policy, namely: taking the perspective of behavioural change, using mixed-method approaches, and using big data. These three innovations are complementary in the sense that a behavioural change perspective urges evaluators to understand and measure changes in companies' R&D and innovation behaviour (and the attribution of any changes to support schemes), which requires the use of a mixed-method approach, and in which data linking or other aspects of big data could play a role.

The report also examines these innovations in the context of the main challenges evaluators have faced and continue to face when seeking evidence of the impact of an innovation policy instrument, namely: 1) skewed effects, 2) lagged effects, 3) paucity of data, 4) low observability (including spillovers), 5) fluidity of companies, and 6) attribution. A more detailed explanation of each challenge is presented in Section 2.

1.3.2 Behavioural change

The perspective of behavioural change is based on psychology, behavioural economics and organisational studies. In the context of the evaluation of R&D and innovation policy, behavioural change refers to: a) the changes in R&D and innovation behaviour of companies (or other actors); and b) the extent to which these changes are induced by R&D and innovation policy (Buisseret et al., 1995; OECD, 2006; Wanzenböck et al., 2013; Roper et al., 2016). Aspects of behaviour include a company's investments in R&D and innovation, its technological and thematic priorities, the level of risk-taking, timing of projects, investing in human capital, propensity for external collaboration, etc.

For policymakers, taking the perspective of behavioural change underlines the importance of *understanding* the incentives and behaviour of 'target groups' such as small, medium-sized and large firms, in specific sectors. As such, behavioural change is one of the heuristics for clarifying the intervention logic of policy. In short, why and how to intervene in order to reach which effects? Understanding the behavioural changes induced by R&D and innovation policy (and other factors) enables changes to be attributed to public policy, which is one of the main evaluation challenges (Den Hertog, 2018).

1.3.3 Mixed-method approaches

The R&D and innovation policy community continues to innovate in individual evaluation methods as well as in combining/mixing evaluation methods. For instance, there are continuous improvements in the Community Innovation Survey, in surveys among beneficiaries of business R&D grants, in econometric modelling and in the use of control groups, and in offline/online engagement of stakeholders (European Commission, 2002; Miles and Cunningham, 2006; Technopolis Group and Manchester Institute of Innovation Research, 2012; What Works Centre for Local Economic Growth, 2015).

Mixed-method approaches can address the attribution challenge and related evaluation challenges such as time lag (see Section 2). Combining methods acknowledges the strengths and limitations of individual methods and allows for a triangulation approach (Denzin, 1989) which permits a broader range of policy effects and impacts to be revealed. In this way, for example, methods for 'measuring how much' and methods for 'understanding how' may be combined to provide a greater depth of understanding how and why a policy measure operates (Technopolis Group and Manchester Institute of Innovation Research, 2012). For instance, the use of econometric modelling techniques, using time series, can reveal whether companies that receive business R&D grants innovate more and increase their productivity, turnover and/or the number of employees. At the same time, accompanying qualitative case studies can indicate how these effects take place and whether the results of the quantitative analysis managed to capture the fluidity of the companies involved.

The need for mixed-method approaches also increases with changes in the policy mix and the growing complexity of the policy objectives. For example, policy instruments may aim to support R&D as well as innovation; support individual actors as well as collaboration in general; and they may aim for immediate effects

as well as 'transitions' (in energy, transport, etc.). Likewise, policy evaluations may need to deliver learning and adaptive policymaking and, at the same time, be accountable to politicians and taxpayers. Consequently, no single method can deliver the data and insights to meet the evolving needs of policymakers (Cunningham, 2018).

1.3.4 Big data, including data linking

The label 'big data' refers to the growing opportunities to collect, process, analyse and use data. In short, big data is about the greater volume, variety and velocity of data (Gartner, 2011: McKinsey Global Institute, 2011; Mayer-Schönberger and Cukier, 2013; Kitchin, 2014).

Currently, in the context of the evaluation of R&D and innovation policy, big data largely concerns data linking rather than using new types of data or new data-analysis tools (Haustein et al., 2014; Technopolis Group et al., 2015; Bakhshi and Mateos-Garcia, 2016). For instance, evaluations can link datasets about companies that participate in support schemes (e.g. scheme-specific datasets) and datasets about the economic performance of companies in general (e.g. open administrative data and (commercial) company databases).

Data linking can help to address the challenge of attribution. For instance, evaluators of business R&D grant schemes can control for the enrolment of companies in *other* support schemes. By applying new types of data, evaluators can also address the evaluation challenge of skewed impact distribution. Using innovation keywords and text-mining project impact reports, company websites and/or company databases, evaluators can collect data about many, if not all, relevant companies. By contrast, methods such as case studies and surveys cover only a sample of companies and may fail to capture those upon which a support scheme has had an impact (Poel, 2017).

1.4 Structure of the report

Section 2 discusses six challenges of designing and conducting evaluation studies: 1) skewed effects, 2) lagged effects, 3) paucity of data, 4) low observability (including spillovers), 5) fluidity of companies, and 6) attribution.

Section 3 elaborates on three methodological innovations (a behavioural change perspective, mixed-method approaches and big data) and how these can help to address evaluation challenges.

Section 4 discusses evaluation studies with one or several innovative elements, as presented during the MLE workshops. The example evaluations were presented by national representatives and by independent evaluators, such as academics, consultants and experts at statistical offices and research institutes.

Section 5 presents the key messages and the current/planned steps of MLE participants, regarding improvements in evaluation studies. This concerns data linking, exploring effects at the level of innovation systems, evaluations of policy mixes, acknowledging and accommodating company variation, and fostering

evaluation communities, including the continuous improvement of evaluation skills.

2 SIX CHALLENGES OF EVALUATING BUSINESS R&D GRANT SCHEMES

2.1 *Business R&D grant schemes*

The provision of direct support for R&D within companies is possibly one of the oldest and most established public policy instruments. The origin of this support can be traced from the immediate post-Second World War period although it has evolved over time, with a shift in focus away from the direct support of single R&D projects within large individual firms towards a focus on direct support to R&D conducted within SMEs (Cunningham, Gök and Larédo, 2015).

The provision of direct support for R&D (either in the form of non-repayable grants or as soft loans) is founded on the underlying rationale that R&D conducted within firms will, directly or indirectly, stimulate innovation leading to the production of new marketable products, processes or services. This view is derived from the linear model of innovation and explains the long history of this type of measure: direct support for R&D follows the classical economic rationale for public intervention, linked to the capacity of firms to appropriate investments made and the relative importance of spillovers associated with their R&D efforts. Direct support measures are thus, in part, intended to compensate for firms' propensity to under-invest due to information asymmetry. At the simplest level, direct support measures basically seek to reduce the risks businesses encounter when innovating (op. cit.).

The shift from large firm support to targeting SMEs, as noted above, has been mainly based on arguments over the comparative efficiency of financing R&D activities in smaller companies, which provides access to a wider range of clients. However, this is at the comparative potential cost of the size of spillovers that may be obtained from the support of larger firms.

As mentioned in Section 1, the MLE for which this report has been written covers business R&D grant schemes as well as other support schemes targeted at businesses. This reflects the fact that multiple incentives can be used to change companies' R&D and innovation behaviour. It also implies that the challenges of evaluating business R&D grant schemes, in different ways, also apply to evaluating other types of support schemes (e.g. tax incentives, vouchers, collaborative R&D and innovation programmes, support for regional clusters, institutes for industrial research, etc.). Note that companies' R&D and innovation behaviour – and any associated drivers or barriers – can also be influenced by support schemes targeted at possible partners, such as universities and research institutes. For instance, these other actors may be given incentives to collaborate with companies.

Although they offer a relatively straightforward policy purpose and a simple modality of action when compared to many other innovation support schemes, the evaluation of direct measures also poses a number of particular problems that are typically encountered in the wider field of evaluation. The remainder of Section 2 discusses six challenges, as identified and described by Miles and Cunningham, 2006; HM Treasury, 2011; Penfield et al., 2014, and others.

2.2   Skewed effects

Although statistical models often assume a 'normal' distribution of observations around a mean, as initially noted by Barber et al. (1994), the impacts of innovation support 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, particularly those reliant on large data sets and varieties of econometric modelling, but also more simplistic descriptive analyses of survey data, seek to estimate the average treatment effect – the mean impact of an intervention on a participant. However, a profile of impacts can be difficult to capture in sample-based analysis and the value for policy learning is minimised, although a highly skewed, long-tailed distribution may indicate that the participant selection process or eligibility criteria may need revision. The presence of a small number of successful recipients may also prompt the use of more qualitative, targeted evaluation approaches to investigate the reasons for success.

2.3   Lagged effects

A second problem is that the desired effects of a policy instrument tend to emerge at various times through a project's lifetime. For evaluators, this poses the dilemma of when is the optimum time to conduct an evaluation and, if more than one evaluation is required, how frequently. For example, if policymakers are concerned about the administration of a scheme, issues concerning uptake and management will emerge soon after implementation. However, if they are interested in the outcomes of a scheme, it may take months or years until prototypes are generated or new products, processes or services introduced to the market, whilst in the initial years following support, returns can appear to be low or even negative.

Likewise, organisational and behavioural changes will take time to generate and become embedded, and their sustainability, along with that of other desired effects, will require even longer time frames. Impacts, particularly on wider actors in the innovation system, or even society and the wider economy, also occur over many years, generally way beyond the duration of support.

2.4   Paucity of data

R&D expenditure, growth, profitability and employment, along with many other anticipated impacts of direct support measures are readily measurable and can lend themselves to the construction of easily obtained quantitative indicators. However, innovation grant programmes support a relatively small number of participants compared to the wider population of firms, and where programmes have different strands or segments, sample-size issues can be significant. The availability of statistically meaningful data may constrain the methodological approach chosen for the evaluation, with a preference for qualitative over quantitative methods. A further constraint is that the types and scales of outcome and impact arising from participation are difficult to ascertain in the absence of counterfactual examples or benchmarks established prior to the establishment of the funding: this exacerbates the data paucity challenge.

One potential option is to mobilise improved monitoring systems which can enable participant cohorts to be tracked to enhance data quality and encourage participation. In addition, multiple survey waves can be used along with selecting sample cohorts across longer time periods to increase sample size.

2.5 Low observability, including spillovers

A further problem is that many of the outcomes and impacts of innovation support are rarely well documented. This problem is exacerbated the further one moves along the intervention logic which encapsulates the programme or scheme's rationale – not least due to the challenge of time lags mentioned previously. Although the primary (project) output, i.e. knowledge, can be embedded in project outputs (e.g. prototypes or products, or patents), which may be captured through monitoring data, surveys, interviews and other approaches, information on less tangible outcomes, such as skills, innovation capabilities and capacities, and spillover effects, etc. is far less amenable to capture. Indeed, many of these spillover impacts are also impossible to predict (and will not appear as identified elements within a logic framework) and difficult to track, observe and measure. Likewise, where the scheme involves multiple and complex objectives, the range of impacts being considered may be wide, with many of the objectives being diffuse and hard to specify.

Similarly, the knowledge produced from participation in a scheme, however it is encapsulated, is capable of moving, often embedded with people, to different companies, industries and applications, thereby creating benefits elsewhere, typically well beyond the scope of the evaluation process.

2.6 Fluidity of companies

Unfortunately for evaluators and policymakers, companies are far from simple organisations that behave in a predictable fashion over time. They resemble organisms in that they are subject to frequent changes, which are generally unpredictable to the external viewer. Such changes may include the introduction of new products or processes, entry to new markets, changes in strategy or leadership and mergers and acquisitions.

Thus, the company that successfully applies for grant or loan funding may not closely resemble the same company when the public support comes to an end. This heterogeneity (both between companies and over time) can be a contributory factor to the skewed responses noted in the first challenge noted above.

2.7 Attribution

Last, and by no means least of the challenges encountered by evaluators, is that of attributing the effects of a particular scheme within a broader environment which includes many other factors such as other public support, alternative sources of finance, the activities of other actors in the innovation system and the broader macroeconomic context (see, for example, HM Treasury, 2011). Thus, innovation support acts as just one part of a complex science and innovation system involving multiple actors and programmes at supra-national, national and sub-national levels. Thus, companies may receive support from several

programmes, provided by multiple organisations. This support may be obtained simultaneously, successively or in an overlapping combination of the two and the attribution of any observed impact to any single intervention can be very difficult, with each programme being necessary but not sufficient on its own to achieve outcomes. Likewise, the success (as measured through evaluation) of one line of support may obscure the relative failure of an accompanying line of support.

Firm size also contributes to the challenge since, as the size of the target firm increases, the direct outcomes of public support may be difficult to distinguish from other forms of support as it will represent but one of several sources of income, etc. and the corresponding outputs will derive from a fraction of the firm's overall innovation portfolio. Even at the level of smaller companies, the process of innovation is in itself highly complex and is dependent on a range of factors and inputs which also militates against using relatively simplistic evaluation tools to determine cause and effect.

Section 3 will explore how three innovations in conducting evaluation studies (a behavioural change perspective, mixed-method approaches and big data) have the potential to address attribution and the other five challenges of conducting evaluation challenges, as discussed above.

3 BEHAVIOURAL CHANGE, MIXED-METHOD APPROACHES AND BIG DATA

3.1 *Aiming for and measuring behavioural change within companies*

3.1.1 *Introduction and definition*

Measuring and understanding behavioural effects (and methods to capture these) were mentioned in the previous MLE as one of the emerging issues in the evaluation of research and development (R&D) and innovation schemes. It is essential that the community of science, technology and innovation (STI) scholars, STI evaluators and STI policymakers have a better understanding of why and how policies and schemes do or do not work; not only what the short-term and long-term (desired) effects of an innovation scheme or programme are, but also the unintended or even undesired behavioural effects and impacts.

We need to know from policymakers what intended behavioural changes they are aiming for through various schemes, including their interaction. From the schemes' beneficiaries, we need to understand in detail how they are benefitting (or not) from, in our case, R&D business grants and how it affects their behaviour at various levels. Over their lifetime, innovating firms embark on an 'innovation journey' where they need to adapt their strategies, capabilities and innovation efforts according to the stage the firm is in, the type of innovations they are seeking and the sectoral or technological innovation system(s) they are (or want to be) part of. This implies that innovation policymakers need to be explicit in terms of the type of behavioural change they wish to facilitate.

3.1.2 *Theoretical underpinning*

In the literature, the topic of behavioural change caused by the use of business R&D grants and/or innovation schemes is mostly associated with the notion of behavioural additionality. Thus, next to measuring input and output additionality when evaluating schemes (especially when using econometric methods), there is a need to measure more precisely how company behaviour has changed as a result of a particular scheme. Additionality is an aspect of assessing the impact of R&D and innovation schemes.

The types and varieties of 'additionalities' discerned over the years have grown. Basically, the concept of additionality concerns the additional or extra R&D or innovation activity that results from public support for R&D and innovation. It is based on the neo-classical assumption that individual actors, due to the existence of knowledge spillovers, tend to underinvest in R&D and innovation. The rationale for public support for R&D and innovation is to prevent this underinvestment. The additionality reflects the additional R&D and innovation activity generated that *would otherwise not have been realised* without public support. For a long time, studies of additionality focused mainly on aspects of both input and output additionality. Input additionality is about the extra R&D and innovation effort or investment made by the firm as a result of public R&D and innovation support. Output additionality relates to the extra outputs generated by the firms as a direct consequence of public R&D and innovation support. Output additionality is harder to measure than input additionality, mainly due to the complexity of the

relationship between inputs and outputs. Hence, its observation requires a deep understanding of how firms behave (OECD, 2006).

Since the 1990s, more attention has been given to behavioural change occurring within companies and other actors as a result of public support for R&D and innovation, especially through the development of the notion of a third level of additionality, that of 'behavioural additionality'. Buisseret et al. (1995) define behavioural additionality (BA) as "the persistent change in the behaviour of the agents, which is exclusively attributable to the policy action, i.e. what difference a policy makes in those it supports". A more recent definition where attribution and persistency also feature prominently is the definition by Gök and Edler (2012) namely: "the persistent change in what the target is doing, how they are doing it and which is attributable to the policy action". Neicu (2016, p. 101) states that "behavioural additionality refers to permanent changes in firm processes and behaviour, such as newly acquired competences, the entry into new business areas or a change in working procedures, occurring because of policy intervention. Such changes may arise due to, among others, learning effects and knowledge spill-overs."

Behavioural additionality is generally less discernible compared to input and output additionality, as it can encompass all the factors related to a firm's innovation capabilities. These factors can be manifold and there is, as yet, no well-accepted set of indicators available for the quantitative measurement of behavioural additionality (although it can be assessed using more qualitative methods). Typical examples of behavioural additionality include: R&D and innovation projects that are started earlier, or completed faster, than originally anticipated; learning from projects (the first three are examples of 'project additionality'); or the performance of more ambitious projects (sometimes referred to as 'scale additionality'). Further examples are increased willingness to collaborate when innovating; the setting of joint R&D and innovation agendas with collaboration partners; and greater willingness to involve potential users in innovation processes ('cooperation additionality')[1].

Another categorisation of behavioural additionality is that presented by Roper et al. (2016, p. 12-17) who differentiate between three types of behavioural additionality using an organisational learning perspective, i.e.

1. congenital additionality: the change in the collection of competences that resides in an organisation due to public support for R&D and innovation and which is mostly assessed through the sum of employees' education and experience;

2. inter-organisational additionality: broadened or deepened innovation linkages of the organisation thanks to public subsidies for R&D and innovation;

---

[1] Categorisation of projects, scale and cooperation additionality is derived from Wanzenböck et al. (2013, pp. 6-7).

3. experiential additionality: the change in the ability to reconfigure routines and processes thanks to public support for R&D and innovation.

Interestingly, Roper et al. (2016) not only assess these three types of behavioural additionality, but also explicitly highlight two complicating factors that may lead to an underestimation of additionality. The first complicating factor is that it is not always clear to what extent the endurance or legacy effects of public R&D and innovation support are taken on board. This is especially key (but not necessarily limited) to behavioural additionality, as the focus is not on one-off project-related decisions to invest more in R&D and innovation, for example (input additionality), but on how firms learn over time, improve their innovation routines and capabilities, and eventually adapt to a changing context in which they operate. These behavioural changes are long-term effects that last longer than the duration of an R&D or innovation project for which a firm receives a grant. The second complicating factor is that the benefits of R&D and innovation support will not be restricted to the beneficiaries of R&D and innovation schemes, and the project outcomes will most likely also be of benefit to non-supported firms, typically through "knowledge spill-overs, technology diffusion, and knowledge exchanges within communities of firms" (Autio et al., 2008, p. 59 quoted in Roper et al., 2016, p. 8).

Some definitions associate behavioural additionality mostly with the scale, scope, level of risk and speed of R&D and innovation projects (i.e. a rather small interpretation of behavioural additionality) rather than organisational learning, changes in routines and capabilities over time (the broader interpretation of behavioural additionality). We also do not like to rule out endurance and the wider spillover effects of R&D and innovation schemes and therefore prefer to use the phrase 'behavioural change' here rather than 'behavioural additionality'. Innovation networks, clusters, triple helix- (industries-governments-universities) type combinations of actors or complete sectoral innovation systems can also change their collective behaviour. Examples include the ways in which R&D and innovation priorities are set or how sustainability (or other societal goals) are addressed through the R&D and innovation efforts taking place in an innovation system. More systemic policy tools attempt to target the network or system level and at this level it is also necessary to identify what behavioural change policy-makers are seeking and what sort of behavioural change is actually realised in practice. The difficulty of discussing behavioural additionality at the systems level is another reason why analyses might be restricted to the determination of 'behavioural change'.

Methods for capturing behavioural additionality (in the wider sense) can be qualitative (fully understanding the motives, context and other details, e.g. via case studies, interviews and surveys) as well as quantitative 'bang for the buck' approaches at both the macro and micro level. There are many examples of evaluations where combinations of methods are used not only to determine typical input and output additionality, but also the behavioural additionality that is induced by a scheme (or not). When reporting the results of evaluations, it is important to not only communicate the bang-for-the-buck type of results (i.e. focusing on input and output additionality), but it is also necessary to report the quantitative attempts to measure behavioural additionality (which in general is more experimental) in combination with detailed and more qualitative narratives

of how schemes are used in practice and impact on firm behaviour. We think these insights into behavioural effects are also needed to see how R&D and innovation schemes impact on their beneficiaries. With the increasing use of systemic instruments, which often comprise multi-goal and multi-method approaches to make changes at the systems level, there is a need to see how innovation systems change as a result of policy interventions. In addition, these higher-level analyses may require mixed-method approaches and the use of big data and linked-data approaches.

3.1.3 Potential for addressing the challenges of evaluating R&D&I policy

Taking the perspective of behavioural change can be effective when addressing the evaluation challenges mentioned in section 1. Addressing behavioural change primarily helps to address the issue of **heterogeneity**. Firms not only differ in sector and size, but also in why and how they use and benefit from schemes. Businesses are not homogeneous in their motivation for using schemes and how they affect their behaviour. It is essential for policymakers to appreciate these differences and to know whether they are reaching the desired population of firms and delivering the changes in behaviour they seek. An example discussed at the second MLE workshop was the evaluation of a Spanish loans scheme operated by the Centre for the Development of Industrial Technology (CDTI). A survey was used to capture the behavioural effects on different types of companies. These insights were used to redesign the scheme. A comparative study, looking at behavioural additionality in Poland, Spain, Portugal and Croatia, also addressed the heterogeneity of companies as well as country differences in the additionalities observed, thereby indicating different institutional set-ups and different policy mixes.

Secondly, investing more effort in understanding behavioural change also addresses the issue of the **time lag** before the effects of interventions start to materialise. Company behaviour does not change overnight and to fully understand the impact of policy schemes it may be necessary to assess the behavioural effects over several years – even after the finalisation of a scheme. To understand the real effects of schemes it is also necessary to understand the extent to which behaviour is affected either temporarily or more permanently. The example of the Support Programme for Seafood industry in Norway (discussed during the second MLE workshop) illustrates that even after the scheme ended, the firms involved continued to exhibit their changed behaviour. This again points to the need to extend the period considered by the evaluation and to leave sufficient 'incubation time' before assessing the scheme's impact.

Thirdly, to assess the wider **spillover and system effects** of schemes requires a fuller understanding of behavioural change both at the firm and wider system level. The development of a framework for assessing the direct, spillover and system-level effects of a scheme would be very welcome. Some countries are clearly experimenting with new evaluation frameworks to assess these. In the UK, the framework for evaluating the Catapult programmes is one example, while in Sweden the evaluation of the Swedish Challenge Driven Innovation (CDI) and Strategic Innovation programmes (SIP) are typical examples of schemes where assessing the behavioural change brought about in sectoral innovation systems is a major ambition. The same goes for the recently evaluated Topsector

Approach (priority sector approach) in the Netherlands. This is an example of a systemic instrument in which it is essential to understand the changed behaviour of multiple actors together with the ways in which they cooperate in order to be able to assess the success of the approach. Here, typically, evaluation benefits from insights acquired in innovation studies. Strangely enough, one of the insights gained during the Stockholm site visit was that thus far evaluation studies are failing to benefit sufficiently from the insights gained from innovation studies. By paying more attention to behavioural change, a logical bridge can be constructed between evaluation studies and innovation studies.

Finally, gaining a more fine-grained understanding as to how R&D business and innovation grants are affecting company behaviour can help tremendously in improving **the policy design** of schemes. By considering how interventions change firm behaviour, policy schemes can be fine-tuned and targeted better thereby reducing the deadweight loss of schemes (i.e. behaviour or changes in behaviour that would have occurred in the absence of the scheme). Closing the policy cycle is key and feeding back the insights gained from *ex-post* evaluations into the policy design phase is possibly one of the most powerful ways to create policy learning.

Finally, the use of mixed-method approaches for evaluation, including developing better ways to benefit from linking various data sets and exploring the use of big data, is essential for gaining a better understanding of how different categories of firms benefit from R&D business grants and wider forms of innovation support and how these forms of support affect their (and possibly others') behaviour.

3.2 Mixed-method approaches: new methods, new combinations

3.2.1 Introduction and definition

One of the main conclusions to emerge from the first cycle of the MLE which took place in 2016 was that, with regards to evaluation methods, there has been and continues to be an increasing trend towards the use of more sophisticated econometric analyses for measuring the impact of direct grants for the support of business R&D. At the same time, 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 also to examine the 'innovation journey of firms', particularly when they use R&D and innovations grants in a simultaneous or sequential manner or combine them with alternative forms of support.

Over many years, an extensive range of evaluation methodologies has been developed to assist policymakers in better understanding the results, outcomes, effects and impacts that arise from their policy support instruments. In addition, these evaluative tools help them to gain a better picture of how such instruments have been implemented, the extent to which they have achieved their objectives and how they might benefit from improvements. For the purposes of this report, we define mixed methods as the combination of multiple evaluation approaches that are applied in tandem to the assessment of policy instrument performance, with a primary focus on the evaluation of direct grant schemes to support R&D.

3.2.2 Theoretical underpinning

This section provides an explanation concerning the need for mixed methods in evaluation by examining the purpose of evaluation and the complexity of the policy instruments under consideration.

Evaluation serves a number of purposes under various rationales. In the past, it was often used primarily for justification, i.e. as a means of checking that the resources invested in a specific programme or scheme were being used appropriately and were generating the anticipated results. Such 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, the major purpose of evaluation now is to inform policy learning at different levels. Typically, three such levels have been defined:

- Operational learning: as a management feedback tool to improve the effectiveness, efficiency and quality of policy intervention, 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: evaluation is used in its 'traditional' sense to determine the outcome and impacts of policy measures and programmes, checking whether, and the extent to which, programmes have achieved their objectives. In this context, evaluation provides a method for policymakers to assess whether the assumptions that 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: evaluations serve to improve the efficiency and effectiveness of national innovation systems by guiding the design and formulation of intervention policies and programmes. They 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, etc. (adapted from VINNOVA, 2004, cited in Miles and Cunningham, 2006).

Evaluation may also be summative, formative or, as frequently occurs, a combination of the two. **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 more outcome-focused than process-focused and tends to be undertaken *ex post*. **Formative evaluation** typically takes place during project implementation with the aim of improving its design and performance. It

complements summative evaluation and provides insights into understanding why a programme works (or not). It also takes account of other factors (internal and external) that can influence 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[2].

Evaluations can also be conducted for a variety of audiences: for instance, programme managers will seek information on the implementation of their policy instruments 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 generated. Added to these, a broader range of policymakers will seek to determine the comparative effectiveness of different policy interventions and any synergistic or contradictory impacts they may exhibit.

Evaluations must also be appropriate to the specific modality of the policy instrument with which they are concerned (that is, how they interact with and influence the behaviour of the targets of their support). Policy instruments exhibit a large range of modalities (and targets), according to the specific purpose for which they have been designed and for the context in which they are implemented. Although in this MLE we are concerned primarily with the evaluation of policies to support business R&D, policymakers 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 support packages (Cunningham, Gök and Laredo, 2015; Edler et al., 2015). Thus, for example, direct grant schemes are now likely to include elements of training or inducements for intra- or inter-sectoral collaboration.

Allied to these developments and 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 and objectives, beyond that of simply stimulating additional RDTI activities (such as improving firms' absorptive capabilities and capacity; strengthening the quality of RDTI activities; supporting collaborative interactions for the production of new knowledge or supporting broader (multiple) interactions – e.g. through clusters or networks). Further context-driven objectives may also be applicable, 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.

The complexity and diversity of the outcomes, effects and impacts of policy interventions also necessitate a range of approaches that can adequately capture

---

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

them. For example, some of the outcomes and impacts may be captured in official statistics, company reports and other documents. Others 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 greater 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 (Denzin, 1984; Miles and Cunningham, 2006).

3.2.3 Potential for addressing the challenges of evaluating R&D&I policy

From the discussion above, we can see that no single evaluation approach is able to provide a fully comprehensive picture of the performance of a policy instrument. Even for policy instruments that have relatively simple modalities and restricted sets of policy goals, it is still desirable to employ several methodologies in combination. This allows us 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.

Over the years, several evaluation 'toolboxes' or guidance manuals have been developed from which policymakers and programme managers may select appropriate evaluation approaches. Examples include:

- The ePUB RTD Evaluation Toolbox (2002)

- DG Enterprise and Industry: Smart Innovation: a practical guide to evaluating innovation programmes (2006)

- HM Treasury, The Green Book: Appraisal and Evaluation in Central Government (2003)

- 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, these manuals serve only as a broad guide to the type of approach available – without careful adaptation to the context of the policy instrument under investigation they will fail to deliver a full understanding of the range of potential impacts on the innovation process policy instruments may have. Moreover, policymakers now seek to answer more complex questions which require increasingly sophisticated approaches to evaluation.

Typically, evaluation methods may be grouped according to the functions they play and the specific information they contribute (Miles and Cunningham, 2006):

- Methods for accessing and generating data (e.g. surveys, interviews or document review)

- Methods for structuring and exploring interventions (e.g. construction of counter-factual sampling approaches or Randomised Control Trials).

- Methods for analysis of data (e.g. descriptive statistics, econometric modelling)

- Methods for drawing conclusions (e.g. impact assessments).

Likewise, the ePub RTD Evaluation Toolbox (2002) distinguishes between quantitative approaches (statistical data analysis, modelling methodologies, etc.) and qualitative approaches (interviews and case studies, cost-benefit analysis, expert panels/peer review, network analysis, etc.).

It is beyond the scope of this report to provide a detailed review of all the available evaluation approaches but, for illustrative purposes, Miles and Cunningham (2006) offer a brief comparison of data-generation methodologies, listing the major advantages and challenges associated with their use (see table 1). Detailed discussion of each type of method is also provided in the report by Miles and Cunningham (2006). In each case, however, the selection of methods is entirely dependent on 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 factors.

Table 1 provides a brief overview of the purpose, advantages and challenges of methods that are used for the evaluation of R&I support schemes.

Table 1: A brief overview of evaluation methodologies

Method	Overall purpose	Advantages	Challenges
Surveys	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	Helps to 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 the 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 for an impression of how a programme operates without interrupting the programme; comes from a review of applications, finances, memos, minutes	Gathers comprehensive and historical information, does not interrupt the programme or client's routine in the 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	Gathers 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, can be expensive
Focus groups	Explore a topic in-depth through group discussions, 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 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	Usually quite time consuming to collect, organise and describe; represents depth of information rather than breadth

		programme to outsiders	

Finally, in terms of addressing the generic challenges set out in Section 2, it must be emphasised again that no single approach can provide all the answers. It is only through the careful selection of how and when to apply a tailored combination of methodologies that policymakers can attempt to ameliorate or minimise the effects of these ongoing challenges.

3.3 Big data

3.3.1 Introduction and definition

Big data is a label, or even a hype, that emerged around 2010. The term is shorthand for the growing opportunities to collect, process, analyse and use data. This includes structured and unstructured data, established and new data sources, and the potential for data linking and using new data analytical methods such as pattern recognition by means of machine learning (Gartner, 2011; McKinsey Global Institute, 2011; Mayer-Schönberger and Cukier, 2013; Kitchin, 2014).

Big data is enabled by ICT/digital and computing innovations. Examples of big data enablers are cheaper, smaller and better sensors (installed in products and in production systems); greater precision of earth-observation systems (e.g. using satellites for tracking and tracing); high-performance (cloud) computing for data storage and data processing; and software for data analytics and visualisation.

Another important enabler is the rise in internet applications and mobile apps. Online, more and more data are available about companies, products, consumers and employees. By means of web scraping, social media mining and other text-mining applications, data can be collected about a company's (new) products, partners, clients, number of employees, vacancies, etc. These opportunities further increase when governments open up administrative data such as business registries and the list of beneficiaries of subsidies (cf. open data). Moreover, the internet contains ever-more articles and papers in which companies and, in particular, researchers describe the results of their research and innovation activities. Opportunities for analysing these online publications (emerging topics, co-authors, location of authors, citations, etc.) are discussed under the label of altmetrics (Galligan and Dyas-Correia, 2013; Haustein et al., 2014).

Definitions of big data focus on the so-called 'Vs', including volume, variety and velocity of data (Gartner, 2011; Mayer-Schönberger and Cukier, 2013). This implies that big data is not only about volume ('more data') but also about a greater variety of data sources (sensor data, transaction data, administrative data, text on websites, surveys, etc.) and high-frequency or even real-time collection and processing of data (Kitchin and McArdle, 2016). Data linking contributes to the volume and variety of data. Different types of data about companies (or other actors) can be linked, using unique identifiers. Note that new data sources and data linking also create challenges. One such is veracity:

ensuring validity and data quality instead of using readily available yet non-relevant or poor data. Another challenge is visualisation: using more data while conveying clear information (Mayer-Schönberger and Cukier, 2013; Kitchin, 2014).

The definition from Taylor, Schroeder and Meyer (2014, p.1) is compatible with others that focus on the 'Vs' although it also makes it clear that data sources always have a link to objects. Moreover, the definition refers to data analytical tools and to the evolution from small to big data:

> *"Big data is a step change in the scale and scope of the sources of materials (and tools for manipulating these sources) available in relation to a given object of interest."*

This definition has been effective in a study about the state of the art and challenges in the use of big data for policymaking (Technopolis Group et al., 2015).

3.3.2 Theoretical underpinning

To discuss the potential and the challenges of big data for the evaluation of business R&D grants (and R&D and innovation policy in general), a brief reflection on evaluation theory is required.

At the heart of evaluation theory is the intervention logic, meaning the rationale behind a policy intervention and the logic about the mechanisms via which a policy intervention leads to the desired effects (the 'theory of change'). As mentioned in Section 3.1, the perspective of behavioural change is one way of looking at the intervention logic. For example, policymakers, companies and other stakeholders may conclude that companies' R&D and innovation behaviour should change: more R&D, larger or more risky projects, more collaboration, addressing a societal challenge, etc. What is hindering these companies? How can (which) policy interventions change the behaviour of (which) companies? Is it required that other actors, such as universities, also change their behaviour?

Subsequently, policymakers and evaluators can select the most relevant indicators to track if and how behaviour is changing and whether these changes are persistent. The next steps are to develop the most efficient and effective data-collection strategy and, linked to this, to design the evaluation strategy (methods, use of control groups, timing of the evaluation, objectives of the evaluation, etc.).

Big data, including new data sources and data linking, increases the options when developing the data-collection strategy of an evaluation. In general, this is positive news. However, big data also creates a challenge (or even a risk) of using data that is merely readily available rather than being relevant (Strassheim and Kettunen, 2014; Technopolis Group et al., 2015; Poel, 2017). Moreover, big data creates a challenge/risk of using low-quality data that can be collected efficiently (e.g. using open administrative data, web scraping and text mining) rather than investing in the collection of high-quality, robust data (Kitchin, 2014). These risks or trade-offs will change as big data becomes mature. For instance, text mining

is a typical example of an approach that requires trial and error, including triangulation with other methods.

In short, big data should not lead to (only) using data that is poorly linked to the intervention logic and is of low quality. Nor should big data lead to (only) using data analysis methods, such as pattern recognition, profiling and predictive modelling that are not proven in the context of policy evaluation. Ideally, big data approaches should be combined with established methods and datasets. Note that the results of policy evaluations can mean that support schemes (using public money) are continued, stopped or adapted.

The risks mentioned above explain why the use of big data by national and international policymakers and public agencies has only recently increased (Technopolis Group, Oxford Institute and CEPS, 2015; Bakhshi and Mateos-Garcia, 2016). This also holds true for the evaluation of R&D and innovation policy. The main uses of big data, so far, are data linking (e.g. Fraunhofer ISI et al., 2009; Gal et al., 2016; Technopolis Group and SEO, 2016; Research Council of Norway, 2016; Hertog et al., 2016) and analysing online publications (Harle et al., 2016; Prins et al., 2016; Bornmann et al., 2017).

3.3.3 Potential for addressing the challenges of evaluating R&D&I policy

Big data has the potential to address two of the major challenges of evaluating R&D and innovation policy: attribution and skewed impact distribution.

Big data, especially data linking, allows for the inclusion of more variables about companies in one dataset. Using unique identifiers for companies (such as Chamber of Commerce registration numbers and VAT numbers), it is possible to link data taken from multiple sources, such as business registries, national innovation surveys, targeted surveys, company websites and (commercial) company databases (Bloomberg, Capital IQ, Bureau van Dijk's Orbis database, etc.). Controlling for more variables increases the possibilities to attribute changes in company behaviour to policy interventions. These possibilities further increase when the dataset contains data about support provided by different schemes (regional, national and European, financial and non-financial). Note that data linking requires a number of checks regarding privacy, confidentiality and ownership of data.

Big data, especially new data sources and text mining, allows for analysing many if not all companies that received business R&D grants or other types of support. For instance, text mining of the update reports and (final) impact reports submitted by beneficiaries reveals, for example, which beneficiaries self-report what types of effects. Along the same lines, web scraping can reveal which companies launched which new products. As such, text mining addresses the evaluation challenge of skewed impact distribution. To summarise this challenge: when providing R&D&I support, few companies or projects can report substantial impact, while many report small or no impact. High-impact projects can be overlooked when using surveys (with low response rates) and sampling a small number of case studies.

4 MLE PARTICIPANTS' KEY MESSAGES AND NEXT STEPS

4.1 Introduction

As mentioned in Section 1, this report has been prepared for a Mutual Learning Exercise in the European Policy Support Facility. During the three MLE workshops, participants discussed how their ministries and agencies are evaluating business R&D grant schemes and other schemes for supporting companies' R&D and innovation. Table 2 below summarises 14 evaluations which are examples of the methods used by the participating countries (detailed descriptions are presented in the Annex). During the exercise, the participants elaborated on the methods used in their recent and planned evaluations. At the end of each workshop, participants were asked which key messages would be taken back home.

To improve the overview of national approaches, participants responded to a survey at the start of this MLE, and one at the end. The 'ex-post survey' also included a question about key messages.

Section 4 describes five key messages in terms of important, current developments in the evaluation of R&D and innovation support schemes. The five messages concern: data linking; assessing effects at the innovation system level; evaluation of the policy mix; (better) acknowledging the heterogeneity of companies; and building or fostering an evaluation community.

Table 2: Summary of recent evaluations of the MLE participants

Country	Evaluated scheme	Examples of a behavioural change perspective, mixed-method approach and big data
Austria	Innovationscheck programmes (vouchers)	Using data linking and a mixed-method approach, the evaluation assessed whether the vouchers led to persistent changes in the R&D&I behaviour of SMEs
Brussels Capital-Region	Doctiris: collaborative R&D projects between doctoral researchers and enterprises	The first evaluation of the Doctiris programme used different qualitative methods and explored how the programme leads to which types of effect
Croatia	Evaluation of the Croatian innovation system	Collaborative behaviour was promoted by means of agenda setting, using a stakeholder engagement strategy (Entrepreneurial Discovery Process)
France	Competitiveness clusters policy	Data linking enabled the creation of a control group and an assessment of input additionality
Germany	Innovative Regional Growth Cores	A mixed-method approach was taken 14 years after the introduction of the programme. This allowed for an assessment of persistent changes in company behaviour
Lithuania	Inno-vouchers LT scheme	Web scraping was used to collect additional data about beneficiaries and other companies
Norway	Multiple schemes	Linked data platform for monitoring and evaluating support schemes

Country	Evaluated scheme	Examples of a behavioural change perspective, mixed-method approach and big data
Norway	Technical Industrial Institutes in Norway	Three datasets were linked, and a quasi-experimental approach was taken to attribute effects to the intervention
Poland	Operational Programme (OP) on improving enterprise competitiveness	A mixed-method approach was used to fully understand how the programme led to effects on companies
Spain	Business R&D grants and other schemes	Comparing the effect of public support, before and during the recent economic crisis. Linking internal and external datasets and using a control group
Sweden	Entrepreneurial experimentation and collaboration	Detailed study of the complex and long-term behavioural effects which contribute to enhanced firm performance
Turkey	Technoparks	Using the linked data in the Entrepreneur Innovation System and taking a mixed-method approach
UK	Smart Scheme	A three-phase approach addressed the issue of time-lagged outcomes and effects
UK	Digital Catapult UK	The evaluation experiments with Agent Based Modelling, including the use of large (un)structured datasets, data linking and text mining of business registries

4.2 Data linking

The MLE's first survey and first workshop/site visit addressed big data. While MLE participants do not yet use new data sources, such as web-scraped data, or new data analytical tools, such as pattern recognition and profiling, they do use data linking. Nearly all MLE participants mentioned how their evaluation studies and, in some cases, data platforms make use of data linking. The example of Innovation Norway's data platform is discussed in Section 3. The UK is developing a similar platform, coordinated by Research Councils UK and Innovate UK. The UK system emphasises sharing information about research projects (activities and results) although the system can also be used to analyse which organisations receive which types of public support. Gradually, data about more support schemes is added to the system and more data is fully disclosed online (open data). Sweden (VINNOVA), France (France Strategie) and other countries are moving in this direction, too. The discussion also touched upon the cost involved in developing integrated, shared data platforms and, in particular, in keeping the data up to date.

With reference to data linking in individual, specific evaluations, MLE participants mentioned this is likely to become the norm. For instance, statistical offices in the UK, Norway and other countries that are involved in evaluations can link data from public agencies (e.g. about beneficiaries) to survey data and official (micro) data about companies. The mixed-method approaches discussed in Section 3 also concern data linking. See, for example, the evaluations from France, Norway and Spain. One of the red threads through the three MLE workshops was that mixed-

method approaches, involving data linking and/or triangulation, help to attribute behavioural change to policy interventions.

One data-linking challenge discussed in the Norway site visit concerns the development of (inter)nationally shared ontologies for R&D and innovation support schemes, actors and effects: in short, using the same classification and terminology for support schemes, SMEs and other types of actors and for effects on product innovation, productivity, employment, etc. This facilitates data linking, benchmarking and evaluations of the interaction between different support schemes (e.g. looking at several support schemes that address SMEs). The OECD and EU REITER project seeks to develop such a taxonomy.

Other data-linking challenges, mentioned during the MLE, are managing data ownership, data confidentiality and privacy (cf. the European General Data Protection Regulation), data integrity and security (with several actors contributing to or using linked datasets).

4.3 Effects at the level of innovation systems

In each of the three site visits, MLE participants mentioned how support schemes that address companies (e.g. business R&D grant schemes) can have indirect effects on the innovation systems in which these companies are active. Innovation systems can be defined in different ways. Examples include national innovation systems, technological innovation systems (for emerging technologies such as photonics), systems that address transitions or missions (such as renewable energy), sectoral innovation systems, and regional innovation systems (a concept that overlaps with the concept of regional economic clusters).

As mentioned in Section 2, business support schemes may have requirements regarding external collaboration partners and the technological, economic, social and environmental challenges that must be addressed. Here, policymakers explicitly or implicitly aim for changes at the level of innovation systems (e.g. more collaboration or a shared agenda). For instance, one of the evaluations commissioned by VINNOVA included a survey in which companies were asked about any indirect, spillover effects on their project partners and other actors (effects on innovation, turnover, employment, etc.).

At a broader level, the R&D and innovation behaviour of individual companies can trigger changes in the behaviour of other actors, such as universities and research institutes – for example, R&D-performing SMEs that call upon research institutes for technical advice. These indirect effects at the aggregated level of innovation systems are likely to increase when other, complementary support schemes address universities, research institutes and public-private collaboration (see Section 5.4 on the policy mix).

In the site visits which focused on behavioural change and mixed-method approaches, MLE participants mentioned how policymakers can put more emphasis on change at the level of innovation systems. For instance, they should try to steer the agenda, the composition of social networks and the knowledge base in the direction of environmental challenges. Thus, this systems-level

perspective influences the design of business R&D grant schemes and other schemes targeted at companies.

MLE participants also stressed the importance of analysing and influencing the impacts and effects on behaviour at the level of individuals. In all organisations, it is the people who become aware of opportunities, change their routines (e.g. become more open for collaboration and risk-taking), develop tacit and codified knowledge and change their skills sets. People collaborate, change jobs in innovation systems and thus influence those innovation systems.

4.4 Evaluation of the policy mix

Although this MLE primarily looks at individual R&D and innovation grants, a challenge obviously requiring attention concerns the evaluation of policy mixes. First, R&D business grants are part of a wider set of individual policy schemes that firms benefit from either at the same time or subsequently. It is key to see at what stage in a firm's innovation journey an R&D business grant is needed to spur R&D and innovation and what is the expected behavioural change that the scheme aims to induce. Firms differ in the extent to which they are in need of formalised R&D and also at what stage of their development they can benefit most from such a scheme. If an R&D business grant is intended to change behaviour, it must be targeted at those firms that will most likely change their behaviour due to the additional support (rather acting as a recurring option to those firms that have already benefited from performing R&D). Increasing insights into the behavioural change caused by an R&D business grant can help position an individual scheme in the wider policy mix.

Secondly, there is an increasing trend for policy schemes themselves to be delivered as policy packages, often addressing several objectives or goals. In these mixes, support for R&D is only one of the elements included in the scheme which may also incorporate, for example, voucher schemes (to enable access to external advisory services), networking schemes (to promote cross-fertilisation of ideas between projects), public procurement (to address users' demands for innovation) and awareness activities (to promote recognition of the benefits of innovation, etc.). Again, understanding what types of behavioural change are induced by individual schemes is essential. At the same time, policymakers have to clearly articulate the types of behavioural change they expect their policies to deliver. Such understanding is required to be able to design the appropriate policy mixes. Similarly, systemic policies, which address the need of a broader spectrum of innovation actors, are also policy packages that can be customised not only to the needs of individual firms, but also to the particular needs of sectoral innovation systems. Thus, it is crucial not only to address the direct effects of policy schemes, but also their systemic and wider spill-over effects.

Thirdly, policy mixes are important since firms seldom benefit from a single scheme; they may use multiple schemes, either concurrently or over a period of time. Therefore, there is a need to develop a better understanding of policy mixes and to be able to attribute which schemes induce which effects in company behaviour. The challenge of attribution also requires evaluators, when evaluating an individual scheme, to know from what other schemes a firm is benefiting or has benefitted from. This requires much richer data sets. Big data, and data

linking offers new possibilities to derive a better understanding of the policy mixes from which firms are benefitting and, more precisely, for attributing the effects of individual schemes. This cannot be done only by assessing quantitative datasets with increasingly advanced econometric methods, but clearly also involves understanding the innovation journeys that various categories of innovating firms are experiencing. In this case, mixed-method approaches are needed.

4.5   Acknowledging the heterogeneity of companies

Businesses exhibit an enormous range of variations, differing not only in size, but also by region, sector and structure, as well as objectives and motivations. How do policymakers account for this heterogeneity when designing, implementing and evaluating new schemes? In this particular MLE, the starting point was R&D grant schemes but discussions have diverged more widely to cover other schemes that spur innovation. This is because many of the schemes reviewed entail issues and problems that apply to the evaluation of business R&D grants and offer generic lessons that can be used in several contexts. In all three workshops, one of the issues to emerge concerned how effective schemes were in addressing the needs of various categories of firms and whether there was a need to focus on particular subsets of firms in order to better understand any potential variations in scheme impact. It became evident that the heterogeneity of firms (both participants and 'control groups') is not always sufficiently well addressed. Some of the key problems that arose in the discussions are mentioned below.

Do programmes target, and subsequently reach, the 'right' categories of beneficiaries? Do we overemphasise the role of start-ups or spin-offs and large firms to the detriment of middle-sized firms? Are we preoccupied with high growth (technology-based) firms at the expense of more established 'regular' SMEs with incremental growth? Do we ignore some categories of firms by focusing on specific technology groups? To what extent would we address different categories of firms, if *societal challenges* are taken as the starting point of schemes?

It is also relevant to intensify efforts, in evaluation practice, to control sufficiently for the various factors that contribute to a heterogeneous population of firms. For instance, the effects of support schemes on individual firms may depend not only on the size, age and technology-intensity of firms, but also on their geographic location and existing collaborations – in short: control for more variables. As mentioned during the London site visit, many statistical models assume a level of homogeneity where sufficient controls cannot be included – and this may lead to (ambiguous) bias in the results of the analyses performed.

Finally, it was again emphasised that the notion of the 'innovation journeys' of individual firms has significant implications for their motivation and choices underlying their use of R&D and innovation schemes. This links to the policy-mix aspect. At every stage of the innovation journey, a different mix of policy schemes may be appropriate (if no appropriate schemes are available, then this is a clear message to policymakers). For example, voucher and tax credit schemes are often seen as 'easy', low-effort, entry instruments from which less-experienced innovators can benefit, whereas R&D programmes can be more suitable for

experienced firms (or those with more time resources to overcome the more complex entry requirements). For high-tech start-ups, the picture is different yet again: for example, being located on campuses or technoparks, working with universities, industrial doctorates, etc. It is possible to construct logical paths or 'innovation careers' within which firms avail themselves of different combinations of schemes through time. When designing new schemes to enter the policy mix, policymakers can be more explicit about how these schemes fit the innovation journey of specific types of firms.

4.6   Building an evaluation community

Countries that are more experienced in terms of the evaluation of R&D and innovation support schemes often benefit from the presence of an active evaluation community. These communities take various shapes and may comprise policymakers, public agencies, statistical offices, consultants, academics, industry associations and other stakeholders.

As an example, in the UK, there is a strong evaluation culture, engendered by the long-standing recognition of the importance of policy learning that evaluation can deliver. Efforts are made to formalise and disseminate the UK's evaluation practices throughout government by actors such as HM Treasury, through their production of guides such as the Green Book and the Magenta Book. Consequently, government funding agencies (like the Research Councils and Innovate UK) routinely undertake evaluations of their support programmes. Departments such as the Department for Business, Energy and Industrial Strategy (BEIS) encourage dialogue and learning about evaluation developments through regular meetings of evaluation users and practitioners.

At an intermediate level, actors such as the What Works Centre play an important role in gathering, translating and distributing knowledge and insights about evaluation, particularly at the local level, while Nesta acts as a focal point for emerging issues in evaluation activity. Both these actors also endeavour to foster greater levels of experimentation in evaluation practice itself. Moreover, the UK, like Austria, Australia, Finland and Norway, is engaged in the Innovation Growth Lab, an international community that develops and tests different approaches to support innovation, entrepreneurship and growth.

On the practitioner side, a large number of private consultancies operate in the UK and have developed close linkages with their (public sector) client base which enables a fruitful dialogue on the use of novel approaches for evaluation.

Finally, the UK hosts a sizeable and highly engaged academic sector which focuses on policy issues, including evaluation. Academics widely disseminate the outcomes of their research through scholarly articles and through more targeted grey literature and direct interactions with the policy community, whilst some develop practical experience via active participation in evaluations in cooperation (and competition) with the private consultancy sector.

Of particular importance in fostering an active and effective evaluation community is the notion of openness. Policy learning can be maximised when the results and outcomes of public funding scheme evaluations are made available to

a broad audience, beyond the immediate owners or sponsors of a policy instrument. Here, the efforts of supra-national bodies, such as the European Commission, play a significant role. One example is that of the EC-supported SIPER (Science and Innovation Policy Evaluations Repository) initiative which, *inter alia*, seeks to collect, categorise and analyse evaluation reports from a broad range of international sources and to make them freely available online to all interested parties[3]. Another example, of course, is the Commission's Policy Support Facility, of which this particular MLE and its predecessor are examples of efforts to share experiences.

The UK is by no means alone in developing a strong evaluation culture. Other European examples are the Nordic countries and, in the Netherlands, where national statistical offices offer their assistance to policymakers and evaluators by making their micro datasets available for evaluations. Particular efforts include linking data on the use of instruments with statistical data which offers major opportunities for more advanced econometric types of evaluations. Another example is France Stratégie, a centre of expertise under the authority of the French prime minister which hosts the National Commission for the Evaluation of Innovation Policies (CNEPI).

Other initiatives that are helping to build a community include agencies that coordinate evaluations and develop communities of evaluation practice (such as TAFTIE) or the activities of evaluation societies (such as the Austrian Platform for Research and Technology Policy Evaluation, fteval or the German Evaluation Society, DeGEval)[4].

Whilst the creation of linkages between those concerned and interested in the practice of evaluation is a very important factor in the creation of an evaluation community, a key ingredient concerns the types of skills required by the members of that community. These will vary according to the people involved, who may have different roles in the evaluation and policy process. Thus, the required skill sets will range from an understanding of policy governance processes, through to highly technical capabilities encompassing the application of advanced econometric techniques, data processing and visualisation, and social science approaches to information gathering and analysis. Policymakers will also need to adapt to new approaches, such as the use of real-time monitoring, data sciences, and stakeholder and policymaker engagement in evaluations, together with broader concepts such as 'networked governance' wherein several public and private actors co-fund and manage support programmes. At all levels, however, the need for increased dialogue and dissemination of good practice (and lessons learned) will play a crucial role in developing evaluation communities.

---

[3] http://si-per.eu/ and https://rio.jrc.ec.europa.eu

[4] http://www.taftie.org/; https://www.fteval.at/ and https://www.degeval.org/en/home/

5 REFERENCES

Bakhshi, H. and Mateos-Garcia, J. (2016). *New Data for Innovation Policy*. Nesta Working Paper.

Barber, J., B. Brinkley and P. Hills (1994). *An evaluation of the Small firms Merit Award for Research and Technology (SMART)*. London: Department of Trade and Industry Research and Technology Policy Assessment Unit.

Bornmann, L., R. Haunschild and W. Marx (2017 *forthcoming*). Policy documents as sources for measuring societal impact: How often is climate change research mentioned in policy-related documents?

Cunningham, P., A. Gök and P. Larédo (2013). *The impact of direct support to R&D and innovation in firms*. Nesta Working Paper 13/03.

Cunningham, P.N., A. Gök and P. Larédo (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.

Cunningham P., P. den Hertog and V. Peter (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. (2018). *Combining Mixed Approaches to Evaluation*. Thematic report prepared for the Horizon 2020 Policy Support Facility, Mutual Learning Exercise on Evaluation of Business R&D Grant Schemes. Brussels: European Commission.

Den Hertog, P. (2018). *Capturing Behavioural Change*. Thematic report prepared for the Horizon 2020 Policy Support Facility, Mutual Learning Exercise on Evaluation of Business R&D Grant Schemes. Brussels: European Commission.

Denzin, N.K. (1989). *The Research Act: A Theoretical Introduction to Sociological Methods*. Englewood Cliffs, NJ, Prentice Hall.

Edler, J., Cunningham, P., A. Gök and P. Shapira (2015). *Handbook of Innovation Policy Impact*, Edward Elgar Publishing Limited, 2015, ISBN 1784711845, 9781784711849.

Fraunhofer ISI, Idea Consult and SPRU (2009). *The Impact of Collaboration on Europe's Scientific and Technological Performance*. Final Report, Karlsruhe, Brussels, Brighton, March 2009.

Gal, D., W. Glanzel and K.R. Sipido (2016). Mapping cross-border collaboration and communication in cardiovascular research from 1992 to 2012. *European Heart Journal,* 2016-0, 1-10.

Galligan, F. and Dyas-Correia, S. (2013). Altmetrics: Rethinking the way we measure. *Serials review*, *39*(1), 56-61.

Gartner Inc. (2011). *Information management in the 21st century*.

Gök, A. and Edler, J. (2012). The use of behavioural additionality in innovation policy making, *Research Evaluation* 21, pp. 306-318.

Gök, A., A. Waterworth and P. Shapira (2015). Use of web mining in studying innovation. *Scientometrics* 102, 653-671.

Harle, C.A., J.R. Vest and N. Menachemi (2016). Using Bibliometric Big Data to Analyze Faculty Research Productivity in Health Policy and Management. *Journal of Health Administration Education*, 33(2), 285-293.

Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., H. Shema and J. Terliesner (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics*, 101(2), 1145-1163.

HM Treasury (UK) (2011). *The Magenta Book: Guidance for evaluation*.
Available at:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/220542/magenta_book_combined.pdf

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Los Angeles: Sage.

Kitchin, R. and McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society* 3(1): 1-10.

Mayer-Schönberger V. and Cukier, K. (2013). *Big Data: A Revolution that Will Transform How We Live, Work, and Think*. New York: Houghton Mifflin Harcourt.

McKinsey Global Institute (2011). *Big data: The next frontier for innovation, competition, and productivity*.

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

Neicu, D. (2016). *Learning how to behave. An analysis of the behavioural changes induced by public support for R&D*. PhD thesis, University of Leuven.

O'Doherty, D. and Arnold, K. (2003). *Understanding innovation: The need for a systematic approach*. Seville: JRC-IPTS.

OECD (2006). *Government R&D Funding and Company Behaviour: Measuring behavioural additionality*. Paris: OECD.

Penfield, T., Baker, M.J., R. Scoble and M.C. Wykes (2014). Assessment, evaluations, and definitions of research impact: A review. *Research Evaluation*, Volume 23, Issue 1, pp. 21-32.

Poel, M. (2017). *Big data: data linking, new data sources and new data analytics methods.* Thematic report prepared for the Horizon 2020 Policy Support Facility, Mutual Learning Exercise on Evaluation of Business R&D Grant Schemes. Brussels: European Commission.

Prins, A., Costas, R., van Leeuwen, T.N. and Wouters, P.F. (2016). Using Google Scholar in research evaluation of humanities and social science programs: A comparison with Web of Science data. *Research Evaluation*, 25(3), 264-270.

Research Council of Norway (2016). *Impact analysis of the technical-industrial research institutes in Norway*. Study prepared by Technopolis Group.

Roper, S. and Hewitt-Dundas, N. (2016). *The legacy of public subsidies for innovation: input, output and behavioural additionality effects*, Enterprise Research Centre, Warwick.

Strassheim, H. and Kettunen, P. (2014). When does evidence-based policy turn into policy-based evidence? Configurations, contexts and mechanisms. *Evidence & Policy: A Journal of Research, Debate and Practice* 10(2): 259-277.

Taylor, L., Schroeder, R. and Meyer, E. (2014). Emerging practices and perspectives on Big Data analysis in economics: Bigger and better or more of the same? *Big Data & Society,* July-December 2014, 1-10.

Technopolis Group, Oxford Internet Institute and CEPS (2015). *Data for Policy: A study of Big Data and Other Innovative Data-Driven Approaches for Evidence-Informed Policymaking: Report about the State-of-the-Art*. Prepared for the European Commission.

Technopolis Group and SEO (2016). *Evaluatie Stichting voor de Technische Wetenschappen* (STW) (in Dutch). Report for the Dutch Ministry of Economic Affairs and NWO.

Thelwall, M., Haustein, S., V. Larivière and C.R.Sugimoto (2013). Do Altmetrics Work? Twitter and Ten Other Social Web Services. *PLOS ONE* 8.

Wanzenböck, I., T. Scherngell and M.M. Fisscher (2013). *How do firm characteristics affect behavioural additionalities of public R&D subsidies? Evidence for the Austrian transport sector*. AIT, MRPA Paper, No 77552.

What Works Centre for Local Economic Growth (2015*). Innovation: grants, loans and subsidies*. Evidence Review 9. London: What Works Centre for Local Economic Growth.

6 ANNEX: EVALUATIONS PRESENTED BY MLE PARTICIPANTS AND INVITED EXPERTS

6.1 Interim evaluation of the Austrian Innovation voucher programme

6.1.1 Introduction

The Innovationscheck programmes (I-Scheck programmes) are SME-oriented grants which aim to encourage these companies to participate in regular R&D and innovation activities, to promote the transfer of knowledge between SMEs and the scientific sector, to support the efficiency and effectiveness of R&D and innovation projects, and to bring their results to market maturity more quickly. The programme – with a smaller innovation voucher (EUR5,000, introduced 2007) and a larger one (EUR 10,000, introduced 2011) is aimed in particular at smaller SMEs which do not undertake regular innovation activities and are without their own R&D personnel. The programme was evaluated in 2016 by an external evaluator[5]. During the evaluation period, 4,500 vouchers were allocated to more than 3,700 firms. As a consequence of the evaluation, both types of vouchers were discontinued and a new EUR 10,000 (plus a 25% own contribution) voucher was introduced from 2018 onwards.

6.1.2 Methodology

A mixed-method approach was used, combining desk research, administrative data and interviews and focus groups with stakeholders. For the first time, data about the innovation vouchers users was coupled with R&D data from Statistics Austria, which helped to enhance the quality and precision of the evaluation. The scheme was evaluated quite positively: 68% of the funded firms were newcomers and had not received any previous governmental R&D or innovation funding (thereby contributing to broadening the base of R&D-performing firms in Austria). Of these newcomers, a quarter continued with R&D and innovation projects, half of them with more complex projects. Some 10% of the newcomers subsequently made use of more complex R&D and innovation schemes.

6.1.3 Evaluation challenges

From the evaluation (particularly the interviews and focus groups) it was evident that new networks emerged between SMEs and knowledge institutes, and that firms were able to bring innovations (closer) to market. It was also evident from these qualitative methods that, without the vouchers, R&D and innovation projects would not have been executed or would have been performed at a slower pace and that the vouchers helped to raise project quality. The scheme also improved SMEs' readiness to experiment with new ideas. The relatively low number of newcomers that subsequently made use of more complex or advanced R&D and innovation schemes was seen as an indication that there is a gap between the typically accessible voucher scheme and more advanced and

---

[5] Jud, Th., Handler, R., S. Kupsa and S. Puhn-Weidiger (2017), Evaluierung der Innovationsscheck-Programme, Convelop, Graz/Wien

complex schemes. The evaluation also showed that applications for the smaller innovation voucher were declining and that no R&D project could be supported with just EUR 5,000. One scenario proposed was to stop the smaller innovation voucher and keep the larger version, but with a mandatory contribution from the SME itself to ensure its commitment to the proposed project. The combination of methods clearly demonstrated how the scheme was used in practice thereby increasing the behavioural insight into the use of the innovation vouchers.

6.2 Evaluation of the Doctiris programme in the Brussels Capital-Region

6.2.1 Introduction

The Doctiris instrument, operated by Innoviris (the Brussels Institute for the encouragement of scientific research and innovation) was created in 2011. It entails a collaborative R&D project between doctoral researchers and enterprises, in which an enterprise hosts the doctoral candidate (for at least 50% of the time). Its objective is twofold: promotion of bilateral collaboration between academia (university college, university, collective research centre) and enterprises, and reinforcement of the innovation potential of the industrial fabric in Brussels through knowledge transfer to the enterprises (via the research content, thesis and researcher) and vice versa. Public support accounts for 100% of the PhD salary, paid via the university. Budgets range between EUR 100k and EUR 400k. Since 2011, the programme has funded 27 projects for a total budget of EUR 7.1m. Nine projects have been finished successfully and six have been abandoned.

IDEA consult was commissioned by the Regional Science Policy Council to evaluate the knowledge transfer achieved between research institutes and enterprises (and the indirect benefits for the Brussels Capital-Region) and to propose prospective scenarios for optimising the type of valorisation that could be expected.

6.2.2 Methodology

The study was based on qualitative research methods and a literature review. The knowledge transfer achieved was evaluated through a number of case studies. Semi-structured interviews were conducted with supervisors in the enterprises (including SMEs as well as large companies), promoters and doctoral students. Eight of the 27 funded projects participated. All projects were at a sufficiently advanced stage (at least in year 3, if not finished) and one abandoned project was included. Projects had different levels of technological intensity. Tangible and intangible knowledge transfer were inventoried, and results, barriers and areas for improvement were identified. Benchmarking with similar programmes (Denmark, France, Flanders) was carried out and, finally, prospective scenarios were developed, taking into account the lessons learned and possible routes for extension (considering the implications of non-profit and public hosts).

6.2.3 Evaluation challenges

Each type of actor involved in the programme (promoter, enterprise, doctoral candidate) was interviewed, and a sufficient variation between SMEs and large enterprises and in the projects' actual progress was achieved. Nevertheless, the limited number of research projects and actors involved in the evaluation (due to a limited budget) inhibited any generalisation about the effectiveness of the programme. The selection of the research projects and actors was not random and was based on an estimation by the internal programme manager. Furthermore, the majority of projects were still running, thus preventing a true '*ex-post*' evaluation of the project. Moreover, the assessed projects were the longest running, excluding those which started after 2014.

Thus, the final results of the evaluation are non-representative and do not allow for attribution. Nevertheless, the evaluation was helpful as an intermediary monitoring tool. It underscored the programme's added value and helped Innoviris to identify in a very tangible manner a significantly broad spectrum of motivations for project submissions, and the nature of the types of transfer between enterprises and academic parties. The study also enabled the barriers and financial, intellectual and research output returns to be determined for both academic and industrial actors. The additional benchmark put the results in perspective and supported the development of prospective scenarios for adjustments to the future programme.

6.3 Evaluation of the Croatian Innovation System through EDP (Entrepreneurial Discovery Process)

6.3.1 Introduction

Prior to designing Croatia's National Smart Specialisation Strategy (S3), an extensive background report on the Croatian Innovation System was prepared to evaluate its current situation and set-up and to address future S3 priorities. The report was based on a series of events with stakeholders. The EDP was part of this (*ex-ante*) evaluation.

6.3.2 Methodology

Over a two-year period, four rounds of partnership consultations were organised through five regional workshops, and focus groups with about 800 participants, including representatives from universities, research institutes, business support organisations, regional development agencies, local governments and representatives of the business sector (SMEs and large companies, including clusters). In addition, a questionnaire was designed for each of the focus groups as a basis for discussions and to facilitate relevant conclusions.

6.3.3 Evaluation challenges

The evaluation gave an overview of the complexity of the Croatian innovation system and also pinpointed the missing links within the system (lack of communication, possibilities to collaborate, funding gaps, certain needs of example companies, etc.). It further helped discussions on the missing links and, following the methodology of EDP (RIS3 guideline), it provided a better

understanding of what the investments would look like until 2020. A plan was developed to launch business R&D grant schemes, supporting partnerships between businesses and universities. The evaluation of these schemes has yet to be done.

Two challenges emerged during the EDP process. The first was to narrow down the areas of interest of particular groups and guide them in the direction of collaboration and discovery of mutual interests and priorities. The evaluation also pointed to the need for collaboration, not only on projects but also at the strategic level, i.e. to change behaviour and develop joint projects and shared agendas. The second challenge was to maintain and involve a large variety of stakeholders during the overall period of the EDP and development of the S3 strategy, in order to tackle the risk of bias towards the interests of specific groups (e.g. large enterprises). The revision and interim evaluation of the S3 are planned for the beginning of 2019.

6.4 *Evaluation of the competitiveness clusters policy in France*

6.4.1 *Introduction*

In 2017, the National Commission for the Evaluation of Innovation Policies (CNEPI) published an assessment of the French competitiveness poles policy (CNEPI)[6],[7]. Launched in 2004-2005, this scheme aims to foster collaborative R&D projects between companies, public research labs and higher education institutions in order to develop innovative business activity in approximately 70 French clusters ("competitiveness pôles"). However, nearly two thirds of cluster companies do not invest in R&D activities, and their membership (in return for a financial contribution) is probably motivated not only by potential spill-over effects but also by the possibility to benefit from the range of services offered by the cluster development teams.

6.4.2 *Methodology and main results*

The econometric study aimed to measure the impact of membership of a competitiveness pole on individual beneficiaries (firms). It mainly relied on the difference-in-differences evaluation method which controls for selection biases due to differences in characteristics between member and non-member companies (the counterfactual). The existence of unique business identifiers enabled several datasets to be linked, making it possible to consider (and control for) the major channels through which agencies or ministries try to boost private

---

[6] Commission nationale d'évaluation des politiques d'innovation (2017), *Avis sur la politique des pôles de compétitivité*, February: http://www.strategie.gouv.fr/sites/strategie.gouv.fr/files/atoms/files/avis_pole2017annexe_02.02.pdf

[7] Ben Hassine, H. and Mathieu, C. (2017), *Évaluation de la politique des pôles de compétitivité : la fin d'une malédiction ?*, Document de travail de France Stratégie N° 2017-03, February: http://www.strategie.gouv.fr/sites/strategie.gouv.fr/files/atoms/files/dt-poles_de_competitivite-ok.pdf

R&D activity, notably via a generous research tax credit. The econometric study looked at company-level data from 2006 to 2012.

The study found that belonging to a competitiveness pôle could be correlated with a significant increase in terms of total R&D spending by 2007 and self-financed R&D by 2009. Each euro of government funding received in 2012 generated an average of EUR 3 in R&D spending, nearly EUR 2 of which is self-financed. SMEs enjoyed an obvious boost, but the effect was less pronounced for larger companies. This additionality effect on R&D expenditure among beneficiary SMEs is important as it represents a departure from the seminal study concerning the clusters support schemes in France[8]. The 2014 study was based on the same evaluation method as the 2017 study, but only examined the first phase (2006-2008) of the cluster policy. The combined results of the 2014 and 2017 studies suggest that the effectiveness of the programme has increased over time, at least in terms of *input additionality*. Similarly, the 2017 study found that, on average, cluster companies had 27.5% more R&D employees compared to non-member firms. Neither the 2014 nor the 2017 study could find the (expected) effect that companies belonging to a cluster score better on *output and outcomes indicators*, such as patent filings, sales, value added, investment, exports, total staff or exports. The explanation could be due to time-lag effects.

6.4.3 Evaluation challenges

The challenge of addressing the time lag between the policy intervention and measuring effects was overcome for input additionality but not for output and outcomes additionality. A follow-up study might be effective in addressing this challenge. The use of a control group addressed the challenge of attributing changes in companies' R&D investments to the cluster policy, especially for SMEs. Because most large companies are members of these clusters, the difference-in-differences method does not allow for the construction of a control group for those that invest significantly in R&D (i.e. more than EUR 16m a year), even though they are major beneficiaries of the policy.

The two evaluation studies did not capture spill-over effects. Particularly in those regions where clusters are located, cluster policy leads to knowledge spillovers to non-participating companies. Thus, it is not sufficient to consider the impact of this cluster policy solely on individual beneficiaries. Hence, France Stratégie and the General Commission for Territorial Equality (CGET) have commissioned a study to investigate the impact of the French competitiveness cluster policy on their region. This study[9] is based not only on econometric regressions and on interviews with stakeholders but also on network analysis, notably in order to analyse possible behavioural changes in terms of the propensity to collaborate.

---

[8] Bellégo, C. and Dortet-Bernadet, V. (2014), "L'impact de la participation aux pôles de compétitivité sur les PME et les ETI", *Économie et Statistique*, N°471, pp. 65-83: https://www.insee.fr/fr/statistiques/fichier/1377928/ES471C.pdf

[9] Eurolio and Technopolis (2018), Impacts économiques et territoriaux des pôles de compétitivité selon les territoires, forthcoming.

6.5 Evaluation of the Innovative Regional Growth Cores programme in Germany

6.5.1 Introduction

The Innovative Regional Growth Cores (Innovative regionale Wachstumkerne) funding programme is run by the German Federal Ministry of Education and Research (BMBF). It focuses on regional alliances involving companies, universities and research institutes that already have a joint technology platform in their region and boast unique selling points in their area of competence. The programme aims to effectively use and develop the competences and resources available in East Germany. Funding is provided for collaborative R&D projects in companies (mainly SMEs) with research institutions, bound together by a joint strategy. Since the start of the programme in 2001, it has supported 55 initiatives and, at the end of 2017, the approved funding amounted to EUR 384m.

6.5.2 Methodology

In 2015/2016, an evaluation was conducted by iit-Berlin, 14 years after the introduction of the programme. The entire programme period (2001-2014) was analysed. It was therefore possible to look at long-term effects such as persistent collaborations, behavioural changes in companies and continued investments in R&D following the end of programme participation. The evaluation also looked at economic effects such as the changing market position of companies or research organisations that created spin-offs.

The evaluation combined an online survey of all participants in the entire funding period, qualitative interviews with different stakeholders (programme owners, funding agency, experts involved in selection process, consultancies which offered support for funded companies) and 10 case studies of funded projects across the entire funding period. Some 350 companies and research institutes funded by the programme participated in the survey.

6.5.3 Evaluation challenges

The long period covered by the evaluation led to several challenges. The possibility to look at impacts of projects which had finished some 6-10 years previously was a unique opportunity because the indirect effects of funding could be observed, an advantage which is often unachievable in standard evaluations.

There were also some limitations. The response rate was lower for projects which had ended earlier and higher for the most recent ones, largely due to changes in the personnel concerned. However, the absolute number of questionnaires for projects in the first programme phase was sufficiently high to provide a reasonable interpretation of the data. In addition, case studies showed that the ability of the former participants to remember details of projects realised 10 years ago was weakening. Another challenge concerned causal attributions: the impact of participating in the programme was harder to describe for projects from the early period because, year on year, more external factors influenced the companies' development ("diluting effect").

A second challenge was linked to the use of a mixed (qualitative/quantitative) methods approach. The use of different methodologies was useful in gaining insights into different aspects of the programme which could be analysed individually by the most appropriate methodological approach. Programme outputs, outcomes and impacts could be generalised by extrapolating the individual survey data. A deeper understanding of the (positive) effects of an intensive preparation phase for each project was only possible by using qualitative interview techniques, although different (qualitative/quantitative) perspectives on the same dimension also led to sometimes contradictory results. The quantitative survey, for example, showed a high success rate for the programme's specific selection process, with participating companies responding quite positively to this particular aspect. The same dimension was discussed more critically in the qualitative case study interviews, where participants stressed challenges regarding the cost/benefit ratio of this phase. The interpretation was that a positive bias towards favourable answers could be expected for online survey results, while a negative bias toward a critical discussion might be possible for the face-to-face interview situation, concentrating on details and challenges of this specific phase.

6.6 Evaluation of the Lithuanian Inno-vouchers LT scheme

6.6.1 Introduction

The Inno-vouchers LT scheme is part of the overall Operational Programme for Economy Growth 2007-2013 and has the specific objective of promoting innovation in SMEs. The vouchers scheme provides small credits (*de minimis* aid) which small business can use to buy R&D expertise or knowledge from research and educational institutions. Each innovation voucher was valued at EUR 5,792 (with participants expected to cover at least 20% of the project costs). Supported activities include industrial or applied research, technological development (experimental or development, design and technological works), and technical feasibility studies.

6.6.2 Methodology

The evaluation was carried out between September 2016 and February 2017 by Visionary Analytics. The objective was to evaluate the relevance, efficiency, effectiveness and impact of the policy instrument (between 2012-2014).

The evaluation, consisting of a typical mixed-methods approach, used theory-based impact evaluation and counterfactual impact evaluation methods. The following data-collection methods were also used: desk research, case studies of four other EU countries, survey, interview, web scraping, statistical and graphical analysis, and two focus groups. A counterfactual analysis using difference-in-differences techniques was also employed. Of particular interest to this MLE was the use of a web-scraping approach. This enabled the automated collection of data from two automotive industry-related web pages (http://imones.lrytas.lt/, http://rekvizitai.vz.lt/) and the web page of the Lithuania Statistics Office (http://www2.stat.gov.lt:8777/imones/sektor.html). The data collected included SMEs NACE codes, company age, number of employees, turnover, debt and city

of registration. The data was used to analyse SMEs and policy instrument descriptive statistics from a number of perspectives.

6.6.3 Evaluation challenges

The web-scraping approach addressed the paucity of data regarding the participants and the (counterfactual) non-participants. In addition, to some extent, the combination of qualitative mixed methods enabled the low observability of effects to be overcome. As the amount of funding provided by the innovation vouchers is relatively small, it was anticipated that its impact on business indicators would be hard to discern in the context of the impact of externalities on business productivity and competitiveness (the challenge of attribution). However, subjective perceptions of the scheme's impact indicated that it had a positive effect on the development of new products and new competences, the creation of new products or services, business competitiveness and business productivity – although these findings should be treated with some caution.

One finding indicated that the Inno-vouchers LT scheme had led to behavioural additionality effects and had a positive impact on new science business cooperation links. According to the survey, SMEs without cooperation experience with Public Research Organisations (PROs) were more positive about such cooperation in the future than SMEs which did not receive the funding. Around 8% of those SMEs funded which lacked prior cooperation experience with PROs before the project had started, cooperated with PROs after the innovation voucher projects.

6.7 Innovation Norway: one data platform for support schemes

6.7.1 Introduction

Innovation Norway is the Norwegian government's agency for supporting innovation and development of Norwegian enterprises and industry. In collaboration with other agencies, the Research Council of Norway, the Ministry of Trade, Industry and Fisheries and the Ministry of Education and Research, Innovation Norway recently developed a data platform. One of the strengths of this platform is that it links the datasets of several agencies and councils which provide support to companies (and other actors). Data about participation of Norwegian organisations in European programmes is also included.

Using this integrated and shared data platform, agencies and evaluators in Norway can analyse which companies receive multiple types of public support and how this changes over time, as companies grow and/or their R&D and innovation activities change. The data in figure 1 is for illustration purposes only because relevant agencies are in the process of uploading their information.

Figure 1: Norwegian data platform for monitoring support schemes
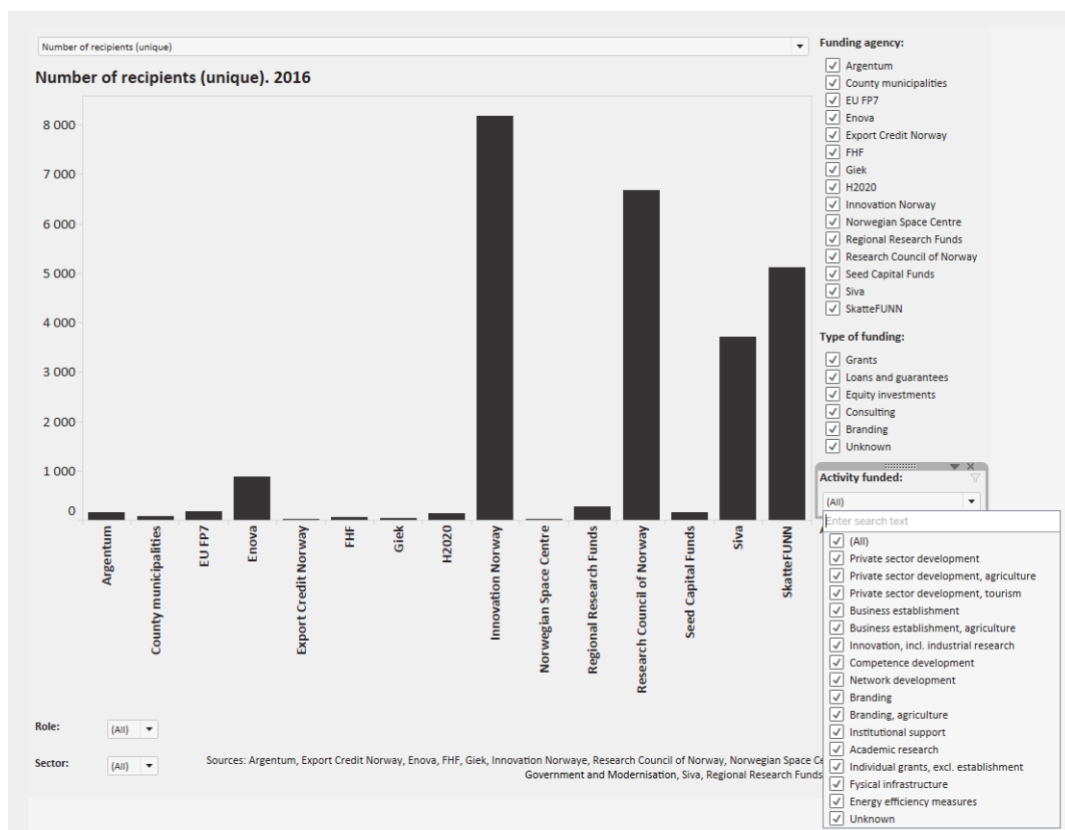
Source: Innovation Norway and Norwegian Ministry of Trade, Industry and Fisheries

For Innovation Norway, this joined-up initiative will complement an internal initiative which links the data across all of Innovation Norway's support schemes. Staff members have easy access (using Microsoft's PowerBI dashboard/front-end) to data about companies that receive(d) support from Innovation Norway. This provides basic checks and descriptive statistics and allows users to explore how Innovation Norway supports companies across their 'innovation journey' (e.g. starting with small, national grants and moving to European consortia in Horizon 2020).

6.7.2 Evaluation challenges

The data platform is now operational and the content can be used in evaluations of individual support schemes or sets/mixes that target specific types of companies, sectors and themes (circular economy, industry 4.0, etc.). The potential of the data platform lies in addressing the evaluation challenges of attribution and time lags.

Because the data relating to a range of support schemes is linked, evaluators can assess whether the effects of one support scheme are influenced by the extent to which companies also use other schemes. In collaboration with Statistics

Norway, Innovation Norway can also link/add company-level data about innovation and economic parameters. This linkage is facilitated by the use of unique identifiers of companies, used in all official communications between them and Norwegian government agencies.

The data platform described above also enables the time lag to be addressed between companies receiving public support and any emerging effects (partly) as a result of receiving public support. The data platform, in essence, is a monitoring platform which makes it possible to track which companies are beneficiaries of which support schemes, over a period of five years or more. This provides a solid basis to conduct *ex-post* evaluations of support schemes, even several years after a company received support or a scheme was stopped.

6.8 *Evaluation of the Technical-Industrial Institutes in Norway*

6.8.1 *Introduction*

The R&D activities of the Technical-Industrial (TI) Institutes in Norway underpin important economic impacts through both commissioned work and collaborative R&D with industry. The institutes provide expertise and research capacity in areas including industrial processes, materials and chemistry and ICT, marine technology, energy, petroleum, nuclear technology, geoscience and technology and society.

Technopolis Group was commissioned to conduct an impact analysis by the Research Council of Norway (RCN). This formed part of the background material for an evaluation of the TI institutes conducted by an international panel of experts appointed by the RCN. The impact analysis involved a mixed-method approach and data linking and was published in 2016 (Research Council of Norway 2016).

6.8.2 *Methodology*

The impact analysis used a range of methods including a web survey and interview data, bibliometric analysis and an in-depth economic analysis. The economic impact of the TI institutes was explored through four different impact streams: (i) direct economic value creation; (ii) indirect and induced economic impact; (iii) economic value created through licensing, patenting and spin-off companies; and (iv) wider economic impact.

The study employed econometric techniques and data linking in an attempt to quantify the wider impact on industry users. It used a quasi-experimental approach to estimate what would have happened (to the industry user's performance) in the absence of the TI institutes' support. This required complementing the analyses of the change in users' performance over time with an analysis of the performance of non-users, i.e. similar companies that did not collaborate with a TI institute (control group).

The study linked data from RCN's data warehouse, (i.e. monitoring data on the interaction of industrial users with the TI institutes), RCN's SkatteFUNN database (which contains information on companies granted tax relief for investing in R&D

and which enabled identification of R&D active companies for the control group) and Eniro's database (which contains financial information on all companies based in Norway). The latter dataset contains information on half a million companies covering over 15 years. The linking was done using unique identifiers and fuzzy matching (of names). All the data were combined to run a counterfactual analysis using several techniques: propensity score matching, difference-in-differences, and panel data with fixed effects. The industrial users' turnover formed the analysis variable.

6.8.3 Evaluation challenges addressed

The methodology accounted for attribution effects due to the use of a quasi-experimental approach to estimate the counterfactual scenario. By using a 15-year panel dataset, the evaluation also accounted for time lags in its modelling, allowing two years for the effect of interaction with the TI institutes to materialise. It also measured how many years the positive effect lasts.

Figure 2 shows the estimated aggregated (grossed-up) turnover development for the industrial users. It also provides a counterfactual scenario for the same companies had they not collaborated with TI institutes. The scenario was calculated by making a prediction at company level, based on the results obtained from the econometric analysis. The difference between the curves thus illustrates the additional turnover attributable to collaboration with TI institutes.

Figure 2: Effect on turnover for the industrial users (billion NOK[10], real prices 1998-fixed)
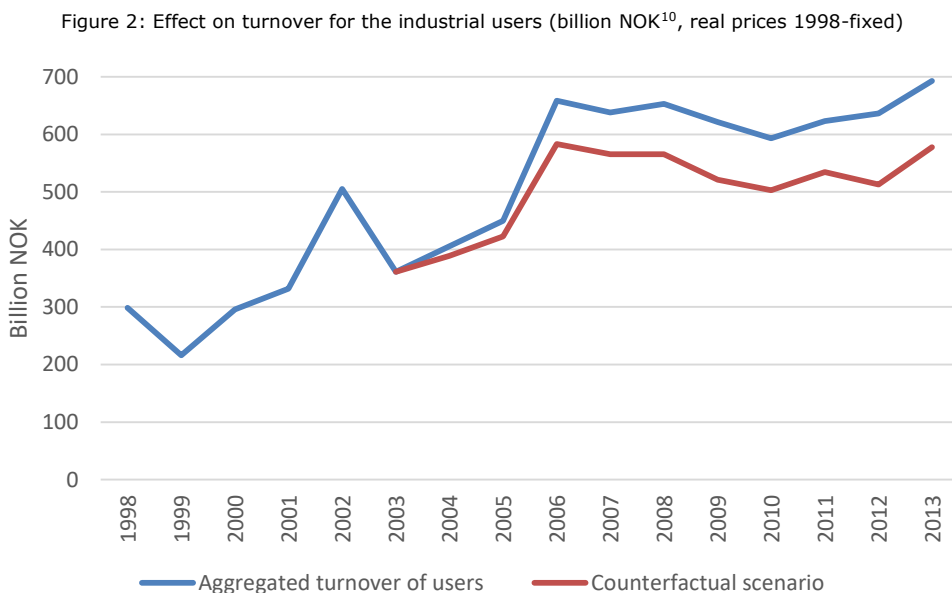
Source: Research Council of Norway (2016)

---

[10] 1 EUR=8.283708 NOK

6.9   Evaluation of SME support at the regional level in Poland

6.9.1   Introduction

The evaluation examined the effects of selected measures from the Integrated Operational Programme (OP) on Regional Development (ZPORR) and the Sectoral OP 'Improving enterprise competitiveness' (SPO WKP) on SMEs in the Polish Zachodniopomorskie region. The selected measures included schemes aimed at: promoting entrepreneurship and the creation of new micro-enterprises; simplifying access to specialist consulting and increasing investment for new micro-firms; facilitating access to external investment financing sources for entrepreneurs, including the provision of micro-loan and guarantee funds; stimulation of SME competitiveness through improved financial access to advisory services; and supporting innovation activities in young start-ups and growth companies.

6.9.2   Methodology

The *ex-post* evaluation was commissioned by the regional marshal's office and conducted in 2010 by an external contractor selected via a competitive tendering process. The main objective of the evaluation was to assess the success of regional SME support policy in the Zachodniopomorskie region in the period 2004-2006 regarding ZPORR and SPO WKP. More specifically, it determined the extent to which the measures had met their objectives and the level of sustainability of the effects, and the influence of the loan and guarantee funds on participating SMEs. It also sought to identify examples of best-practice projects and to draw recommendations for their use in the follow-up Regional Operation Programme for 2007-2013.

The study made use of the following research methods: document and literature analysis, phone questionnaires with beneficiaries, in-depth interviews with beneficiaries, case studies, interviews with experts, and focus groups.

6.9.3   Evaluation challenges

The methodological approach, which linked qualitative and quantitative research results, enabled the consultants to better evaluate the effects of the individual measures, in some way overcoming the issue of attribution which might have resulted from such a complex mix of policy measures. The qualitative approaches allowed the evaluators to take into consideration the perspectives of a range of respondent groups (beneficiaries of the measures and sub-measures, entrepreneurs, scheme administrators and external peer experts). In addition, the methods used were sensitive to the fact that the SMEs in the region were likely to exhibit a large variation in the level of innovation and entrepreneurial development.

However, one problem for the evaluators was the low availability of monitoring data and other information supporting the evaluation process (e.g. contact details).

6.10 The impact of public support for business R&D in Spain

6.10.1 Introduction

The Centre for the Development of Industrial Technology (CDTI) is the main public agency in Spain that grants financial aid to companies for the execution of R&D projects. CDTI provides companies – SMEs and large firms – with grants and loans. In 2017, CDTI launched a study to compare the effect of public support for business R&D on technological inputs and outputs before and during the recent economic crisis. Specifically, firms supported through CDTI programmes for the periods 2002-2005 and 2010-2012 were considered[11].

6.10.2 Methodology

CDTI's data on the beneficiaries of its business R&D support schemes was linked to that about non-supported firms from the Spanish Technological Innovation Survey (the Spanish version of the CIS). This CIS data, in Spain, is owned by the Spanish Institute of Statistics (INE). CDTI data and CIS data were merged by using companies' tax codes as unique identifiers.

Although INE's database is essential to build a proper control group, access to it is restricted by confidentiality rules. Therefore, the merging procedure was carried out by INE staff and CDTI researchers had to work at INE's offices.

One constraint was that the information covering both periods could not be merged as the empirical analysis was undertaken at the INE premises at different times. In addition, the set of variables on non-supported firms provided by INE differed between the two periods analysed, as each of the control samples was available only under distinct anonymisation procedures.

The objectives of the CDTI programme remained constant in recent decades, but funding conditions partially changed between the periods 2002-2005 and 2010-2012. These differences were taken into account when interpreting the results.

Impact assessment was conducted using econometric matching procedures. This methodology assumes that the conditional independence assumption holds – i.e. all firm characteristics explaining selection for a public funding programme are observed. Following this, a probit model is estimated for each period to obtain the propensity scores enabling construction of the counterfactual. To address potential endogeneity, most explanatory variables are included, lagged by one time unit period.

---

[11] A detailed description of the evaluation study: "Public Support to Business R&D and the Economic Crisis: Spanish Evidence" by Ascensión Barajas (*Unit of Impact Assessment, CDTI);* Elena Huergo and Lourdes Moreno (*GRIPICO - Department of Economic Analysis, Universidad Complutense de Madrid)* https://www.researchgate.net/publication/320323742_Public_Support_to_Business_RD_and_the_Economic_Crisis_Spanish_Evidence_Public_support_to_business_RD_and_the_economic_crisis_Spanish_evidence

The results confirm that participation in CDTI programmes positively affected technological inputs (internal R&D intensity, innovation intensity, R&D personnel intensity, and fixed capital intensity), both before and during the economic crisis. Regarding innovative outputs, CDTI support clearly increased the probability of applying for patents, although the effect on process and product innovations differed according to the period considered. Participation in CDTI programmes during the crisis increased the probability of achieving product innovations but not process innovations.

6.10.3 Evaluation challenges

By means of data linking, a control group could be created which enabled the researchers to address the challenge of **attributing effects** to public support measures. The econometric approach took into account the time lag between companies receiving public support and the emerging effects. To some extent, the study succeeded in addressing data limitations (e.g. the size of grants and loans was not taken into account) and the heterogeneity of companies (e.g. company size and sector). Another limitation of the study is that the features of CDTI programmes changed between the periods 2002-2005 and 2010-2012, thus, the relation between the effect of public support and the economic cycle should be treated with caution.

6.11 Entrepreneurial experimentation and collaboration in Sweden

6.11.1 Introduction

The impact analyses carried out by the Swedish Governmental Agency for Innovation Systems (VINNOVA) aim to evaluate and understand the effects of VINNOVA's efforts towards sustainable growth, social benefits and the development of innovation systems. R&D investments for SMEs are given particular importance. In 2017, VINNOVA published a study by researchers from Lund University in which the impact of various VINNOVA programmes on Swedish SMEs was evaluated[12]. The evaluation combined various research methods which aimed to understand in a much more detailed way the innovation journeys of SMEs that had benefitted from various VINNOVA schemes for the promotion of R&D and innovation. The evaluation study is unique as it not only takes a long-term perspective, but also looks into the schemes' spillover and systemic effects, in addition to the direct effects.

---

[12] Lindholm Dahlstrand, A., M. Nilsson and D. Politis (2017), Entreprenöriellt experimenterande och samverkan i innovationssystem, SMF-Effektanalys 2016-2017, Lunds Universitet, Lund. In 2015, VINNOVA published a methodological framework for impact analysis of its R&D programmes for SMEs, see VINNOVA (2015), FoU-program för Små och Medelstora Företag - Metodologiskt ramverk för effektanalyser Författare, VINNOVA Analys 2015-05, VINNOVA, Stockholm

6.11.2 Methodology

The evaluation was a typical mixed-method approach with a focus on behavioural change in the firms involved, i.e. the innovation journey.

As a preliminary step, VINNOVA was responsible for the overall mapping and quantitative analysis of the 1,341 SMEs supported by VINNOVA's SME programme between 2001 and 2015. This included data on sector, size, possible industrial group membership, age at first participation, key figures at group level, firm dynamics (such as bankruptcies, mergers and liquidations), ownership (Swedish/foreign), number of participations and total commitment volume.

In a second step, a smaller population was chosen for a cross-sectional analysis, where these quantitative data were combined with qualitative information through interviews with SMEs directly involved. The cross-sectional analysis consists of companies that have received both targeted SME support (1,341 SMEs) and a consortium-based support (737 SMEs), giving a total of 199 companies. The sample selection for this cross-sectional analysis was further limited to those companies that answered VINNOVA's questionnaire from the 2016 impact study (88 SMEs).

Ultimately, 60 firms were interviewed by telephone. The interviews included questions about both business, collaborative and system effects, their purpose being to understand the innovations of companies that receive VINNOVA funding outside the targeted SME initiative. In a third step, a small number of SMEs (12) were selected for case studies which examined the effects at business, collaborative and system level. A core goal of this stage was to develop a better understanding of VINNOVA's role in the innovation journey and to provide a characterisation of system impacts. The case studies included analyses of companies that have received VINNOVA support and actors (private and public, large and small companies, etc.) which may have been affected indirectly by the VINNOVA-funded project.

The study found that a small share of the innovation journeys and companies account for a large proportion of the direct economic-value creation. In most cases, the projects had relatively modest effects in terms of revenue growth and job creation, and in almost a third of the companies there were no identifiable impacts at all. However, one fifth of the firms have been able to report significant growth and job creation, and a small number show strong growth as a direct result of project funding (underlining the importance of outliers). However, it was observed that the SMEs using VINNOVA's schemes collaborate with other actors which, in the majority of cases, resulted in both the creation of economic value and jobs. Over half the SMEs experienced collaborations that resulted in financial spill-over impacts for their interaction partners.

As for system effects, typical system functions include knowledge development, entrepreneurial experimentation, market formation, resource mobilisation, legitimacy and development of positive externalities. It was shown that this group of SMEs has a frequent and extensive influence on different components of the system, particularly in terms of knowledge development and knowledge

orientation. Up to two thirds of companies in the cross-sectional study indicated they had an impact on the functioning of the innovation system.

Furthermore, the SMEs themselves are important examples of entrepreneurial experimentation, whilst also influencing and strengthening this system component through the creation of new firms or through former employees creating their own firms. The study concludes that there seems to be a clear link between VINNOVA's funding and the company's ability to develop and influence the system functionality. It appeared that 60% of the firms that benefitted from both types of VINNOVA support are university-based SMEs. These were shown, on average, to have been granted more and higher amounts of project support than other companies. They are also slightly larger than other companies, tend to make a profit more often and participate more frequently in various types of ownership changes than non-university-based SMEs. They also have an impact on more of the system's functions than non-university companies, especially entrepreneurial experimentation and the creation of new markets and business models.

An important policy implication of the study is that in the design of public SME support, technical experimentation and market experimentation must be considered: these form the basis for the system's overall entrepreneurial experimentation. This in turn underlines the importance of a well-designed institutional framework for successful entrepreneurial experimentation. The study concluded that the creation of new industries usually requires a very-long-term policy perspective.

6.11.3 Evaluation challenges

The evaluation deals with new ways in assessing spill-over and systemic effects together with the direct effects of R&D and innovation schemes aimed at SMEs. Methods for performing such analyses are as yet underdeveloped, thus the evaluation faced various challenges.

The first concerned the relationship between primary study subjects (i.e. funded projects), secondary study objects (i.e. funded companies) and the wider innovation system. Business dynamics and dynamics in the ownership of project results entail significant challenges in following the development of the projects over a sufficiently long period of time to enable conclusions to be drawn about different types of effects. High-quality evaluations and policymaking require methodological approaches which can handle these dynamic challenges satisfactorily.

A second issue is selection bias (including survival bias) in the cross-sectional analysis since it was only possible to work with those firms that participated in the 2016 impact study and which were also still present and willing to cooperate in the current evaluation. A third evaluation challenge was the temporal scope of the project, which precluded the implementation of any longitudinal studies. A fourth challenge was the selection of the case study firms. Here, the decision was made to analyse the outliers rather more than the 'average' firm: VINNOVA identified a group of potential companies that were assumed to have a particularly large realised or potential system impact.

6.12 Evaluation of technoparks in Turkey: data linking in the Entrepreneur Information System (EIS)

6.12.1 Introduction

Technoparks in Turkey host and support high-tech companies (cf. incubators). They are located on a university campus, which allows them to benefit from academic knowledge and research by universities or research institutes. The relevant legislation conveys important advantages for companies that are located in technoparks; those that conduct R&D and design projects can benefit from tax exemptions and exceptions. Moreover, these companies can obtain budgets for R&D and design projects, via universities and other institutions that support such projects. The main incentives for universities to support technoparks are to transfer knowledge to the technoparks and individual high-tech companies. They also benefit from the companies' knowledge when cooperating with companies and managing the research programmes. As of early 2018, there were 56 active technoparks with 4,486 operating companies. About 82% of employees work as researchers, which underlines the high-tech nature of the companies.

6.12.2 Methodology

A mixed-method approach was used combining quantitative and qualitative methods. Quantitative methods, notably econometrics, included a difference-in-differences approach, using statistical data and data obtained from online surveys. The difference-in-differences analysis was conducted by the experts of the Impact Evaluation Department at the Ministry of Science, Industry and Technology. The use of qualitative methods (in-depth interviews) and the online survey were outsourced.

Conducting difference-in-differences analysis requires data for both an intervention group and a control group. The Ministry of Science, Industry and Technology has a database called the Entrepreneur Information System (EIS), which consists of data obtained from eight institutions. EIS includes micro-level, firm-level data from between 2006 and 2017. This covers financial data (sales, total assets, R&D expenditures, etc.), the number of employees, the number of intellectual properties and the amount of grants and incentives received from government institutions. Using EIS data, the difference-in-differences analysis demonstrated that location on a technopark has a positive impact on companies' R&D expenditure, sales and employment levels.

Online surveys and in-depth reviews were used to analyse the relevance, effectiveness, efficiency, impact and (financial) sustainability of technoparks. Based on this analysis, together with the results of the difference-in-differences analysis, policy recommendations were formulated. Among other things, these address how technoparks can become more sustainable and can continue to stimulate positive behavioural change of innovative, high-tech entrepreneurs.

6.12.3 Evaluation challenges addressed

By using data from the EIS and applying a difference-in-differences approach, the evaluation addressed the attribution and time-lag challenges of conducting evaluations. Two challenges for the future evaluation of technoparks are the low

observability of effects (especially spillovers) and the heterogeneity of companies. Although the EIS and the technopark databases contain very relevant information, they hold little data about commercialisation by companies. Moreover, the EIS and technopark databases do not (yet) contain data about the spillover effects on their region and in specific sectors. The difference-in-differences analysis did not address differences between sectors because the sample size per sector was insufficient. Larger samples per sector and a detailed analysis at sectoral level would help policymakers and technopark management to improve their facilities and services for companies in specific high-tech sectors.

6.13 Evaluation of the UK's Smart scheme

6.13.1 Introduction

Operated by Innovate UK, Smart has been one of the UK's longest-running publicly funded innovation support instruments. It offers grant co-funding of up to £250,000 to support UK-based pre-starts, start-up micro businesses and other SMEs to undertake projects from which successful new products, processes and services could emerge. Three types of project are supported: proof of market, proof of concept, and development of prototype. Its ultimate intended outcomes are to contribute to the UK economy through the creation of wealth and jobs. Its rationale is meant to overcome the risk and uncertainty preventing optimal levels of investment in R&D by private firms.

6.13.2 Methodology

The study, conducted by SQW, was carried out in two phases. The first, a retrospective assessment, included large-scale surveys of businesses that had been awarded Smart grants and a comparison group of businesses that applied for but did not receive one (note the typical use of a counterfactual again), using difference-in-differences analysis. Furthermore, it featured in-depth case studies.

A second longitudinal phase of the evaluation was ongoing at the time of writing. This has adopted a similar mixed-methods approach, although with two rounds of large-scale surveys, the addition of a data-linking component (using firm-level data from administrative sources), and a qualitative element focused on understanding how spill-over effects came about. A final impact evaluation will be reported in 2018.

As such, the evaluation provides a valuable example of the mixed-methods approach to evaluate a range of impacts of Smart on its target population.

6.13.3 Evaluation challenges addressed

High levels of variance in the dataset posed challenges to the econometric analysis of the survey data, thus it was important to run analyses on different segments of the dataset and use this alongside self-reported evidence to provide an assessment of additionality and outcomes. The use of a two-phase approach addressed the issue of time-lagged outcomes and effects, whilst the challenge of low observability was managed by asking the direct beneficiaries whether any customers, suppliers or competitors would have benefited from the project. The survey also asked what form those benefits took, to build a typology of different

spill-overs, which was subsequently tested in the second phase. Contact details were also used in order to interview the indirect beneficiaries as a follow-up exercise. Some of the problems encountered included the difficulty of contacting indirect beneficiaries which limited the analysis to a non-representative qualitative view.

6.14 Evaluation of the Digital Catapult UK: Agent Based Modelling

6.14.1 Introduction

Innovate UK provides financial support for Catapult Centres, 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 Digital Catapult commissioned an evaluation study that tests the potential of an Agent Based Model (ABM). The main objective of the evaluation was to explore the potential impacts of new digital technologies. These can be incremental and disruptive changes, both expected and unexpected. As such, the evaluation design does not – beforehand – define all types of impact. Nor does the evaluation select a small set of mechanisms via which the Digital Catapult Centre has impact on participating companies and other companies. To a large extent, the mechanisms and types of impact are identified in the qualitative, first phase of the ABM exercise. Moreover, the evaluation makes use of several, linked datasets. The relatively open, explorative approach, using several datasets, means this evaluation can be positioned as a big-data approach. The evaluation also explicitly explores how companies change their behaviour having participated in the Digital Catapult.

6.14.2 Methodology

ABM is increasingly seen as a solution to the challenges of modelling emergent and disruptive change. It allows 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, ABM have previously not experienced widespread adoption as the economic community seems to be more locked into existing statistical models.

ABM is one way of modelling complex systems. Steps include defining the boundaries of the system and defining rules to classify the objects/actors in the system. One assumption is that agents are identical (or at least, that there is a small set of agents – with specific rules – who interact 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 and also for management and strategy

development in general. Internally, data is collected about inputs and activities in addition to basic company data (using existing datasets so as 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, while in order to assess impact, they used modelling, looking at value added, productivity and societal benefits.

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 it represents 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). Note that this evaluation is considered as an experiment, to be during 2018.

6.14.3 Evaluation challenges that were addressed

The ABM approach taken in the evaluation of the Digital Catapult has the potential to address the evaluation challenge of low observability. Using ABM, the analysis is not constrained to direct beneficiaries and predefined types of outcomes and impacts.

Instead, ABM opens up the analysis to a broad range of outcomes and impacts, increasing the chances that direct and indirect effects are captured by the evaluation. Also note that ABM allows for assessing (by means of subjective statements as well as data) the order of magnitude of effects and expected effects.

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 report has been prepared for a Mutual Learning Exercise (MLE) on the evaluation of business R&D grant schemes. It focuses on three incremental innovations in the evaluation of support schemes for business R&D and innovation namely: the added-value of taking a behavioural change perspective and measuring and understanding how the R&D and innovation behaviour of companies changes in response to policy measures; recent advances in mixed-method approaches, including econometrics, the use of control groups and qualitative methods, in the evaluation of the impact of business R&D support measures; and the opportunities and challenges of big data in policy evaluations, including data linking.

MLE participants provided examples of recent evaluations and they stressed current and planned activities for improving their evaluations of R&D and innovation support schemes.

Studies and reports

Publications Office